PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

The Croonian Lecture, 1996: Endogenous Damage to DNA

T. Lindahl

Phil. Trans. R. Soc. Lond. B 1996 **351**, 1529-1538
doi: 10.1098/rstb.1996.0139

THE CROONIAN LECTURE, 1996: Endogenous damage to DNA

T. LINDAHL

Imperial Cancer Research Fund, Clare Hall Laboratories, South Mimms, Herts EN6 3LD, U.K.

SUMMARY

Although DNA is the carrier of stable genetic information, this giant molecule exhibits slow turnover in cells as a consequence of endogenous damage. DNA lesions result from hydrolysis, and from exposure to active oxygen and reactive metabolites. These major forms of damage to the heterocyclic bases and to the DNA backbone structure are now well characterized. Most DNA repair enzymes have apparently evolved to prevent genomic instability caused by endogenous lesions, the only exception being those that counteract ultraviolet light damage inflicted by the sun. Despite the efficiency of DNA repair pathways, some forms of endogenous DNA damage still cause mutagenic alterations and may result in human disease.

1. INTRODUCTION

The preceding two Croonian Lectures discussed the structure and dynamics of populations. In this one, I describe the structure and dynamics of the DNA molecule. The 1701 bequest of the Croonian Lecture was 'for the advancement of natural knowledge on local motion', so I shall report on the 'local motion' of DNA. With this introduction in mind, most colleagues might assume that I will discuss the Watson–Crick model of DNA, 'breathing' and the replication of the double helix. However, this will not be the main topic of this lecture. Instead, I largely dwell on the intrinsic lability of the covalent structure of DNA.

The correct covalent structure of this important molecule (figure 1) was proposed by Dan Brown and a former President of the Royal Society, Lord Todd, in a paper submitted in 1951 to the *Journal of the Chemical Society* with the unpromising title 'Some observations on the structure and chemical behaviour of the nucleic acids' (Brown & Todd 1952). As is now well known, the covalent structure of DNA is a linear, non-branched chain of nucleotides, with the deoxyribose residues joined together by 3′, 5′ phosphodiester bonds, and the four bases bound by glycosyl bonds to the C′1 residues of the sugar molecules. The bases are stacked on top of each other, like a pile of pennies. The DNA is a salt at neutral pH, with K^+ and Mg^{2+} ions, polyamines, and lysine and arginine residues of proteins neutralizing the charged phosphate residues. Moreover, DNA is a hydrated macromolecule, with 8–10 tightly bound water molecules per nucleotide residue.

This structure of DNA does not tell you that DNA is the carrier of genetic information, nor that two of these chains form a double helix with hydrogen bonds between adenine (A) and thymine (T) residues, and guanine (G) and cytosine (C) residues. Still, it is clear that without this precise knowledge of the covalent structure of DNA, the three-dimensional conformation could not have been established two years later by Watson and Crick.

2. STRUCTURE AND STABILITY OF THE DNA BACKBONE

Valuable genetic information must be retained in cells; consequently, in those early days, molecular geneticists tacitly assumed that DNA should be an extremely stable molecule, perhaps protected from damage by being covered by histones and other proteins. Watson and Crick proposed in 1953 that mutations might arise by rare tautomeric shifts between the keto and enol forms of the heterocyclic bases, which would give reduced coding specificity during replication; they did not consider the alternative possibility that damage to the covalent DNA structure might be mutagenic. The discovery of nucleosomes showed that DNA is not contained within a protective shell or tube of cellular proteins, and there are also good reasons why the backbone of DNA must not be too stable. Two examples are shown in figure 2.

The difficult task of correct replication of a single-stranded template is performed by a DNA polymerase activity with an accuracy of only about 10^4; that is, one incorporation mistake is made for every ten thousand nucleotides replicated. The problem arises because some alternative base pairs can be formed readily, such as that between G and T or between two purine nucleotides when one of them is in the *syn* conformation, and also because the DNA polymerase itself is a slightly error-prone enzyme. This level of accuracy is not sufficient for faithful replication of a genome, so replicative DNA polymerases have an intrinsic editing function: they can immediately reverse and cleave the

Phil. Trans. R. Soc. Lond. B (1996) **351**, 1529–1538
Printed in Great Britain

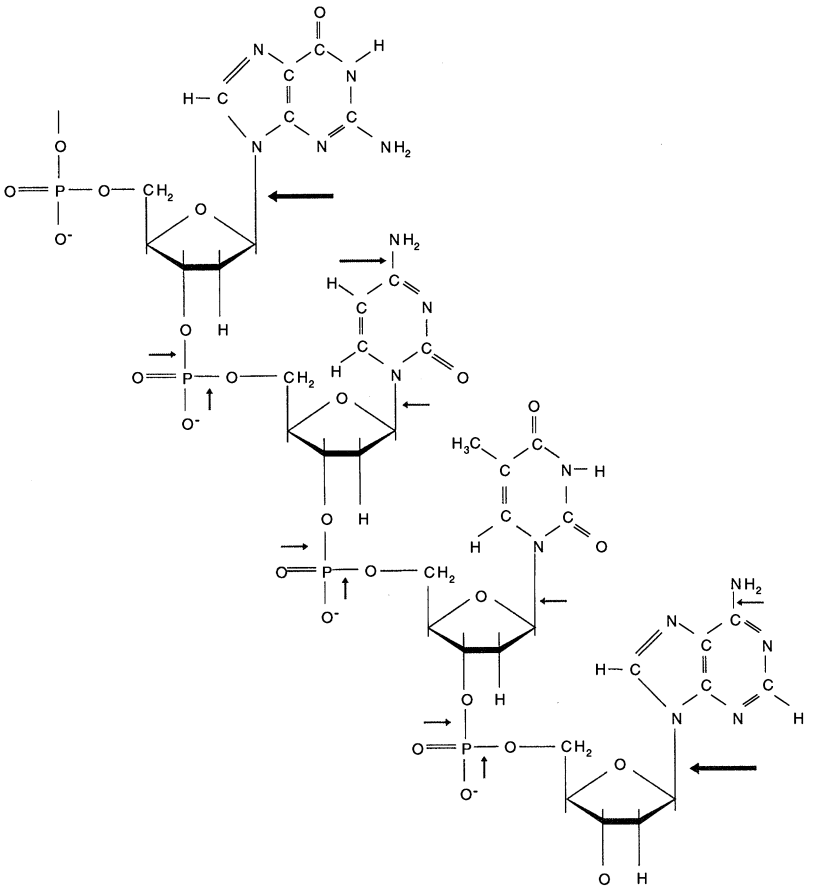

Figure 1. Covalent structure of DNA. Arrows indicate major sites of hydrolytic attack.

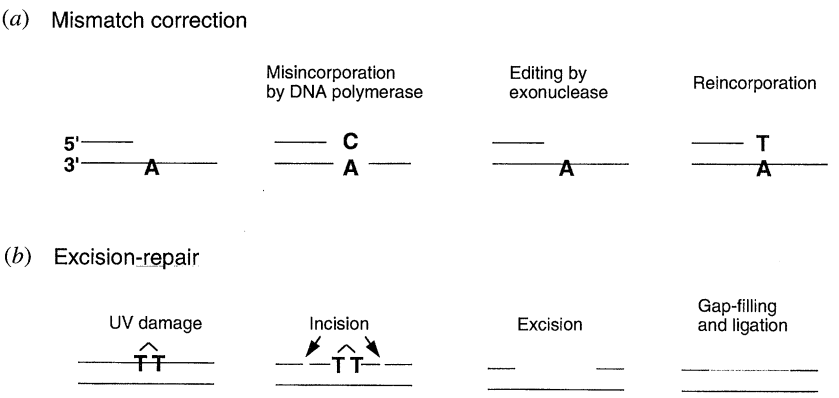

Figure 2. (*a*) Editing of misincorporated nucleotides in DNA. (*b*) Excision-repair of DNA damage: a pyrimidine dimer generated by exposure to ultraviolet light is removed by DNA repair enzymes.

phosphodiester bond they have formed at a mismatch. This improves their copying accuracy from about 10^4 to 10^7, which is essential for precise DNA replication. A further thousandfold accuracy is gained by a mismatch repair system, which screens newly synthesized DNA for errors in the daughter strand and excises those errors. Furthermore, damaged nucleotide residues have to be excised and replaced by special DNA repair enzymes. If the DNA backbone were an exceptionally stable structure, these obligatory corrections would be very difficult and energy-requiring processes. Instead, by employing relatively labile diesters in the backbone, the chain is readily susceptible to hydrolytic cleavage. Enzymes are catalytic agents that typically speed up the rate of a reaction 10^{10}–10^{11} times, so the very slow but relevant spontaneous cleavage of a phosphodiester bond in aqueous solution can be exploited by rapid removal of mismatched or damaged nucleotides by simple hydrolytic nuclease activities. This is a great advantage of the DNA backbone structure compared with a chemically more stable alternative.

Could a synthetic organic chemist devise a better backbone for DNA today? The answer seems to be that possibly something equivalent could be generated, but

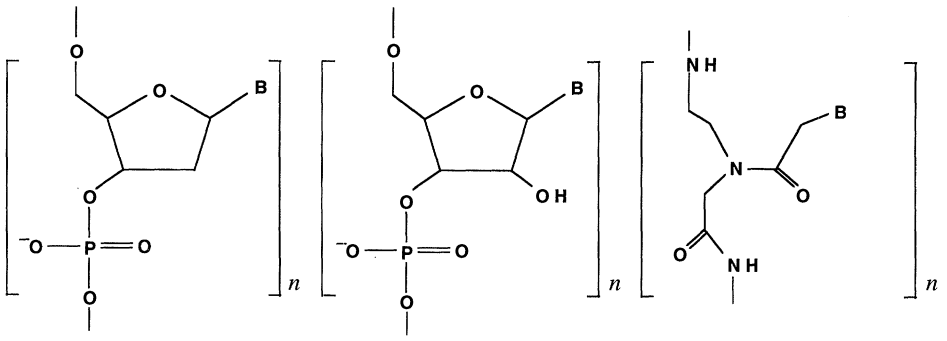

Figure 3. Backbone structures of DNA (deoxyribose phosphate diesters), RNA (ribose phosphate diesters) and peptide nucleic acid (aminoethylglycine peptides).

nothing clearly better (figure 3). The question has been discussed in detail by Frank Westheimer (1987). Most esters are easily hydrolysed, and phosphate esters are unusual in their slow reactivity. Silicic acid is more abundant in nature than phosphoric acid, but its esters are far too easily hydrolysed, so a DNA molecule with silicate instead of phosphate would decompose quite rapidly. The same holds true for most other diesters. Another important property of the chain structure is that it, and its precursors, should be ionized (Davis 1958). Neutral molecules usually have some solubility in lipids and can leach through a simple lipid membrane, whereas ionized molecules are lipid-insoluble; this must have been an important factor in the origin of cells to prevent loss of essential compounds into the surrounding water. Thus, divalent instead of trivalent alternatives, such as sulphate instead of phosphate, in DNA would clearly be inferior, because the corresponding diester would not carry a charge.

Despite the relatively good stability of phosphate esters compared to other esters, they still undergo hydrolysis. RNA is a poor carrier of substantial amounts of genetic information, because the 2′-hydroxyl group of ribose allows formation of a 2′, 3′ cyclic phosphate that facilitates hydrolysis of the chain. Instability of the RNA chain can be counteracted by reduction of the ribose to deoxyribose. My own experiments on this issue have shown that, in the presence of physiological concentrations of Mg^{2+}, the phosphodiester bonds in a DNA chain are 200 times more stable than those in an RNA chain at neutral pH. Although RNA can carry genetic information in viruses, the backbone stabilization of DNA must have been essential for the evolution of higher forms of life. It is noteworthy that the biosynthesis of deoxyribose from ribose by direct reduction of the sugar moiety in DNA precursor molecules is chemically extremely difficult and could only be achieved by one of the most complex and fascinating of all enzymes, ribonucleotide reductase, which contains a stable organic free radical in its active site (Reichard 1995). The creation of ribonucleotide reductase may well have been a bottleneck for the evolution of multiplying cells.

Because esters hydrolyse readily, amides could provide an alternative for backbone structures that are generally more stable (Westheimer 1987); peptide bonds show high stability in aqueous solution. However, once a strategy had evolved for reduction of ribose to deoxyribose, further stabilization of the chain would be unnecessary and would make editing, proofreading and repair more difficult. Recently, nucleic acid chain analogues with the common bases united by peptides instead of deoxyribose phosphates have been synthesized (Nielsen 1995). These 'peptide nucleic acid' structures (figure 3) promise to be important reagents and allow for specific Watson–Crick base pairing with authentic DNA chains. Charged amino acid derivatives could provide an interesting alternative for a backbone structure that was not employed as a long-term solution during evolution.

3. SPONTANEOUS DEPURINATION OF DNA

An unfortunate side effect of the reduction of ribose to deoxyribose in nucleic acids is that the glycosyl bonds become more susceptible to hydrolysis. Thus, degradation of DNA in acid mainly occurs by initial cleavage of glycosyl bonds at purine residues, followed by chain rupture at the weakened abasic sites, rather than by direct cleavage of phosphodiester bonds between two intact nucleotides. The bases also are not entirely stable. In this regard, it is apparently extremely difficult to synthesize base analogues of improved chemical stability without compromising the accurate base-pairing specificity of the present four heterocyclic bases; to my knowledge such analogues have not been made.

The problem I set out to investigate in the early 1970s was to determine whether the intrinsic instability of DNA was of any physiological relevance, or whether the hydrolytic reactions I have mentioned only occurred at significant rates under non-physiological conditions, for example at extreme pH values. It was already known at that time that external agents such as ultraviolet light, ionizing radiation, and alkylating agents could damage intracellular DNA, and that cells had active DNA repair mechanisms to correct that damage. However, it was not known whether spontaneous DNA damage occurred in cells not exposed to external mutagens.

The experiments to investigate that question were straightforward: they involved growing bacterial mutants with appropriate defects in nucleotide metab-

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

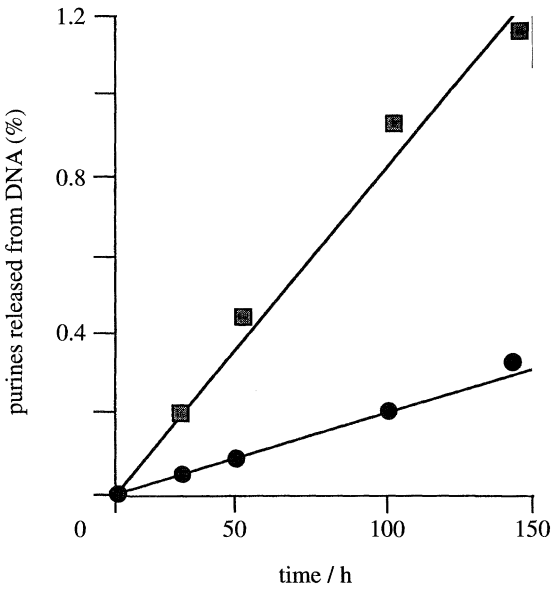

Figure 4. Rate of hydrolytic depurination of DNA at 70 °C. ^{14}C-purine-labelled DNA was incubated in 100 mM KCl, 50 mM Hepes·KOH, 10 mM $MgCl_2$, 1 mM EDTA, pH 7.4, and the amounts of released purines determined at various times. Data are shown for double-stranded DNA (circles) and for single-stranded DNA (squares). Guanine and adenine are released at similar rates, and the reaction proceeds 15-fold more slowly at 50 °C than at 70 °C (from Lindahl & Nyberg 1972).

olism in the presence of a ^{14}C-labelled base, followed by isolation of DNA from the bacteria. The DNA was then incubated for long periods of time in sealed ampoules in buffers of various composition and pH, and the rates of base loss were determined. To reduce incubation times to days rather than weeks or months, a series of elevated temperatures were used; this allowed determination of the activation energies of the reactions, so the rate of base loss at the lower temperature of 37 °C could be accurately estimated. A typical experiment is shown in figure 4.

Base release occurs mainly by a mechanism of specific acid-catalysed hydrolysis even in the neutral pH range, so the buffer composition in the experiments turned out to be irrelevant. As established for acid hydrolysis of deoxyribonucleosides (Zoltewicz *et al.* 1970), protonation of the base is followed by direct cleavage of the purine-deoxyribose glycosyl bond. The rate constant for depurination of double-stranded DNA at 37 °C and pH 7.4 was estimated to be $k = 3 \times 10^{-11}$ s^{-1} (Lindahl & Nyberg 1972). This is a very slow reaction, which most chemists would consider so sluggish as to be irrelevant. However, it does correspond to the loss of 10^4 purines per day from the human genome under conditions *in vivo*, so a long-lived cell might lose several per cent of the purine residues in DNA during the decades of life of an individual. This clearly would be unacceptable, so we searched for a DNA repair enzyme that could correct apurinic sites in DNA.

No enzyme seems to exist that can re-insert the missing purines directly, but a nuclease was found in mammalian cell extracts that incises DNA specifically at apurinic sites (Lindahl & Andersson 1972; Ljungquist & Lindahl 1974). The same type of enzyme was also found by Walter Verly. He correctly reasoned that delayed toxicity of alkylating agents on bacteriophage T7 might reflect slow loss of *N*-alkylated purines from DNA, and he detected an enzyme in the bacterial host that could incise at the secondary alkylation lesions, that is, apurinic sites (Paquette *et al.* 1972). The endonuclease for apurinic sites triggers an excision-repair event that results in replacement of a base-free deoxyribose phosphate residue with the correct nucleotide. So, abasic sites in DNA generated by hydrolytic depurination (figure 1) can be repaired rapidly and efficiently in all living cells. In conclusion, intracellular DNA is not entirely stable but is undergoing constant slow turnover due to spontaneous hydrolytic damage (Lindahl 1993). This conclusion was initially disputed, but the objections (for example, that DNA in chromatin may be protected from water) have largely been dismissed.

Thermophilic bacteria growing at 90–100 °C have a 1000-fold higher rate of depurination of DNA than cells growing at 37 °C, but cell survival may still be explained by efficient DNA repair. A challenge to this concept came from a paper published in *Nature* in 1983 (Baross & Deming 1983), reporting on bacteria from a submarine hot spring that grow at 250 °C. At such a high temperature, hydrolytic depurination of DNA should cause most of the purines to be released in less than a minute; such massive damage could not be effectively repaired. Moreover, it is unclear how hydrogen bonds between the DNA strands could be retained at a temperature at which tin melts. To my relief, a subsequent paper from another group showed that the apparent bacterial growth at 250 °C could be ascribed to a high-temperature chemical artefact, although this interpretation has been rejected by the original authors (Trent *et al.* 1984; Deming & Baross 1993).

4. DEAMINATION OF CYTOSINE TO URACIL IN DNA

In addition to cleavage of base–sugar bonds, some deamination of DNA bases was known to occur both in strong acid and alkali. Shapiro showed that the free ribonucleoside cytidine is deaminated more rapidly than the purine nucleosides, adenosine and guanosine, in weak acid (Shapiro & Klein 1966), but it was unclear whether this reaction was of relevance for DNA. I investigated the rates of cytosine deamination in DNA at neutral pH using the same general approach as for depurination measurements. DNA containing ^{14}C-labelled cytosine residues was incubated in solution for prolonged times at elevated temperatures and hydrolysed enzymatically to nucleosides; the amount of radioactive deoxyuridine generated was determined by chromatography. Cytosine deamination was shown to occur at a significant rate in DNA under physiological conditions (figure 1). The double-helical structure provided considerable protection against hydrolytic deamination, so the reaction was less frequent than depurination (Lindahl & Nyberg 1974). More sensitive methods have been devised recently by others

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

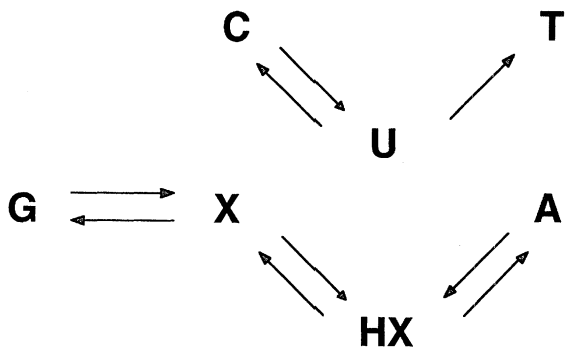

Figure 5. Lack of direct conversion of one common base in DNA to another. To prevent direct decomposition of one base to another, at least two interconversion steps are required. Demethylation of thymine to uracil is unlikely to occur because of the stability of the carbon-carbon bond. HX, hypoxanthine (the nucleoside form is deoxyinosine); X, xanthine.

to measure cytosine deamination in DNA under conditions *in vivo*; and the results are the same (Frederico *et al.* 1990).

The observed rates of deamination indicated that an active cellular DNA repair mechanism should exist to prevent the potentially mutagenic effect of conversion of cytosine to uracil residues. In a search for such an activity, a surprising finding was made: uracil was enzymically released as a free base from DNA containing dUMP residues (Lindahl 1974; Lindahl *et al.* 1977). The enzyme catalysed the hydrolysis of uracil–deoxyribose glycosyl bonds, and was unable to cleave phosphodiester bonds, so it was not a nuclease. Microbial mutants deficient in this DNA glycosylase were later shown to have a mutator phenotype with an increased frequency of C → T transitions; this increase is the expected effect of replicating unrepaired deaminated cytosine residues (Duncan & Miller 1980).

The spontaneous conversion of cytosine to uracil, with a change in coding specificity, and the necessary active removal of uracil residues from DNA, provides an explanation for the presence of uracil with a distinct methyl tag, i.e. thymine, instead of unsubstituted uracil in DNA (Lindahl *et al.* 1977). The genomes of RNA viruses are small and tolerant to a high mutation frequency with continuous selection of progeny, so the simultaneous presence of cytosine and uracil can be accepted in this case. However, a general rule for DNA is that none of the common bases can be directly converted to one of the other three (figure 5). Thus, deoxyinosine residues instead of deoxyguanosine would be inherently unsuitable in DNA because, in addition to the reduced coding specificity, deoxyadenosine residues can be directly deaminated to deoxyinosine.

The three-dimensional structure of uracil–DNA glycosylase was solved last year by two groups (Savva *et al.* 1995; Mol *et al.* 1995). The enzyme apparently scans DNA for minor irregularities in the double-helical structure, and can flip out and cleave deoxyuridine residues at such sites. The abasic site is then repaired by excision-repair, as for sites generated by non-enzymic depurination (figure 6). The entire repair reaction has been reconstituted with purified bacterial enzymes (Dianov & Lindahl 1994) and in recent work

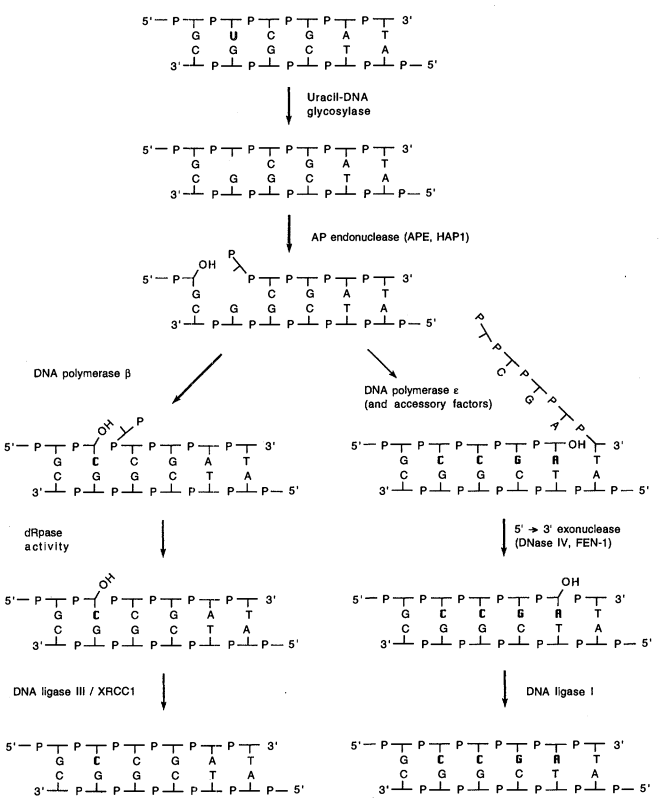

Figure 6. The base excision-repair pathway. The major route is shown on the left and results in the replacement of a single nucleotide residue; a minor alternative pathway for the latter steps is shown on the right. The human enzymes that may account for the separate steps are indicated. From Lindahl *et al.* (1995).

Lesion	Mutations
Apurinic site	A•T → T•A G•C → T•A
Cytosine deamination	C•G → T•A
5-Methylcytosine deamination	5MeC•G → T•A
8-oxo-G formation	G•C → T•A
O^6-Methylguanine formation	G•C → A•T
O^4-Methylthymine formation	T•A → C•G

Figure 7. Mutation patterns of common spontaneous lesions in DNA.

with human enzymes (Kubota *et al.* 1996). The initial steps of the pathway in human cells are the same as in bacteria; the only differences concern details of the gap-filling procedure. Because several different DNA glycosylases have now been identified, the pathway is called 'base excision-repair' since the initial step is the removal of a damaged or non-conventional base from DNA (Friedberg *et al.* 1995).

DNA contains a fifth distinct base, 5-methylcytosine, which forms a base pair with guanine. Enzymic methylation of cytosine to this minor base at CpG sequences in mammalian DNA is largely used for silencing gene expression; the process is essential for normal development (Li *et al.* 1992). The methyl group creates a problem of mutagenesis, because 5-methylcytosine can be directly deaminated to thymine, in contravention to the rule shown in figure 5. A mammalian DNA glycosylase exists that recognizes G·T base-pairs and excises the thymine residue without strand specificity (Neddermann & Jiricny 1993). This enzyme seems very slow and inefficient, perhaps because it cannot be allowed to act on newly replicated DNA before replication errors in the daughter strand have been corrected by editing and mismatch repair. As a consequence, G·C → A·T transitions at sites of cytosine methylation account for about one third of the single-site mutations that cause inherited disease in humans; the same base change is frequent in mutated p53 tumour suppressor genes found in many human cancers (Rideout *et al.* 1990). The other two thirds of point mutations may result as a consequence of several other forms of spontaneous DNA damage (figure 7) as well as from rare mistakes in DNA replication.

5. DNA DAMAGE BY ACTIVE OXYGEN

In addition to its reactivity with water, DNA is also readily damaged by oxygen free radicals (Breen & Murphy 1995). Disruption of the 3′, 4′ bond of deoxyribose may result and lead to a strand interruption. DNA pyrimidines are particularly sensitive to ring saturation, fragmentation, and condensation induced by active oxygen; several non-coding degradation products may be generated, such as thymine glycol, *N*-substituted ureas, and hydantoins. These base derivatives are all actively removed by the action

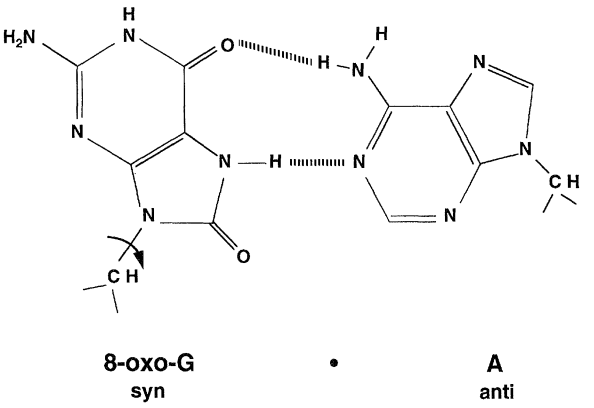

Figure 8. Base-pairing between adenine and the oxygen free-radical-induced lesion 8-oxy-7,8-dihydroguanine in DNA. The arrow indicates rotation of the base from the *anti* to the *syn* conformation of the nucleotide residue.

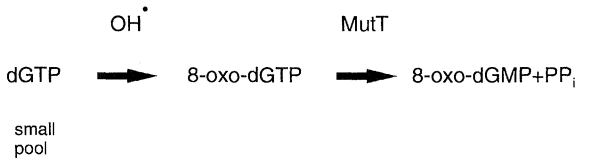

Figure 9. Active removal of oxidized dGTP residues from the intracellular pool by MutT-catalysed cleavage.

of a DNA glycosylase specific for oxidatively damaged DNA (Demple & Linn 1980; Breimer & Lindahl 1980). Another form of oxidative DNA damage involves fragmentation or saturation of the imidazole ring of purines; yet another DNA glycosylase specializes in the removal of such lesions (Chetsanga & Lindahl 1979; Boiteux *et al.* 1987). The most interesting substrate of the latter enzyme is a G with a saturated imidazole ring, 8-oxo-7,8-dihydro-dG (Kasai & Nishimura 1984). This residue in DNA readily assumes a *syn* conformation, which can base-pair with an A residue and thus result in transversion mutations on replication (figure 8). This is the major form of mutagenic base damage caused by active oxygen and occurs so frequently that a number of cellular correction pathways have evolved to minimize oxygen-induced mutagenesis. The defence mechanisms are not restricted to base excision-repair of altered DNA. Because the dGTP precursor is also susceptible to oxidation, the cellular dGTP pool is kept smaller than that of other deoxynucleoside triphosphates, and oxidized dGTP residues are rapidly cleaved by a specific nucleoside triphosphatase, the MutT protein (figure 9) (Maki & Sekiguchi 1992). Absence of MutT protein leads to greatly increased spontaneous mutation frequencies.

In eukaryotic cells, the delegation of oxygen metabolism to mitochondria has been an excellent strategy for protection of DNA in the cell nucleus from damage induced by oxygen free radicals. The small mitochondrial DNA molecules, however, may be susceptible to more damage; mitochondria are known to have a high intrinsic mutation rate.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

6. DNA ADDUCTS GENERATED BY REACTIVE METABOLITES

The effects of water and active oxygen on DNA structure are now fairly well understood, but there are many other reactive small molecules in cells; accidental formation of covalent DNA adducts by those compounds is a relatively unexplored field of research. Certain cells, such as macrophages, produce reactive compounds (NO, oxygen free-radical species) that are damaging to adjacent cells. An example of an intrinsic DNA-damaging agent is *S*-adenosylmethionine, which is the reactive methyl group donor in many cellular transmethylation reactions. Owing to its reactivity, *S*-adenosylmethionine is a weak alkylating agent; its main effect on DNA is the generation of cytotoxic 3-methyladenine lesions (Rydberg & Lindahl 1982). Approximately 600 such residues may be formed per generation in a proliferating mammalian cell. The same rate of DNA alteration would be achieved by continuous exposure of cells to 20 nM methyl methanesulphonate, a strong alkylating agent. Presumably for this reason, all living cells have an efficient DNA glycosylase that removes 3-methyladenine from DNA; this enzyme also contributes to the relative resistance of cells to killing by external alkylating agents (Riazuddin & Lindahl 1978; Laval 1977).

Another type of endogenous DNA damage involves the generation of exocyclic adducts on bases, for example the formation of a third ring on a purine. These adducts can result from exposure to endogenous lipid peroxidation products such as acrolein, crotonaldehyde and malondialdehyde (Nath *et al.* 1996; Chaudhary *et al.* 1994) or from oxidized hydroquinone (Chenna *et al.* 1995). The exocyclic adducts are again removed by base excision-repair, and at least one DNA glycosylase seems to be involved (Hang *et al.* 1996). A current list of DNA glycosylases is shown in table 1; with the exception of one enzyme found only in certain very UV-resistant microorganisms, they are all primarily concerned with the excision of endogenous DNA lesions.

A highly mutagenic DNA lesion of endogenous origin is O^6-methylguanine, which forms base-pairs with T rather than with C residues. The lesion is of particular interest for two reasons. First, it is not produced in significant amounts by alkylating agents acting by an S_N2 mechanism such as *S*-adenosylmethionine, so the occurrence of this lesion indicates the generation by cellular metabolism of small amounts of endogenous nitrosamides. Bacterial mutants deficient in repair of the lesion exhibit an increased spontaneous mutation frequency, particularly in starving cells, so O^6-methylguanine is definitely generated endogenously (Rebeck & Samson 1991). The endogenous alkylating agent has not been identified, but enzymatic nitrosation of a cellular catabolite apparently plays a role in its formation (Taverna & Sedgwick 1996). Second, this DNA adduct is *not* removed by a DNA glycosylase; instead, the O^6 methyl group is transferred to a protein cysteine residue, causing direct reversal of the damage without excision of a nucleotide residue in a unique form of DNA repair (Olsson & Lindahl 1980).

The abundance of DNA repair enzymes acting on endogenous DNA lesions indicates that the main driving force in the evolution of DNA repair has been the inherent lability of DNA under conditions *in vivo*. The only exception to this rule is the extensively investigated repair of DNA damage caused by exposure to ultraviolet light (Wood 1996). The recently completed DNA sequences of small bacteria show that repair enzymes for several types of endogenous DNA lesions as well as for sunlight-induced damage were already present at an early stage of microbial evolution (Fraser *et al.* 1995). Exposure of living cells to ionizing radiation or environmental chemical mutagens has been too recent or insignificant to result in selection of specific DNA repair processes, although microorganisms may be able to induce certain repair functions in response to external DNA-damaging agents. Human cellular defence capacity against such agents, however, is probably only a side effect of the continuous correction of intrinsic DNA damage. In

Table 1. *DNA glycosylases*

Excised base	*E. coli*		Human cells
	presence	gene	
uracil-	+	*ung*	+
thymine glycol-/urea-/other oxidized pyrimidine derivatives	+[a]	*nth* and *nei*	+
8-oxoguanine- (opposite cytosine)/formamidopyrimidine-	+	*fpg* (*mutM*)	+
adenine-(opposite 8-oxoguanine)	+	*mutY*	+
3-methyladenine-(also hypoxanthine-/ethenoadenine-/5-formyluracil-)	+[b]	*tag* and *alkA*	+
thymine-(opposite guanine)	?	–	+
hydroxymethyluracil-	–	–	+
cyclobutane pyrimidine dimer-	(+)[c]	*T4 denV*	–

[a] Two distinct enzymes with slightly different substrate spectra are present.
[b] One constitutively expressed enzyme, and one inducible enzyme with broader substrate specificity.
[c] Only in bacteriophage T4-infected cells.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

consequence, low exposure to most environmental mutagens might be expected to have little or no biological effect: a small or moderate increase in DNA damage above the considerable basal level of endogenous lesions can still be processed effectively by the versatile DNA repair systems. On the other hand, mutations in genes encoding the repair enzymes themselves might prevent correction of certain spontaneous DNA lesions and cause a cellular mutator phenotype. In higher cells, this would be equivalent to the loss of a tumour suppressor gene and result in genomic instability and increased propensity for malignant conversion (Loeb 1991).

7. ANTEDILUVIAN DNA

I conclude this lecture with some comments on a topic that may be of lesser scientific importance than endogenous DNA damage in living cells but is of current public interest; that is, the recovery and sequencing of DNA from fossils. The polymerase chain reaction (PCR) made it possible to retrieve and amplify minute amounts of DNA fragments from ancient material, and sensational preliminary reports appeared in the leading scientific journals. The subsequent realization that several of the apparent recoveries of DNA sequences, such as those from dinosaur bones or chloroplasts in magnolia leaves many million years old, were due to contamination rather than authentic old DNA has dampened the initial enthusiasm, and a mood of experimental stringency with extensive controls is now characteristic of the best groups in the field (Stoneking 1995). There is general agreement that short DNA sequences can be retrieved from occasional fossils 10^4–10^5 years old, although most experiments are unsuccessful and recovery is very difficult or impossible with material from warm climates (Höss *et al.* 1996). The identification by two groups of short authentic DNA sequences from mammoth fossils 10^5 years old from permafrost deposits in Siberia is convincing (Höss *et al.* 1994; Hagelberg *et al.* 1994). These fragments of mitochondrial DNA sequences of mammoth origin are at present the oldest known verified DNA sequences.

The most interesting recent development in this field is that suitable references for DNA degradation are being developed, in particular by Svante Pääbo and co-workers. Thus, racemization of L-aspartic acid residues in proteins to a mixture of the L- and D-enantiomers is a slow hydrolytic reaction proceeding at a very similar rate to DNA depurination. It is therefore not possible to retrieve DNA sequences from fossil material where extensive amino-acid racemization has occurred; several early reports on ancient DNA can be dismissed for this reason (Poinar *et al.* 1996). A great advantage of this procedure is that only milligram amounts of an irreplaceable fossil are required for chromatographic analysis of amino-acid racemization. A similar sensitive detection method for oxidative damage is not yet available. Direct DNA base analysis by gas chromatography – mass spectrometry on the minute amounts of DNA recovered from several grams of fossil bones has, however, shown that there is an inverse relation between the amount of oxidation of DNA pyrimidine residues to hydantoins and the ability to amplify DNA sequences by PCR (Höss *et al.* 1996). This result could be expected because the 5-membered hydantoin rings, which are generated as major and stable active oxygen products by saturation and condensation of DNA pyrimidine rings (Breimer & Lindahl 1985), are not likely to be copied by a DNA polymerase during PCR.

A remaining enigma in the ancient DNA field is whether it will be possible to recover authentic DNA sequences from million-year-old insects entombed in amber. Initial reports on apparently successful DNA retrieval by PCR have remained unconfirmed for several years, in spite of attempts by other laboratories. Amber fossils are protected from water and do not undergo significant aspartic-acid racemization (Poinar *et al.* 1996), so DNA depurination may be similarly suppressed. However, two major problems remain unanswered. Desiccation of DNA generates double-strand breaks, and DNA fragmentation caused by dehydration is a major environmental threat to microorganisms unable to sporulate (Mattimore & Battista 1996). It is not known whether cellular desiccation by amber entombment similarly causes extensive decay of the DNA backbone as a function of time. Furthermore, amber is somewhat permeable to oxygen, and it is not known whether the DNA pyrimidine residues have been oxidized to hydantoins during a time frame of several million years. In the absence of experimental data, neither of these objections necessarily excludes the possibility that short DNA sequences might be amplified from ancient insects in amber. Perhaps an analytic reference procedure, such as a microtest for slow oxidation of methionine or tryptophan residues in proteins, could be developed, analogous to the aspartic-acid racemization test for hydrolytic damage. This might allow an evaluation in quantitative terms as to whether it seems realistic to attempt to retrieve DNA sequences from ancient amber fossils.

REFERENCES

Baross, J. A. & Deming, J. W. 1983 Growth of 'black smoker' bacteria at temperatures of at least 250 °C. *Nature, Lond.* **303**, 423–426.

Boiteux, S., O'Connor, T. R. & Laval, J. 1987 Formamidopyrimidine-DNA glycosylase of *Escherichia coli*: cloning and sequencing of the *fpg* structural gene and overproduction of the protein. *EMBO J.* **6,** 3177–3183.

Breen, A. P. & Murphy, J. A. 1995 Reactions of oxyl radicals with DNA. *Free Rad. Biol. Med.* **18**, 1033–1077.

Breimer, L. & Lindahl, T. 1980 A DNA glycosylase from *E. coli* that releases free urea from a polydeoxyribonucleotide containing fragments of base residues. *Nucl. Acids Res.* **8**, 6199–6211.

Breimer, L. & Lindahl, T. 1985 Thymine lesions produced by ionizing radiation in double-stranded DNA. *Biochemistry* **24**, 4018–4022.

Brown, D. M. & Todd, A. R. 1952 Some observations on the structure and chemical behaviour of the nucleic acids. *J. Chem. Soc.* **1952** (1), 52–58.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

Chaudhary, A. K., Nokubo, M., Reddy, G. R., Yeola, S. N., Morrow, J. D., Blair, I. A. & Marnett, L. J. 1994 Detection of endogenous malondialdehyde-deoxyguanosine adducts in human liver. *Science, Wash.* **265**, 1580–1582.

Chenna, A., Hang, B., Rydberg, B., Kim, E., Pongracz, K., Bodell, W. J. & Singer, B. 1995 The benzene metabolite *p*-benzoquinone forms adducts with DNA bases that are excised by a repair activity for human cells that differs from an ethenoadenine glycosylase. *Proc. natn. Acad. Sci. U.S.A.* **92**, 5890–5894.

Chetsanga, C. J. & Lindahl, T. 1979 Release of 7-methylguanine residues whose imidazole rings have been opened from damaged DNA by a DNA glycosylase from *E.coli*. *Nucl. Acids Res.* **6,** 3673–3684.

Davis, B. D. 1958 On the importance of being ionized. *Archs Biochem. Biophys.* **78**, 497–509.

Deming, J. W. & Baross, J. A. 1993 Deep-sea smokers: windows to a subsurface biosphere. *Geochim. cosmochim. Acta* **57**, 3219–3230.

Demple, B. & Linn, S. 1980 DNA N-glycosylases and UV repair. *Nature, Lond.* **287,** 203–208.

Dianov, G. & Lindahl, T. 1994 Reconstitution of the DNA base excision-repair pathway. *Curr. Biol.* **4**, 1069–1076.

Duncan, B. K. & Miller, J. 1980 Mutagenic deamination of cytosine residues in DNA. *Nature, Lond.* **287**, 560–561.

Fraser, C. M. *et al.* 1995 The minimal gene complement of *Mycoplasma genitalium*. *Science, Wash.* **270**, 397–403.

Frederico, L. A., Kunkel, T. A. & Shaw, B. R. 1990 A sensitive genetic assay for the detection of cytosine deamination: determination of rate constants and activation energy. *Biochemistry* **29**, 2532–2537.

Friedberg, E. C., Walker, G. C. & Siede, W. 1995 *DNA repair and mutagenesis*. Washington, D.C.: ASM Press.

Hagelberg, E., Thomas, M. G., Cook, C. E. Jr, Sher, A. V., Baryshnikov, G. F. & Lister, A. M. 1994. DNA from ancient mammoth bones. *Nature, Lond.* **370**, 333–334.

Hang, B., Chenna, A., Rao, S. & Singer B. 1996 1,N^{6-}Ethenoadenine and 3,N^{4-}ethenocytosine are excised by separate human DNA glycosylases. *Carcinogenesis* **17**, 155–157.

Höss, M, Pääbo, S. & Vereshchagin, N. K. 1994 Mammoth DNA sequences. *Nature, Lond.* **370**, 333.

Höss, M., Jaruga, P., Zastawny, T. H., Dizdaroglu, M. & Pääbo, S. 1996 DNA damage and DNA sequence retrieval from ancient tissues. *Nucl. Acids Res.* **24**, 1304–1307.

Kasai, H. & Nishimura, S. 1984 Hydroxylation of deoxyguanosine at the C-8 position by ascorbic acid and other reducing agents. *Nucl. Acids Res.* **12**, 2137–2145.

Kubota, Y., Nash, R. A., Klungland, A., Schär, P., Barnes, D. E. & Lindahl, T. 1996 Reconstitution of DNA base excision-repair with purified human proteins: Interaction between DNA polymerase β and the XRCC1 protein. *EMBO J.* (In the press.)

Laval, J. 1977 Two enzymes are required for strand incision in repair of alkylated DNA. *Nature, Lond.* **269,** 829–831.

Li, E., Bestor, T. H. & Jaenisch, R. 1992 Targeted mutation of the DNA methyltransferase gene results in embryonic lethality. *Cell* **69,** 915–926.

Lindahl, T. 1974 An N-glycosidase from *E. coli* that releases free uracil from DNA containing deaminated cytosine residues. *Proc. natn. Acad. Sci. U.S.A.* **71,** 3649–3653.

Lindahl, T. 1993 Instability and decay of the primary structure of DNA. *Nature, Lond.* **362**, 709–715.

Lindahl, T. & Andersson, A. 1972 Rate of chain breakage at apurinic sites in double-stranded DNA. *Biochemistry* **11**, 3618–3623.

Lindahl, T., Ljungquist, S., Siegert, W., Nyberg, B. & Sperens, B. 1977 DNA N-glycosidases. Properties of uracil-DNA glycosidase from *E. coli*. *J. biol. Chem.* **252**, 3286–3294.

Lindahl, T. & Nyberg, B. 1972 Rate of depurination of native DNA. *Biochemistry* **11**, 3610–3618.

Lindahl, T. & Nyberg, B. 1974 Heat-induced deamination of cytosine residues in deoxyribonucleic acid. *Biochemistry* **13,** 3405–3410.

Lindahl, T., Satoh, M.S. & Dianov, G. 1995 Enzymes acting at strand interruptions in DNA. *Phil. Trans. R. Soc. Lond.* B **347**, 57–62.

Ljungquist, S. & Lindahl T. 1974 A mammalian endonuclease specific for apurinic sites in double-stranded DNA. *J. biol. Chem.* **249,** 1530–1540.

Loeb, L. A. 1991 Mutator phenotype may be required for multistage carcinogenesis. *Cancer Res.* **51**, 3075–3079.

Maki, H. & Sekiguchi, M. 1992 MutT protein specifically hydrolyses a potent mutagenic substrate for DNA synthesis. *Nature, Lond.* **355**, 273–275.

Mattimore, V. & Battista, A. 1996 Radioresistance of *Deinococcus radiodurans*: Functions necessary to survive ionizing radiation are also necessary to survive prolonged desiccation. *J. Bact.* **178**, 633–637.

Mol, C. D., Arvai, A. S., Slupphaug, G., Kavli, B., Alseth, I., Krokan, H. E. & Tainer, J. A. 1995 Crystal structure and mutational analysis of human uracil-DNA glycosylase: Structural basis for specificity and catalysis. *Cell* **80**, 869–878.

Nath, R. G., Ocando, J. E. & Chung, F.-L. 1996 Detection of 1, N^2-propanodeoxyguanosine adducts as potential endogenous DNA lesions in rodent and human tissues. *Cancer Res.* **56**, 452–456.

Neddermann, P. & Jiricny, J. 1993 The purification of a mismatch-specific thymine-DNA glycosylase from HeLa cells. *J. biol. Chem.* **268,** 21218–21224.

Nielsen, P. E. 1995 DNA analogues with nonphosphodiester backbones. *A. Rev. Biophys. biomolec. Struct.* **24**, 167–183.

Olsson, M. & Lindahl, T. 1980 Repair of alkylated DNA in *E. coli*: Methyl group transfer from O^6-methylguanine to a protein cysteine residue. *J. biol. Chem.* **255**, 10569–10571.

Paquette, Y., Crine, P. & Verly, W. G. 1972 Properties of the endonuclease for depurinated DNA from *Escherichia coli*. *Can. J. Biochem.* **50,** 1199–1209.

Poinar, H. N., Höss, M., Bada, J. L. & Pääbo, S. 1996 Amino acid racemization and the preservation of ancient DNA. *Science, Wash.* **272**, 864–866.

Rebeck, G. W. & Samson, L. 1991 Increased spontaneous mutation and alkylation sensitivity of *Escherichia coli* strains lacking the *ogt* O^6-methylguanine DNA repair. *J. Bact.* **173,** 2068–2076.

Reichard, P. 1995 To be there when the picture is painted. *A. Rev. Biochem.* **64**, 1–28.

Riazuddin, S. & Lindahl, T. 1978 Properties of 3-methyladenine-DNA glycosylase from *E. coli*. *Biochemistry* **17**, 2110–2118.

Rideout, W. M., Coetzee, G. A., Olumi, A. F. & Jones, P. A. 1990 5-Methylcytosine as an endogenous mutagen in the human LDL receptor and p53 genes. *Science, Wash.* **249,** 1288–1290.

Rydberg, B. & Lindahl, T. 1982 Nonenzymatic methylation of DNA by the intracellular methyl group donor S-adenosyl-L-methionine is a potentially mutagenic reaction. *EMBO J.* **1,** 211–216.

Savva, R., McAuley-Hecht, K., Brown, T. & Pearl, L. 1995 The structural basis of specific base-excision repair by uracil-DNA glycosylase. *Nature, Lond.* **373**, 487–493.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

Shapiro, R. & Klein, R. S. 1966 The deamination of cytidine and cytosine by acidic buffer solutions. Mutagenic implications. *Biochemistry* **5**, 2358–2362.

Sobol, R. W., Horton, J. K., Kühn, R., Gu, H., Singhal, R. K., Prasad, R., Rajewsky, K. & Wilson, S. H. 1996 Requirement of mammalian DNA polymerase-β in base-excision repair. *Nature*, *Lond.* **379,** 183–186.

Stoneking, M. 1995 Ancient DNA: How do you know when you have it and what can you do with it? *Am. J. hum. Genet.* **57**, 1259–1262.

Taverna, P. & Sedgwick, B. 1996 Generation of an endogenous DNA methylating agent by nitrosation in *Escherichia coli*. *J. Bacteriol.* **178**, 5105–5111.

Trent, J. D., Chastain, R. A. & Yayanos, A. A. 1984 Possible artefactual basis for apparent bacterial growth at 250 °C. *Nature*, *Lond.* **307**, 737–740. [Reply from Baross, J. A. & Deming, J. W., *Nature*, *Lond.* **307**, 740.]

Watson, J. D. & Crick, F. H. C. 1953 Genetical implications of the structure of deoxyribonucleic acid. *Nature*, *Lond.* **171**, 964–967.

Westheimer, F. H. 1987 Why Nature chose phosphates. *Science*, *Wash.* **235**, 1173–1178.

Wood, R. D. 1996 DNA repair in eukaryotes. *A. Rev. Biochem.* **65**, 135–167.

Zoltewicz, J. A., Clark, F. O., Sharpless, T. W. & Grahe, G. 1970 Kinetics and mechanism of the acid-catalyzed hydrolysis of some purine nucleosides. *J. Am. Chem. Soc.* **92**, 1741–1750.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES