

THE CHEMICAL BASIS OF BIOLOGICAL MUTATION

D. M. BROWN

University Chemical Laboratory, Cambridge, England

Since the realisation that the genetic material in living organisms is deoxyribonucleic acid (DNA) [or ribonucleic acid (RNA) in RNA-containing viruses], the way has, in principle, been open for an attack on the question of the chemical basis of biological mutation. In this paper I wish to draw some conclusions respecting one mutagen, but, more importantly, try to demonstrate the nature of the problems involved and the kind of questions that must be asked.

As background, it is necessary first, to recall that the understanding of the genetics of micro-organisms and the molecular basis thereof, has expanded enormously in recent years; this is particularly so of bacterial virus (bacteriophage) genetics¹. The T even phages contain DNA in the form of a linear double-helical molecule containing of the order of 200 000 base-pairs. Using purely genetic methods Benzer^{2, 3} was able to make a fine structure map of the *rII*-gene in T4, to which much more detail has since been added⁴. It is a map consisting of many points or regions in linear array and for each point or region, there is a T4r mutant (defective in the function that the *r* region controls). Map distances are, in fact recombinational frequencies derived from crossing mutants pair-wise in a suitable host bacterium. Although the scale may be badly distorted at short map distances⁵, nevertheless the view has become established with considerable certainty that the points on the map represent in many cases single base-pairs and the regions more or less extensive parts of the bacteriophage DNA. Of the several hundred *rII* mutants (and the many thousand in other genetic material) that have been studied some arose spontaneously and others were chemically induced. From the standpoint of chemical mutagenesis bacteriophages are valuable experimental material since they can be treated with the reagent in the test-tube and then, in the absence of excess reagent, replicated in the host bacterium allowing mutations, if any, to be expressed.

As a result of genetic and chemical considerations, several types of mutational change have been recognised. These are listed in *Table 1* and have

Table 1. Types of mutational change

-
1. Large deletions, insertions.
 2. Small deletions, insertions
(frame shift mutants).
 3. Substitutions
 - (i) transitions
 - (ii) transversions.
-

been discussed in greater detail elsewhere^{4, 6, 7}. Some mutants result from deletions or insertions of relatively large pieces in the DNA chain². Others come from the loss or gain of one or a small number of base-pairs (frame-shift mutants)⁸. A third class evidently comes about by the substitution of one base-pair by another. It is clear that, as Freese was first to point out⁹, substitution can occur in several ways as shown in *Figure 1*. Two sub-classes

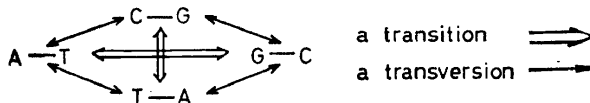


Figure 1. Base-pair substitution.

can be discerned; transitions in which a purine-pyrimidine pair in the standard type organism is replaced by the other purine-pyrimidine pair in the mutant, and transversions in which the replacement is by a pyrimidine-purine pair. Mutants have been recognised which fall into each of these categories. It would be impossible, here, to detail the evidence (see refs. 4, 7). Suffice it to say that all classes arise spontaneously; large deletions can be induced by nitrous acid by a mechanism which is still obscure¹⁰; frame-shift mutants can be induced by phage replication in presence of certain acridine derivatives^{11, 12}. Transversions appear spontaneously¹³ but have not been, recognisably, induced or reverted chemically, a fact that poses an interesting problem for the future. Transitions are evidently the simplest form of genetic alteration—at least conceptually. They are induced by a variety of mutagens. For example, some nitrous acid mutations arise from $AT \rightarrow GC$, and others from $GC \rightarrow AT$ substitutions¹⁴, a reasonable explanation being that the primary chemical events are deamination of adenine and cytosine respectively. It is important to recognise that the chemical event is only expressed as a mutational change when the DNA is replicated. There are, in fact, two questions that have to be asked about the mechanism of chemical mutagenesis: what is the initial chemical event and why is the modified base so produced replicated erroneously? From this standpoint it would be valuable to have a mutagen which was highly specific in that its action was limited to one of the four bases present in DNA. It could then be used in reversion studies to establish the nature of the mutational change produced by another mutagen; it would, in this sense, provide a primary standard. Nitrous acid, in this respect, does not qualify. However, hydroxylamine and methoxyamine¹⁵ appear to be such mutagens, the former having been the subject of a large number of studies^{4, 16-19}. One noticeable feature is that $T4r$ mutants induced by hydroxylamine are not, in general, reverted to the standard type by hydroxylamine¹⁶. This indicates, at once, a high degree of base-specificity which has been confirmed in more recent studies using the small single-stranded DNA bacteriophage, S-13¹⁹.

In the remainder of my talk I wish to give you the evidence that this base is cytosine and discuss the nature of the chemical modification by hydroxylamine and why the modification is mutagenic.

THE CHEMICAL BASIS OF BIOLOGICAL MUTATION

First of all we should be quite clear about a major difficulty inherent in this sort of study. When a phage suspension is mutagenized the chemical event leading to the mutation may effect only one base out of perhaps a thousand base-pairs in the relevant part of the genome. The event, then, cannot be directly observed. Having consideration for the fact that most chemical reactions are accompanied by side-reactions, it cannot be said that the event leading to mutation is necessarily or even probably the main reaction. We may add to this the fact that many events which are potentially mutagenic may not be phenotypically expressed as a consequence of the degeneracy of the genetic code or they may be lethal. In other words a bacteriophage system, although much simpler than a mammalian one, is yet too complex for a complete mechanistic study; a model replicating system is needed which will link the chemical and biological observations and in which quantitative relationships can be established.

Hydroxylamine does not react with the purine bases, but it does so, around neutrality, with cytosine, and to a lesser extent with uracil; the latter is more reactive under basic conditions²⁰⁻²². The pH optima for mutagenesis in T4 and for reaction with cytosine and its N¹-derivatives show some correlation¹⁸. The reactions with cytosine are shown in somewhat simplified form in *Figure 2*. The present view is that two simultaneous reactions occur²³⁻²⁵, an

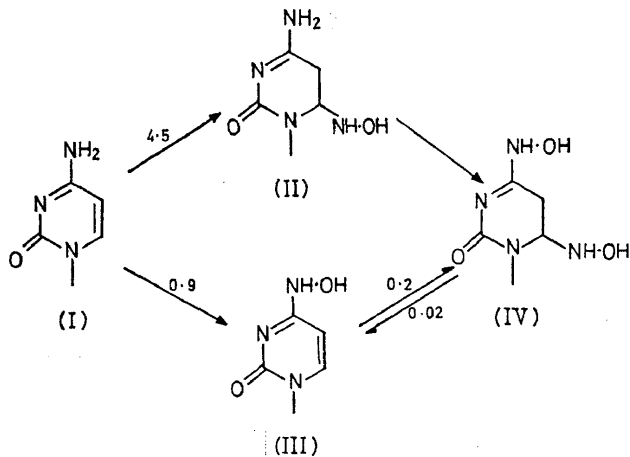


Figure 2. First order rate constants (h^{-1}) measured at 3.5 M NH_2OH ; pH 6.5; 37°C (data from Brown and Hewlins²⁴)

addition to the 5,6-double bond and a nucleophilic displacement of the amino by a hydroxylamino group. This is clearly established by kinetic studies which show that the reactions which lead to saturation of the 5,6-double bond are about five times as fast as the displacement reaction. The mono-adduct (II) does not appear to have an appreciable life-time so that (IV) rapidly accumulates. However, in polycytidylic acid, while the rates of the initial reactions are about the same, the mono-adduct (II) does build up to a low steady state concentration²⁶. By contrast, 5-hydroxymethylcytosine gives no evidence of adduct formation^{24, 27}. Although the matter has not

been entirely clarified it appears that only the displacement reaction is important. It will be recalled that the DNAs of T even bacteriophages contain hydroxymethylcytosine while those of phage S 13 and *B. subtilis* contain cytosine. All are mutagenized by hydroxylamine. We therefore have to consider the adduct (II) and the exchange product (III) as possible sources of replication error. The di-adduct (IV), and other transformation (e.g. hydrolysis) products of (II) and (III) are not considered to be serious contenders for this role²⁸, as I hope to make clear shortly.

At this point, I wish to introduce some enzymatic studies which are important to the development of the problem. When polycytidylic acid (poly-C) is used as template with guanosine triphosphate (GTP) as substrate with the enzyme RNA polymerase from *Micrococcus lysodeiکتicus*, the complementary polynucleotide, poly G, is formed²⁹ (Figure 3). When the copolymer poly-CU is used as template poly-GA is the product, given the requisite triphosphate substrates. Other triphosphates (e.g. CTP, UTP) are no

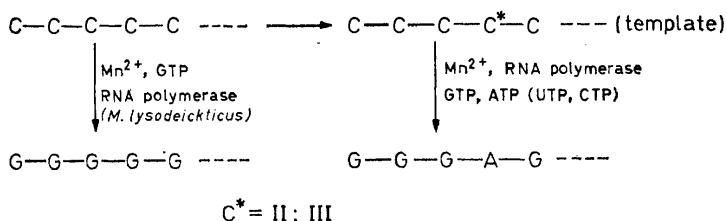


Figure 3

incorporated. The enzyme, in fact, has a high degree of specificity. When the copolynucleotide containing cytosine and N⁴-hydroxycytosine (N⁴-OHC) residues (20:1) is synthesized (by enzymatic copolymerization of the corresponding ribonucleoside-5' diphosphates) and used in this system it is found that adenylate (but not uridylylate or cytidylate) residues are incorporated,

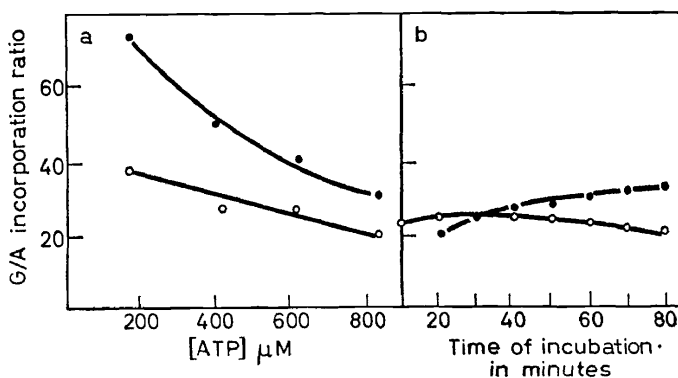


Figure 4. Ratio of incorporation of GMP and AMP directed by poly CU (20:1) ○—○, and poly (C + N⁴OHC) (20:1) ●—●, in presence of *M. lysodeiکتicus* RNA polymerase, 0.0033 M Mn²⁺, and (a) 833 μM GTP, 24°C, 60 min, or (b) 833 μM GTP, 833 μM ATP, 24°C. (Data from G. R. Banks *et al.*³⁰)

i.e. that poly-GA is again formed³⁰. Thus this synthetic copolymer emulates quite closely a poly CU of similar composition. Comparisons are made in *Figure 4*, which show the G/A ratios of the synthesized polymers formed in the presence of RNA polymerase, both as a function of time of synthesis and of ATP concentration. We may, I think, draw the clear conclusion that N⁴-hydroxycytosine residues in the template are directing adenylate incorporation with considerable efficiency. Nearest neighbour frequencies show that the adenylate residues are present internally in the poly-G. The question has been argued in more detail elsewhere³¹, but, if the polyribonucleotide—RNA polymerase system—is a valid model for bacterial or phage replication then the experiments provide good ground for believing that the modification of hydroxymethylcytosine residues which leads to the N⁴-hydroxy derivative should in turn account for the observed C→T transitions in hydroxylamine mutagenesis.

Essentially the same type of enzymatic study can be made in order to determine the nature of the base which is complementary to the hydroxylamine or methoxyamine adduct (III). These experiments carried out by Dr. J. H. Phillips were much more arduous than the above and have already been published in detail^{28, 32}. Briefly, as is seen in *Figure 5*, for small extents of methoxyamine treatment, longer treatment leads to greater adenylate incorporation, when the modified templates are used in the polymerase system. Other experiments using [¹⁴C]-methoxyamine permit the small amount of adduct (II), formed in the poly-C template to be measured. The relationship between the number of residues altered in the template (at

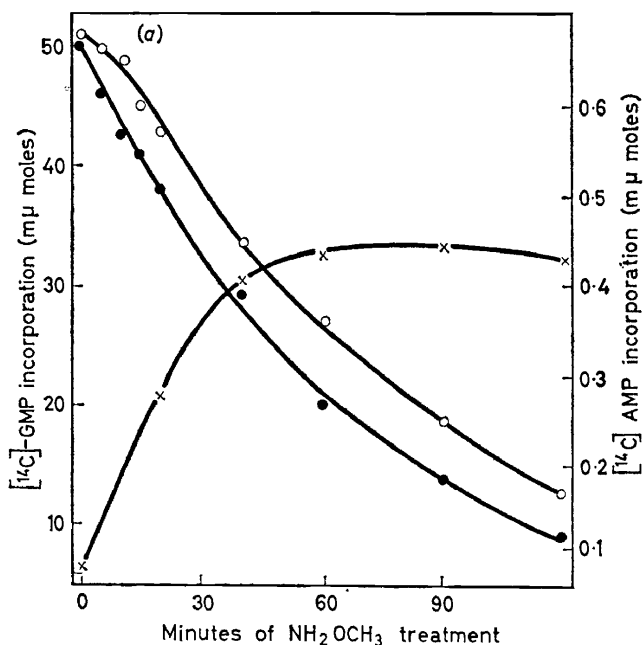


Figure 5(a)

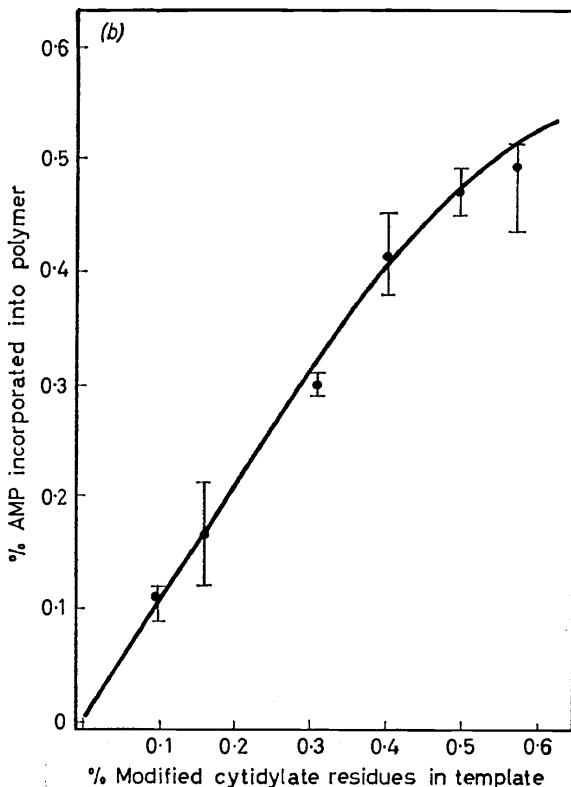


Figure 5(b)

Figure 5. (a) Incorporation of GMP in the presence (○) and absence (●) of ATP_i and incorporation of AMP in the presence of GTP (×), directed by poly-C pre-treated with methoxyamine (2.0 M, pH 5.5, 37°C).

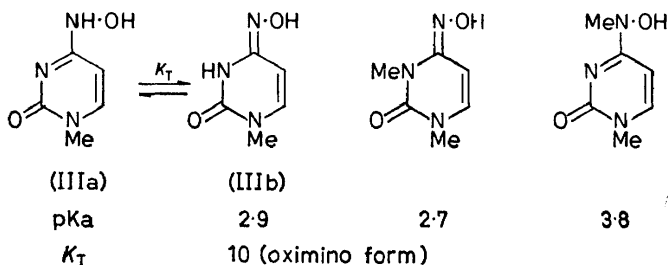
(b) Percentage of adenylate found in poly-AG synthesised using poly C templates pre-treated with methoxyamine. (Data from Phillips *et al.*³²)

varying times of reaction) and the number of adenylate residues incorporated using these modified templates may then be established. The answer is rather clear-cut; every modified cytosine residue directs the incorporation of one adenylate residue. No such observation is made when hydroxylamine treated poly-A or poly-U are used as templates. The specific effect is only obtained with poly-C. Thus, so far as evidence from this *in vitro* system is concerned, the hydroxylamine addition reaction should be mutagenic and account for the transitions observed in S 13¹⁹ and in the *B. subtilis* transforming system¹⁶.

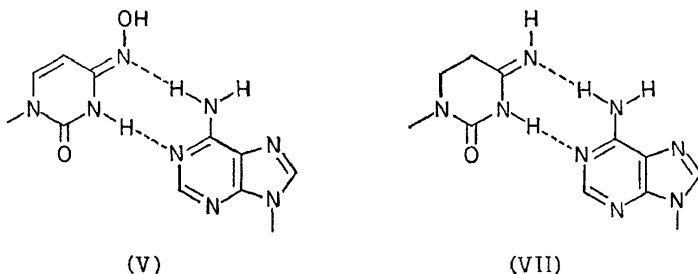
What explanation can we advance for these results? In their simplest terms the experiments say that both the exchange product (II) and the adduct (III) resemble uracil (or thymine) in their base-pairing potentialities. The simplest explanation would be that (II) and (III) correspond to uracil and not to cytosine in tautomeric form³³. The chemistry of N⁴-hydroxycytosine and a number of related compounds has been investigated

THE CHEMICAL BASIS OF BIOLOGICAL MUTATION

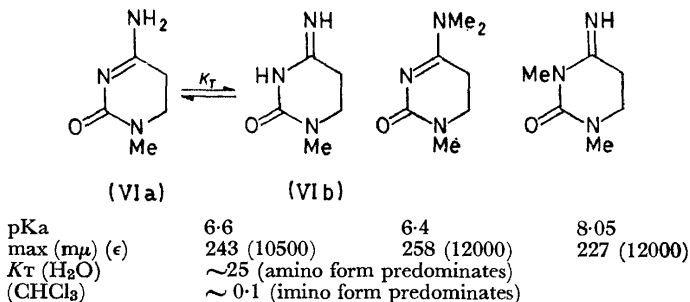
in some detail²⁴ and the conclusion is quite clear; N⁴-hydroxy (and methoxy)-cytosine and its N¹-derivatives exist in the oximino form (III b), as judged from u.v. and i.r. spectra and from pKa measurements. From the latter a tautomeric constant of 10 in aqueous solution may be derived. One can go a little further. A number of workers³⁴⁻³⁶ have studied the interaction of



derivatives of the nucleic acid bases in non-polar solvents, using i.r. methods. These conditions favour inter-base hydrogen-bonding and diminish stacking interactions. The conclusion emerges that the Watson-Crick base-pairs have higher formation constants than the other possible ones, including those due to self-association. When the interactions between N⁴-methoxy-1-methylcytosine, and 9-ethyladenine were investigated by Dr. Hewlins in carbon tetrachloride solution, clear evidence for base-pairing was found²⁴. The value for the association constant obtained was 120 l. mole⁻¹ as compared, for example, with 40 l. mole⁻¹ found for the adenine-adenine self-interaction (cf. refs. 34, 36). No interaction with 9-ethylguanine was observed. Thus both at the chemical and the enzymatic level N⁴-hydroxycytosine residues show a strong resemblance to uracil (or thymine), and this suggests that replication error due to mis-pairing of the type (V) should occur.

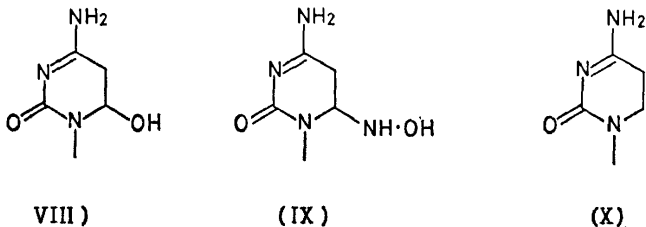


The position with respect to the hydroxylamine adduct (II) is more difficult. As yet it has not been possible to isolate it. We considered that 5,6-dihydrocytosine derivatives (VI) would serve as useful models. In general these are more readily prepared by ring synthesis than by the reduction of cytosines³⁷⁻³⁹. It is clear from the data³⁹ shown above that 1-methyl-5,6-dihydrocytosine is an amino-compound (VIa) in water. The tautomeric constant of 25 should, however, be compared with the value of 10⁵ for cytosine⁴⁰, a very large shift toward the imino-form (VIb) which, in fact,



is the predominant species in chloroform. It is probable that the hydroxylamine adduct (II) will have a tautomeric constant of the same order of magnitude. If the conditions at the point of replication are essentially hydrophobic with a low dielectric constant then, probably, the replication error due to (II) can be ascribed to mis-pairing in its imino-tautomeric state, with adenine as in (VII). Such a conclusion must be very tentative indeed at present. An alternative hypothesis, that the protonated adduct mis-pairs with adenine, has been discussed⁴². [The pKa of (II) is probably quite close to 5.6, the measured value for 6-hydroxy-5,6-dihydrocytosine (VIII)⁴¹]. It is equally tentative and discussion is rendered difficult since the question of the pKa values of individual bases in native DNA is a matter of great uncertainty^{43, 44}.

Nevertheless, although a final explanation may still be to find, the formation of reduced residues in DNA—(VIII) by photochemical hydration^{45, 46},



(IX) by hydroxylamine addition and possibly (X) by hydrazine reduction⁴⁷—appears to be potentially mutagenic.

ACKNOWLEDGEMENT

I wish to acknowledge my indebtedness to my colleagues who have taken part in the work I have described—Dr. R. Schell, Dr. J. H. Phillips, Dr. G. R. Banks, Dr. M. J. E. Hewlins and Professor L. Grossman.

References

- ¹ G. S. Stent. *Molecular Biology of Bacterial Viruses*. W. H. Freeman, San Francisco, Calif. (1963).
- ² S. Benzer in *The Chemical Basis of Heredity* (Ed., W. D. McElroy and B. Glass), Johns Hopkins Press, Baltimore, p. 70, 1957.
- ³ S. Benzer. *Proc. U.S. Natl. Acad. Sci.* **41**, 344 (1955); **45**, 1607 (1959); **47**, 403 (1961).
- ⁴ D. R. Krieg in *Progress in Nucleic Acid Research*, (Eds. J. N. Davidson and W. E. Cohn), Academic Press, New York, **2**, 125 (1963).

THE CHEMICAL BASIS OF BIOLOGICAL MUTATION

- ⁵ I. Tessman. *Genetics* **51**, 63 (1965).
- ⁶ E. Freese in *Molecular Genetics*, Part 1 (Ed. J. M. Taylor), Academic Press, New York, p. 207 (1963).
- ⁷ L. E. Orgel. *Advan. Enzymol.* **27**, 289 (1965).
- ⁸ S. Brenner, L. Barnett, F. H. C. Crick and A. Orgel. *J. Mol. Biol.* **3**, 121 (1961).
- ⁹ E. Freese. *J. Mol. Biol.* **1**, 87 (1959); *Proc. U.S. Natl. Acad. Sci.* **45**, 622 (1959).
- ¹⁰ I. Tessman. *J. Mol. Biol.* **6**, 442 (1962).
- ¹¹ A. Orgel and S. Brenner. *J. Mol. Biol.* **3**, 762 (1961).
- ¹² B. Ames and H. Whitfield. *Cold Spring Harbour Symp. quant. Biol.* (1966).
- ¹³ S. Brenner, A. O. W. Stretton and S. Kaplan. *Nature* **206**, 994 (1965);
C. Yanofsky, E. C. Cox and V. Horn. *Proc. U.S. Natl. Acad. Sci.* **55**, 274 (1966).
- ¹⁴ F. H. C. Crick, Royal Society Croonian Lecture, 1966.
- ¹⁵ V. F. Chubukov and S. G. Tatarinova. *Zh. Mikrobiol. Epidemiol. Immunobiol.* **42**, 80 (1965).
- ¹⁶ E. Freese, E. Bautz and E. Bautz-Freese. *Proc. U.S. Natl. Acad. Sci.* **47**, 845 (1961);
E. Freese and H. B. Strack. *Proc. U.S. Natl. Acad. Sci.* **48**, 1796 (1962).
- ¹⁷ H. Schuster and W. Von Veilmutter. *J. Chim. Phys.* **53**, 1005 (1961);
H. Schuster and H.-G. Wittmann. *Virology* **19**, 421 (1961).
- ¹⁸ S. P. Champe and S. Benzer. *Proc. U.S. Natl. Acad. Sci.* **48**, 532 (1962).
- ¹⁹ I. Tessman, R. K. Poddar and S. Kumar. *J. Mol. Biol.* **9**, 352 (1964).
- ²⁰ H. Schuster. *J. Mol. Biol.* **3**, 447 (1961).
- ²¹ D. M. Brown and P. Schell. *J. Mol. Biol.* **3**, 709 (1961); *J. Chem. Soc.* 208 (1965).
- ²² D. W. Verwoerd, W. Zillig and H. Kohlhage. *Z. Physiol. Chem.* **332**, 184 (1963).
- ²³ P. D. Lawley. *J. Mol. Biol.* **24**, 75 (1967).
- ²⁴ D. M. Brown and M. J. E. Hewlins, *J. Chem. Soc., C.*, 1922 (1968); *Nature*, **221**, 656 (1969).
D. M. Brown, M. J. E. Hewlins, and P. Schell, *J. Chem. Soc., C.*, 1925 (1968).
- ²⁵ N. K. Kochetkov, E. I. Budowsky, E. D. Sverdlov, R. P. Shibaeva, V. N. Shibaev and G. S. Monastirskaya, *Tetrahedron Letters* 3253 (1967).
- ²⁶ D. M. Brown and J. H. Phillips. *J. Mol. Biol.* **11**, 663 (1965).
- ²⁷ C. Janin and D. Shugar. *Acta Biochim. Polon. (Engl. Transl.)*, **12**, 338 (1965).
- ²⁸ J. H. Phillips, D. M. Brown, R. Adman and L. Grossman. *J. Mol. Biol.* **12**, 816 (1965).
- ²⁹ C. F. Fox, W. S. Robinson, R. Haselkorn and S. B. Weiss. *J. Biol. Chem.* **239**, 186 (1964).
- ³⁰ G. R. Banks, D. M. Brown and L. Grossman, unpublished work.
- ³¹ J. H. Phillips and D. M. Brown in *Progress in Nucleic Acid Research* (Eds. J. N. Davidson and W. E. Cohn), Academic Press, New York, **7**, 349 (1967).
- ³² J. H. Phillips, D. M. Brown and L. Grossman. *J. Mol. Biol.* **21**, 405 (1966).
- ³³ P. D. Lawley and P. Brookes. *J. Mol. Biol.* **4**, 217 (1962).
- ³⁴ Y. Kyoguka, R. C. Lord and A. Rich. *J. Amer. Chem. Soc.* **89**, 496 (1967).
- ³⁵ J. Pitha, R. N. Jones and P. Pithova. *Can. J. Chem.* **44**, 1045 (1966).
- ³⁶ E. Küchler and J. Derkosch. *Z. Naturforsch.* **21b**, 209 (1966).
- ³⁷ C. C. Cheng and L. R. Lewis. *J. Heterocyclic Chem.* **1**, 260 (1964).
- ³⁸ A. E. Kretov and N. D. Borodavko. *Zh. Organ. Khim.* **2**, 367 (1966).
- ³⁹ D. M. Brown and M. J. E. Hewlins. *J. Chem. Soc., C.* 2050 (1968).
- ⁴⁰ G. W. Kenner, C. B. Reese and A. R. Todd. *J. Chem. Soc.* 855 (1955).
- ⁴¹ H. E. Johns, J. C. LeBlanc and K. B. Freeman. *J. Mol. Biol.* **13**, 849 (1965).
- ⁴² J. Ono, R. G. Wilson and L. Grossman. *J. Mol. Biol.* **11**, 600 (1965).
- ⁴³ L. F. Cavalieri and A. L. Stone. *J. Amer. Chem. Soc.* **77**, 6499 (1955).
- ⁴⁴ C. Zimmer and H. Venners. *Biopolymers* **4**, 1073 (1966).
- ⁴⁵ J. W. Drake. *J. Mol. Biol.* **6**, 268 (1963).
- ⁴⁶ K. B. Freeman, P. V. Hariharan and H. E. Johns. *J. Mol. Biol.* **13**, 833 (1965).
- ⁴⁷ D. M. Brown, A. D. McNaught and P. Schell. *Biochem. Biophys. Res. Commun.* **24**, 967 (1966).