

CHAPTER III

DNA

F. H. C. CRICK

*M.R.C. Unit, Cavendish Laboratory
University of Cambridge
Cambridge, England*

THE STRUCTURE OF DNA

THE STRUCTURE of DNA consists of two long polynucleotide chains, wound together around a common axis. A base attached to one chain is joined by hydrogen bonds to the base opposite it on the other chain, but only certain pairs of bases can be united in this way. They are:

Adenine with Thymine
or
Guanine with Cytosine.

Along any one chain any sequence of the bases is possible, but if the sequence along one chain is given then the sequence along the other is automatically determined, because of the pairing rule.

It is important to notice that the two chains run in opposite directions: if the sequence of atoms along one of them is considered to run up, then along the other it runs down.

This structure is now supported by a mass of physicochemical evidence, and especially by the careful and detailed work of Wilkins and his colleagues, using x-ray diffraction. It is also compatible with the analytical rule, first pointed out by Chargaff, that in any sample of DNA the molar amount of adenine equals that of thymine, and that of guanine equals that of cytosine (or analogue).

Both the x-ray and the chemical results have been shown to

apply to DNA from many different sources; the only certain exception is the DNA from the small virus ϕ X174 (or the related S13) which appears to be single-stranded, and which does not give 1:1 base ratios (Sinsheimer, 1959; Tessman, 1959).

THE PROPOSED REPLICATION SCHEME

It is now widely believed that DNA is the most important part of the genetic material and that the genetic information is contained in the precise sequence of the bases along the molecules, so that one would expect that the cell, in duplicating its DNA, would exactly reproduce the sequence of the bases. The fact that DNA normally consists of two complementary chains, which fit precisely together, immediately suggests a scheme for specific replication (Watson and Crick, 1953).

In outline one imagines that in some way the two chains of the double helix are separated, and that each chain then acts as a "template" to guide the formation of a new companion chain. If we call the two chains A and B, then chain A will guide the formation of a new chain B, and the old chain B will guide the formation of a new A, so that we shall end up with two AB's, where we only had one before.

This scheme is obviously very plausible. We must now ask how far it is supported by recent work.

EVIDENCE FOR THE REPLICATION SCHEME

Do the two chains come apart?

There are three types of experiments which suggest that the two chains of the duplex are separated in ordinary duplication.

A. The Meselson-Stahl Experiment

This uses the technique of density gradient centrifugation introduced by Meselson, Stahl and Vinograd (1957). If a solution of a heavy salt, such as caesium chloride, is spun in an ultracentrifuge for two days, a gradient of density is set up. If some DNA is included in the cell it will move until its (effective) density is the same as that of the surrounding salt solution. For DNA of high molecular weight, which diffuses slowly, the band will be fairly narrow. The position of the band in the cell will

depend on the density of the DNA. If DNA is obtained from bacteria grown on a source of N^{15} (instead of N^{14}) it will be correspondingly heavier and will band in a different place. Thus by this technique one can measure not merely the average density, but the *density spectrum* of a sample of DNA.

The actual experiment consisted in growing *E. Coli* for many generations in an N^{15} medium until all the DNA was heavy. Then, at time zero, the medium was changed to N^{14} , and N^{14} bases were also added, so that DNA synthesized before time zero was heavy and that made after time zero was light (Meselson and Stahl, 1958).

At time zero, naturally, one heavy peak was found in the ultracentrifuge. After one generation of growth there was no heavy peak left; there was only one peak of intermediate density. After two generations there were two equal peaks, one light and one intermediate.

These results are obviously exactly what one would expect from the proposed replication scheme, but they are not enough to prove it. They indicate clearly that each DNA molecule is a duplex; that the two equal parts separate, and that the new DNA consists of one old part and one new part. They do not show that these two parts are the two chains of the double helix.

If the DNA of intermediate density is heated at 100° for 15 minutes, its structure collapses, the molecular weight drops to half, and this material now gives *two* equal peaks in the gradient centrifuge, one light and one heavy. It is thus very likely that the two chains come apart during replication, but so far the evidence falls short of proof.

B. The Phage—Star Experiment

This experiment (see Levinthal and Thomas, 1957) was, in time, the first of those described in this section. Levinthal developed a special autoradiographic technique for studying the tracks which come from P^{32} atoms, using a special photographic emulsion. His results on the T-even Phages of *E. Coli* also suggest that a big piece of DNA consists of a duplex, the two halves of which separate on replication, but the interpretation depends

on the existence of the "big piece." As this is to some extent controversial the experiment will not be discussed further here.

C. Chromosome—Autoradiograph Experiments

These fall into two classes. The first, done by Plaut and Mazia (1956) used C^{14} . This does not permit fine resolution in the autoradiography. Their results appeared to show unequal labelling after only one division, under conditions where equal labelling might have been expected. However, this contradicts the rather more definitive experiment to be described next. This contradiction is not yet explained.

Dr. Herbert Taylor and his colleagues (Taylor, Woods and Hughes, 1957) have developed an autoradiograph technique using tritium labelling. This has the advantage that the range of the β particle in the emulsion being very short (about 1μ) high resolution autoradiography is possible. They worked with an organism—the broad bean—with large chromosomes, and were easily able to see whether an individual chromosome, or even a part of a chromosome, was labelled or not.

The cells chosen for examination were the rapidly growing cells of the root tip. The roots were immersed in tritium-labelled thymidine for 12 hours, and then placed in an unlabelled solution, containing colchicine, for various lengths of time. The colchicine, while allowing the replication of the chromosomes, prevents nuclear division, so that all the daughter chromosomes stay in one cell. It is immediately obvious how many replications a particular set of chromosomes has undergone, since one merely has to see how many chromosomes are present. Cells are frequently found in which the two daughter chromatids, though separate along most of their length are still linked at the centromere, and the labelling of two such daughter chromatids can be easily compared.

The thymidine can be shown to go only into DNA, which is synthesized during interphase. It is found that after *one* division the two daughter chromosomes are always equally labelled. After a second division (in general) one is labelled and the other unlabelled. These results again show the DNA to be a duplex,

though now at the chromosome level rather than at the molecular level.

Occasionally, in the second division, one chromatid is unlabelled for part of its length, and the rest labelled, while the other sister chromatid is labelled, except for the latter part. This suggests that during or after replication the two sister chromatids broke, interchanged, and rejoined.

The results and the interpretation of these experiments have been challenged by LaCour and Pelc (1958), but recent careful work by Woods and Schairer (1959) has made it likely that these criticisms are without foundation.

The earlier remarks about identifying the duplex as the two chains of the DNA apply even more strongly in this case. The results also suggest interesting models for the structure of chromosomes, but this is outside the present discussion (but see the paper by Freese, 1958).

THE ENZYMATIC SYNTHESIS OF DNA

The brilliant work of Dr. Arthur Kornberg and his colleagues (see, for example, Bessman *et al.*, 1958, and Lehman *et al.*, 1958) on the enzymatic synthesis of DNA supports to some extent the proposed replication scheme. The enzyme system was obtained from *E. Coli*, and purified several thousand fold. It needs as precursors the nucleoside triphosphates (e.g. the deoxy analogue of ATP). The corresponding diphosphates or the ribose triphosphates will not serve. To achieve net synthesis (for all cases except one) the enzyme requires *all four precursors* and a "primer" of DNA; it can synthesize up to 10 or 20 times the amount of DNA used as primer. Physico-chemical studies have shown that the newly-made DNA has a high molecular weight (comparable, though usually a little less, than that of the primer) and that it has the rigidity associated with the double-helix structure.

The base composition of the new DNA always obeys the pairing rule ($A = T$ and $G = C$) and the ratio of A/G is similar to that of the DNA used as a primer; it does not appear to depend upon the relative concentration of the four precursors.

If uridine triphosphate is added to the incubation mixture thymide triphosphate can be omitted, and the new DNA will

contain uracil rather than thymine. Uracil cannot be substituted in this way for the other three bases. By this technique it can be shown which base a given base analogue will replace. In all the cases where an unusual base is supplied as the nucleoside triphosphate one can predict by the base-pairing mechanism which base it will replace. Thus as far as it goes this evidence gives support for the proposed base-pairing.

Kornberg suggests that the reason that uracil is not found in DNA is that the enzyme which makes thymidine triphosphate cannot handle the uracil analogue, and so the proper precursor is not available to the cell.

So far all the results described are at least qualitatively compatible with the replication scheme; however it is found that if the enzyme is given only the adenine and thymine precursors it will, after a long lag, synthesize a DNA-like polymer containing only A and T. The mechanism by which this is produced is not known.

If this A-T polymer is now extracted and used as a primer in the usual system, containing all four precursors, synthesis starts immediately without any lag and the new material contains A and T. This looks like very strong support for the replication mechanism, but it remains to be shown by direct experiment that in this case, as well as the others, the two chains are coming apart, and that material is not merely being added to the ends of chains.

The DNA from the virus ϕ X174 is believed to be single-stranded (Sinsheimer 1959). This DNA can also act as a primer in the Kornberg system, (Kornberg, personal communication). The relative efficiency of different sorts of DNA acting as primer is a complicated subject which is still under active study by Dr. Kornberg.

GENETIC FINE STRUCTURE

In recent years it has been possible with a number of organisms to carry out genetic mapping within the genetic functional unit. (See the brief review by Pontecorvo, 1958.) The most detailed and successful study has been that of Benzer, using *Coli* Bacteriophage T4 (Benzer, 1957).

We can rearrange Benzer's results in the r_{II} locus briefly as follows:

(1) There are many distinct sites within the r_{II} locus, probably several hundred.

(2) Some mutants appear to map as points: these almost always have a finite rate of back mutation. Others map as extended regions; these never back mutate.

(3) All the sites can be mapped in a linear order. The mapping was done partly in the usual way, and partly by an ingenious method of "deletion mapping."

(4) The r_{II} locus has two regions ("cistrons"), A and B. Mutants defective only in region A can complement (in mixed infection) mutants defective only in region B. There is a point on the map such that the A region is entirely to one side of it and the B region entirely to the other.

(5) The smallest distance so far found between two point mutations is very small but it is still some ten times greater than the genetic resolving power of the system; if the genetic material of the phage is assumed to be DNA then a rough calculation suggests that this smallest observed distance is only one or a few base pairs in extent.

It is obvious that these results fit very well into the picture of a "gene" (meaning here the functional unit) as part of a molecule of DNA.

MUTAGENESIS

This subject is in a state of rapid development. It suffices to say here that at first sight the results do not fit in a simple way into our model. Benzer discovered that mutations at certain points, which he called "hot-spots," occurred much more frequently than the average. It was also discovered by Benzer and Freese (1958) that the mutagen bromouracil, which can increase the overall mutation rate by one hundred fold, produces new hot-spots. (See also Brennet et al. 1958.)

It is tentatively assumed that the hot-spots occur by a special mechanism, but its character is quite unknown.

CONCLUSION

Although there are certain difficulties, our picture that DNA is the essential part of the chromosome and that DNA is replicated by a process using specific base-pairing is standing up very well. It is also encouraging that both in the intact cell, as in Benzer's work, or in cell-free systems, such as Kornberg's, we have techniques which are still capable of much further exploitation.

REFERENCES

- Benzer, S. (1957) in *The Chemical Basis of Heredity* ed. McElroy and Glass, Johns Hopkins Press, Baltimore.
- Benzer, S. and Freese, E. (1958) *Proc. Nat. Acad. Sci.*, **44**, 112.
- Bessman, M. J., Lehman, I. R., Adler, J., Zimmerman, S. B., Simms, E. S. and Kornberg, A. (1958) *Proc. Nat. Acad. Sci.*, **44**, 633.
- Bessman, M. J., Lehman, I. R., Simms, E. S. and Kornberg, A. (1958) *J. Biol. Chem.*, **233**, 171.
- Brenner, S., Benzer, S. and Barnett, L. (1958) *Nature* **182**, 983.
- Freese, E. (1958) *Cold Spring Harbor Symp.*, **23**, 13.
- Freese, E. (1959) *Proc. Nat. Acad. Sci.*, **45**, 622.
- La Cour, L. F. and Pelc, S. R. (1958) *Nature* **182**, 303.
- Lehman, I. R., Zimmerman, S. D., Adler, J., Bessman, M. J., Simms, E. S. and Kornberg, A. (1958) *Proc. Nat. Acad. Sci.*, **44**, 1191.
- Levinthal, C. and Thomas, C. A. (1957) in *The Chemical Basis of Heredity* ed. McElroy and Glass, Johns Hopkins Press, Baltimore.
- Meselson, M., Stahl, F. W. and Vinograd, J. (1957) *Proc. Nat. Acad. Sci.*, **43**, 581.
- Meselson, M. and Stahl, F. W. (1958) *Proc. Nat. Acad. Sci.*, **44**, 671.
- Plaut, W. and Mazia, D. (1956) *J. Biophys. Biochem. Cyt.* **2**, 573.
- Sinsheimer, R. L. (1959) *J. Mol. Biol.*, **1**, 37 and 43.
- Taylor, J. H., Woods, P. S. and Hughes, W. L. (1957) *Proc. Nat. Acad. Sci.*, **43**, 122.
- Tessman, I. (1959) *Virology* **7**, 263.
- Watson, J. D. and Crick, F. H. C. (1953) *Nature* **171**, 964.
- Woods, P. S. and Schairer, M. V. (1959) *Nature* **183**, 303.

Discussion

DR. BARTON CHILDS: I wonder if Dr. Crick would comment on crossing over.

DR. CRICK: Yes. I could make one or two comments along these lines.

Attempts have been made to make a uniform mechanism of genetic recombination. As far as I know these have all been unsuccessful.

There are two basic mechanisms of recombination of which we can conceive, although there may be others. One is that of breakage and rejoin, which is the one customarily assumed. The other one is known as copy choice. You start copying. When you have made some new material copying along one molecule, you skip across and copy along the other.

There are reasons for believing, although they are complicated and not very convincing, that you do have copy choice in phage genetics. Let me say that the rate of recombination in phage is very, very much higher per length of DNA than it is in higher organisms. That is one of the reasons you can do the genetics so easily.

It is now thought that you cannot make a unitary theory but you have to invoke both types of mechanism. The reason is that in higher organisms you do not always recover the parental types; if you had only a copy-choice mechanism which never broke the chromosome, you should recover them.

So the picture we have is that at the molecular level it is, possibly, a copy-choice mechanism, but in higher organisms there is also breakage and rejoin. This would be certainly an attractive hypothesis if Taylor's result that you get breakage and rejoin in mitosis stands up. He has to show this is not merely due to the effects of the radiation and possibly he has done this. If that is so, it means that we have evidence that chromosomes can break and rejoin, although, of course, we would like to have the evidence for meiosis.

The other question is: Where do they break and rejoin?

I would say here that the evidence isn't enough to produce more than very tentative hypotheses and possibly the one due to Freese (Freese, E., *Cold Spring Harbor Symp.* 23:13, 1958) is as good as any that you have DNA molecules joined together by protein molecules, which are attached in a rather special way.

The chromosome would then consist of DNA molecules of about 10 million molecular weight, each of which we would believe (from transforming factor evidence) to carry several genes, and these DNA molecules would be joined together by special proteins.

Breakage and rejoin could possibly happen *between* DNA molecules rather than in the middle of them. This is the tentative picture we have.

Now, you may ask, what exactly happens in the phage system, where you are clearly getting very fine mapping which certainly isn't due to breaks between whole DNA molecules?

We used to think that when the DNA molecule replicated, it did so by making two new chains—the two new chains you needed—at the same time. The idea was you did not take the two chains apart first. You unwound them during the synthesis of the two new ones. We believe now that this is implausible because, if you remember, the actual chemical sequence of these two chains runs in opposite directions and this mechanism would imply that one chain is growing in one direction chemically while the other chain is growing in the opposite direction chemically.

We suspect now that it is more reasonable to have a method of growth in which you only make one chain at a time. If you assume this, you can explain both the little virus I mentioned, which has single-chain DNA, and also genetic recombination. The reason is that if you are going to have genetic recombination between two DNA molecules at this very fine level, it must be done by the molecules coming together in some way. We do not think you can do that by having two and two chains coming together to make a four-chain situation. We would much prefer to use the usual base-pairing of two single chains.

If you had the old system of replication, where you never have single chains, it is very difficult to do that. On the other hand, when you make just one chain at a time, it is not too difficult to devise theories of various sorts in which the single-chain regions can be used for recombination.

DR. SEYMOUR GELFANT (Syracuse University): How do you reconcile the diagram you put on the board showing blocks of protein appearing alternately along the chain of DNA, with electromicrographs of chromosomes (H. Ris in *The Chemical Basis of Heredity*, McElroy and Glass, The Johns Hopkins Press, Baltimore, 1957) which indicate that the chromosomal microfibril has an inner core of DNA surrounded by a continuous shell of protein. Moreover, physiological changes related to chromosome duplication can be correlated with an increase in diameter of the microfibrils.

DR. CRICK: I would say most of those EM pictures are very difficult to interpret. When you look at the lamp brush chromosome, for example, you see loops and other complicated things.

I do agree with you that this model is incomplete and does not explain the apparent diameter of the "DNA." That is why I said that

I only want you to regard this in the most tentative way and why I did not put it into the lecture. I think actually you can make a lot of models of a chromosome and they are all very speculative. The only reason I described one was to get away from the model that had just one enormous chain of DNA, which I think is less likely. I think the loops they see could be the DNA molecules. Don't you think so?

DR. GELFANT: The loops could still contain protein along the entire length, though.

DR. CRICK: Yes, that is certainly true. But it would not affect the argument, which was whether you got breakage and rejoin *within* the DNA molecules or at the point where the molecules joined. I am trying to avoid discussing chromosome structure because I think it is too difficult a subject at the moment.

DR. WALTER VINCENT (Syracuse): I would like to ask why you tend to ignore the very delightful hot spot on the model that you have shown here in your attempts to analyze what the mechanism of mutation is.

DR. CRICK: This is a matter of your point of view. If you are interested in what mutations occur, the hot spot is the important one because it occurs more often than the others. One is interested in it, but one does not see what one can do about it. It is a question of what technique you should use to study it.

On the other hand, if you are interested in the mechanism—the basic mechanism—of replication, then any type of mutation may give you information. The thought is that some are more easily studied than others—not that they are necessarily more important.

Let me put it another way. A mutation is a way of making a mistake. Obviously you can make mistakes in a number of different ways. Some will be more interesting than others, as we all know! It depends on your point of view.

DR. VINCENT: You made one suggestion which seemed to me to make them likely for study. This is that perhaps at this area one has an unusual set of base sequences which one might get at. Now, it strikes me that this might be the one approach that you could use to study this particular area:

One might expect—say on just a simple type of hypothesis—that this “hot spot” is an unusual sequence of bases in terms of more than usual occurrence of a particular type of base. One might expect then that you could either increase or decrease the number of mutations occurring at that point by say making certain analogues that you

throw into this area. This would give you some idea, I think, regarding the mechanism of mutation.

DR. CRICK: Analogues would be looked upon as specific mutagens in such cases. This approach has not given any clear results so far. If you study a spontaneous hot spot and then see what happens with bromuracil, though the *average* mutation rate is increased a hundred times by bromuracil this mutagen does not appear to increase the mutation rate at this particular hot spot. But I do not know what you do with this fact.

The way we think we *might* get at it is to find the protein that corresponds to a particular gene and find the hot spots in that; we could then find the point in the amino acid sequence of the protein—which ideally is accessible to us by experiment—which corresponds to the hot spot, and we might, if we knew how the coding went, deduce the base sequence from the amino acid sequence. But this is obviously some time ahead. That is why I say you cannot do it too easily.

I am actually concealing the fact—because it is unpublished—that Freese has got some very interesting results on which you can make a tentative theory for mutagenesis. (Now published; Freese, E., *Proc. Nat. Acad. Sci.* 45:622, 1959.)

DR. GELFANT: I was interested in your assumption concerning the specificity of tritium labelled thymidine (for DNA) and the observation recently made by Brachet (J. Brachet, *Exptl. Cell Research* 14: 650, 1958) showing that tritium labelled thymidine can also be found in RNA and the cytoplasm. Would you care to comment on this?

DR. CRICK: Isn't it true that the majority of it goes into DNA? You may get a small amount of it into something else, but I thought most people had done controls by chemical methods in their material. I did not know that Brachet had shown that any went into anything but DNA.

DR. GELFANT: That is right. It is still under investigation, but I was wondering how seriously this might affect your conclusions concerning the replication of DNA.

DR. CRICK: I did not know the matter was in question, I confess. I thought the result was all right.

DR. INGRAM: Could you make one more comment on the DNA of Sinsheimer's small phage, ϕ X 174, the single stranded DNA, and tell us a little about it?

DR. CRICK: As you probably know—those of you who read *The New York Times*—this DNA is from a little virus known as ϕ X 174, or

S 13, which is a small virus, rather like a small RNA virus. The DNA has a molecular weight of about 2 million and appears from physical-chemical studies to be single stranded. (Sinsheimer, R. L., *J. Mol. Biol.* 1:37, 43, 1959)

The base ratios are not 1:1. Of course, you could have single stranded DNA and have base ratios of 1:1 if you had equal mixtures of two chains which at one time were paired. So this shows if there are two sort of chains, at least they are not present in equal amounts.

What Dr. Kornberg has done and has allowed me to mention is to show that this DNA will act as a primer in his system—quite a good primer. You put in single chain DNA as a primer and you end up with two-chain DNA. In other words, the enzyme system makes two-chain DNA using single chain as a primer.

How could you explain using these ideas how you could get single chain DNA?

As I have said earlier, it is difficult on the old mechanism. It is not too bad when you think of the new type of mechanism of one chain being made at a time. What would happen is that the single chain would go into the cell and you could get a new one made alongside it. Then you might duplicate that, and get two-chain DNA for some time—until such a stage as you got the protein of the virus being made by some independent mechanism.

You now have to make two postulates. One is that you tend to start the replication at one end of the DNA rather than the other. To achieve this, one might assume that you have a lot of adenine-thymine pairs at that end of the DNA and a lot of guanine-cytosine ones at the other. We know from the work of Doty and his colleagues that the guanine-cytosine pair, with three hydrogen bonds, is more stable than the adenine-thymine pair with two. The other assumption is that in the later stages of infection the protein or the virus comes and sits on the single chain and captures it, before it can act as a primer to make a double chain.

Thus by having a process which throws off single chains, and by having a capturing process, due to the protein, and by starting replication preferentially at one end, you can get a virus which has a single chain and which does not have 1:1 base ratios. The same type of mechanism could also be postulated for an RNA virus.

There is, I think, a general point to be made here; that it is perhaps not reasonable to think of the double structure as essential for DNA replication. It may happen to be there for other reasons. All you may need for most of these replication processes is a single chain structure

and an enzyme. It may not matter what the configuration of the rest of the chain is except at the point where the enzyme is acting.

This would lead us to a rather different concept. Instead of the idea of a template, you would have the idea of a tape (a word first used by Commoner); that is to say, a lot of elements joined in a fixed, lineal order which can take up any configuration as long as it has a fixed configuration as it is fed into the machine.

From this point of view, therefore, the DNA double helix would not be the active form. It would be one tape sitting on another tape in order to keep each other quiet.

I do not know whether this idea is valid. It would imply that the gene cannot act while it is in the double helical form; to act it would have to be activated, possibly from one end, possibly by a special protein, as suggested to me by Dr. Francois Jacob. You would not need a special mechanism for turning genes *off*; but you would have to turn them on.

These are all rather speculative ideas but they do perhaps give us a slightly varied concept: the idea of the tape instead of the template. This may be true in protein synthesis as well, as suggested by Koshland.

DR. VINCENT: Dr. Crick, could we generalize that a little further?

I am intrigued by your concept here of the attachment of protein onto DNA, which essentially inhibits the formation of a double chain. Could this be generalized to the concept of RNA in a cell? Of course, here we apparently have a single chain. Possibly this is the reason RNA is never found in a double chain—because it is always hooked up with a protein so rapidly that it cannot get into the double chain configuration.

DR. CRICK: That, basically, is the idea—that all the operations that take place using nucleic acid may be of this sort. For protein synthesis it would be like having a little machine into which you feed a tape of RNA and from which a new chain of protein comes out. You know these copying machines that are used with high speed computers? You feed in one tape and it copies it and you make another. That is exactly the type of situation.

There are three contexts: The making of DNA on DNA; the making of RNA on DNA; and the making of protein on RNA—and, whether it applies in all those contexts, I do not know. The one I would particularly like to know about—and the one I am most uncertain about—is the making of RNA on the DNA template.

DR. LEONELL STRONG, Roswell Park: I would like to go back to

the idea of the hot spot, (excessively high mutability at certain loci), comparing it with the mechanism of mutability with these very low background mutation rates.

It is a fact that in genetic material these are not the only biological states in mutability; they are merely the two extreme classes. You also have a genetic state in which apparently some loci do not mutate at all, at least there is no evidence that they do since there are obviously more loci as indicated by ultra-violet studies than are represented by the known mutations. So that you are not having a different mechanism in the hot spot and the low grade mutability rate, but you may have a gradient factor which must be taken into consideration.

Would you like to discuss this point?

DR. CRICK: I am not sure that I fully understand you. You certainly get in this system of Benzer's all sorts of mutation rates including some that are so unstable that you cannot study them. You get mutation rates of—say—50% on the one hand, right through to ones which mutate very slowly—say 10^{-10} .

When you have no mutation rate at all, however, one is inclined possibly to believe that that is because you have a dud form of the gene; in other words, that you have a deletion. I think it would be difficult to get high stability in any other way. The best way to be stable, if you are a gene, is to be dead.

DR. INGRAM: I can see reasons for 5 bromuracil being a mutagen. But, why is a proflavine a mutagen?

DR. CRICK: We do not know, but a possible explanation would be that the proflavine combines with the copying site in such a way that it holds the two chains an unnatural distance apart so that instead of putting in a purine-pyrimidine pair, you would put in a purine-purine pair. This is the type of idea that is going around. But the answer is—nobody knows.

It might be amusing if we ask ourselves how much tape there is: how much information you could carry in one of your haploid sets of chromosomes. The figures are all straight-forward, but they are sometimes a little surprising.

It turns out that you have about 3×10^9 base pairs in your haploid set of DNA, though we don't know of course that it is all genetically significant. There are some amusing calculations which can be done along these lines.

First of all, if you put all the DNA molecules end-to-end, what would be their total length? Remember that the nucleus is just a few μ across. The answer is that you have about a hundred centimeters

of DNA in each of your haploid sets. This means that if you took all the DNA molecules in *your body* and strung them out end-to-end, they would certainly reach to the sun.

This may be a surprise to you—that you can take a little bit of yourself and spin it out so fine that it can get to the sun. You might think the proper thing to say is, "That's because the molecule is so thin." That is the wrong point of view. It is because you are so fat. In other words, you are all very complicated organisms.

You can look at it another way (one due to Kornberg); if you take a haploid set of DNA from every member of the human race—and now I am not trying to spread it out but trying to make it as small as possible—what volume would it occupy? The answer is that it would go into a pin head. On this picture the whole genetic information of the human race can be put in a pin's head.

There is, I think, a more meaningful and useful calculation: suppose you took the haploid amount of DNA from one person and used it to write a few books. You might take three base pairs for a letter of the English alphabet. How many different books could you write using this amount of information? The answer is you can write roughly 500 different large textbooks. So, even though it is a very small space, if the information is carried at the molecular level (as we have been arguing from Benzer's work), you can carry an enormous amount of it.

Certainly you would think that was enough. But I must remind you we have no method of estimating how much information (how many proteins, let us say) we need to make a higher organism. We can make a rough estimate for *E. Coli* but how much is needed to make your hand, for example, we just do not know at all. But it does look as if there is plenty available.