On the Genetic Recombination of Bacteriophage ØX174 DNA Molecules

Thesis by

Robert Michael Benbow

In Partial Fulfillment of the Requirements
for the Degree of
Doctor of Philosophy

California Institute of Technology
Pasadena, California

1972
(Submitted May 19, 1972)

ACKNOWLEDGMENTS

I especially wish to thank Robert L. Sinsheimer for his guidance and expert criticism.

I also wish to thank the following colleagues:

Professors Ed Lewis, Jerome Vinograd, Norman Davidson, Max Delbruck, and Sterling Emerson provided fascinating discussions and valuable advice.

James Bonner provided his assistance and much-needed support on several crucial occasions.

Gloria Davis and Alma Jean Shafer worked with me during numerous experiments.

Tony Zuccarelli endured and tolerated my frenetic suggestions for several years.

The ØX174 group endured me.

I wish to thank the Caltech Biology Division for providing generous financial assistance.

Finally, I wish to express my deepest gratitude to Paul Howard Flanders for making this opportunity possible.

ABSTRACT

Genetic recombination between two ØX174 parental replicative form DNA molecules primarily occurs by the following sequence of events: (i) a single strand scission in one RF DNA molecule; (ii) displacement synthesis; (iii) unimolecular branch migration; (iv) attack; (v) formation of a hydrogen bonded joint molecule; (vi) bimolecular branch migration and further DNA (repair?) synthesis; (vii and viii) covalent bond formation and single strand scission (order not certain); (ix) restoration of recombinant DNA molecules to the parental RF configuration. The net result is an asymmetric non-reciprocal recombination event yielding one parent and one recombinant. The average net DNA synthesis is less than 600 nucleotides; breakage and reunion occurs; over 50% of all genetic exchanges involve regions less than 400 nucleotides in length. Single strand insertion is (probably) the most common recombination event. Recombinant formation is complete at or shortly after the initiation of progeny RF DNA synthesis; the completed recombinant DNA molecule and surviving parental molecule then resume normal ØX174 DNA replication.

A (minor) secondary recombination mechanism exists in which recombinant formation occurs between ØX174 progeny RF DNA molecules.

PREFACE

In the same year that I came to Caltech to work in Dr. Sinsheimer's laboratory, William Hayes wrote: "Recombination is the central phenomenon of genetics and the principal tool of genetic analysis. It is therefore paradoxical that its nature remains one of the few dark corners of cellular behavior which the recent developments in molecular biology have, so far, done little to illuminate" (Hayes, 1968).

In the following thesis I shall present a model for the genetic recombination of bacteriophage ØX174 DNA molecules based upon research that I have carried out during the four years since Dr. Hayes wrote his book. The pathway for the formation of recombinants is basically complete except for the precise action of the rec A enzyme. The data upon which this model is built are described in some detail in the text. These data are taken from a series of publications and manuscripts on the nature of ØX174 genetic recombination written by myself and colleagues in Dr. Sinsheimer's laboratory. These papers are:

(1) Benbow, R.M., Hutchison, C.A., Fabricant, J.D., and Sinsheimer, R.L., "Genetic Map of Bacteriophage ØX174" J. Virology 7, 549-558, 1971.

(2) Benbow, R.M., Mayol, R.F., Picchi, J.C., and

Sinsheimer, R.L. "Direction of Translation and Size of Bacteriophage ØX174 Cistrons" J. Virology, in the press, 1972.

(3) Benbow, R.M., Eisenberg, M.G., and Sinsheimer, R.L., "Multiple Length DNA Molecules of Bacteriophage ØX174" Nature, in the press, 1972.

(4) Benbow, R.M., Eisenberg, M.G., and Sinsheimer, R.L., "The Structure of Multiple Length DNA Molecules of Bacteriophage ØX174", to be submitted, J. Mol. Biol., 1973.

(5) Zuccarelli, A.J., Benbow, R.M., and Sinsheimer, R.L., "Deletion Mutants of Bacteriophage ØX174" Proc. Nat. Acad. Sci., in the press, 1972.

(6) Benbow, R.M., Zuccarelli, A.J., Davis, G., and Sinsheimer, R.L. "Genetic Recombination of Bacteriophage ØX174", to be submitted, J. Mol. Biol., 1973.

(7) Benbow, R.M. and Sinsheimer, R.L. "A Role for Single Strand Breaks in ØX174 Genetic Recombination", to be submitted, J. Mol. Biol., 1973.

(8) Benbow, R.M., Schafer, A.J. and Sinsheimer, R.L. "Parental DNA Exchange During Genetic Recombination Between ØX174 DNA Molecules", to be submitted, J. Mol. Biol., 1973.

(9) Benbow, R.M., and Sinsheimer, R.L. "High Negative Interference in Bacteriophage ØX174", to be submitted, Genetics, 1973.

(10) Benbow, R.M., Zuccarelli, A.J., and Sinsheimer, R.L., "Genetic Recombination Between ØX174 DNA Molecules", to be submitted, Proc. Nat. Acad. Sci., 1973.

TABLE OF CONTENTS

INTRODUCTION

Genetic recombination is defined simply as the question how "genetic markers on homologous (DNA) molecules come to be located on the same molecule" (Bodmer and Darlington, 1971). More specifically, genetic recombination is "an interaction between homologous DNA molecules to yield DNA molecules having a nucleotide sequence partly that of one parent molecule and partly that of the other" (Meselson, 1964). In this thesis, the genetic recombination of bacteriophage ØX174 DNA molecules is studied as a model system for examining the general process of genetic recombination.

To establish the pathway or pathways by which ØX174 recombinants are formed, I have analyzed the ØX174 recombination process in terms of each of the major existing models of genetic recombination. By correlating the observed properties of, and isolated intermediates in, ØX174 genetic recombination with the specific predictions of each model, I have characterized the primary pathway for the formation of ØX174 genetic recombinants.

(i) Models of Genetic recombination

A priori two general mechanisms exist for the genetic recombination of DNA molecules: copy choice and breakage and reunion (Hayes, 1968). The critical distinction between these two theories is that copy choice "yields recombinant

chromosomes which are synthesized from new material and inherit only genetic information from the parental structures" (Hayes, 1968). In contrast, breakage and reunion results in recombinant chromosomes which contain physical material from both parental chromosomes. Most current molecular models of genetic recombination are based on breakage and reunion mechanisms (Whitehouse, 1969).

(ii) The Stadler-Towe Model

Nevertheless, copy choice is used as the basis for a recent model to account for the uniformly non-reciprocal recombination of allelic cysteine mutants in Neurospora crassa (Stadler and Towe, 1963). Regions of switching (Freese, 1957), fixed pairing regions (Stahl, 1961; Murray, 1961), and fixed recombination regions called polarons (Lissouba, Mousseau, Rizet, and Rossignol, 1962) are incorporated into this model.

A switching region is the point where a new DNA strand is replicated from the DNA template of the other of the two parental chromosomes; presumably this also corresponds to a region of close chromosomal pairing (Pritchard, 1960).

A fixed pairing region implies that chromosomal "effective" pairing (Pritchard, 1955; 1960) occurs in relation to a fixed chromosomal discontinuity. This concept circumvents numerous objections to copy choice models raised by gene conversion and by recombination of two or

more markers in the same gene. (As initially proposed, however, fixed pairing regions are independent of whether the model is based on copy choice or breakage and reunion (Stahl, 1961)).

The polaron, originally defined for specific use with a copy choice model of genetic recombination, is a fixed recombination region in which only a single switch can occur, resulting in non-reciprocal segregation ratios from that point to the end of the polaron. A corollary of the polaron hypothesis of possible relevance to ØX174 genetic recombination is that only non-reciprocal recombination events can occur between two markers in the same polaron.

The Stadler-Towe or modified polaron model for genetic recombination incorporates two modifications into Lissouba's initial polaron model. The major features of the modified polaron model as applied to ØX174 recombination are shown in figure 1. Copy choice is used to generate one non-reciprocal recombinant within the polaron; one parental genome survives intact.

However, unlike the initial polaron model, double switches (switchbacks) frequently occur a short distance beyond the initial switch point. This presumably predicts the high negative interference observed by Benbow, Hutchison, Fabricant, and Sinsheimer (1971) during ØX174 genetic recombination. In addition, "crossover option" points occur at

Figure 1. The Stadler-Towe Modified Polaron Model for Genetic Recombination

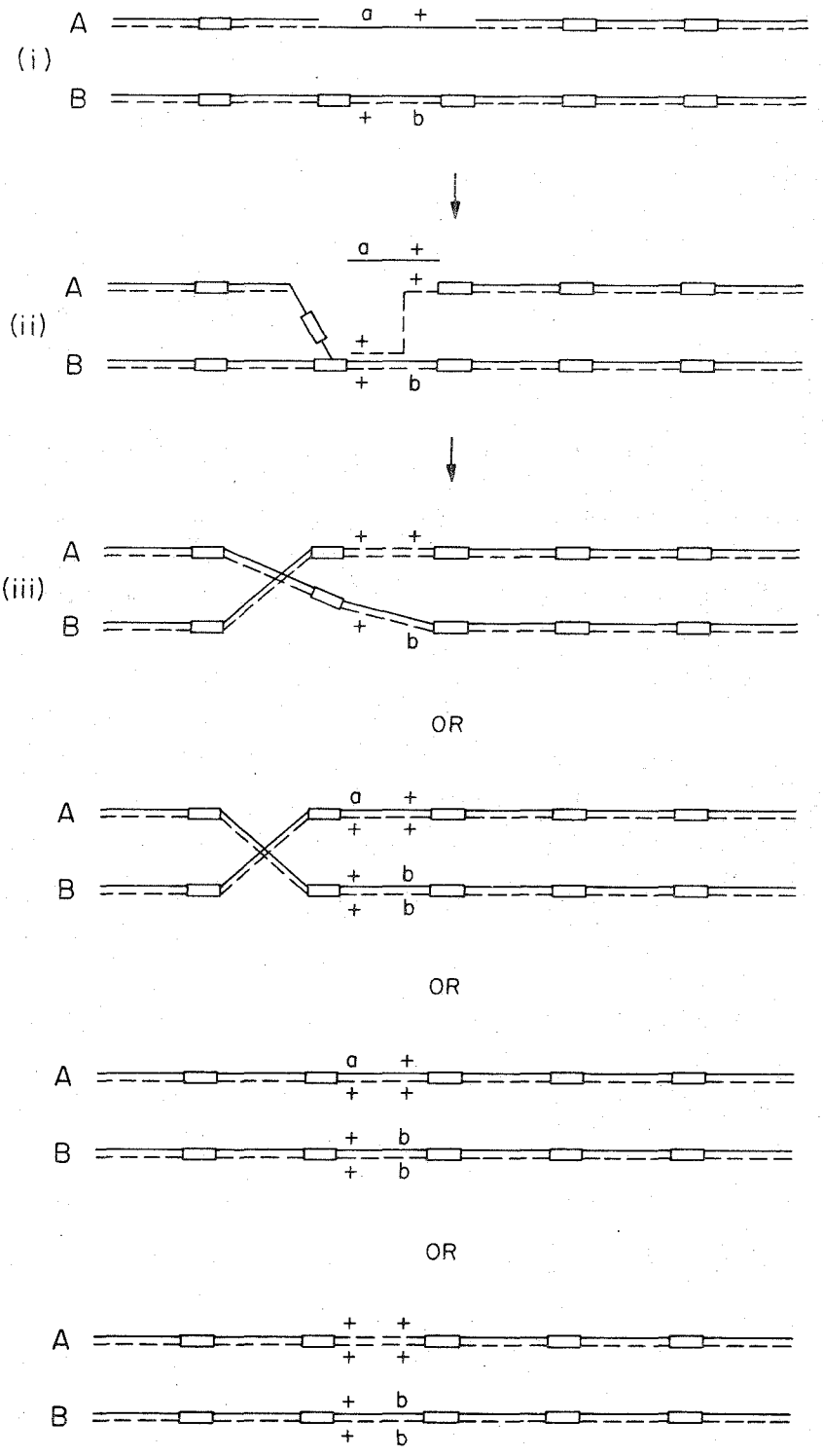
Two ØX174 replicative form (RF) DNA molecules are shown with genetic markers a and b (step i). The solid line represents the parental (plus) strand of each RF molecule, and the dotted line represents the newly synthesized complementary (minus) strand. The single strand region in RF molecule A represents a polaron which has not yet replicated---an essential feature of this model since polarons are the only regions which undergo non-reciprocal recombination. It is not known how many polarons are present in each RF molecule. The open rectangles represent "crossover option" points at which reciprocal recombination occurs; these crossover events are "operationally indistinguishable from breakage and reunion." The number of "crossover option" points per ØX174 RF DNA molecule is also unknown.

The unreplicated polaron (step i) is partially copied from its normal template until a switch point is reached (step ii). Then, the other RF molecule is used as a template until a "crossover option" point is reached (step iii).

Four possible wt recombinant progeny are predicted by this model. If the uncopied region of the polaron is excised, then one parent and one recombinant are

generated (step iii---choice 1 and 4). If the newly synthesized DNA region reanneals with the unreplicated region of the polaron, one heteroduplex DNA molecule and one parental molecule are generated (step iii---choice 2 and 3). Since reciprocal recombination events are rare during ØX174 genetic recombination (Benbow, Zuccarelli, Davis, and Sinsheimer, 1973), step iii---choices 1 and 2 are unlikely.

THE STADLER-TOWE MODEL



the end of polarons which allow reciprocal recombination to occur at fixed breakage points. In principle, therefore, the Stadler-Towe model uses copy choice only to explain non-reciprocal recombination events and resorts to breakage and reunion as a mechanism for reciprocal crossover events. In addition, it is necessary to assume (ad hoc) that non-reciprocal recombinants are generated only in polarons which initially fail to replicate. Note that one parent and one recombinant are recovered from each non-reciprocal recombination event according to the predictions of this model.

(iii) The Boon-Zinder Model

A breakage and copy mechanism was proposed by Boon and Zinder (1969) to account for their observation that a single recombination event between two bacteriophage f1 DNA molecules results in the recovery of one parent and one recombinant. Three stages in the recombination process were proposed: initiation; replication; and return. The Boon-Zinder model as applied to ØX174 recombinant formation is drawn in figure 2.

Initiation requires a single strand break in one RF molecule and a double strand break in roughly the same region of the other RF molecule. "Although the breaks could be independent of each other, it seems possible that some of the breaks could be promoted and specified by the

Figure 2. The Boon-Zinder Model for Genetic Recombination

The Boon-Zinder breakage and copy model for generating one bacteriophage f1 parent and one recombinant is redrawn for bacteriophage ØX174. Complete circular molecules are not shown to minimize confusion. Steps i, ii, and iii represent the stage of initiation.

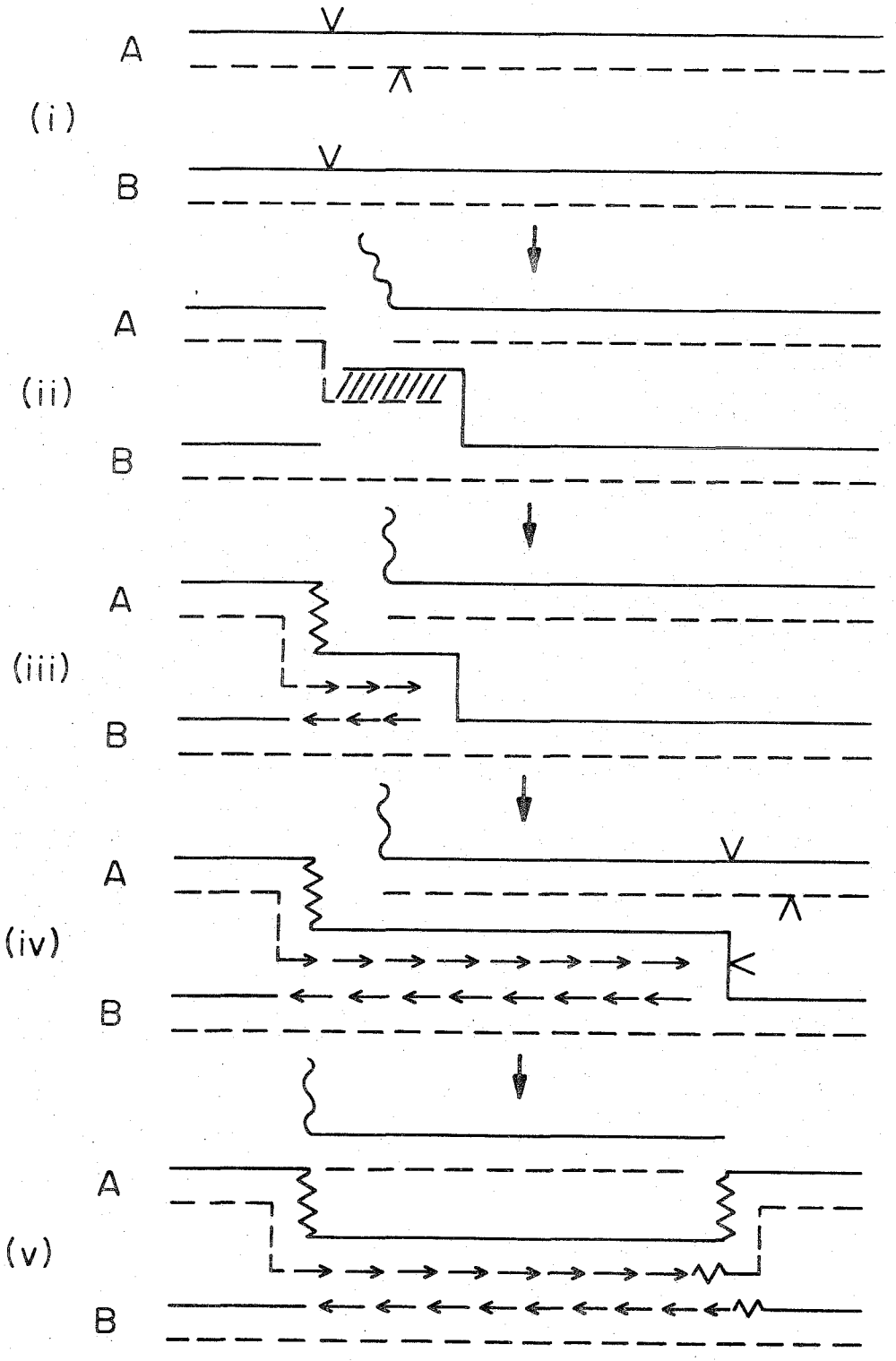
In step i a total of three single strand breaks are made, perhaps in a concerted mechanism. The viral (plus) strand from one molecule uncoils and anneals with the complementary (minus) strand from the other RF molecule to form a hydrogen bonded joint molecule (step ii). DNA repair synthesis links the annealed strand with the appropriate broken strand on the other RF molecule. DNA replication begins, presumably using the other strand as a primer.

Replication (→→→), the second stage of the Boon-Zinder model, proceeds via a replicating fork until a return event occurs (step iv).

Return, the final stage of their model, involves three single strand breaks in the same three strands broken initially (for a correct return), and must occur before the replication fork travels around the circular genome to the initial single strand breaks. DNA repair synthesis links the appropriate strands and a fragment is released (to be degraded?).

The products of the recombination event are a recombinant RF (which was replicated into the DNA molecule containing two single strand breaks) and one parental RF (which was maintained by DNA repair in the other molecule which contained one single strand break).

BOON-ZINDER MODEL



other breaks and by the pairing." The Boon-Zinder model makes a specific prediction: the RF DNA molecule with the double strand break is the extended recombinant molecule whereas the repaired molecule with the single strand break remains parental.

Replication via a Cairn's-type replicating fork (Cairns, 1963) extends one RF molecule by copying a single strand of the other RF molecule. The intact circular strand of the non-extended RF molecule is copied (repair?) to regenerate a parental genotype. Boon and Zinder point out that restriction of DNA replication to only one direction introduces the polaron concept (Lissouba, Mousseau, Rizet, and Rossignol, 1962) into their model.

Return requires single strand breaks in the same three strands in order to generate a unit length DNA molecule. If the other strand of the replication fork is broken (an incorrect return), this model predicts the formation of double length recombinant DNA molecules with an accompanying reciprocal recombination event.

A final assumption of the Boon-Zinder model is that the probability of a return event depends only on the type of DNA molecule (ØX174 in this case) and the distance (l) from the initiation point (x). They define R_1 as the fraction of chromosomes which have not yet returned at distance l. Thus:

$$dR_1/dl = -kR_1$$

where k is a specific constant for ØX174. Solving, they obtain:

$$R_1 = e^{-kl} .$$

If k is high, returns close to the initiation point are highly probable, and high negative interference is predicted.

(iv) The Holliday Model

Models of genetic recombination by breakage and reunion were independently formulated by Holliday (Holliday, 1962; 1964) and by Whitehouse (Whitehouse, 1963; 1964; Hastings and Whitehouse, 1964; Whitehouse and Hastings, 1968). Their models introduce the concept of a hybrid DNA region and suggest the possibility of base pair mismatch correction either by DNA repair enzymes or by copying in the hybrid region. Since both models have greatly influenced subsequent studies of genetic recombination, the application of each model to ØX174 genetic recombination is outlined below.

Many of the basic assumptions used in the Holliday and Whitehouse models were found to be true for the formation of ØX174 recombinants. Thus, the genetic material of the ØX174 chromosome (genome) does consist of s single DNA double helix (Sinsheimer, Starman, Nagler, and Guthrie, 1962; Benbow, Zuccarelli, and Sinsheimer, 1973). ØX174

nonsense mutations (and presumably temperature sensitive mutations as well) involve single base pair substitutions (Benbow, Hutchison, Fabricant, and Sinsheimer, 1971).

In addition, following Holliday (Holliday, 1964) I shall assume that general genetic pairing---the pairing which brings the ØX174 genomes into close contact along their entire length---occurs when two ØX174 RF molecules of different genotypes become attached to "essential bacterial sites" in the same cell (Yarus and Sinsheimer, 1967). Likewise, specific or effective genetic pairing (Pritchard, 1955) I shall assume to be the events in which one (or more) DNA strands of one ØX174 RF molecule become hydrogen bonded (or otherwise joined but not covalently linked) to one or more of the DNA strands of the other ØX174 RF molecule.

The Holliday model (Holliday, 1962; 1964) is shown in figure 3. The initial step, termed strand separation, is postulated to begin "at a defined point, perhaps a linker between contiguous DNA molecules." While it has been shown that bacteriophage ØX174 does not contain any non-DNA linkers (Fiers and Sinsheimer, 1963), nevertheless, the existence of a strand-specific break has been documented (Francke and Ray, 1972), which may be associated with the normal mode of ØX174 recombinant formation (Benbow, Zuccarelli, Davis, and Sinsheimer, 1973). Holliday's model

Figure 3. The Holliday Model for Genetic Recombination

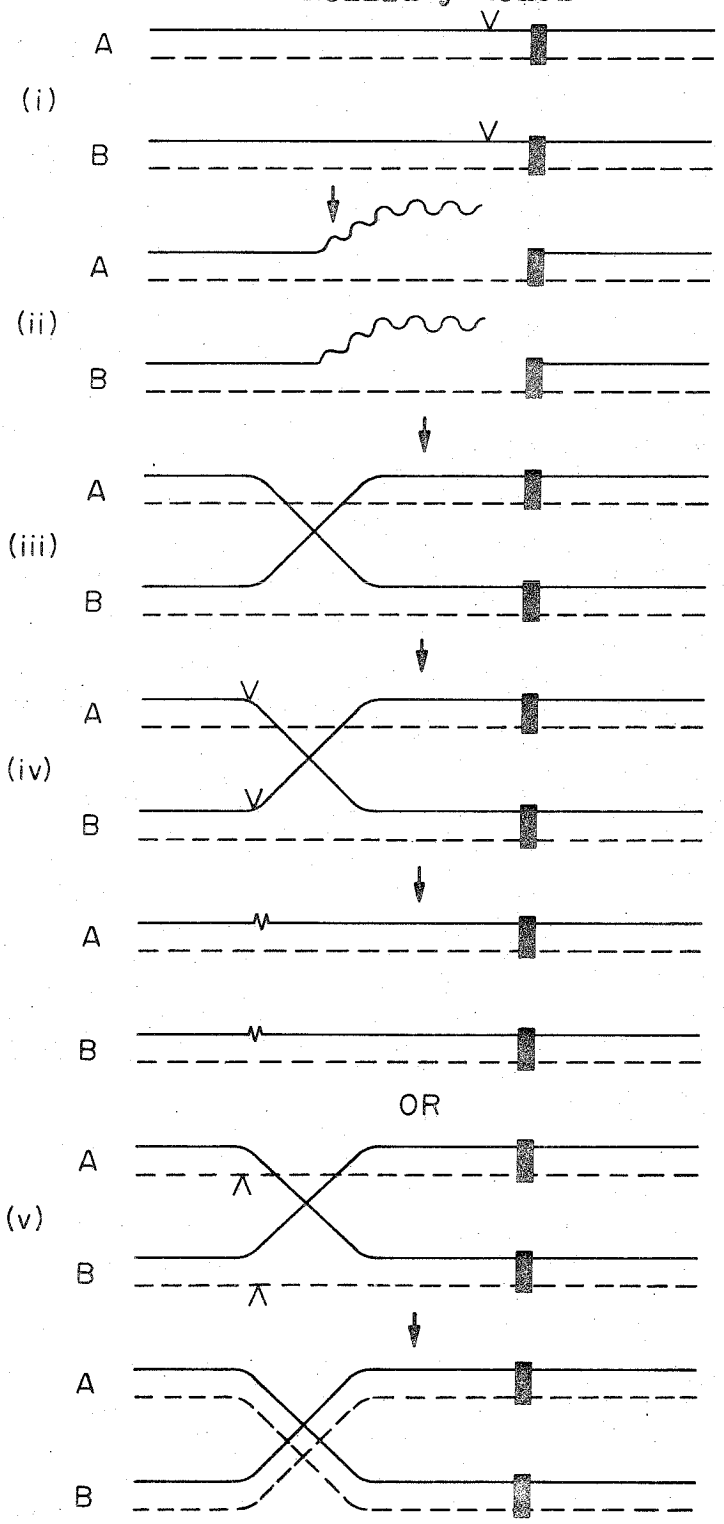
The original Holliday model (Holliday, 1964) was designed to provide a mechanism for gene conversion and non-reciprocal recombination in fungi. This mechanism, which is incorrect in several details for bacteriophage ØX174 recombinant formation, is drawn here for the value of the concepts it introduces.

Single strand breaks (arbitrarily shown here in the plus strand) occur at a defined point in the same strand of each of the two RF DNA molecules (step i). These DNA strands unwind (step ii) a random distance and each anneals to the complementary (in this case minus) strand of the other RF molecule forming a hydrogen bonded joint molecule (step iii). Single strand breaks again occur---either in the same strands as initially (step iv), or in the previously unbroken strands (step v).

The result of step iv is a single strand insertion event; the result of step v is a reciprocal crossover event. The most important point is that either step generates a hybrid DNA region. Base pairing mismatches are then corrected by DNA repair enzymes to yield non-reciprocal gene conversion with or without a reciprocal crossover event. Assuming that excision of one of the two bases in the mismatched base pair is more probable---for example, T may be easier to excise than G---this

mechanism also predicts marker-specific recombination effects (Norkin, 1970).

Holliday Model



requires both ØX174 RF molecules to be broken in the same strand (either plus or minus) at (or very near) the same site. Strand separation then proceeds along each RF molecule, presumably in the same direction by unwinding, but not necessarily ending at the same point.

At the point of initiation of strand separation, precise breakage must have occurred according to Holliday's initial model (Holliday, 1964). Our current understanding of DNA repair processes makes the stringency of this postulate unnecessary. Each of the separated strands then pairs with the complementary strand on the other RF molecule, creating a hydrogen bonded joint molecule (Anraku and Tomizawa, 1965a; 1965b).

The next step requires breakage in two strands of the same polarity (i.e. either two plus or two minus strands). If the initially broken strands are again broken, the result is two single strand insertion events. If the two previously unbroken strands are broken, the result is a single crossover event.

A single crossover event, however, results in a ØX174 double length RF DNA molecule. Benbow, Eisenberg, and Sinsheimer (1972) have shown that less than 0.5% of double length ØX174 RF DNA molecules are genetic recombinants. Assuming the procedure used to prepare double length molecules does not inadvertently select against

recombinant molecules, this implies that a large majority of ØX174 recombinants must be formed by some other mechanism.

Although the stages outlined above are sufficient to account for reciprocal crossover events, Holliday suggests that an additional mechanism operates to account for non-reciprocal recombination events such as those observed in fungal tetrad analysis. He points out that, according to his model, a hybrid DNA region extends from the first site-specific break to the second random break in each of the two genomes. If one of the two RF DNA molecules contains a mutation within this region, then each of the hybrid regions formed will contain a mispaired base.

Holliday postulates that "this condition of mispaired bases is unstable" (Holliday, 1964). (Fresco and Alberts (1960) previously had shown that unpaired bases could swing in and out of a Watson-Crick hydrogen bonded double helix without destroying its integrity). An enzyme-mediated exchange reaction---presumably by DNA repair enzymes---was suggested in order to account for correction of the mispaired regions. Depending on the direction, time of action, and base preference of the correction enzymes, various types of non-reciprocal recombinants are generated by this mechanism.

(v) The Whitehouse Model: Nonhomologous Breakage Points

Whitehouse (1963; 1964; 1965; 1968) and Whitehouse and Hastings (1964; 1965) have proposed detailed models of genetic recombination involving breakage and reunion, hybrid DNA regions, and correction by copying of base-pair mismatches. Three of these models---based on nonhomologous breakage points, homologous breakage points, and fixed homologous breakage points (the latter is often called the polaron hybrid DNA model of genetic recombination) ---are discussed below.

In the Whitehouse model of nonhomologous breakage points the events depicted in figure 4 are postulated. A single strand break occurs in complementary strands (i. e. the plus strand of one RF molecule and the minus strand of the other) of each of the two RF DNA molecules near to, but not exactly at, the same site. Following these breaks, the broken DNA strands unwind and anneal with each other using Watson-Crick base-pairing specificity. New DNA synthesis---presumably either by DNA repair or DNA replication---fills in the regions left vacant by the annealed complementary strands. Then the newly synthesized DNA strands unwind and reanneal with each other, while the unused, unpaired DNA strands are degraded by an unspecified mechanism.

In essence, therefore, this model is a breakage and

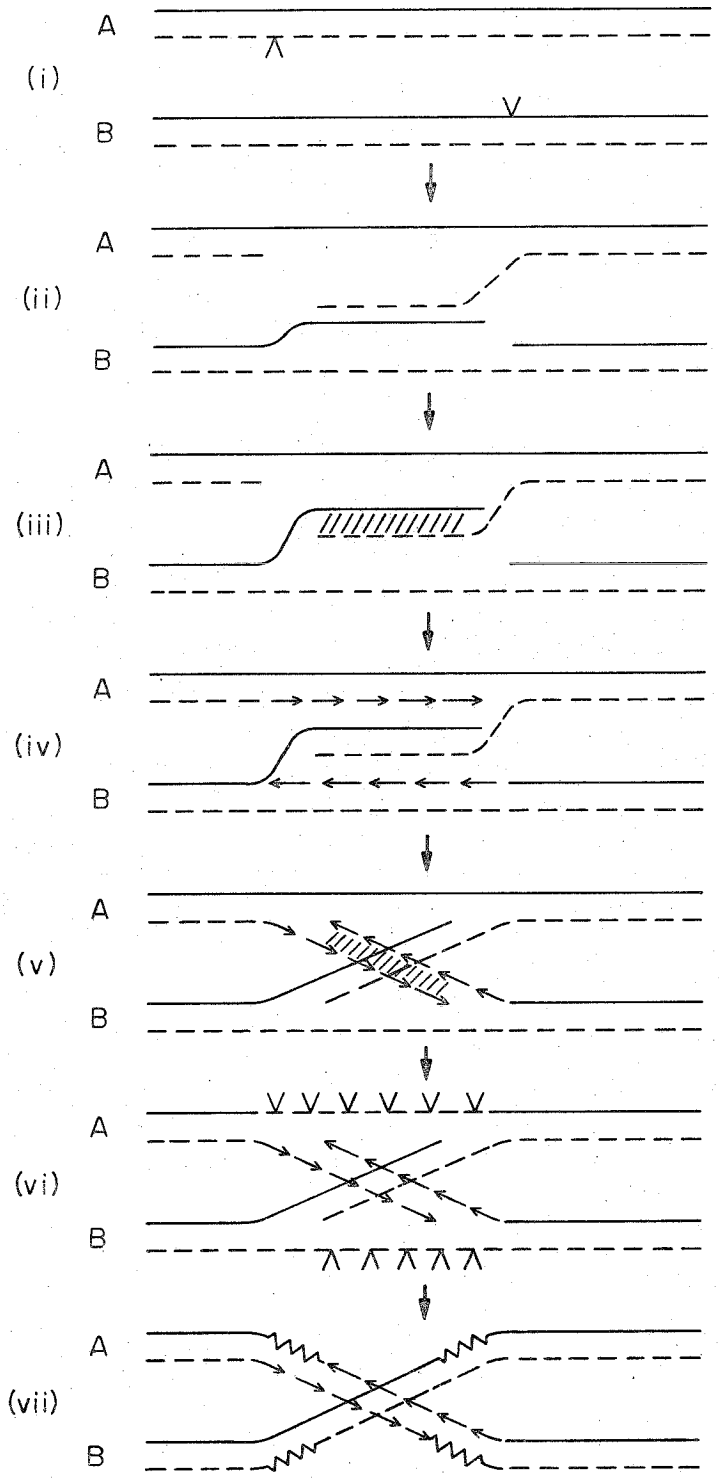
Figure 4. The Whitehouse Model: Non Homologous Breakage Points

The Whitehouse models are based on breakage and reunion coupled with a copy mechanism for the generation of non-reciprocal recombinants during gene conversion.

In step i, single strand breaks are formed in complementary strands of the two RF DNA molecules (i.e. in the plus strand of one RF molecule and the minus strand of the other). The single strands uncoil in opposite directions (step ii) and anneal with each other to yield a hydrogen bonded joint molecule (step iii). New DNA synthesis---either by DNA repair enzymes or DNA replication enzymes, but not invoking a Cairns-type replication fork---copies the original strands in the regions vacated by the now hydrogen bonded DNA strands (step iv). The newly synthesized chains also uncoil and anneal with each other (step v). Unspecified degredation enzymes remove the unpaired single strand regions remaining in the original parental RF molecules (step vi). Repair enzymes join the inserted and the parental DNA strands (step vii).

The result is a reciprocal recombinant which contains hybrid DNA regions. Non-reciprocal recombination events occur when one region of one of the paired chains is spontaneously deleted, then replaced by using the other DNA strand in the hybrid region as the template.

WHITEHOUSE MODEL I



copy model, but with an entirely different sequence of events and different consequences than the breakage and copy model proposed by Boon and Zinder (1969). One obvious prediction of the Whitehouse model---the generation of ØX174 double length recombinant RF DNA molecules---is not fulfilled (Benbow, Eisenberg, and Sinsheimer, 1972).

Whitehouse explains non-reciprocal recombination and gene conversion by the spontaneous deletion of a region of one of the paired chains followed by copying of this region from the remaining chain. Obviously a mismatched base pair in the original paired region is corrected in the copied replacement.

(vi) The Whitehouse Model: Homologous Breakage Points

A modification of the initial Whitehouse breakage and copy model incorporating homologous breakage points is shown in figure 5. Introduction of single strand breaks into the complementary strands of ØX174 RF DNA molecules <u>at the same site</u> is followed by spontaneous unwinding of both strands. Since the single strand breaks are at the same point, complementary strands are unable to pair by hydrogen bonding.

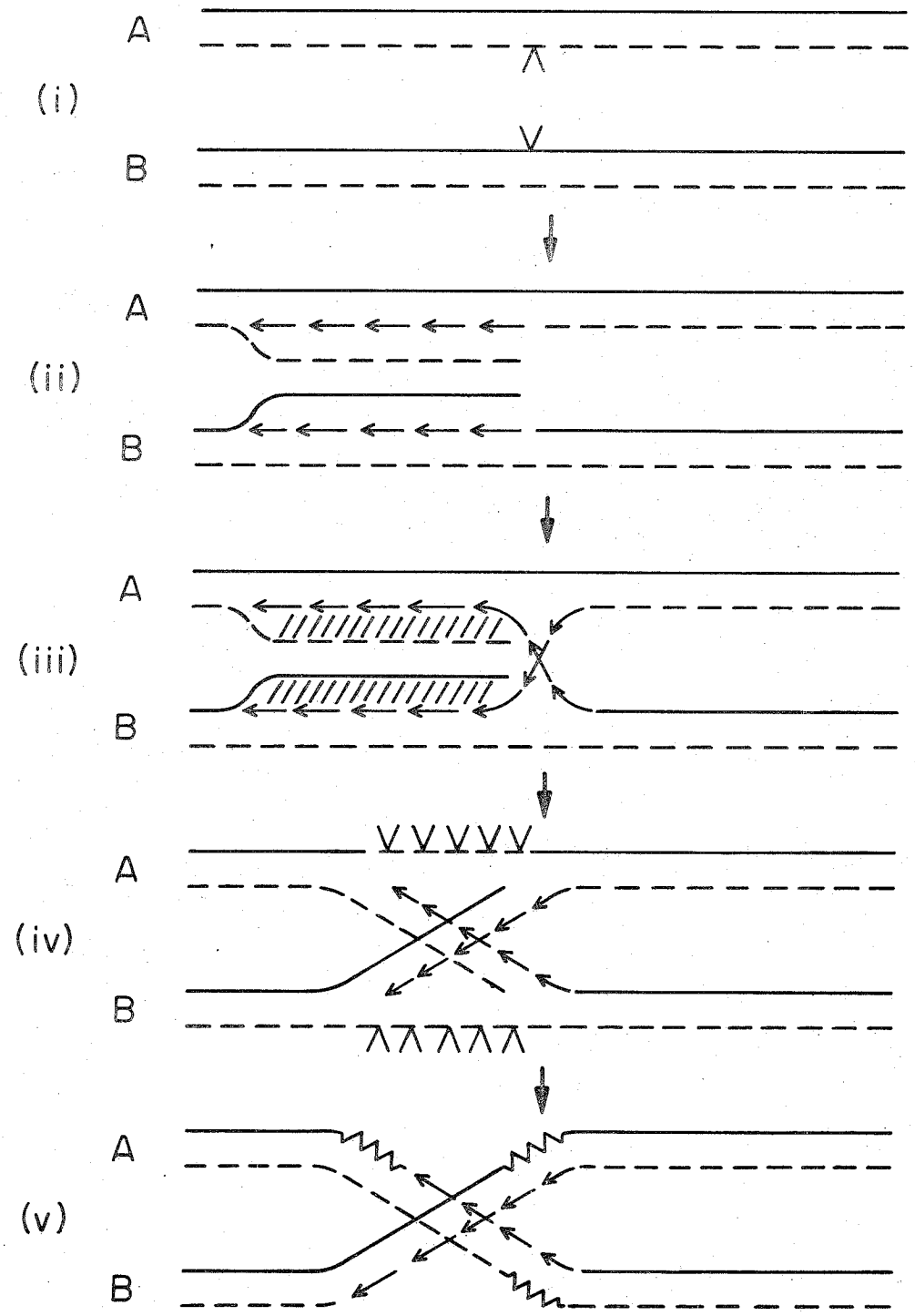
DNA synthesis commences on each genome using the single strand end as a primer and the intact DNA strand as a template. The newly synthesized DNA strands then unpair and reanneal with the <u>preexisting</u> complementary strands.

Figure 5. The Whitehouse Model: Homologous Breakage Points

The second Whitehouse model, involving homologous breakage points, is drawn below as applied to ØX174 genetic recombination. The third Whitehouse model, employing fixed homologous breakage points, differs from the second Whitehouse model only in that the single strand breaks are site-specific.

Breakage of the complementary strands of the two ØX174 RF DNA molecules at the same point (and for the third Whitehouse model at the same site) is shown in step i. The broken DNA strands uncoil; DNA replication recopies the vacated single strand regions (step ii). The newly synthesized DNA strands uncoil and simultaneously pair with the previously uncoiled parental complementary strands forming hydrogen bonded joint molecules (step iii). Parental single strand regions are degraded by unspecified enzymes (step iv); DNA repair enzymes complete the joining process (step v).

WHITEHOUSE MODEL 2 and 3



The net result is a reciprocal crossover event.

As in the previous Whitehouse model, gene conversion and non-reciprocal recombination occur by copying a deleted region of one of the strands in the hybrid region. Whitehouse also considers the possibility that the breakage points lie in the same (non-complementary) strands of the two RF DNA molecules, but concludes that his model strongly favors the occurrence of single strand breaks only in complementary strands.

(vi) The Hastings-Whitehouse Model: Fixed Homologous Breakage Points

To account for the fact that non-reciprocal recombination events can be ordered to give a genetic map based on recombination frequencies, and to suggest a reason why non-reciprocal recombination within a given gene conversion locus is polarized (i.e. why gene conversion frequencies are higher at one end of a locus than at the other), Hastings and Whitehouse (1964) added to their earlier model the concept of <u>fixed</u> homologous breakage points. The Hastings-Whitehouse model is often termed the polaron hybrid DNA model of genetic recombination.

The sequence of events in this model follows those outlined in figure 5 except that the homologous breakage points are in fixed regions of the RF DNA molecules (analogous to Stahl's fixed pairing regions (1961)). Using this

model a genetic map may be constructed from non-reciprocal recombination frequencies because the lengths of hybrid DNA regions are randomly variable. Thus, because a site close to the fixed breakage point is corrected more frequently than one far away, "the frequency of (non-reciprocal) recombination between two markers...depends on the frequency with which a length of hybrid DNA comes to an end between them." Polarity of non-reciprocal recombination presumably occurs because hybrid DNA regions consistently are extended further on one side of the fixed breakage point than on the other.

Hastings and Whitehouse (1964) also suggest that fixed breakage points are recognized by endonucleolytic enzymes which read specific nucleotide sequences (perhaps analogous to the ØX174 strand specific cistron A endonuclease (Knippers, personal communication)), and that the fixed breakage points may correspond with the origin (s) of DNA replication. In this context, it is of interest that the origin of ØX174 DNA replication (Baas and Jansz, 1972) corresponds to a high recombination region which is the proposed site of a strand- and site-specific break.

Although additional thousands of models of genetic recombination have been devised, the four discussed above invoke most of the important general features of copy choice, breakage and copy, and breakage and reunion mechan-

isms. As will be apparent, the pathway for the formation of ØX174 recombinants (Benbow, Zuccarelli, and Sinsheimer, 1973) incorporates features from each of these models. In addition, we consider the phenomenon of branch migration (Lee, Davis, and Davidson, 1970) to be indispensible for the occurrence of genetic recombination. For this idea we owe a great debt to similar conclusions about genetic recombination of T4 DNA molecules reached by Broker and Lehman (1971).

(viii) Genetic Recombination

Genetic recombination mechanisms in all organisms exhibit certain general properties (Hotchkiss, 1971). (1) Reciprocal genetic exchanges usually occur between two parental genomes (crossover events). (2) Recombination frequencies usually can be ordered to form a genetic map (additivity). (3) Recombination events can occur within a single gene (divisibility of the gene). (4) Measured recombination frequencies between distant markers are frequently smaller than the total map distance between the markers (map contraction). (5) Occurrence of a crossover event or genetic exchange often greatly increases the probability of another nearby genetic exchange (high negative interference). (6) Excessive recombination frequencies often are observed for particular markers and/or specific regions of the genome (map expansion). (7)

The particular genetic markers used in a cross may alter the observed recombination frequencies (marker-specific effects). (8) The products of single recombination events are often non-reciprocal (non-reciprocal recombination). (9) Non-reciprocal recombination events are frequently observed in close association with reciprocal crossover events (gene conversion).

Recombinant formation between bacteriophage ØX174 DNA molecules exhibits most of these properties except for the occurrence of reciprocal recombination events (Benbow, Zuccarelli, Davis, and Sinsheimer, 1973). Reciprocal recombination products are found in populations of ØX174 DNA molecules, but are rarely if ever seen as the products of a single recombination event in a single cell.

The remaining sections of this introduction are used to describe the ØX174 genome and to define the concepts of a genetic map, of synapsis, of genetic exchange, of hybrid DNA regions, and of high negative interference.

(ix) The Bacteriophage ØX174 Genome

The chromosome (genome) of bacteriophage ØX174 physically is well defined. Under the growth conditions used in these experiments, the single strand DNA molecule of the intact virus is converted to a double stranded replicative form (RF) DNA molecule immediately after infection (Cairns

and Denhardt, 1969). Each RF molecule contains a viral (parental or plus) strand and a complementary (minus) strand. The molecular weight of a wt RF molecule is 3.2×10^6 daltons (Sinsheimer, Starman, Nagler, and Guthrie, 1962); its contour length after being spread by an aqueous Kleinschmidt procedure (Davis, Simon, and Davidson, 1971) is 1.685 microns (Benbow, Eisenberg, and Sinsheimer, 1972).

ØX174 replicative form DNA molecules exist in (at least) two forms: RF I ("closed") and RF II ("nicked"). Neither strand of an RF molecule contains a non-DNA linker (Fiers and Sinsheimer, 1962). In addition, exonuclease digest experiments suggest that, with the possible exception of a small (40-100 base pairs) region of self-complementary nucleotide sequence, the entire genome consists of unique, non-repetitive sequences (Fiers and Sinsheimer, 1962).

The bacteriophage ØX174 chromosome genetically is the most densely marked of any organism. Temperature sensitive (ts), cold sensitive (cs), amber (am-UAG codon), ochre (och-UAA codon), opal (op-UGA codon), and deletion (del) mutants defining a total of ten cistrons are known (Benbow, Mayol, Picchi, and Sinsheimer, 1971). The chemical basis of most of these mutations is known (Hayes, 1968). Furthermore, isolated, purified, genetically

marked DNA molecules can be analyzed and/or directly employed to initiate the infection process using a spheroplast assay (Guthrie and Sinsheimer, 1962; Benbow, Zuccarelli, Davis, and Sinsheimer, 1973).

In this thesis I have attempted to exploit these straightforward physical and genetic properties in order to probe the process of ØX174 genetic recombination.

(x) Genetic Maps

The first genetic map in any organism was constructed in Drosophila by Sturtevant (1913). A circular genetic map for bacteriophage S13 (a very close relative of bacteriophage ØX174) was constructed by Baker and Tessman (1967). In this thesis I use the genetic map of ØX174 to ask a slightly different question.

Does a genetic map represent---in a quantitative fashion---the number of nucleotides between genetic markers in the DNA molecule? For small regions (Benzer, 1959) and for single genes (Yanofsky, Carlton, Guest, Helinski, and Henning, 1964) of large genomes, proportional relationships between genetic and physical distances have been found. Often, however, these relationships break down on close examination---as in the case of marker-specific recombination effects in Escherichia coli (Norkin, 1970). In this thesis the ØX174 genome is mapped in detail and its gene products are identified: as a result the genetic

and physical maps for an entire genome can be compared.

(xi) Genetic Recombination: Synapsis

Any model of genetic recombination must have two additional a priori features (Bodmer and Darlington, 1969): synapsis---"the alignment of chromosome segments"; and exchange---"the recombining of information or substance carried by the synapsed homologs". ØX174 synapsis is directed by base pairing. Hydrogen bond formation between nucleotides in the complementary strands of the DNA molecules is the only well documented process with the necessary precision. In this thesis synapsis via hydrogen bond formation using Hoogsteen and reversed Watson-Crick pairing is suggested as an alternative to Watson-Crick base pairing.

Another key question about synapsis for which ØX174 recombinant formation suggests a solution is whether there exist large regions of continuous pairing with a low probability of recombination, or whether there exist small discontinuous pairing regions with a high probability of recombination. In experiments characterizing bacteriophage T4 heterozygote formation, Steinberg and Edgar (1962) have shown that the second alternative (called multiple discontinuous switches) is required in the region of pairing as initially proposed by Pritchard (1955; 1960). Note, however, that these short pairing regions in T4 may be only single strand switches (i.e. only one of the two

DNA strands needs to be copied and inserted). Evidence presented in this thesis suggests that only short discontinuous regions synapse; when effective pairing does occur, the probability of genetic recombination is very high.

(xii) Genetic Recombination: Exchange

Does each parental chromosome contribute part of its physical material to the recombinant molecule or only its information content? For bacteriophage lambda Meselson (Meselson and Weigle, 1961) and Kellenberger (Kellenberger, Zichichi, and Weigle, 1961) have shown that parental density labeled DNA is found both in unrecombined parental molecules and in unreplicated recombinant molecules. Furthermore, the amount of labeled material found in the recombinant DNA molecules corresponded roughly to the parental contribution expected on the basis of the genetic map positions of the markers. In this context it should be noted that Kaiser (1962) has correlated the physical and genetic maps of bacteriophage lambda using a DNA transformation assay.

In later work Meselson (Meselson, 1964) showed that very little DNA synthesis is required for the formation of recombinant DNA molecules. Meselson's results apparently exclude for lambda recombinant formation the breakage and copy mechanism proposed by Boon and Zinder (1969) for

bacteriophage f1 recombinant formation. The initial recombinant lambda DNA molecule composed of unreplicated pieces of two parental duplexes was termed the primary recombinant (Meselson, 1964). It should be noted that these experiments do not provide a satisfactory molecular model to account for the way double strand DNA molecules partially unwind and exchange physical material.

In this thesis I show that ØX174 parental DNA molecules undergo similar breakage and reunion during rec A "enzyme" mediated genetic recombination (Benbow, Schafer, and Sinsheimer, 1973).

(xiii) Hybrid DNA Regions

If the breakage sites on each of the two recombining DNA molecules are not perfectly aligned during the formation of recombinants, hybrid DNA regions are formed. "At the site of joining, the complementary polynucleotide chains of a given parental segment terminate at different levels, separated by approximately 10^3 nucleotide pairs, so that complementary chains of different parentage overlap one another to form a hybrid region" (Meselson, 1966).

Estimates of the size of the DNA hybrid regions formed during recombinant formation have been made for several organisms including T4: 100 to 5,000 base pairs by Drake (Drake, 1967), and 10^3 base pairs or four map units by Barricelli (Barricelli and Doermann, 1960); lambda: 1.5×10^3

base pairs by Amati and Meselson (1965); and bacterial DNA transforming fragments: 450 base pairs minimum size by Cato (Cato and Guild, 1968).

Excision mechanisms (DNA repair type) for correction of hybrid DNA region mismatches have been suggested by Taylor (Taylor, Haut, and Tung, 1962), Holliday (1962; 1964), Whitehouse (1963), and Meselson (1964; 1965).

The occurrence of joint molecules (Anraku and Tomizawa, 1965a; 1965b) during recombinant formation in bacteriophage T4, supports the idea that recombination involves the "association of homologous DNA molecules by base pairing in regions of hybrid overlap, followed by repair of single-stranded gaps and formation of covalent bonds" (Bodmer and Darlington, 1969). Note, however, that Tomizawa's experiments (and those of Kozinski (1969) as well) do not directly correlate genetic recombinants with physical structures. Their experiments are subject to the criticism implicit in Broker and Lehman's (1971) demonstration that almost any T4 DNA structure can be seen if you look hard enough.

In this thesis the size of the average hybrid DNA regions involved in ØX174 recombinant formation is estimated to be roughly 375 nucleotides (Benbow, Schafer, and Sinsheimer, 1973).

(xiv) High Negative Interference

High negative interference (Hayes, 1968)---"a crossover in one region tends to be accompanied by a second in the adjacent region"---is another phenomenon initially observed in higher organisms for which ØX174 provides an excellent model system. The work of Meselson and Amati (Amati and Meselson, 1965) with high negative interference during bacteriophage lambda recombination is the best documented study of high negative interference in bacteriophages.

High negative interference during bacteriophage ØX174 recombination was demonstrated by Benbow, Hutchison, Fabricant, and Sinsheimer (1971) during three- and four-factor crosses. Apparently the small circular genome makes high negative interference an almost obligatory accompaniment to genetic exchange. Furthermore, since repeated or multiple genetic exchanges are uncommon during ØX174 recombination, it is easy to observe unobscured high negative interference.

The mean length of a high negative interference cluster region corresponds roughly to the length of a gene in bacteriophage T4 (Stahl, Edgar, and Steinberg, 1964). Taking advantage of our knowledge of absolute distances on the ØX174 genome, the average length of genetic exchange regions in ØX174 may be estimated

physically using the genetic phenomenon of high negative interference.

The high negative interference during genetic recombination can be measured by the index of interference, i, where i is defined as:

$$i = R_d/R_{12}R_{23} .$$

R_d is the measured frequency of double exchanges in the region between markers 1 and 3; R_{12} and R_{23} are the measured two-factor recombination frequencies between markers 1 and 2 or 2 and 3 (Amati and Meselson, 1965). Since the distances between markers (in nucleotides) are known for ØX174, it is possible to calculate from R_d the percentage of double exchanges which occur in the known distance between markers 1 and 3. By varying the distances between, and locations of, markers 1, 2, and 3, and by determining the percentage of double exchanges occurring in each region, a distribution of the size of <u>all</u> genetic exchanges over the entire ØX174 genome may be plotted as a function of distance.

In this thesis the size of the average genetic exchange region is estimated; this estimate is consistent with our proposed model of the pathway of formation of ØX174 recombinants.

THE HYPOTHESIS

A general model for genetic recombination between two DNA molecules is drawn in figure 6. This model represents a conscious attempt to use the fewest, simplest, best documented biochemical events that result in genetic exchanges, i.e. to use a biological principle of least action.

Figure 6. The Benbow-Sinsheimer Model of Genetic Recombination

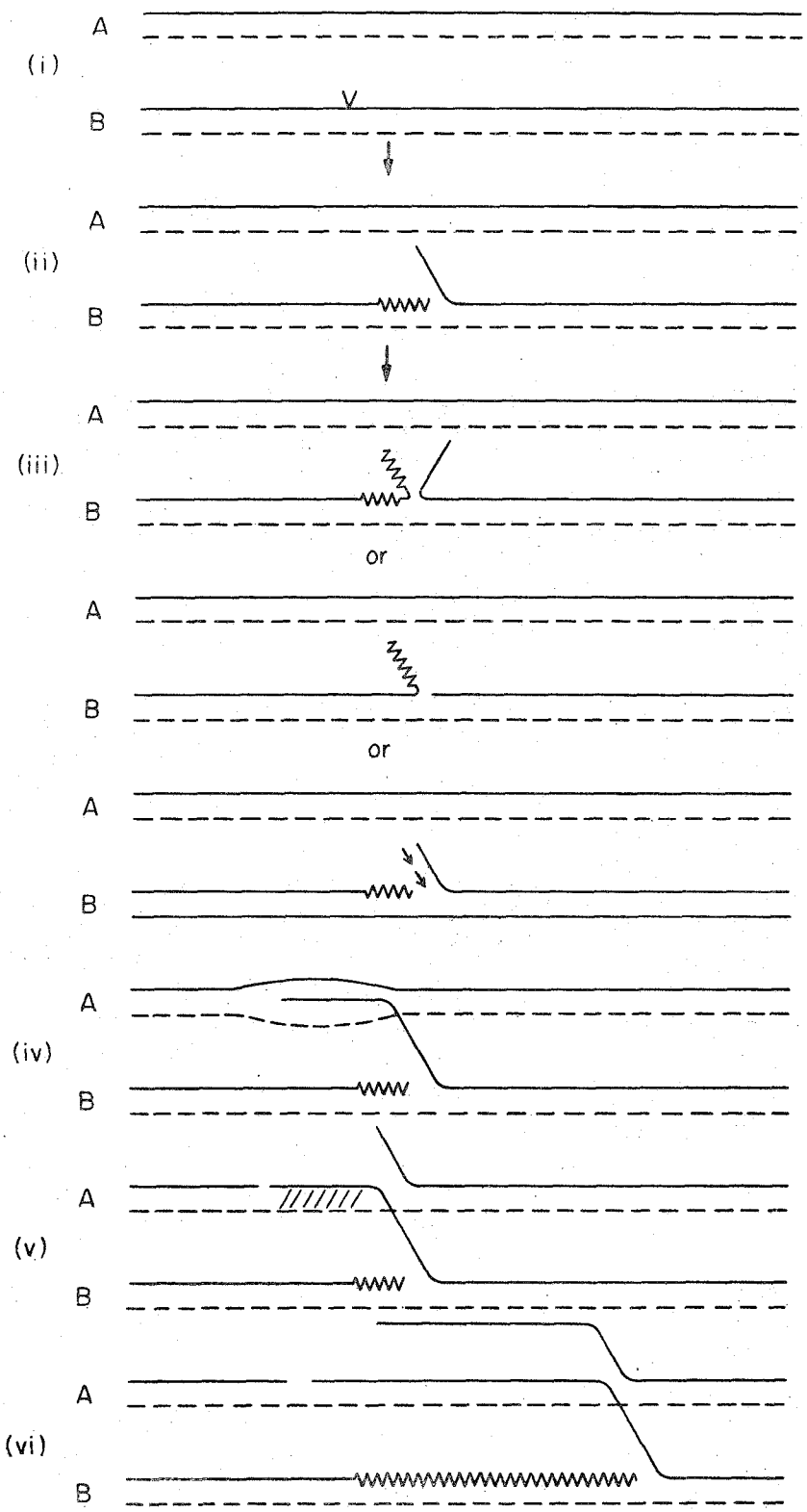
A hypothetical general model for the primary mechanism of genetic recombination between two ØX174 DNA molecules was used as the rationale for the experiments described in this thesis.

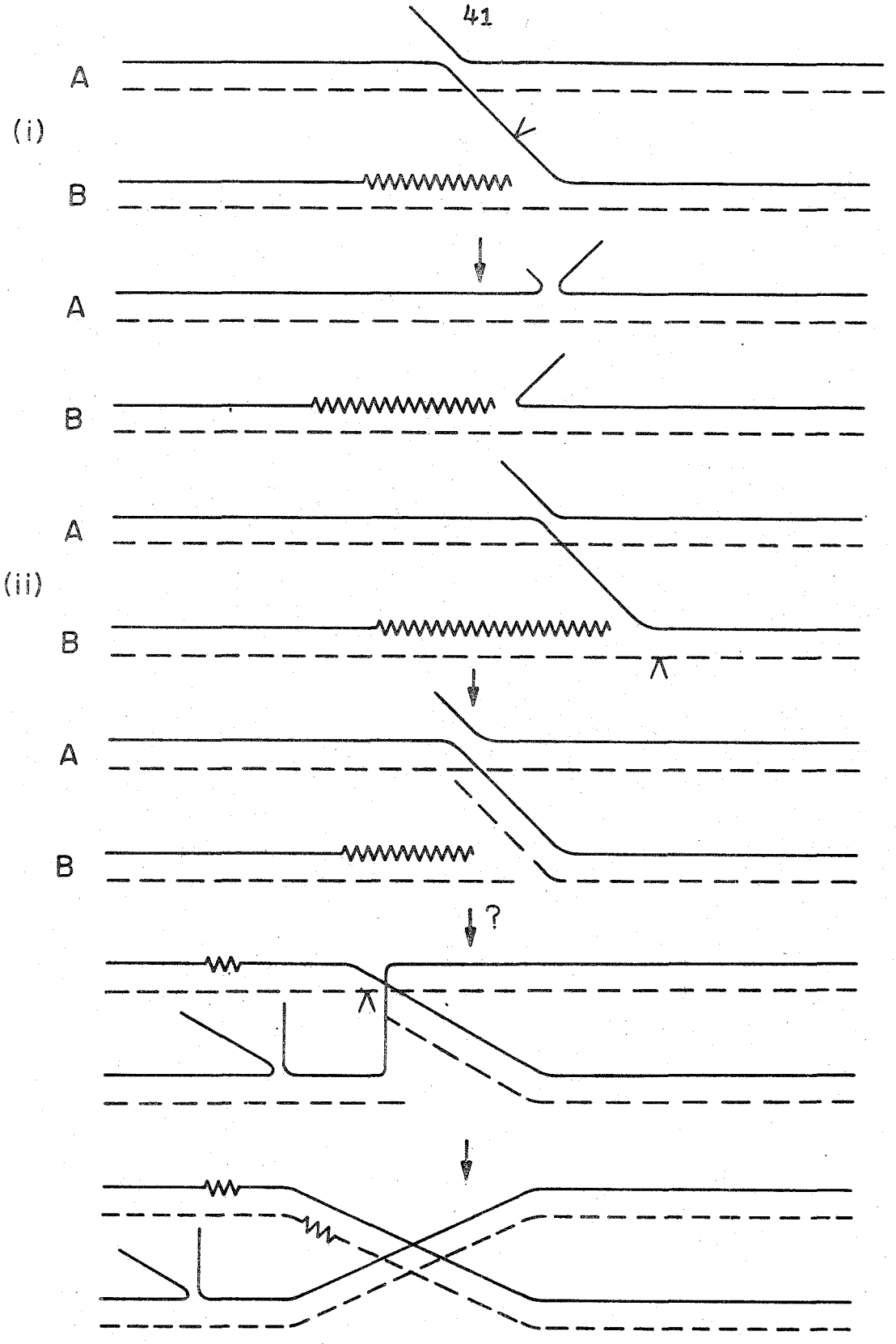
A single strand "break" occurs in one of the two strands of one of the two RF DNA molecules (step i). Displacement synthesis results in the extension of a single stranded DNA "tail" (step ii). Branch migration (step iii), possibly coupled with further DNA synthesis (step iii, choice 3), results in any one of several recombination proficient structures. Most commonly, the extended single stranded DNA "tail" attacks the intact RF DNA molecule, possibly using reversed Watson-Crick and Hoogsteen base pairing to form a triple helix (step iv). A single strand scission is induced in one of the two DNA strands: one of the most likely structures to be formed is a hydrogen bonded joint molecule. Further branch migration occurs creating an extensive DNA hybrid region.

Two predicted consequences of the Benbow-Sinsheimer model are shown in figure 6b. If a second single strand scission occurs in the initially broken strand, the net result is a single strand insertion event (step i).

I shall assume that the remaining "tails" can easily be eliminated---perhaps by a single strand assimilation mechanism (Cassuto and Radding, 1971)---to regenerate circular parental RF DNA molecules.

Alternatively, if a second single strand scission occurs in the complementary strand of the initially broken RF or in the complementary strand of the other RF, one likely sequence of events (step ii) leads to a reciprocal recombination event. This occurrence is less likely for small genomes like that of ØX174 (see section vi).


A
(i)
V
B
A
(ii)
B
A
(iii)
B
or
A
B
or
A
B
A
(iv)
B
A
(v)
B
A
(vi)
B


A
(i)
B
A
B
A
(ii)
B
A
B
?

METHODS AND MATERIALS

The exact procedures used in each section are described in the legends to each figure.

Pedigrees of the bacterial strains used in this thesis are given in Benbow, Zuccarelli, Davis, and Sinsheimer (1973). In general, strains were derived from one of three separate parents: E. coli C, the standard non-permissive host strain (Sinsheimer, 1959); E. coli HF4704, a non-permissive, excision defective host strain (Lindquist and Sinsheimer, 1968); and HF4714, a permissive (for am mutations) host strain (Benbow, Hutchison, Fabricant, and Sinsheimer, 1971).

The ØX174 mutants used in this thesis are described in table 1.

The procedure used to measure genetic recombination between ØX174 DNA molecules is extremely sensitive to quantitative variations. For this reason, only those effects which resulted in order of magnitude changes in recombinant formation were studied. The exact procedure for measuring the genetic recombination of bacteriophage ØX174 is described below (Benbow, Hutchison, Fabricant, and Sinsheimer, 1971).

HF4714 (or an isogenic derivative) is grown in KC broth at 37 ^{o}C with aeration to a concentration of 1 x 10^{8} cells/ml as measured in a Petroff-Hauser counter. KC broth is 0.5% KCl, 1% tryptone, and 0.001 M $CaCl_2$. The cells are

Table 1

The Bacteriophage ØX174 Genome

Cistron	Function	MW	Location	Mutants	Location
A	RF DNA Replication	13,500	0-325	am 86	24
				am 50	48
				ts 128	105
				am 33	165
				am 8	240
				am 30	283
				am 35	294
				am 18	305
B	SS DNA Replication; Virion protein(?)	25,000	326-928	ts 9	351
				am 16	374
				am 14	640
				ts 116	825
				och 5	900
C (+I)	SS DNA Replication	34,000 (+7,000)	929-1917	och 6	1050
D	Major SS DNA Replication	14,500	1918-2267	am H81	2002
				am 42	2002
				am 10	2039
E	Host Cell Lysis	17,500	2268-2689	am 27	2413
				am 3	2485
J?	Structural Component of Virion	9,000	2690-2906	am 6	2714

Cistron	Function	MW	Location	Mutants	Location
F	Capsid	50,000	2907-4112	op 9	2916
				op 6	2931
				tsh 6	3345
				am H57	3884
				am 88	3920
				am 87	3944
				ts 41D	3993
				am 89	4058
G	Spike	20,500	4113-4606	am 9	4137
				ts Y	4205
				ts 79	4450
				am 32	4525
H	Spike	37,000	4607-5500	am 90	5100
				am 23	5185
				am N1	5235
				am 80	5295
				ts 4	5330

The material in this table is calculated from the data given in Benbow, Hutchison, Fabricant, and Sinsheimer, 1971; Benbow, Mayol, Picchi, and Sinsheimer, 1972; and Sinsheimer, 1968.

These numbers do not represent actual nucleotide position assignments since they were calculated on the basis of ØX174 protein molecular weights assuming a 5,500

nucleotide genome with no nonessential gene regions. Instead, these numbers are aids for calculating marker separations (figures 8,9, and 20); nevertheless, the nucleotide distances should be accurate to 5% in most cases, particularly in the region from cistron D through cistron H.

made 0.003 M KCN and aerated for 10 minutes at 37 °C. Phage stocks at about 5×10^{10} plaque forming units (pfu) per ml in 0.05 M borate, pH 8.0, are freshly diluted to 2×10^{8} pfu/ml in KC broth: 0.003 M KCN. 0.5 ml of each of two phage stocks plus 0.2 ml of the KCN-treated bacteria are mixed in an adsorption tube in an ice bath; phage are allowed to adsorb for 15 minutes at 37 ° C. 0.2 ml of the phage-cell complexes are diluted into 20 ml of KC broth at 37 °C (or at 32 °C for crosses using temperature sensitive mutants) and aerated for 90 minutes. The resulting bursts are titered immediately under appropriate conditions after dilution through Denhardt starvation buffer (Denhardt and Sinsheimer, 1965).

RESULTS

(i) The Genome of Bacteriophage ØX174

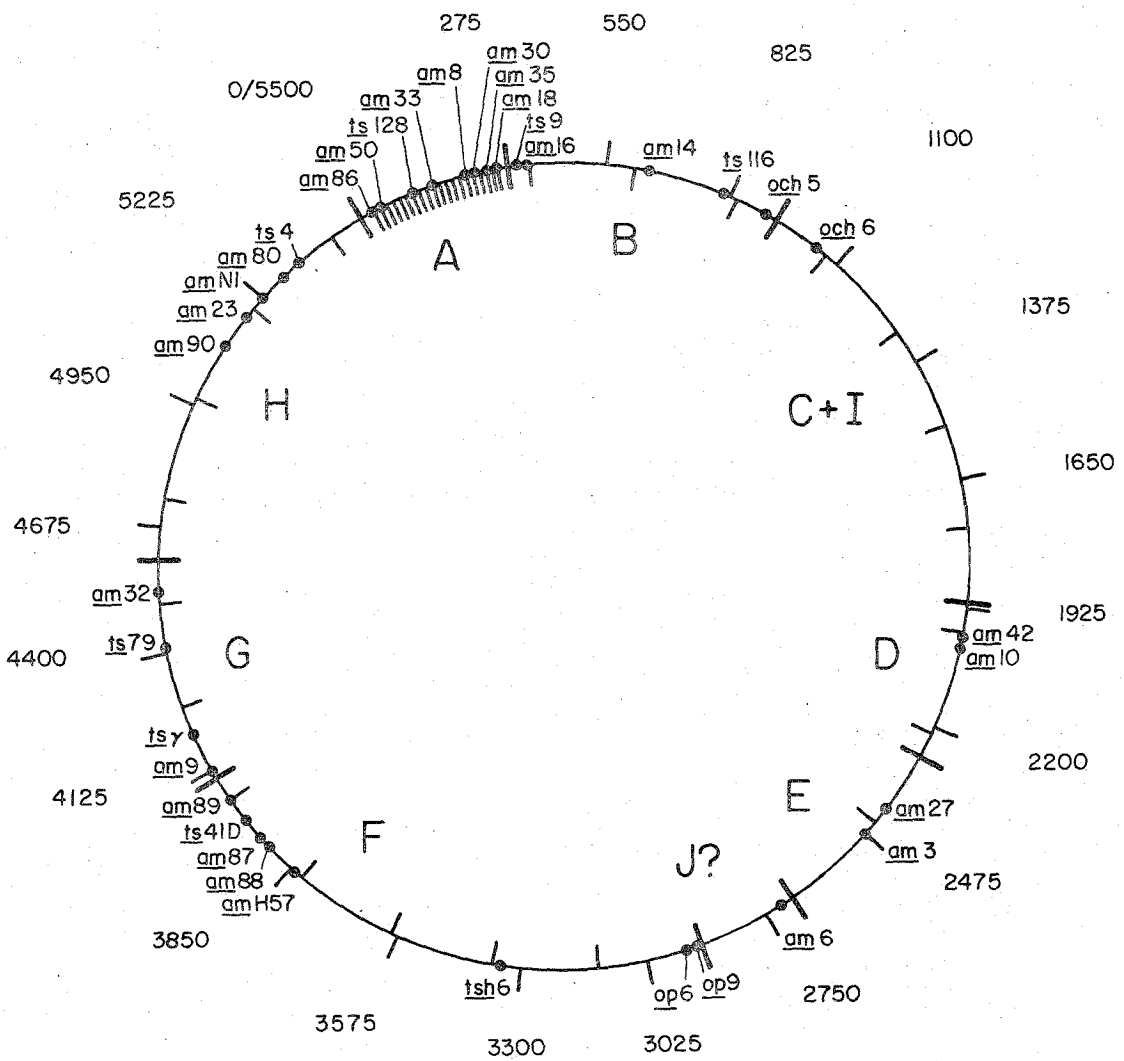
A genetic map of bacteriophage ØX174 was constructed from the results of two and three factor genetic crosses by Benbow, Hutchison, Fabricant, and Sinsheimer (1971). The protein coded by each of the ten ØX174 cistrons is identified in SDS-polyacrylamide gel electrophoresis patterns by Benbow, Mayol, Picchi, and Sinsheimer (1972). The genetic map, redrawn to include accurate cistron boundaries based on these protein molecular weight determinations and corrected for marker-specific recombination effects, is presented by Benbow, Zuccarelli, Davis, and Sinsheimer (1973).

This genetic map, reflecting actual physical distances (i.e. the number of nucleotides between two genetic markers), is presented in figure 7. The genetic recombination frequency between any two markers separated by less than 30^{o} arc may be read directly off this map, assuming neither of the markers is subject to a large marker-specific recombination effect (see below). The units are wt recombinants per total progeny phage. An estimate of the physical location of each genetic marker and the molecular weight of each cistron product is given in Table 1.

The phenomenon of map contraction occurs for markers separated by more than 30^{o} arc; as a result, the measured

Figure 7. The Genetic Map of Bacteriophage ØX174

The order of cistrons is established by the three factor crosses described in Benbow, Hutchison, Fabricant, and Sinsheimer, 1971. Cistron boundaries are drawn using the cistron product molecular weights in Benbow, Mayol, Picchi, and Sinsheimer, 1972. One map unit (divisions on inner circumference of the genetic map) represents 1×10^{-4} wt recombinants per total progeny phage. The physical location (in nucleotides) was estimated for am 42(D), am 10(D), am 3(E), am 27(E), am 6(J?), op 6(F), am H57(F), am 88(F), am 87(F), am 89(F), am 9(G), am 32(G), and am N1(H) on the basis of polypeptide fragments in SDS-polyacrylamide gels. Distances in nucleotides are indicated by the divisions on the outer circumference of the genetic map. Cistron A is arbitrarily selected as the origin. The number of nucleotides in the complete genome is assumed to be 5,500 (Sinsheimer, 1968). The total length of the bacteriophage ØX174 genome is 24.4×10^{-4} wt recombinants per total progeny phage.

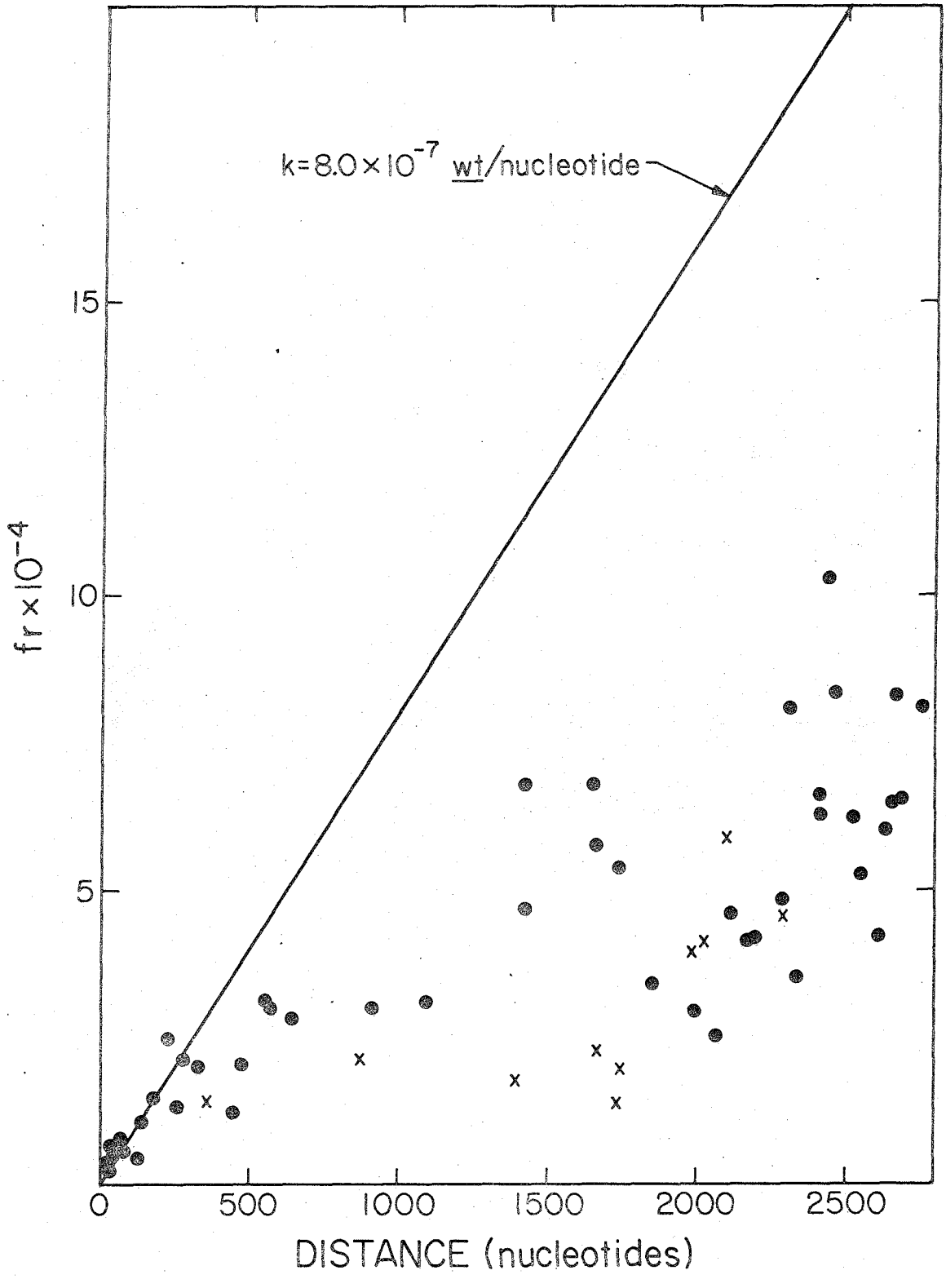

275
550
825
1100
1375
1650
1925
2200
2475
2750
3025
3300
3575
3850
4125
4400
4675
4950
5225
0/5500
am30
am8
am35
am33
am18
ts128
ts9
am50
am16
am14
ts116
am86
och5
och6
ts4
am80
amN1
am23
am90
am32
ts79
ts7
am9
am89
ts41D
am87
am88
amH57
tsh6
op6
op9
am6
am27
am3
am42
am10
A
B
C+I
D
E
F
G
H
J?

recombination frequency is lower than the value read off the genetic map (Edgar, Feynman, Klein, Lielausis, and Steinberg, 1962). As shown in figure 8, map contraction in bacteriophage ØX174 two and three factor genetic crosses can occur with markers separated by only 300 nucleotides, and is universal for markers separated by more than 500 nucleotides. This extreme map contraction presumably results from the high negative interference observed by Benbow, Hutchison, Fabricant, and Sinsheimer, (1971), and from the necessity for <u>two</u> formal genetic exchanges to occur in order to regenerate a unit length circular genome (Baker and Tessman, 1967).

The direction of translation of the bacteriophage ØX174 genome follows cistron order D-E-J?-F-G-H-A-B-C. To establish this, the position of a nonsense mutation on the genetic map was compared with the physical size (molecular weight) of the appropriate protein fragment generated in non-permissive cells (Benbow, Mayol, Picchi, and Sinsheimer, 1972). This direction of translation is confirmed by the complementation behavior of several strong polar mutations (<u>op</u> 6 in cistron F and <u>am</u> 9 in cistron G), and by the molar ratios of ØX174 proteins synthesized <u>in vivo</u> (cistron D protein is synthesized in the largest molar quantity, followed by the proteins of cistrons J?, F, G, and H in progressively decreasing

Figure 8. Map Contraction in Bacteriophage ØX174 Genetic Recombination

Recombination frequencies (wt recombinants per total progeny phage) measured in two-factor genetic crosses (Benbow, Hutchison, Fabricant, and Sinsheimer, 1971) are plotted against the physical distance between the markers (in nucleotides). For distances less than 300 nucleotides, recombination frequency and distance are directly proportional. Thus, $f_r = kd$ with $k = 8.0 \times 10^{-7}$ wt recombinants per nucleotide. Note that not all available data are plotted in this region due to crowding. For marker separations greater than 300 nucleotides, the measured recombination frequency is considerably less than the map distance. In addition, slight marker-specific effects on genetic recombination, for example the depression of recombination frequencies in crosses with the marker am 10(D) (crosses signified by an x) , are also apparent. For any given marker, however, two-factor recombination frequencies are a function (although nonlinear) of distance from other markers.


k=8.0×10⁻⁷ wt/nucleotide
fr×10⁻⁴
15
10
5
0
500
1000
1500
2000
2500
DISTANCE (nucleotides)

amounts).

The contour length of bacteriophage ØX174 monomeric replicative form DNA measured in the electron microscope is 1.685 ± .051 microns (Benbow, Eisenberg, and Sinsheimer, 1972; Kleinschmidt, Burton, and Sinsheimer, 1963). To obtain mutants of bacteriophage ØX174 whose appearance (contour length) is distinguishable when viewed in the electron microscope, the dispensibility of the cistron E protein (the "lysis" protein; Hutchison and Sinsheimer, 1966) was exploited. Zuccarelli, Benbow, and Sinsheimer (1972) isolated a population of deletion mutations of bacteriophage ØX174 which: (i) are genetically defined within the cistron E region; (ii) are physically defined as individual deletions with an average size of 7% of the ØX174 genome (about 390 nucleotides); (iii) are able to reproduce as efficiently as ØX174 wt or am 3 (E); and (iv) possess an essentially normal viral coat structure and all viral functions except host cell lysis. These ØX174 del E mutants represent the first genetically and physically characterized deletion mutants of a small DNA bacteriophage.

The ten putative cistron products of bacteriophage ØX174 require the entire genome coding capacity (95%; Benbow, Mayol, Picchi, and Sinsheimer, 1972). Zuccarelli, Benbow, and Sinsheimer (1972) show that there are no non-

essential regions in the bacteriophage ØX174 genome of more than 50-100 nucleotides: ØX174 deletion mutant viral (plus) strands were hybridized with ØX174 wt complementary (minus) strands. The resulting heteroduplex DNA molecules were examined in the electron microscope using the procedure of Davis, Simon, and Davidson (1971). No multiple deletion loops were detected out of more than 500 molecules examined. Since our population of deletion mutants was strongly selected on the basis of extensive DNA losses, and since deletion loops of more than 50-100 nucleotides may be visualized with this technique, an upper limit for the size of non-essential regions of the genome is established.

Distances on the bacteriophage ØX174 genetic map and distances on a physical map constructed from the molecular weights of ØX174 proteins and protein fragments are proportional over most of the genome with the exception of a high recombination region within cistron A (Benbow, Mayol, Picchi, and Sinsheimer, 1972). This rough proportionality of the genetic and physical genomes implies that the genetic recombination frequencies measured in two factor genetic crosses and the segregation ratios measured in three factor genetic crosses approximately reflect the number of nucleotides between the markers---marker specific effects excepted. Thus, the recombination frequency between

two markers may, to a first approximation, be expressed as a linear function of the distance between the two markers:

$$f_r = kdz \qquad \text{(equation 1)}$$

where f_r is the recombination frequency in units of wt recombinants per total progeny phage; $k = 8.0 \times 10^{-7}$ wt recombinants per nucleotide; d is the number of nucleotides between the two markers; and z is the percentage of exchanges which do not involve a second exchange between the two markers (i.e. a correction for map contraction or high negative interference). A corollary of this relationship is that a physical distance such as the contour length of a replicative form DNA molecule viewed in the electron microscope can be used to determine the approximate location on the genetic map of a physical marker.

Marker-specific effects in genetic recombination (Norkin, 1970) are seen in two and three factor genetic crosses with several bacteriophage ØX174 markers. A marker-specific effect is defined by the criterion that the recombination frequency is primarily determined by the base substitutions at the mutant (and neighboring) sites. Distance (number of nucleotides between the two markers) is not correlated with recombination frequency over the range of 10 to 5,000 nucleotides for markers exhibiting marker-specific recombination effects. Marker-specific

effects for several bacteriophage ØX174 nonsense mutants are shown in figure 9. Thus am 23 (H), a marker in cistron H, lies within 50 nucleotides of several other cistron H nonsense mutations. These other cistron H mutants, am N1 (H) for example, exhibit "normal" recombination frequencies with outside markers which are a function of distance as shown in figure 8. The recombination frequencies measured when am 23 (H) is crossed with outside markers are invariably much lower. A similar depression of recombination frequencies is seen for the och 6 (C) mutation in cistron C.

In addition, most ØX174 mutations appear to show slight marker-specific effects. For example, as shown in figure 8, crosses involving am 10 (D) typically exhibit slightly depressed recombination frequencies which, nevertheless, accurately reflect distances on the genetic map. To account for these, equation 1 may be rewritten:

$$f_r = m_1 m_2 kdz \qquad \text{(equation 2)}$$

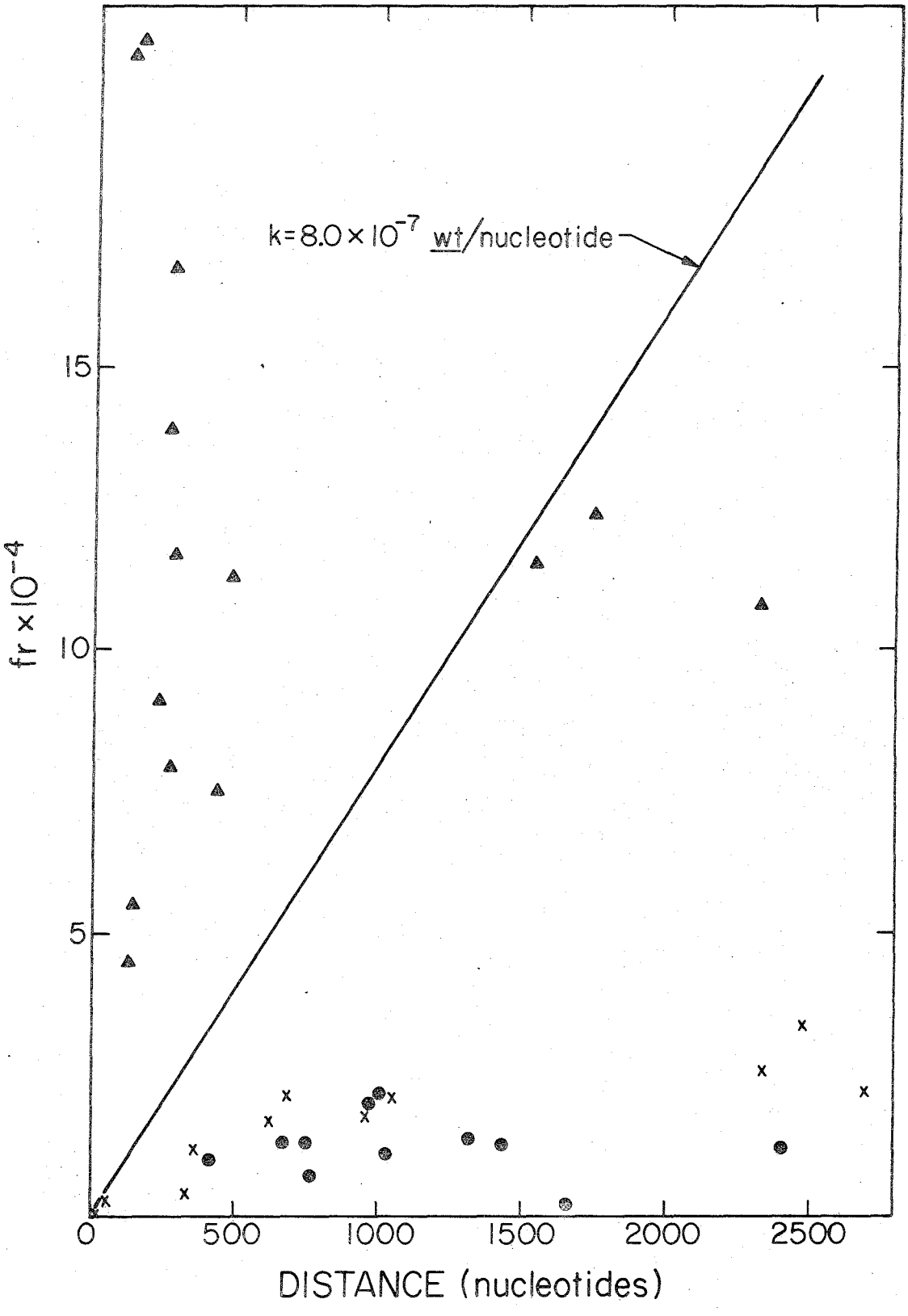
where m_1 and m_2 represent the slight marker-specific effects of markers 1 and 2. The marker-specific effects, m_n, are defined as:

$$m_n = 1 - e^{-c_n d}$$

where c_n is a constant for marker n. The constants, c_n, for each marker n have values in the range of 6-10 x 10^{-3} for slight marker-specific effects, and less than 1 x 10^{-3}

Figure 9. Marker-Specific Effects and Map Expansion in Bacteriophage ØX174 Genetic Recombination

Recombination frequencies (wt recombinants per total progeny phage) are plotted against the distance between markers as in figure 8. Two-factor crosses between am 23(H) (x) or och 6(C) (●) and outside markers illustrate marker-specific depression of recombination frequencies. Two-factor crosses between mutants in the high recombination region of cistron A (▲) and outside markers or other cistron A mutants demonstrate the phenomenon of map expansion. Note that both marker-specific recombination effects and site-specific recombination effects mask the distance dependence of recombination frequencies over the entire ØX174 genome.


k=8.0×10⁻⁷ wt/nucleotide
fr×10⁻⁴
15
10
5
0
500
1000
1500
2000
2500
DISTANCE (nucleotides)

for pronounced marker-specific effects. Since the distances, d, in two factor crosses with our ØX174 mutants typically range from 250 to 2,500 nucleotides, the m_n's for slight marker-specific effects may be represented by constants. These definitions also agree with the finding that at large distances (1×10^6 nucleotides) marker-specific effects are no longer observed (Norkin, 1970).

The only assumptions involved in these definitions are that (1) each base substitution has a constant effect which depends only on the substitution and the surrounding nucleotides; and (2) this effect varies as a function of the distance in nucleotides from the base substitution. These definitions are easily extended to account for the observation that a substitution at a locus near to, but distinct from the marker site also can alter the recombination frequency.

The high recombination region observed in the cistron A region (Benbow, Hutchison, Fabricant, and Sinsheimer, 1971) probably results from an entirely different effect. As shown in figure 9, intracistronic recombination frequencies within cistron A, and all recombination frequencies involving markers in the cistron A high recombination region (see figure 11) exhibit the phenomenon of map expansion (Holliday, 1964). An hypothesis to account for these site-specific recombination effects is presented

below based on the postulated existence of a site (and strand) specific "break" related to the origin of ØX174 replicative form DNA replication. Using similar assumptions to those employed in the definitions of marker-specific effects on recombination, a site specific factor, A_n, can be defined for each marker:

$$A_n = 1 + a_n e^{-wd_{an}}$$

where a_n is a general constant of order of magnitude 1×10^6. This value of a_n corresponds to the assumption that the site-specific "break" is roughly 1×10^6 times more probable than any other random single "break". The distance d_{an} is the number of nucleotides from the site-specific "break" to the marker n. w is a constant related to the effective stimulatory range of the site-specific "break".

The specific mapping function resulting from this modification of equation 2 for two markers is:

$$f_r = (1-e^{-c_1 d})(1-e^{-c_2 d})(1+e^{-wd_{a1}})(1+e^{-wd_{a2}})kdz \quad \text{(equation 3)}$$

It is of interest to examine the physical meaning of these proposed constants. $k = 8 \times 10^{-7}$ wt recombinants per nucleotide represents the simple probability of an exchange occurring between two adjacent nucleotides. It should be a universal constant for all organisms. $c_n = 10 \times 10^{-3}$ (or 1×10^{-3} for pronounced marker-specific effects) implies that a specific base substitution in association

with the surrounding nucleotides alters the simple exchange probability by a factor of 100 (or 1000). $a_n = 1 \times 10^6$ corresponds to the idea that the nucleotide directly adjacent to the site-specific "break" is 10^6 times more likely to participate in a recombination event than if there were no site-specific "break". w (value not determined) is a measure of the relative distance over which the site-specific "break" exerts its influence; I propose that w corresponds to the amount of displacement synthesis and/or branch migration which occurs during the initiation of recombination. Finally, z (variable) is a measure of the length of the average genetic exchange region, i.e. the average distance between two formal crossover events.

Thus, while the genetic map drawn in figure 7 roughly reflects the number of nucleotides between genetic markers, map contraction, high negative interference, marker-specific recombination effects, and map expansion make simple interpretation of purely genetic data rather difficult.

Nevertheless, several general conclusions emerge out of these studies of the bacteriophage ØX174 genome which are important for my examination of the process of genetic recombination between ØX174 DNA molecules. The entire ØX174 genome is very densely marked (one mutation

per 100 nucleotides on the average) with well characterized genetic markers. The physical distances in nucleotides between many of these genetic markers is accurately (<5%) known. And genetically and physically marked deletion mutants exist which behave as apparently normal virions during ØX174 genetic recombination, yet which can be distinguished in mixed infections by measuring contour lengths obtained by electron microscopy.

(ii) Genetic Recombination of Bacteriophage ØX174

Previous research on the genetic recombination of bacteriophage ØX174 and the closely related bacteriophage S13 is reviewed by Benbow, Zuccarelli, Davis, and Sinsheimer (1973). The primary conclusion to emerge out of these earlier studies is that there are at least two pathways for the formation of bacteriophage ØX174 genetic recombinants. These were termed the primary recombination mechanism and the secondary recombination mechanism (Baker, Doniger, and Tessman, 1971). The chief evidence for two independent mechanisms is that rec^+ (rec A "enzyme" mediated) ØX174 genetic recombination is stimulated by ultraviolet irradiation of the input phage, while recombinant formation in a rec A host is not (Tessman, 1968).

To systematically examine the ØX174 recombination process under a variety of conditions, a series of test crosses were devised (Table 2). The circularity test cross consists of the minimum number of three-factor crosses required to establish the circularity of the ØX174 genome. Not all control crosses are included since the full proof of circularity is given in Benbow, Hutchison, Fabricant, and Sinsheimer (1971).

The cistron product test cross consists of two- and three-factor genetic crosses which, when performed under both permissive and non-permissive conditions, establish

TABLE 2

The Circularity Test Cross	Order Deduced
(1) am N1(H) ts Y(G) x ts 79(G)	N1-(79-Y)
(2) am 88(F) ts 79(G) x ts Y(G)	(79-Y)-88
(3) am 3(E) ts 79(G) x ts Y(G)	(79-Y)-3
(4) am 3(E) ts 79(G) x am 88(F)	79-(88-3)
(5) am 88(F) ts 79(G) x am 3(E)	79-(88-3)
(6) am 3(E) ts 79(G) x am 27(E)	79-(3-27)
(7) am 3(E) ts 9(B) x am 27(E)	(3-27)-9
(8) am 3(E) ts 79(G) x am 42(D)	79-(3-42)
(9) am 3(E) ts 9(B) x am 42(D)	(3-42)-9
(10) am 3(E) ts 9(B) x ts 116(B)	3-(116-9)
(11) am 33(A) ts 116(B) x ts 9(B)	(116-9)-33
(12) am 33(A) ts 116(B) x am 86(A)	116-(33-86)
(13) am 33(A) ts Y(G) x am 86(A)	(33-86)-Y
(14) am33(A) ts Y(G) x ts 79(G)	33-(79-Y)
(15) am 33(A) ts Y(G) x am N1(H)	(33-N1)-Y
(16) am N1(H) ts Y(G) x am 33(A)	(33-N1)-Y

The Cistron Product Test Cross	Cistron
(1) am 86(A) x am 18(A)	A
(2) am 50(A) x am 33(A)	A
(3) am 14(B) x am 16(B)	B
(4) ts 9(B) x ts 116(B)	B
(5) am 10(D) x am H81(D)	D
(6) am 27(E) x am 3(E)	E

The Cistron Product Test Cross	Cistron
(7) *am* 87(F) x *am* 89(F)	F
(8) *op* 6(F) x *am* 88(F)	F
(9) *am* 9(G) x *am* 32(G)	G
(10) *ts* 79(G) x *ts* Y(G)	G
(11) *am* N1(H) x *am* 80(H)	H
(12) *ts* 4(H) x *ts* 28(H)	H

The Genetic Distance Test Cross	Span
(1) *am* 10(D) x *am* 9(G)	2039 to 4137
(2) *am* 3(E) x *am* 9(G)	2485 to 4137
(3) *am* 3(E) x *am* 87(F)	2485 to 3944
(4) *am* 27(E) x *am* N1(H)	2413 to 5235
(5) *am* 87(F) x *am* N1(H)	3944 to 5235
(6) *am* 9(G) x *am* 86(A)	4137 to 24
(7) *am* N1(H) x *am* 33(A)	5235 to 165
(8) *am* N1(H) x *am* 35(A)	5235 to 294
(9) *am* N1(H) x *am* 16(B)	5235 to 374
(10) *am* 86(A) x *am* 16(B)	24 to 374
(11) *am* 86(A) x *am* 35(A)	24 to 294
(12) *am* 86(A) x *am* 10(D)	24 to 2039
(13) *am* 35(A) x *am* 10(D)	294 to 2039
(14) *am* 16(B) x *am* 10(D)	374 to 2039
(15) *am* 16(B) x *am* 3(E)	374 to 2485

The Defined Marker Test Cross

(1) *am* 87(F) x *am* N1(H)

(2) *am* 9(G) x *am* N1(H)

(3) *am* 10(D) x *am* 9(G)

(4) *am* 3(E) x *am* 9(G)

whether a ØX174 cistron product is required for ØX174 genetic recombination. For example, if two amber mutations am A and am B, each of which is linked to the same ts marker (i.e. am A ts C and am B ts C), are crossed at both permissive (30 °C) and non-permissive (40 °C) temperatures, it can be established by spheroplast assay whether the cistron containing the ts C mutation is essential for the formation of ØX174 wt recombinants. Similarly, if two markers in the same cistron (i.e. am A x am B, where A and B are in the same cistron) are crossed in a permissive (Su^+) and a non-permissive (Su^-) host, the requirement of that cistron product for the production of viable wt recombinants can be assessed.

The genetic distance test cross is a series of two- and three-factor genetic crosses which measure map distances in overlapping regions of the genome. The purpose of this test cross is to determine if any set of conditions (different host strains, ultraviolet irradiation, etc.) alters recombination frequencies specifically in one region of the genome (i.e. a selective effect on recombinant formation in one or more particular regions).

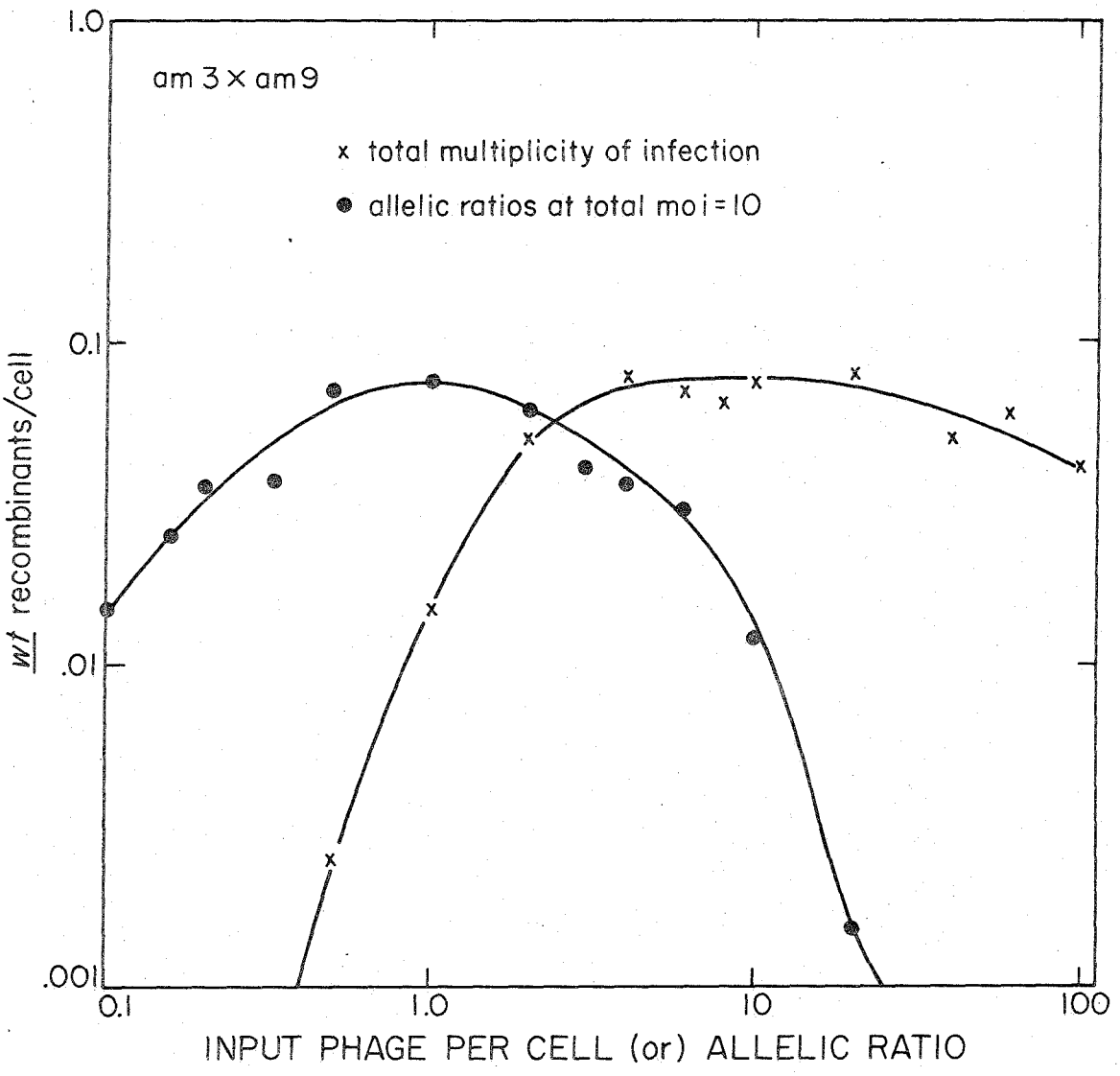
The defined marker test cross is a series of two- and three-factor crosses in which the physical location of the markers on the genome is believed to be known as well

as the number of nucleotides between the markers to an accuracy of better than 5%. In addition, the physiological defects introduced by these mutations are believed to be understood, marker-specific recombination effects are minimal, and all these markers are located in the region of the genome where we have found no recombination anomalies. In short, these crosses are the "safest" crosses and are used to assay the effects of different physiological and genetic conditions of ØX174 genetic recombination. I shall assume that any effect which is demonstrated with the defined marker test cross is typical of genetic recombination between ØX174 DNA molecules in general.

The effects of varying the allelic ratio ot the multiplicity of infection on the recombination frequencies measured with the defined marker test cross are shown in figure 10. The usual multiplicity of infection in our experiments is 5 of each of the two infecting phage. However, the frequency of wt recombinant formation is roughly independent of the total multiplicity of infection in the range 2 to 20 or more. Similarly, the measured recombination frequencies or segregation ratios are independent of the allelic ratio as long as it does not exceed 6:1. From these data I conclude that recombination events probably occur between two randomly selected ØX174

Figure 10. ØX174 wt Recombinant Formation as a Function of Total Multiplicity of Infection or Allelic Ratio

wt recombinants per cell for the two-factor genetic cross am 3(E) x am 9(G) are plotted versus the total multiplicity of infection (am 3(E) + am 9(G)) or against allelic ratio (am 3(E)/am 9(G)). Other markers in the defined marker test cross yield similar results. The two-factor crosses were carried out as described in Methods and Materials except that the infecting phage concentrations were varied as indicated. Total burst sizes were unaffected by different allelic ratios, but were decreased at multiplicities of infection above 40 or below 1.0.


1.0
0.1
.01
.001
am 3 × am 9
x total multiplicity of infection
● allelic ratios at total moi = 10
wt recombinants/cell
0.1
1.0
10
100
INPUT PHAGE PER CELL (or) ALLELIC RATIO

genomes and that few (or no) subsequent or multiple matings occur. As long as at least one of each type of genome is inside the cell, and as long as there is a reasonable probability of their interaction (i.e. general genetic pairing), then recombination can occur. It should be noted that this differs from the multiple mating events observed in bacteriophage T4 recombination (Hayes, 1968).

The effect of host genes on bacteriophage ØX174 genetic recombination is summarized in table 3 based on data from a defined marker test cross in each of the indicated host strains. The host rec A mutation reduces ØX174 genetic recombination by roughly 8-fold (2 to 100 fold depending on the markers used). This effect led Tessman (1966) to define rec A "enzyme" mediated recombination as the primary recombination mechanism. It should be noted that rec A reduces host recombination frequencies by a factor of 1000 (Clark and Margulies, 1965). The host rec B and rec C mutations alone have no detectable effect on ØX174 recombination frequencies; in double mutants with rec A they effect a further 5-10 fold reduction in ØX174 genetic recombination.

Other host markers tested had little or no effect on ØX174 recombination frequencies: uvr A, uvr B, uvr C, pol A, ts 7 (ligase), and permutations of these mutations did not

TABLE 3

am 10(D) x *am* 9(G)

Host	Genotype	wt rec./cell x 10^2
HF4714	*wt*	6.0
HF4712	*rec* A	0.7
RMB 110	*rec* B *uvr* A	6.5
HF4704	*uvr* A	8.5
ts 7	*ligase*$^-$	4.5
1000	*pol* A	5.0
NH4547	*rec* A *rec* B *uvr* A	0.09 *
Su 2_{am}	*wt*	3.5
Su 2_{och}	*wt*	4.0
H502	*uvr* A *endo* I$^-$	3.0
HF4705	*uvr* B	7.5
HF4706	*uvr* C	7.0

* estimated from spheroplast assay.

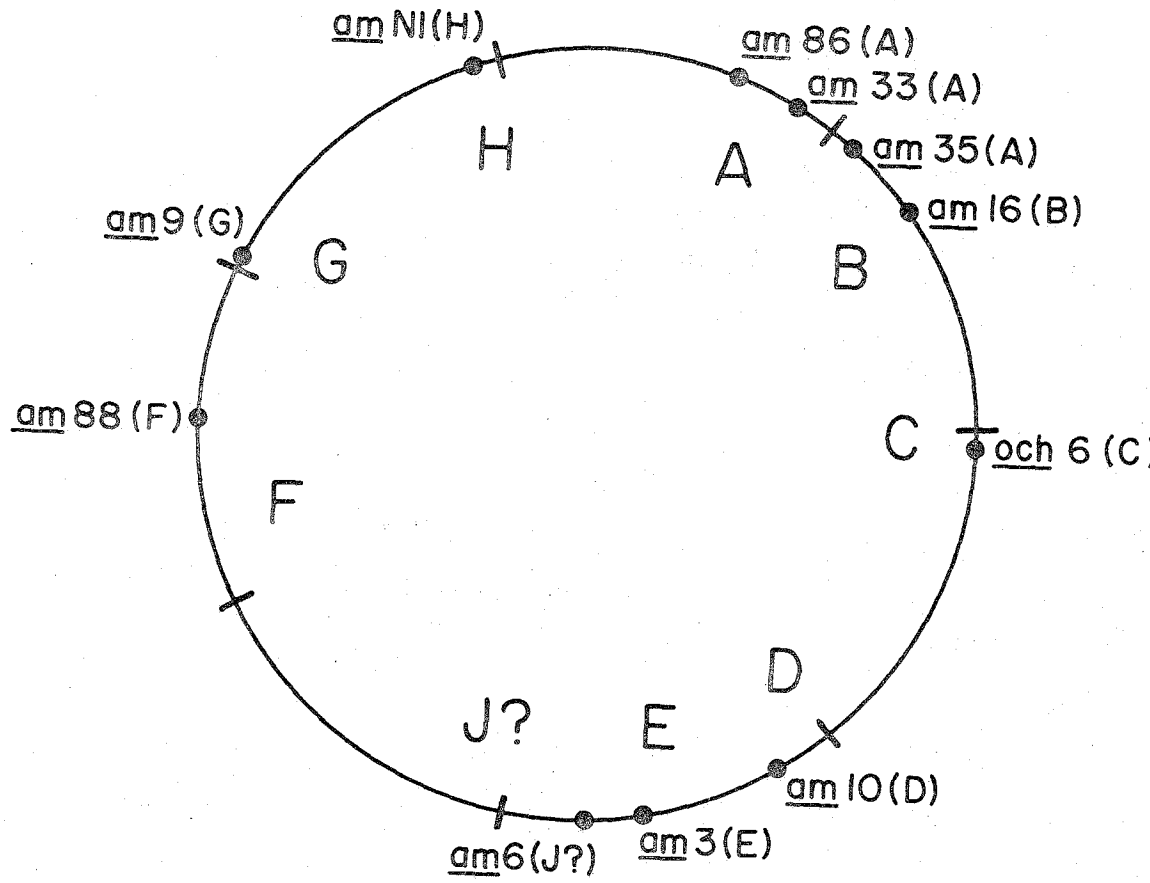
appreciably decrease or increase ØX174 genetic recombination frequencies. The sbc A and sbc B genes isolated by Clark's group were shown (personal communication) to have little detectable effect on ØX174 genetic recombination. In summary, only the host rec A "enzyme" appears to play a major role in the primary mechanism of ØX174 genetic recombination.

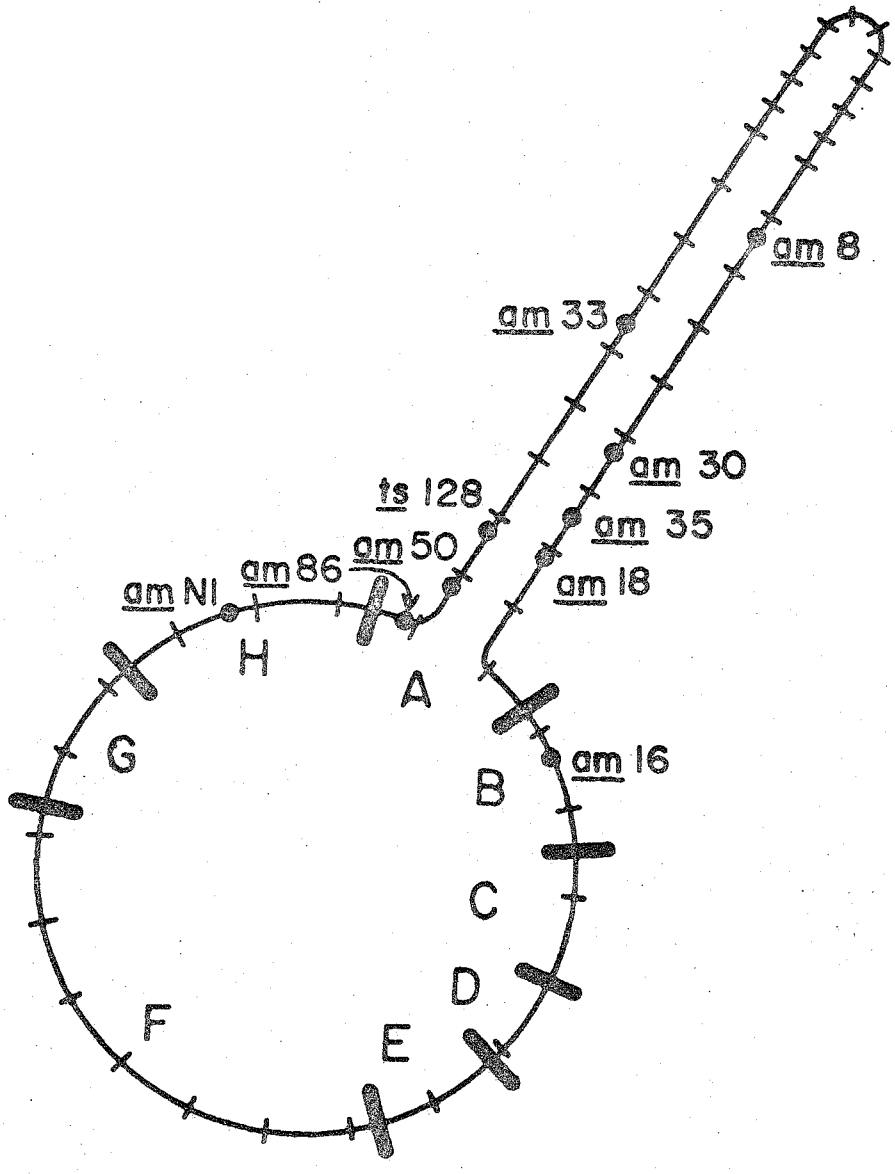
The circularity test cross and the genetic distance test cross were carried out in HF4712, a rec A host presumed to be isogenic with the standard rec^+ host, HF4714. The resulting genetic map is presented in figure 11a (Benbow, Zuccarelli, Davis, and Sinsheimer, 1973). The rec A genetic map is circular and apparently represents a minimap which resembles the map in a rec^+ host. One striking feature, however, is the lack in the rec A map of a high recombination region within cistron A.

Benbow, Hutchison, Fabricant, and Sinsheimer (1971) have shown that intracistronic recombination frequencies within cistron A are abnormally high in a rec^+ host as shown in figure 11b. However, cistron A recombination frequencies in a rec A host are consistent with the rest of the genome. An explanation of this recombination anomaly is presented below based on the hypothesis that there is a strand-specific single strand "break" in cistron A which stimulates genetic recombination mediated

Figure 11. Genetic Map of Bacteriophage ØX174 in a rec A Host

The frequency of wt recombinants per total progeny phage in two-factor crosses carried out in HF4712 rec A are represented schematically. The two-factor crosses were carried out as described in Methods and Materials. For reference purposes a similar genetic map constructed in the rec^+ host HF4714 is also shown; this map differs from figure 7 in that it is not corrected for marker- and site-specific recombination effects. Note that the cistron A region in the rec^+ host exhibits very high recombination frequencies: to draw the map it is necessary to assume that cistrons B and H are closely linked, i.e. that recombination events within the cistron A "loop" usually occur in pairs so that the cistron A region is inert genetically with respect to the rest of the genome.


am NI(H)
am 86 (A)
am 33 (A)
am 35(A)
am 16 (B)
am 9 (G)
am 88 (F)
och 6 (C)
am 10 (D)
am 3 (E)
am 6 (J?)
H
A
B
G
C
F
D
J?
E


am 8
am 33
am 30
ts 128
am 35
am 50
am 86
am N1
am 18
H
A
G
am 16
B
C
D
F
E

by the rec A "enzyme".

The cistron product test cross was carried out in a rec^+ and in a rec A host. As shown in table 4, no bacteriophage ØX174 cistrons are required for wt recombinant formation to occur by the primary mechanism in the rec^+ host. Thus, the number of wt recombinants generated under conditions where a cistron product is lacking are equal to (or greater than) the number of wt recombinants formed when the cistron product is present.

In the case of the rec A host, however, cistron A product is required. Note that in the rec A host the secondary mechanism of genetic recombination is responsible for recombinant formation.

At least two possible mechanisms for the role of cistron A product in the secondary mechanism might be envisioned: cistron A product might be involved directly in the recombination process per se; or something that cistron A does might alter indirectly the number of wt recombinants formed. Baker, Doniger, and Tessman (1971) have suggested the latter mechanism is involved in bacteriophage S13 recombination. They postulate that, since cistron A mutants are blocked in the formation of progeny replicative form DNA molecules, the secondary mechanism primarily acts on the progeny RF molecules--- as well perhaps as the parental ones. The primary

TABLE 4

Cross (Type 1)	Cistron	rec$^+$Su$^+$	rec$^+$Su$^-$	rec ASu$^+$	recASu$^-$
am 18(A)xam 86(A)	A	7.6	18.8	0.08	0.003
am 33(A)xam 50(A)	A	4.1	11.3	0.05	0.006
am 14(B)xam 16(B)	B	1.8	1.5	0.21	0.33
am 10(D)xamH81(D)	D	0.9	1.3	0.10	0.15
am 3(E)xam 27(E)	E	0.6	0.2	0.07	0.02
am 88(F)xop 6(F)	F	13.5	16.8	1.75	2.05
am 9(G)xam 32(G)	G	1.3	0.9	0.30	0.21
amN1(H)xam 80(H)	H	0.4	0.5	0.07	0.10

Cross (Type 2)	30°C	40°C
am 3(E)ts 128(A) x am 9(G)ts 128(A)	6.3	15.2

Cross (Type 3)	rep 3Su$^+$ *	rep 3Su$^-$
am 33(A) x am 50(A)	12.5	27.5
am 14(B) x am 16(B)	2.3	1.6
am N1(H) x am 80(H)	0.3	0.8

(1) wt recombinants per cell x 10^2

(3) wt recombinants/total progeny yielded by spheroplast assay x 10^4

(2) am$^+$ ts 128(A) recombinants assayed by spheroplast assay x 10^4

* rep 3 Su$^+$ is C1704 lysogenic for Ø80 $p_{Su\ III}{}^+$

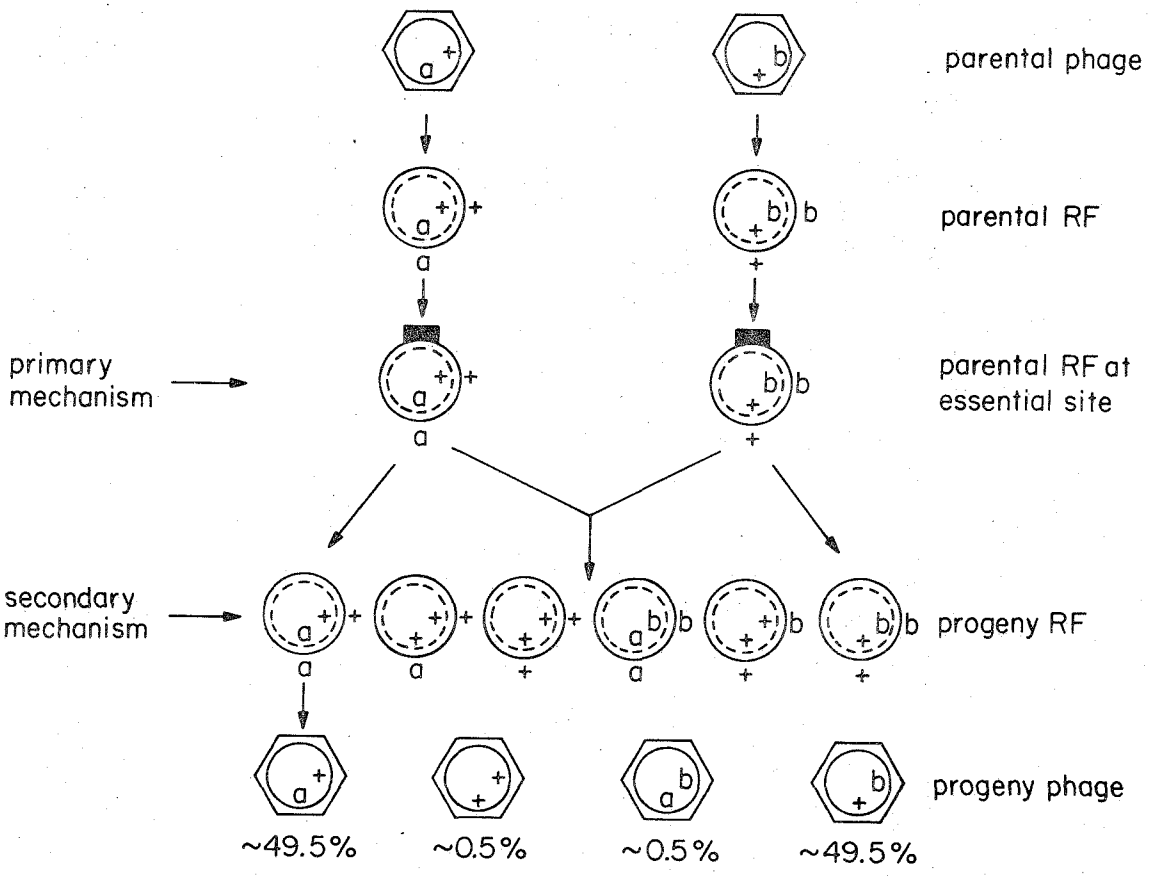
mechanism acts only on parental replicative form DNA molecules. To clarify this proposal, a scheme for bacteriophage ØX174 replication during the growth conditions used in these experiments is drawn in figure 12.

To test Baker, Doniger, and Tessman's (1971) hypothesis, which presumably determines the stage of ØX174 DNA replication during which the primary mechanism of genetic recombination is active, a cistron product test cross was carried out in a rep 3 host. As shown in table 4, the number of wt recombinants is no lower for cistron A mutants in the rep 3 host than for mutants in cistrons B or H. It is thus likely that their interpretation is correct, i.e. that the cistron A product per se is not involved directly in the production of wt recombinants by the secondary mechanism. Note, however, that this is not a clean-cut demonstration since it is necessary to assay extracted RF DNA for the wt recombinants rather than determining the burst of wt progeny phage. This is because a rep 3 host (Denhardt, Dressler, and Hathaway, 1967) is unable to replicate bacteriophage ØX174 DNA beyond the parental replicative form stage.

The time of formation of wt recombinants in the infection cycle of bacteriophage ØX174 was examined by two separate procedures. In the first, single bursts of wt recombinants were plated on non-permissive lawns (which

Figure 12. DNA Replication Stages of Bacteriophage ØX174

The stages of DNA replication of bacteriophage ØX174 are redrawn from Sinsheimer (1968). The stages of parental RF, parental RF at an essential site, progeny RF, and single stranded progeny phage are confirmed by the data and electron micrographs in this thesis. The proposed primary and secondary mechanisms of genetic recombination are believed to act at the indicated stages. Note that phage heterozygotes are impossible although heterozygous RF can exist. The two RF heterozygotes indicated in the progeny RF stage as results of the recombination event (+/a,+/+ and +/+,+/b) in reality may only exist as parental RF.

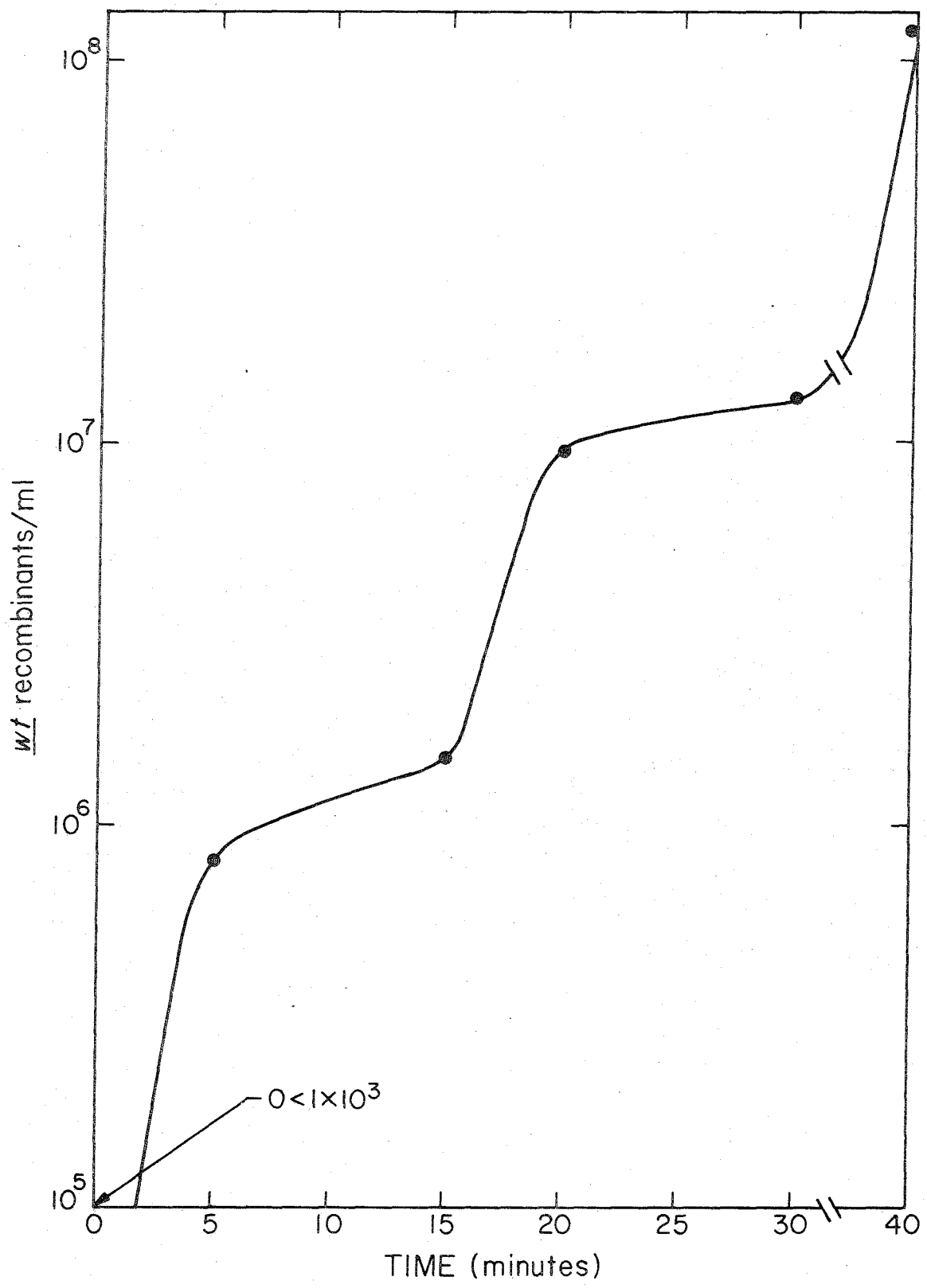

parental phage
parental RF
primary mechanism
parental RF at essential site
secondary mechanism
progeny RF
progeny phage
~49.5%
~0.5%
~0.5%
~49.5%

do not plate total progeny phage from a cross): most single bursts of recombinants are large as shown in table 5. In contrast, single bursts of revertants are small (Denhardt and Silver, 1966). In terms of the replication scheme presented in figure 12, this implies that most wt recombinant formation occurs early in infection, certainly before single strand DNA synthesis begins. Recall that cistron A mutants in non-permissive cells form normal numbers of wt recombinants, even though they only replicate the infecting parental DNA strand and do not make progeny replicative form DNA. These two observations imply that the primary mechanism of recombination takes place very shortly after infection, and that most genetic recombination does not involve progeny replicative form DNA molecules.

Direct confirmation for this hypothesis was obtained by measuring wt recombinant production in extracted RF DNA molecules as a function of time. One experiment of this type is shown in figure 13. Although other interpretations are possible, the simplest explanation for the observed wt recombinant formation is that wt recombinants are formed by five minutes after the initiation of DNA replication. These wt recombinants act as parental RF, turning out progeny RF---which subsequently produce single stranded recombinant virions. Note that the

Figure 13. The Time Course of Appearance of wt Recombinants During Bacteriophage ØX174 Infection

The production of wt recombinants in extracted RF DNA molecules was assayed by the spheroplast assay procedure described in the legend to figure 13. The experimental design is as follows: HF4704 cells growing under the conditions described in the legend to figure 18 were synchronized with cyanide. At t = 0, equal numbers of phage (am 9(G) and am 10(D)) were added at a total multiplicity of infection of 10. RF DNA was extracted at the indicated times as described in the legend to figure 18. At t = 15 minutes cyanide was removed and ØX174 DNA synthesis commenced. By t = 20 and 30 minutes (5 and 15 minutes after the start of DNA synthesis) wt recombinant formation was apparently nearly complete. At t = 40 minutes (25 minutes after the start of DNA synthesis) progeny RF and recombinant progeny RF are being generated.


10^8
10^7
10^6
10^5
wt recombinants/ml
$0<1\times10^3$
0
5
10
15
20
25
30
40
TIME (minutes)

initial events in wt recombinant formation probably occur during the adsorption period while DNA replication is blocked by cyanide.

In the cross am A x am B, what genotypes survive the recombination event to produce progeny phage particles? Are the ØX174 DNA molecules surviving a recombination event reciprocally recombinant? The reciprocality of recombinant formation can be examined on several levels. On a population level, genetic recombination in bacteriophage ØX174 appears to be reciprocal. This conclusion is based on the observation that double mutants typically are observed at roughly the frequency at which wt recombinants are generated in the same cross---within a factor of two or three. For example, the double mutant am N1(H)ts Y was selected by screening roughly 10,000 single plaques from the cross am N1(H) x ts Y (G). Two double mutants were isolated; the frequency (2×10^{-4}) is comparable to the observed recombination frequency (3.1×10^{-4} wt recombinants per total progeny phage). This has not been tested rigorously in view of the tedium involved in testing several thousand plaques in order to confirm a generally accepted point.

On the level of a single cell, however, evidence from bacteriophages T4 (Hershey and Rotman, 1947) and f1 (Boon and Zinder, 1971) suggests that individual recombination

events are non-reciprocal. Technically it is difficult to examine single bursts of total progeny phage in crosses where recombination has occurred since wt recombinants are generated in only one out of every one hundred cells. Therefore the wt recombinants and total progeny produced in non-permissive cells were examined: only cells which contain a wt recombinant will lyse to release phage if the two mutants are in the same cistron. The data from nine observed single bursts are given in table 5.

A single burst from a cell in which ØX174 genetic recombination has taken place yields one parent and one recombinant---in our case only the wt recombinant. Thus, genetic recombination between ØX174 DNA molecules is asymmetric and non-reciprocal. If an am 88(F)op 6(F) double mutant were generated it would form plaques on the double suppressor strain used to examine the total yield of progeny phage; therefore, no am-op double mutants were generated in the nine single bursts in which wt recombinants were generated. A similar conclusion was reached for the filamentous bacteriophage f1 by Boon and Zinder (1971), although we were unaware of their results at the time of our experiments.

One of the chief problems faced by earlier investigators of ØX174 genetic recombination (Hutchison, personal communication) was the lack of a suitable spheroplast

TABLE 5

SINGLE BURST EXPERIMENT

Cross	Host	*wt*/burst
*am*88(F) × *op*6(F)	Pm^+	18.5
*am*88(F) × *am*88(F)	Pm^+	1.3
*op*6(F) × *op*6(F)	Pm^+	0.9
*am*88(F) × *op*6(F)	Pm^-	26.6

	am	op	+	am - op
(1)	47	3	51	0
(2)	62	5	19	0
(3)	2	97	31	0
(4)	0	35	31	0
(5)	34	1	39	0
(6)	87	2	4	0
(7)	67	3	31	0
(8)	1	94	13	0
(9)	0	85	21	0

The nine single bursts examined took place in non-permissive (Su^-) host cells; thus, only those crosses which yielded some *wt* recombinant phage were observed.

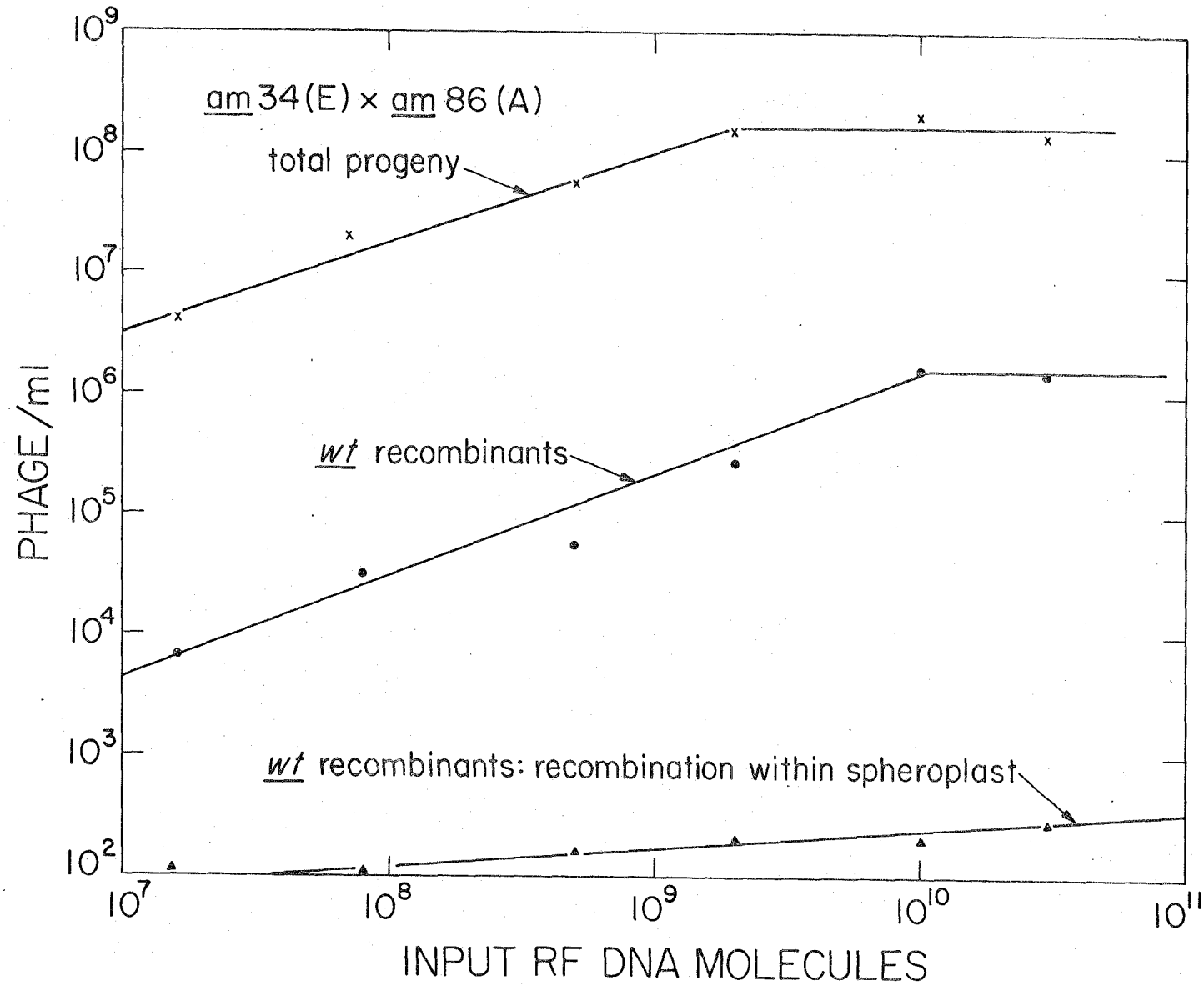
assay for wt recombinants in replicative form DNA perparations. The existing spheroplast assays for isolated RF DNA molecules were notoriously inefficient; and evidence was obtained that genetic recombination occurred within the spheroplast during the assay. This obviously complicated earlier results. Therefore, a series of E. coli K12 recombination-deficient mutants (a gift of Dr. Paul Howard-Flanders) were tested in order to develop a recombination free spheroplast assay.

NH4557, a rec A rec B uvr A derivative was selected because it gave the lowest background of recombination within the spheroplast coupled with the highest yields of total progeny phage. It is apparently necessary to have both the rec A and the rec B mutation in the same strain. The rec A mutation is essential for reduction of wt recombinant formation. However, the rec B mutation is also essential in order to obtain a good yield of total progeny phage. Presumably the rec B endonuclease (Oishi, 1969) otherwise degrades the input RF DNA. The spheroplast assay procedure is that of Guthrie and Sinsheimer (1962) except that NH4557 is substituted for K12W6.

A typical spheroplast assay for wt recombinants is given in figure 14. The background of recombination within the spheroplast assay is of the order of magnitude of 1×10^{-5} or less. In experiments involving the

Figure 14. A Recombination Free Spheroplast Assay for Bacteriophage ØX174 RF DNA Molecules

Replicative form DNA molecules were prepared from HF4704 either singly or mixedly infected with am 34(E) and am 86(A) by the procedure described in the legend to figure 16. Spheroplast assays were carried out as described by Guthrie and Sinsheimer (1962). Total bursts from each assay tube were plated on HF4714 (a permissive strain to assay for progeny) and on C (a non-permissive strain to assay for wt recombinants). Note that the wt phage produced by a mixture of RF from singly infected cells (▲) are several orders of magnitude fewer than the wt recombinants found in RF from mixedly infected cells (●). Total progeny phage yields from the mixture of RF from singly infected cells (not shown) were identical to the yields from the mixedly infected RF preparations (X).


am 34 (E) × am 86 (A)
total progeny
wt recombinants
wt recombinants: recombination within spheroplast
PHAGE/ml
INPUT RF DNA MOLECULES
10^9
10^8
10^7
10^6
10^5
10^4
10^3
10^2
10^7
10^8
10^9
10^10
10^11

assay of double stranded RF DNA molecules, protamine sulfate was sometimes added to increase the efficiency of infection (Benzinger, Kleber, and Huskey, 1971). Protamine sulfate does not alter the efficiency of single strand DNA infections; also, spheroplast assays with and without protamine sulfate give similar results except that the total yields are higher in the presence of protamine sulfate. I conclude that a recombination-free spheroplast assay using the strain NH4557 can be used to examine the formation of wt recombinants at any level above 1×10^{-5} wt recombinants per total progeny phage.

Other physiological effects on ØX174 genetic recombination have been examined (Benbow, Zuccarelli, Davis, and Sinsheimer, 1973). These include the effects of temperature, media, starvation synchronized infection, and time in cyanide on the observed recombination frequency. None of these conditions has any significant effect on the formation of wt recombinants. This suggests that the conditions I use to examine ØX174 genetic recombination closely parallel "real life" growth conditions except that the recombination frequencies are elevated because cyanide is used to synchronize infection. In unsynchronized infection, superinfection exclusion (Hutchison, 1968) lowers recombination frequencies to near the level of spontaneous reversion frequencies.

The experiments presented in this section and the preceding one are obvious applications of previously used genetic techniques to bacteriophage ØX174. This thesis represents one of the few cases where all of these techniques have been applied to one organism in one laboratory. At this point (both in time and in this paper) bacteriophage ØX174 is the most thoroughly genetically characterized organism in the world per unit mass (tongue in cheek).

In the following sections two biophysical studies and a genetic probe into the process of genetic recombination of ØX174 DNA molecules are described.

(iii) The Role of Single Strand "Breaks" in Genetic Recombination between ØX174 DNA Molecules

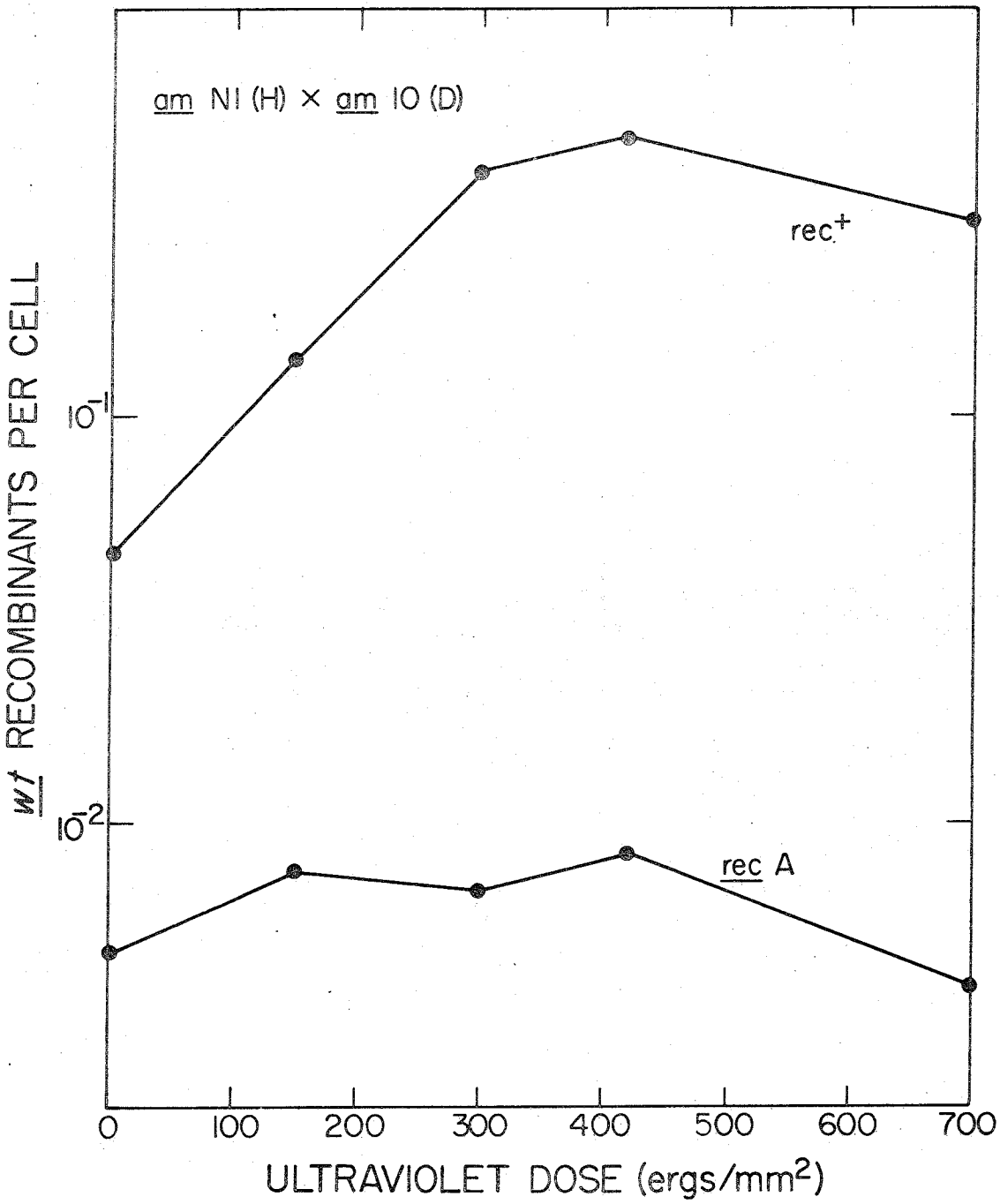
Genetic recombination of bacteriophage S13 (a very close relative of ØX174) was initially demonstrated after ultraviolet irradiation of the parental phage prior to infection (Tessman and Tessman, 1961); in general, it is well established that ultraviolet irradiation stimulates genetic recombination (Hayes, 1968). Furthermore, Tessman (1968) has shown this stimulatory effect for bacteriophage S13 genetic recombination is greater in rec^+ cells than in rec A cells. In order to establish the basis for this stimulatory effect, the structure of ultraviolet damaged ØX174 replicative form DNA molecules has been examined. The structures observed suggest a possible recombination intermediate in the primary mechanism of genetic recombination for bacteriophage ØX174 DNA molecules.

Ultraviolet irradiation of ØX174 parental phage prior to infection stimulates ØX174 recombinant formation: to show this the defined marker test cross was carried out in a rec^+ and in a rec A host following exposure of the parental phage to ultraviolet irradiation. As can be seen in figure 15 for the two-factor genetic cross am N1(H) x am 10(D), the number of wt recombinants per rec^+ cell increases about 15 fold with increasing dose up to about three "hits" per parental phage, then decreases slowly.

Figure 15. Ultraviolet Stimulation of Bacteriophage ØX174 Genetic Recombination

ØX174 am N1(H) and am 10(D) were ultraviolet irradiated and crossed as follows: a mixture of 50 microliters of each phage in 0.05M borate, pH 8.0, at 2 x 10^{10} phage per milliliter was ultraviolet irradiated at a dose rate of 20 ergs/mm^2/sec in a 10 cm plastic petri dish. 10 microliter aliquots were removed at appropriate intervals and diluted into 1.0 ml KC-0.003M KCN. Two-factor genetic crosses were carried out as described in Methods and Materials using the ultraviolet irradiated parental phage.

wt recombinants per cell (plated on E. coli C) are plotted against the incident ultraviolet dose (in ergs/mm^2). The one "hit" (1/e) dose is 140 ergs/mm^2. Total yields of progeny phage (plated on HF4714) were constant at roughly 100 phage per cell except at 700 ergs/mm^2 where a slight (20%) decrease was observed. The total multiplicity of infection (based on plaque forming units of unirradiated phage) is 10.


am N1 (H) × am 10 (D)
rec+
rec A
wt RECOMBINANTS PER CELL
10⁻¹
10⁻²
0
100
200
300
400
500
600
700
ULTRAVIOLET DOSE (ergs/mm²)

For other markers the stimulation ranges from 10 to 50 fold. In the rec A strain, however, no increase in the number of wt recombinants is observed at any dose for any pair of markers.

Benbow and Sinsheimer (1973) establish that the stimulatory effect in the rec^+ cell primarily is due to an ultraviolet lesion in one of the two parental replicative form DNA molecules: it is not necessary that both parental genotypes be irradiated to obtain nearly maximal stimulation. This implies that one RF structure probably can "attack" another. Thus, it is less probable that both RF molecules initially open and unwind in complementary regions for pairing and exchange as suggested by Whitehouse (1963) and by Holliday (1964).

The structure of ultraviolet damaged ØX174 replicative form DNA molecules was examined under conditions where parental RF are unable to replicate. Sucrose velocity sedimentation patterns of extracted DNA molecules demonstrate a change from a predominance of covalently closed RF I to a predominance of "nicked" RF II molecules with increasing ultraviolet dose to the parental DNA (Benbow and Sinsheimer, 1973). After alkaline sucrose velocity sedimentation of the RF II, the ultraviolet-irradiated parental (viral) plus strands remain circular; linear DNA molecules primarily are (complementary) minus strands and

complementary strand fragments. Furthermore, the size distribution of the complementary strand fragments after alkaline sucrose velocity sedimentation suggests that the ultraviolet lesions serve as a block to DNA replication. The ultraviolet damage does not simply induce "nicks" or "gaps".

Direct confirmation of this hypothesis was obtained by application of aqueous and formamide spreading techniques for the visualization of single and double stranded DNA in the electron microscope (Davis, Simon, and Davidson, 1971). The sizes of complementary strand fragments are easily measured from electron micrographs of ultraviolet damaged replicative form DNA molecules. Figure 16 shows several examples of ultraviolet damaged RF DNA molecules. These structures suggest that either single strand "breaks" or long single stranded (unpaired DNA) regions are responsible for the observed increase in ØX174 genetic recombination induced by ultraviolet irradiation.

To distinguish between these alternatives, another method of introducing single strand "breaks" into ØX174 replicative form DNA molecules was employed. Thymine starvation of a thymine-requiring host stimulates ØX174 genetic recombination in a rec^+ but not in a <u>rec</u> A host in the same manner as ultraviolet irradiation of parental DNA molecules (Benbow and Sinsheimer, 1973).

Figure 16. Electron Micrographs of Ultraviolet Damaged and Thymine Starvation Damaged ØX174 Replicative Form DNA Molecules

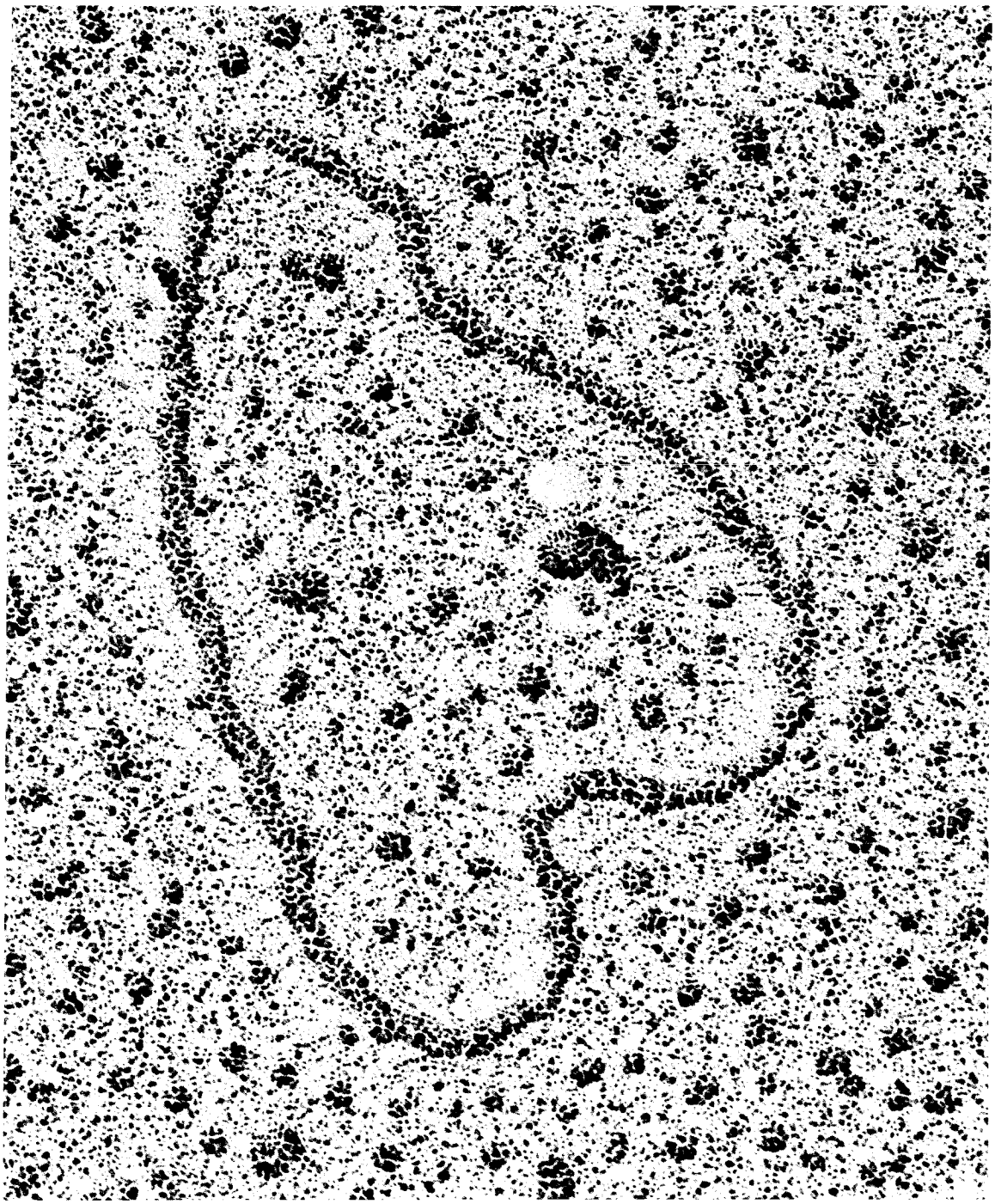
ØX174 RF DNA molecules were purified as described below and prepared for electron microscopy by the procedure of Davis, Simon, and Davidson,(1971). Typical RF DNA molecules are shown resulting from a) unirradiated parental phage; b) parental phage ultraviolet irradiated to 140 ergs/mm^2 (1 "hit" per molecule); c) parental phage ultraviolet irradiated to 280 ergs/mm^2; d) parental phage ultraviolet irradiated to 420 ergs/mm^2; e) parental RF damaged by 15 minutes thymine starvation of the thymine requiring host C1704 (Calendar, Lindquist, Sironi, and Clark, 1970) in the presence of an energy source and amino acids.

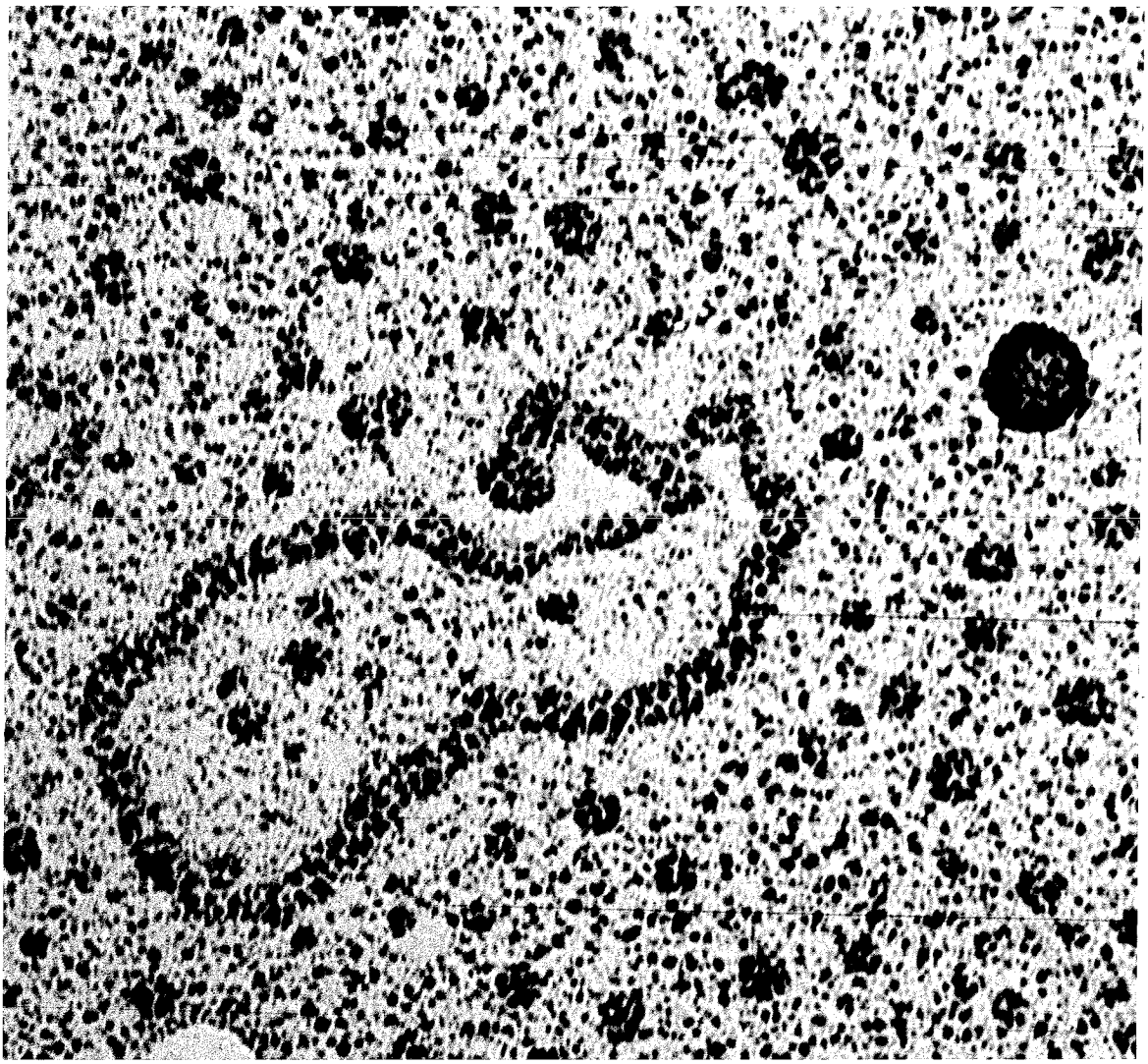
C1704, a rep 3 thy derivative of HF4704 (Lindquist and Sinsheimer, 1967) was grown to 4 x 10^8 cells/ml in TPG-AA plus 2 micrograms/ml thymine. Ultraviolet irradiated ØX174 am 3(E) was added at a multiplicity of infection of 20 along with 2 microcuries/ml 14C thymidine. Input phage had previously been label ed with 3H thymidine. After 20 minutes at 37 ^{o}C ØX174 replicative form was extracted as described in Komano and Sinsheimer (1968).

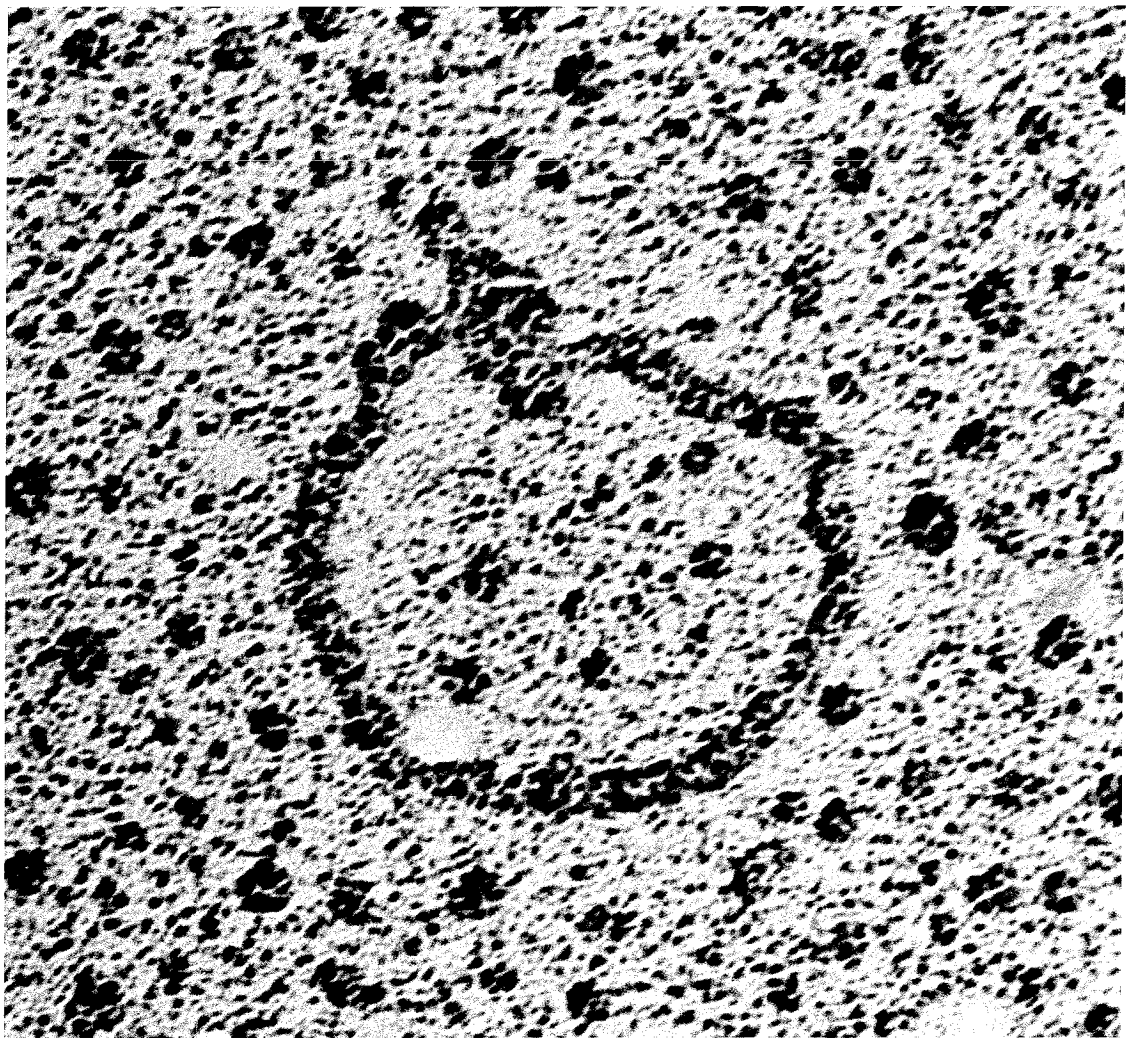
Velocity sedimentation in the SW 41 rotor through 5-20% sucrose (1.0M NaCl, 0.05M Tris, pH 7.5, 0.005M EDTA) for 8 hours generally resulted in clean separation of

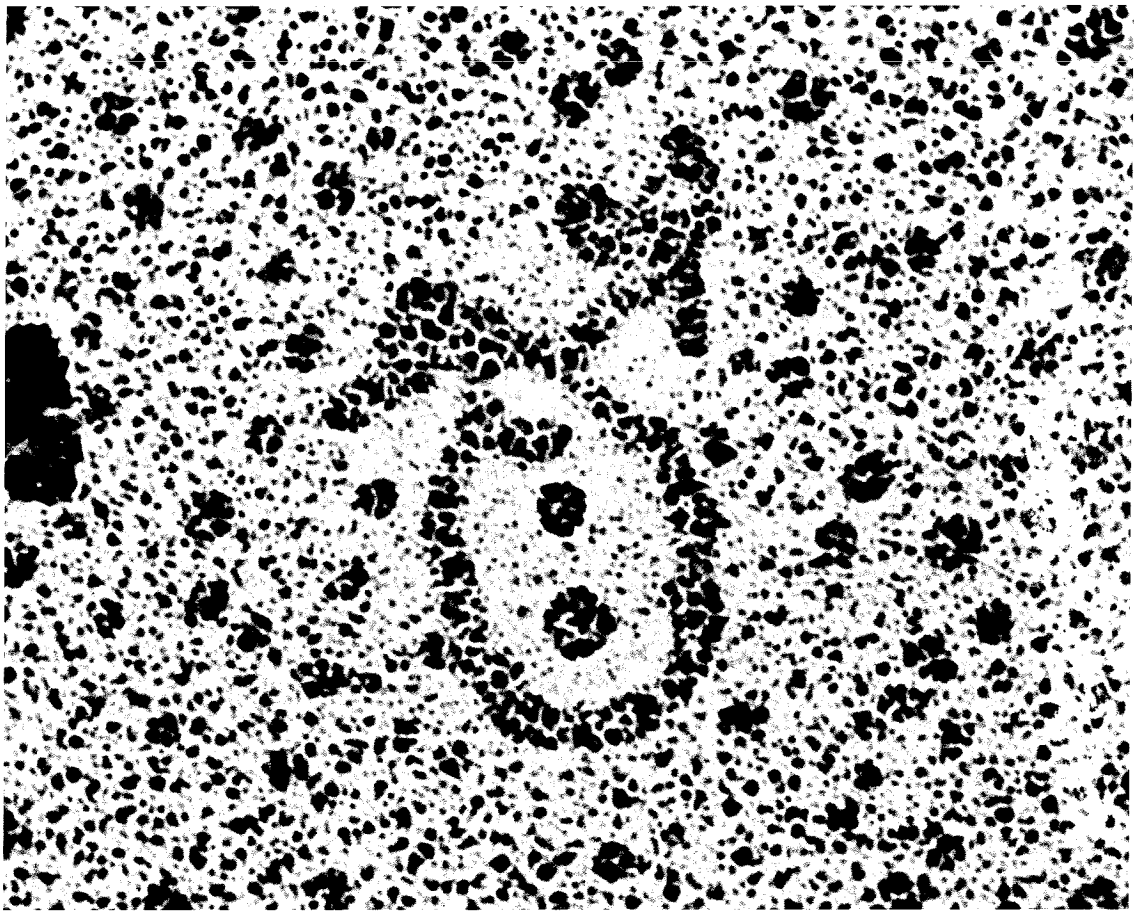
labeled RF I and RF II DNA molecules. The region from roughly 10S to 40S (which presumably contains all the RF DNA molecules) was pooled, dialyzed against 0.05M Tris, pH 8.0, and examined in the electron microscope.

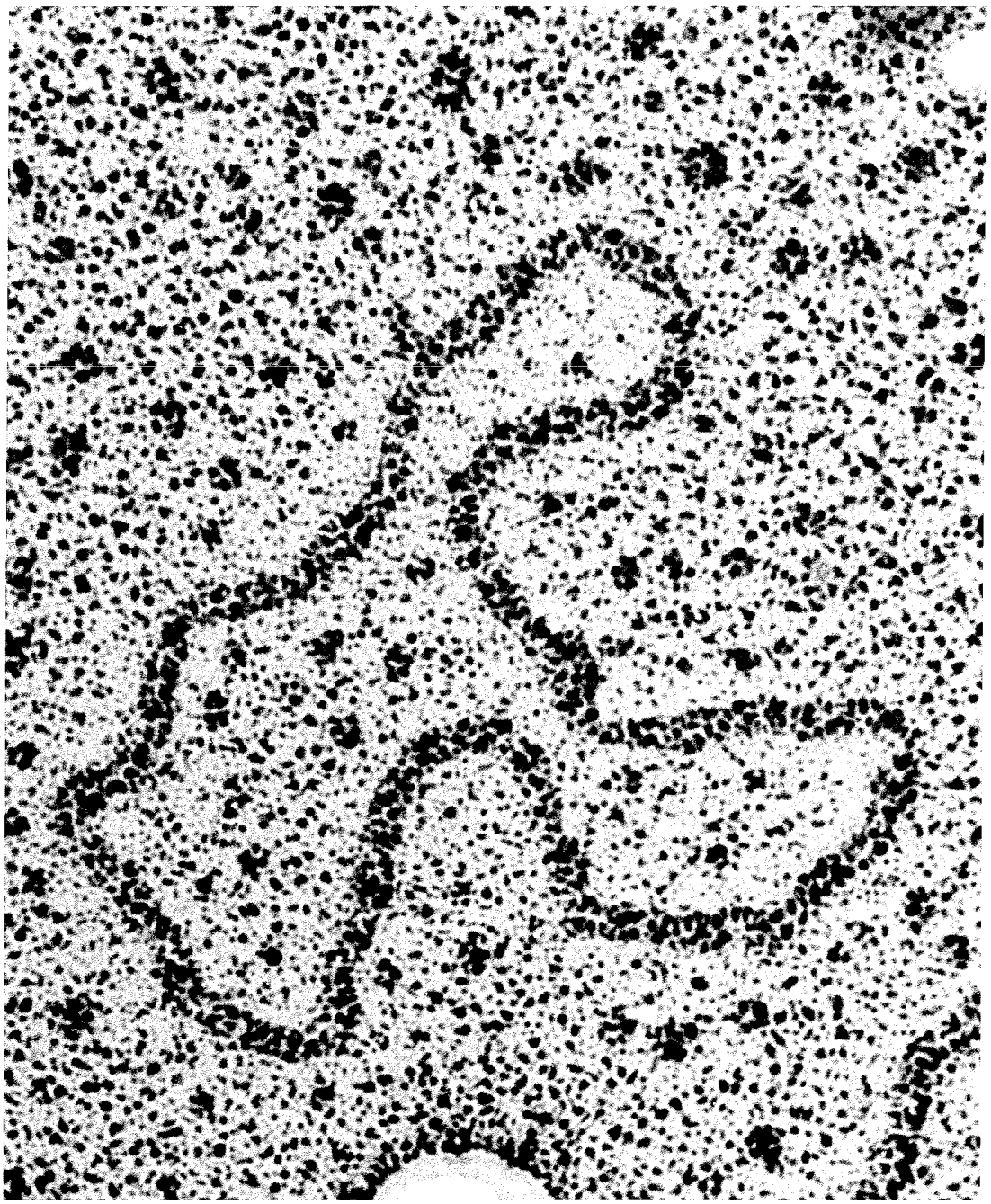
The thymine starvation experiment was carried out analogously except that the infected cells (after 10 minutes for adsorption) were centrifuged 7000 rpm for 20 minutes and resuspended in TPG-AA without thymine. Note that the rep 3 mutation has been shown by Denhardt, Dressler, and Hathaway (1967) to block ØX174 DNA replication beyond the parental RF stage.











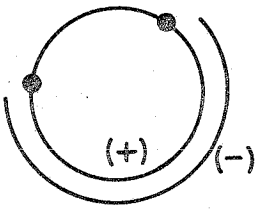
The structure of these DNA molecules was examined by neutral and alkaline sucrose velocity sedimentation. The results obtained indicate that single strand "breaks" are introduced into both complementary (minus) and viral (plus) DNA strands (Benbow and Sinsheimer, 1973; see also Francke and Ray, 1971a). However, these "breaks" do not dramatically decrease the size distribution of the linear DNA molecules suggesting that extensive degradation has not occurred. Thus, as shown in the electron micrographs of figure 16, these thymine-starvation-damaged DNA molecules appear as nearly normal length RF II DNA molecules after being prepared by an aqueous spreading procedure (Davis, Simon, and Davidson, 1971).

The structures for ultraviolet-damaged and thymine-starvation-damaged ØX174 RF DNA molecules are pictured schematically in figure 17. The common feature appears to be the introduction of single strand "breaks"---though whether it is the free DNA strand ends adjacent to the single strand region or some other feature is unclear. The term single strand "break" in this thesis refers to a discontinuity of unspecified length in one of the two strands of a ØX174 RF DNA molecule. Obviously, the rec A "enzyme", which is of major importance in the primary mechanism of recombination of ØX174 (and of its host, E. coli.) thrives upon these single strand "breaks".

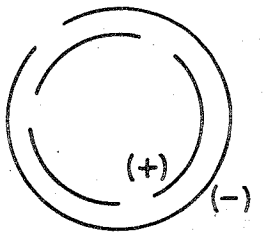
Figure 17. Recombination Proficient ØX174 Replicative Form DNA Molecules

The proposed structures for ultraviolet-damaged and thymine-starvation-damaged ØX174 replicative form DNA molecules are derived from data given in the text and in figure 16. A hypothetical RF molecule containing a strand-specific "break" in the cistron A region of the genome is also shown. It is proposed that DNA molecules with random single strand "breaks" (which would look similar to one of the structures drawn) are normal intermediates in the primary mechanism of bacteriophage ØX174 genetic recombination.

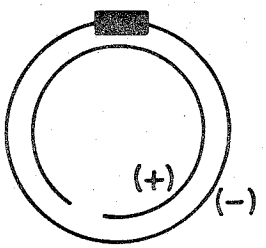
RECOMBINATION PROFICIENT STRUCTURES



Parental RF
containing pyrimidine
dimers in viral strand
(ultraviolet irradiation)



Parental RF containing gaps
in either strand
(thymine starvation)



Parental RF containing strand specific
"break" (or other structure) in
cistron A region
(cistron A endonuclease)

This preference of the rec A "enzyme" for single strand "breaks" suggests an explanation for the anomalously high recombination region observed within cistron A in a rec^+ host. As shown in figure 9, cistron A recombination frequencies typically are very high in the rec^+ host. Recombination frequencies across cistron A not involving markers within the high recombination region of cistron A are relatively normal (figure 8), but recombination frequencies between outside markers and markers within the high recombination region are also elevated (figure 9). In a rec A host, however, no evidence is seen for a high recombination region in cistron A; this is apparent from the genetic map in a rec A host shown in figure 11.

One obvious explanation of this anomaly is that the cistron A region of the genome contains a site-specific "break" which stimulates recombination in this region in a rec^+ but not in a rec A cell. In this context, it is of interest that Baas and Jansz (1972) have located the origin of ØX174 RF DNA replication within cistron A at the center of the high recombination region as drawn in figure 11 (Benbow, Hutchison, Fabricant, and Sinsheimer, 1971). Furthermore, Francke and Ray (1971b, 1972) have shown that cistron A product does in fact introduce a strand-specific break into ØX174 parental replicative form DNA molecules although the location of the "break" is not

yet mapped.

In order to explain the high recombination also observed between cistron A mutants in the absence of cistron A protein (which is presumed to make the specific break), it is necessary to introduce the ad hoc hypothesis that lack of cistron A forces ØX174 DNA molecules to remain as parental RF and thus "at risk" with regard to recombination for a longer time. This idea is (weakly) supported by the observation (table 4) that the recombination frequency observed between outside markers when each genome lacks cistron A function is 2-3 times the normal recombination frequency observed for these same markers in the presence of cistron A function.

The primary conclusion to be drawn from this section is that artificially induced single strand "breaks" preferentially stimulate ØX174 genetic recombination mediated by the rec A "enzyme". I extrapolate from this observation to propose that ØX174 parental RF DNA molecules containing similar random single strand "breaks" are normal intermediate structures in the primary pathway for ØX174 recombinant formation.

(iv) Parental DNA Exchange During Genetic Recombination Between ØX174 DNA Molecules

The occurrence of breakage and reunion during genetic recombination has been convincingly demonstrated only for bacteriophage lambda (Meselson and Weigle, 1961; Kellenberger, Zicchi, and Weigle, 1961). However, several different recombination pathways are active during lambda recombination: the lambda red, int, xis, and the host rec A and rec B pathways are known so far. Since I have shown that most bacteriophage ØX174 genetic recombination takes place under direct control of the rec A "enzyme", it is possible to test whether breakage and reunion occurs during ØX174 recombinant formation via the rec A pathway. Furthermore, the size of the recombinant regions and the size of regions involving DNA synthesis or DNA repair may be determined by this approach. The chief disadvantages are the low recombination frequencies (not too serious) and the fact that only one of the two parental replicative form DNA strands can easily be density labeled (which limits the resolution).

Recombination occurs efficiently between parental RF DNA molecules as shown in table 4 by genetic crosses between mutants defective in cistron A (which therefore do not make progeny RF). Indirectly this also constitutes a formal proof that recombinant RF is made, since

RF which are defective in cistron A can not be rescued by complementation (Sinsheimer, 1968). Two independent experimental designs were employed to look for parental DNA exchange (breakage and reunion) during the genetic recombination of bacteriophage ØX174.

The first design involves recombinant formation between two "heavy" phages, followed by an assay for wt recombinants before the parental RF are able to replicate. This can be done either genetically by blocking further parental RF replication, or physiologically by looking for wt recombinants at an early time. The results of a typical experiment are shown in figure 18. Benbow, Schafer, and Sinsheimer (1973) show that these same results obtain when a variety of conditions are used to block parental RF DNA replication.

In figure 18 a peak of wt recombinants is found one fraction to the "light" side of the fully "heavy" DNA molecules. Fully "heavy" in this case actually refers only to the parental DNA strand; the density of the RF molecule, including the newly synthesized complementary strand is "heavy-light". The finding that the majority of the wt recombinants are nearly "heavy" implies either that most recombinants are generated by a breakage and reunion process between two "heavy" DNA molecules, or that recombination involves only a very small region of one "heavy"

Figure 18. Parental DNA Exchange During Genetic Recombination Between Two "Heavy" Bacteriophage ØX174 DNA Molecules

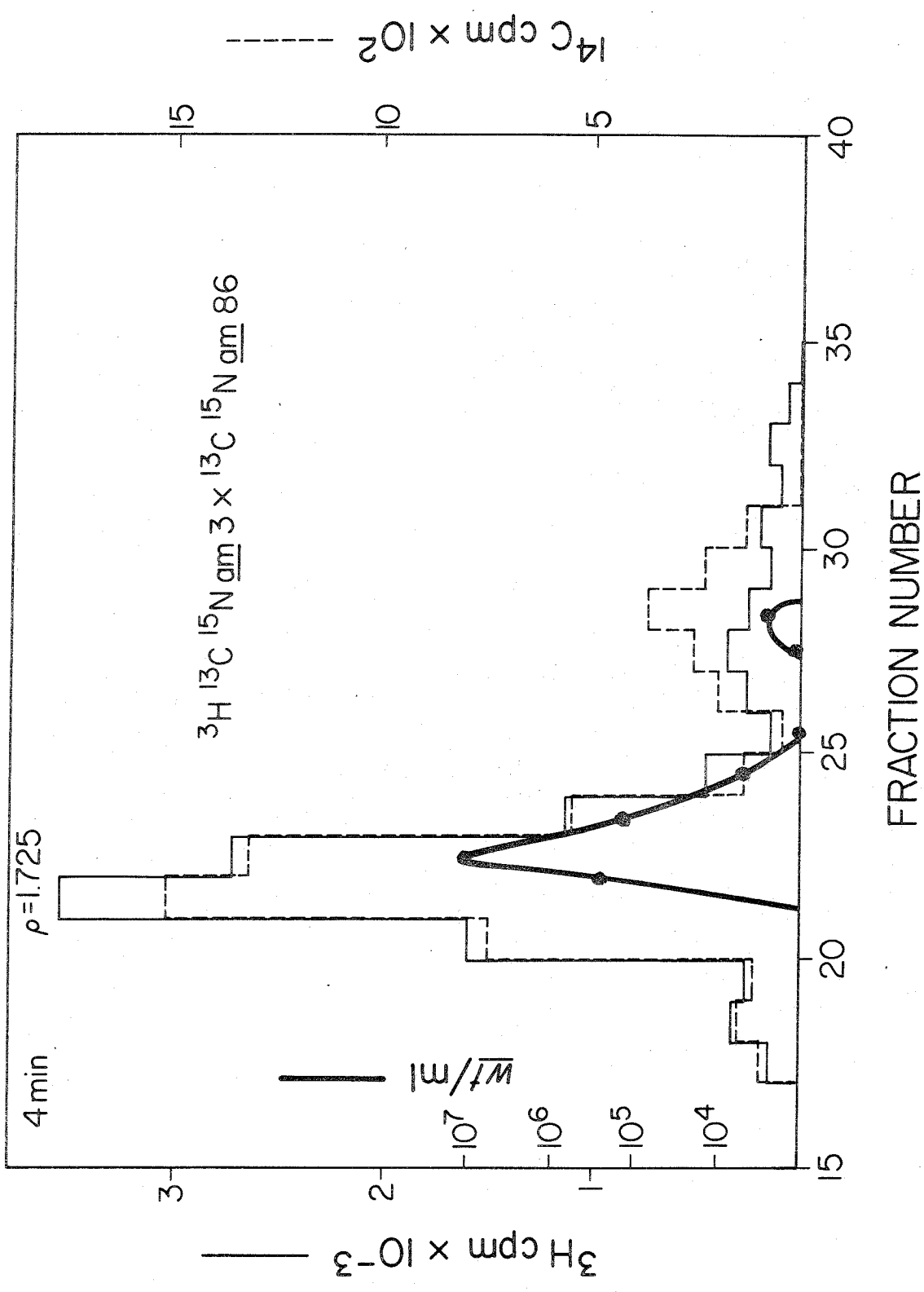
^{3}H cpm (parental strand label), ^{14}C cpm (complementary newly synthesized DNA strand label), and wt recombinants per ml (assayed by the recombination free spheroplast assay described in figure 14) are plotted against fraction number for a CsCl sedimentation equilibrium density gradient.

HF4704, a thy host, was grown at 37°C with aeration in TPG-AA + 2 micrograms/ml thymine to a concentration of 2 x 10^{8} cells per ml. The culture was mode 0.003M KCN and aerated for 10 minutes. $^{3}H^{13}C^{15}N$ am 3(E) and $^{13}C^{15}N$ am 86(A) were added at a multiplicity of infection of 10 along with 2 microcuries/ml ^{14}C thymidine. After 15 minutes at 37°C for adsorption, the culture was pelleted and resuspended in fresh medium. At 4 minutes (and at 20 minutes) an aliquot of infected cells was pelleted; ØX174 replicative form DNA was extracted by the procedure of Komano and Sinsheimer (1968) except that 0.03M KCN was present to prevent further DNA replication. ØX174 RF DNA was purified by sucrose velocity sedimentation in an SW41 rotor as described in figure 16.

The resulting ØX174 RF DNA was sedimented to equilibrium in CsCl in an SW50 rotor for 36-48 hours.

The pattern for t = 4 minutes is shown above.

The pattern for t = 20 minutes, which is not drawn, exhibits a large ^{14}C peak at fully "light" density and a 5-10 fold excess of wt recombinants at the fully "light" density. This is consistent with the proposal that the recombinants which are formed by 4 minutes form substantial quantities of recombinant progeny RF by 20 minutes.


^{14}C cpm x 10^2
FRACTION NUMBER
^3H cpm x 10^{-3}
^3H ^{13}C ^{15}N am 3 x ^{13}C ^{15}N am 86
ρ=1.725
4 min
wt/ml

DNA molecule which is filled in by copy choice or by DNA repair.

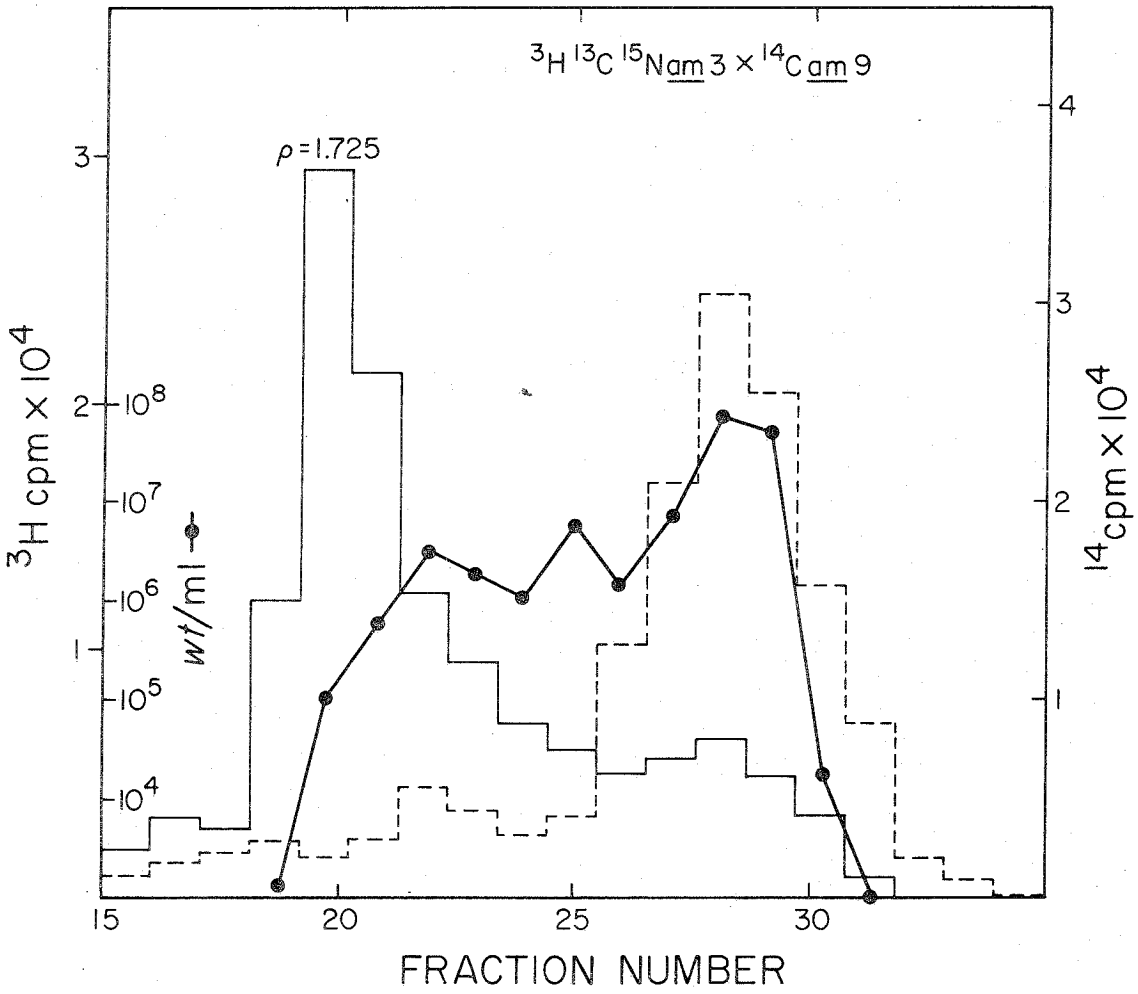
Since the second experimental design (shown in figure 19) proves that recombinant formation can involve substantial (25%) regions of each of the two genomes, the density shift to "light" can be used to estimate the amount of DNA filled in by DNA repair or by new DNA synthesis. Assuming a roughly linear relationship between the density shift and the genome percentage of "light" DNA inserted into the "heavy" genome, newly synthesized regions involve less than 12% of the genome or about 600 nucleotides. This means that about 300 nucleotides or less per single strand break are involved in joining recombinant DNA molecules---assuming we are not measuring contributions from any DNA repair processes elsewhere.

The formation of wt recombinants between one "heavy" and one "light" parental DNA molecule was examined in the second experimental design as shown in figure 19. The prediction: if recombination involves only small regions of the genome then two recombinant peaks will be seen--- at positions one or two fractions removed from fully "heavy" and fully "light" DNA molecules. If substantial portions of the genome participate in breakage and reunion, wt recombinants should be seen as a broad peak across the entire region between "heavy" and "light" DNA molecules.

Figure 19. Parental DNA Exchange During Genetic Recombination Between One "Heavy" and One "Light" Bacteriophage ØX174 DNA Molecule

^{3}H cpm ("heavy" parental DNA label), ^{14}C cpm ("light" parental DNA label), and wt recombinants per ml are plotted against fraction number for a CsCl equilibrium sedimentation density gradient.

The RF DNA molecules were prepared as described in the legend to figure 18 except that the experiment was terminated with 0.03 M KCN at 4 minutes after the initiation of DNA replication.


$^{3}H\ ^{13}C\ ^{15}N$ am 3 × ^{14}C am 9
ρ = 1.725
^{3}H cpm × 10^{4}
^{14}cpm × 10^{4}
wt/ml
10^{8}
10^{7}
10^{6}
10^{5}
10^{4}
1
2
3
4
15
20
25
30
FRACTION NUMBER

In fact a broad peak is seen, but with two (or three?) maxima as shown in figure 19. This implies that most recombination events involve the insertion of regions of less than 25% of the genome, but that a significant percentage of large (25-50%) exchanges can occur.

These experiments do not answer the question whether one or two strands of one of the DNA molecules is inserted into the other molecule in order to produce the recombinant.

Once a wt recombinant DNA molecule has been formed, it begins to direct the formation of wt recombinant progeny RF DNA molecules. Evidence for this was obtained by Benbow, Schafer, and Sinsheimer (1973). They show that the majority of wt recombinant DNA molecules in a cross between two "heavy" DNA molecules are predominantly "heavy" as shown in figure 18, but that wt recombinant DNA molecules formed in the same cross 20 minutes after the initiation of DNA replication are predominantly "light" (legend to figure 18).

The data presented in figures 18 and 19 are most simply explained by the idea that breakage and reunion takes place during rec A mediated ØX174 genetic recombination. A model utilizing these results is given in section vi. No attempt was made to investigate the secondary mechanism of genetic recombination (the residual ØX174

recombination which occurs in rec A cells).

(v) High Negative Interference in Bacteriophage ØX174 Genetic Recombination

A purely genetic examination of high negative interference during ØX174 recombinant formation yields some indirect conclusions which are not easy to obtain by direct biophysical measurements. First, the average size of a genetic exchange region is estimated by examining segregation of a third marker which is closely linked to two markers for which recombinants are selected. One class of the segregants are molecules which have undergone a second exchange. By varying distances between the external markers, one can estimate the percentage of all crosses that involve regions of defined length. This is simply a measurement of high negative interference, but with the important advantage that we know the physical distance (in nucleotides) between the genetic markers.

Secondly, the frequency of multiple exchanges (recombination events involving molecules which previously have recombined) can be measured. The only limitation is the large number of agar plates required to determine the segregation of markers in four-factor crosses.

Data demonstrating high negative interference in bacteriophage ØX174 genetic recombination are shown in Benbow, Hutchison, Fabricant, and Sinsheimer (1971).

As previously established for bacteriophage lambda (Amati and Meselson, 1965), the frequency of double exchanges occurring between closely linked markers is considerably higher than the product of the appropriate two-factor recombination frequencies.

From Amati and Meselson (1965) we note that i, the index of interference, is defined as:

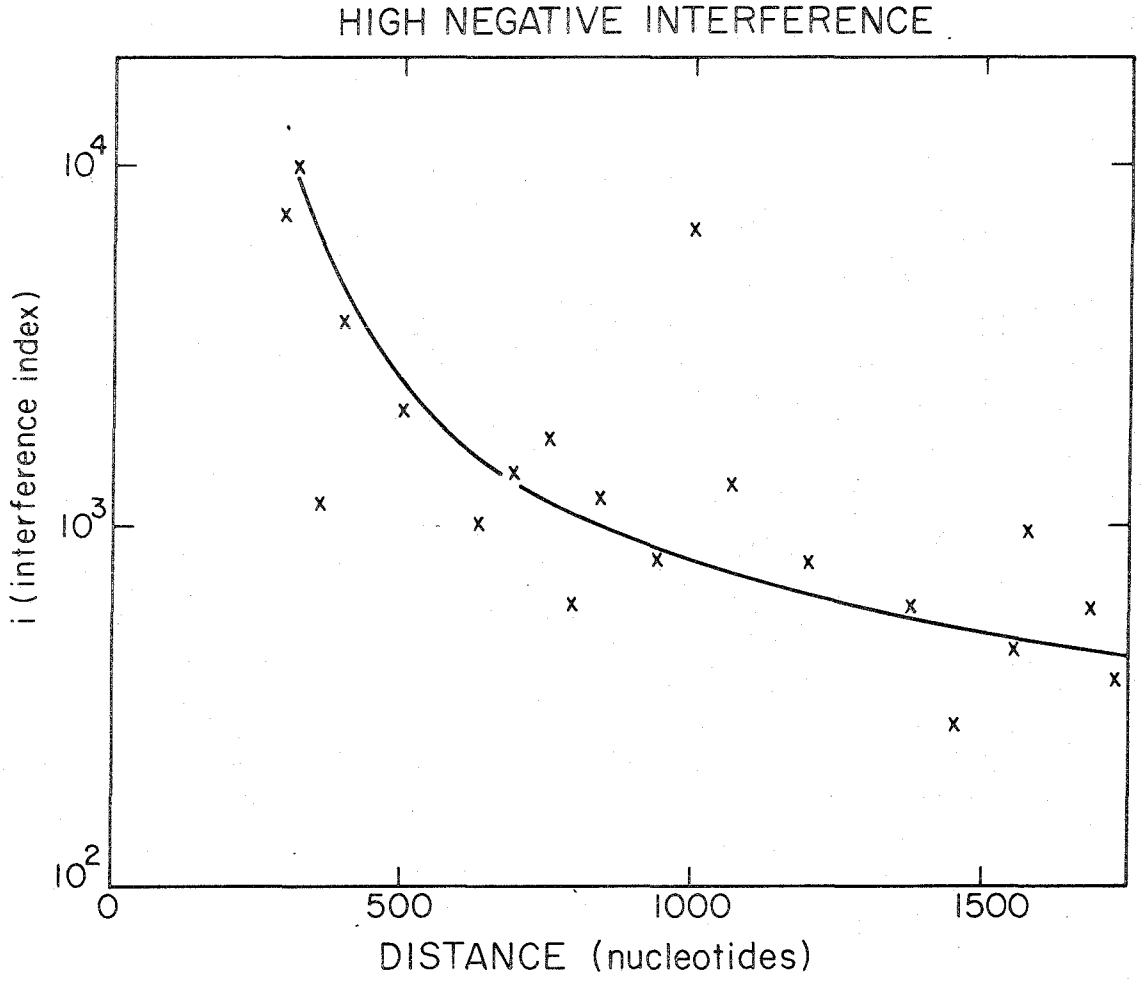
$$i = R_d/R_{12}R_{23}.$$

A plot of i against $R_{12} + R_{23}$ for bacteriophage lambda shows that i rises from 2 or 3 for the most distant markers to 70 for the most closely linked markers. For bacteriophage ØX174 a similar plot shows that i varies from 70 for the most distant markers to over 2000 for closely linked markers (Benbow and Sinsheimer, 1973).

A more interesting plot of i against the physical distance between the external markers is shown in figure 20. The physical distances are estimated from the data in figure 7 and from the results of Benbow, Mayol, Picchi, and Sinsheimer (1972). From these data I estimate that well over 50% of all genetic exchanges involve regions shorter than 400 nucleotides. This is consistent with our measurements of the exchange of parental DNA during ØX174 genetic recombination, and with the extreme map contraction observed for marker separations greater than 500 nucleotides.

Figure 20. High Negative Interference in Bacteriophage ØX174

The interference index, i, is plotted against the distance (in nucleotides) between markers. A sample calculation follows: Let us consider the cross *am* 88(F)-*ts* 79(G) x *ts* 4(H): the two-factor cross *am* 88(F) x *ts* 79(G) yields 1.5×10^{-4} *wt* recombinants per total progeny phage; *ts* 79(G) x *ts* 4(H) gives 1.8×10^{-4} *wt* recombinants per total progeny phage. The three-factor cross *am* 88(F)-*ts* 79(G) x *ts* 4(H), selected for ts^+ recombinants, gives 1.8×10^{-4} ts^+ recombinants, 96% of which are also am^+ and 4% *am* 88(F). The measured number of double exchanges in the marker region *ts* 4(H) to *am* 88(F) is $(4 \times 10^{-2} \times 1.8 \times 10^{-4})$. Therefore, $i = 7.2 \times 10^{-6} / 2.70 \times 10^{-8}$ or 2.7×10^{2}. The distance between *am* 88(F) and *ts* 4(H) is roughly 1410 nucleotides (figure 7). The other points in figure 20 were calculated in a similar manner.


HIGH NEGATIVE INTERFERENCE
i (interference index)
10^4
10^3
10^2
0
500
1000
1500
DISTANCE (nucleotides)

Repeated genetic exchanges are extremely rare during ØX174 genetic recombination (less than 1 in 10^4 recombination events). Apparently multiple mating events as well as repeated matings between the same two genomes are excluded. This was determined by examining the recombinants from the multiple cross am 3(E)och 6(C) x am 3(E)ts 79(G) x och 6(C)ts 79(G). A search for wt multiple recombinants (i.e. a wt must arise either by a reversion event followed by a recombination event or by one recombination event with one genome followed by another recombination event with a different genome) showed that the number of wt recombinants did not exceed the spontaneous background level of about 10^{-7}. (The spontaneous level is higher than might be predicted because the och 6(C) mutation is somewhat leaky).

(vi) A Model for Genetic Recombination Between ØX174 DNA Molecules

The proposed sequence of events leading to ØX174 genetic recombination following infection of cyanide treated rec^+ host cells by two ØX174 virions of different genotypes is outlined in figure 21. During cyanide synchronized infections the input phage particles attach, adsorb, and eclipse (Newbold and Sinsheimer, 1969; 1970a; 1970b). The single stranded DNA molecules are converted to double stranded replicative form DNA molecules. Since single stranded DNA molecules typically are never found free inside the cell, I shall assume that recombination involves only double stranded RF molecules: this has not been proven rigorously.

The initiating event is the introduction of a single strand "break" into either strand of one of the RF DNA molecules. This can occur in numerous ways. Random breaks are generated by copying errors which are excised, by random endonuclease attacks, by failure of ligase to immediately join two DNA strands, and by numerous other mechanisms. In addition, one or more site-specific breaks may occur in one or more of the DNA strands, particularly in the cistron A region. Furthermore, single strand scissions can be introduced artificially by ultraviolet irradiation of the parental DNA before

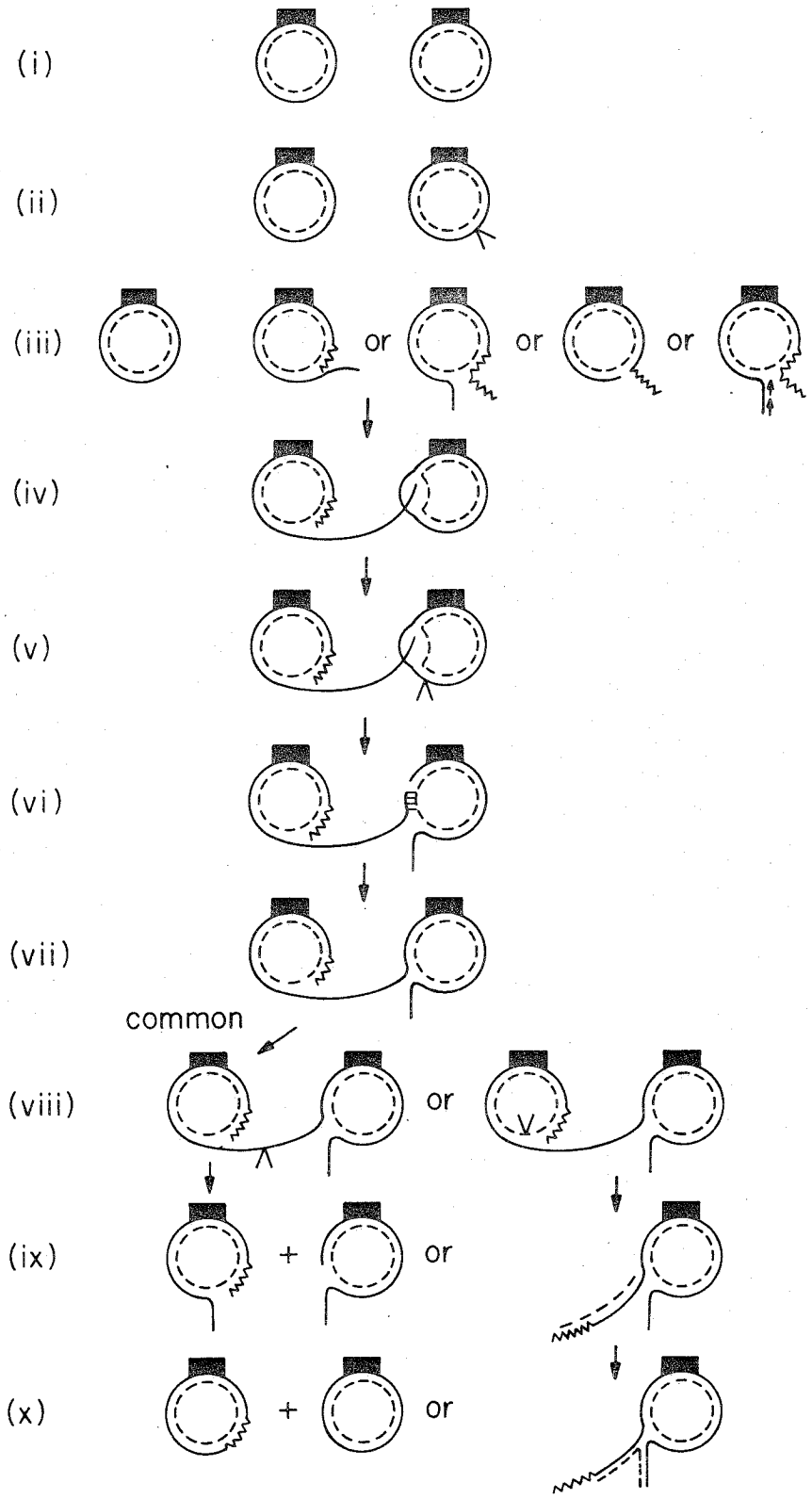
Figure 21. A Model of the Pathway for the Primary Mechanism of Genetic Recombination Between ØX174 DNA Molecules

Two RF I DNA molecules of different genotypes are attached to essential bacterial sites (step i). A single strand scission occurs in one of the two RF DNA molecules (step ii). A "tail" is extended by displacement synthesis which is subject to branch migration (step iii). The extended "tail" attacks the other RF DNA molecule (step v) resulting in a single strand scission in the attacked molecule and formation of a hydrogen bobded joint molecule (step vi). Further branch migration and DNA synthesis may occur and the two RF DNA molecules may become covalently linked (step vii). If a second single strand scission occurs in the same strand initially broken, a single strand insertion event occurs (step viii---common). The resulting RF molecules are one parent and one heteroduplex, which presumably yield one parent and one recombinant in the progeny phage burst.

If the second single strand scission occurs in the complementary strand, the result is shown in step viii. DNA structures of this type are fairly common (1-3%) in ØX174 RF DNA preparations.

The details of ring closure and the elimination of "tails" presumably follow those previously established

for conventional DNA repair mechanisms (Howard-Flanders, 1968; Richardson, 1969).


(i)
(ii)
(iii)
or
or
or
(iv)
(v)
(vi)
(vii)
common
(viii)
or
(ix)
+
or
(x)
+
or

infection, or by thymine starvation of a thymine deficient host immediately after infection. A single strand "break" in one of the two RF molecules is sufficient.

There is no need for two breaks to occur at specific linker points on the same strands as proposed by Holliday (1964). There is no need for both RF molecules simultaneously to be cut in fixed homologous regions as postulated by the Hastings-Whitehouse model (1964). There is no need for unreplicated polarons as proposed by Stadler and Towe (1963). A single random single strand "break" is all that is required.

Since ØX174 genetic recombination occurs in only 0.1 to 1% of all multiply infected cells, the observed frequency of random breaks---for example, the observed number of RF II molecules per cell---is more than adequate to account for the recombination frequencies observed.

A second essential requirement for the primary mechanism of ØX174 genetic recombination is the presence of the rec A "enzyme". Recombination proficient structures (figure 17) survive in rec A cells for up to several hours; these damaged genomes may be induced to recombine up to several hours later following introduction of the rec^+ gene by transduction (Hertman and Luria, 1966; Benbow, Devoret, and Howard-Flanders, 1972). Furthermore, Howard-Flanders, Rupp, Wilkins, and Cole (1968) demonstrated that

these structures survive for up to two hours in rec A cells, but less than 20 minutes in rec B or rec$^+$ cells. The presence of the rec A "enzyme" causes irreversible loss of these structures with a half-life at 37 ^oC of about 7.5 minutes (Cole, personal communication). I wish to speculate that these structures are "branch-migrated" into irretrievably "tangled" structures.

The precise mechanism of action of the rec A "enzyme" is not known; the effect, however, is clear. The recombination proficient structure is destroyed in such a manner that conventional DNA repair mechanisms cannot rescue it. Thus, photoreactivation and excision repair enzymes are unable to rescue a photodamaged circular DNA molecule after it has been acted upon by the rec A "enzyme", i.e. after it has initiated a recombination event.

The next essential event in ØX174 recombinant formation is displacement synthesis and the extension of a "tail" from the RF molecule containing a single strand "break". These "tails" have been observed under a wide variety of conditions; the most appealing demonstration from the point of view of my model was the recent paper by Francke and Ray (1972) presenting electron micrographs of these structures found under conditions where ØX174 DNA replication was blocked by the host rep 3 mutation. Branch migration can occur freely in these structures with

the results pictured in figure 21.

I propose that the next step is the attack of the other RF by the recombination proficient structure. This differs from other models in that it is not a random alignment of two unwound single strands, but rather is the formation of a unique hydrogen bonded structure. In this context it should be noted that there are two hydrogen bonding mechanisms with sufficient precision to form a recombinant structure. The usual Watson-Crick hydrogen bonding structures form complementary double stranded helices. In addition, however, reversed Watson-Crick, and normal and reversed Hoogsteen base pairing can be used to form triple helical structures of great stability. A prediction of interest for recombination mechanisms is that the normal Watson-Crick type of pairing should favor initiation of recombination in G-C rich regions, whereas pairing via triple helices should strongly favor recombination in A-T rich regions (Benbow, unpublished manuscript). No data on this point exist for ØX174, but recombination was reported to occur preferentially for bacteriophage lambda in A-T rich regions (Cordone and Radman, 1970).

A single strand scission is then made in the other (attacked) RF, in either strand, and in the approximate region of the initial single strand "break" due to the strain caused by the attacking action of the extended

"tail". Following this, a hydrogen bonded but not covalently bonded structure is formed between the two RF molecules. These steps are shown in figure 21; a possible hydrogen bonded joint molecule (which looks just like a catenane) is shown in figure 22a.

The consequences of a ØX174 recombination event depend in part on the branch migrated structure that initiates it. The structures which favor a single strand insertion mechanism and a double strand insertion mechanism are drawn with their consequences in figure 21 (see also figure 6b for the reciprocal double strand insertion event).

Completion of a single strand insertion event regenerates RF DNA molecules with "tails". Completion of the first join in a double strand integration event generates structures like the possible recombination intermediate in figure 22b. Depending on the sequence of events, either single or double strand insertion events can give rise to either free RF DNA molecules or catenated RF DNA molecules. Proof that catenanes can contain genomes of two different genotypes is presented by Benbow, Eisenberg, and Sinsheimer, (1972). A recombinant catenane is shown in figure 22c.

Our model predicts that small genomes should favor single strand insertion events whereas larger genomes should favor double strand integration events (the presence of another single strand "break" in either of the two

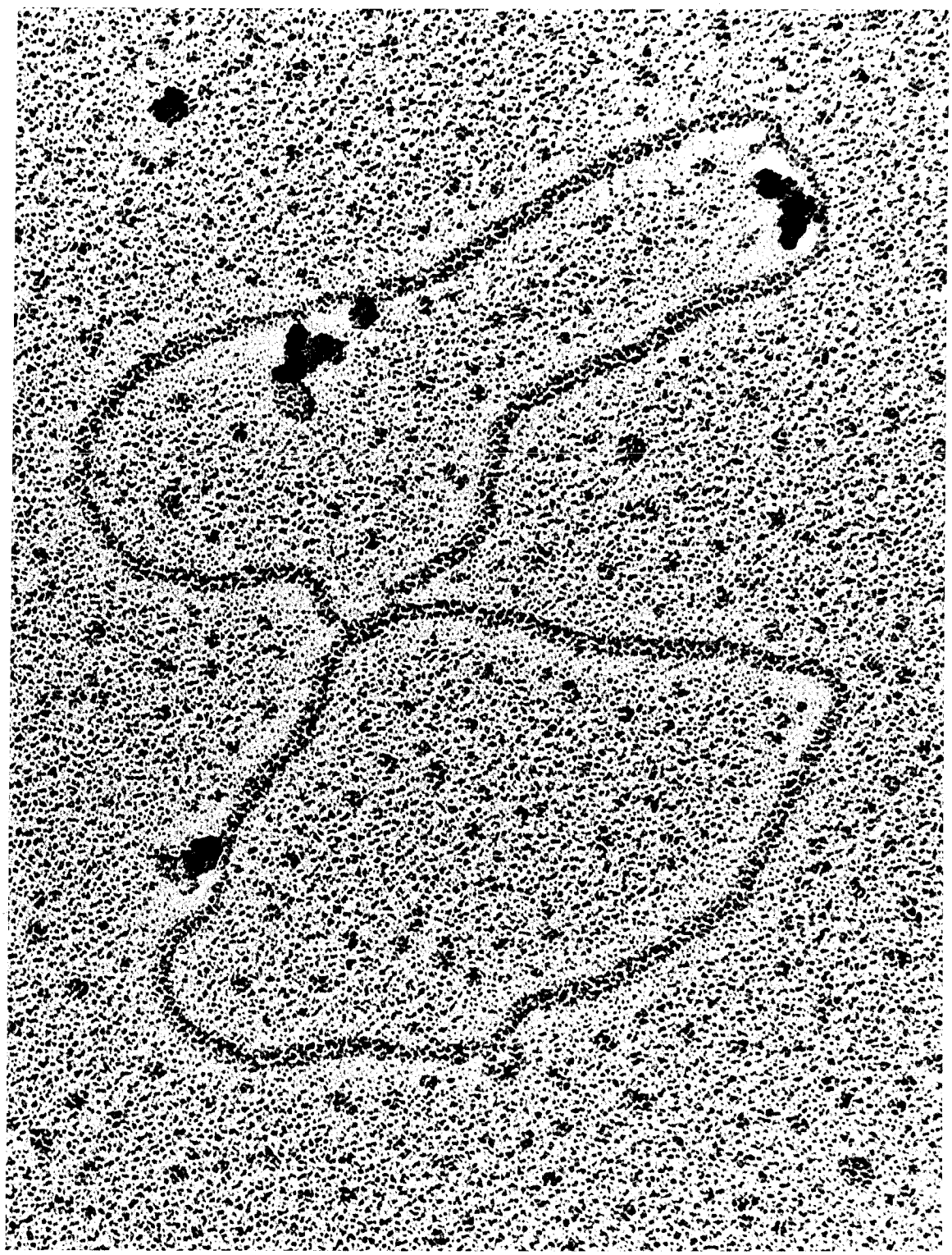
Figure 22. DNA Structures Observed During ØX174 Recombinant Formation

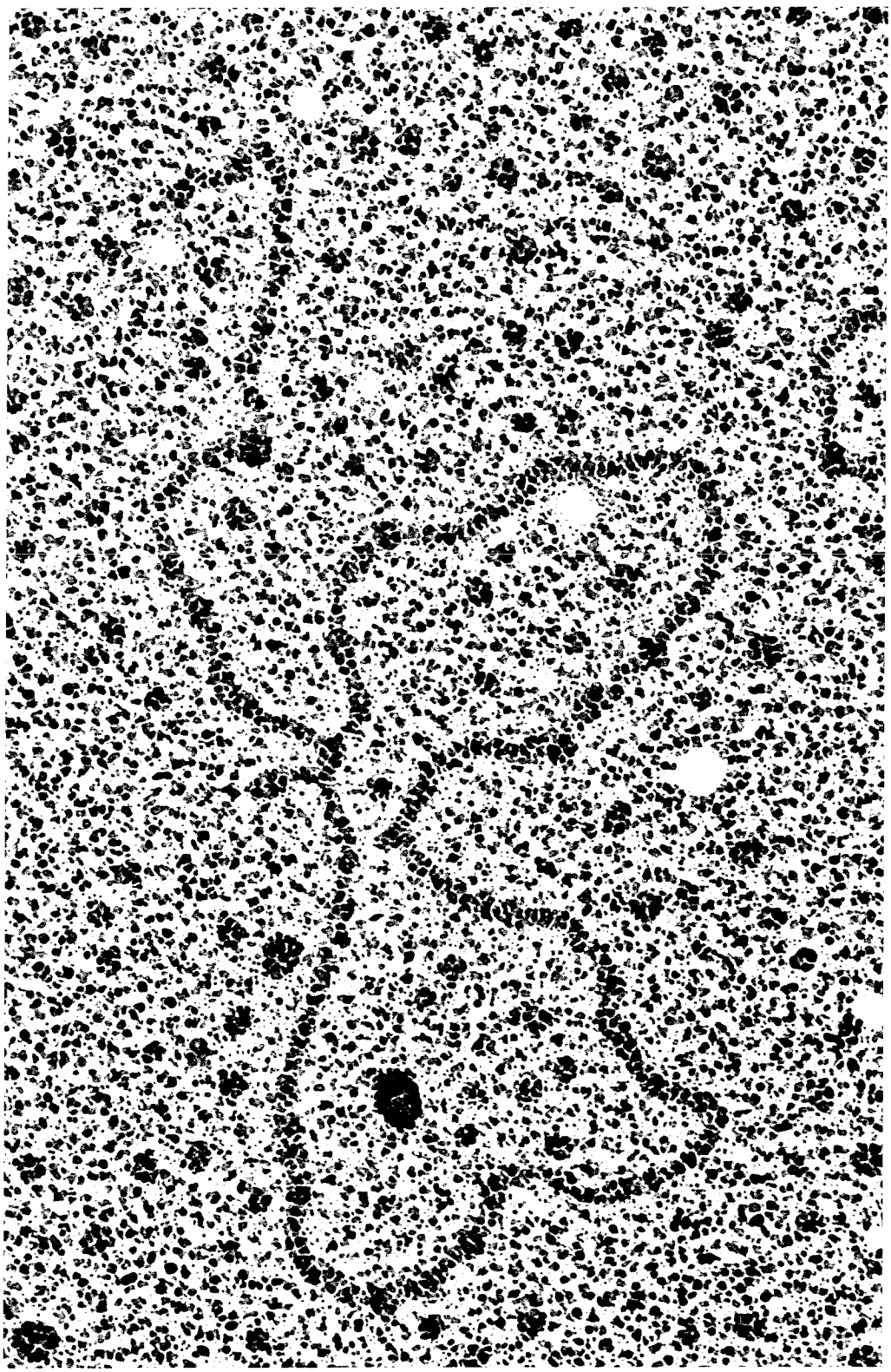
Replicative form DNA molecules were extracted as described in the legend to figure 16 and prepared for electron microscopy as described by Davis, Simon, and Davidson (1971).

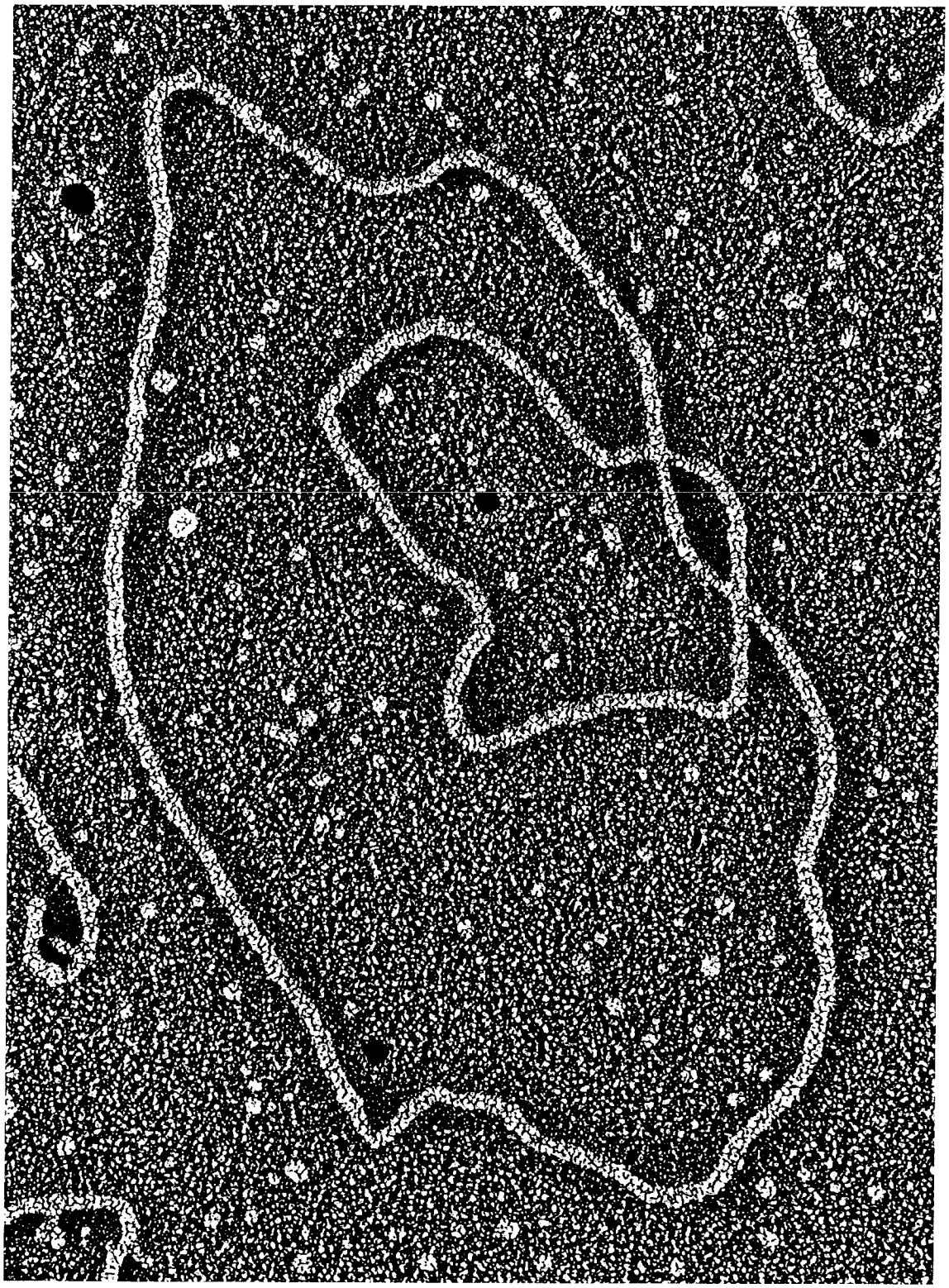
(a) A presumed hydrogen bonded joint molecule isolated from the 27S region of a neutral sucrose gradient run in the SW 41 rotor as described in the legend to figure 16. These molecules are presumed to be hydrogen bonded since 30-40% of the 27S material resediments as RF II after dialysis against 1M NaCl, 0.05M Tris, pH 8.0 at room temperature for 24 hours.

(b) A possible recombination intermediate diagrammed in figure 22. Other proposed recombination intermediates, for example "tailed" RF molecules, have been published previously (Francke and Ray, 1972).

(c) A recombinant 2:1 catenane is shown from Benbow, Eisenberg, and Sinsheimer (1972). A double length am 3(E) genome interlocked with a single length del E 25 genome can not possibly occur simply by DNA replication.







branch migrating molecules should favor the integration of double stranded DNA as long as the second "break" is not in the same strand as the initial "break"). Since ØX174 has a small circular genome (with presumably few breaks per molecule) my model predicts that most recombinants are formed from the integration of single strands up to unit length forming extensive heteroduplex regions. Furthermore, since specific single strand "breaks" with the exception of the proposed site-specific "break" in cistron A are presumably infrequent, my model predicts that the predominant result of a ØX174 genetic cross will be one parent and one recombinant.

Since I have failed to purify the rec A "enzyme", it is at present impossible to test my model by generating recombinants in vitro. I have followed an alternative course: the intermediate and final structures predicted by other models of genetic recombination were not found within the cells which form ØX174 genetic recombinants; the structures predicted by my model have all been observed (figures 16, 22; Benbow, Eisenberg, and Sinsheimer; see also Francke and Ray (1972) for structures that I call recombination intermediates---although the authors suggest that they are DNA replication intermediates.)

Experiments done by Kim, Sharp, and Davidson (1972) with one of my multiple length DNA preparations have shown

that branch migration occurs in ØX174 DNA molecules; its rate is adequate to account for the formation of recombinants by my model.

It has not been possible to devise an experiment using ØX174 which critically distinguishes between the Boon-Zinder model and my model. My arguments against their model (both weak) are that I (and Meselson, 1964) have not measured very much DNA synthesis during the formation of wt recombinants (i.e. 12% seems small in view of the presumably extensive replication their model demands); and that my model readily accounts for the diversity of reciprocal as well as non-reciprocal data in other organisms.

In summary, I have described a mechanism for the genetic recombination of bacteriophage ØX174 DNA molecules, tested most of its specific predictions except for the action of purified rec A "enzyme", and argued that this model has features which make it preferable to existing models of genetic recombination.

REFERENCES

Amati, P. and Meselson, M., Genetics 51, 369, 1965.

Anraku, N. and Tomizawa, J., J. Mol. Biol. 11, 501, 1965a.

Anraku, N. and Tomizawa, J., J. Mol. Biol. 12, 805, 1965b.

Baas, P. and Jansz, H., J. Mol. Biol. 63, 569, 1972.

Baker, R., Doniger, J., and Tessman, I., Nature New Biology 230, 23, 1971.

Baker, R. and Tessman, I., Proc. Nat. Acad. Sci. 58, 1438, 1967.

Barricelli, N.A. and Doermann, A.H., Virology 11, 136, 1960.

Benbow, R.M., Devoret, R., and Howard-Flanders, P., Mol. Gen. Genet., in the press, 1972.

Benbow, R.M., Eisenberg, M, and Sinsheimer, R.L., Nature, in the press, 1972.

Benbow, R.M., Hutchison, C.A., Fabricant, J.D., and Sinsheimer, R.L., J. Virol. 7, 549, 1971.

Benbow, R.M., Mayol, R.F., Picchi, J.C., and Sinsheimer, R.L., J. Virol., in the press, 1972.

Benbow, R.M., Schafer, A.J., and Sinsheimer, R.L., in manuscript, 1973.

Benbow, R.M. and Sinsheimer, R.L., in manuscript, 1973.

Benbow, R.M., Zuccarelli, A.J., Davis, G., and Sinsheimer, R.L., in manuscript, 1973.

Benbow, R.M., Zuccarelli, A.J., and Sinsheimer, R.L., in manuscript, 1973.

Benzer, S., Proc. Nat. Acad. Sci. 45, 1607, 1959.

Benzinger, R., Kleber, I., and Huskey, R., J. Virol. 7 646, 1971.

Bodmer, W.F. and Darlington, A.J., in Caspari, E.W., and Ravin, A.W., Genetic Organization 1, 223, 1969.

Boon, T. and Zinder, N.D., Proc. Nat. Acad. Sci. 64, 573, 1969.

Boon, T. and Zinder, N.D., J. Mol. Biol. 58, 133, 1971.

Broker, T.R. and Lehman, I.R., J. Mol. Biol. 60, 131, 1971.

Cairns, J., J. Mol. Biol. 6, 208, 1963.

Cairns, J. and Denhardt, D., J. Mol. Biol. 43, 458, 1970.

Calendar, R., Lindquist, B., Sironi, G., and Clark, A.J., Virology 40, 72, 1970.

Cassuto, E. and Radding, C.M., Nature New Biology 229, 13, 1971.

Cato, A. and Guild, W.R., J. Mol. Biol. 37, 157, 1968.

Clark, A.J. and Margulies, A.D., Proc. Nat. Acad. Sci. 53, 451, 1965.

Cordone, L. and Radman, M., Virology 41, 166, 1970.

Davis, R., Simon, M., and Davidson, N., Methods in Enzymol. 21D, 518, 1971.

Denhardt, D.T., Dressler, D.H., and Hathaway, A., Proc. Nat. Acad. Sci. 57, 813, 1967.

Denhardt, D.T. and Silver, R.B., Virology 30, 10, 1966.

Denhardt, D.T. and Sinsheimer, R.L., J. Mol. Biol. 641, 1965.

Drake, J.W., Proc. Nat. Acad. Sci. 58, 962, 1967.

Edgar, R.S., Feynman, R.P., Klein, S., Lielausis, I., and Steinberg, C.M., Genetics 47, 179, 1962.

Fiers, W. and Sinsheimer, R.L., J. Mol. Biol. 5, 408, 1962.

Francke, B. and Ray, D.S., Virology 44, 168, 1971a.

Francke, B. and Ray, D.S., J. Mol. Biol. 61, 565, 1971b.

Francke, B. and Ray, D.S., Proc. Nat. Acad. Sci, 58, 1344, 1972.

Freese, E., Genetics 42, 671, 1957.

Fresco, J.R. and Alberts, B.M., Proc. Nat. Acad. Sci. 46, 311, 1960.

Guthrie, G. and Sinsheimer, R.L., Biochem. Biophys. Acta 72, 290, 1963.

Hastings, P.J. and Whitehouse, H.L.K., Nature 201, 1052, 1964.

Hayes, W., The Genetics of Bacteria and their Viruses, Second Edition, John Wiley and Sons, 1968.

Hershey, A.D. and Rotman, R., Genetics 34, 44, 1949.

Hertman, I. and Luria, S., J. Mol. Biol. 23, 117, 1967.

Holliday, R., Genet. Res. 3, 472, 1962.

Holliday, R., Genet. Res. 5, 282, 1964.

Hotchkiss, R.D., Adv. Genet. 16, 325, 1971.

Howard-Flanders, P., Ann. Rev. Biochem. 37, 175, 1968.

Howard-Flanders, P., Rupp, W.D., Wilkins, B.M., and Cole, R.S., Cold Spring Harbor Symp. Quant. Biol. 33, 195, 1968.

Hutchison, C.A., Thesis, California Institute of Technology, 1968.

Hutchison, C.A. and Sinsheimer, R.L., J. Mol. Biol. 18, 429, 1966.

Kaiser, A.D., J. Mol. Biol. 4, 275, 1962.

Kellenberger, G., Zicchi, M.L., and Weigle, J.J., Proc. Nat. Acad. Sci. 47, 869, 1961.

Kim, J.S., Sharp, P.A., and Davidson, N., Proc. Nat. Acad. Sci., in the press, 1972.

Kleinschmidt, A.K., Burton, A., and Sinsheimer, R.L., Science, 142, 961, 1963.

Komano, T. and Sinsheimer, R.L., Biochim. Biophys. Acta 155, 295, 1968.

Kozinski, A.W. and Kozinski, P.B., J. Viorl. 3, 85, 1969.

Lee, C.S., Davis, R.W., and Davidson, N., J. Mol. Biol. 48, 1, 1970.

Lindquist, B. and Sinsheimer, R.L., J. Mol. Biol. 28, 87, 1967.

Lissouba, P., Mousseau, J., Rizet, G., and Rossignol, J.L., Adv. Genet. 11, 343, 1962.

Meselson, M., in Moore, J.A., Ideas in Modern Biology, 3, 1965.

Meselson, M., J. Mol. Biol. 9, 734, 1964.

Meselson, M., in Brink, R.A., Heritage from Mendel, U of Wisconsin Press, Madison, p. 81, 1966.

Meselson, M. and Weigle, J.J., Proc. Nat. Acad. Sci. 47, 857, 1961.

Murray, N.E., Genetics 46, 886, 1961.

Newbold, J.E. and Sinsheimer, R.L., J. Mol. Biol. 49, 49, 1969.

Newbold, J.E. and Sinsheimer, R.L., J. Mol. Biol. 50, 23, 1970.

Newbold, J.E. and Sinsheimer, R.L., J. Virol. 5, 427, 1970.

Norkin, L.C., J. Mol. Biol. 51, 633, 1970.

Oishi, M., Proc. Nat. Acad. Sci. 64, 1292, 1969.

Pritchard, R.H., Heredity 9, 343, 1955.

Pritchard, R.H., Genet. Res. 1, 1, 1960.

Richardson, C.C., Ann. Rev. Biochem. 38, 795, 1969.

Sinsheimer, R.L., J. Mol. Biol. 1, 37-42, 1959.

Sinsheimer, R.L., Prog. Biophys. Nucleic Acid Res. 8, 115, 1968.

Sinsheimer, R.L., Starman, B., Nagler, C., and Guthrie, S., J. Mol. Biol. 4, 142, 1962.

Stadler, D.R. and Towe, A.M., Genetics 48, 1323, 1963.

Stahl, F., J. Chim. Phys. 58, 1072, 1961.

Stahl, F.W., Edgar, R.S., and Steinberg, J., Genetics 50, 539, 1964.

Steinberg, C.M. and Edgar, R.S., Genetics 47, 187, 1962.

Sturtevant, A.H., J. Exp. Zool. 14, 43, 1913.

Taylor, J.H., Haut, W.F., and Tung, G., Proc. Nat. Acad. Sci. 48, 190, 1962.

Tessman, I., Biochem. Biophys. Res. Commun. 22, 169, 1966.

Tessman, I., Science 161, 481, 1968.

Tessman, E. and Tessman, I., Virology 7, 465, 1961.

Whitehouse, H.L.K., Nature 199, 1034, 1963.

Whitehouse, H.L.K., Proc. 11th Int. Congr. Genet. 2, 87, 1965.

Whitehouse, H.L.K., Nature 211, 708, 1966.

Whitehouse, H.L.K., Towards an Understanding of the Mechanism of Heredity, 2nd edition, St. Martin's Press, New York. 1969.

Whitehouse, H.L.K. and Hastings, P.J., Genet. Res. 6, 27, 1965.

Yanofsky, C., Carlton, B.C., Guest, J.R., Helinski, D.R., and Henning, U., Proc. Nat. Acad. Sci. 51, 266, 1964.

Yarus, M.J. and Sinsheimer, R.L., J. Virol. 1, 135, 1967.

Zuccarelli, A.J., Benbow, R.M., and Sinsheimer, R.L., Proc. Nat Acad. Sci., in the press, 1972.