Massey University Library. Thesis Copyright Form

Title of thesis: Prophecy and Gospel Restoration in Enoch 1 and Jubilees

(1) (a) I give permission for my thesis to be made available to readers in the Massey University Library under conditions determined by the Librarian.

(b) I do not wish my thesis to be made available to readers without my written consent for _____ months.

(2) (a) I agree that my thesis, or a copy, may be sent to another institution under conditions determined by the Librarian.

(b) I do not wish my thesis, or a copy, to be sent to another institution without my written consent for _____ months.

(3) (a) I agree that my thesis may be copied for Library use.

(b) I do not wish my thesis to be copied for Library use for _____ months.

Signed T. M. Doherty

Date 27/3/87

The copyright of this thesis belongs to the author. Readers must sign their name in the space below to show that they recognise this. They are asked to add their permanent address.

NAME AND ADDRESS	DATE

PROPHAGES AND GENETIC RESTRICTION IN *ESCHERICHIA COLI*, STRAIN W.

A thesis presented in in partial fulfilment of the requirements for the degree of Master of Science in Microbiology at Massey University.

Timothy Mark Doherty

1987

ABSTRACT.

E.coli strain W (ATCC 9637) is lysogenic for 2 phages which propagate on either E.coli C or E.coli K12. Each phage adsorbs specifically to only one of these hosts and the phages are therefore designated W_c and W_k respectively. These phages display a large number of similarities to phage Mu but restriction endonuclease digestion and probing with radioactively labelled DNA show that the W-phage are not identical to phage Mu.

Lysogens of E.coli C carrying either of the W-phage exhibit strong restriction towards phages P_1 and Lambda; lysogens of E.coli K12 carrying the same phage have not been shown to exhibit such restriction.

Conjugation experiments between E.coli W, E.coli C, and E.coli K12 allowed the mapping of the insertion sites of the 2 W-prophages on the E.coli W genome

ACKNOWLEDGEMENTS.

I would like to thank the technical and academic staff of the Microbiology and Genetics Department, who all gave so me so much help. Most of all however, I would like to thank my Supervisor, Professor Don Bacon, for his inspiration, patience and guidance and also Florence Bacon for all the careful work she did on my behalf.

I would also like to thank all my friends in the Department who made the laboratory such an enjoyable place to work, and whose company also filled so much time, so well, outside the University.

T M Doherty.

TABLE OF CONTENTS.

LIST OF FIGURES

LIST OF PLATES.

INTRODUCTION.

This thesis is primarily concerned with prophage-encoded genetic restriction systems expresed in E.coli strain W. Phage restriction is the expression of one of a number of systems which may prevent a phage from initiating a successful infection of a bacterial host cell. The term genetic restriction applies to the degradation by specific endonucleases, of DNA not native to the restricting cell. These restriction endonucleases recognise specific sequences of nucleotides in the DNA and cleave at, or near these sequences. This not only causes abortive infection by phage, but also restricts the transfer of cellular DNA by conjugation. The synthesis of restriction endonucleases may be directed by the host genome, by plasmids carried by the host, or by an endogenous prophage. Different strains may produce restriction endonucleases of differing activities.

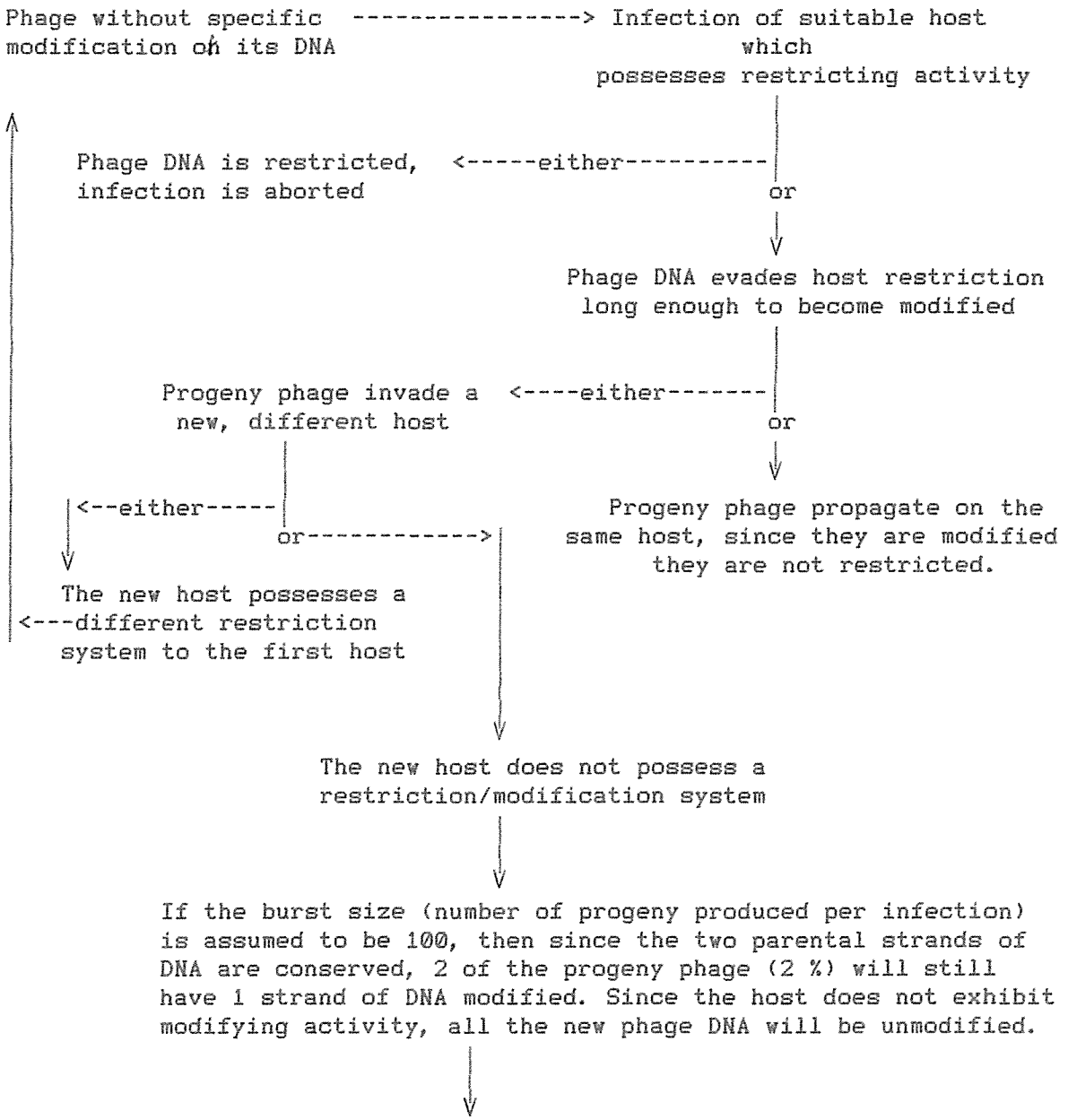
Restriction is almost invariably associated with modification. This may involve either methylation or glucosylation of the restriction enzyme recognition sites. Again, the genes coding for methylation may be located on the host genome, on a plasmid or a prophage. The modification system serves to protect the cell's own DNA from endonucleolytic attack, but will also protect the DNA of an invading phage that evades the effect of the restriction system.

Modification of DNA is not a genetic change, and is therefore not inheritable, although the genes which specify the enzymes responsible for modification obviously are. Phage which propagate on a modifying host retain this feature; each new generation of phage is modified as its DNA is replicated. However, a few generations of growth in a non-modifying

host will cause the modified DNA to be diluted out of the population [Arber (1965), Arber and Linn (1969), Hattman et al. (1966), Revel and Luria (1970)]; this can be best conveyed by the following diagram [see figure 1]

Genetic restriction involves the endonucleolytic scission of DNA; infection of a restricting host by a phage carrying DNA labelled with P^{32} [Dussoix and Arber (1962), Arber et al (1963), Smith and Pizer (1966), Glover et al. (1967)] has been shown to give rise to a number of breakdown products. The major DNA products of this process are acid-insoluble fragments, oligonucleotides, and inorganic phosphates (which are acid-soluble). Breakdown occurrs in two steps, first, relatively infrequent cuts along the length of the genome (dictated by the specific recognition sites) mediated by the restriction enzyme, then a slower degradation of the DNA by various, non-specific DNAases. The progress of restriction can be monitored by the increase in the proportion of label (P^{32}) which becomes acid-soluble with time. After a brief time, however, the restricting cell will begin to utilise the liberated phosphorus, so the level of label in the acid-soluble fraction will decrease.

The evidence for this two-step process in restriction includes the observation that there is some minimal expression of early phage functions in the restricting host, presumably before the DNA becomes totally degraded [Terzi (1968)]. Also, unmodified phage genetic markers can be retrieved in a restricting host by superinfection with a modified phage [Dussoix and Arber (1962)].



If the burst size (number of progeny produced per infection) is assumed to be 100, then since the two parental strands of DNA are conserved, 2 of the progeny phage (2 %) will still have 1 strand of DNA modified. Since the host does not exhibit modifying activity, all the new phage DNA will be unmodified.

Continued propagation in the non-modifying host. If 50 % of the progeny give rise to successful infections, 1 of these is likely to derive from the original modified parent. This new generation of phage will contain 50 x 100 new progeny, only 1 of which (0·02 %) is still modified. This "dilution" of the parent strand will continue until no modified DNA is left in the population. This is an ideal case of course; in reality, attrition due to inefficiencies in mating and breakdown of old DNA will remove modified parent DNA even more swiftly.

FIGURE 1 : DILUTION OF MODIFIED PARENTAL DNA OUT OF A POPULATION BY MAINTAINANCE IN A NON-RESTRICTING HOST

Above the molecular level, restriction is primarily visible as a decrease in the efficiency of plating of the restricted phage, or a decrease in the transmission of markers by conjugation. The level of restriction in any system depends on the physiological status of the restricting cells and also on the activity of the particular restriction endonuclease. Related to this latter factor is the nature of the of the genome which is being restricted. Some restriction enzymes respond to a recognition sequence of six nucleotides. These enzymes cut the DNA relatively infrequently, about once every 1024 nucleotides. Other restriction enzymes may recognise a sequence of four or five nucleotides. These enzymes, by simple random chance, will have a greater opportunity of binding to a recognition sequence and will therefore cut the DNA more frequently. Obviously, the greater the length of unmodified DNA exposed to the restriction enzyme, the greater the chance that recognition sequences will be available for the restriction enzyme to initiate an attack on the DNA. Arber has shown [Arber and Linn (1969)] that the loss of one recognition site from the genome of phage fd by mutation increases its efficiency of plating on the restricting host E.coli B from 7×10^{-4} (wild type, 2 recognition sites) to 3×10^{-2} (mutant, 1 recognition site). Statistically, the chances for a length of DNA, whether bacterial or phage in origin, with three or more recognition sites to survive exposure to restriction will be low. However such phages do propagate on restricting hosts (albeit at low frequencies). In these cases, survival is probably due to cells which are poor in their ability to restrict due to their physiological state. These weakly restricting cells still retain their modifying activity and the progeny phage may therefore be modified and able to initiate a spreading infection. Mutants totally lacking in

restrictive ability do exist, but they form a negligible part of the natural population and are less likely to provide centers of infection than the "weak" cells already mentioned. PB1395 (used in the work described this thesis) is an example of one of these restrictionless mutants which retains its ability to modify.

Various environmental factors such as the age of the culture, the composition of the media, and the length of incubation will influence the physiological state of the cells, and thus the efficiency of the restriction enzymes [Schell and Glover (1966), Arber and Linn (1969)]. If these factors are constant, the efficiency of plating of a given unmodified phage on a specific strain of bacteria is constant, and characteristic for that system.

In view of the relatively greater length of bacterial DNA, it is only reasonable to assume that there is a correspondingly greater number of potential restriction sites. This being so, conjugation between a restricting F^- (recipient) cell and an unmodified F^+ (donor) cell would seem unlikely to result in the transferral of much genetic information. However it has been observed by Glover, Copeland and other workers that approximately 20 minutes after mating commences the restriction system of the recipient appears to lose its effectiveness. It has been suggested that this is due to the saturation of the restricting endonucleases with unmodified DNA from the donor. The influence of restriction on bacterial mating has been thoroughly explored by Wood [Wood (1966)] and Arber [Arber and Morse (1964)].

Present evidence appears to suggest that the restriction endonucleases are located in the periplasm between the cell wall and the cell membrane. The evidence for this includes conversion of cells to spheroplasts

[Moholt and Frazer (1965)] and treatments which are known to remove cell membrane peripheral proteins, such as osmotic shock or treatment with EDTA [Schell and Glover (1966)]. Such treatment has been shown to decrease or eliminate restricting ability. Some workers [Schell and Glover (1966)] have infected restricting cells with unmodified phage at a temperature which decreased the efficiency of restriction. When the cells were resuspended in a hypertonic medium at 37°C restricting ability was restored, but a significantly increased number of phage were able to initiate sucessful infections, indicating that the restriction system was ineffective once the phage had penetrated the cell membrane. Control cells treated in the same way (barring pre-adsorption of phage at the lower temperature) showed no alteration in the efficiency of plating at 37°C when compared to untreated cells, so the increase in the number of sucessful infections was not due to damage to the restriction system of the treated cells.

These results are particularly germane to E.coli W, because the W-prophage-mediated restriction system does not appear to possess an associated modification system. It is the only modification-less restriction system covered in the literature. It is obvious that the genomes of both the phage and its host must be protected from digestion by their own enzymes by some other strategy. Furthermore, this method must be independant of any feature peculiar to E.coli W since the W-phage can successfully lysogenise other strains of E.coli. Strict compartmentalisation of the endonuclease in the periplasm of the cell is a likely method of protection. While it is possible that the W-phage might lack a recognition site for its own restriction enzyme, it is not plausible that the necessary site does not exist along the length of a

bacterial genome. In addition to this, the very marked effect that restriction has on bacterial conjugation observed in matings between sub-strains of E.coli W, indicates that the phage restriction is active against the E.coli W genome. Since there is reason to believe that a host factor is necessary for the activity of the phage-coded restriction system, the compartmentalisation of this hypothetical host factor could serve the same purpose as compartmentalisation of the enzyme and act as a further safeguard. E.coli W has its own chromosomally-encoded restriction and modification system of the hsd type but the modification by this system does not give protection from the W-phage mediated restriction endonuclease. The restriction systems operating in E.coli W are more fully described in the Historical Review and in Experimental section 1.

The restriction system of bacteria is a mechanism which protects the genetic "integrity" of the cell. It is ideally suited to its role as a first line of defence against phage, since its efficiency is greatest toward a phage which the cell has not previously encountered, a stage when neither cell or phage has evolved any mechanism of co-existance. Although restriction also inhibits the free exchange of genetic material, by conjugation, transduction or transformation and would therefore seem to be a limiting factor to the evolution of bacteria, some mechanism to preserve genetic isolation may not be a disadvantage.

The modification associated with most restriction systems has a protective role. DNA which survives restriction activity and is modified by the host is not restricted when it again enters cells of that strain. Since any restriction enzyme requires that both strands of a piece of DNA be unmodified before it can initiate an attack, the semi-conservative nature of DNA replication ensures that newly synthesised DNA will not be

restricted. The new DNA, being thus protected from restriction, will eventually become modified, and so will protect its own newly synthesised DNA when it in turn is replicated.

There are at least two different forms of modification. The first type involves specific methylation of either cytosine or adenine at the recognition site of the restriction enzyme, rendering that particular restriction endonuclease unable to bind to the DNA and cleave it. The crucial role of methionine in many modifying systems was suggested by the fact that modification was reduced in strains which were starved for methionine [Arber 1965]. Modification may also be reduced in cells which have been pre-infected with U.V.-irradiated phage T_3, which causes cleavage of s-adenosyl methionine, which is the substrate for the specific DNA methylase [Hirsch-kauffmann and Sauerbier 1968]. While these observations are strong indications of an important involvement of methionine in modification, they cannot be considered conclusive, since both involve gross alteration of the state of the modifying cells physiological state, and thus directly affect the growth of the phage.

A more direct link is provided by the determination of the relative degree of methylation of modified and unmodified DNA. It must be noted that most phage and bacterial DNA is non-specifically methylated. The relatively small proportion of extra methylations required to modify the specific recognition sites of the DNA is not easily distinguished. The DNA of phage fd however, is not heavily methylated. Using labelled 6-methyl-aminopurine (6-MAP) it was found that unmodified phage fd DNA has approximately 1 methylated adenine base per 4000 nucleotides. After modification by E.coli B the phage fd DNA carried 2 methylated adenine base per 4000 nucleotides. Since a single-stranded length of phage fd DNA

is about 6000 nucleotides, the B-specific modification added 2 methylated adenine bases. This agrees with the current evidence that unmodified phage fd DNA carries two B-specific sites on its DNA [Mamelak and Boyer 1970].

A second type of modification has been demonstrated with the T-even phage [Revel and Luria 1970]. These phage have hydroxy-methyl-deoxy-cytidylate (dHMP) instead of deoxycytidylate (dCMP) in their DNA. The hydroxymethyl group is glucosylated by glucosyl transferases using uridine diphosphoglucose (UDPG). This reaction is carried out by host enzymes induced and directed by the phage genome. If either the glucosyl transferases or UDPG are unavailable, modification of the phage DNA does not proceed, and the phage genome remains vulnerable to restriction.

There is a significant functional difference between these two mechanisms of modification, beyond the fact that they involve different enzymes and substrates. The bacterial modification system serves as a protective device against the cells own restriction enzymes should compartmentalisation prove insufficient. The bacterial modification system is therefore specifically intended to complement the bacterial restriction system.

The glucosylation of the T-even phage is a phage strategem unrelated to bacterial restriction/modification systems (except insofar as it is designed to circumvent them). It serves to preserve the DNA of the infecting phage from restriction.

Restriction-deficient (Res^-) mutants do occur in both bacterial and phage encoded systems [Arber and Linn 1969]. A significant number of Res^- mutants also lack the strain-specific modifying activity (Mod^-). With the exception of E.coli W, no Res^+ Mod^- mutants have been described in the

literature. A number of possibilities may account for this. It has been assumed that restriction without its concurrent modification is likely to prove fatal to the cell concerned. There is the possiblity that laboratory selection of mutants has focussed on changes in restrictive behaviour since there are no techniques available to directly select cells whose modification has been altered.

It has been shown that various strains of E.coli (E.coli K12, E.coli B, E.coli A, et al.) can lose all of their restrictive ability at a frequency compatible with a single mutation. This implies that there is a single locus carrying a gene for a single restricting endonuclease. The existence of this single locus has been confirmed by conjugation and transduction experiments. The locus has been designated the Host Specificity Restriction (hsdR) locus. Together with the genes for the DNA modifying (hsdM) and sequence-recognition (hsdS) functions, this gene forms a closely linked group located about 1 minute clockwise of threonine on the map of the E.coli chromosome in the different strains of restricting E.coli (E.coli K12, E.coli B, E.coli A, and E.coli W) [Arber and Linn 1969, Bachmann 1983]. These genes have been shown to be allellic by their ability to be transduced with the serine B or thymine R markers. The great similarity of these genes and their positions on the gene map suggest that the hsd genes of the restricting E.coli may be derived from a common ancestor.

The hsd locus includes 3 closely linked genes ; the hsdR, which codes for the restricting endonuclease activity, the hsdM, which codes for the modifying enzyme activity, and the hsdS, which codes for the protein which binds specifically to the recognition site that the preceding gene products act from.

Modification/ restriction systems can be divided into three groups on the basis of protein structure and their mode of action [Lewin 1983]. The simplest and most common (it occurs in about one in three bacterial strains) modification and restriction system is the type II. The type II restriction/ modification system has 2 genes coding for a restriction endonuclease and a modification methylase. Each of these enzymes, which is totally independant of the other, includes an active site which recognises the specific recognition site. The recognition site is often palindromic and is usually 5 - 6 nucleotides in length. The restriction endonuclease cuts at or near this recognition site, by breaking one bond in each strand of the DNA. Some enzymes may produce staggered cuts, giving rise to so-called "sticky" ends, others may make parallel cuts, giving rise to "blunt" ends. Because these enzymes are highly specific in the patterns in which they cut DNA, they are used extensively to examine the structure of DNA. EcoR1 is a good example of the basic type II restriction enzyme.

A subgroup of the type II restriction endonucleases, (possibly a step forward in evolutionary terms) are enzymes such as Hph, Mbo2, and Hga1, which cleave the DNA at a site slightly removed from the recognition site [Yuan 1981].

The generalised outline given above, of a 3-gene system, applies to the type I restriction/modification systems. This system is present in E.coli B and E.coli K12 (and the other restricting E.coli strains) and was the first restriction system explored at the genetic level. The association of three gene products was first suggested by complementation studies involving mutants of known phenotype. For example the diploid res$^-$ mod$^-$ / res$^-$ mod$^-$ was used as the recipient of an F-prime known to be

carrying *res*$^-$ *mod*$^-$ genes. Where complementation occurred, normal restriction and modification took place. It was shown [Arber and Linn 1969] that restriction and modification were restored simultaneously. This is most easily explained if the genotypes involved were *hsdR*$^-$ *hsdM*$^-$ *hsdS*$^+$ and *hsdR*$^+$ *hsdM*$^+$ *hsdS*$^-$.

The gene products are associated in a pentamer, with 2 restriction endonuclease , 2 modifying methylase , and a single site-recognition protein. Once the pentamer is bound to the DNA, either the restricting or the methylating activities can occur. These activities are mutually exclusive. Cells deficient in the *hsdM* or *hsdS* activities are *res*$^-$ *mod*$^-$; mutations in the *hsdR* gene are *res*$^-$ *mod*$^+$. There appears to be some mechanism to prevent a potentially lethal *res*$^+$ *mod*$^-$ phenotype. ATP is required for the type I enzyme complex to release from an already methylated recognition site or to modify or restrict an unmodified recognition site and it increases the efficiency of modification of the sister strand at a site which is hemi-methylated. The type I restriction enzymes cleave the DNA more than 1000 nucleotides from the recognition site. The exact mechanism is unknown, but electron-microscopic studies reveal looping of the DNA as it is being restricted, implying that the DNA is moved through the enzyme complex until a suitable site for cleavage is found. There is no information yet on what constitutes a suitable site for a type I restriction endonuclease to cleave, but it is not entirely random, since restriction occurs preferentially in some regions of DNA. Cleavage is a 2-step process. Initially, 1 strand is broken, then a second break is made in the other strand at a site nearby.

The type III restriction/modification system is coded for by 2 genes. One specifies a protein with *hsdR* activity, the other a protein with 2

domains carrying both hsdM and hsdS activities. The two proteins are associated in a dimer which requires ATP to bind to the DNA. Once bound, both restricting and modifying activities are expressed simultaneously, and each competes to complete its task before the other can. Modification takes place at the recognition site, while restriction takes place 24-26 base pairs away from the recognition site, presumably because of steric interactions. One interesting point in which type III systems differ from the other 2 types of restriction system is that the DNA is only hemimethylated by its modification system, that is, only the recognition site on one strand of the DNA is methylated. How modification is maintained during the replication of the DNA is unknown, but it has been suggested that the modification enzymes are linked to replication.