STADLER SYMP. Vol. 10 (1976) University of Missouri, Columbia 65

PHASE VARIATION IN SALMONELLA: ANALYSIS OF THE CONTROLLING ELEMENT OF *H2* GENE EXPRESSION

(gene expression, recombinational control, phase variation, DNA inversion)

JANINE ZIEG, MICHAEL SILVERMAN, MARCIA HILMEN & MELVIN SIMON

Department of Biology (B-022)
University of California, San Diego
La Jolla, California 92093

SUMMARY

In the phase variation system of Salmonella, the alternative expression of H1 and H2 flagella is controlled by a region of DNA adjacent to the H2 structural gene known as the Phase Determinant. The Phase Determinant regulates H2 gene activity via a site specific recombinational event. Electron microscopic evidence and restriction endonuclease site mapping indicate that the recombinational event results in an inversion of a region of DNA 800 bp (base pairs) in length. The inversion process does not depend on the RecA recombinational pathway of E. coli. Plasmids have been constructed in which the expression of non-related genes appear to be under phase variation control. These plasmids have provided evidence concerning the direction of transcription of the H2 structural gene and the position of the H2 promoter.

INTRODUCTION

There appear to be various mechanisms by which organisms regulate the expression of their genetic material. One type of control mechanism, first described by B. McCLINTOCK in *Zea Mays*, involves specific rearrangements of genetic material which result in the activation of certain genes (1957, 1978). Observations in other systems, including *Drosophila* (GREEN, 1977, 1978), *Escherichia coli* (STARLINGER & SAEDLER, 1972), and *Saccharomyces cerevisiae* (HICKS, STRATHERN, & HERSKOWITZ, 1977) may also be explained by invoking mechanisms involving insertions, inversions, and transpositions of DNA. Recently, with the introduction of molecular cloning techniques, evidence for this dynamic property of DNA and its role in the regulation of gene expression has been presented. For example, TONEGAWA has shown that the construction of a mature immunoglobin gene in mice may involve transposition events (TONEGAWA, BRACK, HUZUMI, 1977). Furthermore, the inversion of a specific region of the bacteriophage mu genome, the G region, appears to be correlated with the formation

of viable phage particles under certain conditions (CHOW & BUKHARI 1977).

A clear example of a specific rearrangement of genetic material which is involved in regulating gene expression is found in the phase variation system of *Salmonella* (ZIEG, SILVERMAN, HILMEN & SIMON 1977a, 1977b). *Salmonella* strains have two antigenically distinguishable flagellar types, H1 and H2. These flagellar types differ in approximately 13% of the amino acids which compose the flagellin protein, the structural subunit protein of the flagellar filament (McDONOUGH 1965). The genes which code for H1 and H2 flagellin are located in different regions of the *Salmonella* genome (LEDERBERG & EDWARDS 1953). Their alternative expression appears to be controlled by two genetic elements linked to the H2 gene: the Phase Determinant (PD) (LEDERBERG & IINO 1956), and *rh1*, which must repress expression of the H1 gene (FUGITA, YAMAGUCHI & IINO 1973). It has been proposed that when the PD promotes expression of H2 flagellin, the *rh1* gene product is concomminantly synthesized, and H1 expression is repressed. When the PD does not promote H2 expression, the repressor is also not produced, and H1 is then expressed. A suggested mechanism by which the PD controls H2 gene expression involves a metastable property of the DNA in the region adjacent to the H2 structural gene (ENOMOTO & STOCKER 1975).

Isolation and Characterization of the Phase Determinant

To determine the nature of the metastable change in the *Salmonella* genome, DNA regions carrying the H1 and H2 genes were isolated by two molecular cloning techniques: (1) Eco RI restriction of the genome and ligation onto plasmid vehicle Col E1 (HERSHFIELD, BOYER, YANOFSKY, LOVETT & HELINSKI 1974), and (2) mechanical shearing of the genome, tailing with poly(dC), and annealing to plasmid pAO1 which had been treated with Pst I endonuclease and tailed with poly(dG) (A. OTSUKA, personal communication). The latter method would have the effect of manufacturing Pst I sites at the ends of the sheared fragments. After transformation into an *E. coli* host deficient in synthesis of its own flagellin molecule (Hag^-), colonies that were motile were selected and tested with specific antiserum to determine whether they were expressing H1 or H2 flagella. The clones identified as expressing H2 flagella were also shown to exhibit a type of phase variation between H2-specific motility and non-motility. Thus, it was concluded that the PD must be present and functional on the cloned fragment. When plasmid molecules carrying this region were linearized with a restriction endonuclease, denatured, renatured, and examined by electron microscopy (DAVIS, SIMON & DAVIDSON 1971), a region of nonhomology corresponding to approximately 800 bp was observed (Fig. 1). A direct correlation between the frequency of these "bubble" structures and the phase state of the DNA was established (ZIEG, SILVERMAN, HILMEN & SIMON 1977a). Therefore, it appeared that some specific rearrangement of the DNA was responsible for phase variation in *Salmonella*. Fig. 2 indicates the position of the "bubble" region and restriction endonuclease sites on some of the cloned fragments and derivatives. The small derivative, pJZ110, no longer carries a functional H2 structural gene, thereby mapping the H2 gene to the region of the DNA shown in Fig. 2.

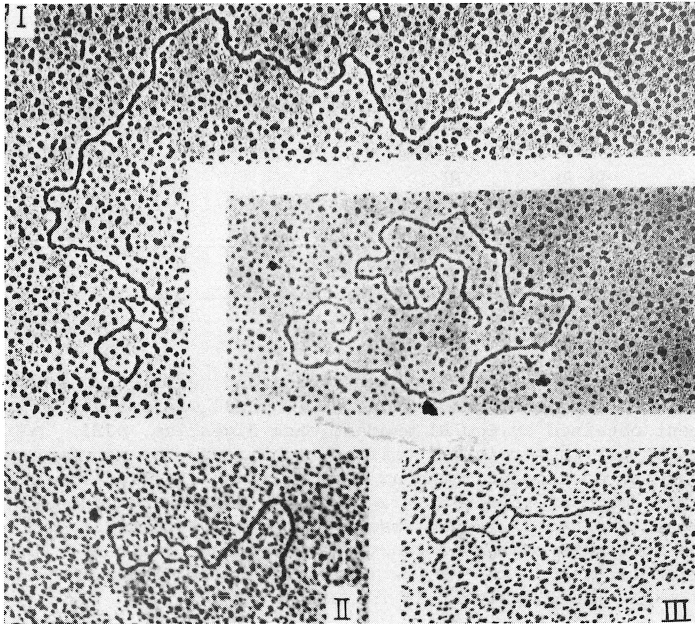


FIGURE 1. Electron micrographs of renatured DNA fragments which carry the H2 structural gene. The 15 kb fragment obtained by Eco RI endonuclease digestion (I), pJZ1, has a region of nonhomology near one end of the molecule. Internal Sal I endonuclease sites allowed the original fragment to be reduced in size (II and III). The smaller derivative (III), pJZ60, is only 3.75 kb in length, yet still carries a functional H2 gene.

The Inversion Model

Electron microscopic examination of the hybrid plasmids using slightly altered renaturation conditions, suggested the nature of the genetic rearrangement. In the structures shown in Fig. 3, the central region is double-stranded while the arms are nonhomologous and, thus, unable to pair. These "H" structures, reminiscent of those seen by HSU & DAVIDSON in the analysis of the G-inversion of bacteriophage mu (1974), are formed by heteroduplexes between strands carrying opposite orientations of this 800 bp region of DNA. "H" structures shown in Fig. 3A were formed by Pst I digested pJZ110 creating 0.1 and 0.15 micron arms flanking the 0.27 micron central region. "H" structures were also seen with heteroduplexes between Pst I treated pJZ110 and Eco RI digested pJZ60 (B). Some of these had the expected 0.1 and 0.15 micron arms of pJZ110 and the 0.5 micron arms of pJZ60. When heteroduplexed structures such as those shown in (B) associate and anneal with each other, the entirely double-stranded structures shown in (C) are observed. These heteroduplexes are all consistent with the notion that the genetic rearrangement involved in phase variation is an inversion of a region of DNA 800 bp in length adjacent to the H2 structural gene.

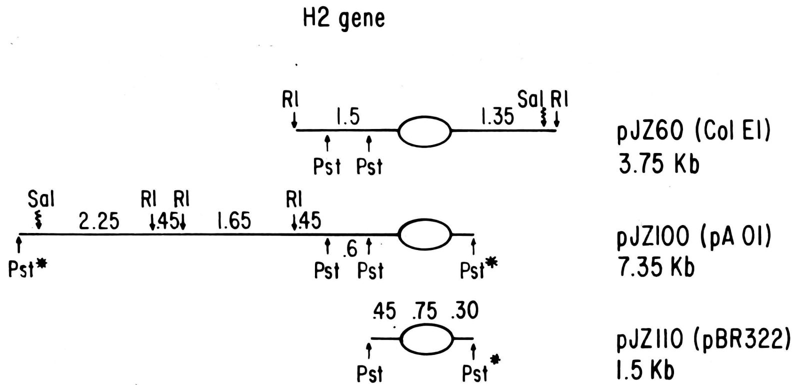


FIGURE 2. Position of the "bubble" region and restriction endonuclease sites on cloned fragments and derivatives. pJZ60 was derived from the original fragment obtained by Eco RI endonuclease digestion, pJZ1. pJZ100 was obtained by the poly(dC)-poly(dG) tailing technique. The manufactured Pst I endonuclease sites are indicated as Pst I* sites. The Pst I* site to the right of the "bubble" region as drawn is only 450 bp from the end of the "bubble" region. The Pst I derived fragment of pJZ110, pJZ110 is only 1.5 kb in length and no longer carries a functional H2 structural gene. The H2 gene must, therefore, lie in the region indicated on the figure. The plasmid vehicle for each cloned fragment is indicated.

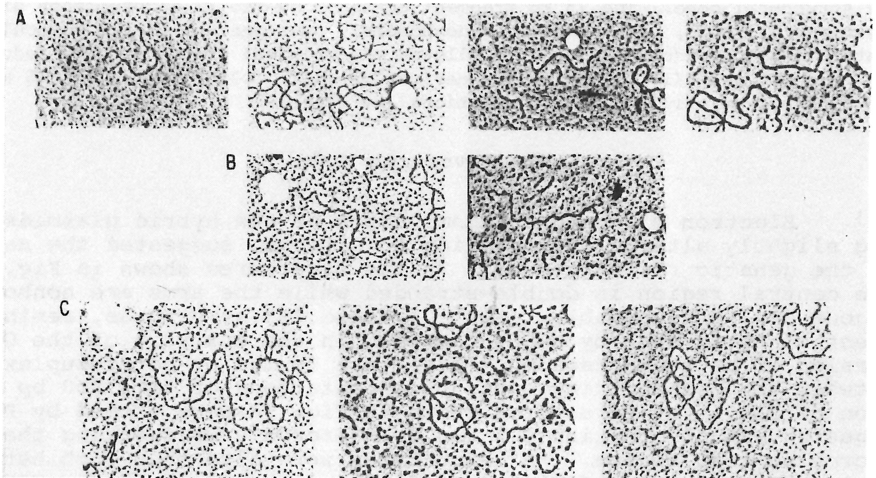


FIGURE 3. "H" structures formed by heteroduplexes between DNA strands carrying opposite orientations of the central 800 bp region. In (A), "H" structures were formed with Pst I digested pJZ110. In (B), heteroduplexed Pst I treated pJZ110 and Eco RI digested pJZ60 also form "H" structures with the predicted arm lengths. The "double-H" structures shown in (C) are composed of molecules such as those in (B) which have associated with each other. The arms from the two "H" molecules have annealled, forming totally double-stranded molecules with two double-stranded central regions.

Restriction Site Analysis

The inversion model was supported by restriction endonuclease site mapping of the cloned fragment carried on pJZ110. Two derivatives of pJZ110, pJZ120 and pJZ140, were used in the restriction site analyses. These derivatives contain identical and inverted arm sequences flanking the inversion region. The formation of these derivatives may have occurred by a recombinational event within multimer molecules in preparations of pJZ110. Fig. 4. shows that digestion of pJZ110 with Pst I endonuclease yields minor bands (low intensity) slightly larger (1.7 kb) and slightly smaller (1.4 kb) in size than the cloned 1.5 kb fragment. Comparison of Fig. 4 (A and B) with Fig. 4 (D and E) illustrates that the intensity of the minor bands reflects the amount of multimer

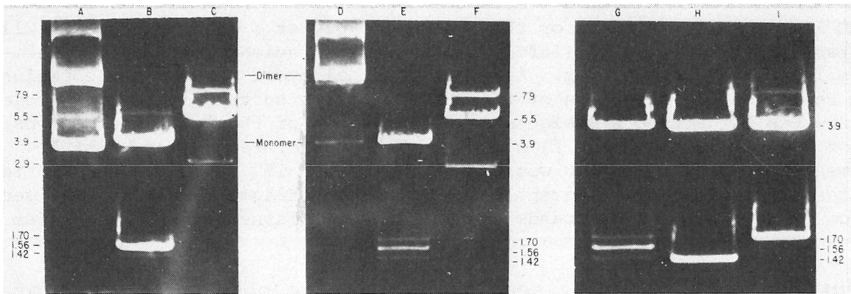


FIGURE 4. Analysis of pJZ110 by 1% agarose gel electrophoresis. In A and D, supercoiled DNA samples of two preparations of the plasmid DNA reveal different amounts of monomer, dimer, and multimer molecules. B and E show the Pst I digestion patterns of the pJZ110 preparations. Different intensities of the 1.70 kb and 1.42 kb Pst I minor bands can also be seen between the two preparations. Eco RI digestion patterns of the two preparations are shown in C and F. Again, the amounts of 7.9 kb and 2.9 kb Eco RI minor bands vary with the preparation. The 1.42 kb and 1.70 kb Pst I minor bands from pJZ110 (G) were isolated and cloned separately on pBR322, and they are shown in H (pJZ140) and I (pJZ120). The size of the restriction fragments is given in kilobase pairs.

forms in the DNA preparation. Eco RI restriction similarly yielded minor bands (Fig. 4C and F). The Pst I minor bands were isolated from the agarose gel and cloned separately on pBR322 (Fig. 4H and I); they are pJZ120 and pJZ140, respectively. Fig. 5 illustrates a possible mechanism for the generation of these minor bands. Within a dimer molecule, a recombinational event between inversion regions oriented in opposite directions (left) would effectively invert the region between the cross-over points. The resulting molecule (right) would give fragments that are slightly larger and smaller in size than the original cloned fragment upon Pst I digestion. To confirm that the pJZ120 and pJZ140 derivatives did contain inverted repeat arm sequences, these plasmids were examined by electron microscopy following Pst I endonuclease digestion, denaturation, and rapid renaturation.

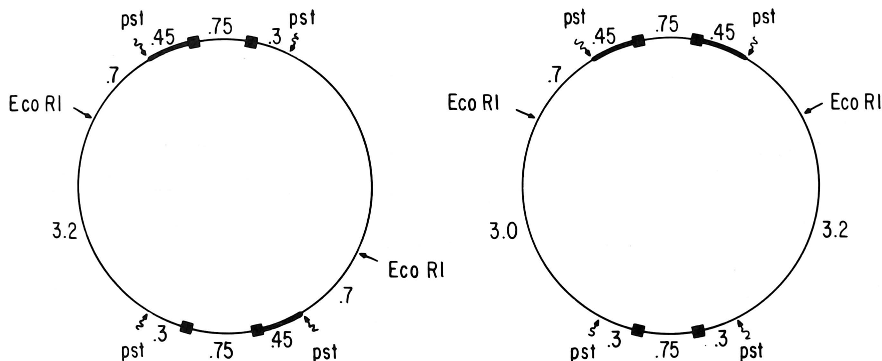


FIGURE 5. Possible scheme for the generation of Pst I minor bands of pJZ110. In a tandem dimer of pJZ110 (left), the inversion regions could orient themselves in opposite directions. A recombinational event between these regions would result in the formation of the molecule shown on the right. This molecule now has two 450 bp arm sequences flanking one of the inversion regions, and two 300 bp arm sequences flanking the other inversion region. Pst I treatment of this recombinant would yield fragments 150 bp longer and shorter than the original pJZ110 fragment - the minor band fragments. Also, the generation of the Eco RI minor bands could also be explained by this mechanism.

Structures such as those seen in Fig 6 were observed. Most of the molecules are the result of rapid intramolecular reannealing. The double-stranded stems measure 0.1 micron or 300 bp in length for pJZ120 (A) and 0.15 micron or 450 bp in length for pJZ140 (B). This corresponds to the length of the DNA that is duplicated on each end of the molecule.

These duplicate arm plasmids were then used to determine the distribution of restriction endonuclease sites within the inversion region and the 300 bp and 450 bp arm sequences (Fig. 7). The pJZ120, pJZ110, and pJZ140 fragments were digested with Hae II, Hae III, and Hpa II restriction endonucleases and the products were examined by polyacrylamide electrophoresis. The maps are consistent with the notion that the central 800 bp region can exist in either orientation. The restriction endonuclease maps can also be used to determine the nucleotide sequence of the region. Valuable information concerning the nature of the end sequences of the inversion region, the beginning of the H2 gene transcript, etc., may be obtained from the nucleotide sequence.

RecA Independence

The sharpness of the bands observed on polyacrylamide gel electrophoresis and the constancy of measurements of "bubble" and "H" structures in the electron microscope suggest that the inversion process is occurring through some site-specific recombinational event at the end points of the inversion region. To determine if the inversion event occurring at these sites is dependent on the host cell's major recombinational pathway, hybrid

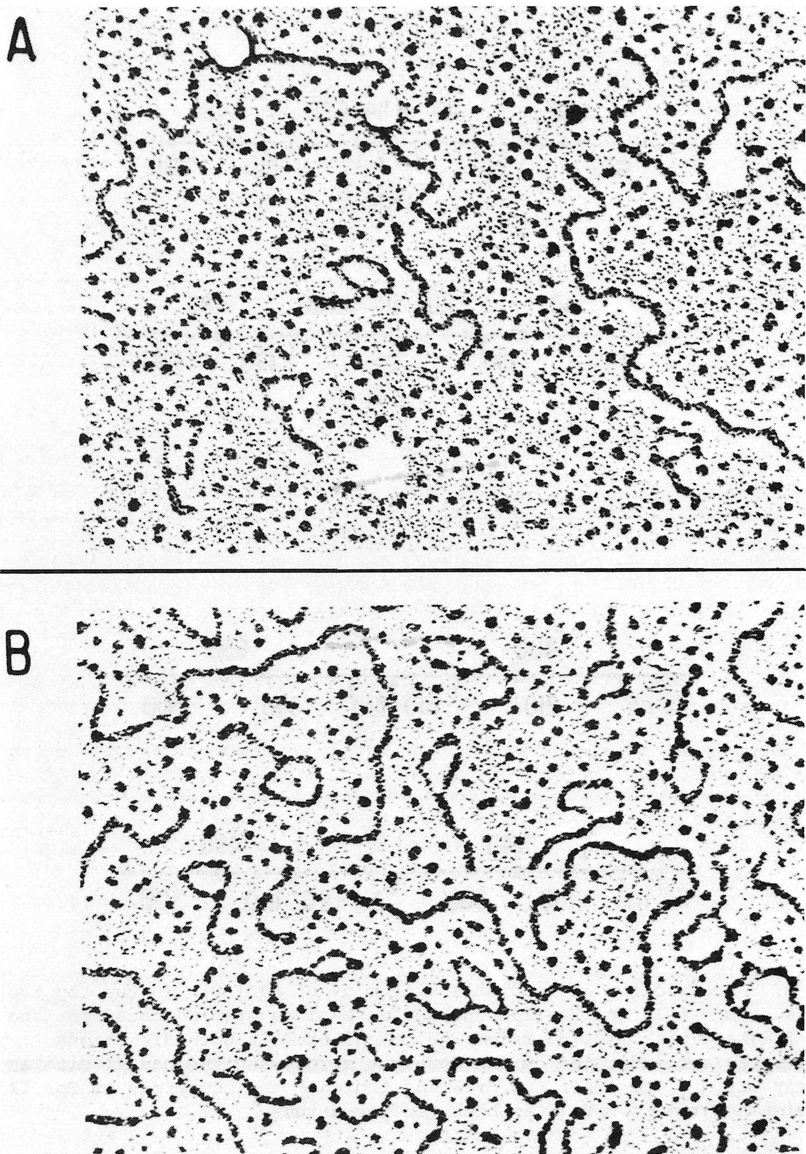


FIGURE 6. Electron micrographs of rapidly reannealled pJZ120 and pJZ140 fragments. The double-stranded stems of structures formed from pJZ120 (A) measure .30 kb, the length of the inverted arm sequences. The single-stranded circle at the end of the stem corresponds to the length of the inversion region. Similarly, in (B), the double-stranded stems correspond to the length of the inverted arms of pJZ140, .45 kb.

plasmids were transformed into RecA$^-$ strains HB101 and MS726. Plasmid DNA prepared from these strains was denatured, renatured, and spread for electron microscopy. "Bubble" and "H" structures

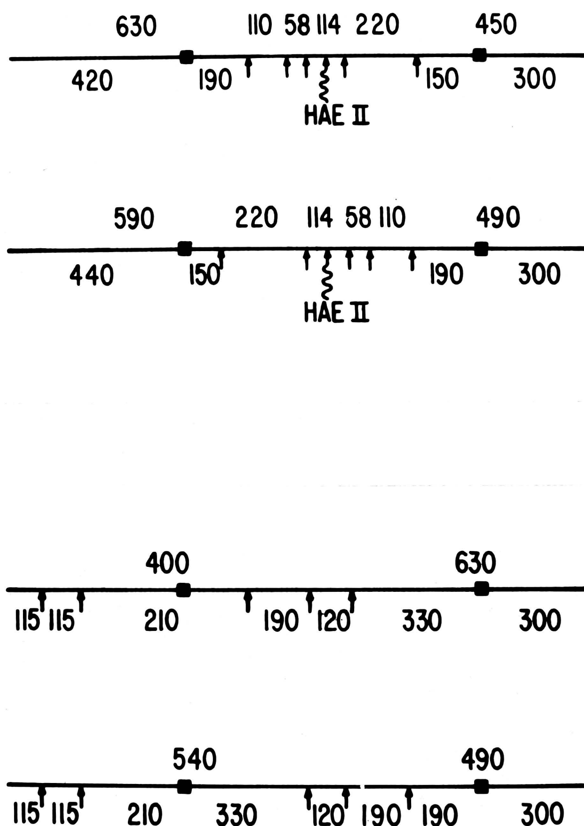


FIGURE 7. Restriction endonuclease site maps of pJZ110. In the top two drawings, the Hae II and Hae III endonuclease sites are indicated for the pJZ110 fragment with the inversion region oriented in both directions. The Hae III sites are indicated by the straight arrows and the Hae II site as the curly arrow. The lower two drawings indicate the positions of Hpa II endonuclease sites on pJZ110 in both configurations.

were observed at frequencies very similar to those observed with DNA prepared from Rec^+ strains. Thus, the inversion event proceeds via some enzymatic system other than the RecA system of *E. coli*. It is possible that the inversion region codes for an enzyme that catalyzes the inversion process. The region has sufficient coding capacity for a polypeptide of molecule weight 24,000-30,000. There may also be host enzyme systems other than the RecA system which catalyze the inversion process.

VARIOUS CONSTRUCTIONS INVOLVING *H2*

If the ends of the inversion regions have sites necessary for inversion, it should be possible to obtain mutations at these sites which would prevent change in orientation of the region. In order to select for these "fixed" mutants, a phenotype which can be easily scored would be desirable. It is to this end that we have constructed a number of recombinant molecules in which different genes appear to be under the control of the phase variation switch. In one such construction, illustrated in Fig. 8, the phase variation switch has been inserted adjacent to the *trpB* gene from *Salmonella*. This hybrid molecule (pJZ200) transformed into an *E. coli* strain carrying a *trpB-E* deletion allowed the transformed cells to grow on minimal medium supplied with indole, using the *trpB* function on the hybrid plasmid. Furthermore, the indole utilizing cells appeared to be undergoing phase variation to the Indole$^-$ phenotype while maintaining the antibiotic resistance of the plasmid vehicle. Thus, it appears that the *trpB* function is being controlled by the phase variation switch and from a promoter activity within the H2 gene fragment. It is our hope that the Trp$^+$ phenotype will provide a stringent selection through which mutants fixed in the ability to utilize indole may be found.

The pJZ200 hybrid plasmid also suggested the direction of the H2 gene transcription. A similar construction was made in which the phase variation switch was inserted adjacent to the *cheW* gene from *E. coli* on a hybrid lambda vehicle in the same configuration as in pJZ200 (SILVERMAN, unpublished observations). This hybrid confirmed the direction of transcription as being from the inversion region toward the arm carrying the H2 gene. By examining the restriction products of lambda vehicles carrying these genes on polyacrylamide gels, it was shown that when the H2 gene was "on", H2 flagellin and the *cheW* gene products were produced. When the H2 gene was "off", neither H2 flagellin nor the *cheW* gene product was synthesized.

These constructions also imply that the fragment carrying the inversion region contains the promoter activity for transcription of the H2 gene and genes downstream. Recent experiments in which specific substitutions of *gal* DNA into parts of the inversion region were constructed, suggest that the promoter lies within the inversion itself (SILVERMAN, unpublished observations).

CONCLUSIONS

Genetic studies in a number of systems have suggested that the expression of genes may be regulated through rearrangements of the genetic material. The phase variation system of *Salmonella* offers an opportunity to study the nature of these types of regulatory mechanisms at the molecule level. A site-specific inversion of a region of DNA 800 bp in length appears to be responsible for phase transition. This inversion process is RecA independent, yet may proceed through some other host enzymatic system or by a product coded by the inversion region itself. It appears that the promoter that initiates transcription of the H2 gene is included in the inversion region. When the region is oriented in one direction, this promoter is coupled to the H2

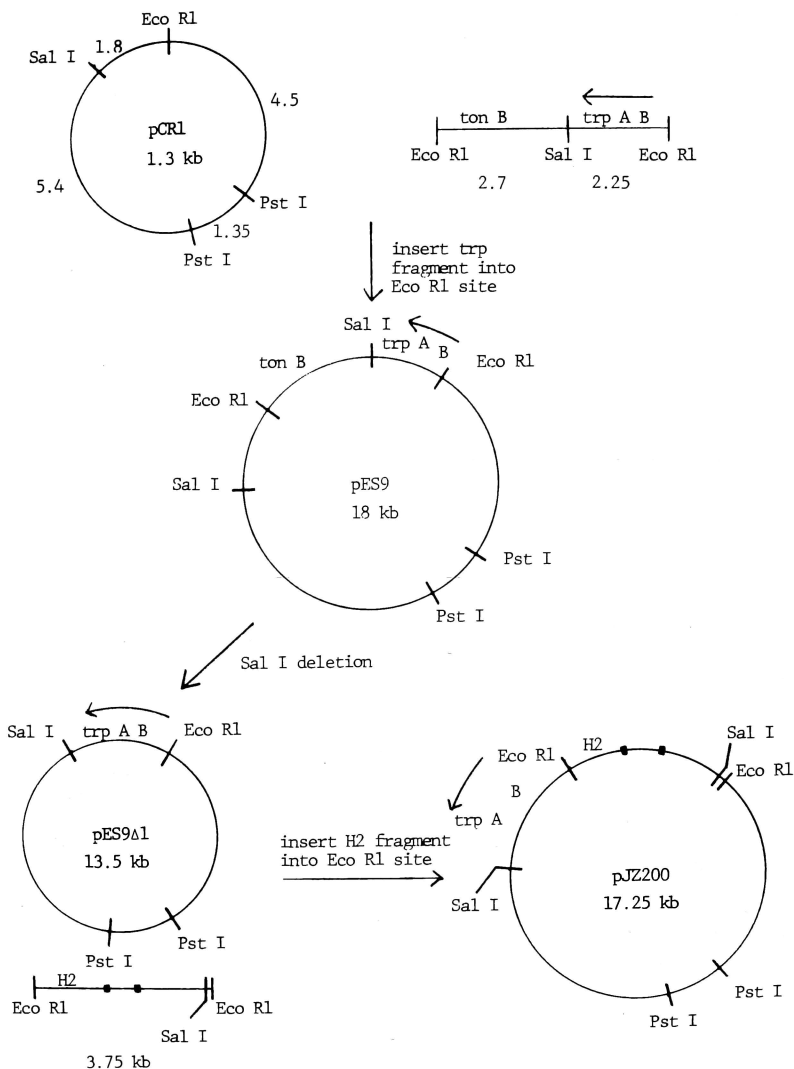


FIGURE 8. Construction of the TrpB-H2 hybrid plasmid. Plasmid vehicle pCR1 was used as a recipient for an Eco RI fragment carrying the TonB and TrpA and B genes from *Salmonella*. This hybrid plasmid, pES9, carries the Kan^R determinant of the plasmid vehicle. The Sal I site inside the inserted Eco RI fragment was used to delete the Sal I fragment forming plasmid *pES9Δ1* which has a single Eco RI site. This plasmid was then used as a recipient for the 3.75 kb Eco RI fragment from pJZ60 carrying the inversion region (defined by the small boxes) and the H2 structural gene. The hybrid plasmid, pJZ200, is 17.25 kb in length and expresses both the H2 gene and Trp B genes from the phase variation control region.

structural gene; while in the opposite orientation, they are uncoupled and the H2 gene is not transcribed.

ACKNOWLEDGEMENTS

This work was supported by grants from the National Science Foundation (PCM 76-17197) and the National Institutes of Health (USPHS AI 13008-02).

LITERATURE CITED

CHOW, L. and A. BUKHARI 1977 Bacteriophage Mu Genome: Structural Studies on Mu DNA and Mu Mutants Carrying Insertions. Pp. 295-306. DNA Insertion Elements, Plasmids, and Episomes. (Bukhari, A.I., J.A. Shapiro and S.L. Adhya, Eds.) Cold Spring Harbor.

DAVIS, R., M. SIMON and N. DAVIDSON 1971 Electron Microscope Heteroduplex Methods for Mapping Regions of Base Sequence Homology in Nucleic Acids. Pp. 413-428. Methods of Enzymology XXI.

ENOMOTO, M. and B.A.D. STOCKER 1975 Integration at *hag* or elsewhere of H2 genes transduced from *Salmonella* to *Escherichia coli*. Genetics 81:595-614.

FUGITA, H., S. YAMAGUCHI and T. IINO 1973 Studies on H-O variants in *Salmonella* in relation to phase variation. J. Gen. Microbiol. 76:127-134.

GREEN, M.M. 1977 The Case for DNA Insertion Mutations in *Drosophila*. Pp. 437-445. DNA Insertion Elements, Plasmids, and Episomes. (Bukhari, A.I., J.A. Shapiro and S.L. Adhya, Eds.) Cold Spring Harbor.

GREEN, M.M. 1978 The Genetic Control of Mutation in *Drosophila*. Stadler Genetics Symposia. Vol. 10 (G. Rédei, Ed.) University of Missouri.

HERSHFIELD, V., H. BOYER, C. YANOFSKY, M. LOVETT and D. HELINSKI 1974 Plasmid Col El as a molecular vehicle for cloning and application of DNA. Proc. Natl. Acad. Sci. USA 71: 3455-3459.

HICKS, J., J. STRATHERN and I. HERSKOWITZ 1977 The Cassette Model of Mating-Type Interconversion. Pp. 457-462. DNA Insertion Elements, Plasmids, and Episomes. (Bukhari, A.I., J.A. Shapiro and S.L. Adhya, Eds.) Cold Spring Harbor.

KONDOH, N. and H. HOTANI 1974 Flagellin from *E. coli* K-12: Polymerization and molecular weight in comparison with *Salmonella* flagellins. Biochem. Biophys. Acta 336:117-

LEDERBERG, J. and P. EDWARDS 1953 Serotypic recombination in *Salmonella*. J. Immunol. 71:232-240.

LEDERBERG, J. and T. IINO 1956 Phase variation in *Salmonella*. Genetics 41:743-757.

McCLINTOCK, B. 1957 Controlling elements and the gene. Cold Spring Harbor Symp. Quant. Biol. 21:197-216.

McCLINTOCK, B. 1978 Mechanisms that Rapidly Reorganize the Maize Genome. Stadler Genetics Symposia. Vol. 10 (G. Redei, Ed.) University of Missouri.

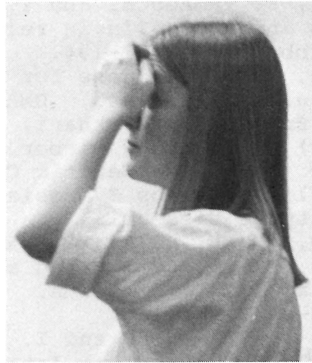
McDONOUGH, M. 1965 Amino acid composition of antigenically distinct *Salmonella* flagellar proteins. J. Mol. Biol. 12:342-355.

SAEDLER, H. and P. STARLINGER 1967 O° mutations in the galactose operon in *E. coli*. Mol. Gen. Genet. 100:178-89.

TONEGAWA, S., C. BRACK, N. HOZUMI and R. SCHULLER 1977 Cloning of an immunoglobulin variable region gene from mouse embryo. Proc. Natl. Acad. Sci. USA 74:3518-3522.

ZIEG, J., M. SILVERMAN, M. HILMEN and M. SIMON 1977a Recombinational switch for gene expression. Science 196:170-172.

ZIEG, J., M. SILVERMAN, M. HILMEN and M. SIMON 1977b The mechanism of phase variation. Pp. 25-35. Molecular Approaches to Eukaryotic Genetic Systems. (Wilcox, Abelson, Fox, Eds.) ICN-UCLA.



Janine Zieg