1988-02-01

A Genetic and Structural Analysis of P22 Lysozyme: A Thesis

Dale Rennell
University of Massachusetts Medical School

A GENETIC AND STRUCTURAL ANALYSIS OF P22 LYSOZYME.

A Thesis Presented

by

Dale Rennell

Submitted to the Faculty of the
University of Massachusetts Medical School in partial
fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY IN MEDICAL SCIENCES

February 1988

Molecular Genetics and Microbiology

A GENETIC AND STRUCTURAL ANALYSIS OF P22 LYSOZYME.

A Thesis Presented

by

Dale Rennell

Approved as to style and content by:

Michael Volkert, Chairman of Committee

Duanne Jenness, Member

Paul Dobner, Member

Martin Marinus, Member

Robert Sauer, Member

Thomas Miller, Dean Graduate School

ANTHONY POTEETE, ADVISOR

February 1988

MOLECULAR GENETICS AND MICROBIOLOGY

ACKNOWLEDGEMENTS

I would like to thank Tony Poteete who has patiently taught and guided me. I would also like to acknowledge the help and friendship of my co-workers.

I also would like to thank my husband, Paul and my mother, Doris Rennell for the love and encouragement that they gave me during the recent years.

ABSTRACT

GENETIC AND STRUCTURAL ANAYLYSIS OF P22 LYSOZYME

P22 lysozyme, encoded by gene *19*,is an essential phage protein responsible for hydrolyzing the bacterial cell wall during lytic infection. P22 lysozyme is related to T4 lysozyme in its mode of action, substrate specificities, and in its structure. Gene *19* was located on the phage genome, subcloned, and then sequenced. Lysozyme was produced in large quantities and purified for biochemical characterization and for crystallograpic studies. Gene *19* consists of 146 codons, and encodes a protein with a molecular weight of 16,117.

Amber mutations were created in gene *19* by in vitro primer-directed mutagenesis. The mutations were crossed by homologous recombination onto the phage genome. The phages bearing the amber mutations in gene *19* were screened for the ability to grow on six different amber suppressor strains. Amino acid substitutions that resulted in nonfunctional or less functional lysozyme were determined. Of 60 possible amino acid substitutions at 11 different sites in P22 lysozyme, 20 are deleterious. The phage bearing amber mutations in gene *19* that failed to grow on given suppressor strains were reverted and second site intragenic revertants were obtained. The mutations were sequenced.

A substitution of serine for glutamine at residue 82 is

compensated for by changing residue 46 from serine to leucine. This single change enables the phage to form a plaque at 30^0C but not at 40^0C. When the triple change asn42->lys; ser46->leu; and ser43->pro is present the lysozyme produced is no longer temperature sensitive. The crystal structure of P22 lysozyme is not yet solved. Assuming that the structures of T4 lysozyme and P22 lysozyme are similar, one can examine the positions of equivalent residues in the T4 lysozyme structure. The spatial arrangement of the residues changed by the secondary site mutations and the original substitution can then be visualized. The mutations discussed above all map far from the original mutation on the T4 three dimensional model.

A substitution of leucine for tyrosine at position 22 is compensated for by the double mutation of arg18->ser and ser23->lys. When the equivalent residues are mapped on the T4 three dimensional model the changes map in close proximity to the original mutation.

TABLE OF CONTENTS

LIST OF TABLES

TABLE

LIST OF FIGURES

FIGURE

STRAIN LIST

Strain	Pertinent Genotype	Source
W3110	*E.coli*	Lederberg, 1960
W3110 *lacI*^{q}	*E.coli lacI*^{q} L8	Brent and Ptahne 1981
C600	*E.coli* Su$^+$	
DB7000	*Salmonella typ* LT2 *leu*Aam414 Sup0	Miriam Susskind
MS1363	*Salmonella typ* LT2 *leu*Aam414 supE	Miriam Susskind
MS1868	*Salmonella typ* LT2 *leu*Aam414 r-m+ sup0	Miriam Susskind
GM1675	*dam*-4 del(*lac-pro*) *thi*-1 *sup*E *rel*A1 /*F'*lac*I*^{q} del(M15) *pro*+	Martin Marinus
MS1362	*Salmonella typ* LT2 *leu*Aam414 *sup*D	Miriam Susskind
MS1363	*Salmonella typ* LT2 *leu*Aam414 *sup*E	Miriam Susskind
MS1364	*Salmonella typ* LT2 *leu*Aam414 *sup*F	Miriam Susskind
MS1365	*Salmonella typ* LT2 *leu*Aam414 *sup*G	Miriam Susskind
MS2310	MS1868 bearing the plasmid pKM101 amp^r	Miriam Susskind
MS1387	*Salmonella typ* LT2 *sup*Q *pro* (del*attP22*) *cys*B *his leu*D *fol*101 /F'*lac pro*	Miriam Susskind
CV112	*Salmonella typ* LT2 *pol*A*ts*	Van Beveran, Ph.D. thesis, Tufts University, 1985 was provided by A. Wright
TP278	MS2310 with *sup*E	this study
TP279	MS2310 with *sup*F	this study
TP280	MS2310 with *sup*G	this study
TP282	MS2310 with *sup*D	this study

Strain	Pertinent Genotype	Source
TP284	MS1868 lysogenized P22 *sieA*44 *Ap*2 del 7283	this study
TP292	W3110 lysogenized with P22 Kn321 *sieA*44 *m*44	this study
TP241	MS1363 lysogenized with P22 Kn321 *sieA*44 *m*44.	this study
TP308	MS2310 bearing pDR463 (su^+ phe)	this study
TP309	MS2310 bearing pDR464 (su^+ cys)	this study
TP320	MS1868 bearing pDR105 (T4 lysozyme producer)	this study
TP315	W3110 lysogenized with P22 *sieA*44 *m44* Kn469	this study

CHAPTER I

INTRODUCTION

CHAPTER I

INTRODUCTION

The Question of How a Protein's Structure is Determined

An intriguing question for biologists for decades has been how a protein assumes its tertiary structure. In the 1930's, Corey and Pauling started x-ray crystallographic studies of amino acids and polypeptides to obtain a set of standard bond distances and angles for amino acids and small peptides. By 1951, before a complete amino acid sequence had been determined for any protein, they were able to predict that given the conformations that peptide bond-linked amino acids could assume, alpha helixes and beta-pleated sheets could be structural elements of proteins. In 1953, Fredrick Sanger determined the amino acid sequence of insulin. This was the first demonstration that a protein has a precisely defined amino acid sequence. Once amino acid sequences were elucidated, it was hoped that the structures of proteins could be predicted. The famous studies of Haber and Anfinsen on the denaturation and renaturation of bovine ribonuclease showed conclusively that all the information needed to obtain the proper structure was contained in the primary amino acid sequence. This laid to rest a hypothesis that some type of biological template determined the structure that a protein would assume. This discovery did not simplify the task of

predicting structures though, since in a simple protein of just 166 amino acids there is estimated to be 1×10^{79} possible conformations if each residue could assume three different conformations (Levinthal, 1968). Since proteins have precisely determined amino acid sequences, and the unique structures of every protein must be related to the amino acid sequence, the question became how amino acid sequences dictate structures.

Background. In 1964, Davies tried to correlate the amino acid composition of 15 proteins and the alpha helix content of each protein to predict how amino acid composition affected alpha helix content. This was done at a time when the crystal structure and amino acid sequence of only one protein was known: sperm whale myoglobin. The helix content for the rest of the proteins was determined by their optical properties in aqueous solution. By statistical methods, he was able to correlate low helix content with high mole percentage of the following amino acids: serine, threonine, valine, isoleucine, and cysteine. He determined that when these amino acids constituted greater than 25% of the protein's amino acids they somehow destabilized helixes (Davies, 1964). Many statistical predictive methods for determining a protein's secondary structure from amino acid sequences have been published (Chou and Fasman, 1974a,b;

Garnier et al., 1978; Lim, 1974a,b; Robson and Pain 1974a,b; Kabsch and Sander,1984; Cohen et al., 1983). These algorithms have been formulated based on the availability of protein sequences and the co-ordinates of crystal structures. Even today, the amount of knowledge about the crystal structure of proteins is limited by the number of structures that have been solved. The Protein Information Resource Data Bank contains about 6000 protein sequences but the Brookhaven data bank contains co-ordinates of only 300 different protein structures of which only about 100 are for nonhomologous proteins (Blundell et al., 1987, Claverie and Bricault, 1986). Even though the predictive accuracy of these algorithms has increased, we still do not know how particular secondary structures will align themselves in relation to one another to obtain the tertiary structure. To date a tertiary structure of a protein has not been accurately predicted (Ponders and Richards, 1987). Attempts have been made (Levitt and Chothia, 1976) to categorize groupings of secondary structures into 4 classes. The authors used schematic two-dimensional representations termed topology/packing diagrams. The proteins fall into four classes, which are designated; (I) all-helix, (II)all-beta, (III) helix + beta (contains helixes and beta structures that do not mix but tend to segregate along the polypeptide chain), (IV) helix/beta (contains mixed or approximately alternating segments of helical and beta-strand

secondary structure).

Folding Pathways. Some methods for predicting protein structure are based on thermodynamics. That a protein folds by a limited number of pathways rather than searching randomly through all possible conformations was recognized early on since the amount of time that a random trial and error method would take is inconsistent with observed time of synthesis of native proteins (Anfinsen, 1973; Levinthal, 1968; Wetlaufer, 1973). Two models dominate the field today. The first is random search nucleation and chain propagation. This hypothesis states that nucleation centers in an unfolded polypeptide chain would initiate the folding process and direct it to a kinetically favorable pathway (Wetlaufer,1973; Lewis et al., 1971; Creighton, 1974). The nucleation center is postulated to be between 8 and 18 residues in length. The difficulty with this model is once the nucleation center is formed the rest of the residues in the polypeptide chain must search randomly to sample which conformation is most thermodynamically stable. The second working model of protein folding is the diffusion collision theory of Karplus and Weaver (1979). This states that local native-like structures form "microdomains" which diffuse and collide with other microdomains and coalesce into the native structure. The authors do not feel that there is any given order to

coalescence of the microdomains. Levitt and Chothia (1976) found that secondary structural elements represented in their topology/packing diagrams that are adjacent along the polypeptide chain are also in three dimensional contact more often than secondary elements that were not adjacent along the polypeptide chain. This would seem logical given the diffusion and collision model since the microdomains formed from adjacent residues would be closest.

Harrison and Durbin (1985) make an analogy between protein folding pathways and assembly of a jigsaw puzzle. They feel that the diffusion, collision theory fits most experimental data and that there is no one discrete pathway to a final structure. Sometimes a local secondary structure will form first while at other times another will. If incorrect structures coalesce with one another then the complex will fall apart due to lack of stabilization by other secondary structures. They feel it is this internal editing rather than a pathway defined by the amino acid sequence, that makes folding nonrandom.

Protein Mutants. An important prediction of Harrison and Durbin is that mutants kinetically blocked in any given pathway will have alternative pathways available so it will be unlikely to find any protein folding mutations. Apparent contradictions to this prediction would be mutations

isolated as *tss* (temperature sensitive for synthesis). Temperature-sensitive mutant proteins when formed at permissive temperature and then exposed to nonpermissive temperatures behave normally (Streisinger et al., 1961). It was thought that such mutants represented folding mutations, since it was only during synthesis that high temperature was detrimental. There are many examples of *tss* mutants (eg. *E. coli lac*I, Sadler and Novick, 1965; P22 tail protein, King, 1987). Harrison and Durbin note that *tss* mutations have been found only in proteins that undergo oligomerization or assembly into a functional unit, and speculate that these mutations are not folding mutations but assembly mutations. If their model is correct then mutations that kinetically block folding will not be isolated.

The study of protein mutants has yielded much information on the chemistry of active sites of enzymes. Yet information about how the three dimensional structures form and are maintained has been harder to obtain. Mutant proteins, particularly those that exhibit a change in stability are being studied to determine how final structures of proteins are stabilized. This might be the only route to decipher how amino acid sequences dictate structures if, as Durbin and Harrison propose, true folding mutations will not be found. Matthews and colleagues (1987) are examining mutations that affect the stability of protein structure. For instance, they

reasoned that a decrease in configurational entropy of the unfolded state would stabilize the protein toward reversible and irreversible thermodenaturation. They reasoned that it takes more free energy to hold a glycine residue in a given conformation than it does to maintain an alanine residue since glycine lacks a *beta*-carbon and has more backbone conformational flexibility than alanine. By replacing glycine 77 with an alanine they were able to enhance T4 lysozyme's thermostability. Other investigators are attempting to increase protein stability by adding disulfide bonds. This has been tried with T4 lysozyme (Perry and Wetzel, 1984), dihydrofolate reductase (Villafranca et al., 1983), ribonuclease A (Lin et al., 1984), chicken lysozyme (Udea et al., 1985), and bacterial protease subtilisin BPN'(Pantoliano et al., 1987). Mutations such as these which enhance protein stability do not really address the question of protein folding since these mutations result in a larger free energy difference between the folded and unfolded protein and therefore a more stable protein would be expected to result unless the new disulfide bonds interfered in some other interaction in the final structure.

Nearly 100 mutant human hemoglobins have been characterized. The mutants were examined in conjunction with an atomic model of hemoglobin based on x-ray analysis of horse hemoglobin (Perutz and Lehmann, 1968). It was found that the

molecule was extremely sensitive to changes of internal non-polar contacts, especially those near the hemes. The molecule was relatively insensitive to replacements of most amino acid residues on its surface. Studies such as these allowed one to understand some of the interactions that take place in the tertiary structures.

T4 lysozyme mutants have also been studied in detail (Alber et al., 1987a,b; Matthews et al., 1987; Streisinger et al, 1961) with the crystal structures solved for a large number of the mutant proteins. Information obtained from these studies has allowed investigators to formulate some generalities. For instance in studying temperature sensitive mutations in T4 lysozyme the investigators determined that most of the mutations that affected the stability of the protein were at sites with low mobility and low solvent accessibility in the folded protein (Alber et al., 1987).

Another example of mutations whose primary defects are destabilization of the tertiary structures are Cro repressor mutants that have been recently isolated (Pakula et al., 1986). Some of the cro repressors that had decreased stability had nearly normal specific activity. The half life of the molecules was drastically reduced. These were cases where a protein was able to fold but the equilibrium between the folded state and the unfolded state was changed so that the protein spent more time in the unfolded state. Most of these mutations affected residues in the tightly packed protein

interior.

Second Site Revertants. The question of how amino acid sequences determine structures has proved to be hard to approach by genetics. Part of the problem might be that the mutant proteins studied have been only slightly defective. By this I mean that if the protein had a severe mutation that prevented it from folding into a proper structure then that protein would probably never be isolated since denatured proteins are often attacked by proteases and degraded rapidly. The use of second site revertants might overcome this difficulty. Using second site revertants for studying proteins is not a new idea. Helinski and Yanofsky described studies of second site revertants of the A protein of *E. coli* tryptophan synthetase in 1963. Using genetic analysis and peptide mapping they were able to determine that a second site intragenic change of a tyrosine to a cysteine would compensate for an original mutation that substituted glutamic acid for glycine. This second change when isolated from the original mutation yielded a functionally ineffective A protein. It was also an allele specific suppressor since it failed to suppress several other A protein mutations, including a change of glycine to arginine at the site of the original mutation. Since it was known that several amino acids could substitute at the position of the original

mutation with no loss of function, it is thought that the mutation probably affected the protein's three dimensional structure and did not change the residues in the active site because substitutions are not often found in an enzyme's catalytic site.

Second site revertants have also been isolated in Staphylococcal nuclease A (Shortle and Lin, 1985). They have characterized three intragenic "global suppressors" of nuclease-minus mutations. A global suppressor is a mutation which is capable of suppressing the nuclease-minus phenotype of different alleles. It was postulated that these mutations increased the stability of the protein, thus overcoming a decrease in protein stability introduced by the original *nuc*- mutations.

Double mutants were also utilized by Carter et al., (1984) to determine the structural effect of a primary mutation that increased the affinity of tyrosyl tRNA synthetase for ATP. They were able to determine that the original mutation acted by distorting the polypeptide backbone, thus altering the interaction of the side chain of another residue with the substrate ATP.

Second site revertants have also been used to investigate the interactions between proteins, and between proteins and DNA (Jarvik and Botstein, 1973; Hochschild et al.,1983; Youderian et al., 1983; Gardella, pers. comm.).

The use of second site reversion could allow for the characterization of protein mutants that might not be obtainable by other means. For instance, if a primary mutation created a protein that was insoluble or degraded that molecule would be unavailable for detailed analysis, but if a second site mutation, which compensated for the original defect, allowed a protein to fold it could be characterized. Then by examining the double mutant one might be able to deduce the defect of the single mutant. Second site revertants also allow one to determine which amino acids within the protein interact with one another either directly or indirectly.

There are several technical problems associated with the isolation of large numbers of second site revertants however: first, there is the task of distinguishing second site revertants from primary site revertants; second, if one wishes to study interactions that take place within the protein one must be sure that the second site revertants are intragenic; and third, since second site revertants might be fairly rare you need to have a strong selection for reversion events.

Scheme for isolating second site revertants of missuppressed amber mutations in P22 lysozyme. Bacteriophage P22 lysozyme presented a nice system in which to isolate second site revertants. First, P22 is a genetically pliable

organism; it can undergo homologous recombination at high frequency, it can effect the transduction of plasmid or bacterial genomes, and it can be grown lysogenically or lytically. Second, P22 lysozyme has several useful properties for this type of study. It is a small protein with a molecular weight of about 16,000 but large enough to have a two domain structure. The gene encoding P22 lysozyme consists of only 438 bases so it is easy to sequence mutations. It is monomeric, and apparently does not interact with other proteins. This makes it unlikely that second site revertants will be in other genes. P22 lysozyme is easily assayed in vitro, permitting quantitative studies of mutant proteins. It is an essential P22 gene for lytic growth, but phage lacking lysozyme function can be propagated as prophage and recovered after induction by lysing the cells artificially, or by complementing the phage with another source of lysozyme. The conditional essentiality of the gene makes selection of revertants simple.

The approach I took was to create primary mutations in P22 lysozyme and then isolate second site revertants that compensated for the original mutations. The primary mutations were missuppressed amber mutations. I use the term "missuppressed" to designate amber mutations that, when placed in a strain containing a particular amber suppressor, yield a nonfunctional or less active protein. The amber mutations were created by oligonucleotide-directed mutagenesis by using

a plasmid-borne lysozyme gene as template. When the amber mutations had been identified on the plasmid, the mutant genes were transferred to P22 by homologous recombination, this step was facilitated by setting up a strong selection for the recombinant phage. The resultant phage bearing amber mutations were plated on six amber suppressor strains to determine the suppression patterns. Each amber mutation thus allowed me to determine the affect of five or six single amino acid substitutions. Missuppressed ambers, that is phage bearing amber mutations, when grown on particular suppressor strains produce nonfunctional lysozymes, were reverted. The revertants were screened to determine if they were primary or secondary site revertants. This screen is a simple one due to the nature of the primary mutation. If an amber mutation undergoes any change it is no longer an amber, therefore any revertant that is capable of growing on a suppressor-minus strain, by definition is a primary site revertant. The mutations present in the secondary site revertants were sequenced.

CHAPTER II

PHAGE P22 LYSIS GENES: NUCLEOTIDE SEQUENCES AND FUNCTIONAL RELATIONSHIPS WITH T4 AND LAMBDA GENES

This chapter was published as a paper (Rennell, Dale and Poteete, A.R. (1985). Phage P22 Lysis Genes: Nucleotide Sequences and Functional Relationships with T4 and Lambda Genes. *Virology* **143**, 280-289). A.R. Poteete constructed pTP163 and pTP82 and observed that these plasmids would complement a lysozyme minus P22 phage. I carried out all the other work described in this and subsequent chapters.

CHAPTER II

PHAGE P22 LYSIS GENES: NUCLEOTIDE SEQUENCES AND FUNCTIONAL RELATIONSHIPS WITH T4 AND LAMBDA GENES

Introduction

Salmonella phage P22 has two genes that are essential for lysis of the bacterial cell wall, genes *13* and *19* (Botstein et al., 1972). P22 gene *13* is similar to Lambda gene *S*. P22 *13-* shows the same phenotype as Lambda *S-*: failure to lyse, production of viable phage for many hours past normal lysis time, and continued production of lysozyme which can be released by artificially disrupting the cells. Gene *13* and gene *S* map in analogous positions in their respective phage genomes, between genes involved in late regulation and DNA encapsulation. Presumably 13 protein acts like S protein, by disrupting the cell membrane, thereby giving hydrolytic enzymes access to the cell wall (Reader and Siminovitch, 1971). This type of activity is also characteristic of the product of T4 gene, *t* (Josslin, 1970).

P22 gene *19* encodes a protein that has cell wall hydrolyzing activity. A *19-* phage fails to lyse the host cell but produces viable phage. A prominent cell wall hydrolyzing activity present in cells infected with P22 wild type is missing or altered in cells infected with phage bearing mutations in gene *19* (Botstein et al., 1972). Rao and Burma (1971) studied this activity and found that it has substrate specificities like those of T4 lysozyme. They concluded that the P22 enzyme is a lysozyme. T4 lysozyme, the product of

gene *e*, works by hydrolyzing Beta-1,4 glycosidic bonds between alternating units of N-acetylglucosamine and N-acetyl muramic acid in the peptidoglycan (Tsugita et al., 1968). Lambda has two proteins responsible for the cell wall hydrolyzing activity found in phage-infected cells: a transglycosylase, the product of gene *R*; and an endopeptidase, the product of gene *RZ*. The *R* and *Rz* proteins together are known as lambda endolysin (Bienkowska-Szewczyk and Taylor, 1980).

Phage lysozymes are particularly accessible to study by physical and genetic techniques because of two characteristics. First, lysozymes are small proteins with an easily assayable activity. Second, it is easy to obtain and characterize mutant phage with altered lysozymes. These advantages have been exploited in the study of T4 lysozyme. Comparative studies of the three dimensional structures of wild type and mutant T4 lysozymes have considerably enhanced our understanding of lysozyme catalysis (Grutter and Matthews, 1982).

Studies comparing T4 lysozyme with other lysozymes have been informative. For example, it has been found that the three dimensional structures of lysozymes are highly conserved, but the amino acid sequences are extremely variable. Even though T4 lysozyme and hen egg white lysozyme (HEWL) have non-homologous sequences, key regions of their three dimensional structures are superimposible (Rossman and Argos, 1976; Levitt and Chothia, 1976; Remington and Matthews,

1978). The similar activities of P22 and T4 lysozyme encourage further studies comparing the two phage lysozymes. Our assumption, confirmed by results described below, was that the genetic non-relatedness of the two phages would preclude a high degree of sequence homology between the T4 and P22 lysozymes. Hence we are dealing with proteins of dissimilar primary structure that are closely related in function.

We present in this paper the results of experiments that locate gene *13* and gene *19* on the chromosome of P22 and determine their DNA sequences. In addition, we show that plasmids that express T4 lysozyme, P22 lysozyme or lambda endolysin can complement both P22 *19*- and lambda *R*-. Thus these functions are in a sense interchangeable.

Material and Methods

Media, buffers, and enzymes. LB broth, LB agar, M9CAA, soft agar, and DNA buffer were as described previously (Poteete, 1982) RB is 50 mM TrisHCl pH 7.2, 100 mM NaCl, 1 mM EDTA, 3 mM 2-mercaptoethanol, and 10 percent glycerol (v/v). TN25 is 50 mM TrisHCl pH 7.5, 1 mM EDTA, 1.4 mM 2-mercaptoethanol, 5 percent glycerol (v/v), and 25 mM NaCl.

Restriction endonucleases, T4 DNA ligase, T4 polynucleotide kinase and *E.coli* DNA polymerase I large fragment were purchased from New England Biolabs. Polyethylenimine was purchased from Sigma. Ammonium sulfate

(special enzyme grade) was purchased from Schwarz Mann/Inc.

Strains. *E. coli* strain W3110 *lac*I^q L8 (Brent and Ptashne, 1981) was generally used for propagating plasmids and for infections by lambda. C600 was the Su+ *E. coli* strain used. *Salmonella* strains DB7000 (*leu*Aam414 Sup0), MS1363 (*leu*Aam414 supE), and MS1868 (*leu*Aam414 r-m+ sup0) are derivatives of *Salmonella typhimurium* LT2.

P22 *13*-amH101 *c*1-7 was the source of P22 DNA for most plasmid constructions. P22 *19*-amber strains were *19*-amN111 *c*2, *19*-amH447 *c*1-7, *19*-amH1162 *c*1-7, *19*-am1164 *c*1-7, *19*-amH1239 *c*1-7, *19*-amH1365 *c*1-7, and *19*-am1026 *c*1-7. Lambda strains used were lambda *R*-am5, lambda *c*I-857 *S*-am7 and lambda KH54 (*c*I del). T4 DNA for plasmid constructions was from lambda-EIII e+ hybrid phage (Owen et al., 1983).

Plasmids. Descriptions of the plasmids constructed for these studies are in Table 1 and Fig. 1. General methods used for plasmid construction were as described previously (Poteete and Roberts, 1981). *E. coli* and *Salmonella* transformations and plasmid DNA preparations were carried out as described previously (Poteete, 1982).

DNA sequencing. Methods for DNA sequencing were as described by Maxam and Gilbert (1980). The plasmids pTP163, pDR102, pDR103, pDR100 and pDR101 were used as sources of

fragments for sequencing wild type gene *19* and part of gene *13*. The wild type sequence for gene *13* was obtained from DNA fragments isolated from pDR117. Complete sequences of both strands of DNA for gene *13* and *19* were determined. All segments of sequence were ordered by direct overlap (Fig. 2). Partial sequences of one strand of each mutant were determined. The DNA fragments for the mutant sequences came from plasmids pDR110, pDR113, and pDR114. The fragments were purified using the method of Hansen (1981).

Lysozyme assays. Lysozyme activity was determined by decrease in turbidity of $CHCl_3$-treated *E. coli* measured by absorbance at 470 nm as described by Rao and Burma (1971). The lysozyme purification was also monitored by discontinuous SDS-polyacrylamide gel (13 percent) electrophoresis (Laemmli, 1970). Following electrophoresis, gels were stained with Coomassie blue.

Protein purification. P22 lysozyme was purified from MS1868 infected with P22 *13*-amH101 *c*1-7. The cells were grown in 2 liters of LB broth to 1 x 108 cells/ml, phage was added at a multiplicity of 10 and the culture was aerated at 30^0C for 3 hours. The cells were centrifuged at 6000 rpm for 10 minutes in a Sorvall GSA rotor. The cell pellet was resuspended in 30 ml RB then frozen at -20^0C overnight. The frozen cells were placed in a 37^0C water bath until thawed and

put on ice. All subsequent steps were carried out between 0^0 and 4^0C. The cells were sonicated with four bursts of 10 seconds duration. It was determined by microscopy that the suspension had less than 1 percent of the original cell count that failed to lyse. The lysate was centrifuged at 40,000 rpm for 90 minutes in a Beckman 70.1 Ti rotor; 20 ml of RB was added to the supernatant, and polyethyleieimine was added to a final concentration of 0.6 percent. The mixture was stirred at 4^0C for 45 minutes and then centrifuged at 7000 rpm for 30 minutes in a Sorvall SS 34 rotor. Most of the lysozyme activity was in the supernatant. Ammonium sulfate was added to 90 percent saturation and the mixture was stirred overnight at 4^oC. The precipitate was pelleted by centrifugation at 15,000 rpm for 90 minutes in a Sorvall SS 34 rotor. The pellet contained the majority of the lysozyme activity. The pH was monitored thoughout the preceding steps and was always in the range of 7.0 to 8.0. The pellet was redissolved with 2 ml TN25. The solution was applied to a 46 cm x 1.5 cm column of G-50 SF Sephadex which was then developed with TN25; 4 ml fractions were collected. The lysozyme activity was in one peak between the void and included volumes. Pooled peak fractions when examined by SDS polyacrylamide electrophoresis showed only four major bands. The sample was then applied to a 15 cm x 1.5 cm column of DEAE cellulose. The lysozyme was in the flow-through fraction and judged to be 90 percent pure by

SDS-polyacrylamide gel electrophoresis: one other minor band could be seen with Coomassie blue stain when 1.5 micrograms of protein were applied to the gel.

To prepare the sample for amino acid sequence analysis it was concentrated with a membrane concentrator (Millipore immersible CX-10) and then applied to a G-50 Sephadex column that was developed with 0.1 M ammonium bicarbonate. The peak containing the activity was then lyophilized. The yield was 0.04 mg of lysozyme as judged by optical density at 280 nm. (We estimate that a 1 mg/ml solution of P22 lysozyme would have an A280 of 0.837, based on its amino acid composition as revealed by the DNA sequence).

Marker Rescue Experiments. Gene *19* was located on the physical map of P22 by Rutila and Jackson (1981). To localize the gene further, we carried out marker rescue experiments with plasmids described in Table 1. Strain MS1868 bearing plasmids to be tested was grown to 2 x 10^8 cells/ml and plated with lambda soft agar on LB plates containing 20 µg/ml ampicillin. Dilutions of the *19*-am P22 phages were spotted and plaque formation was noted. To obtain more quantitative results plating efficiencies were determined (see Table 2). From the results in Table 2 gene *19* was located on a 1100 bp *Hpa*I restriction fragment (shown in Fig. 1). The *Hpa*I fragment can be divided into two nearly equal parts by cleavage at a *Nae*I site shown in Fig. 1. Plasmids were constructed that had

either the right or left half of the *Hpa*I fragment. Five *19*-amber alleles that were known to occupy different sites in the gene (unpublished results) were used in these marker rescue experiments. All the alleles could be rescued by plasmids containing the entire *Hpa*I fragment while two alleles could be rescued from plasmids containing the right half and the other three alleles could be rescued from plasmids containing the left half of the *Hpa*I fragment.

Results

Sequences of Genes *13* and *19*. The DNA sequence of a 1188 bp segment of the P22 chromosome known to contain gene *19* (see Methods) was determined. The sequence is shown in Fig. 3. There are only two open reading frames present in the sequence large enough to encode for gene *13* whose protein product has an estimated molecular weight of 10,500 (Youderian and Susskind, 1980) and for gene *19* whose protein product has an estimated molecular weight of 15,000 (Youderian and Susskind, 1980) to 20,000 (Rao and Burma, 1971). Translated rightward (with respect to the map in Fig. 1), the first large open reading frame is gene *13*. This gene starts with a single ATG and terminates with a single TAA. It consists of 108 codons and potentially encodes a protein with a molecular weight of 11,520. The last six sense codons of gene 13 overlap the first codons of gene 19 in a different reading frame. Gene *19*

starts with two tandem ATG codons and terminates with a single TGA. It consists of 146 codons and potentially encodes a protein with a molecular weight of 16,117.

The open reading frames depicted in the figure represent gene *13* and gene *19* according to the following observations. (1) Marker rescue experiments described in the Methods section show that gene *19* is contained within the region sequenced. Moreover, the gene straddles the *Nae*I site as predicted by the marker rescue experiments. (2) Amber mutants of gene *13* and gene *19* result in TAG codons in the open reading frames identified for the respective genes. We also sequenced the *h*21 mutation of gene *13*. This mutation is used as a morphological marker since it changes plaque morphology on green plates (Botstein et al.,1972). Youderian and Susskind (1981) showed that the protein product of gene *13-h*21 has an altered electrophorectic mobility on SDS polyacrylamide gels. The sequence shows the change to be a G to T in the third base of codon 57 of gene *13*. This change causes a substitution of cysteine for tryptophan (see Fig. 3). (3) The amino acid sequence of purified P22 lysozyme agrees with that predicted by the DNA sequence. Lysozyme was purified from phage-infected cells by procedures outlined in the Methods section. The amino acid sequence was determined for the first 33 residues of the amino terminal end. This amino acid sequence shows that the protein begins with a single methionine residue. The data concerning the rest of

the sequence (not shown) is consistent with that shown in Fig. 3.

The protein product of gene *13* is basic; it contains 13 basic residues (Lys, His, Arg) and 8 acidic residues (Asp, Glu). The predicted hydrophobicity of the protein was examined by the methods of Cantor and Schimmell (1980). The results of these calculations are called "Z" values. Z values of 0.52 ± .11 are indicative of an intrinsic membrane protein. The Z value of 13 protein is .4362 which suggests an intrinsic membrane protein.

The protein product of gene *19* has a molecular weight of 15,968 (the difference from the molecular weight predicted from the DNA sequence is due to the absence of the first methionine encoded in the DNA sequence from the mature protein). This is in agreement with the electrophoretic mobility of the purified protein in a 13 percent SDS polyacrylamide gel, where it migrates slightly behind HEWL (MW 14,600)(not shown). P22 lysozyme is a basic protein, containing 21 basic residues and 14 acidic residues. The basic character of the protein is apparent in its chromatographic behavior: it binds to carboxy methyl cellulose at pH 7.5 and flows through DEAE cellulose at pH 8.0 in low salt.

Comparison of Lysis Functions of P22, Lambda and T4.

Bacteriophages T4, lambda, and P22 all encode at least two

lysis functions. One of the functions (lambda S, T4 t and P22 *13*) is thought to act on the inner cell membrane. Our results suggest strongly that P22 *13* protein and lambda S protein work in the same way. The DNA sequences of *S* (Sanger et al., 1982) and *13* are 90 percent homologous. The amino acid sequences of the protein products predicted by the DNA sequences are 89 percent homologous (see Fig. 4). These observations strongly suggest that the proteins are virtually the same and would likely be interchangeable. No sequence comparisons can be made to T4 *t* since to our knowledge this gene has not yet been sequenced.

The phages' second lysis function acts on the cell wall. The lambda function is encoded by two genes, *R* and *RZ*. The T4 and P22 genes are *e* and *19* respectively. P22 lysozyme and T4 lysozyme have some amino acid sequence homology. This homology is particularly interesting since it is localized to regions of the protein near and in the active site of T4 lysozyme. The nature and implications of this homology are discussed by Weaver et al. (1984). We have been unable to detect any significant homology between P22 *19* and lambda R or RZ proteins. This is not suprising since R is a transglycosylase and RZ is an endopeptidase. It has been shown previously that T4 lysozyme and lambda transglycosylase differ in enzymatic activities as well as primary structure. (Bienkowska-Szewczyk and Taylor 1980; Black and Hogness, 1969; Tsugita and Inouye, 1968).

Complementation of Lysis defective phages. In infected cells the products of T4 e, lambda *R* and *RZ*, and P22 *19* require the activities of the membrane-disrupting proteins of the respective phages in order to gain access to the cell wall. We carried out an experiment to test the specificity of this requirement. Plasmids were constructed that express lambda *R* and *RZ*, P22 *19* and T4 e. P22 and T4 functions were expressed from plasmids that had similar structures, in which the genes were placed downstream from the P_{tac} promoter in a pBR322 vector. The lambda genes were expressed from lambda P_{late}. A fragment of lambda DNA containing the carboxy terminal region of gene *Q*, gene *S*, gene *R* and gene *RZ* was placed in a pBR322 vector. This plasmid was constructed using DNA from an *S*-am phage. Since the complementation tests were done in Suppressor minus cells only genes *R* and *RZ* are expressed. Expression from P_{late} requires the gene product of *Q* or the homologous protein from P22, the gene product of *23*, to be supplied in trans from the incoming P22 phage. Cells bearing these plasmids were tested to see if they would complement the growth defect of lambda *R*- or P22 *19*-. The results are given in Table 3. All the plasmids complemented both lambda *R*- and P22 *19*-. We conclude from this that the products of T4 e, lambda *R* - *RZ* and P22 *19* are interchangeable.

Discussion

Location and Sequence of the Lysis Genes . We have determined the sequences of P22 genes *13* and *19*, and thus established their precise locations within a small segment of the P22 chromosome. The lysis genes can also be mapped with respect to other P22 sequences by comparison of the sequence in Fig. 3 with that determined by Kroger and Hobom (personal communication). These investigators determined the DNA sequence of a 5944 base pair segment of the P22 chromosome, extending from gene 18 to the start of gene *13*, and overlapping our sequence by 78 base pairs. The sequences differ at one position. The base at position 77 is a C in our sequence and a T in theirs, implying His and Tyr, respectively, as the eighth amino acid residue in *13* protein. These sequencing results, as well as other cloning and marker rescue experiments (unpublished), lead to a map that differs from that proposed by Rutila and Jackson (1981). These authors reported that genes *13* and *19* map to the right of the rightmost *Pst*I site shown in Fig. 3, on the basis of marker rescue experiments. We do not know the reason for this discrepancy.

Mechanism of Action of Gene *13* Protein . We have shown that P22 gene *13* protein is highly homologous to lambda gene S protein. This finding in conjunction with previous

observations by others suggests that the proteins work in the same way, by disrupting the inner cell membrane.

Speculation on the type of interactions that the phage-encoded membrane-disrupting proteins have with their respective lysozymes and with the inner cell membrane leads one to two models. In one model membrane-disrupting proteins attach to the inner cell membrane and then by a specific interaction with the lysozyme allow the lysozyme access to the cell wall. This could be a receptor-type interaction with specific recognition taking place between the membrane-disrupting protein and the lysozyme. In the extreme form of this model, S protein would work only with R and RZ proteins, *13* protein would work only with *19* protein, and t protein would work only with *e* protein. We have shown this is not the case. However, in a less extreme form one could postulate that R, *19* and *e* proteins have in common some structures that S and *13* proteins recognize. These hypothetical common structures are not reflected in any obvious way in the primary structures of the proteins. A second model is that *13* and S proteins act by forming pores or channels in the inner cell membrane in a non-specific manner, creating an opening that excludes only on the basis of size. Our results and those of others (e.g., Wilson, 1982) are fully consistent with this view.

Mechanism of Action of gene *19* Protein. P22 gene *19* encodes a protein that lyses chloroform-treated *E. coli* cells. This was shown by a number of observations. (1). Lytic activity present in cells infected with P22 wild type is missing or altered in cells infected with phage bearing mutations in gene *19* (Botstein et al., 1972). (2). We have purified a lytic activity from phage-infected cells and found that the amino acid sequence of the principal protein species present in the preparation is as predicted by the DNA sequence of gene *19*. (3). We have purified the same activity from cells bearing plasmids containing P22 gene *19* and no other intact phage genes (data not shown). This last observation rules out the possibility that the lytic activity is the product of any P22 gene other than *19*.

TABLE 1

PLASMIDS

Plasmids	
pTP82	*HindII* E fragment (Rutila and Jackson, 1981) of P22 *13*-amH101 *c*1-7 inserted into the *HindII* site of pTR262 (Roberts et al, 1980).
pTP163	2000 bp gene *19* containing *Pst*I fragment from pTP82 inserted into the *Pst*I site of pBR322.
pDR100	1100 bp *Hpa*I fragment containing gene *19* from pTP163 ligated with *Bam*H1 linkers, then digested with *Bam*HI, purified and ligated with *Bam*HI digested pBR322. The orientation is such that gene *13* and *19* are transcribed in the same direction as the *bla* gene of pBR322.
pDR101	1100 bp *Hpa*I fragment containing gene *19* from pTP163 ligated with *Bam*H1 linkers, then digested with *Bam*HI, purified and ligated with *Bam*HI digested pBR322. The orientation is such that gene *13* and *19* are transcribed in the same direction as the tetracycline resistance gene of pBR322.
pDR102	Deletion of DNA in pDR100 between the *Nae*I site present in gene *19* to the *Nae*I site at position 1284 in pBR322, generated by digestion with *Nae*I and religation.
pDR103	Deletion of DNA in pDR101 between the *Nae*I site present in gene *19* to the *Nae*I site at 1284 in pBR322 generated by digestion with *Nae*I and religation.
pDR110	A 189 bp *Msp*I fragment containing an internal coding region of gene *19*, derived from a 1100 bp *Hpa*I fragment of P22 *19*-*am*H1162 *c*1-7, ligated into the *Cla*I site of pBR322.
pDR113	A 423 bp *Msp*I fragment containing an internal coding region of of gene *19*, derived from P22*19*-*am*H1164 *c*1-7, ligated into the *Cla*I site of pBR322.
pDR114	Same construction as pDR113 except that the DNA fragment was isolated from P22 *19*-*am*H1239 *c*1-7.

pDR117 — 1100 bp *Hpa*I fragment from P22 *19-am*H1026 *c*1-7 inserted into the *Eco*RV site of pBR322.

pDR116 — A 550 bp *Hinc*II-*Taq*I fragment from pDR100 ligated to an origin-containing fragment of pBR322 generated by digestion with *Cla*I and *Pvu*II. The pDR100 fragment contains all of gene *19*. The *Cla*I/*Taq*I junction reconstitutes a *Cla*I site.

pDR118 — P_{tac} containing *Pst*I-*Pvu*II fragment from P_{tac} 12 (Amann et al, 1983) ligated with the large fragment generated by digestion of pDR116 with *Cla*I and *Pst*I and filling in the *Cla*I end.

pDR105 — This plasmid contains DNA from 3 sources. One fragment is a *Hind*III-*Pst*I fragment derived from a pBR322 derivative which had the *Pvu*II site converted to a *Hind*III site by the addition of a *Hind*III linker. This fragment, which contains the origin of pBR322, is joined to a *Hind*III-*Ava*II fragment of T4 DNA from lambda-EIIIe+ phage (Owen et al, 1983) that contains gene e. The *Ava*II site is filled in and joined to a *Pst*I-*Pvu*II fragment from P_{tac} 12 containing P_{tac}.

pDR107 — A 2298 bp fragment of lambda *c*1-857 *S-am*7 DNA containing the 3' end of gene *Q*, gene *S* (*S-am*7), gene *R* and gene *RZ* was generated by *Hind*III and *Cla*I digestion. The *Cla*I site was filled in and the fragment was ligated to an origin-containing fragment of pBR322 generated by *Pvu*II and *Hind*III digestion.

TABLE 2

MARKER RESCUE*

P22	no plasmid	pTP163	pDR100	pDR102	pDR103
19-amH1162	1.3x10-6	+	+	2x10-6	.004
19-amH1026	1.1x10-6	+	+	6x10-6	.001
19-amN111	2.1x10-6	+	+	1x10-7	.001
19-amH1239	1x10-7	+	+	.01	1x10-7
19-amH447	4.1x10-6	1.0	1.0	.006	2.0x10-6

* Figures shown are ratios of titers on MS1868 (su^-) bearing the indicated plasmids to titers on the permissive host MS1363(su^+). The "+" symbol means there was efficient plaque formation on a spot test, not quantitated. Plasmid structures are depicted in Fig.1

TABLE 3

COMPLEMENTATION*

Plasmid	Phage: Lambda *S*+ *R*-am	Phage: P22 *19*-am*13*+	Phage: P22 *13*-am *19*+
pDR105 (T4*e*)	+	+	-
pDR118 (P22*19*)	+	+	-
pDR107(lambda) (*R,RZ*)	+	+	ND
none	-	-	-

Complementation. * Plasmid bearing cells (either DB7000 for the P22 infections or W3110 *lacI*^{q} L8 for the lambda infections) were spotted with dilutions of the indicated phage and lysis was noted. Plating efficiencies were determined and in all + cases were found to be 1.0 relative to growth on Su+ cells (MS1363 or C600). ND indicates not done.

Fig. 1 Structures of plasmids used for sequencing and complementation. The top line shows location of P22 genes *23*, *13*, and *19*. The location of gene *23* is established by the sequence of Kroger and Hobom (personal communication). Genes *13* and *19* are located as described in the text. The next 4 lines show the regions of the P22 chromosome present in each pBR322 derivative plasmid. The last three lines diagram the plasmids used for complementation tests. The arrows represent P_{tac} promoters. Plasmid construction is summarized in Table 1.

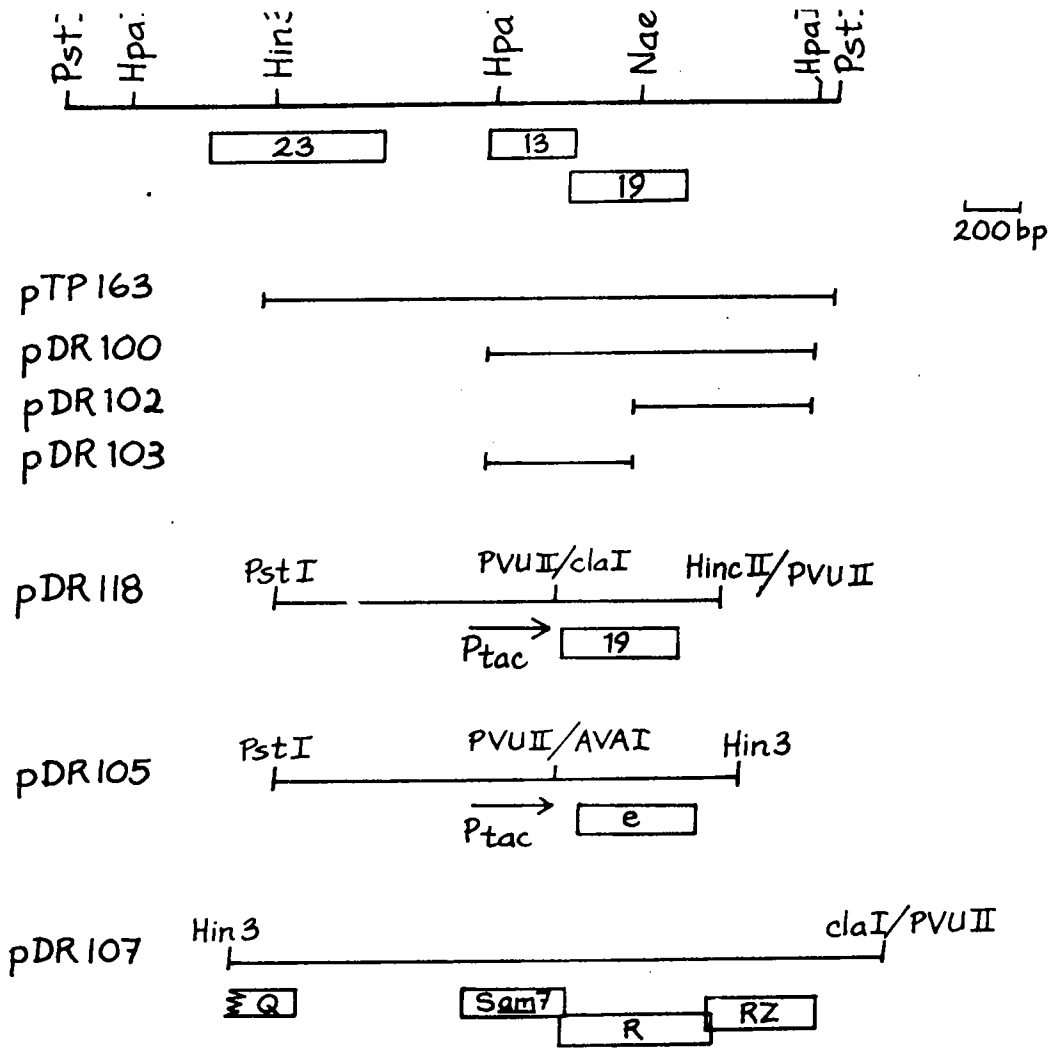

Pst
Hpa
Hin
Hpa
Nae
Hpa
Pst
23
13
19
200bp
pTP163
pDR100
pDR102
pDR103
pDR118
PstI
PVUII/claI
HincII/PVUII
Ptac
19
pDR105
PstI
PVUII/AVAI
Hin3
Ptac
e
pDR107
Hin3
claI/PVUII
Q
Sam7
R
RZ

Fig. 2. DNA sequencing strategy. The bars indicate the location of gene *13* and gene *19* coding sequence. The arrows indicate the orgin and extent of sequence information obtained from different fragments. The arrrows above the map represent sequence determinations from labeled 5' ends; the arrows below, from labeled 3' ends. The enzymes used were: *Eco*RII (E), *Hpa*I (Hp), *Fnud*II (F), *Msp*I (M), *Dde*I (D), *Taq*I (T), *Hinf*I (H), and *Nae*I (N).

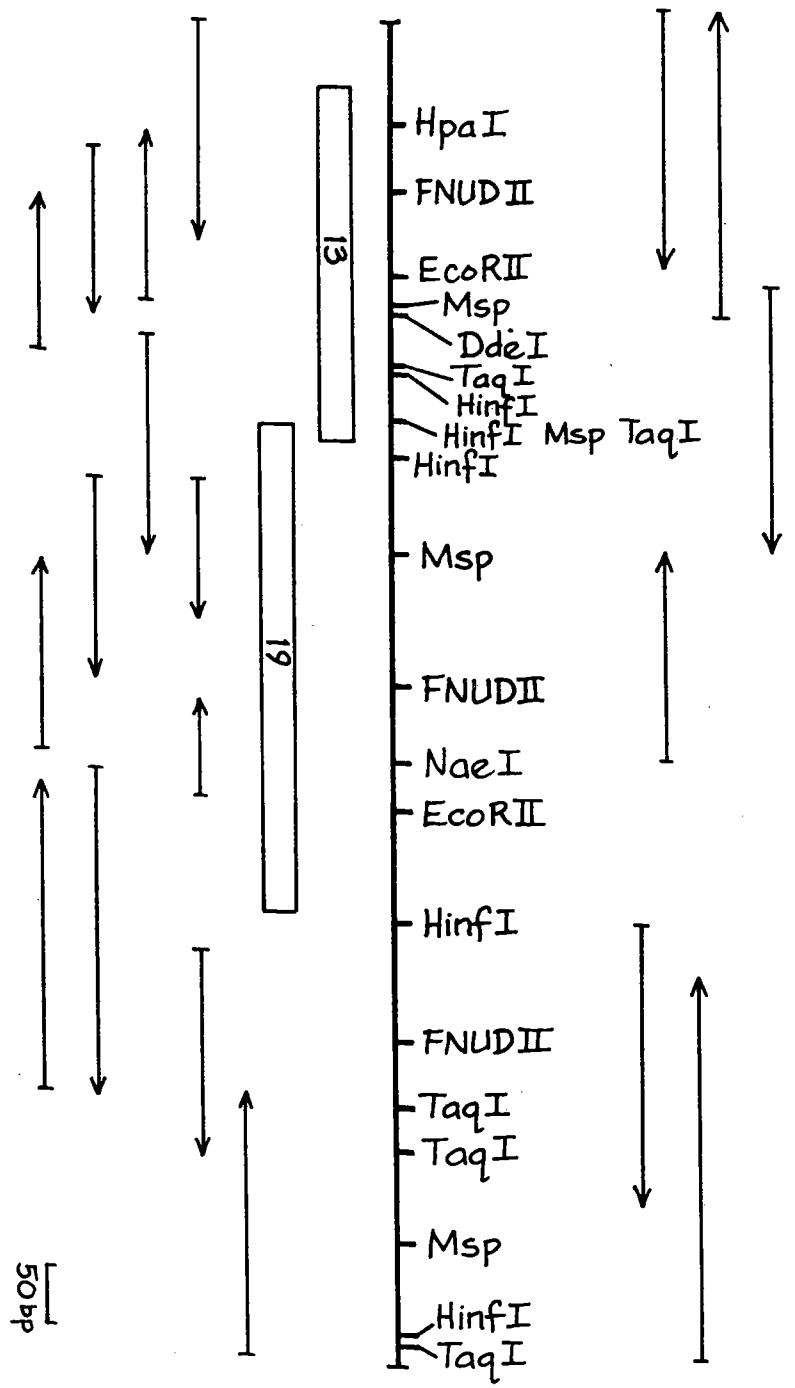

HpaI
FNUDII
EcoRII
Msp
DdeI
TaqI
HinfI
HinfI Msp TaqI
HinfI
Msp
FNUDII
NaeI
EcoRII
HinfI
FNUDII
TaqI
TaqI
Msp
HinfI
TaqI
13
19
50bp

Fig. 3. DNA sequence and amino acid translation of gene *13* and gene *19* of P22. The amino acid translation of gene *13* is shown above the nucleotide sequence and gene *19* translation is shown below the nucleotide sequence. Numbers are shown to the left of the corresponding base. The positions of mutations sequenced are shown by a * above the base. The alleles sequenced were *13-am*H101 (C to T at position 113), *13*-h21 (G to T at position 226), *19*-amH1162 (C to T at position 555), *19*-amH1164 (ACC to CTT at positions 700, 701, and 702), *19-am*H1239 (G to A at position 730); they are shown in order.

```
                                                                  Met Lys Lys Met Pro Glu
  1 CCAATATCAACCAATTCATAACATTGAACAAATCCTCACGGTCGTGAGGTAAGAC ATG AAA AAG ATG CCA GAA

    Lys His Asp Leu Leu Thr Ala Met Met Ala Ala Lys Glu Gln Gly Ile Gly Ala Ile Leu
 74 AAA CAT GAT CTG TTA ACC GCC ATG ATG GCG GCA AAG GAA CAG GGC ATC GGG GCA ATC CTC

    Ala Phe Ala Met Ala Tyr Leu Arg Gly Arg Tyr Asn Gly Gly Ala Phe Lys Lys Thr Leu
134 GCG TTT GCA ATG GCG TAC CTT CGC GGT CGG TAT AAT GGC GGT GCG TTT AAG AAA ACA CTA

    Ile Asp Ala Thr Met Cys Ala Ile Ile Ala Trp Phe Ile Arg Asp Leu Leu Val Phe Ala
194 ATA GAC GCA ACG ATG TGC GCC ATT ATC GCC TGG TTC ATT CGT GAC CTT TTA GTC TTC GCC

    Gly Leu Ser Ser Asn Leu Ala Tyr Ile Ala Ser Val Phe Ile Gly Tyr Ile Gly Thr Asp
254 GGA CTG AGT AGC AAT CTT GCT TAC ATA GCG AGT GTG TTT ATC GGC TAC ATC GGC ACA GAC

    Ser Ile Gly Ser Leu Ile Lys Arg Phe Ala Ala Lys Lys Ala Gly Val Asp Asp Ala Asn
314 TCG ATT GGT TCG CTA ATC AAA CGC TTC GCT GCT AAA AAA GCC GGA GTC GAT GAT GCA AAT
                                                                  Met Met Gln Ile

    Gln Gln ***
372 CAG CAG TAA C GGA ATC ACC AGA TTA AAA CGT GAA GAA GGT GAG AGA CTA AAA GCC TAT
     Ser Ser Asn  Gly Ile Thr Arg Leu Lys Arg Glu Glu Gly Glu Arg Leu Lys Ala Tyr

432 TCA GAT AGC AGG GGG ATA CCA ACC ATT GGG GTT GGG CAT ACC GGA AAA GTG GAT GGT AAT
    Ser Asp Ser Arg Gly Ile Pro Thr Ile Gly Val Gly His Thr Gly Lys Val Asp Gly Asn

492 TCT GTC GCA TCA GGG ATG ACA ATC ACC GCC GAA AAA TCT TCT GAA CTG CTT AAA GAG GAT
    Ser Val Ala Ser Gly Met Thr Ile Thr Ala Glu Lys Ser Ser Glu Leu Leu Lys Glu Asp

552 TTG CAG TGG GTT GAA GAT GCG ATA AGT AGT CTT GTT CGC GTC CCG CTA AAT CAG AAC CAG
    Leu Gln Trp Val Glu Asp Ala Ile Ser Ser Leu Val Arg Val Pro Leu Asn Gln Asn Gln

612 TAT GAT GCG CTA TGT AGC CTG ATA TTC AAC ATA GGT AAA TCA GCA TTT GCC GGC TCT ACC
    Tyr Asp Ala Leu Cys Ser Leu Ile Phe Asn Ile Gly Lys Ser Ala Phe Ala Gly Ser Thr

672 GTT CTT CGC CAG TTG AAT TTA AAG AAT TAC CAG GCA GCA GCA GAT GCT TTC CTG TTA TGG
    Val Leu Arg Gln Leu Asn Leu Lys Asn Tyr Gln Ala Ala Ala Asp Ala Phe Leu Leu Trp

732 AAA AAA GCT GGT AAA GAC CCT GAT ATT CTC CTT CCA CGG AGG CGG CGA GAA AGA GCG CTG
    Lys Lys Ala Gly Lys Asp Pro Asp Ile Leu Leu Pro Arg Arg Arg Arg Glu Arg Ala Leu

792 TTC TTA TCG TGA GTCGTATTAAGGCAATTATTGCGTCTGTCATTATCTGCATCATCGTCTGTCTTTCGTGGGCTG
    Phe Leu Ser ***

867 TTAATCATTATCGTGATAACGCCATCACCTACAAAGAGCAGCGCGATAAAGCCACATCAATCATCGCTGATATGCAGAA

946 GCGTCAACGAGATGTAGCAGAACTCGATGCCAGATACACAAAGGGACTTGCTGATGCTAACGCGACTATCGAAACTCTC

1025 CGCGCTGATGTTTCTGCTGGCGTAAGCGCCTGCAAGTCTCCGCCACCTGTCCAAAGTCAACGACCGGAGCCAGCGGCAT

1104 GGGCGATGGAGAAAGCCCAAGACTTACAGCAGATGCTGAACTCAATTATTACCGTCTCCGAAGTGGAATCGACAGGATA

1183 ACCGCG
```

Figure 3

Fig. 4. An Amino acid sequence (predicted from DNA sequence) comparison between lambda *S* and P22 *13*. P22 *13* is shown on the top line and lambda *S* is given on the lower line. There is 89 percent homology between the two proteins.

MET LYS LYS MET PRO GLU LYS HIS ASP LEU LEU THR ALA MET MET ALA ALA LYS GLU
MET LYS --- MET PRO GLU LYS HIS ASP LEU LEU ALA ALA ILE LEU ALA ALA LYS GLU

GLN GLY ILE GLY ALA ILE LEU ALA PHE ALA MET ALA TYR LEU ARG GLY ARG TYR ASN
GLN GLY ILE GLY ALA ILE LEU ALA PHE ALA MET ALA TYR LEU ARG GLY ARG TYR ASN

GLY GLY ALA PHE LYS LYS THR LEU ILE ASP ALA THR MET CYS ALA ILE ILE ALA TRP
GLY GLY ALA PHE THR LYS THR VAL ILE ASP ALA THR MET CYS ALA ILE ILE ALA TRP

PHE ILE ARG ASP LEU LEU VAL PHE ALA GLY LEU SER SER ASN LEU ALA TYR ILE ALA
PHE ILE ARG ASP LEU LEU ASP PHE ALA GLY LEU SER SER ASN LEU ALA TYR ILE THR

SER VAL PHE ILE GLY TYR ILE GLY THR ASP SER ILE GLY SER LEU ILE LYS ARG PHE
SER VAL PHE ILE GLY TYR ILE GLY THR ASP SER ILE GLY SER LEU ILE LYS ARG PHE

ALA ALA LYS LYS ALA GLY VAL ASP ASP ALA ASN GLN GLN
ALA ALA LYS LYS ALA GLY VAL GLU ASP GLY ARG ASN GLN

Figure 4

CHAPTER III

ISOLATION AND CHARACTERIZATION OF P22 LYSOZYME AMBER MUTATIONS

CHAPTER III

ISOLATION AND CHARACTERIZATION OF P22 LYSOZYME AMBER MUTATIONS

Introduction

Studies described in the previous chapter resulted in the identification of five different amber mutations in the P22 lysozyme gene; three were sequenced and found to be at codons 64, 113, and 122. Subsequently the others were sequenced and found to be at codons 65 and 80. To enlarge this collection, seven more amber mutations were created at six additional codons by mismatched primer-directed mutagenesis in vitro. Oligonucleotides that incorporated amber codons were used as primers of DNA synthesis on single stranded plasmid templates bearing gene *19*. The amber mutations were subsequently transferred to P22 by homologous recombination in vivo. In this chapter I describe the procedures used to create the amber mutations. The phage bearing the amber mutations in gene *19* were plated on six suppressor strains to determine which amino acid substitutions yielded functional lysozyme.

Materials and Methods

Enzymes, buffers and media. T4 DNA ligase, restriction endonucleases, and *E. coli* DNA polymerase I large fragment were purchased from New England Biolabs. AMV reverse transcriptase, *Hpa*I restriction endonuclease, and *Sal* I restriction endonuclease were purchased from Boehringer-

Mannheim. P22 tail protein was a gift from Peter Berget. Hen egg white lysozyme (HEWL) was dissolved in water (1 mg/ml), it was used at a final concentration of 20 to 50 µg/ml.

Restriction buffers, kinase buffer, and ligase buffer were as recommended by the suppliers. PCM was 10 mM Pipes pH 6.8, 10 mM $CaCl_2$, and 10 mM $MgCl_2$. 10 x sample buffer was 0.25% bromophenol blue, 0.25% xylene cyanol and 25% glycerol. Sequencing gel loading buffer was 80% deionized formamide, 50 mM Tris-borate (pH 8.3), 1 mM EDTA, 0.1% (w/v) xylene cyanol, and 0.1% (w/v) bromophenol blue. Precipitation buffer was 20 mM Tris HCl pH 8.0, 10 mM NaCl, 2 mM EDTA, and 0.5 M NH_4Ac. 5 x AB buffer was 250 mM Tris HCl, pH 8.3, 300 mM NaCl, and 50 mM DTT. 5 x RT was 250 mM Tris HCl, pH 8.3, 300 mM NaCl, 50 mM DTT, and 150 mM $Mg(OAc)_2$. 5 x dNTP was a mixture of all four dNTP's at 2 mM in 1 x AB. 5 x ddNTP was (one mixture for each of the four ddNTP's) each ddNTP at 250 mM in 1 x AB. Buffer C was 200 mM Tris HCl pH 7.6, 100 mM MgCl, 500 mM NaCl, and 10 mM DTT. Buffer D was 200 mM Tris HCl pH 7.6, 100 mM MgCl. DNA buffer was 10 mM Tris HCl (pH 7.4), 5 mM NaCl, and 1 mM EDTA. BS (buffered saline) was 90% (% vol) 0.85% NaCl and 10% SB. SB was 0.4 M Na_2HPO_4, and 1.1 M KH_2PO_4 pH approximately 7.0.

LB broth contained 10 g tryptone, 5 g yeast extract, 5 g NaCl, and 1 ml 1 M NaOH per liter. LB agar contained, in addition, 15 g agar per liter. LB medium was supplemented with appropriate antibiotics at the following concentrations;

10 to 20 μg tetracycline/ml, 50 to 100 μg ampicillin/ml, or 20 μg/ml kanamycin for growth of plasmid-bearing strains. Soft agar contained 10 g tryptone, 2.5 g NaCl and 9 g agar per liter. Minimal medium was M9CAA (Smith and Levine, 1964) with the casein hydrolysate omitted. M9 plus glucose was minimal media with 1% glucose added. Lambda plates contained 10 g tryptone, 2.5 g NaCl and 11 g agar per liter of H_2O.

Plasmids. Plasmid constructions were done by standard methods (Maniatis et al., 1982). Transformation of *Salmonella* was carried out as described by Lederberg and Cohen (1974).

Plasmids used for constructions are derivatives of pBR322 (Bolivar et al., 1977). The details of the constructions are given in Table 4. Plasmids pGFIB:phe and pGFIB:cys (Normanly et al., 1986); pDR105, pDR118, pDR110, pDR100, and pDR116 (Rennell and Poteete, 1985;Table 1 and Fig 1); pTP352 (Knight et al., 1987); pTP30 (Berget et al., 1983); pZ152 (Zagursky and Berman, 1983); and pKM101 amp^r (Youderian et al., 1982) were previously described. Plasmid pPB20::Tn5-13 was generously supplied to us by Peter Berget. This plasmid has the pBR322 *Hind*III-*Bam*HI origin of replication fragment; a 4800 bp segment of bacteriophage T4 DNA containing all of gene *7*, part of gene *6*, and part of gene *8*; and Tn5 inserted in the T4 DNA portion. Plasmids pDR281 and pDR392 were used for the site directed mutagenesis. Both plasmids

contain an M13 origin of replication and P22 gene *19*. pDR281 confers ampicillin resistance and has gene *19* under P_{tac} control. Plasmid pDR392 confers tetracycline resistance and has gene *19* under P_{lac} control. pBR322 oc4 is pBR322 containing an ochre mutation in the 4th codon of beta lactamase. This plasmid was generously supplied to us by David Botstein. pDR321 was used to construct the defective phage P22 Kn321. It contains a 5.04 kilobase internal *Hpa*I fragment of Tn5 (Rothstein et al., 1980) obtained from pPB20::Tn5-13, inserted into the middle of the coding sequence of gene *19*. This Tn5 DNA encodes resistance to kanamycin (Jorgensen et al., 1979) and is unable to transpose because it lacks essential parts of the transposon's insertion sequences. Plasmid pDR321 also has P22 DNA flanking gene *19* in order to allow for homologous recombination between the phage and the plasmid.

Plasmids pGFIB:phe and pGFIB:cys, encode amber suppressor tRNA's that insert phenylalanine and cysteine respectively; both confer resistance to ampicillin as well (Normanly et al., 1986). I found that *Salmonella* strains bearing these plasmids were unstable, tending to lose the plasmids during prolonged growth in liquid culture in spite of the addition of ampicillin to the medium. To give the plasmids a more strongly selectable trait they were modified by the addition of the tetracycline resistance determinant of pBR322 (See Table 4).

Bacteria. *Escherichia coli* W3110 *lacI*^{q} L8 (Brent & Ptashne, 1981) was used for propagation of plasmids and for growth of lambda phage. W3110 (Lederberg, 1960) was used as host to prophage P22 Kn321 *sie*A44 *m*44. Strain GM1675 *dam*-4 del(*lac-pro*) *thi*-1 *sup*E *rel*A1 */F'*lacI^q del(M15) *pro*+, used for propagating phage f1 IR1 and generating single-strand plasmids, was provided by Martin Marinus. All other strains are derivatives of *Salmonella typhimurium* LT2. Strains MS1362, MS1363, MS1364, MS1365 (all *leu*Aam414, bearing the amber suppressor alleles *sup*D, *sup*E,*sup*F, *sup*G, respectively), DB7000 (*leu*Aam414), MS1868 (*leu*Aam414 r^- *m*+), MS2310 (MS1868 bearing the plasmid pKM101 amp^r), and MS1387 *sup*Q *pro*(del *att*P22) *cys*B *his leu*D *fol*101 /F'*lac pro* were from Miriam Susskind. Strain CV112 (*polAts*) (Van Beveran, Ph.D. thesis, Tufts University, 1985) was provided by Andrew Wright. TP278, TP279, TP280, TP282 are MS2310 with *sup*E,*sup*F, *sup*G, and *sup*D respectively. These strains were created by transducing with P22 *HT* the amber-suppressors from MS1363, MS1364, MS1365 and MS1362 into MS2310 and selecting leu$^+$ transductants on minimal medium containing ampicillin. TP284 is MS1868 lysogenized with P22 *sieA*44 *Ap*2 del 7283. TP292 and TP241 are W3110 and MS1363 respectively, lysogenized with P22 Kn321 *sieA*44 *m*44. TP308 and TP309 are MS2310 bearing pDR463 and pDR464 respectively. TP320 is MS1868 with plasmid pDR105.

Phage. Lambda phage *R*am5 *c*I- was used for complementation tests. Phage f1 *IR*1 (Enea and Zinder, 1982) was used to package single stranded plasmids. The remainder of the phages used were P22. Synthetic amber mutations were introduced into P22 *sie*A44 *m*44. P22 *HT* was used for transductions. A lysogen of P22 Ap2 *sie*A44 *mnt-ts*1 (deletion Ap2-gene 20) was used to titer defective oversized P22 particles (Weinstock et al., 1979). The phage bearing amber mutations in gene *19* were either clear derivatives that had a plaque morphology marker (*h*21) as described in chapter 2 (P22 19-amH1162 *h*21 *c*1-7, P22 *19*-amH1164 *h*21 *c*1-7, P22 *19*-amH1239 *h*21 *c*1-7, P22 *19*-amN111 *c*2, and P22 *19*-amH1026 *h*21 *c*1-7) or *c*+ phage as described in this chapter (P22 *sie*A44 *m*44 *19*am12, P22 *sie*A44 *m*44 *19*am22, P22 *sie*A44 *m*44 *19*am39, P22 *sie*A44 *m*44 *19*am61, P22 *sie*A44 *m*44 *19*am65, P22 *sie*A44 *m*44 *19*am82, and P22 *sie*A44 *m*44 *19*am83). The five amber alleles described in chapter 2 have been renamed to agree with the nomenclature used for the rest of amber alleles . *19*-amH1162, *19*-amH1164, *19*-amH1239, *19*-amN111, and *19*-amH1026 will be known as *19*am64, *19*am113, *19*am122, *19*am80, and *19*am65 respectively. I will not mention the *sie*A44, *m*44, *h*21, *c*2, or *c*1-7 alleles below when describing the *19*am phage. The phage used in immunity tests were P22 *vir*3, P22 *c*2-5, and P22 *c*1-7. The defective P22 phage used to cross mutant alleles of gene *19* from plasmids to phage was P22 Kn321 *sie*A44 *m*44.

I had many problems with P22 *19*am122(trp). The

suppression pattern showed that it was *ts* with all the suppressor strains. The *ts* phenotype was attributed to the lysozyme gene, since phage plated on lysozyme-producing cells at 40^0C would yield a wild type plaque size indicating that only lysozyme was defective at 40^0C. The reversion rate of this *ts* phenotype was such that in a typical stock there would be 10% non-*ts* phage that still contained amber mutations. These revertants did not yield wild type phage in a cross with the *19*am122(trp) parent indicating that the original amber mutation was retained. The *19*am22 allele was "backcrossed" into P22 by constructing a plasmid containing this allele and then allowing the plasmid to undergo homologous recombination with P22 Kn321. The recombinant phage that resulted did not have the *ts* phenotype, but still contained *19*am122(trp). I reasoned that the original isolate of this phage had a second mutation that made the phage *ts*. This mutation was crossed out when I backcrossed the amber mutation. The new non-*ts* phage would not recombine to yield wild type with the original phage, therefore both phage contained the same amber mutation.

Construction of P22 Kn321 *sie*A44 *m*44. A defective P22 phage was constructed by crossing P22 *sieA*44 *m*44 with a plasmid that carried a selectable marker (See Fig. 5). MS1868 was transformed with pDR321. The plasmid-bearing strain was then infected with P22 *sie*A44*m*44 at a multiplicity of 10 and aerated at 37^0C for two hours. The culture was treated with

chloroform and the lysate was used to transduce kanamycin resistance to a *pol*A-ts *Salmonella* strain (CV112). The transduction was done by infecting, at a multiplicity of 10, a 5 ml culture of the *pol*A strain which was at a cell density of 1 x 10^8/ml. After 20 minutes for adsorption of the phage, 0.5 ml of the infected cells was spread on a plate that contained 20 µg/ml kanamycin. Plates were incubated at 30^0C and at 37^0C. The plates that were grown at 30^0C had about 100 colonies, and when tested, most of the colonies had both kanamycin and ampicillin resistance. This indicated that in these colonies the entire plasmid was present. There were at least 10 fold fewer colonies on the plates that were grown at 37^0C. The majority of colonies that were tested were kanamycin resistant and ampicillin sensitive, indicating that the plasmid was not present.

The kanamycin resistant, ampicillin sensitive colonies were further characterized by testing for immunity to P22 infection and ability to yield plaque-forming phage. In the immunity test, colonies were cross streaked against P22 *c*1-7, P22 *vir*-3, and P22 *c*2-5. These phages were suspended in BS buffer and 0.1 ml streaks containing 1 x 10^8 plaque-forming units were placed on lambda plates. The colonies to be tested were streaked over the phage so that the last phage to be encountered would be P22 *vir*-3. The plates were incubated overnight at 37^0C. In this test, lysogens are sensitive to

P22 *vir*-3 and resistant to P22 *c*1-7 and *c*2-5. In the "yielder" test the recombinant lysogen should not be able to release plaque-forming phage upon induction for two reasons. First the 5.04 kb Tn5 insertion makes the phage genome oversized. Therefore, no individual phage will have a complete genome or have the terminal redundancy necessary for circularization. (The recombinant phage is able to form a lysogen, however: two or more phage infecting the same cell can recombine and circularize.) Second, the recombinant phage is unable to produce lysozyme, since the coding sequence of gene *19* has been interrupted by the Tn5 insertion. The phage yielding test was done by making a small patch of the candidate lysogen on a lambda plate, then exposing the plate to an inducing dose of UV light. A drop of soft agar containing a small inoculum of TP278 was placed over the patch. The plate was incubated at 30^0C for 12 hours. Colonies that failed to yield phage and were immune to *c*1-7 and *c*2-5 were tested further. As a control, I carried out "yielder" tests with colonies which were resistant to both kanamycin and ampicillin as well as immune. All these colonies yielded plaque-forming phage upon induction. This observation suggests that these colonies contained both intact plasmid and phage. This could occur if the cell had been infected with both a transducing particle and a phage particle. The *Pol*A-ts allele might be "leaky"

allowing some plasmid replication. Another interpretation of this observation would be the presence of a cointegrate of the phage and plasmid that resolved upon induction.

Candidate lysogens were grown as liquid cultures (20 ml) in LB plus kanamycin to a density of 2 x 10^8/ml and induced with 2.4 μg/ml mitomycin C. The induced cultures were aerated at 37^0C for 3 hours and then 0.5 ml of chloroform and 0.5 ml HEWL (1 mg/ml) were added to lyse the cells. When the culture cleared, debris was removed by low-speed centrifugation. The phage particles were pelleted by a high speed spin at 27,000 g for 90 min. and resuspended with 2 ml of BS. This stock of defective phage was titered on TP284. The incoming phage recombine with the resident defective prophage, in this way overcoming the lack of terminal repetition due to the large Tn5 insertion, and thus being able to circularize (Weinstock et al., 1979). The resident prophage has an intact lysozyme gene, and so can complement incoming *19*- phage.

The P22 Kn321 stock was used to lysogenize MS1363. A kanamycin resistant colony was chosen and checked for the presence of a defective prophage as described above. This strain (TP241) was used to cross plasmid-borne *19*-amber alleles onto the phage genome (see Fig. 6).

When I started developing these methods I used the *Salmonella* strain TP241, but found this to be cumbersome since

to transform this strain with the amber-bearing plasmid, plasmids had to be prepared by CsCl equilibrium gradients. This was necessary since the efficiency of transformation of *Salmonella* is at least 1000 fold lower than that of *Escherichia coli*. Therefore, the defective prophage was moved to an *Escherichia coli* strain by mating. The *Escherichia coli* strain carrying the defective P22 prophage (TP292) was used in the same manner as the *Salmonella* strain (TP241), except it could be transformed with DNA isolated from minilysates. The *E. coli* lysogen was made by A. Poteete. Briefly, the steps were: lysogenization of an auxotrophic F' *Salmonella* strain (MS1387, *sup*Q *pro*(del *att*P22) *cys his leu*D *fol*101 /F'*lac pro*) with P22 Kn321, then mating this lysogen with an *E. coli* F- prototroph (W3110), and selecting a kanamycin resistant prototroph.

Mutagenic primers. The primers were 15 base oligonucleotides synthesized by the University of Massachusetts DNA synthesis facility. They were designed to introduce amber codons into P22 gene *19*. The sequences of the primers are:

am12(lys)	5'AGATTATAGCGTGAA 3'
am22(tyr)	5'GCCTAGTCAGATAGC 3'
am39(val)	5'GGAAAATAGGATGGT 3'
am61(glu)	5'CTGCTTAAATAGGAT 3'

am65(trp) 5'TTGCAGTAGGTTGAA 3'

am82(gln) 5'CTAAATTAGTATGAT 3'

am83(tyr) 5'AACCAGTAGGATGCG 3'

The mismatched bases were at least 5 bases in from the 5' end. The primers were supplied as lyophilized mixtures of the 15-mer and all the products of incomplete incorporation. This mixture was resuspended with 1 ml of water which normally yielded a solution of 4mg/ml. The primers were phosphorylated. The kinase reaction mixture contained:

75 µl of the primer mixture (about 300 ug)

15 µl 10 x kinase buffer

20 µl ATP (100 mM)

2 µl T4 DNA kinase

38 µl of H_2O

The mixture was incubated at 37^0C for 30 minutes. Oligonucleotides were then repurified by phenol extraction and ethanol precipitation. The pellet was resuspended with 75 µl DNA buffer. The DNA was then electrophoresed after the addition of 7.5 µl sample buffer, at 150 volts on a 20% acrylamide, 1% bis-acrylamide gel until the bromophenol blue migrated 17 centimeters. To visualize the DNA, the gel plates were separated and the plate with the gel was placed over a sheet of Kodak Eastman "chromogram sheet" (a thin layer of cellulose with a fluorescent indicator) and then illuminated from above with a short wave UV light. The DNA created

visible shadows. The 15-mer was easily distinguished from the other bands which formed a ladder from 1-mers to 14-mers. The 15-mer band was cut out of the gel by using a razor blade. The gel slice was finely diced with a razor blade and placed in 1 ml of 0.5 M ammonium acetate, 1 mM EDTA. This was incubated with shaking at 37^0C for 12 hours. The solution was then spun by low speed centrifugation through a large (1 ml) pipet tip and the eluate collected in a plastic 3ml vial. The pipet tip trapped the gel fragments. The fragments were then washed twice with 200 µl of the ammonium acetate solution. The resultant 1.4 ml was brought up to a volume of 10 ml with H_2O and further purified on a "Sep-Pak" C_{18} cartridge (Waters). The Sep-Pak column was prepared by washing with 10 ml of ethanol, then 10 ml of 10 mM ammonium acetate, followed by 10 ml of H_2O. The sample was applied to the column and pushed gently through. The column was washed with 10 ml of 10 mM ammonium acetate and the DNA was eluted with 25% ethanol, 75% 30 mM ammonium acetate. Most of the 15-mer eluted in the first 1 ml, but 1.5 ml was routinely collected. This suspension was frozen and lyophilized. The powder was resuspended with 0.5 ml of DNA buffer. The absorbance at 260 nm was determined and the concentration of DNA was estimated assuming that the A_{260} of 20 µg oligonucleotide/ml is 1.0.

Isolation of single strand template. GM1675 was transformed to ampicillin resistance with pDR281. GM1675 was

used since it has a mutation in *dam*; plasmid or phage grown in this strain are not methylated at *dam* sites (GATC). When single strand plasmid isolated from this strain is used as template for in vitro second strand synthesis, both strands of the resulting duplex are unmethylated at *dam* sites. When unmethylated plasmid DNA is introduced into *E. coli*, the cell's mismatch repair system will repair any mismatches randomly, rather than repairing the newly synthesized unmethylated strand preferentially as it would if methylated single stranded plasmid DNA was used as template. (Radman and Wagner, 1984).

An individual transformant was streaked for single colonies on an LB plate containing ampicillin. After overnight incubation at 37^0C, a single colony was used to inoculate 5 ml of LB broth (containing ampicillin) and the culture was grown to saturation. A 0.1 ml portion of the saturated culture was used to inoculate 2 ml of LB broth that contained 100 µg/ml ampicillin. 1×10^{11} f1 *IR*1 phage were added. The culture was aerated at 37^0C for five to eight hours. The cells were pelleted by spinning 1.5 ml in the microfuge for 5 minutes. The supernatant was added to 0.15 ml of 20% PEG 2 M NaCl, and incubated on ice for 15 minutes. The precipitated phage were pelleted by a 10 minute spin in the microfuge. The supernatant was poured off and the tube was wiped clean. The pellet was resuspended with 0.2 ml of DNA buffer and extracted with an equal volume of phenol. Two drops of 8 M ammonium

acetate were added to the aqueous phase and the solution was extracted with 0.5 ml of ether. The tube was filled with ethanol and placed at -70^{0}C for 30 minutes and the DNA was pelleted by spinning in a microfuge for 10 minutes. The pellet was dried under vacuum and then resuspended with 20 µl of DNA buffer.

The ratio of single strand phage DNA to single strand plasmid DNA was estimated by visualizing the DNA after electrophoresing 2 µl on a 0.7 percent agarose gel and staining with ethidium bromide. The phage is 6 kb and the plasmid is 4.5 kb. The ratio of plasmid to phage DNA varied between 1 and 0.01.

The reason for the large variation in the amount of plasmid DNA recovered has not been elucidated. Some possible reasons were investigated. One possibility was the state of the cells at the time of infection. However, varying the procedure by infecting cells in early, mid, and late log phase produced no consistent variation in the ratio of plasmid to phage DNA recovered. Another possibility was the length of time of infection. I varied the length of infection from 4 hours to overnight but again there was no consistent change in the ratio recovered. It was noted that infections done for 6 hours seemed to yield the most total DNA. Another possibility was that the culture lost the plasmid at a high frequency, since the plasmid was detrimental to the cells: it has the lysozyme gene under control of P_{tac}. To investigate whether

this was the problem, pDR392 was constructed because this has the less active promoter P_{lac}; in addition, pDR392 confers tetracycline resistance. This was tried because when ampicillin resistant cells are grown in liquid culture, the secreted β-lactamase breaks down the ampicillin in the medium, thus decreasing the drug concentration and thereby allowing nonresistant cells to grow. Tetracycline resistance, on the other hand, is mediated by activation of a membrane pump that actively transports tetracycline out of the cells; the level of drug in the medium remains constant, and a strong selection for plasmid-bearing cells is maintained (Benveniste and Davis, 1973). The plasmid conferring tetracycline resistance did not change the variability of the ratio of phage to plasmid DNA seen previously using the plasmid encoding ampicillin resistance.

The only factor, that was found to affect the ratio of phage to plasmid DNA recovered, was the length of time that the plasmid had been resident in the cells at the time of infection. The best procedure appeared to be to transform the host cell with purified plasmid DNA, and to select transformants, grow liquid cultures, and infect in quick succession, never allowing the plasmid-bearing cells to experience less-than rapid growth for any significant period of time. Adherence to this schedule typically resulted in ratios between 0.5 and 1.

In vitro extension and ligation reactions. In vitro extension-ligation reactions were done by mixing 0.6 to 1 pmol of single strand target plasmid (1 µl of template prepared as described above) with 15 to 17 pmol of mutagenic primer (15 base oligonucleotide), 10 pmol of a primer designed to anneal to pBR322 sequence in the plasmid such that it's 3' end is 13 bases 5' of the *EcoR*I site (5'GTATCACGAGGCCCT 3' purchased from New England Biolabs), 1 µl of buffer C, and 5 µl of H_2O. This was incubated at 65^0C for 5 minutes, then placed at room temperature for 30 minutes. To this annealing reaction, I added 10 µl of an enzyme mixture which contained 1 µl 10 X buffer D, 1µl of dNTP (2.5 mM each of dATP, dGTP, dTTP, dCTP), 1 µl 100 mM ATP, 1 µl 100 mM DTT, 2 µl (20 units) *Escherichia coli* DNA polymerase I large fragment and 2 µl of T4 DNA ligase. The mixture was incubated at 15^0C overnight. Control reactions contained everything except for the mutagenic primer.

I used three variations of this protocol to construct the amber alleles. The variations consisted of different single strand templates. The templates were derived from pDR281, pDR292, or pDR401. The use of the different templates did not result in a change in the frequency of obtaining amber mutations which was usually between 2 and 14%.

The plasmid pDR401 was used for the gapped circle

mutagenesis (Fig.7). The DNA species in the reaction consisted of 0.015 pmol of pDR401 which had been digested with *Pst*I, 0.1 pmol of a *BamH*I to *EcoR*I pBR322 origin- containing fragment, and 10 pmol of the mutagenic primer (This reaction should have had 1 pmol of each pDR401 and the *BamH*I-*EcoR*I fragment). Plasmid pDR401 is pBR322 oc4 with a fragment containing a P_{lac} promoter transcribing gene *19* inserted into the *Eco*RI site. The plasmid is tet^r, amp^s, and lysozyme+ when it is grown on a non ochre-suppressing strain. The double strand DNA in the reaction was denatured by heating to 100^0C for 3 minutes then the mutagenic primer was added and the reaction was heated at 65^0C for 10 minutes. The DNAs were then annealed by allowing the tube to set at room temperature for 20 minutes at which time the *Escherichia coli* DNA polymerase I large fragment and T4 DNA ligase were added.

This reaction was designed to increase the efficiency of obtaining the mutant (See Fig. 7). The idea was to decrease the amount of second strand synthesis that would be needed to form a double strand plasmid. This reaction was also designed to increase the chance of *E. coli* mismatch repair system correcting the heteroduplex in such a way as to conserve the amber mutation. The host for pDR401 was GM1675 (*Dam*-), so that the newly formed heteroduplex would be unmethylated at *dam* sites on the strand that contained the *ochre* mutation and the lysozyme-plus allele, whereas the other

strand would be at least 3/4 methylated at these sites, since the pBR322 that the *BamHI-EcoRI* fragment was isolated from was grown in a *dam+* cell. The chances of maintaining the lysozyme amber allele (introduced by the primer) and the ampicillin resistance wild type allele (introduced on the pBR322 fragment) would therefore be increased.

Only two products of the in vitro reaction would plate on tetracycline and ampicillin plates, both which would be a result of correction in the heteroduplex formed. One would be the result of correcting the ochre in the beta-lactamase gene but leaving intact the lysozyme gene either because the mutagenic primer did not anneal or because the mismatch repair system corrected the amber. The second product that would plate was the product desired, a plasmid with a wild type beta-lactamase and an amber in lysozyme gene. The parents (the pBR322 fragment and the pDR401) were tet^s or amp^s. The reaction conditions were the same as outlined above, except that the extension ligation reaction was done for only 2 hours.

Screening for lysozyme mutants. The reaction mixtures were used to transform W3110 *lacI*^{q} to ampicillin resistance. The transformations were done by adding 1 to 2 µl of reaction mixture and 200 µl of PCM to 300 µl of competent *E. coli* incubating at 0^0C for 30 minutes, then at 37^0C for 2 minutes,

then at room temperature for 10 minutes. The transformation mixtures were added to 2 ml LB broth and aerated at 37^0C for 30 minutes, following which 0.5 ml was spread on a LB ampicillin plate. I normally obtained between 50 and 100 transformants per plate. On the control plates (no mutagenic primer) I usually had about the same number of transformants.

Ampicillin resistant colonies that arose from the transformation were screened for the introduced amber mutation by a complementation test. Approximately 1×10^8, 1×10^6, and 1×10^4 plaque forming units of lambda *cI-Ram* were streaked in lines on a lambda plate and the transformants were streaked across the lines from the least dilute to the most concentrated using a toothpick. The transformants that failed to lyse contained plasmids that were defective in lysozyme, possibly because an amber mutation had been introduced. The colonies from the control plates (no mutagenic primer added) would consistently lyse indicating that the lysozyme gene was intact. The percentage of experimental colonies that lysed varied, depending upon the efficiency of the extension ligation reaction. The range normally seen was between 1 and 40%. This did not seem to be a function of an individual primer, since for any given primer the range would vary between different reactions. Occasionally a colony would give an intermediate result. If the positive control streak (a colony known to produce lysozyme) showed lysis when streaked

across 1 x 10^4 phage and a negative control colony (a colony containing pBR322) had no lysis even when streaked across 1 x 10^8 phage, then an intermediate colony would show lysis when streaked across 1 x 10^6 and possibly a small amount of lysis when streaked across 1 x 10^4 phage. I reasoned that these colonies arose because they contained two populations of plasmids. One possibility was that the heteroduplex plasmid that resulted from the in vitro reaction was replicated before it was corrected by the mismatch repair system. To determine if this was the case, I streaked such an intermediate colony for single colonies, then picked several single colonies and streaked those for single colonies then tested several of the single colonies. Three distinct populations of ampicillin resistant colonies resulted: intermediate, lysozyme plus, and lysozyme minus. This result showed that the cells did contain more than one type of plasmid. Normally I discarded the intermediate colonies because it was too time consuming to get the plasmids to segregate.

Crossing amber alleles into P22. Plasmid DNA (isolated from colonies that were scored as lysozyme minus), prepared by CsCl equilibrium gradients or from minilysates, was used to transform strain TP241 or strain TP292 to ampicillin resistance (See Fig. 6). CsCl-purified plasmid DNA was used to transform TP241, whereas DNA isolated from minilysates was used to transform TP292. Transformed clones were purified by

streaking. Single colonies were used to inoculate 5 ml of LB containing both ampicillin and kanamycin. The cultures were grown with aeration at 37^0C until they had reached a cell density of about 2 x 10^8/ml. Mitomycin C (2.4 µg/ml) was added and the cultures were aerated for an additional 4 hours. Chloroform (0.5 ml) and HEWL were added to lyse the cultures. The debris was spun out by a 5 minute spin at 5900 g and the resulting lysate was used to infect a *Salmonella su+* strain. Before the lysate was used for infection, the phage were tailed by adding 1 x 10^{11} P22 tails. It has been noted that induced P22 prophages are deficient in tails. If the *Escherichia coli* defective lysogen was used, then the *Salmonella su+* strain was also *r-* *m+* so that unmodified phage would not be restricted.

The *Salmonella su+* strain used for the infection depended upon which amino acid was originally encoded by the codon that was changed to an amber. If the amber mutation substituted for a codon of an amino acid for which a suppressor tRNA strain was available, then that strain was used, since a functional lysozyme would be produced. On the other hand, if the amber mutation was in a codon that previously specified an amino acid for which no suppressor tRNA was available, then the lysate was used to infect all the *su+* strains.

The mitomycin C-induced lysates from TP292 or TP241,

transformed with plasmids bearing amber mutations in gene *19*, contain several different types of phage: (1) defective, oversized, lysozyme minus phage, (2) defective, lysozyme minus phage that are no longer oversized because they have undergone a deletion either of the inserted Tn5 DNA or a deletion of part of the P22 genome, (3) *19* amber phage (See Fig. 6), and (4) wild type phage. Type 1 are induced P22 Kn321 with no changes. Type 2 are no longer oversized because of a spontaneous deletion of either Tn5 or P22 DNA, but they are lysozyme minus due to the interruption of the coding sequence of gene *19*. Even if Tn5 DNA was deleted, there would still be linker sequences present that would prevent lysozyme from being produced since the correct reading frame would not be restored. Type 2 could also be phage that still contain the Tn5 DNA but have deleted at least 2 kb of phage genome so that they are no longer oversized. Type 3 are phage that have undergone homologous recombination with the plasmid such that the amber mutation was crossed onto the phage. These phage will plate if grown on a suppressor strain that inserts an amino acid that is compatible at the position of the amber; or if they are plated on a strain that provides a source of lysozyme. Type 4 are phage that have undergone homologous recombination with the plasmid at a point such that the amber mutation was not crossed onto the phage genome. These phage will plate on *su-* strains.

When the lysates were plated on suppressor plus strains the selection for the recombinant phage is a double selection based on the lysozyme-plus phenotype and upon the ability of the phage to form a plaque upon single infection. In cases where a source of lysozyme would be supplied (if none of the available suppressors yield functional lysozyme), the ability to form a plaque upon single infection would still in effect, since the phage would still be oversized and thus unable to form a plaque upon single infection. In this manner, the rare recombinant phage present in the induced lysate was easily found. In most cases if 100 µl of a 200 fold dilution of the induced lysate was plated on a *su+* strain, about 50 plaques were formed. All of these plaques represented recombinant phage. To determine if a recombinant contained an amber mutation, the phage was plated on *su-* cells by transferring phage from the plaque on the *su+* lawn to a fresh lawn of *su-* cells using a sterile toothpick. Those phage that failed to grow were considered presumptive recombinant phage that had acquired the plasmid-borne *19*-amber mutation. Normally about half of the recombinant phage contained amber mutations. When 100 µl of 200 fold dilution of induced lysate was plated on TP284 (the strain used to titer the defective particles), there was confluent lysis since all the defective phage could now plate.

The phage bearing amber mutations were plaque purified

by taking a plug of agar containing the phage from the *su+* strain and resuspending it in 1 ml BS with a drop of chloroform. A 0.1 ml portion of a 10^{-4} dilution of this single-plaque lysate was plated on a permissive *su+* strain. An individual plaque was picked, and the single-plaque lysate was used to streak for a single plaque on *su+* permissive cells. The single plaque was used to infect 30 ml of LB broth that had been inoculated with 1 ml of saturated culture of a permissive *su+* strain. This stock culture was incubated at 25^{0}C for 24 hours with aeration. Chloroform (0.5 ml) was added and the culture was further shaken for 20 minutes. The cells and debris were then pelleted by centrifugation for 5 minutes at 5900 g. The resulting lysate was spun at 27,000 g for 90 minutes to pellet the phage. The pellet was resuspended with 2 ml of BS. The phage stock was titered on *su+* permissive cells at 30^{0}C. If there were any variations in the plaques, such as smaller or pinpoint plaques among the rest, the stock would be titered on the lysozyme producing strain (TP320). In this way, I could tell if the *su+* strain I was using for my permissive strain was yielding an EOP of 1. This was necessary since some of the amber mutations were not at codons for which we had suppressor strains. By supplying lysozyme from a plasmid, and thereby complementing the phage in trans, it was possible to determine the actual titer.

TP320 could be used to plate phage which had an amber

mutation which resulted in nonfunctional lysozyme when plated on any of the suppressors. The plasmid present in this strain, pDR105, produces T4 lysozyme from the P_{tac} promoter. This plasmid has virtually no homology at the DNA level with P22, therefore homologous recombination between the plasmid and the phage is nil. This plasmid allows a lysozyme-defective P22 to plate with an efficiency of 1.

I did not have any phage bearing amber mutations which failed to grow on all the suppressor strains, but I tried using TP320 to see how it would work if I did need it. In theory TP320 should work fine but in practice I had an unexpected difficulty. TP320 when infected with P22 results in a high transduction frequency of the resident plasmid (pDR105). This transduction frequency of the plasmid is as high as one normally sees when P22 infects a strain bearing a plasmid that has known homology to the phage. There is no significant homology that we know about between P22 and pDR105. This transduction of pDR105 produces misleading results when screening phage that arise from an infection of TP320. For instance, I plated an induced lysate, from TP241 bearing the plasmid that carried am12, on TP320. The plaques were screened by stabbing into DB7000 and TP320. Phage that were able to grow on TP320 but not on DB7000 were candidates for recombinant phage bearing the amber mutation. There were 4 phage out of 32 tested, that failed to grow on DB7000.

These phage were then stabbed from TP320 into DB7000, TP320 and MS1363. The surprising result was that now only 1 of these 4 phage failed to grow on DB7000. By passaging through TP320 for a second time, three phage acquired the ability to grow on DB7000. To test the possibility that transduction of pDR105 was responsible I prepared single plaque lysates from the DB7000 plate. These lysates were then used to infect DB7000 and infected cells were spread on ampicillin plates. There was a high percentage of transduction of the pDR105 amp^r plasmid. This meant that by growing phage on TP320, I was obscuring my results.

Construction of Seven Amber Mutations. The alleles am12, am65, and am82 were made using single strand pDR392 as the template strand. The transformation of W3110 *lacI*^{q} with the am12 reaction yielded approximately 100 colonies. These were screened for the ability to complement lambda *c*I- *R*am. Of the 100 colonies screened, 2 failed to complement. Those clones were streaked and single colonies were isolated. DNA was recovered from minilysates that were prepared from 5 ml overnight cultures grown in LB plus ampicillin. This DNA was used to transform TP241 to ampicillin resistance. An ampicillin/kanamycin resistant colony was used to inoculate 5 ml of LB plus ampicillin and kanamycin. A 0.5 ml portion of these overnight cultures was used to inoculate 50 ml of LB. This was grown at 37^0C until a cell density of 2 x 10^8/ml was

attained, then 0.4 ml of 0.4 mg/ml solution of mitomycin C was added and the incubation was continued overnight. Chloroform and HEWL was added and the debris was then spun out. The resulting lysate was used to infect TP320. The plaques that arose were plated on TP320 and DB7000 to distinguish the phage which contained amber mutations. I found 1 plaque out of 50 that would grow on TP320 and MS1363 and not on DB7000. The plaque on the MS1363 plate was used to make a single plaque lysate. The phage was plaque purified and a stock was grown on MS1363.

Amber 65 (tryptophan) was made by site directed mutagenesis and also by hydroxylamine mutagenesis. I will describe the amber mutation made by site directed mutagenesis, since this is the allele that I worked with the most (I did not realize until after this allele had been created that one of the chemically induced amber mutations in the collection had the same change, because these amber mutations were not sequenced until I was far along in the site directed mutagenesis part of the project). I chose to work with the site directed mutant, since it was less likely to contain other extragenic mutations and because the phage was isogenic with the rest of my collection.

The transformation of W3110 *lacI*^{q} with the extension ligation reaction for am65 yielded approximately 100 ampicillin resistant colonies. Of these colonies, 2 failed to

complement lambda *c*I-*R*am phage. DNA was prepared from minilysates and was used to transform TP241 to ampicillin resistance. A colony was chosen and used to grow a culture, which was induced with mitomycin C and lysed with chloroform and HEWL. The lysate was supplemented with tail proteins and used to infect MS1363 . The resulting plaques were screened by plating on DB7000. Of 12 plaques, 3 were phage bearing amber mutations. The phage bearing the amber mutations were plaque purified and a stock was grown on MS1363.

After the mutagenesis reaction for am82 (glutamine), W3110 *lac*I^q was transformed to ampicillin resistance. The resulting colonies were screened for the ability to complement lambda *c*I-*R*am. Of 100 colonies, 3 failed to complement. DNA was isolated and used to transform TP241 to ampicillin resistance. A colony was selected This was grown at 37^0C until a cell density of 2 x 10^8 was attained, then mitomycin C was added and the culture continued to aerate for 3 hours. The lysate was recovered after a 15 min. incubation with chloroform and HEWL. This lysate was used to infect MS1363. The resulting plaques were screened for the amber mutation by plating on DB7000. Of 13 plaques, 2 contained amber mutations. They were plaque purified and stocks were grown on MS1363.

Ambers 22 (tyrosine), 61 (glutamate), and 83 (tyrosine) were created using single strand pDR281 as template. Of 48 ampicillin resistant colonies from the am22

reaction, 3 were lysozyme-defective. There were 7 ampicillin resistant colonies from the am61 reaction and 1 was lysozyme-defective. Of the 22 colonies tested from the am83 reaction, 1 was lysozyme-defective. Plasmid DNA was prepared from each of these colonies. A 10 µl sample of plasmid DNA was used to transform TP241. Ampicillin resistant colonies were selected and grown in 50 ml of LB with ampicillin and kanamycin, then induced for four hours with mitomycin C. After addition of chloroform and HEWL, the lysates were cleared of debris and used to infect suppressor strains. The infections were done using 100 µl of the induced lysates and 0.3 ml of cells which were at a density of $3x10^8$/ml. MS1364 was used for am22 and am83 while MS1363 was used for am61. The resulting plaques were screened to see if they would plate on DB7000. Of the 49 plaques tested from the lysate of am22, 3 were contained amber mutations. Of the 49 plaques tested from the lysate of am61, 28 were phage that had amber mutations. The *19*am83(tyr) lysate yielded 42 plaques and 33 of these were phage with amber mutations.

The amber 39 (valine) mutation was created by using a gapped heteroduplex for template as previously. The in vitro reaction mixture was used to transform W3110 and ampicillin resistant colonies were screened for ability to complement lambda *R*-. 1 of the 80 colonies tested failed to complement. This colony was streaked for singles and the singles were retested for noncomplementing activity. They all tested

lysozyme-defective. A single colony was used to grow a 5 ml culture and DNA was prepared from a minilysate.

The DNA was used to transform TP292, the *E. coli* strain bearing P22 Kn321, to amp^r and tet^r. A resistant colony was selected and streaked for singles. A single colony was used to inoculate 50 ml of LB plus ampicillin and tetracycline. This culture was grown to a density of 2 x 10^8/ml and induced with mitomycin C for four hours at 37^0C. Chloroform and HEWL were added and the debris was spun out. The lysate was used to infect TP278. The resultant plaques were plated on MS1362, MS1363, and DB7000. Of the 14 plaques tested, 6 were phage that had incorportated amber mutations. One of these mutant phages was selected and a phage stock was grown on MS1363.

Sequencing the phage bearing amber mutations. DNA was extracted from 0.4 ml of phage stock in a microfuge tube by adding 2 µl diethylpyrocarbonate, 10 µl 10% SDS and 50 µl 2 M Tris-HCl 0.2M Na_2EDTA (pH 8.5), and incubating uncovered at 70^0C in a fume hood for 5 minutes. Protein and SDS were precipitated by addition of 50 µl 5 M potassium acetate, followed by incubation on ice for 30 min, and centrifugation in a microfuge for 10 minutes. The supernatant was transferred to a new tube, and the DNA was precipitated with 2.5 volumes of ethanol. The DNA was spun onto a sealed glass

pasteur pipet and transferred to a clean microfuge tube containing 200 µl precipitation buffer. The tip of the pipet was washed with the buffer repeatedly, until all the DNA went into solution. Following extraction with 3 drops of phenol, the aqueous phase was removed and extracted with 1 ml ether. The ether phase was removed, and the DNA was precipitated with 5 volumes of ethanol, pelleted by centrifugation in a microfuge, dried under vacuum, and redissolved with 50 µl DNA buffer (It does not go into solution quickly; drawing it repeatedly into a pipetman tip helps). The usual yield from this procedure was about 50 ug DNA.

The phage DNA was digested with *Hind*III (25 µl in 50 µl, with 20 units *Hind*III) overnight at 37^0C. The digestion was checked by running 2 µl in an agarose gel. The crudely purified DNA occasionally resists cutting. In these cases, complete digestion can usually be achieved by phenol and ether extraction of the reaction mixture, followed by precipitation, washing and drying as described above, and repetition of the digestion reaction. The reaction mixture can be used directly in the sequencing reactions.

The sequencing primers were purified in the same manner as the mutagenic primers, except the sequencing primers were not kinased before purification. The primers used were:

(1) 5' CTCGATTGGTTCGCT 3'

(2) 5' GTTGGGCATACCGGA 3'

(3) 5' AGTAGTCTTGTTCGC 3'

(4) 5' GCTTTCCTGTTATGG 3'

After gel electrophoresis, elution from the Sep-Pak column, and lyophilization, the DNA was dissolved with 1 ml of DNA buffer and the absorbance at 260 nm was determined. The primers were phosphorylated in a reaction mixture that contained:

1 µl primer at 10 pmol/µl

1 µl DNA buffer

1 µl 10 x kinase buffer

8 µl *gamma*-^{32}-ATP (3000 Ci/mmol, 10 mCi/ml)

1 µl T4 polynucleotide kinase (NEB 8 U/µl)

The reaction was carried out at 37^{0}C for 45 min, then stopped by incubation at 90^{0}C for 3 min. 40 µl DNA buffer was added, the reaction was phenol extracted, followed by ether extraction, and then the ether was boiled off at 90^{0}C. This preparation was used directly for sequencing. It remains usable for at least two weeks at -20^{0}C.

The annealing and primer extension reaction mixture contained:

3 µl kinased primer (0.6 pmol)

2 µl *Hind*III-digested P22 DNA (0.04 pmol)

2.4 µl 5 x AB

4.6 µl water

This mixture was placed in a microfuge tube and spun in a microfuge for 2 seconds. This was incubated at 90^{0}C for

3 min., spun to collect in the bottom of the tube, frozen with dry ice, then thawed on wet ice. 2 µl portions of the annealing mixture were added to tubes labeled G,A,T, and C, each containing 1 µl dNTP mix, 1 µl the appropriate ddNTP, and 1 µl RT mix. These were spun to collect the liquid at the bottom of the tubes, then incubated at 50^0C for 15 min., spun again, then 1 µl more RT mix was added. Incubation continued for an additional 15 min. at 50^0C. The tubes were spun again, and the reaction was stopped by adding 5 µl sequencing dye buffer and heating to 90^0C for 3 min.. After another spin the tubes were placed on ice and 2 µl samples were loaded on a sequencing gel. Sequences were usually readable after a 24 hr. exposure of the dried-down gel to film at -80^0C. An intensifying screen was used.

The *19*am phage were sequenced to confirm the presence of the presumed amber mutation. Three *19*am phage were sequenced by using the chemical modification sequencing protocol of Maxam and Gilbert. These phage; am64, am113, and am122 were described in the previous chapter. The rest of the *19*am phage were sequenced using the dideoxy chain termination protocol.

Phage *19* am61, am65, am81, am82, and am83 were sequenced using primer #2. The codons sequenced were 40 through 102. Phage *19* am12, am22, and am39 were sequenced using primer #1. The sequenced DNA represented codons 99 of gene *13* through codon 50 of gene *19*. In all cases the amber

mutation was confirmed and no other changes were present. The gene *19* alelle am65 was created twice, once by hydroxylamine mutagenesis and once by site directed mutagenesis.

Recombination tests. Recombination experiments were performed between alleles of P22 *19*am phage to determine if mutations were in the same nucleotide or neighboring nucleotides. The experiments were done by spotting 10 µl of BS containing 1 x 10^5 phage onto a freshly poured bacterial lawn. The bacterial lawn consisted of an LB plate with 2.5 ml of soft agar and 0.5 ml of a culture of TP266 at a density of 2 x 10^8. The spot was allowed to dry at room temperature for 5 minutes and then the second phage was spotted directly over the dried spot. The second phage spot was allowed to dry and the plate was incubated at 37^0C overnight. If the phage tested had mutations that were within 3 to 6 bases of the original mutation then the spot would not show signs of lysis. On the other hand if the phage had mutations that were separated by at least 6 bases then recombinant phage would lyse the spot on the bacterial lawn. Negative control spots were parental self-crosses done using 20 µl of the phage suspensions. Positive control spots were done using phage that were known to contain mutations that were separated by at least 6 bases.

Determining Suppression Patterns. Suppression patterns

were determined for *19*am phage by plating serial dilutions of phage stocks with the suppressor strains. The strains used were TP278, TP279, TP280, TP282, TP308, TP320 and TP309. The strains were kept as glycerol stocks at -20^{0}C. A working stock would be prepared by streaking 10 µl of the glycerol stock on an LB plate and incubating it overnight at 30^{0}C. Single colonies were selected from these plates and streaked on minimal plates that were supplemented with either ampicillin (TP278, TP279, TP280, TP320 and TP282) or ampicillin plus tetracycline (TP308 and TP309). The plates were incubated at 30^{0}C for 24 hours. Single colonies were used to inoculate 5 ml of M9 plus and antibiotics. It was important to grow liquid cultures of TP308 and TP309 in M9-glucose minimal medium supplemented with antibiotics because the plasmids that encoded the suppressor tRNA's were readily lost if selection was not maintained. These cultures were aerated at 30^{0}C until they reached a cell density of about 1 x 10^8/ml. The cultures were then used to plate P22 *19*am phage. The infections were done by adding 100 µl of phage suspended in BS to 0.5 ml of the bacterial culture. The phage were allowed to adsorb to the cells for 15 minutes at room temperature and then 2.5 ml of soft agar at 50^{0}C was added, the suspension was gently vortexed and poured onto lambda plates. Infections using TP308 or TP309 were plated on tetracycline-containing LB plates. The soft agar was allowed to harden, about 5 minutes at room temperature and then the plates were incubated overnight at

either 30 or 40^{0}C. Usually 0.1 ml of the dilutions used contained 100 to 1000 phage, unless it was previously known that a given mutant phage would not plate efficiently on a particular suppressor strain. In such a case between 1000 and 100,000 phage were used for the infections.

Results

Site Directed Mutagenesis. Seven amber mutations were introduced into P22 gene *19* by site directed mutagenesis, replacing codons 12 (lysine), 22 (tyrosine), 39 (valine), 61 (glutamate), 65 (tryptophan), 82 (glutamine) and 83 (tyrosine). Three variations of the basic protocol outlined in the Materials and Methods section were employed. The amber mutations were introduced by extension of mutagenic primers hybridized to plasmid-borne gene *19*. Once the mutant plasmids were identified, the mutations were crossed onto the P22 genome. A defective P22 phage which was oversized and *19*- was used for this cross. Recombinant phage were selected and the genes containing the amber mutations in gene *19* were sequenced.

Suppression patterns of phage with gene *19* amber mutations. The suppression patterns of the eleven mutant phage were determined at 30^0C and 40^0C. The following section shows the individual suppression patterns for each of the mutant phages. The ability of each phage to form a plaque is indicated by + and - signs. If the phage forms normal sized plaques with an efficiency of greater than 0.5, relative to its plaque forming ability on TP320 then it is designated by ++; if the phage forms pinpoint plaques or plates with an efficiency of 0.01 to 0.5 then it is designated by +; if the phage plates with an efficiency between 10^{-5} and 0.01 then it is designated by +/-; if the phage plates with an efficiency of less than 10^{-5} then it is designated by -. The suppressor

strains used were TP278(gln), TP279(tyr), TP280(leu), TP282(ser), TP308(phe), and TP309(cys). They are designated below by the amino acid inserted by the suppressor tRNA. See Table 5 for a compilation of all the suppression patterns.

Suppression Pattern for Phage with *19*am12(lys) Mutation.

	glu	tyr	leu	ser	phe	cys
40^0	+	++	++	++	+/-	-
30^0	++	++	++	++	+	+/-

The suppression pattern of P22 *19*am12 (lys) indicates that of the 6 different amino acids substituted at residue 12, the phenylalanine and cysteine substitution have the greatest effects on the function of lysozyme. The substitution of glutamine at 40^0C results in slightly less lysozyme activity since the plaque size is affected but the efficiency of plating is normal. Lysine residues are usually found at a molecular surface, though sometimes they are found in internal salt bridges or are involved in catalysis. It is not surprising that phenylalanine can not substitute for lysine, since phenylalanine is usually found at the inside of protein molecules. Leucine is a large nonpolar residue that is normally found inside, but it can substitute quite well for lysine at this position. The suppression pattern indicates that the substitution of cysteine for lysine results in the most severe defect of the six substitutions.

Suppression Pattern for Phage with *19*am22(tyr) Mutation.

	gln	tyr	leu	ser	phe	cys
40^0	+	++	+	+	++	-
30^0	+	++	++	++	++	-

The suppression pattern for P22 *19*am22 reveals the effect of five different amino acid substitutions for tyrosine. Of the five substitutions, only phenylalanine yields a protein that is comparable to wild type at 30^0C and 40^0C.

Suppression Pattern for Phage with *19*am39(val) Mutation.

	gln	tyr	leu	ser	phe	cys
40^0	++	++	++	++	++	+/-
30^0	++	++	++	++	++	+

The suppression pattern of P22 *19*am39 (val) indicates that five of the six substituted amino acids yield functional lysozyme. The substitution of cysteine for valine yields a protein with reduced activity. It is possible that the cysteine forms crosslinks with other cysteines blocking formation of the normal structure. This is unlikely though since the interior of the bacterial cell is thought to have a reducing environment. This position does not appear to be very critical to lysozyme.

Suppression Pattern for Phage with *19*am61(glu) Mutation.

	gln	tyr	leu	ser	phe	cys
40^0	++	++	++	++	++	+
30^0	++	++	++	++	++	++

The suppression pattern for P22 *19*am61 shows that this glutamate is not a critical residue for the protein. Five of the six substitutions yield a lysozyme with normal activity.

Suppression Pattern for Phage with *19*am64(gln) Mutation.

	gln	tyr	leu	ser	phe	cys
40^0	++	++	++	++	++	++
30^0	++	++	++	++	++	++

The suppression pattern for P22 *19*am64 (glu) demonstrates that the glutamine at position 64 is not a critical residue for P22 lysozyme. All five substitutions yield functional lysozyme. The cysteine in this case does not seem to interfere with the protein structure or function as is the case with most of the other positions studied.

Suppression Pattern for Phage with *19*am65(trp) Mutation.

	gln	tyr	leu	ser	phe	cys
40^0	++	++	++	++	++	+
30^0	++	++	++	++	++	+

The suppression pattern determined for P22 *19*am65

(trp) indicated that this residue is another noncritical residue for P22 lysozyme. Only the cysteine has a significant effect on the activity of the protein. The suppression pattern was the same for both the site directed and the chemically induced amber mutations at this site.

Suppression Pattern for Phage with *19*am80(gln) Mutation.

	gln	tyr	leu	ser	phe	cys
40^0	++	-	++	++	-	-
30^0	++	-	++	++	-	+/-

The suppression pattern for P22 *19*am80 (gln) demonstrates that this position in P22 lysozyme has some constraints as to what type of residues will be tolerated. Once again the cysteine substitution is not tolerated well. Additionally there is an intolerance for the two aromatic amino acids. It is possible that the lack of flexibility or size of the aromatic rings is unacceptable at this position.

Suppression Pattern for Phage with *19*am82(gln) Mutation.

	gln	tyr	leu	ser	phe	cys
40^0	++	-	++	+/-	-	-
30^0	++	-	++	+	-	-

The amber mutation at position 82 (gln) of P22 *19* allows us to investigate the effect of five different amino acid substitutions at this position. Of these five

substitutions, only leucine yields a functional protein at high and low temperature.

Suppression Pattern for Phage with *19*am83(tyr) Mutation.

	gln	tyr	leu	ser	phe	cys
40^0	+	++	-	-	++	-
30^0	++	++	+	+/-	++	-

The suppression pattern of P22 *19*am83 (tyr) indicates that only the aromatic phenylalanine can be substituted for tyrosine and still yield a completely functional protein. A glutamine substitution yields a temperature sensitive protein.

Suppression Pattern for Phage with *19*am113(gln) Mutation.

	gln	tyr	leu	ser	phe	cys
40^0	++	++	++	++	+	+
30^0	++	++	++	++	++	+

The glutamine at position 113 of P22 lysozyme does not appear to be a critical residue since all five substitutions yield effective proteins. The phenylalanine and cysteine substitutions cause a change in plaque size but not in the efficiency of plating. This indicates that the protein is not as efficient as wild type but does still function. The phage bearing the amber muation at position 133 also has an amino acid change at position 112 (see chapter 2). The question of the importance of the change at position 112 that resulted in

a tyrosine codon being changed to a serine codon was investigated. Five other independently isolated amber mutations mapped to the same location in gene *19* by recombination tests *19*am-H736, *19*am-H1195, *19*am-H1212, *19*am-H1299, and *19*-amH1344. All of these phage have the same suppression patterns as P22 *19*am113 that was sequenced. It would be highly unlikely that all had undergone the same type of double mutation.

Suppression Pattern for Phage with *19*am122(trp) Mutation.

	gln	tyr	leu	ser	phe	cys
40^0	++	++	++	++	+	-
30^0	++	++	++	++	++	+

The suppression pattern for P22 *19*am122 (trp) indicates that at 30^0C five of the six amino acid substitutions yield lysozymes with normal activity. The phenylalanine and cysteine substitutions are temperature sensitive.

Discussion

By examining the phenotypes of proteins with known substitutions one can determine which residues of the protein are neutral, in the sense that a substituted amino acid has no discernible effect on the protein's function. The functional

test used in this study is the ability of P22 to form a plaque of normal size and with normal efficiency. We have some idea of how much lysozyme activity is needed to form a plaque. A hybrid P22 which has bacteriophage T4 lysozyme gene substituted for the P22 lysozyme gene forms a normal plaque. The amount of lysozyme produced by this hybrid during a single cycle of infection was compared to the amount produced by wild type P22. Wild type P22 produced 30 fold more lysozyme activity. Since the hybrid phage makes a wild type size plaque, then we can assume that P22 normally produces excess lysozyme. In order to see an effect of an amino acid substitution we must reduce the specific activity more than 30 fold (L. Hardy, personal communication).

From the suppression patterns presented in Table 5, it is apparent that the majority of the positions studied for P22 lysozyme can tolerate amino acid substitutions. Of 60 possible substitutions at 11 positions, only 18 reduced the efficiency of plating of the phage at 30^0C. When the temperature was raised to 40^0C, 7 more substitutions reduced plating efficiency. Only 7 of the 60 possible substitutions result in a defect in lysozyme severe enough that the phage plates with the same efficiency on su+ as on su- when the plating is done at 30^0C. This demonstrates a large degree of tolerance, since the substituted amino acid side chains are branched and nonpolar (leucine), large, aromatic, and nonpolar

(phenylalanine), polar and neutral (cysteine, serine, and glutamine), and polar and aromatic (tyrosine). Since amino acids such as leucine or phenylalanine, which are usually found buried within a protein, affect the function of lysozyme very little when substituted for amino acids that are normally found on the molecular surface of a protein, such as lysine (position 12) or glutamate (position 61), it appears that a protein can make adjustments for very different types of amino acids at certain positions.

By examining the frequency of amino acid changes between corresponding proteins of different species "exchange groups" have been defined (G.E. Schulz and R.H. Schirmer, 1984). The amino acids within such a group exchange preferentially with each other. There are four exchange groups: 1. aromatic side chains-phenylalanine, tyrosine, and tryptophan. 2. positively charged side chains- lysine, arginine, and histidine. 3. large aliphatic nonpolar side chains-valine, leucine, isoleucine, methionine, and cysteine. and 4. small side chains regardless of polarity- threonine, aspartate, asparagine, glycine, alanine, glutamine, glutamate, and proline. If one examines the substitutions that were neutral (had less than a 30 fold decrease in activity) in P22 lysozyme, it becomes obvious that amino acid substitutions from different exchange groups can be tolerated, even when a substitution from the same exchange group is not acceptable.

For instance, both cysteine and valine belong to the same exchange group, yet when the *19*am39 (val) is substituted with a cysteine, a marginal lysozyme results. Yet this valine can be substituted with members of the other three exchange groups (glutamine, tyrosine, serine, and phenylalanine) with no reduction in activity.

These examples demonstrate the necessity of a more comprehensive study of protein structure. The rules that have been formulated to date clearly have many exceptions.

Possibly by undertaking a more extensive study of which substitutions are allowed at any given position in a protein, rules could be deduced that would have fewer exceptions. A survey, such as the one outlined here, and the one done on *lac* repressor (Miller et al, 1979) could be easily broadened using the methods developed in this survey.

The site directed mutagenesis procedures outlined allow one to create an amber mutation at any given location in a gene. A complementation test to distinguish the mutant transformant was helpful since the efficiency in recovering the mutation was only in a range of 1 to 10%. This efficiency can be increased using a gapped circular ds DNA molecule as template in the extension ligation reaction (L. Hardy, personal communication).

Amber mutations allow one to insert any amino acid for which a suppressor tRNA is available. Suppressor tRNAs for all the amino acids might be available in the near future since

Abelson and Miller are currently trying to construct in vitro all the possible suppressors (Normanly et al 1986). This would mean that all substitutions could be examined just by growing an phage bearing an amber mutation on a battery of suppressor strains.

TABLE 4

PLASMIDS

Plasmids

pTP289 Ligation of the *Hind*III-*Bam*HI replication origin fragment from pBR322 with the *Bam*HI-*Nae*I fragment from pDR100 containing the carboxy-terminal portion of gene *19*, and a fragment containing the amino-terminal portion of gene *19* generated by digesting pDR110 with *Eco*RI, filling in the ends, and digesting with *Hind*III. A *Bgl*II linker (5'CAGATCTG 3') is inserted between the *Nae*I and filled-in *Eco*RI ends. This plasmid was constructed by D. Herrick.

pDR321 This plasmid was used to construct a P22 defective lysogen. Insertion of the large internal *Hpa*I fragment of Tn5 into the site created by digesting pTP289 with *Bgl*II and filling in.

pDR281 This plasmid was used as template for in vitro mutagenesis. Ligation of *Eco*RI-*Bam*HI lysozyme containing fragment from pDR118 and *Eco*RI-*Bam*HI origin containing fragment of pZ152 (Berman).

pDR392 This plasmid was used as template for in vitro mutagenesis. Three fragments ligated together: 1. *Cla*I (filled in)-*Bam*HI lysozyme-containing fragment of pDR116; 2. *Pst*I-*Pvu*II P_{lac} fragment from pTP30; 3. *Pst*I-*Bam*HI M13 origin of replication containing fragment from pTP352.

pDR463 This plasmid produces Phe sup tRNA and is tet^R and amp^R. pGFIB:Phe (Normanly et al, 1986) was digested with *Cla*I and filled in then ligated with *Eco*RI (filled in)-*Pvu*II tet^R fragment from pBR322.

pDR464 This plasmid produces Cys sup tRNA and is tet^R and amp^R. pGFIB:Cys (Normanly et al, 1986) was digested with *Cla*I and filled in then ligated with *Eco*RI (filled in)-*Pvu*II tet^R fragment from pBR322.

pDR401 A plasmid used in gapped circle mutagenesis. It is tet^R, amp^S, $lyso^+$. pDR392 *Eco*RI-*Bam*HI (both ends filled in) fragment containing lysozyme gene was ligated into the filled in *Eco*RI site of pBR322 bearing an ochre mutation in the *β-lactamase* gene.

TABLE 5

SUPPRESSION PATTERNS OF PHAGE BEARING AMBER MUTATIONS IN GENE *19*

	gln	tyr	leu	ser	phe	cys
am12	+	++	++	++	+/-	-
(lys)	++	++	++	++	+	+/-
am22	+	++	+	+	++	-
(tyr)	+	++	++	++	++	-
am39	++	++	++	++	++	+/-
(val0	++	++	++	++	++	+
am61	++	++	++	++	++	+
(glu)	++	++	++	++	++	++
am64	++	++	++	++	++	++
(gln)	++	++	++	++	++	++
am65	++	++	++	++	++	+
(trp)	++	++	++	++	++	+
am80	++	-	++	++	-	-
(gln)	++	-	++	++	-	+/-
am82	++	-	++	+/-	-	-
(gln)	++	-	++	+	-	-
am83	+	++	-	-	++	-
(tyr)	++	++	+	+/-	++	-
am113	++	++	++	++	+	+
(gln)	++	++	++	++	++	+
am122	++	++	++	++	+	-
(trp)	++	++	++	++	++	+

Table 5. Suppression pattern of P22 phage bearing amber mutations in gene *19*. The amino acids inserted by TP278 (gln), TP279(tyr), TP280 (leu), TP282 (ser), TP308 (phe), and TP309 (cys) are shown across the top. The *19* alleles are shown on the left side with the original amino acid at that position shown in parenthesis. The top line for each allele represents titers at 40^{0}C. The bottom line represents titers at 30^{0}C. The symbols represent titers on the suppressor tRNA strain relative to the titers on a lysozyme-producing strain (TP320). A ++ symbol indicates normal plaque size and a ratio of titer of su+/ TP320 of greater than 0.5. A single + sign indicates pinpoint sized plaques or a ratio of between 0.01 and 0.5. A +/- symbol indicates an a ratio of titers between 0.0001 and 0.01. A minus symbol means that the titer on the su+ strain was the same as on the su- strain.

FIGURE 5

CONSTRUCTION OF P22 Kn321

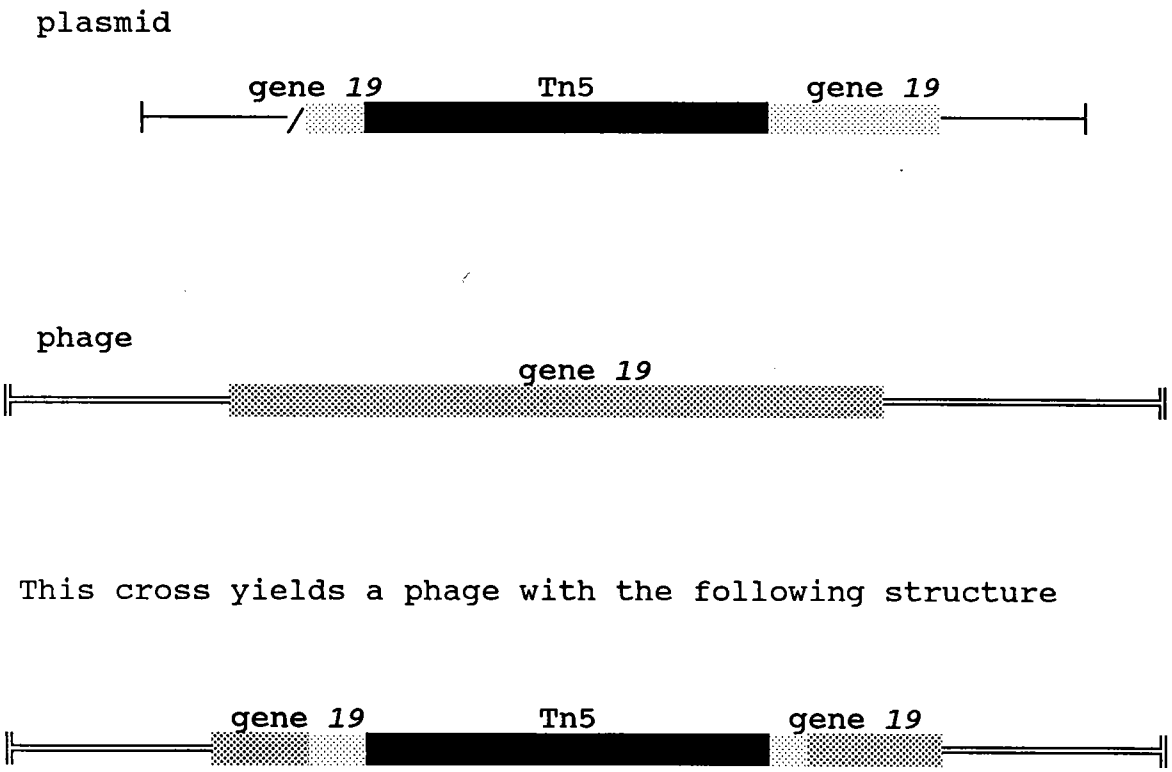


Figure 5. Homologous recombination between plasmid and phage to create a defective P22 phage. The Tn5 insertion is 5 kb and it interrupts the coding sequence of gene *19*.

FIGURE 6

RECOMBINATION BETWEEN AMBER-BEARING PLASMIDS AND P22

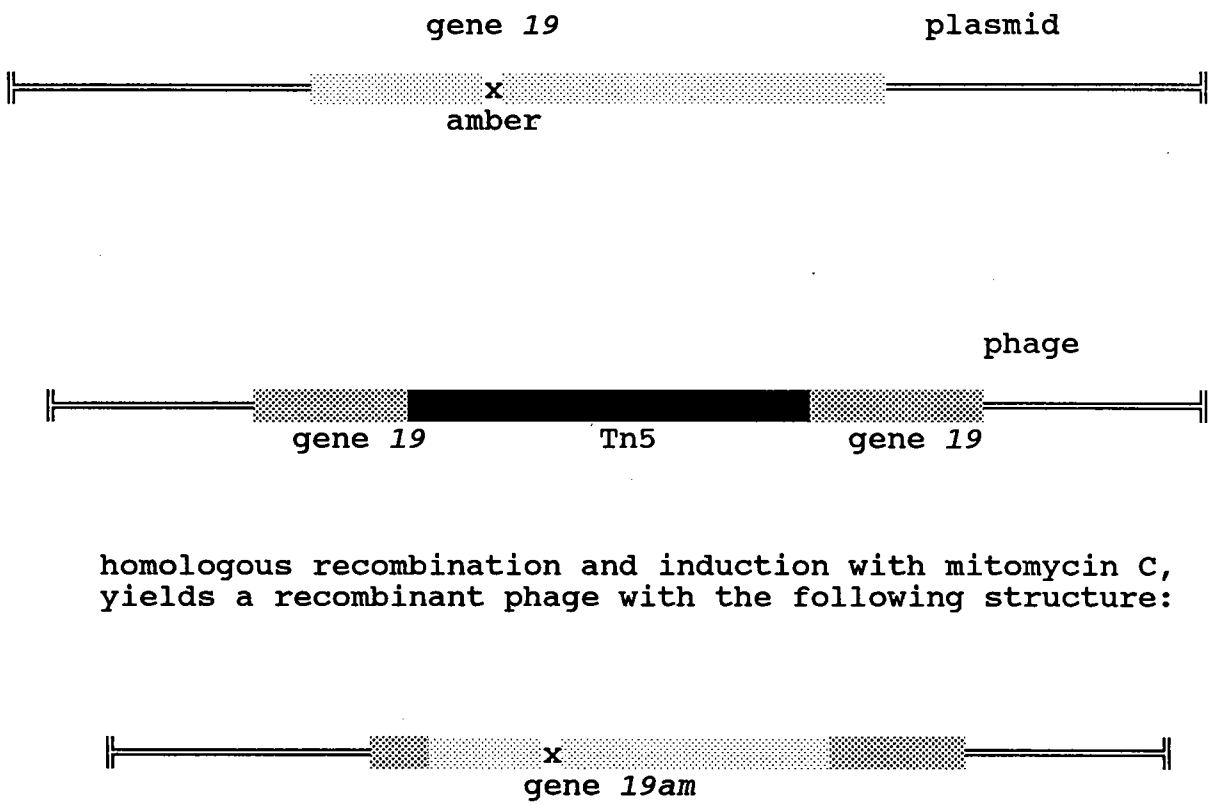


Figure 6. The recombination between the amber bearing plasmid and the defective prophage to yield an amber bearing phage. The phage has lost the Tn5 insertion and has acquired the amber mutation. If the first crossover point occurred 3'to the amber mutation then a wild type phage would result.

FIGURE 7

GAPPED CIRCLE MUTAGENESIS

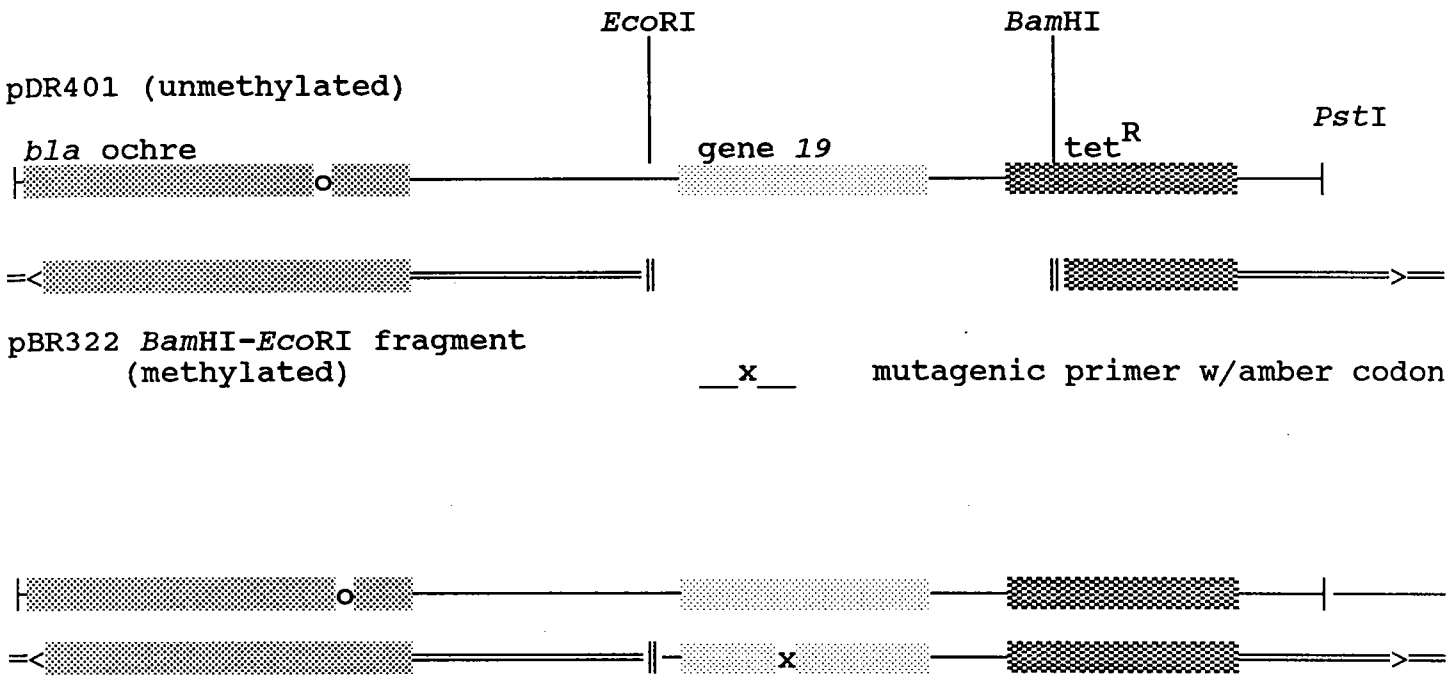


Figure 7. Gapped circle mutagenesis. The substrates shown on the top half of the page are denatured and reannealled. An extension ligation reaction is done resulting in the structure shown on the bottom of the page. This heteorduplex has mismatches in *β-lactamase* and in gene *19*. Since the strand containing the amber in *19* is methylated, then upon transformation the cell's repair system should selectively correct to this sequence, thus yielding a double strand plasmid with a functional amp and tet resistance gene and an amber in *19*.

The other structures formed will be parental and have either an ochre in *β-lactamase* or a truncated tet^R gene. These will not be viable when selected on amp + tet plates.

CHAPTER IV

SECOND SITE REVERTANTS OF MISSUPPRESSED P22 GENE *19* AMBER MUTATIONS

CHAPTER IV

SECOND SITE REVERTANTS OF MISSUPPRESSED P22 GENE *19* AMBER MUTATIONS

Introduction

This chapter describes how second site intragenic revertants of "missuppressed" amber mutations in P22 gene *19* were isolated and characterized. I use the term "missuppressed" to designate phage bearing amber mutations that, when used to infect a strain containing a particular amber suppressor, fail to form normal sized plaques at normal efficiencies due to detrimental affects of amino acid substitutions at the position of the amber mutation in P22 lysozyme.

As discussed in the previous chapter, the 60 amino acid substitutions at 11 different positions of P22 lysozyme yielded a considerable amount of information about which residues were critical to the structure/function of lysozyme. As interesting as that information is, it does not allow one to understand how a protein assumes or stabilizes its tertiary structure. Since all the information most proteins need to obtain their final structures is contained in the primary amino acid sequence, it stands to reason that these blueprints must operate by specific interactions between the amino acids. By studying which interactions occur, understanding of the blueprints might come about. One approach to deciphering these interactions is to identify individual interactions that are critical for the protein's structure/function. By

isolating second site intragenic revertants of missuppressed amber mutations, pairs of amino acids that interact with one another in direct or indirect fashions are identified.

The survey of amber mutations in P22 lysozyme yielded 18 combinations of site and suppressor that resulted in reduced or undetectable lysozyme activity. Revertants that restored activity were isolated from these missuppressed amber mutations. The revertants were screened to determine whether the site of the reversion was primary or secondary. This screen, which is critical to the success of the project, was fast and efficient due to the nature of the primary mutations. The revertants of missuppressed amber mutations either still contain amber mutations (second site revertants) or no longer contain the amber mutation (primary site revertants). Hence, if the revertant phage requires a suppressor tRNA to grow, the reversion was a second site reversion. On the other hand, if the revertant phage can grow in the absence of a suppressor tRNA, the reversion event was a primary site reversion. Secondary site revertants were identified and the the sites containing the mutations were sequenced. Some interactions that take place between amino acids of P22 lysozyme were identified.

Materials and Methods

Plasmids. Plasmids were constructed using standard in vitro DNA techniques (Maniatis et al., 1982). Plasmids pDR469, pTP400, and pTP399 are described in detail in Table 6. Plasmid pDR469 was used in construction of a defective P22. This plasmid has a 5.04kb fragment of Tn5 DNA encoding kanamycin resistance inserted between P22 DNA that consists of the 3' half of gene *13*, the first 9 codons of gene *19* and the last 13 codons of gene *19* with 250 bases of DNA 3' of gene *19*. Other plasmids were described in chapter 3.

Enzymes, chemicals, media, and buffers. The enzymes, chemicals, media, and buffers are as described in Chapter 3.

Construction of P22 *sie*A44 *m44* Kn469. A defective P22 prophage that lacked gene *19* was used to cross mutant alleles from plasmids onto unmutagenized P22. Plasmid pDR469 was used to transform *Salmonella su- r- m+* (MS1868) to kanamycin resistance. This strain was then grown to a cell density of 1 x 10^8/ml in LB broth suplemented with ampicillin and kanamycin at 37^0 C. The culture was then infected with P22 *sie*A44 *m44* at a MOI of 10 and aerated for 2 hours at 37^0C. Chloroform was added and the debris was removed by centrifugation at 5900 g for 5 minutes. The resulting lysate was titered on TP284. A

Salmonella polA-ts strain (CV112) grown to a cell density of 1 x 10^8/ml in LB broth at 30^0 C was infected with the lysate (MS1868/pDR469 + P22 *sie*A44 *m*44) at a MOI of 10, the phage were allowed to adsorb for 15 minutes and the infected cells were spread on kanamycin LB plates and grown at 37^0C. The colonies were screened for ampicillin sensitivity, P22 immunity, and the ability to yield plaque-forming phage. Colonies which were amp^s, P22 immune, and could not yield phage, were selected. The procedures used for the above tests are described in Chapter 3. A liquid culture of the selected clone was grown in LB broth supplemented with kanamycin to a cell density of 1 x 10^8/ml at 37^0 C and then mitomycin C was added and the culture continued to be aerated for 3 hours at 37^0 C. A lysate was obtained after treating the cells with chloroform and HEWL and then centrifuging at 5900 g for 5 min, to spin the cell debris out. The lysate was used to infect a *Salmonella* F'-bearing strain (MS1387 [del *att*P22] *cys*B *his leu*D *fol*101 \ F' *lac pro*), in which the only P22 attachment site is episomal. Lysogens were selected by spreading the infected cells on kanamycin plates. The colonies were screened for P22 immunity and inability to yield plaque-forming phage. A colony that passed these tests was used to mate with *E. coli* (W3110). The defective P22 prophage was thus transferred to an *E. coli* strain. The lysogen was selected on minimal plates supplemented with kanamycin.

Colonies were tested in the same manner as described above.

Isolation of Revertants. Revertants of missuppressed amber mutations were isolated in three different ways. Spontaneous revertants were recovered by plating phage bearing amber mutations on suppressor strains that yielded lysozyme that had less than wild type activity. The number of phage plated to detect spontaneous revertants depended upon the plating efficiency of the given phage on that particular suppressor strain. For instance, if a *19*am phage stock had 1 x 10^{12}pfu/ml but would only form plaques on a given suppressor strain at 1 x 10^{-6}, then about 2 x 10^{8}pfu would be plated. This would yield about 200 plaques. On the other hand, if a stock had 1 x 10^{12}pfu/ml but would plate with an efficiency of 1 on a given suppressor strain, forming pinpoint sized plaques, then 500 phage would be plated.

Any larger plaques from the background of pinpoints would be screened for primary or secondary site reversion. Such plaques would be replated by transferring phage from the center of the plaque to plates that had fresh bacterial lawns of the suppressor strain, a suppressor-minus strain and TP320. If the plaque being tested yielded phage that failed to grow on a suppressor-minus strain, grew on TP320, yet plated better than the original mutant phage on the suppressor plus strain, then it was a candidate for a second site revertant. The T4 lysozyme-producing, suppressor-minus strain TP320 was used to

screen for phages that had reverted the original amber mutation in gene *19*, yet picked up other amber mutations in other essential P22 genes. These phages will not grow on TP320.

Only combinations of mutant phage and suppressors strains that exhibited plating efficiencies of less than 1×10^{-3} as compared to the same phage's titer on a permissive strain were useful for finding spontaneous revertants.

Revertants were also isolated from phage that had undergone UV mutagenesis. The UV irradiation was performed by suspending 1×10^{10} phage in 1 ml of BS in an open glass petri dish and exposing it to 3500 ergs/mm^2 UV light at 254 nm as measured by "Black Ray" dosimiter from Ultra Violet Products. The phage were transferred to foil wrapped glass tubes and stored at 4^0C. This treatment typically reduced the titer of the phage by a factor of 10^4 to 10^5, with 1-4% clear plaque formers among the survivors. The UV-irradiated phage were plated (enough live phage to yield no more than 200 plaques) on UV-stimulated (300 ergs/mm^2) suppressor strains which bore pKM101 amp^R (Youderian et al., 1982). This is the procedure I shall refer to as "direct plating". An alternative method was "cycling" or passaging phage though one suppressor strain and using the lysate to infect another suppressor strain. Cycling was done by infecting 1 ml cultures pKM101 amp^r-bearing, UV-stimulated suppressor cells at a density of 1×10^9/ml, with 1

$x10^9$ UV-irradiated phage, allowing 20 minutes for adsorption, then diluting into 50 ml of LB broth and growing at 30^0C or 37^0C for 2-6 hours. Chloroform (1 ml) and HEWL was added and the debris was pelleted by centrifugation at 5900 g for 5 min. The resulting lysate was plated on missuppressing suppressor strains on lambda plates. Revertants were screened in the same manner as previously described. Strains bearing pKM101 were UV-stimulated before infection so that *muc*A and *muc*B would be induced. The induction of these genes produced 10-40 fold increases in the titer of the UV-irradiated phage stocks. This reactivation was dependent on the presence of the pKM101 plasmid. Appendix 1 shows typical results of UV-stimulated mutagenesis.

Sequencing the Revertants. The sequencing was done in the manner described in the previous chapter. The parent phage bearing an amber mutation in gene *19* was sequenced side by side with the revertant phage. All revertants were sequenced from 32 bases 5' of the start site of gene *19* to 25 bases 3' of gene *19*. The original amber codon was confirmed in all cases.

Backcrossing the Mutations. The mutations present in gene *19* were subcloned and then backcrossed to an unmutagenized P22. DNA was extracted from 0.4 ml of phage stock by the procedure outlined in Chapter 3. The DNA was

digested with *Hpa*I endonuclease in a reaction that contained 50 µl DNA (0.5 µg/µl), 14 units of *Hpa*I, 10 µl 10 x *Hpa*I buffer and 47 µl of H_2O. The digestion was carried out at 37^0C for four hours and then checked on 0.7% agarose gels. If the digestion went to completion, the reaction was terminated by phenol extraction, ether extraction, and ethanol precipitation. The DNA was resuspended in 40 µl of DB, and after the addition of 5 µl sample buffer, the mixture was electrophoresed on a 5% bis-acrylylcystamine-crosslinked polyacrylamide gel (Hansen, 1981) at 150 volts until the bromophenol blue migrated to the end of the gel. The *Hpa*I fragment of 1.1 kb (band G) that contained gene *19* was excised from the gel and the DNA was extracted following the protocol of Hansen (1981). The *Hpa*I fragment was ligated with pBR322 that had been digested with *EcoR*V. The ligation reaction mixture was used to transform W3110 to ampicillin resistance. The colonies were checked for tetracycline sensitivity. Tetracycline sensitive colonies were picked and 5 ml liquid cultures were grown in LB plus ampicillin. DNA was isolated from minilysates and the size of the plasmids was determined by electrophoresis on a 0.7% agarose gel using pBR322 as a size standard. Plasmids that were larger than pBR322 by about 1 kb were identified.

These plasmids were then used to transform TP315 to

ampicillin resistance. TP315 is W3110 with a P22 *sie*a44 *m*44 Kn469 prophage. A clone was selected and grown in 50 ml of LB until a cell density of 2 x 10^8/ml was reached and then the culture was induced with 40 µl of mitomycin C (0.4 mg/ml) for three hours at 37^0C. Chloroform and HEWL was added and the aeration at 37^0C continued for 20 minutes. The debris was spun out by centrifugation at 5900 g for 5 minutes.

The resulting lysates with 1 x 10^{10}/ml P22 tails added were used to infect *Salmonella* suppressor strains which were permissive for the amber being tested. The infections were done using 0.1 ml of the lysate or 0.1 ml of a 100-fold dilution of the lysate and 0.5 ml of cells at 2 x 10^8/ml. To dertermine linkage of the secondary phenotype 50 single plaque lysates were made from the plaques that resulted. Dilutions of the single plaque lysates were spotted on suppressor-minus cells, suppressor cells that missuppressed the original mutant phage, and permissive suppressor cells. The spots were scored for the secondary phenotype. Suppression patterns and plaque morphologies of these reconstructed second site revertants were then compared to those of the original second site revertants and to the original mutant phage.

Results

All combinations of P22 *19*am phage and suppressor strains that resulted in reduced lysozyme activity were screened for second site reversion. It became apparent that second site reversion was much rarer than primary site reversion. This result was presumably not due to any particular targeting of the amber codons by the mutagenesis procedure, since UV mutagenesis used in conjunction with pKM101 has the broadest spectrum of any mutagen available (Coulondre and Miller, 1977; Youderian et al., 1982). Therefore, I feel that the scarcity of second site revertants for certain alleles reflects the lack of sites that can compensate for the original substitutions. The relative frequencies of primary and secondary site revertants among plaque-formers on the missuppressing hosts varied greatly, second site revertants represented only 0.1-20% of total revertants depending upon which allele was being reverted (see Table 2).

Spontaneous Revertants. The first method used to isolate second site revertants was by plating unmutagenized *19*am phage directly with the missuppressing strains. Spontaneous revertants that grew were screened for primary or second site reversion. Since reversion rates for most of the *19*am phage ranged between 10^{-5} or 10^{-6}, and no more than 1000

plaque-forming phage could be plated on a single plate, the only combinations that were tried in this fashion were ones that had an EOP of less than 0.001. All the revertants were primary site revertants. The combinations and the results were as follows: The ratios represent second site revertants/primary site revertants.

		suppressor strains		
phage		TP280(leu)	TP308(phe)	TP309(cys)
am12(lys)	40^{C}_{0}	0/0		0/5
am22(tyr)	$30^{0}C$			0/50
am22(tyr)	$40^{0}C$			0/50
am39(val)	$40^{0}C$			0/0
am80(gln)	$40^{0}C$	0/50	0/5	

Normally on any given plate I would have between 5 and 20 revertants for the combinations given above. If I tried to plate more phage per plate, the bacterial lawn would die before the phage grew. The number of phage per plate depended upon the combination of phage and suppressor being plated, but the normal range for the alleles described above was between 10^5 and 10^7 phage per plate.

Revertants Obtained by Direct Plating of UV-Irradiated Phage. Revertants were isolated after UV-irradiation and direct plating. These conditions made it possible to recover revertants with less plating. The most common problem I encountered in these experiments was that the "dead" phage in the dilutions would inhibit the growth of the bacterial lawn.

This was probably due to the large number of phage particles adsorbing to the bacterium, even though the particles had defective genomes due to the mutagenesis or the formation of lysogens. When the reversion rate of the amber mutation was high enough, then revertants could be recovered by plating less total (dead or alive) phage and the bacterial lawn would be intact. A summary is given in table 7.

P22 *19*am12(lys) was plated on TP278(gln), TP308(phe), and TP309(cys). The TP308(phe) and TP309(cys) platings yielded about 50 plaques each, and all were primary revertants. Obtaining revertants on TP278(gln) was more difficult since the parent made a small plaque. Only 32 revertants were tested and all were primary revertants.

P22*19*am22 (tyr) was plated on TP278(gln), TP280(leu), TP282(ser), and TP309(cys). All revertants tested (350 each condition except for TP309(cys), for which only 100 were tested) were primary or primary plus a secondary site extragenic revertants. The secondary site extragenic revertants were ambers in other P22 essential genes that would grow on the missuppressing strains. These were isolated in combination with primary revertants in gene *19*. They were distinguished from intragenic secondary site revertants by plating on TP320. Because this strain is su- but lysozyme+, it does not plate the extragenic ambers.

P22 *19*am39(val) was plated on TP309(cys) and 14 of 50 revertants were secondary site revertants. Two of these

secondary site revertants (*19*am39sr575 and *19*am39sr581) were sequenced and both had changes only in the ribosome binding site of gene *19*. These secondary site revertants were not further characterized.

I did not try to isolate revertants for P22 *19*am61(glu), *19*am64(gln), and *19*am65(trp) since the suppression patterns indicated that these phage grew fairly well in the suppressor strains.

Revertants were isolated from P22 *19*am80(gln) on TP279(tyr), TP308(phe), and TP309(cys). From each condition 100 revertants were screened. All were primary revertants.

P22*19*am82(gln) was plated on TP279(tyr), TP282(ser), TP308(phe), and TP309(cys). On TP279(tyr), 750 revertants were screened they were all either primary or primary plus extragenic revertants. There were 600 revertants screened from platings on TP282(ser), all were primary site revertants. I screened 50 revertants from TP308(phe) and TP309(cys). These were all primary site revertants.

Revertants of P22 *19*am83(tyr) were obtained on TP278(gln), TP280(leu), and TP282(ser). Of the 500 revertants screened from TP278(gln) all were primary revertants. Of the 350 revertants from TP280(leu) all were primary revertants. There were 200 revertant from TP282(ser) that were screened and all were primary revertants.

Revertants Obtained by "Cycling" UV-Irradiated Phage. Phage were subjected to UV mutagenesis and then used to infect UV-stimulated permissive or semi-permissive suppressor strains. The infected cells were incubated for at least 2 hours so that multiple rounds of infection would take place. Lysates were prepared from these cultures and used to infect missuppressing strains. Revertants that arose on the missuppressing strains were screened for primary or secondary site reversion. Revertants obtained in this manner are not independent; therefore when secondary site revertants were found in a given lysate, only one would be characterized, unless there were differences in plaque morphology between individual revertants. A summary is given in Table 8.

Phage *19*am22(tyr) was passaged on TP282(ser) and TP278(gln). The lysates were plated on TP280(leu) and TP278(gln). Revertants were screened for secondary site reversion. The lysate passaged though on TP278(gln) yielded only primary site revertants, 150 on both TP280(leu) and TP278(gln). The lysate passaged on TP282(ser) yielded second site revertants when plated on TP280(leu). There were 16 second site revertants from the 50 plaques that were tested. Of the 16 second site revertants, 2 were picked for further study: *19*am22(tyr)sr517 and *19*am22(tyr)sr520. Only primary revertants (18) were found when the lysate was plated on TP278(gln).

Phage *19*am82(gln) was irradiated with UV and then passaged on TP282(ser), TP279(tyr), and TP278(gln). Phage lysates were prepared from these infected cultures and used to plate for revertants on TP279(tyr) and TP282(ser). From the lysate passaged on TP279(tyr) which was plated on TP282(ser), 50 revertants were screened, all of which were primary revertants. When this phage was passaged on TP278(gln), a permissive strain, 50 revertants were selected on both TP279(tyr) and TP282(ser). All were primary site revertants. A lysate passaged on TP282(ser) was used to select revertants on TP279(tyr) and TP282(ser). The 150 revertants on TP279(tyr) were primary revertants. From the 100 revertants on TP282(ser) there were 5 secondary site revertants. The plaque morphology of at least three of the secondary site revertants differed enough from each other to consider them individually (*19*am82(gln)sr512, *19*am82(gln)sr513, and *19*am82(gln)sr516). Two additional secondary site revertants were chosen: *19*am82(gln)sr514 and *19*am82(gln)sr515.

Phage *19*am83(tyr) was passaged on TP278(gln) and TP280(leu). The lysate passaged on TP278(gln) was used to infect TP280(leu). Two hundred revertants were screened; 29 secondary site revertants were found. The one chosen to characterize was *19*am83(tyr)sr542. This lysate was also used to infect TP278(gln). Fifty revertants screened, and 18 were secondary site revertants. The lysate passaged on TP280(leu) was used to infect both TP280(leu) and TP278(gln). Fifty

revertants were screened for both conditions, 7 secondary site revertants were found out of the 50 tested when the lysate was plated on TP278(gln). Of the 50 plaques tested from the TP280(leu) plates, 12 were secondary site revertants. One revertant was chosen to characterize: *19*am83(tyr)sr532.

The collection of secondary site revertants isolated consists of *19*am82(gln)sr512, *19*am82(gln)sr513, *19*am82(gln)sr514, *19*am82(gln)sr516, *19*am22(try)sr517, *19*am22(tyr)sr520, *19*am83(tyr)sr532, *19*am83(tyr)sr542, *19*am39(val)sr575, and *19*am39(val)sr581.

Suppression patterns of the second site revertants. The suppression patterns of the second-site revertants are shown in Table 9. The changes seen in all of the secondary site revertants are improvements in the plaque-forming ability of the phage. In no case did a second site change decrease the phage's ability to form a plaque with a given substitution. In theory, one could expect some secondary changes to improve the plating ability of the phage on the missuppressing strain but at the same time decrease the ability of the phage to plate on a previously permissive strain. It was thought that a change which increased and decreased the plaque-forming ability of the phage, depending upon which amino acid was being substituted, would demonstrate that the secondary site change was an allele-specific change

and not just a change which increased the protein's overall stability. The suppression patterns of the secondary site revertants indicate that the revertant phage can plate on missuppressing strains other than the one it was isolated on or cycled on, yet the improvements are not across the board since there are still missuppressing strains in which the phage show no improvement. For instance, *19*am83(tyr)sr532 was a result of cycling *19*am83(tyr) in TP280(leu) and then picking a revertant from TP280(leu), yet it shows a marked improvement in its plating ability on TP278(gln). The same is true for *19*am22(tyr)sr517. It was cycled and picked on TP282(ser), yet shows improvement on TP278(gln).

Sequences of the revertants. I had three classes of revertants: 1. nucleotide changes that resulted in an amino acid change in lysozyme 2. nucleotide changes in gene *13* that resulted in a change in the vicinity of the ribosome binding site of gene *19* and in some cases caused an amino acid change for p13, and 3. No nucleotide changes found in the DNA sequenced.

Most of the revertants had more than one change (see Table 10). The revertants that had changes in the ribosome binding site were presumed to be regulatory mutants and were not further characterized. These included *19*am39(val)sr575, *19*am39(val)sr581, *19*am83(tyr)sr532, and *19*am83(tyr)sr542. Two

second site revertants, *19*am22(tyr)sr517 and *19*am82(gln)sr515 had no nucleotide changes in gene *19*. The revertants that showed changes in gene *19* were *19*am82(gln)sr512, *19*am82(gln)sr513, *19*am82(gln)sr514, *19*am82(gln)sr516, and *19*am22(tyr)sr520.

Backcrossing the second-site revertants. Since the revertant phage had undergone extensive UV-irradiation, it was possible that the phenotypes observed were not due to the changes present in gene *19*, but were due to changes elsewhere in the phage genome. Since lysozyme appears to act independently as a monomeric protein I did not expect to recover revertants that had compensating changes in other phage proteins with which lysozyme acts. It was possible, however, that changes in the amount of transcription from P_{Late}, changes in the stability of the mRNA, or changes in an unknown gene, might increase synthesis of the missuppressed lysozyme or bypass the need for lysozyme. Two classes of the sequenced revertants are apparently extragenic: the revertants with changes in the ribosome binding site and revertants with no changes found. To rule out any affects of extragenic mutations from the sequenced intragenic revertants, I reconstructed the secondary site revertants by removing a 1.1 kb *Hpa*I fragment containing part of gene *13* and all of gene *19* from the revertant phage and cloning it into a plasmid. The

resulting plasmids were then used to cross gene *19* alleles onto P22 *sie*A44 *m*44 using P22 Kn469 *sie*A44 *m*44 (see Fig. 8). Recombinant phage from these crosses were selected and linkage of the amber mutations to the mutations responsible for their altered plating properties was determined.

The recombinant phage derived from *19*am82(gln)sr512 exhibited the secondary phenotype (ability to plate on additional suppressor strains) in 50 out of 50 plaques tested. The plaque morphology of the reconstructed phages was the same as that of original secondary site revertant phage isolated and as the parental *19*am phage. The reconstructed phage derived from *19*am82(gln)sr513 and *19*am82(gln)sr516 exhibited the secondary phenotype in 50 out of 50 plaques tested. For these two phages, however there were two plaque morphologies segregating in the crosses. In both cases there were normal turbid plaques and hazy turbid plaques. The hazy turbid plaques represented 20% of the recombinants. The hazy phenotype segregated independently from the secondary phenotype. The original mutant phage bearing an amber mutation in gene *19* and the secondary site revertant phage had normal turbid plaques. It is possible that a mutation elsewhere in the secondary site revertant phage influenced the plaque morphology such that the hazy phenotype was masked. Thus, the hazy mutation was only seen when crossed into a clean background. To investigate this mutation further the original mutant phage and the backcrossed phages (derived from

the second site revertant phage) both the normal turbid and the hazy turbid, were reverted on suppressor-minus cells and compared to wild type. The plaques of the original amber revertant, wild type, and the backcrossed normal turbid revertant were identical but the backcrossed hazy revertant maintained the hazy phenotype. I think that the hazy phenotype mutation lies in gene *13* where other morphology markers are known to occur. This was not investigated further.

Recombinant phage were obtained for *19*am22(tyr)sr520 and 50 out of 50 plaques tested exhibited the secondary phenotype.

Recombinant phage were obtained for *19*am83(tyr)sr532 and of the 50 plaques tested 46 exhibited the plating phenotype of the secondary site revertant. This phage has a ribosome binding site change which lies outside of the deletion that is present on P22 Kn469, therefore it was expected that in this case the mutation responsible for the secondary phenotype would segregate independently from the amber mutation, although it is closely linked. The secondary site phenotype was recovered 88% of the time. This is in very good agreement with the position of the secondary site mutation relative to the amount of homology present between the phage and the plasmid. The phage and the plasmid have 305 bases of homology before the point of nonhomology starts, the

secondary site mutation is located at base 295. Thus the secondary site mutation is 85% of distance from the start of the homology. Since this cross shows clearly that crossovers can be detected, by examining 50 single plaque lysates, when they occur only 46 bases from the point of nonhomology and all the secondary phages were sequenced beyond 50 bases from the point of nonhomology, then I can conclude that the secondary mutations responsible for the phenotypes observed in the other backcrossed phage are the mutations that were within gene *19*. I conclude that the changes present in gene *19* of *19*am82(gln)sr512, *19*am82(gln)sr513, *19*am82(gln)sr516, and *19*am22(tyr)sr520 that were sequenced are responsible for the ability to plate on the missuppressing strains.

Discussion

Amber suppression. The use of nonsense suppressors to study proteins has a long history. Miller and co-workers in 1979 studied 400 mutant *lac* repressor proteins each with a known amino acid substitution by using nonsense suppressors. RNA polymerase subunit *beta* was studied in a like fashion by Nene and Glass (1980). These studies yielded information concerning the efficiency of suppression as a by-product. The efficiency of suppression was found to be dependent not just

upon the specific suppressor allele being used but also upon the context in which the nonsense codon was found. Bossi (1983) and Miller and Albertini (1983) studied context affects of suppression in *Salmonella* and *E. coli*, respectively. The results were similar for suppressor tRNAs of both species. Since the studies I am carrying out are done in *Salmonella*, I will disscuss the results of Bossi. He examined the efficiencies of supD, supE, supF, and supJ, the amber suppressors characterized and isolated by Winston et al., (1979). Bossi found a range of efficiencies for each of the suppressors; supD ranged from 8.4% to 57.3%, supE 1% to 20%, supF 9.5% to 79.4%, and supJ from 14.8% to 51.9%. It was found that the bases preceeding the amber codon had no effect yet the base in the immediate 3' position correlated with efficiencies. The lowest efficiencies result when the 3' base is a C; an exception is when the triplet following the amber is CUN where N stands for any base. The highest efficiencies are seen when the 3' base is an A.

The efficiencies of suppression must be taken into account when looking at missuppressed amber mutations and isolating second site revertants. If a second site revertant changed the codon following a missuppressed amber mutation then the phenotype of the revertant could be due to a change in the protein or a change in the efficiency of translation. Of the eleven amber mutations used in this study six sites

originally coded for amino acids for which suppressor tRNAs are available. In all six cases the phage bearing the amber mutations formed normal size plaques when grown in a suppressor strain that inserted the wild type amino acid. None of the eleven sites are followed by a C.

Possible affects of the second site revertants. Of 25 combinations of P22 *19*am phage and suppressor strains that result in less active lysozymes, revertants were isolated for 18 combinations. From over 4000 revertants screened, 82 secondary site revertants were isolated. Of these 82 secondary site revertants, 10 were fully characterized.

By examining the suppression tables of the phage bearing the primary amber mutations to see how many combinations of phage and suppressor strains could theoretically yeild second site revertants and then comparing this to the combinations for which second site revertants were actually isolated, it becomes apparent that secondary site mutations are extremely rare, and for some combinations, were not obtainable. Second site revertants proved to be the most difficult to isolate from the mutant phages that exhibited the most severe defects on a particular suppressor strain, such that the the titers were the same on the suppressor strain as on the suppressor-minus strain. The combinations of mutant phage and suppressor that were moderately defective yielded

the most secondary site revertants. The combinations that resulted in a mild defect were difficult to work with and the numbers of revertants screened was too low to judge the frequency of isolating secondary site revertants.

Putative regulatory mutations accounted for 4 of the secondary site mutations. All of these were in the area of the gene *19* ribosome binding site. Due to the overlap of gene *13* and gene *19* all but one of these regulatory mutations also caused amino acid changes in gene *13*. It is therefore difficult to say with certainty if the increase in plating ability is due to a change in gene *13* gene product or an increase in translation of gene *19*. It is known that increasing the rate of translation of T4 gene e in a hybrid P22 phage can allow some missuppressed amber mutants to form plaques (Knight et al., 1987). This can be due to a greater amount of a less efficient lysozyme as seen when the phage is grown on a missuppressing strain. The absolute amount of translation increases so the absolute amount of read-though of the amber mutation increases even though the ratio of read-though to termination and suppression stays the same. This leads to more lysozyme being produced even though the amino acid at the position of the amber mutation is not known. This effect can be seen when phage bearing amber mutations is grown on a suppressor-minus strain (Knight et al., 1987). It is not

thought that gene *19* is translationally regulated as is the case of T4 gene *e*. Yet a better ribosome binding site could conceivably increase the amount of translation. I tried to see if this was the case by radiolabeling protein synthesized during phage infection and comparing amount of lysozyme synthesized. P22 lysozyme migrates with a host band during electrophoresis and is not readily quantified this way.

One secondary site mutation that mapped near the ribosome binding site did not cause an amino acid change in gene *13*, and yet it still suppressed the missuppressed amber mutation. P22*19*am39(val)sr575 has a CGC->CGT (Arg->Arg) change. The rest of the regulatory mutants share the same change GCC->ACC (Ala->Thr). These changes do not seem to bring the sequence closer to a consensus ribosome binding site (see Fig. 9).

The secondaries that show no change in gene *19* (*19*am22(tyr)sr517 and *19*am82(tyr)sr515) are less obvious. These indicate that some change in P22 can compensate for a less efficient lysozyme. Since it is thought that the P22 lysozyme acts independently, I do not think it is in a protein that interacts directly with lysozyme, but rather a consequence of a mutation that indirectly affects the level of lysozyme made. P22 probably does not have an accessory lysis gene analogous to the *RZ* gene of phage lambda (Garret and

Young, 1982). P22 deletion mutants have been created that delete sequences downstream of gene *19*. These deletions have no effect under a variety of plating conditions on the plaque-forming ability of P22 (A. Poteete, personal communication). Since P22 genes related by function are arranged in clusters, it seems unlikely that an accessory lysis gene would map far from genes *13* and *19*. It is more likely that the mutation is one that increases transcription from P_{Late} by changing the promoter or by destroying a terminator such as t_1 or t_2 so that P_{Late} transcript can be made earlier than normal. It could be a mutation that increases the stability of the mRNA. The same type of extragenic non-ribosome binding site secondary site mutations have been found by L. Hardy in our laboratory for a P22 T4e hybrid phage. He is in the process of mapping these. To date, it is known that the extragenic mutations in these phages do not map between P_{Late} and gene *19*.

The second site mutations in gene *19* might be of two different types. One type would be a "global suppressor" (Shortle et al., 1986). A global suppressor is thought to be a change which increases the stability of the protein. This type of mutation is non-allele specific, suppressing many different primary mutations. The other type of mutation would be an allele specific suppressor that made a compensatory change in the protein's structure to accommodate the primary mutation. There could be two types of allelic specific

suppressors; one that affects the final structure or one that affects an interaction present only in a folding intermediate. It is likely that the *19*am22(tyr)sr520 is an allele specific suppressor that affects the final structure of lysozyme. The revertants *19*am82(gln)sr512, *19*am82(gln)513, and *19*am82(gln)516 could be global suppressors or allelic specific suppressors that effect either the final structure or a folding intermediate. I do not have any indication what type of mutation these are from their postulated position in the three dimensional structure.

APPENDIX

Mutagenesis. The UV mutagenesis was monitored in a variety of ways. The formation of clear mutants and the increase in the reversion rate of amber mutations were indications of the degree of mutagenesis. The degree of reactivation of UV-irradiated phage was used to follow the induction of *muc*A and *muc*B, transcribed from pKM101 present in the suppressor strains. The fall in phage titer after UV-irradiation was also used to monitor the degree of mutagenesis.

The induction of the pKM101-bearing strains was accomplished by a growing the strains in minimal medium plus antibiotics until a cell density of 1 x 10^8/ml was attained. At that point 5 ml would be placed in the bottom of a glass petri dish and exposed to UV light for 10 seconds. The distance and light I used delivered 300 ergs/mm^2 when the exposure was 10 seconds. The cells were then used directly for infection with UV-irradiated phage. The infections were performed by adding the phage to be tested to 0.5 ml of culture in a glass tube, letting the phage adsorb for 20 minutes, then adding 5 ml of non UV-irradiated cells that were at a cell density of 2 x 10^8/ml. The cells were then plated by adding 2.5 ml of soft agar and pouring the mixture onto lambda plates. I found that if the cells to be UV stimulated were at higher density I did not see as much reactivation of

UV-irradiated phage.

The tests to follow the mutagenesis were not performed for every phage but were usually performed for one phage per experiment. A typical result is shown below.

This table show the increase in reversion frequency of an amber mutation, after UV treatment. Strains TP279(tyr) and TP278(gln) carry the pKM101 plasmid and the cells were UV stimulated before infection. TP279(tyr) strain missuppresses P22 *19*am82(gln) to the same degree as an Su-. The TP278(gln) suppressor strain inserts the original amino acid for *19*am82(gln).

allele	nonpermissive TP279(tyr)	permissive TP278(gln)	reversion freq.
*19*am82(gln)	1×10^3	1×10^{10}	10^{-7}
*19*am82(gln) (UV)	2×10^2	1×10^6	10^{-4}

The table below shows killing, and reactivation of UV treated phage. All the suppressor strains insert the original am acid. The suppressor tRNA from MS1363(gln) was transduced into an su- strain bearing pKM101 to create TP278(gln). In this particul experiment killing was 4 logs as judged by the plating on MS1363(gln). The reactivation was 30 fold as indicated by the titer of the UV-irradiated phage on MS136 as compared to the titer on TP278 (UV stimulated). There was some reactivation even without induction of the *muc* system as indicated by the

increased titer on TP278(gln). There were 4% clear plaques in th *19*am82(gln) (UV) stock and no detectable clears in the *19*am82(gln (<1%).

allele			
*19*am82(gln)	1×10^{10}	1×10^{10}	1×10^{10}
*19*am82(gln) (UV)	1×10^{6}	4×10^{6}	3×10^{7}

TABLE 6

PLASMIDS

Plasmids

pDR469 Ligation of 5.04 kb *Hpa*I fragment obtained from pPB20::TN5-13 with pTP399, which had been prepared by digesting with *Hind*III and *Bgl*II, and filling in the ends generated by the digestions. This construction replaces most of gene *19* with TN5 DNA.

pTP399 This plasmid contains DNA from three sources: An origin of replication containing fragment derived from pZ152 (Zagusky and Berman, 1984)(*Pvu*II-*Eco*RI) ligated to P22 DNA that includes DNA from a *Rsa*I site (that has been ligated with *Sal*I linkers to the pZ152 filled-in *Eco*RI) in gene *13* to the first *Dde*I site in gene *19* (the first nine codons of gene *19*). This is followed with a *Hind*III linker, and T4 DNA that encompasses the ribosome binding site of gene e and all of gene e. This is followed by a *Bgl*II linker. After the *Bgl*II linker, P22 DNA continues from 70 bp from the end of gene *19* to a *Hpa*I site 384 bases 3' of gene *19*. The *Hpa*I site has been converted to a *Bam*HI site when ligated with *Bam*HI linkers to *Pvu*II site of pZ152.

pDR512 Plasmid used for backcrossing the mutant alleles to P22. Ligation of 1.1 kb *Hpa*I fragment of P22 am82sr512 with *Eco*RV digested pBR322.

pDR513 This is the same construction as pDR512 except it contains P22 DNA from P22am82sr513.

pDR516 This is the same construction as pDR512 except it contains P22 DNA from P22am82sr516.

pDR520 This is the same construction as pDR512 except it contains P22 DNA from P22am22sr520.

pDR532 This is the same construction as pDR512 except it contains P22 DNA from P22am83sr532.

TABLE 7

Isolation of Revertants with Direct Plating

	Missuppressing Strains						Revertants	
allele	TP278	TP279	TP280	TP282	TP308	TP309	primary	secondary
Am12	+						32	0
					+		50	0
						+	50	0
Am22	+						350	0
			+				350	0
				+			350	0
						+	100	0
Am39						+	50	14
Am80		+					100	0
					+		100	0
						+	100	0
Am82		+					750	0
				+			600	0
					+		50	0
						+	50	0
Am83	+						500	0
			+				350	0
				+			200	0

Table 7. A summary of the isolation of revertants after UV-irradiation and direct plating.

TABLE 8

Isolation of Revertants with "Cycling"

	Suppressor Strains			
allele/Su+	TP278 (gln)	TP279 (tyr)	TP280 (leu)	TP282 (ser)
am22/TP278	0/150		0/150	
am22/TP282	0/18		16/50	
am82/TP278		0/50		0/50
am82/TP279				0/50
am82/TP282		0/150		5/100
am83/TP278	18/50		29/200	
am83/TP280	7/50		12/50	

Table 8. A summary of revertants isolated from UV-irradiated phage that were passaged through semi-permissive su+ cells before being plated on missuppressing strains. The phage allele is shown on the left with the strain used to passage the phage. The numbers shown are the numbers of secondary site revertants over the number of revertants screened. The numbers are in the columns under the suppressor strain used for the platings.

TABLE 9

SUPPRESSION PATTERNS OF SECOND SITE REVERTANTS

Phage	Inserted Amino Acid					
	Gln	Tyr	Leu	Ser	Phe	Cys
am82 (gln)	++	-	++	+/-	-	-
am82sr512	++	+/-	++	++	-	+/-
am82sr513	++	-	++	+	-	-
am82sr514	++	-	++	+	-	-
am82sr516	++	-	++	+*	-	-
am22 (tyr)	+	++	+	+	++	-
am22sr517	++	++	++	++	++	+/-
am22sr520	++	++	++	++	++	+/-
am39 (val)	++	++	++	++	++	+/-
am39sr575	++	++	++	++	++	+
am39sr581	++	++	++	++	++	++
am83 (tyr)	+	++	-	-	++	-
am83sr532	++	++	++	-	++	-
am83sr542	++	++	++	-	++	-

Table 9. Suppression patterns of parent and second site revertant phage. The suppression strains used were TP278, TP279, TP280, TP282, TP308, and TP309. The platings were done at 37°C. The symbols are; ++ , normal plaque size and titer; +, pinpoint plaques and titer reduced by no more than 1000 fold; -, titer reduced by greater than 1000 fold. The titers were standardized by titering on TP320. Plaque size determinations are subjective comparisons of plaque size on the most permissive strain for the given phage. *, indicates slightly better plaque than + but not as good as ++.

Table 10. Sequence changes of second site revertants. The nucleotide changes found in the second site revertants of missuppressed *19*am phage are shown above. The allele number is shown on the left. The nucleotide changes are shown in the next column and the subsequent amino acid changes are shown in the following column. The numbers represent codon numbers of gene *19* unless otherwise indicated.

TABLE 10

SEQUENCE CHANGES OF SECOND SITE REVERTANTS

	Nucleotide changes	Amino Acid changes
am82sr512	AAT TCT->AAA CCT TCA->TTA 42 43 46	ASN SER->LYS PRO SER->LEU 42 43 42 43 46 46
am82sr513	TCA->TTA 46	SER->LEU 46 46
am82sr514	TCA->TTA 46	SER->LEU 46 46
am82sr516	TCA->TTA TCT->ACT GAC->GTC 23 23 56 128	SER->LEU SER->THR ASP->VAL 23 23 56 56 128 128
am22sr517	NONE	
am22sr520	AGA CTA->AGT TTA TCA->AAA 18 19 18 19 23 23	ARG LEU->SER LEU SER->LYS 18 19 18 19 23 23
am39sr575	CGC->CGT 97(GENE 13)	ARG->ARG 97(GENE 13)
am39sr581	GCC->ACC 103(GENE 13)	ALA->THR 103(GENE 13)
am83sr532	GCC->ACC TTA->TTG 103(GENE 13) 11 11	ALA->THR LEU->LEU 103(GENE 13) 11 11
am83sr542	GCC->ACC 103(GENE 13)	ALA->THR 103(GENE 13)

FIGURE 8

BACKCROSSING THE MUTANT ALLELES TO P22

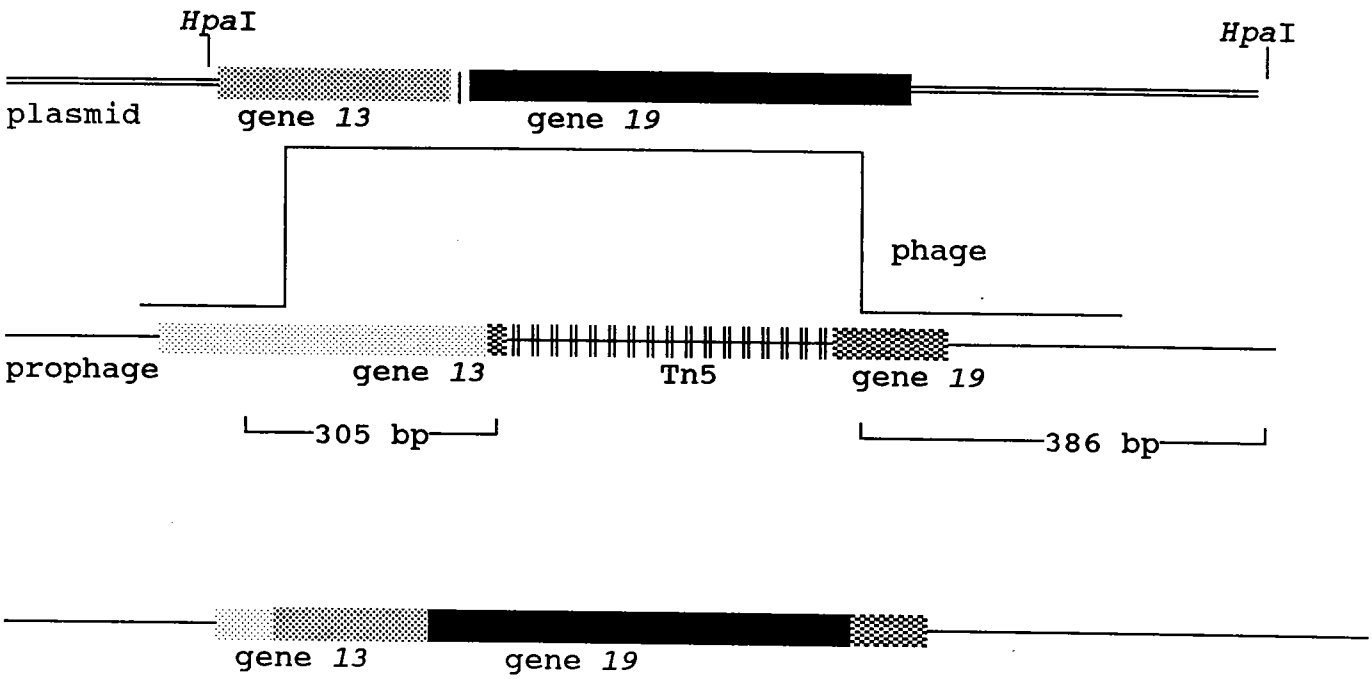


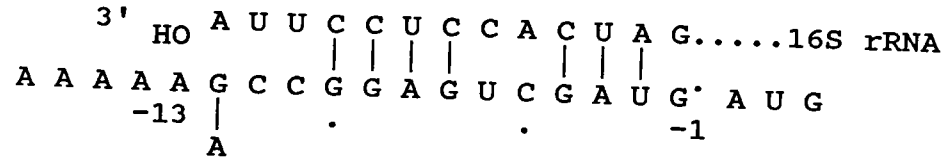
Figure 8. The cross between P22Kn469 and pDR512-pDR520 series of plasmids. Non-homology between the plasmid and the phage in gene *19* is represented by ||||||. Gene *19* has been deleted from the prophage between codon 9 and codon 123 and substituted with a large insertion. The number of base pairs available in each crossover interval are shown below the prophage

FIGURE 9

RIBOSOME BINDING SITE MUTATIONS

A.

Wild type P22 gene *19* ribosome binding site



B.

Wild type P22 gene *13* 3' sequence

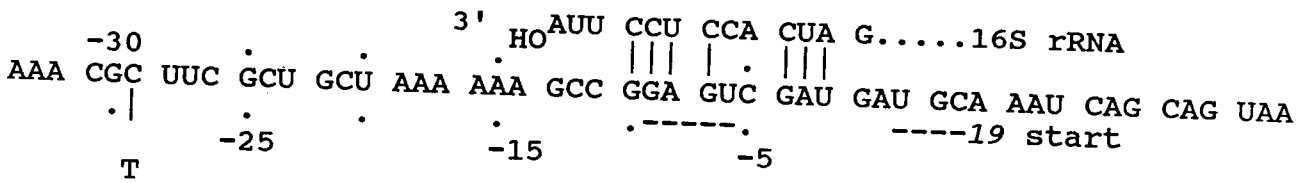


Fig. 9. Panel A. The sequence in the area of P22 ribosome binding site is shown below the sequence of the 3' end of the 16S rRNA. The presumed bases that pair are marked. The secondary change found in 3 of the regulatory mutants is designated below the P22 sequence. Panel B. The 3' sequence of gene *13* is shown with the start site and ribosome binding site of gene *19* indicated by dashed lines. The secondary change that does not change an amino acid in gene *13* is indicated below the line. The sequence was searched for possible stem and loop structures and none were found that would involve the changed base.

CHAPTER V

CONCLUSIONS

CHAPTER V

CONCLUSIONS

The question "How does a protein attain and maintain its tertiary structure as dictated by its amino acid sequence?" has been a significant question asked by biologists for decades (Anfinsen, 1964). The work presented in this thesis explores an approach to this question by developing a system that examines defective proteins which have compensating changes. Second site reversion has been used as a method to study proteins for more than 20 years but only in recent years with the technological advances that have been made has it become feasible to do this on a significant scale (Helinski and Yanofsky, 1963; Shortle and Lin, 1986).

In this thesis I describe 11 amber mutations in P22 gene *19*. All of the amber mutations created non-functional truncated proteins when grown on su- strains ranging in length from 12 residues to 122 residues. Using six different suppressor tRNA strains it was possible to study 60 amino acid substitutions in P22 lysozyme. Of 60 amino acid substitutions only 18 resulted in defective lysozymes. This was not a surprising result since many investigators have found that proteins exhibit a large degree of tolerance to amino acid substitutions (Miller et al, 1979; Perutz and Lehmann, 1968; and Streisinger et al., 1966).

Amino acid substitutions that yield defective lysozymes

are not predictable by current methods (Chou and Fasman, 1974; Schulz and Schirmer, 1984). Tables 11 and 12 show how the substituted amino acid at each residue position corresponds to exchange groups as defined by Schulz and Schirmer or helix propensities as defined by Chou, respectively. Neither index predicted which substitutions would result in nonfunctional lysozyme. For instance at position 61 where the original amino acid was glutamate, the substitution of tyrosine yields a functional lysozyme when both indexes would indicate that tyrosine would not substitute well for glutamate; at position 82 where the original amino acid was glutamine both indexes indicate that serine would be a neutral substitution and yet it is a detrimental one. The efficiencies of individual substitutions must be determined empirically.

The missuppressed amber mutants were used to isolate second site revertants. These revertants enable interactions between amino acids to be mapped. The relationships between the second site mutations and the primary site mutations are not always apparent from the positions of the mutations in the amino acid sequence. The mutations must be mapped in context of the three dimensional structure. Only in this way can one see the relationships, but even then the effects of the secondary change on the primary mutation are not always readily apparent. The wild type, mutant protein, and revertant protein should be

crystallized and the structures determined by x-ray diffraction patterns. This is an enormous task. If the wild type protein has a known crystal structure then the changes present in the structure of the revertant protein can be predicted by modeling (eg. Ebright et al., 1984; Burley and Petsko, 1985). When the substitutions map near each other on the three dimensional model it is sometimes possible to envision how the affected residues interact by way of Van der Waals forces or hydrogen bonding. The difficulties come when the mutations are distant from one another. To explain the ability of one mutation to suppress the other one must invoke totally different structures or changes transduced long distances through the protein's backbone. Changes transduced long distances have been seen in studies done with single mutant T4 lysozymes (Alber et al., 1986). Most mutant proteins that have had crystal structures solved though show only very localized conformational changes (Alber et al., 1987; and Matthews et al., 1987). Other than direct interaction between residues one can postulate that the interaction does not take place in the final protein structure but is a transient interaction that takes place during protein folding. This is much harder to decipher since folding intermediates if they exist have not been isolated except in cases where disulfide bonds are involved (Beasty et al., 1986; Anderson and Welaufer, 1976; and

Freidman et al., 1966). A type of suppressor which is not allele specific has been characterized; global suppressors are changes which do not interact directly with the primary mutation but increase the overall stability of the protein thereby suppressing a defect which resulted in an unstable protein. Shortle and Lin (1986) isolated and characterized mutations of this kind for Staph A nuclease.

Since the structure of P22 lysozyme has not yet been solved, it is difficult to interpret the second site mutations. Some information is available though, since it is postulated that the structure of P22 lysozyme will be very similar to that of T4 lysozyme (Weaver et al., 1986). If we examine residues of T4 lysozyme that correspond to residues of P22 lysozyme we can determine if they are predicted to be close enough in the tertiary structure of the protein to have an effect on one another. Figure 10 shows how the residues of T4 lysozyme line up with those of P22 lysozyme. Figure 11 indicates the locations of the mutations present in the second site revertants P22 *19* am82(gln)sr512, *19*am82(gln)sr513 and *19*am82(gln)sr516 superimposed on the structure of T4 lysozyme. As shown the revertants do not seem to interact directly with the primary substitution. They map on the opposite side of the molecule in residues that are not predicted to have the same conformation as in P22 lysozyme. These mutations might be

global suppressor mutations which stabilize the protein thereby increasing the activity of the mutant. These particular mutations were not isolated in conjunction with any other primary mutations. The experiments that would answer whether these mutations are global would be to isolate the second site mutations away from the primary site mutation and then combine the second site mutations with other primary site substitutions. Also stability studies should be done on purified lysozyme to see if the temperature of denaturation has increased in the mutant proteins.

The P22 *19*am82(gln)sr512 mutation which has additional changes as compared to the P22 *19*am82(gln)sr513 demonstrates that even if the ser -> leu at residue 46 is a global suppressor, the additional changes at residues 42 and 43 enables the phage to grow well at 40^{0}C on serine suppressor while the original change at position 46 does not. The P22 *19*am82(gln)sr512 phage can also grow slightly on tyrosine and phenylalanine-inserting suppressors while the P22 *19*am82(gln)sr513 phage shows no growth. This could be a case of additive global effects. Figure 12 shows the location of the mutations present in second site revertant, P22 *19*am22(tyr)sr520, superimposed on the T4 lysozyme structure. Since these changes map very close to the original mutation on both the primary and tertiary structure maps it is speculated that these changes represent allele

specific suppression of the substitution at position 22. It appears that the two changes map on either side of a surface loop.

Significance. The study described in this thesis was designed to develop a system that would allow the characterization of second site revertants of missuppressed amber mutations. The system was set up so that amber mutations could easily be created in any codon and transferred to phage in a dependable manner. The subsequent screening for second site revertants and isolation of several demonstrates that this system is workable. When this work was first undertaken it was hoped that the crystal structure of P22 lysozyme would be solved. This has not occurred as of yet, even though it is currently being worked on by Brian Matthews. Without the crystal structure available it is difficult to draw conclusions about how the mutations effect the structure of P22 lysozyme. Yet, the work described here does have some significance since the methods are currently being used to create amber mutations in every codon of T4 lysozyme gene which is present in a P22 hybrid phage. Second site revertants isolated in that system will be readily interpretable in terms of the three dimensional structure. It is hoped that some further understanding of how an amino acid sequence dictates

structure will come out of these studies.

It is more important today than ever before to understand how amino acid sequences determine structures, since with the molecular biology technologies available, genes could be designed to incorporate changes that would make proteins more efficient or stable or even have unique enzymatic properties if we could understand the relationship between amino acid sequences and structures.

Table 11

Amino Acid Substitutions Exchange Values

original residue		Substituted residue gln	tyr	leu	ser	phe	cys
lys	12	+;11	+;3	+;6	+;10	-;3	-;4
tyr	22	-;2	+;137	+;6	+;3	+;66	-;1
val	39	+;9	+;3	+;15	+;9	+;7	-;11
glu	61	+;10	+;1	+;6	+;11	+;3	+;5
gln	64	+;21	+;2	+;6	+;11	+;3	+;5
trp	65	+;2	+;1	+;3	+;1	+;41	-;1
gln	80	+;21	-;2	+;6	+;11	-;3	-;5
gln	82	+;21	-;2	+;6	-;11	-;3	-;5
tyr	83	+;2	+;137	-;6	-;3	+;66	-;1
gln	113	+;21	+;2	+;6	+;11	+;3	-;5
trp	122	+;2	+;1	+;3	+;1	+;41	-;1

Table 11. This table shows the exchange values for each amino acid substitutions at the positions of the amber mutations. The original amino acid and codon numbers are given on the left. The substituted amino acids are shown across the top. The symbols + and - indicate if there was any defect in the plating ability of the phage bearing amber mutations on that suppressor strain at 30^{0}C. The number following gives a ratio; the probability that an amino acid pair will occur at a given position in proteins of common ancestry divided by the probability that it will occur by chance. All ratios are mutiplied by a factor of 10.

TABLE 12

Helix Propensities of Amino Acid Substitutions

	Substituted residue					
	gln h	tyr b	leu H	ser i	phe h	cys i
original residue						
lys,I 12	+;h,I	+;b,I	+;H,I	+;i,I	-;h,I	-;i,I
tyr,b 22	-;h,b	+;b,b	+;H,b	+;i,b	+;h,b	-;i,b
val,h 39	+;h,h	+;b,h	+;H,h	+;i,h	+;h,h	-;i,h
glu,H 61	+;h,H	+;b,H	+;H,H	+;i,H	+;h,H	+;i,H
gln,h 64	+;h,h	+;b,h	+;H,h	+;i,h	+;h,h	+;i,h
trp,h 65	+;h,h	+;b,h	+;H,h	+;i,h	+;h,h	-;i,h
gln,h 80	+;h,h	-;b,h	+;H,h	+;i,h	-;h,h	-;i,h
gln,h 82	+;h,h	-;b,h	+;H,h	-;i,h	-;h,h	-;i,h
tyr,b 83	+;h,b	+;b,b	-;H,b	-;i,b	+;h,b	-;i,b
gln,h 113	+;h,h	+;b,h	+;H,h	+;i,h	+;h,h	-;i,h
trp,h 122	+;h,h	+;b,h	+;H,h	+;i,h	+;h,h	-;i,h

Table 12. This table shows the propensities of finding any given amino acid in a helix. The original amino acid and codon numbers are given on the left. The substituted amino acids are shown across the top. The symbols + and - indicate if there was any defect in the plating ability of phage bearing amber mutations on that suppressor strain at 30°C. The letter following represents Helical asssignments: H, strong helix former; h, helix former; I, weak helix former; i, helix indifferent; b, helix breaker; B, strong helix breaker (Chou and Fasman, 1974).

FIGURE 10. The sequences are arranged for maximum sequence homology. There are three gaps introduced into the P22 sequence: a 12 residue gap between 47 and 48, a 5 residue gap between 71 and 72, and a 1 residue gap after 102. The residue that show identity are underlined. The residues that correspond to amber mutations are double underlined and residues correspond to second site changes are in bold print.

FIGURE 10

Alignment of the Amino Acid Sequences of T4 Lysozyme and P22 Lysozyme

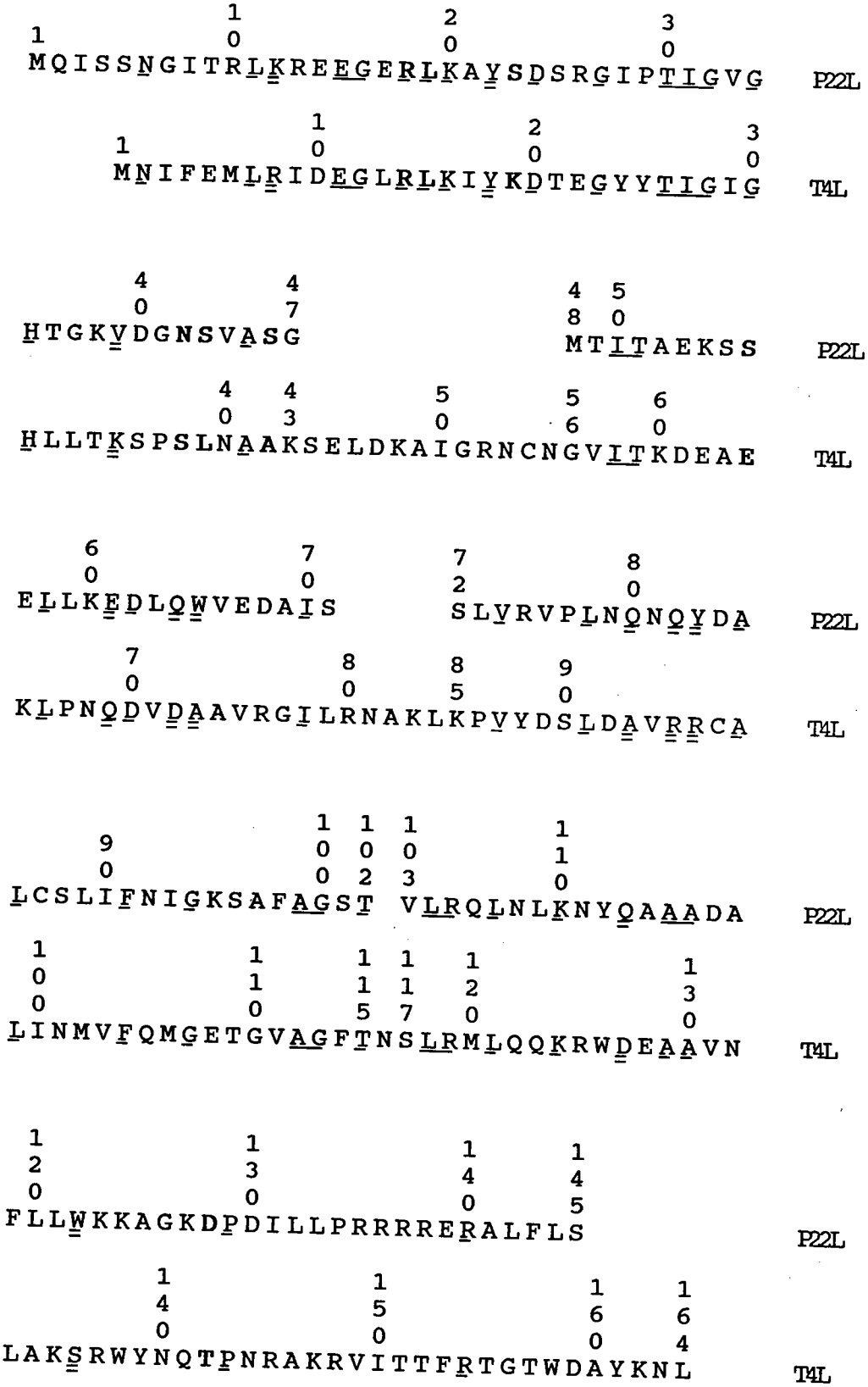


Figure 11. The locations of the mutations present in P22 *19am*82sr513 and *19am*82sr516 are shown superimposed on the structure of T4 lysozyme.

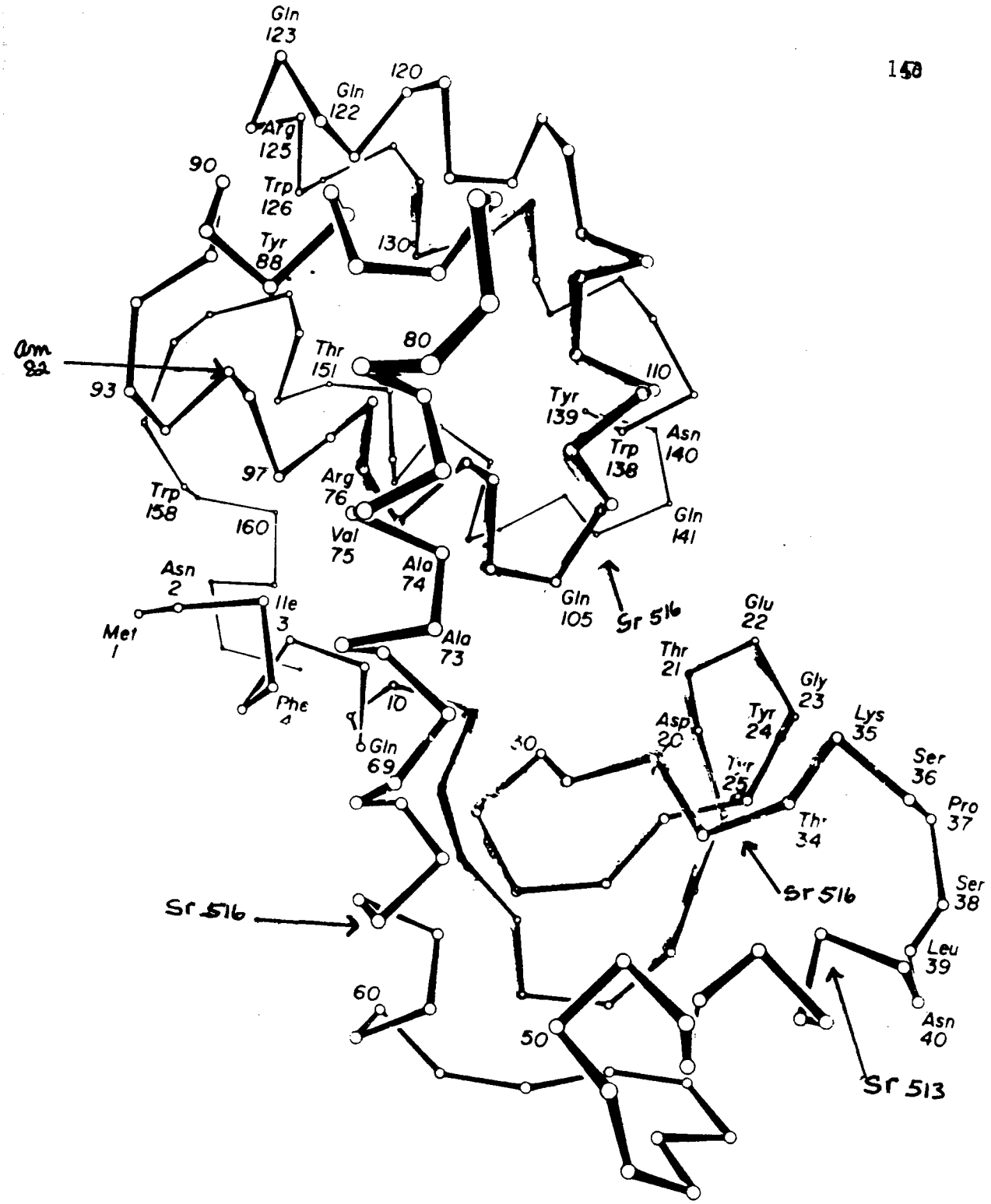

Gln 123
120
Gln 122
Arg 125
90
Trp 126
Tyr 88
130
Arm 82
Thr 151
80
110
93
Tyr 139
Asn 140
Trp 138
97
Arg 76
Trp 158
160
Val 75
Gln 141
Asn 2
Ala 74
Gln 105
Sr 516
Glu 22
Ile 3
Met 1
Ala 73
Thr 21
Gly 23
Phe 4
10
Asp 20
Tyr 24
Lys 35
Gln 69
30
Tyr 25
Ser 36
Pro 37
Thr 34
Sr 516
Sr 516
Ser 38
Leu 39
60
50
Asn 40
Sr 513

Figure 11

Figure 12. The location of the mutations present P22 *19*am22sr520 are shown supreimposed on the structure of T4 lysozyme.

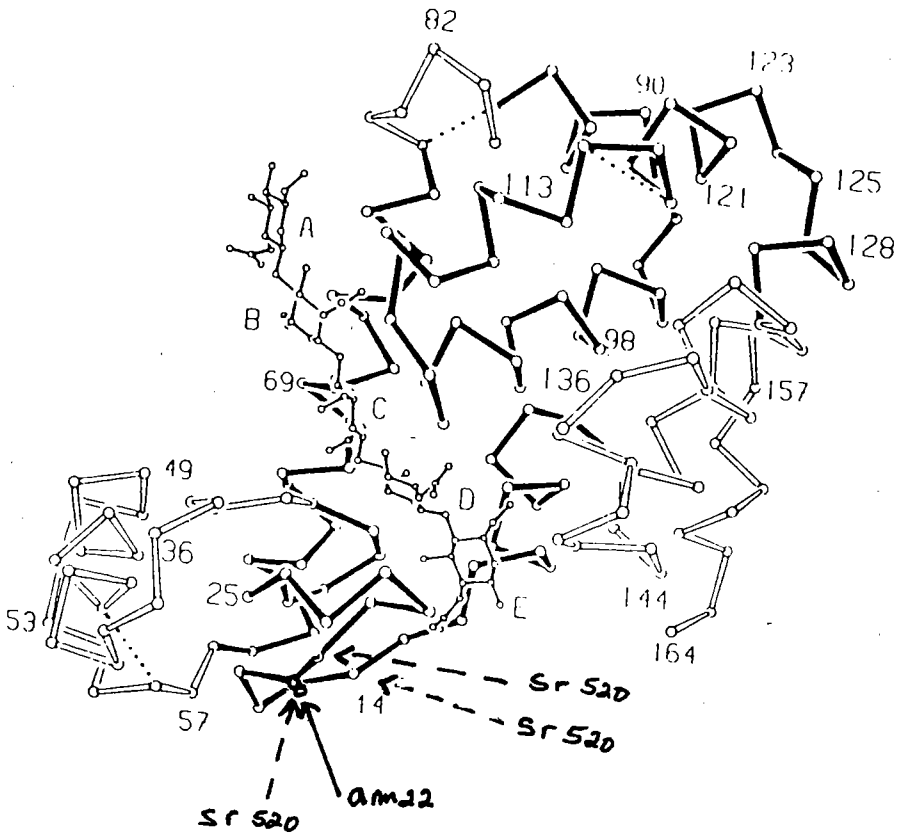

82
90
123
113
121
125
128
A
B
98
69
136
157
C
49
D
36
25
E
144
53
164
14
Sr 520
Sr 520
57
amid
Sr 520

Figure 12

References

Alber, T., Dao-pin, S., Wilson, K., Wozniak, J. A., Cook, S.P., and Matthews, B.W. (1987a). *Nature, 330*, 41-46

Alber, T. K., Dao-pin, S., Nye, J. A., Muchmore, D.C., and Matthews, B.W. (1987b). *Biochemistry, 26*, 3754-3758.

Amann, E., Brosius, J., and Ptashne, M. (1983). *Gene, 25*, 167-178.

Anderson, W.L., and Wetlaufer, D.B. (1975). *J. Biol. Chem., 251*, 3147-3153.

Anderson, W. L., and Wetflaufer, D.B. (1976). *Journal of Biol. Chem., 251*, 3147-3153.

Anfinsen, C.B. (1973). *Science, 181*, 223-230.

Anfinsen, C.B., and Haber, E. (1961). *Journal of Biological Chemistry, 236*, 1361-1363.

Beasty, A.M., Hurle, M.R., Manz, J.T., Stackhouse, T., Onugger, J.J. and Matthews C.R. (1986). *Biochem., 25*, 2965-2974.

Benveniste, R., and Davies, J. (1973). *Ann. Rev. of Biochem., 42*, 471-506.

Berget, P.B., Poteete, A.R., and Sauer, R.T. (1983) *J. Mol.Biol., 164*, 561-572.

Bienkowska-Szewczyk, K. and Taylor, A. (1980). *Biochemica et Biophysica Acta, 615*, 489-496.

Black, L.W., and Hogness, D.S. (1969). *J. Biol. Chem., 244*, 1968-1975.

Blundell, T.L., Sibanda, B.L., Sternberg, M.J.E., and Thornton, J.M. (1987). *Nature, 326*, 347-352.

Bolivar, F., Rodriguez, R.L., Greene, P.J., Betlach, M.C., Heyneker, H.L., and Boyer, H.W. (1977). *Gene, 2*, 95-113.

Bossi, L. (1983). *J. Mol.Biol., 164*, 73-87.

Botstein, D., Chan, R.K., and Waddell, C.H. (1972). *Virology ,49*, 268-282.

Brent, R. and Ptashne, M. (1981). *Proc. Natl. Acad. Sci. USA, 78*, 4204-4208.

Burley, S.K. and Petsko, G.A. (1985). *Science, 229*, 23-28.

Cantor, C.R. and Schimmel, P.R. (1980) Biophysical Chemistry. Part 1: The conformation of biological macromolecules. W.H. Freeman and Company. San Francisco pp 46-59.

Carter, P.J., Winter, G., Wilkinson, A.J., and Fersht, A.R. (1984). *Cell, 38*, 835-840.

Chou, P. Y., and Fasman, G.D. (1974). *Biochemistry, 13*, 222.

Chou, P.Y., and Fasman, G.D. (1974). *Biochemistry, 13*, 211-221.

Chou, P.Y., and Fasman, G. D. (1978). *Ann. Rev. Biochem., 47*, 251-276.

Claverie, J.M., and Bricault, L. (1986). *Proteins: Structure, Functions and Genetics. ,1*, 60-65.

Coulondre, C., and Miller, J.H. (1977). *J. Mol. Biol., 117*, 577-606.

Creighton, T.E. (1974). *J. Mol. Biol., 87*, 579-602.

Creighton, T.E. (1974). *J. Mol. Biol., 87*, 603-624.

Creighton, TE. (1974). *J. Mol. Biol., 87*, 563-577.

Creighton, T.E. (1978). *Prog. Biophys. Molec. Biol., 33*, 231-297.

Davies, D.R. (1964). *J. of Mol. Biol., 9*, 605-609.

Ebright, R.H., Cossari, P., Gicquel-Sanzey, B., and Beckwith, J. (1984). *Proc. Natl. Acad. Sci. USA , 81*, 7274-7278.

Enea, V., and Zinder, N.D. (1982) *Virology, 122*, 222-226.

Friedman, M.E., Scheraga, H.A., and Goldberger, R.F. (1966). *Biochem.,5*, 3770-3778.

Garnier, J., Osguthorpe, D.J. and Robson, B. (1978). *J. Mol. Biol., 120*, 97.

Garret, J.M. and Young, R. (1982). *J. Virology, 44*, 886-892.

Glass, R.E., Nene, V., and Hunter, M.G. (182). *Biochem. J., 203*, 1-13.

Grutter, M.G., and Matthews, B.W. (1982). *J. Mol. Biol., 154*, 525-535.

Hansen, J.N. (1981). *Anal. Biochem., 116*, 146-151.

Harrison, S.C., and Durbin, R. (1985). *Proc. Natl. Acad. Sci. USA, 82*, 4028-4030.

Helinski, D.R., and Yanofsky, C.(1963). *Journal of Biol. Chem., 238*, 1043-1048.

Hothschild, A., Irwin, N., Ptashne, M. (1983). *Cell, 32*, 319-325.

Jarvik, J., and Botstein, D. (1975). *Proc. Nat. Acad. Sci. USA, 72*, 2738-2742.

Jorgensen, R.A., Rothstein, S.J. and Reznikoff, W.S. (1979). *Mol. Gen. Genet., 177*, 65-72.

Josslin, R. (1970). *Virology, 40*, 719-726.

Karplus, M., and Weaver, D.L. (1976). *Nature, 260*, 404-406.

Knight, J.A., Hardy, L.W., Rennell, D., Herrick, D., and Poteete, A.R. (1987). *Journal of Bacteriology, 169*, 4630-4636.

Kuntz, I.D. (1975). *J. of the American Chemical Society, 97*, 4362-4368.

Laemmli, U.K. (1970). *Nature (London), 227*, 680-685.

Lederberg, E.M., and Cohen, S.N. (1974). *Journal of Bacteriology, 119*, 1072-1074.

Lederberg, J.M. (1960). *Symp. Soc. Gen. Microbiol., 10*, 115-131.

Levinthal, C. J. (1968). *J. Chim. Phys., 65*, 44-45.

Levitt, M., and Chothia, C. (1976). *Nature, 261,* 552-558.

Lewis, P.N., Momany, F.A., Scheraga, H.A. (1971). *Proc. Natl. Acad. Sci. USA, 68*, 2293-2297.

Lin, S.H., Konish, Y., Denton, M.E., Scheraga, H.A. (1984). *Biochem, 23*, 5504-5512.

Lim, V.I. (1974a). *J. Mol. Biol., 88* , 857-872.

Lim, V.I. (1974b). *J. Mol. Biol., 88*, 873-894.

Matthews, B.W., Nicholson, H. and Becktel, W.J. (1987). *Proc. Natl. Acad. Sci. USA, 84*, 6663-6667.

Matthews, B.W.(1987). *Biochemistry, 26*, 6885-6891.

Maxam, A., and Gilbert, W. (1980). *Meth. Enzymol. 65,* 499-559.

Miller, J.H., and Albertini, A.M. (1983). *J. Mol. Biol., 164,* 59-71.

Miller, J.H. (1985). *J. Mol. Biol., 182,* 45-68.

Normanly, J., Masson, J.M., Kleina, L.G., Abelson, J., and Miller, J.H. (1986). *Proc. Natl. Acad. Sci. USA, 83,* 6548-6552.

Owen, J.E., Schultz, D.W., Taylor, A. and Smith, G.R. (1983). *J. Mol. Biol., 165,* 229-248.

Pakula, A.A., Young, V.B., and Sauer, R.T.(1986). *Proc Natl. Acad. Sci. USA, 83,*8829-8833.

Pantoliano, M.W., Ladner, R.C., Bryan, P.N., Rollence, M.L., Wood, J.F., and Poulos, T.L. (1987). *Biochemistry, 26,* 2077-2082.

Perry, L.J. and Wetzel, P. (1984). *Science, 226,* 555-557.

Perutz, M.F., and Lehmann, H. (1968). *Nature, 219,* 902-909.

Ponder, J.W., and Richards, F.M. (1987). *Nature, 193,* 775-791.

Poteete, A.R. (1982). *Viology, 119,* 422-429

Poteete, A.R., and Roberts, T.M. (1981). *Gene,* 13, 153-161.

Radman, M. and Wagner, R. (1984). in T. Trautner (ed) Methylation of DNA. Springer Varlay Berlin.

Rao, G.R.K., and Burma, D.P. (1971). *J. Biol. Chem., 246,* 6474-6479.

Reader, R.W., and Siminovitch, L. (1971). *Virology, 43,* 607-622.

Remington, S.J. and Matthews, B.W. (1978) *Proc. Natl. Acad. Sci. USA, 75,* 2180-2184.

Rennell, D. and Poteete, A.R. (1985). *Virology, 143,* 280-289.

Robson, B. and Pain, R.H. (1974a). *Biochem. J., 141,* 883-897.

Robson, B. and Pain, R.H. (1974b). *Biochem. J., 141,* 869-882.

Roberts, T.M., Swanberg, S.L., Poteete, A., Riedel, G. and Backman, K. (1980). *Gene, 12,* 123-127.

Robson, B. (1976). *J. Mol. Biol., 107,* 327-356.

Rose, G.D., Geselowitz, A.R., Lesser, G.J., Lee, R.H., and Zehfus, M.H. (1985) *Science, 229,* 834-838.

Rossmann, M.G., and Argos, P. (1976). *J. Mol. Biol., 105,* 75-96.

Rothstein, S. J., Jorgensen, R.A., Postle, K., and Reznikoff, W.S.(1980). *Cell, 19,* 795-804.

Rutila, J.E., and Jackson, E.N. (1981). *Virology, 113,* 769-775.

Sanger, F., Nicklen, S., and Coulson, R. (1977). *Proc. Natl. Acad.Sci. USA., 74,* 5463-5467.

Schulz, G.E., and Schirmer, R.H. (1984). in Principles of Protein Structure. Springer-verlag Charles Canter (ed).

Shortle, D., and Lin, L. (1985). *Genetics, 110,* 539-555.

Smith, H.O. and Levine, M. (1964). *Proc. Natl. Acad. Sci. USA, 52,* 356-363.

Sternberg, M.J.E. (1986). *Anti-Cancer Drug Design, 1,* 169-178.

Streisinger, G., Mukai, F., Dreyer, W.J., Miller, B., Horivch, S. (1961). *Cold Spring Harbor Symp. Quant. Biol., 31.*

Tsugita, A., and Inouye, M. (1968). *J. Mol. Biol., 37,* 201-212.

Tsugita, A., Inouye, M., Terzaghi, E., and Streisinger, G. (1968). *J. Biol. Chem., 243,* 391-397.

Villafrasnca, J.E., Howell, E.E., Oatley, S.J., Xuong, N., and Kraut, J. (1987). *Biochemistry, 26,* 2182-2189.

Weaver, L.H., Rennell, D., Poteete, A.R., and Matthews, B.W. (1985). *Journal of Mol. Biol., 184,* 739-741.

Wells, J.A., and Powers, D.B. (1986). *Journal of Biol. Chem., 261,* 6564-6570.

Wetlaufer, D.B. (1973). *Proc. Natl. Acad. Sci. USA, 70,* 697-701.

Weinstock, G.M., Susskind, M.M., and Botstein, D. (1978). *Genetics, 92,* 685-710.

Walker, G.C. (1984). *Microbiological Reviews, 48,* 60-93.

Wilson, D.B. (1982). *J. Bact., 151,* 1403-1410.

Winston, F., Botstein, D. and Miller, J.H. (1979). *J. Bacteriology, 137*, 433-439.

Youderian, P. and Susskind, M. (1980). *Virology, 107*, 270-282.

Youderian, P., Bouvier, S., Susskind, M.M. (1982). *Cell, 30*, 843-853.

Youderian, P., Vershon, A., Bouvier, S., Sauer, R.T., and Susskind, M.M. (1983). *Cell, 35*, 777-783.

Zagursky, R.J., and Berman, M.L. (1984). *Gene, 27*, 183-191.