Copyright © 1987 by the Genetics Society of America

Identification of Sites Influencing the Folding and Subunit Assembly of the P22 Tailspike Polypeptide Chain Using Nonsense Mutations

Bentley Fane and Jonathan King

Department of Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139

Manuscript received March 21, 1987
Revised copy accepted June 19, 1987

ABSTRACT

Amber mutations have been isolated and mapped to more than 60 sites in gene *9* of P22 encoding the thermostable phage tailspike protein. Gene *9* is the locus of over 30 sites of temperature sensitive folding (*tsf*) mutations, which affect intermediates in the chain folding and subunit association pathway. The phenotypes of the amber missense proteins produced on tRNA suppressor hosts inserting serine, glutamine, tryosine and leucine have been determined at different temperatures. Thirty-three of the sites are tolerant, producing functional proteins with any of the four amino acids inserted at the sites, independent of temperature. Tolerant sites are concentrated at the N-terminal end of the protein indicating that this region is not critical for conformation or function. Sixteen of the sites yield temperature sensitive missense proteins on at least one nonsense suppressing host. Most of the sites with *ts* phenotypes map to the central region of the gene which is also the region where most of the *tsf* mutations map. Mutations at 15 of the sites have a lethal phenotype on at least one tRNA suppressor host. For nine out of ten sites tested with at least one lethal phenotype, the primary defect was in the folding or subunit association of the missense polypeptide chain. This analysis of the tailspike missense proteins distinguishes three classes of amino acid sites in the polypeptide chain; residues whose side chains contribute little to folding, subunit assembly or function; residues critical for maintaining the folding and subunit assembly pathway at the high end of the temperature range of phage growth; and residues critical over the entire temperature range of growth.

THE structures of over 150 proteins have been determined to atomic resolution and the complete amino acid sequences of all these polypeptide chains are precisely known. Yet it is still not clear how the sequence of amino acids determines the spatial conformation of the polypeptide chain (ANFINSEN and SCHERAGA 1975; KIM and BALDWIN 1982; KING 1986). A variety of experimental approaches, including sequence comparisons of proteins with homologous structure such as hemoglobins and ribonucleases (DICKERSON and GEIS 1983; KREBS, SCHMID and JAENICKE 1983), indicate that many positions in a chain can tolerate considerable variation without altering the folding of the protein.

Deciphering the rules through which amino acid sequences determine conformation will require identifying those residues which carry conformational information, and those which do not. This requires an experimental system in which genetic defects in folding and maturation can be distinguished from defects in function and stability. The P22 tailspike endorhamnosidase is one such protein.

The tailspike polypeptide chain is the product of gene *9* of P22 (BOTSTEIN, WADDELL and KING 1973). The chain has 666 amino acids (SAUER *et al.* 1982), and exists as a trimer in its native state (GOLDENBERG, BERGET and KING 1982). Six trimeric spikes attach to phage heads to convert them into infectious phage particles (ISRAEL, ANDERSON and LEVINE 1967). The tailspike has a glycosidic activity which hydrolyzes the O-antigen of *Salmonella typhimurium* (IWASHITA and KANEGASAKI 1976). Though its 3-dimensional structure has not yet been solved, Raman spectroscopy reveals that the secondary structure of the tailspike is more than 50% beta sheet (G. THOMAS, personal communication). The native spike is thermostable up to 80° (GOLDENBERG and KING 1981).

As a result of the unusual stability of the native tailspike, and the physiology of late protein synthesis in phage infected cells, it has been possible to characterize the *in vivo* folding and subunit association pathway for the tailspike (GOLDENBERG and KING 1982). As shown in Figure 1, newly synthesized chains form a partially folded early intermediate. This yields a species capable of specific chain recognition, which associates to form the protrimer species. The protrimer, a relatively long lived intermediate, folds further to form the native spike. The early intermediate is thermolabile (GOLDENBERG, BERGET and KING 1982), while the terminal protrimer to native transition is cold sensitive. As a result, intermediates are well separated both energetically and kinetically from the native structure.

Gene *9* is the locus of more than 30 sites of temper-

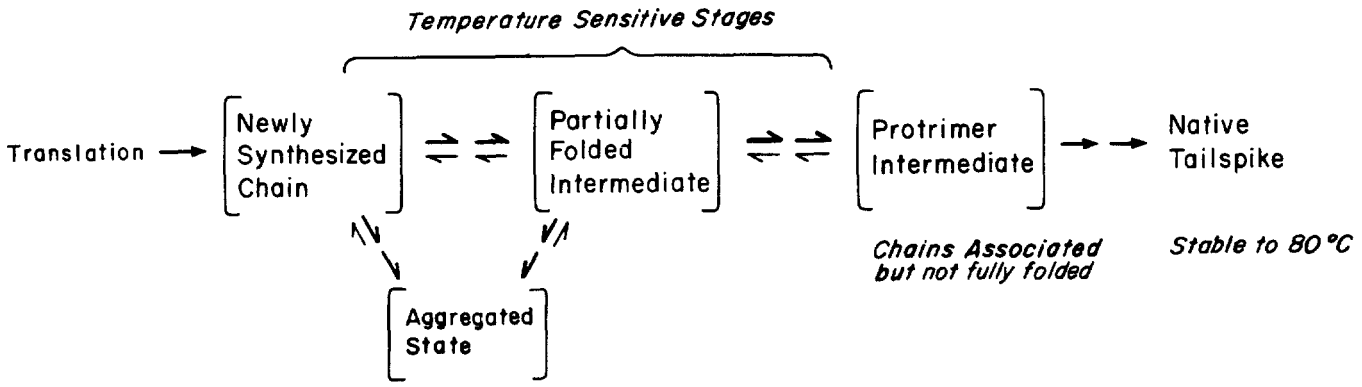

FIGURE 1.—Pathway for the folding and subunit association of the gene 9 tailspike endorhamnosidase. Newly synthesized polypeptide chains fold partially into a conformation that has enough structure for chain-chain recognition (GOLDENBERG, SMITH and KING 1983). These species interact to form the protrimer, in which the chains are associated but not fully folded (GOLDENBERG and KING 1982). The protrimer folds further to form the native spike, which is stable above 80° and protease and detergent resistant (GOLDENBERG, BERGET and KING 1982). The early intermediates are thermolabile and are probably the loci of the *tsf* mutations. Chain folding processes occur both before and after chain association, so that there is no species corresponding to a "native" monomer.

ature sensitive folding mutants (*tsf*) (SMITH, BERGET and KING 1980; KING *et al.* 1986). The *tsf* mutants are a class of *tss* mutants (SADLER and NOVICK 1965). These mutations specifically destabilize intermediates in the folding pathway but do not alter the functions or thermostability of the native protein once correctly folded at the permissive temperature (SMITH and KING 1981; GOLDENBERG and KING 1981; GOLDENBERG, SMITH and KING 1983). The predominance of *tsf* phenotypes presumably reflects two features of the system: the existence of thermolabile maturation intermediates for this thermostable native protein (GOLDENBERG, SMITH AND King 1983) and the wide temperature range (18–42°) over which phage and bacterial proteins must mature productively. Mutations altering the function of the tailspike without altering its folding pathway are rare; only a single such mutation has been reported (BERGET and POTEETE 1980).

The amino acid substitutions and local sequence at many of the *tsf* sites have been determined by DNA sequencing (YU and KING 1984; KING *et al.* 1986; R. VILLAFANE and J. KING, personal communication). The *tsf* sites are predominantly substitutions of glycines and hydrophilic residues. Electrophoretic characterization of purified native forms of mutant proteins indicate that more than 40% of the *tsf* substitutions occur at the surface of the protein and may mark sites of turns in the polypeptide chain (KING, HAASE and YU 1987; M.-H. YU and J. KING, personal communication). At restrictive temperature the substitutions interfere with the formation of the correct local conformation. Once past the thermolabile intermediate in the folding pathway, the surface location permits the accommodation of the mutant residue in the native structure.

Many of the classic studies in bacterial genetics focused on mutants which maintained a subset of the activities of the mature protein such as CRM^+ (CRAWFORD and JOHNSON 1963). Mutant polypeptide chains completely lacking biological activity have received only limited attention (KING and YU 1986). The *tsf* mutations prevent the expression of all the activities of the polypeptide chain at restrictive temperature (SMITH and KING 1981), consistent with a block in chain folding or association. This suggests that absolute lethal mutations might identify residues critical for chain folding or association over the entire temperature range of phage growth.

The analysis of the missense phenotypes of nonsense mutations provided a useful approach (MILLER *et al.* 1979). For nonsense mutations the lethal phenotype on the su^- host is a property of the translational stop and does not depend on the amino acid interactions in the polypeptide chain. If a mutant is recovered on the su^+ host which inserts the wild-type residue at a site of interest, very little selection will have been exerted with respect to either folding or function. These strains can then be tested on host strains inserting non-wild-type residues at the nonsense codon. The insertion of a non-wild-type residue might yield lethal, ts, other missense, or no defective phenotypes.

In those cases where the amber mutants have been recovered on hosts that do not insert the wild-type amino acid, there will be selection for sites which can tolerate the inserted residue. These sites, however, can still be tested for their phenotype with other amino acids at the site.

Miller and co-workers used nonsense mutations (MILLER *et al.* 1977; COULONDRE and MILLER 1977; MILLER, COULONDRE and FARABAUGH 1978) to genetically dissect the role of local regions of the polypeptide chain in the various functions of the *lac* repressor (SCHMITZ, COULONDRE and MILLER 1978; SOMMER *et al.* 1978, MILLER *et al.* 1979). The missense proteins synthesized on different tRNA suppressor hosts were assayed for the various activities of the

TABLE 1

***Salmonella typhimurium* strains used in these studies**

Salmonella strain	Genotype	Reference or source
DB7136	*leu A414(am), his C525(am)*	WINSTON, BOTSTEIN and MILLER (1979)
DB7154	*leu A414(am), his C525(am), supD10(ser)*	WINSTON, BOTSTEIN and MILLER (1979)
DB7155	*leu A414(am), his C525(am), supE20(gln)*	WINSTON, BOTSTEIN and MILLER (1979)
DB7156	*leu A414(am), his C525(am), supF30(tyr)*	WINSTON, BOTSTEIN and MILLER (1979)
DB7157	*leu A414(am), his C525(am), supJ60(leu)*	WINSTON, BOTSTEIN and MILLER (1979)
GW1	*his⁻, gal⁻, pKM101(ampʳ)*	G. WALKER
HMB54	*leu A414, his C525, supD10, pKM101*	This paper
HMB55	*leu A414, his C525, supE20, pKM101*	This paper
HMB56	*leu A414, his C525, supF30, pKM101*	This paper
PB7577	*leu A414, ampʳ, delpMC5*	P. BERGET
PB7580	*leu A414, ampʳ, delpMC2*	P. BERGET
PB7582	*leu A414, ampʳ, delpMC13*	P. BERGET

repressor. Many of the substitutions affected specific functions. These experiments, however, were not designed to differentiate defects in the folding of a domain with defects in domain function.

Here, we present the isolation, mapping, and preliminary phenotypic characterization of nonsense mutations in gene *9*. By taking advantage of properties of the tailspike, we have been able to determine whether the missense proteins synthesized in tRNA suppressor *Salmonella* hosts were defective in chain folding and association, or in the functions of the mature protein. The nonsense mutations also provide starting points for the further isolation of am^+ pseudorevertants and second site suppressors of the missense phenotypes (JARVIK and BOTSTEIN 1975). The results indicate that such mutants can be used to map interactions among different local sequences involved in directing the chain folding and subunit association pathway.

MATERIALS AND METHODS

Phage and bacterial strains: Bacterial strains are all derivatives of *Salmonella typhimurium* LT2. The genotypes of the bacterial host strains are summarized in Table 1.

The HMB *Salmonella* strains are the progeny of mating the corresponding su^+ hosts (DB7154, DB7155 and DB7156) with GW1. The protocol is that described by DAVIS, BOTSTEIN and ROTH (1980). Recombinant cells were recognized by their ability to grow on minimal media, supplemented with ampicillin, with galactose as the carbon source.

P22 *cl-7* was the reference phage for most experiments. The preexisting amber mutants in gene *9*, *amN110* and *amN108*, are described in BOTSTEIN, CHAN and WADDELL (1972) and SMITH, BERGET and KING (1980). The mutants *amH1200*, *amH1291*, *amE1017* and *amH1014* are described in POTEETE and KING (1977) and SMITH, BERGET and KING (1980). BERGET and POTEETE (1980) further characterized the suppressed missense proteins produced by *amH1014*. YU and KING (1984) have determined the sequences for *amE1017* and *amH1200* and the sequence for an am^+ts revertant of *amH1291* (M.-H. YU and J. KING, personal communication).

Media and chemicals: The media for these experiments are those described in SMITH, BERGET and KING (1980). Radiolabeled (^{14}C) protein markers were purchased from New England Nuclear.

Isolation of new P22 amber mutants: Standard procedure for plating phage consisted of adding two drops (approximately 0.1 ml) of exponentially growing cells concentrated to 2.0×10^9 cells/ml suspended in Luria broth and placed into 2.0 ml of soft agar. A 10-ml aliquot of P22 (*cl-7*) phage, suspended in M9 salts and 2 mM $MgSO_4$ at a concentration of 2.0×10^8 phage/ml, was irradiated in a 15-ml petri dish for 120 sec at a fixed energy output of 1.5 joules/sec/m^2. Phage survival was 0.05% on HMB (pKM101) strains. To increase the survival of irradiated phage, the host was irradiated prior to infection (WALKER 1978). HMB cells were grown to a concentration of 1.0×10^8 cells/ml, concentrated twofold, resuspended in M9 salts and 2 mM $MgSO_4$ and then irradiated for 10 sec. This treatment resulted in 50% of the cells surviving. Survival of irradiated phage under these conditions was 1.0%. This protocol performed on P22 c^+ phage yielded 2.9% clear mutants among the surviving phage.

We utilized the distinctive plaque morphology of amber mutants on double layer plates to aid in the identification of such mutants. The plates consisted of a bottom su^- lawn incubated for three hours at 30° before the addition of the upper su^+ lawn. The upper su^+ lawn contained 100 infectious centers/ml (irradiated phage allowed to absorb to irradiated HMB hosts) and 2.0×10^8 of the corresponding su^+ cells (without pKM101) before incubation. The plates incubated for 14 hr at room temperature. We screened the plates for turbid plaques resembling the morphology of known gene 9^- amber mutants. These were then stabbed into a set of indicator lawns; su^-, su^+ and su^- with exogenously added native P22 tailspikes. The ability of native tailspikes to rescue amber phage *in situ* was the criterion for identifying 9^- amber mutants.

We were particularly interested in isolating amber mutations that could be suppressed by aromatic residues. The largest set of amber mutants, 90, were isolated on DB7156 (tyrosine), 11 were isolated on DB7154 (serine) and nine on DB7155 (glutamine). We failed to recover amber mutants on DB7157 (leucine). Since the plates were incubated at room temperature, we suspect that the lack of recovery was a consequence of the cold sensitive mutant tRNA gene of this host (WINSTON, BOTSTEIN and MILLER 1979). The letters "S" (serine), "G" (glutamine) and "T" (tyrosine) found

in the names of the new amber phage mutants indicate the amino acid inserted by the suppressor strain used for the mutant's isolation. Hence, *amUT51* was the 51st amber mutant isolated and it was isolated on DB7156 (*supF*) inserting tyrosine.

Enrichment for 9^- amber mutants: HMB host cells were infected with phage, prepared as described above, at an m.o.i. of 0.01 in LB at 30°. The progeny of this permissive infection was used to infect su^- cells, also at an m.o.i. of 0.01. Those su^- cells infected with 9^- amber mutants yield mature phage heads without tails. Phage can absorb to cells but heads, lacking tails, cannot. Equal volumes of exponentially growing cells and the su^- phage lysates were mixed together and allowed to incubate for 10 min at 37°. The infectious centers were removed by centrifugation. After three rounds of absorption, the phage titer dropped to 0.1% of the starting titer. The heads present were rescued by adding exogenous tailspikes (2.0×10^7 tailspikes/ml) *in vitro* to the enriched lysates.

Suppression patterns: To determine the initial suppression patterns of the 9^- amber mutants the efficiencies of plating of the mutant stocks were measured on each of the four su^+ hosts at 30°, 37° and 39°. For the most restricted mutants, 10^7 pfu were plated. Assays for cold sensitivity were conducted in separate experiments plating at 21°. DB7157 was not included in assays to examine cold sensitivity since the tRNA suppressor itself is cold sensitive (WINSTON, BOTSTEIN and MILLER 1979). In order to reduce variations due to moisture in these plating assays, all utilized plates were between 3 and 5 days old. Plates incubated at elevated temperatures (37° and 39°) were prewarmed to the assay temperatures before incubation. All 12 assay conditions for each mutant (4 su^+ hosts × 3 temperatures) were conducted at the same time to limit variations in media or incubation temperatures that could affect plaque formation and morphology. All assays were conducted with fresh plating bacteria prepared concurrently.

Deletion lysogen mapping: For these experiments, we plated both 4.0×10^8 or 4.0×10^6 pfu/plate on each of the three deletion lysogen hosts at 30°. A titer fivefold higher than that obtained for each mutant on the su^- (DB7136) host served as the standard for rescue. The same precautions, described in the previous section, to reduce experimental variability were also employed in these experiments. Furthermore, a set of sequenced (YU and KING 1984; M.-H. YU and J. KING, personal communication) and/or historically well characterized (SMITH, BERGET and KING 1980) amber mutations served as a positive control for recombination within each deletion interval.

Additional genetic assays: To ensure that the 9^- amber mutants carried no additional *ts* mutations outside of gene *9*, we plated 50–200 pfu of each new amber mutant on the su^- host with exogenously added native tailspikes at 39°. To identify those mutants that might carry an additional 9^- *ts* mutation, we examined the am^+ reversion frequencies at 39°.

Phage crosses: The procedure for phage crosses is described in SMITH, BERGET and KING (1980) except su^+ hosts were used. After phage absorption, the infections were diluted 1/200 and constantly aerated.

Radiolabeled lysates and SDS gel electrophoresis: These protocols are described in SMITH, BERGET and KING (1980).

Hybrid phage: The hybrid phage (mutant genomes with wild-type tailspikes) used for one step growth experiments were constructed by adding equal volumes of $5^-am/13^-am/cl\text{-}7$ su^- phage lysates and 9^- (*am*) su^- phage lysates. Phage were concentrated by centrifugation. Background contamination of $5^-/13^-$ phage was typically 10^{-4} and easily recognized by plaque morphology. In growth experiments, the titer of mature phage heads was determined by plating with exogenously added native tailspikes.

TABLE 2

Codons in gene *9* that can mutate to amber via one base change[a]

Residue	Codon	Amino acid[b]	Residue	Codon	Amino acid	Residue	Codon	Amino acid
21	TCG	ser	163	AAG	lys	406	TTG	leu
31	TAC	tyr	202	TGG	trp	412	TAC	tyr
34	CAG	gln	205	AAG	lys	417	TAC	tyr
45	CAG	gln	207	TGG	trp	421	CAG	gln
49	TAC	tyr	212	CAG	gln	446	AAG	lys
53	GAG	glu	213	TGG	trp	449	TAT	tyr
59	CAG	gln	232	TAT	tyr	464	TAC	tyr
62	CAG	gln	233	CAG	gln	483	AAG	lys
74	TAC	tyr	239	TAC	tyr	486	CAG	gln
85	CAG	gln	283	TTG	leu	489	CAG	gln
92	TAT	tyr	291	AAG	lys	491	TAT	tyr
98	CAG	gln	315	TGG	trp	512	CAG	gln
101	TAT	tyr	317	AAG	lys	555	TAT	tyr
106	TTG	leu	320	TAT	tyr	566	AAG	lys
107	AAG	lys	328	TAT	tyr	569	AAG	lys
108	TAC	tyr	335	CAG	gln	576	TTG	leu
113	TAT	tyr	355	TAT	tyr	577	TAC	tyr
122	AAG	lys	359	GAG	glu	590	TAT	tyr
123	TAT	tyr	365	TGG	trp	631	AAG	lys
130	TAT	tyr	371	TCG	ser	637	TTG	leu
133	TTG	leu	377	TAT	tyr	639	TAT	tyr
134	CAG	gln	384	TCG	ser	640	TGG	trp
150	TAT	tyr	388	TAC	tyr	648	AAG	lys
153	TAT	tyr	391	TGG	trp	663	TTG	leu
156	GAG	glu	405	GAG	glu			

[a] From the complete sequence of gene *9* as determined by SAUER *et al.* (1982).

[b] Totals: 27 tyrosines, 14 glutamines, 12 lysines, 7 tryptophans, 7 leucines, 4 glutamates and 3 serines.

RESULTS

Isolation of new 9^- amber mutants: The sequence of gene *9* reported by SAUER *et al.* (1982) contains 74 codons that can mutate to the amber codon via one base change (Table 2). Previous isolations of P22 amber mutants (BOTSTEIN, CHAN and WADDELL 1972; POTEETE and KING 1977) yielded ten independent sites (SMITH, BERGET and KING 1980). In order to obtain mutations at the majority of the other potential amber codons, we mutagenized P22 particles with ultraviolet light, and screened for amber mutants with double overlay plates. Amber candidates were picked from all four su^+ hosts, though the largest set was from the tyrosine inserting host.

Of 150,000 mutagenized *cl-7* plaques screened, 334 behaved like true amber mutants. To identify phage carrying amber mutations in gene *9*, the amber mutants were further tested for *in situ* complementation with native tailspikes by plating approximately 300 pfu on an su^- host with exogenously added native tailspikes. Forty-five were rescued, indicating that they carried their amber mutations within gene *9* and

did not carry amber mutations in other essential genes.

A second set of mutagenized phage was subjected to an enrichment for gene *9*⁻ amber mutants. HMB host cells were infected with the mutagenized phage. The progeny of this permissive infection was then used to infect su⁻ cells, at low multiplicity. Those cells infected with phage carrying nonsense mutations in gene *9* produce phage heads, lacking tails, which are unable to absorb to host cells. This lysate was incubated with *Salmonella* host cells and the resulting phage/bacteria absorption complexes were removed by centrifugation, enriching the remaining supernatant for phage heads. After three such rounds of absorption and centrifugation purified tailspikes were added to the final lysate to rescue the phage heads remaining. The infectious phage formed were screened as described above. Of 15,000 plaques, 86 turned out to be amber mutants and 66 were in gene *9* by *in situ* complementation with exogenously added tailspikes.

All of the 111 newly isolated *9*⁻ amber mutants had am⁺ reversion frequencies consistent with a single amber mutation (10^{-7}). To exclude strains that carried an additional *ts* mutation outside of gene *9*, the *9*⁻ amber strains were tested for their ability to be rescued by exogenously added tailspikes, when plated at 39° on an su⁻ host. Eight mutants were not rescued in the assay indicating that they carried an additional *ts* mutation outside of gene *9*.

Since gene *9* is a locus of *ts* mutations, it seemed likely that some of the amber mutants carried additional *ts* mutations within gene *9*. To identify such putative double mutants, we examined the reversion frequencies to *am*⁺ at 39°. If the amber strains carried an additional *ts* allele in gene *9*, the strains should exhibit very low reversion frequencies at restrictive temperature. Nine of the strains displayed reversion frequencies less than 3.0×10^{-8} at 39°, but in the normal range at 30°. These strains could either carry *ts* mutations in gene *9*, or they could be single amber mutants in which *ts* pseudoreversions of the amber codon occur at a higher frequency than reversion events to the wild type phenotype.

The number of distinct amber sites: The set of gene *9*⁻ amber alleles was too large to directly determine the number of distinct amber sites by two-factor crosses. As an initial means of distinguishing ambers at different sites, we used the patterns of amber phage growth on *Salmonella* hosts carrying different suppressor tRNAs. Though our primary interest in the phenotypes of the amber mutants on different suppressors was to assess the role of the wild type amino acid residue at the amber site, the growth patterns were very useful for categorizing the mutants. We determined the efficiency of plating of each of the 94 new amber mutant stocks and six of the previously isolated amber mutant stocks on all four tRNA suppressor hosts at 30°, 37° and 39°. The reference phage *cI-7* plated with equal efficiency at the two lower temperatures. At 39° the plating efficiency was reduced by about 50%, due to the inefficiency of tailspike maturation at higher temperatures (GOLDENBERG and KING 1982).

TABLE 3

Deletion lysogen mapping of the *9*⁻ amber mutants in the interval between amino acid residues 216 and 417

Mutant	Percent permissive titer[a]: On PB7577 a.a. 91–215	On PB7582 a.a. 216–417	On DB7136 su⁻ host
amUT51	7.2×10^{-3}	9.1×10^{-5}	4.6×10^{-5}
amUT111	1.6×10^{-2}	$<1.2 \times 10^{-4}$	$<1.2 \times 10^{-4}$
amUT96	1.7×10^{-2}	4.7×10^{-5}	4.7×10^{-5}
amUT52	2.0×10^{-2}	$<2.9 \times 10^{-5}$	$<2.9 \times 10^{-5}$
amUS16	2.0×10^{-2}	8.8×10^{-4}	2.8×10^{-3}
amUT102	2.5×10^{-2}	9.1×10^{-5}	1.3×10^{-4}
amUT97	2.7×10^{-2}	9.1×10^{-5}	1.9×10^{-6}
amUT25	3.6×10^{-2}	2.7×10^{-4}	2.7×10^{-4}
amUS22	4.0×10^{-2}	$<8.2 \times 10^{-5}$	8.2×10^{-5}
amUT42	4.8×10^{-2}	1.8×10^{-4}	3.6×10^{-4}
amUT79	9.8×10^{-2}	3.0×10^{-4}	2.1×10^{-4}
amUS14	1.2×10^{-1}	7.1×10^{-4}	6.0×10^{-4}
amUT86	1.2×10^{-1}	5.1×10^{-5}	$<5.1 \times 10^{-5}$
amE1017	1.2×10^{-1}	3.5×10^{-4}	2.3×10^{-3}

[a] The most permissive titer was determined on each mutant's most permissive host at 30°.

The plating efficiencies of the mutants under these conditions varied over six orders of magnitude. The majority of strains failed to plaque with normal efficiency under at least one of the growth conditions. Many displayed *ts* phenotypes on one or more suppressor hosts. Some strains were fully restricted, so that the phage titer corresponded to the number of *am*⁺ revertants as determined by plating on the su⁻ host. Examples of these growth patterns are summarized in Table 5, below. Strains designated *UT* were initially isolated with the tyrosine inserting strain as the su⁺ host, strains designated *US* were isolated from the serine inserting host, and those designated *UG* were isolated from the glutamine inserting host.

The patterns of growth on the four different tRNA suppressor hosts at the three temperatures defined 38 distinct suppressed phenotypes. This indicated that at least 38 different sites were represented among the 100 amber mutants.

To roughly order the amber mutations, we mapped them against a set of three P22 lysogens carrying deletions with different endpoints in gene *9* (P. BERGET, personal communication). The three strains of *Salmonella* P22 deletion lysogens used in the marker rescue experiments divided the gene into four deletion intervals. The titer of each amber mutant on the su⁻ host served as a control for *am*⁺ reversion and

TABLE 4

Recombination between amber mutants with similar suppressed phenotypes and mapping within the same deletion interval

	Titer/% recombination					
	amUG3	*amUT23*	*amUT26*	*amUT27*	*amUT32*	*amUT36*
amUT36	2.0×10^{-4} (0.040)	8.9×10^{-7} (0.000)	6.5×10^{-8} (0.000)	6.3×10^{-4} (0.13)	4.5×10^{-7} (0.000)	1.4×10^{-6}
amUT32	7.4×10^{-5} (0.015)	3.8×10^{-7} (0.000)	4.1×10^{-7} (0.000)	2.9×10^{-4} (0.059)	5.5×10^{-7}	
amUT27	3.4×10^{-4} (0.068)	9.3×10^{-4} (0.19)	3.9×10^{-4} (0.078)	7.1×10^{-7}		
amUT26	1.4×10^{-4} (0.028)	4.4×10^{-7} (0.000)	4.0×10^{-7}			
amUT23	1.2×10^{-4} (0.023)	1.8×10^{-7}				
amUG3	1.2×10^{-6}					

background plaque formation. Rescue presumably arises from recombination between the mutant phage and the phage genome in the deletion lysogen host. A fivefold increase in am^+ recombinants over the background was the criterion for rescue of the amber mutant by the deletion lysogen. The nine strains that displayed reversion frequencies less than 3.0×10^{-8} at 39° were included in these assays.

For those mutants mapping within a particular deletion interval, the extent of recombination with the proceeding deletion lysogen endpoint provides a rough indication of the ordinal location of the amber mutations within the interval (CHAN and BOTSTEIN 1972). Table 3 presents the experimental data for 14 amber mutations mapping within the deletion interval that spans from amino acid 216 to amino acid 417.

Those alleles that mapped within the same deletion interval and also displayed the same pattern of suppression, were crossed by each other to determine whether the amber mutations were separable by recombination. The nine strains displaying low am^+ reversion frequencies at 39° were also included in these assays since recombination between their amber alleles could be scored independently of any additional *ts* alleles. Five of these mapped to the same interval and displayed identical suppression patterns. The recombination values between two sequenced mutations *amH1200* and *amH1291*, single base changes that lie 12 nucleotides apart (YU and KING 1984; M.-H. YU and J. KING, unpublished data) provided a reference for the sensitivity of the crosses. This cross typically yielded values of recombination of 1.0×10^{-4} which is fivefold over the background of am^+ revertants emerging from the infections of either mutant alone. Table 4 displays the results of two-factor crosses between a set of six amber mutants with similar suppressed phenotypes that map within the first deletion interval, corresponding to the N-terminus of gene *9*. If the crosses did not yield am^+ progeny at a frequency fivefold higher than that obtained from simultaneous infections of each mutant alone, the amber mutations were assigned to a single site.

From this set of crosses we concluded that *amUT23*, *amUT26*, *amUT32* and *amUT36* are not separable by recombination and assigned them to the same amber site. But *amUT27* and *amUG3* appear to represent separate sites from the other cluster.

Similar crosses were carried out for all sets of ambers not resolved by the deletion mapping and suppression patterns. These results taken together indicated that the 109 amber mutations defined 64 different sites in gene *9*. The results are summarized in the genetic map shown in Figure 2.

Suppressed phenotypes of 9^- amber mutants: As noted above, the analysis of the primary suppression patterns indicated at least 38 distinct suppressed phenotypes. Table 5 displays the suppression patterns of the canonical amber mutant at each of the 64 sites within the gene. The su^+ hosts insert either serine, glutamine, tyrosine or leucine at the amber site (WINSTON, BOTSTEIN and MILLER 1979). The suppressed phenotypes range from those amber mutants which grow on all tRNA suppressor hosts at all temperatures (*amUT102*) to those amber mutants that grow only on one tRNA suppressing host (*amUT51* and *amUT105*). A "++" indicates that neither the plaque morphology nor the eop of the mutant was significantly affected by amino acid insertion or temperature ranges. The "+" symbol also signifies that the mutant was permissive for growth under the designated conditions. In these cases, however, the size of the plaque was reduced and/or the eop dropped into the 10^{-1} to 10^{-2} range. A "+−" indicates that the eop dropped below 10^{-2} but remained above the 10^{-3} value. With all the mutants assigned a "+−" symbol the plaque size was considerably reduced under the conditions for which

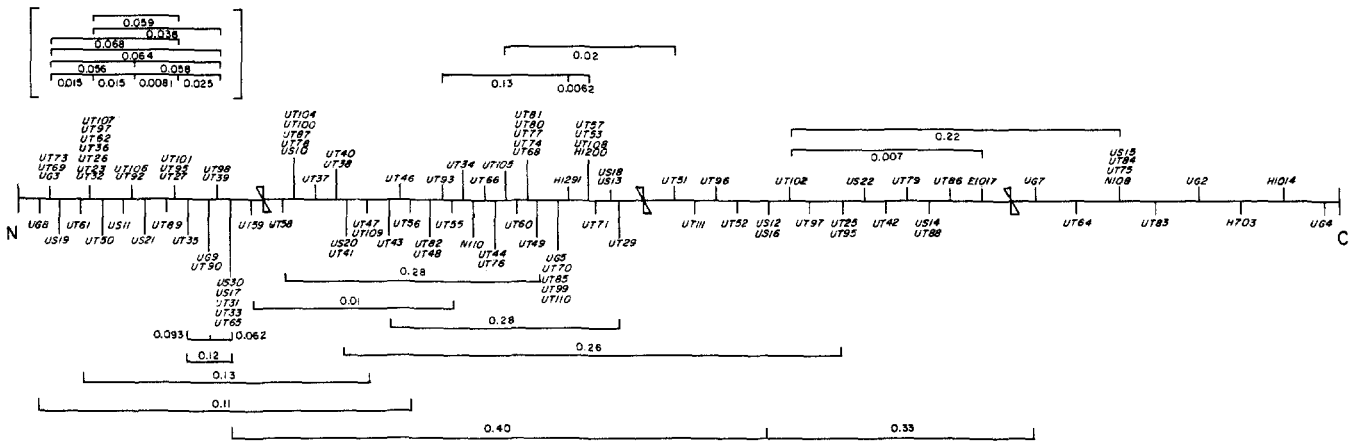

FIGURE 2.—Genetic map of gene *9* amber mutations. This map summarizes three lines of experimentation: deletion lysogen mapping, suppression patterns and two-factor crosses. Mutations appearing at the same place on the map were not separable by recombination and displayed the same suppressed phenotype. The deletion lysogen endpoints are marked with *double triangles*. Mutations falling after the first deletion interval were placed on the map in order of their recombination with the preceding deletion lysogen endpoint. The brackets above and below the map contain the percentages of recombination for a sampling of the two-factor crosses.

the symbol is placed. A "–" indicates an eop of 10^{-3} to 10^{-5}. Those conditions where growth was restricted to the level of the reversion frequency on an su^- host, with eop's less than 10^{-6}, are designated with a "––" symbol.

Figure 3 depicts the range of phenotypes observed on just one su^+ host, DB7154, which inserts serine at the amber site. The efficiencies of plating have been normalized to the highest titer obtained as a function of temperature. The curve of the wild-type P22 (*cl-7*) in Figure 3D demonstrates the temperature-dependent maturation of the wild-type tailspike (GOLDENBERG, BERGET and KING 1982). Figure 3D also displays an example of an amber mutant, *amUT102*, which grows on all su^+ hosts at all experimental temperatures. We refer to such sites as "tolerant." The am^+ reversion frequency, as a function of temperature, served as an internal control for background am^+ revertant phage. The titer of *amUT102* on the su^- host is also presented in Figure 3D.

Figure 3A illustrates the variety of *ts* phenotypes found among the gene *9* amber mutants. The yield of *amUT51* is already severly reduced at 30°, the permissive temperature for most gene *9 tsf* mutants (SMITH, BERGET and KING 1980). We refer to this phenotype as "hyper-ts." *AmUT51* shows a hyper-ts phenotype only when serine is inserted at the amber site. Glutamine and leucine at this site confer a lethal phenotype, independent of temperature (Table 5). Tyrosine insertion suppresses the amber mutation at all temperatures. The efficiency of plating of *amUT93* is maximal at 30°, but is sharply reduced at 37° and 39°. Only tyrosine at the amber site suppresses this mutant at elevated temperatures. In contrast *amUS14* is only restricted at 39°.

Figure 3C illustrates the behavior of two amber mutants, *amUT64* and *amUT83*, which display a lethal phenotype on the serine-inserting su^+ host. At all temperatures the eop was four orders of magnitude lower than the permissive titer, which for these strains was determined on DB7156 (tyrosine), the only host which suppresses these mutants at elevated temperatures.

Figure 3B displays the complex phenotype of *amH703*, both cold sensitive and temperature sensitive. Amber mutants, representing 43 distinct amber sites, were examined for a *cs* suppressed phenotype. *AmH703* and *amUT52* were the only mutants to exhibit *cs* phenotypes. The restriction of growth at lower and higher temperatures was a function of the amino acid inserted at the *amH703* site. Both were alleviated when the mutant was plated on DB7155 (glutamine).

Thirty-three sites are classified as tolerant, with eop's above 10^{-1} (10% of the most permissive titer) on tRNA suppressor hosts inserting serine, glutamine, tyrosine, or leucine at 30° and 37°. A subset of these are more restricted in growth at 39° but the eop's never drop as low as those obtained on the su^- host. Mutants at 16 of the amber sites were tightly restricted on at least one tRNA suppressor host at elevated temperatures. Mutants at nine sites were lethal on one or more tRNA suppressor hosts and finally, six amber mutants had both lethal and temperature sensitive tRNA suppressing hosts.

Identification of amber fragments: To determine more precisely the location of the amber sites, we measured the molecular weight of the amber fragments produced in the su^- host. Lysate samples were electrophoresed through acrylamide gels of varying concentrations, depending on the expected size of the fragments. Autoradiograms were prepared and examined for the presence of a band absent from the gel pattern of the control *cl-7* lysate. Lysates were prepared with the 64 canonical amber mutants. All of the mutant lysates lacked the band corresponding to the tailspike polypeptide chain. In 33 of the lysates an

TABLE 5

Patterns of amber mutant growth as a function of host amber suppressor tRNA and temperature for 64 distinct amber sites within gene9[a]

	Inserted amino acid at the amber site:													
	Serine			Glutamine			Tyrosine			Leucine			Deletion	
Mutant	30°	37°	39°	30°	37°	39°	30°	37°	39°	30°	37°	39°	interval	Alleles
amUT102	++	++	++	++	++	++	++	++	++	++	++	++	216-417	1
amUT27	++	++	++	++	++	++	++	++	+	++	++	++	001-090	3
amUT92	++	++	++	++	++	++	++	++	+	++	++	++	001-090	1
amUS15	++	++	++	++	++	++	++	++	+	++	++	+	438-666	4
amE1017	++	++	+	++	++	++	++	++	+	++	++	++	216-417	1
amUG3	++	++	++	++	++	+	++	++	+	++	++	+	001-090	3
amUT23	++	++	+	++	++	+	++	++	+	++	++	++	001-090	7
amUS10	++	++	+	++	++	+	++	++	+	++	++	+	091-215	5
amUT39	++	++	+	++	++	+	++	++	+	++	++	+	001-090	2
amUT55	++	++	++	++	++	−	++	++	++	++	++	++	091-215	1
amUG8	++	++	++	++	++	++	++	++	+	++	++	−	001-090	1
amUG9	++	++	++	++	++	++	++	++	−	++	++	+	001-090	2
amUS12	++	++	+	++	++	++	++	++	−	++	++	++	216-417	2
amUT35	++	++	++	++	++	++	++	++	−	++	++	+	001-090	1
amUT59	++	++	++	++	++	−	++	++	++	++	++	+	001-090	1
amN110	++	++	++	++	++	++	++	+	−	++	++	+	091-215	1
amUG7	++	++	++	++	++	+	++	++	−	++	++	+	418-437	1
amUS17	++	++	+	++	++	+	++	++	−	++	++	+	001-090	5
amUT38	++	++	−	++	++	+	++	++	+	++	++	+	091-215	2
amUT47	++	++	−	++	++	−	++	++	++	++	++	++	091-215	2
amUS19	++	++	++	++	++	++	++	+	−	++	+	−	001-090	1
amUT49	++	++	+	++	++	−	++	++	++	++	++	−	091-215	1
amUT56	++	++	++	++	++	+	++	++	−	++	++	−	091-215	1
amUT58	++	++	++	++	++	−	++	++	+	++	++	−	091-215	1
amUT61	++	++	−	++	++	−	++	++	++	++	++	+	001-090	1
amUT43	++	++	+	++	++	−	++	++	−	++	++	+	091-215	1
amUT29	++	++	++	++	+	−	++	++	−	++	++	+−	091-215	1
amUT37	++	++	−	++	++	+−	++	++	+	++	++	−	091-215	1
amUT46	++	++	−	++	++	−	++	++	++	++	++	−	091-215	1
amUT79	++	++	−	++	++	−	++	++	−	++	++	++	216-417	1
amUT50	++	++	++	++	+	−	++	++	−	++	++	−	001-090	1
amUS20	++	++	−	++	++	−	++	++	−	++	++	−	091-215	1
amUT25	++	++	−	++	++	−	++	++	−	++	++	−	216-417	2
amUS14	++	++	−−	++	++	+	++	++	−	++	++	+	216-417	2
amUT111	++	++	−	++	−	−−	++	++	++	++	+	−−	216-417	1
amUS21	++	++	++	++	−	−	++	−	−	++	−	−	001-090	1
amUT42	++	−	−	++	−	−−	++	−	−	++	+	−	216-417	1
amUG5	++	−−	−−	++	++	−	++	++	−	++	+	−	091-215	5
amUT44	++	−−	−−	++	+	−	++	++	−	++	++	−	091-215	2
amUT96	++	+	−−	++	−−	−−	++	−	−	+	+	−	216-417	1
amUT97	++	−	−	++	−−	−−	++	++	−	+	−	−	216-417	1
amUS11	++	−−	−−	++	++	++	++	−−	−−	+	+	+−	001-090	1
amUT60	+	−−	−−	++	−−	−−	++	++	−−	++	+−	−−	091-215	1
amUT93	++	−−	−−	+	−−	−−	++	++	++	+	−	−−	091-215	1
amH1291	++	−−	−−	++	−−	−−	++	++	++	+	−−	−−	091-215	1
amUT57	+	−−	−−	++	−−	−−	++	++	−−	++	−−	−−	091-215	4
amUS22	++	−−	−−	++	−−	−−	+	−−	−−	++	−−	−−	216-417	1
amUG2	+	+	−	++	++	+	+−	−−	−−	++	++	+	438-666	1
amUT89	++	−	−−	++	−−	−−	++	−	−	+−	−−	−−	001-090	1
amUT66	++	++	+	−	−−	−−	++	++	+−	+	+	−	091-215	1
amUT64	−	−−	−−	++	−−	−−	++	++	++	+	−−	−−	438-666	1
amUT68	−	−−	−−	++	−−	−−	++	++	−−	++	−−	−−	091-215	5
amUT48	++	++	−	−	−−	−−	++	++	−	−	−	−	091-215	2
amUS13	++	++	+	−	−−	−−	++	−	−	−	−	−	091-215	2
amUG4	+	+	+	++	++	++	−−	−−	−−	++	++	++	438-666	1
amUT52	++	++	+−	−−	−−	−−	++	++	−	+	−	−	216-417	1
amH703	+	−−	−−	++	++	++	−−	−−	−−	++	++	++	438-666	1
amUT34	++	++	++	−−	−−	−−	++	+	−	+−	−	−	091-215	1
amUT71	++	++	++	−−	−−	−−	++	++	−	−	−	−	091-215	1

TABLE 5—Continued

Mutant	Inserted amino acid at the amber site: Serine 30°	Serine 37°	Serine 39°	Glutamine 30°	Glutamine 37°	Glutamine 39°	Tyrosine 30°	Tyrosine 37°	Tyrosine 39°	Leucine 30°	Leucine 37°	Leucine 39°	Deletion interval	Alleles
amUT83	–	–	––	––	––	––	++	++	–	+	–	–	438-666	1
amUT86	++	–	––	––	––	––	++	+	––	–	––	––	216-417	1
amH1014	––	––	––	++	++	++	––	––	––	++	++	++	438-666	1
amUT51	––	––	––	––	––	––	++	++	+	––	––	––	216-417	1
amUT105	––	––	––	––	––	––	++	++	––	––	––	––	091-215	1

[a] A "++" symbol indicates that neither the plaque morphology nor the eop of the mutant is significantly affected. The "+" symbol signifies that the mutant is still permissive for growth but either the plaque morphology is reduced and/or the eop drops into the 10^{-1} to 10^{-2} range. A "+–" indicates that the eop drops below 10^{-2} but remains above 10^{-3}; the plaque size is considerably reduced. A "–" designates an eop somewhere between 10^{-3} and 10^{-5}. These phenotypes are typically leaky. And finally, a "––" denotes a condition where growth is severely restricted with an eop less than 10^{-6}.

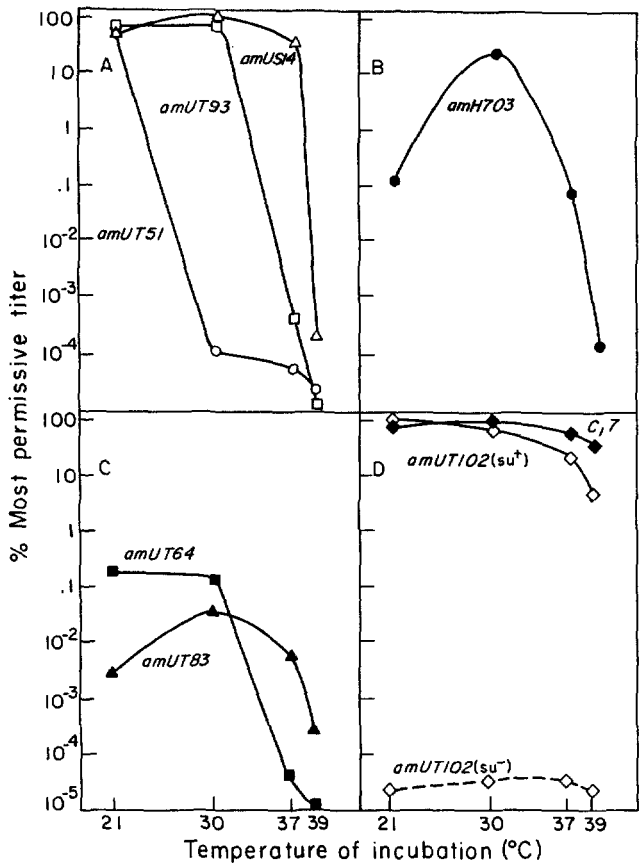

FIGURE 3.—Efficiencies of plating of different gene *9*⁻ amber mutants on the serine inserting suppressing host (DB7154). The efficiencies of plating have been normalized to the titer of each mutant's most permissive host at 30°. (A) The variety of *ts* phenotypes found among the amber mutants. (B) The complex cs/ts phenotype of *amH703*. Both defects are alleviated when this mutant is plated on DB7155 (glutamine). (C) The lethal phenotypes of two suppressed amber mutants. (D) The curve of wild-type *cI-7* and the temperature dependent maturation of the wild-type tailspike. Also shown is the curve of a tolerant amber *amUT102*; its efficiency of plating is host and temperature independent. The *dotted line* in (D) depicts the titer of *amUT102* on the su⁻ host (DB7136) as a function of temperature.

extra band appeared which we identified as the amber fragment.

Figure 4 displays an autoradiogram of an SDS gel with lysates of mutants arranged according to the mobility of the putative amber fragments. The molecular weights of the fragments were calculated using a set of five protein markers, shown in the outside lanes of the gel in Figure 4. The shortest amber fragment resolved was that produced by *amUS20* (autoradiogram not shown), estimated to be 9000 daltons, and only resolved in 20% acrylamide gels.

Of the 31 mutants for which we failed to detect an amber fragment, 16 mapped in the left most deletion interval and eight mapped to the beginning of the next deletion interval. The amber fragments of these mutants were probably smaller than the resolution limit of the smallest pore size gels used, 20%. The remaining seven mutants mapped between mutants with resolved fragments. The amber fragments of these mutants were most likely covered by other bands in the gel.

Phenotype clustering: Combining the amber fragment molecular weights with the genetic mapping data yielded a more physical map of the 64 *9*⁻ amber sites displayed in Figure 5. Mutants for which amber fragments were resolved are placed above the line; mutants placed below the line did not produce resolvable fragments. In the latter cases, the mutations were placed on the map by their extent of recombination with the proceeding deletion lysogen endpoint.

The map in Figure 5 also displays the suppressed phenotypes of the amber mutants. A striking feature of the map is the clustering of tolerant sites (*open circles*) in the amino terminus of the gene. The amber sites that produce missense proteins with strong ts phenotypes (*solid squares*) cluster in the middle of the gene where the majority of *tsf* mutations map (YU and KING 1984; M.-H. YU, R. VILLAFANE and J. King, personal communication). A significant number of the amber mutations which produce lethal missense polypeptides (*filled diamonds* and *solid triangles*) cluster among the most carboxy terminal amber sites. There is also a small lethal cluster in the 24–25,000-dalton region. Some of the mutants producing suppressed lethal polypeptides on one su⁺ host also produce *ts* polypeptides on other su⁺ hosts (*solid triangles*).

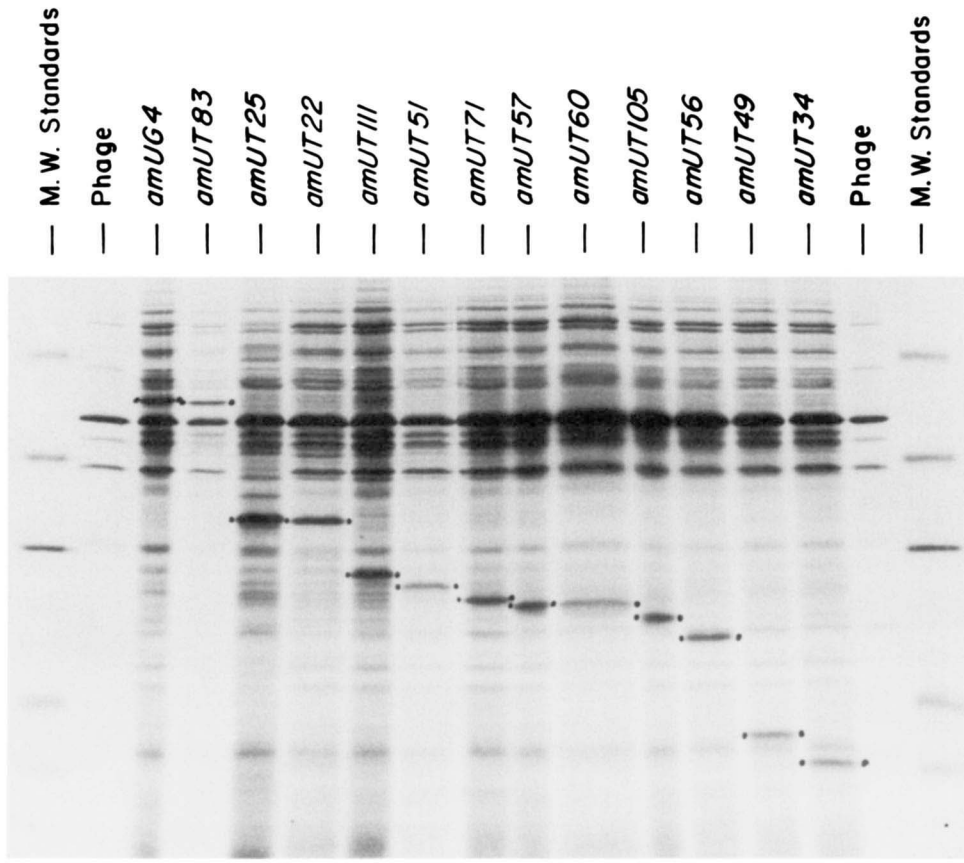

FIGURE 4.—Electrophoresis of ^{14}C-labeled lysates of su$^-$ cells infected with gene 9 amber mutants. *Salmonella* DB7136 (su$^-$) was infected with the amber mutants and incubated with a four minute pulse of ^{14}C-labeled amino acids 40 min after infection. After a 1-min chase, cells were rapidly frozen and then thawed in SDS buffer. Portions were electrophoresed through a 12.5% acrylamide gel. The amber fragments of these thirteen lysates had been identified in other gels and are arranged here according to their mobility. The molecular mass of the putative amber fragments were estimated by a comparison to a set of five ^{14}C-methylated markers, in the outside lanes. They are cytochrome *c* [12.3 kilodaltons (kD)], lactoglobulin A (18.4 kD), carbonic anhydrase (30 kD), ovalbumin (46 kD), and albumin (69 kD). We estimated the molecular mass of the putative amber fragments to be (from left to right) 54.2, 54.0, 34.5, 34.1, 27.6, 25.3, 25.0, 24.5, 24.1, 23.1, 21.3, 15.2 and 11.8 kD.

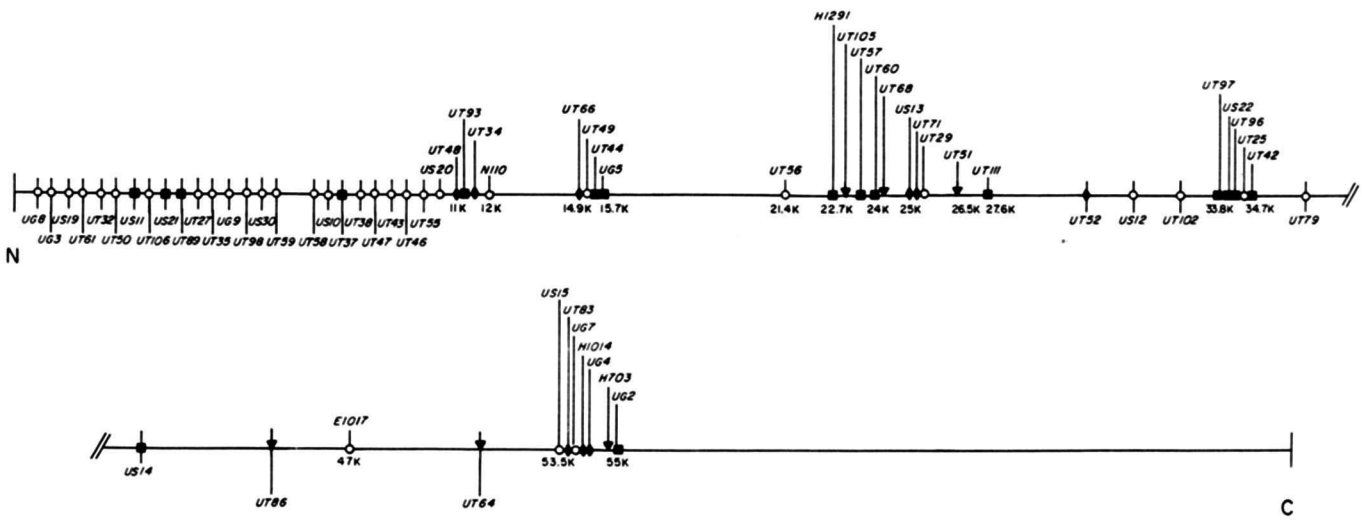

FIGURE 5.—Distribution of amber mutations and their missense phenotypes in gene 9. This map was constructed by combining the amber fragment molecular weights with the genetic mapping data. Mutants placed above the line produced resolvable amber fragments. In those cases where an amber fragment was not resolved, it is placed below the map in its genetic map order. None of the mutants mapping within the first deletion interval or the beginning of the next deletion interval produced resolvable fragments. *Open circles* depict tolerant ambers. *Solid squares* depict those mutants which have a tight *ts* phenotype on at least one su$^+$ host. *Solid diamonds* indicate those mutants which have lethal missense phenotypes on at least one su$^+$ host. *Solid triangles* identify those mutants that have both tight *ts* and lethal phenotypes.

Phenotypic defects: BERGET and POTEETE (1980) described an amber mutant *amH1014*, which yielded a mature but defective tailspike on DB7156. The missense tailspikes were able to bind to phage heads, but were defective in their endorhamnosidase function. The inactive particles produced in such infections could not be rescued by the addition of wild-type tailspikes. In contrast, the heads produced when cells are infected with gene *9* mutants conferring a folding or subunit association defect, were efficiently rescued by adding exogenous tails (SMITH and KING 1981).

The missense polypeptide chains generated by amber mutants growing on the suppressor hosts could be defective in a stage of tailspike maturation, head binding or, like *amH1014*, they could mature, but be defective in functions of the native protein. In the latter case, the mature tailspikes would bind to the phage head forming inactive particles. To distinguish between these alternatives, we determined whether the heads produced in such infections could be rescued by the addition of exogenous wild-type tailspikes.

Exponentially cultured su^+ cells were infected with amber mutants of interest. At various times after infection, samples were taken and the infected cells lysed with chloroform. One portion of each sample was plated directly for viable phage, and another was incubated with tailspikes to measure the presence of mature heads lacking tailspikes. To avoid variations in absorption due to the missense proteins in the infecting particles, we used hybrid phage assembled *in vitro* for these experiments (MATERIALS AND METHODS).

Figure 6 contrasts the behavior of three phenotypically different amber mutants. The curves illustrate the production of mature heads and viable phage particles as a function of time after infection.

The upper left panel shows the accumulation of heads without tails in cells infected with *amUT51*. This is typical of a defect in the maturation of the tailspike protein (SMITH and KING 1981). As a control infection, we used *amE1017*, a tolerant amber. Free heads are somewhat in excess over phage, reflecting the reduction of tailspike production due to the limited efficiency of the tRNA suppression.

The upper right panel shows the growth curve for *amH1014* which yields mature but functionally inactive tailspikes when tyrosine is at the amber site (BERGET and POTEETE 1980). Most of the heads produced in these cells accumulate with defective tails, and are neither viable, nor rescuable by the addition of exogenous tails. Rescue of the heads is inefficient because they already have the *amH1014* protein bound.

The hyper-ts missense gene *9* polypeptides produced in DB7154 cells infected with *amUT51*, on the other hand, are defective in some step in tailspike maturation. The final yield of mature heads is two orders of magnitude greater than that of background phage. *amUT51* behaves in the same manner when grown on DB7155 (data not shown) which results in a "lethal" substitution.

We conducted these experiments with 10 of the 15 amber mutants in su^+ hosts in which lethal and hyper-ts polypeptides were synthesized. With the exception of *amH1014*, all of the mutants displayed the defective folding or subunit assembly phenotype.

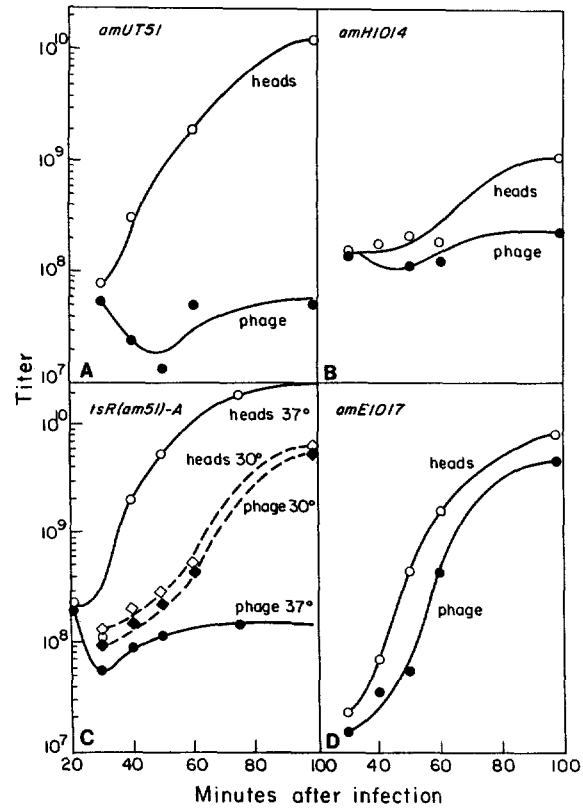

FIGURE 6.—Physiological defects of representative missense mutant proteins. The curves illustrate the production of mature heads and viable phage particles as a function of time after infection of restrictive su^+ hosts. Viable phage were assayed by titering under permissive conditions; mature heads were assayed by titering in the presense of exogenous native tailspikes. (D) The curves for *amE1017*, a tolerant amber, illustrate that mature heads are somewhat in excess over phage, reflecting the decreased efficiency of tailspike production due to the limited efficiency of the tRNA suppression. (A) The curves for *amUT51* (in DB7154) demonstrates the accumulation of heads without tails which is typical of a defect in the maturation of the tailspike protein. (B) The growth of *amH1014* (in DB7156) illustrates the behavior of a mutant which yields mature but functionally inactive tailspikes (BERGET and POTEETE 1980). Most of the heads produced in these cells have mature but functionally defective suppressed missense tailspikes bound to them and cannot be rescued by the addition of native tailspikes. The lower left panel also displays the growth curves for an am^+ts revertant of *amUT51*. At elevated temperatures it behaves like *amUT51* when grown on a nonpermissive su^+ host. Since the tailspike protein is involved in phage absorption, these experiments were conducted with hybrid particles. These particles were assembled *in vitro* from wild-type tailspikes and mature heads carrying the mutant genome. This eliminates variation in the initial infection of the host cells.

Figure 6 also contains the curve of an *am*$^+$*ts* revertant of *amUT51*. At nonpermissive temperatures, it behaves like the original amber mutant when serine is inserted at the amber site.

am$^+$ reversion analysis: For those mutants that yielded a strong *ts* and/or absolutely defective missense protein on one of the su$^+$ hosts, it seemed likely that we would be able to isolate revertants at the amber site to an amino acid generating a *ts* phenotype. We examined the *am*$^+$ revertants of the amber mutants which produce lethal missense proteins by picking revertants to *am*$^+$ at 30° and screening these for *ts* phenotypes at 37° and 39°. We recovered pseudorevertants for 6 of the 15 sites studied: *amUT51*, *amUT64*, *amUT68*, *amUT86*, *amUT105* and *amH703*. These six mutants displayed a tight *ts* phenotype on at least one su$^+$ host in addition to their lethal phenotype.

For a number of these pseudorevertants, we have determined if the defect was in maturation or function. Figure 6C shows the results of experiments with a *ts* revertant of *amUT51*, *tsR(am51)-A*. This pseudorevertant was backcrossed with its *amUT51* parent, and yielded no recombinants above the *am*$^+$ or *ts*$^+$ background indicating that both mutations are at the same site. At 30° the infection was permissive, producing a healthy burst of viable phage, and a few nontailed heads. At the restrictive 37° temperature, however, the phage yield was sharply reduced, and the cells accumulated heads which could be rescued by adding exogenous native tailspikes. Thus both the missense serine containing protein and the *ts* pseudorevertant at the same site are defective in the folding or maturation of the tailspike polypeptide chain at restrictive temperature. The preliminary biochemical characterization of the missense polypeptide produced by an *am*$^+$ *hyper-ts* revertant at this site, *htsR(am51)-E*, indicates that the SDS-resistant native trimer is produced at the permissive temperature (20°) but not at the restrictive temperature (30°) suggesting that mutations at this site confer a defect in the maturation of the tailspike protein and not the attachment of tailspikes to mature phage heads (data not shown). Similar results were also obtained with *am*$^+$*ts* revertants of *amUS11*, *amUT57* and *amUT86*.

DISCUSSION

Of the 111 newly isolated phage strains carrying amber mutations in gene *9*, 94 had the characteristics of single amber mutations. These mutants, along with six previously isolated amber mutants, defined at least 64 distinct sites within gene *9*. This is probably a minimum estimate, since close sites might not have been separated by the two-factor crosses. Most of these amber sites are probably derived by a single base change of the 74 parent codons listed in Table 2. Since the mutants were selected using hosts which insert tyrosine, serine, or glutamine at the amber codon, these sites should be represented.

The mapping and suppression patterns do not provide sufficient information to identify the actual mutant sites. But since the codons which could mutate to *UAG* in one step include 27 tyrosine codons and seven tryptophan codons, a substantial number of these must be represented in the amber set. Seven of the mutants, *amUT49*, *amUT61*, *amUT46*, *amUT111*, *amUT93*, *amUT64* and *amUT51*, are only suppressed by tyrosine at high temperature (Table 5) and are probably among these.

In the case of the tryptophan codons, we know that two mutants, *amH1291* and *amH1200* (allelic to *amUT57*) are at tryptophan sites by direct sequence analysis of the mutant DNA (YU and KING 1984, M.-H. YU and J. KING, personal communication). These mutants are well suppressed only by tyrosine insertion.

A few of the mutations may have arisen from multiple base changes. In the extensive study conducted with nonsense mutations in the *lac* repressor of *Escherichia coli*, MILLER and coworkers reported that nonsense mutations arising from tandem base changes with UV mutagenesis were 1.0% of the total nonsense mutations recovered (MILLER 1985). Codons that can mutate to stop codons by such tandem base changes include those for phenylalanine, methionine, valine, proline, threonine, alanine, cysteine, arginine and glycine. Precise identification of the amber sites awaits sequencing of the gene *9* region from mutant DNA.

Phenotypes of the amber missense proteins: The set of amino acids inserted by the four host suppressors have very different properties; serine is hydrophilic and a hydrogen bond donor, glutamine is very hydrophilic and can act as both an H-bond donor and receptor, leucine is hydrophobic with no reactive groups, and tyrosine is aromatic and hydrophobic, but in addition its hydroxyl can donate an H bond. The gene *9* amber mutants displayed a variety of phenotypes when grown on serine, glutamine, tyrosine and leucine inserting hosts at different temperatures. These ranged from fully permissive under all conditions to temperature sensitive on one or more of the tRNA suppressor hosts to absolute lethal phenotypes on most of the tRNA suppressor hosts.

If all four of these residues are tolerated at the site of the amber codon, we term the site tolerant. Thirty three of the amber sites had this character. Such residues could be spacer residues, whose side chains contribute little to the folding or function of the protein, but are still needed to achieve proper conformation. They may also be preferentially located at the protein surface, where variability is better tolerated (DICKERSON and GEIS 1983).

Over 30 of the amber sites had a temperature

sensitive phenotype when grown on at least one of the tRNA suppressor strains. This was not suprising given the numerous *ts* sites in gene *9* (SMITH, BERGET and KING 1980; KING *et al.* 1986). The *tsf* mutants identify sites in the chain needed to stabilize intermediates in the pathway, or otherwise ensure the correct chain folding and association pathway at the high end of the temperature range for phage growth. All of the *ts* sites in gene *9* described in SMITH, BERGET and KING (1980) have been found to be temperature sensitive folding mutations (SMITH and KING 1981; M.-H. YU, C. HAASE-PETTINGELL and J. KING, personal communication). We presume that the *ts* phenotypes of the amber mutants represent the same or related sites.

The recovery of *ts* pseudorevertants at a number of the am^{ts} sites supports this notion. With *ts* revertants at five of the amber sites, *amH1291*, *amUS11*, *amUT51*, *amUT57* and *amUT86*, the formation of native tailspikes at restrictive temperature has been examined directly. In all cases the mutant tailspike polypeptide chains failed to mature into the SDS-resistant native trimer. Defects in other steps of maturation such as the protrimer to native transition, or binding of the tailspike to the capsid, could be represented at the other sites.

A third set of amber mutants manifested lethal phenotypes when grown on at least one su^+ host. The mutants were unable to plaque regardless of the temperature of growth. In two-thirds of these cases the defects were sufficiently stringent that the efficiencies of plating were reduced to or below the level of revertants in the stock. These amino acid substitutions correspond to absolute lethal mutations.

The analysis of the missense polypeptide chains synthesized in su^+ cells infected with lethal ambers showed that most of them, except for *amH1014*, were defective in the folding or subunit assembly of the mutant chain. These mutations probably mark sites in the polypeptide chain where the wild-type amino acid interactions are maintaining the correct folding and association pathway throughout the entire temperature range of phage growth. Alternatively the lethal substitutions could introduce or accelerate an off-pathway interaction, such as aggregation (GOLDENBERG 1981; YU and KING 1984).

An additional possibility is that some of the mutations specifically prevent the binding of the tailspike to the phage head. P. BERGET (personal communication) has isolated absolute lethal mutations at over 40 sites in gene *9* by selecting for particles lacking tailspikes. This selection was an effort to find mutant tailspikes that could mature but had lost the ability to bind to phage heads. Analysis of the proteins synthesized in mutant infected cells indicated that the tailspike polypeptide chain was synthesized but did not form a mature tailspike suggesting folding or chain association defects (P. BERGET, personal communication). These results, together with the absence of this phenotype among the *tsf* mutants (GOLDENBERG, SMITH and KING 1983), suggest that mutations permitting maturation but specifically blocking head binding are rare.

We were somewhat surprised by the rarity of sites such as *amH1014*, where the mutant protein is mature and capable of binding to heads, but is defective in the endorhamnosidase activity. Particles with the mature *ts* missense tailspikes showed no gross defects in their ability to absorb. The codons that can mutate to amber in one step, however, are only 10% of the total residues, and there are long stretches of the gene lacking these codons. Regions critical for catalytic activity may be located in such regions of the chain. Alternatively, the identity of the residues at the active site may be crucial. If the suppressed missense proteins are not functional, amber mutants at such sites would not be recovered.

Character of the *ts* and *lethal* sites: The *tsf* mutations mark amino acid sites or local sequences needed to stabilize intermediates in protein folding or subunit assembly at elevated temperatures. These regions of the polypeptide chain could be involved in interactions necessary to maintain the productive maturation pathway only at the high end of the temperature range of growth. In this model, these sites would not be critical at lower temperatures, and would therefore not be loci of lethal mutations. For example a number of the *tsf* mutations are substitutions of threonines by isoleucines (YU and KING 1984). The extra hydrogen bond made by the threonine side chain in the wild type might stabilize some intermediate at the restrictive temperature, but be dispensable at permissive temperature.

An alternative model is that *tsf* sites or sequences are involved in interactions crucial for folding at all temperatures, but that the *tsf* amino acid substitutions only disrupt the intermediate at elevated temperature. In this case lethal mutations would appear at the same site.

The amber missense phenotypes suggest both classes of sites exist. *AmUT86* is an example of a crucial site. It is lethal with glutamine inserted, but *ts* with serine inserted. Furthermore, it reverts to an am^+ts phenotype. While a serine insertion can still direct the proper maturation at lower temperatures, it disrupts, or fails to direct, the process at elevated temperatures. Glutamine insertion, however, creates a temperature independent block which identifies this site as crucial. This suggests that a subset of the *tsf* sites are probably also sites for lethal substitutions. *AmUT57*, on the other hand, is temperature sensitive on all of the suppressors, but grows at 30° on all su^+ hosts. This site appears to be important primarily at high temper-

ature. Given the wide temperature range over which phage and bacterial polypeptide chains have to fold, it is not surprising that there may be additional polypeptide chain interactions required at the high end of the temperature range.

Clustering: A distinctive feature of the results is the clustering of phenotypes. As noted above, the amber sites at the N-terminus of the polypeptide chain are tolerant with respect to the amino acid inserted. No *tsf* mutants have been found in the region from aa 1 to 150 (R. VILLAFANE and J. KING, personal communication). This suggests that this region of the polypeptide chain is not critical for the folding or function of the tailspike polypeptide chain.

Newly synthesized polypeptide chains are often assumed to fold sequentially from their N-termini as the chains are coming off the ribosome. There is little experimental evidence, however, on the sequence of *in vivo* maturation events in prokaryotes. Immunological studies with tailspike amber fragments up to two-thirds of the gene *9* length indicate that their antigenicity is much more like that of guanidine denatured protein than that of the native tailspike (SHEA 1977). These observations suggest that the N-terminus achieves its native conformation late in the folding pathway directed by the tertiary structure of the rest of the protein.

The amber sites in the central region of the chain are much more sensitive to the amino acid inserted; 16 of 22 amber mutants manifesting tight *ts* phenotypes map to this region. This is also where most of the *tsf* mutations are located (YU and KING 1984; M.-H. YU, R. VILLAFANE and J. KING, personal communication). This suggests that this region of the polypeptide chain directs the conformation of the thermolabile intermediate in the wild-type maturation pathway (GOLDENBERG, BERGET and KING 1982).

Ts mutations are distributed irregularly among different genes (EDGAR, DENHARDT and EPSTEIN 1964; EDGAR and LIELAUSIS 1964). This probably reflects differences in both the structure and the folding pathway of the polypeptides for which different genes code, rather than variation at the nucleotide level.

We recovered no amber mutations in the most C-terminal portion of the chain. There are 12 codons from amino acid 435–666 which can mutate to amber via one base change. Eight of these are lysines, leucines, and tryptophans, which might not have been found in our screen. Four of the codons are tyrosines, but these require T:A to G:C transversions, which are not as frequent among our sequenced UV induced mutations. The C-terminus of the chain also lacks *tsf* sites. Either this region of the chain is very critical, tolerating few substitutions—our most carboxy ambers have primarily lethal phenotypes—or else this region is dispensable and amber mutations propagate on the su^- host. If the C-terminus were necessary for the nucleation of the productive pathway, that would explain the lack of antigenicity of amber fragments.

CRAWFORD and coworkers have observed similar clustering in the *trpB* gene of *E. coli* which codes for the B protein of tryptophan synthetase (CRAWFORD and JOHNSON 1963, 1964; CRAWFORD et al. 1970). The mutations categorized as "unclassified CRM-less mutations" are candidates for folding defects. These mutations were not suppressed by any nonsense suppressors—amber, ochre or opal—nor did they appear to arise from frameshifts. They were totally defective in function and did not cross-react with antibodies recognizing the mature *trpB* protein.

Second-site suppressors: The amber mutations with *ts* or lethal phenotypes on an su^+ host are excellent material for the isolation of second site suppressors of the missense defects. The presence of the amber makes it easy to distinguish between revertant events at the site of the original mutation and those which occur at a second site. Mutants that carry a second site suppressor retain the amber phenotype. Thus far we have been able to isolate and map second site suppressors for ten mutational sites within gene *9*. The suppressors of two of these sites, *amUT34* and *amUT71*, are extragenic, mapping both to the left of gene *9*, in genes encoding proteins essential in the assembly of the P22 phage head, and to the right of gene *9*, a poorly understood region of the P22 genome. The second site suppressors of eight sites appear to be intragenic. The identification of second site suppressors may indicate those regions or residues of the polypeptide that interact during the chain folding and association pathway.

We thank PETER BERGET, RUTH GRIFFIN SHEA, GEORGE J. THOMAS, JR., ROBERT VILLAFANE, CAMERON HAASE-PETTINGELL and MYEONG-HEE YU for communicating unpublished results, DAVID BOTSTEIN and GRAHAM WALKER for kindly provided bacterial strains. This research was supported by National Science Foundation grant PCM 8402546 (J. K.). B.F. was supported by a National Institutes of Health Institutional Training grant.

LITERATURE CITED

ANFINSEN, C. B. and H. A. SCHERAGA, 1975 Experimental and theoretical aspects of protein folding. Adv. Protein Chem. **29:** 205–300.

BERGET, P. B. and A. R. POTEETE, 1980 Structure and functions of the bacteriophage P22 tail protein. J. Virol. **34:** 234–243.

BOTSTEIN, D., R. K. CHAN and C. H. WADDELL, 1972 Genetics of bacteriophage P22. II. Gene order and gene function. Virology **49:** 268–282.

BOTSTEIN, D., C. H. WADDELL and J. KING, 1973 Assembly and DNA encapsulation in *Salmonella* phage P22. I. Genes, proteins, structures and DNA maturation. J. Mol. Biol. **80:** 669–695.

CHAN, R. K. and D. BOTSTEIN, 1972 Genetics of bacteriophage P22. I. Isolation of prophage deletions which affect immunity to superinfection. Virology **49:** 257–267.

COULONDRE, C. and J. H. MILLER, 1977 Genetic studies of the *lac* repressor: III. Additional correlation of mutational sites with specific amino acid residues. J. Mol. Biol. **117:** 525–567.

CRAWFORD, I. P. and L. M. JOHNSON, 1963 Mutants of *Escherichia coli* defective in the B protein of tryptophan synthetase. I. Selection and classification. Genetics **48:** 725–736.

CRAWFORD, I. P. and L. M. JOHNSON, 1964 Mutants of *Escherichia coli* defective in the B protein of tryptophan synthetase. II. Intragenic position. Genetics **49:** 267–278.

CRAWFORD, I. P., S. SIKES, O. BELSER and L. MARTINEZ, 1970 Mutants of *Escherichia coli* defective in the B protein of tryptophan synthetase. III. Intragenic clustering. Genetics **67:** 201–211.

DAVIS, R. W., D. BOTSTEIN and J. R. ROTH, 1980 *Advanced Bacterial Genetics*. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.

DICKERSON, R. E. and I. GEIS, 1983 *Hemoglobin: Structure, Function, Evolution, and Pathology*. Benjamin Cummings, Menlo Park, Calif.

EDGAR, R. S. and I. LIELAUSIS, 1964 Temperature sensitive mutants of bacteriophage T4D: their isolation and genetic characterization. Genetics **49:** 649–662.

EDGAR, R. S., G. H. DENHARDT and R. H. EPSTEIN, 1964 A comparative genetic study of conditional lethal mutations of bacteriophage T4D. Genetics **49:** 635–648.

GOLDENBERG, D. P., 1981 Genetic and biochemical analysis of the folding and subunit assembly of the bacteriophage P22 tail spike protein. Thesis, MIT, Cambridge, Mass.

GOLDENBERG, D. and J. KING, 1981 Temperature-sensitive mutants blocked in the folding or subunit assembly of the bacteriophage P22 tailspike protein. II. Active mutant proteins matured at 30°C. J. Mol. Biol. **145:** 633–651.

GOLDENBERG, D. and J. KING, 1982 Trimeric intermediate in the *in vivo* folding and subunit assembly of the tailspike endorhamnosidase of bacteriophage P22. Proc. Natl. Acad. Sci. USA **79:** 3403–3407.

GOLDENBERG, D., P. BERGET and J. KING, 1982 Maturation of the tailspike endorhamnosidase of *Salmonella* phage P22. J. Biol. Chem. **257:** 7864–7871.

GOLDENBERG, D. P., D. H. SMITH and J. KING, 1983 Genetic analysis of the folding pathway for the tailspike protein of phage P22. Proc. Natl. Acad. Sci. USA **80:** 7060–7064.

ISRAEL, V. J., F. ANDERSON and M. LEVINE, 1967 *In vitro* morphogenesis of P22 from heads and baseplate parts. Proc. Natl. Acad. Sci. USA **57:** 284–291.

IWASHITA, S. and S. KANEGASAKI, 1976 Enzymic and molecular properties of baseplate parts of bacteriophage P22. Eur. J. Biochem. **65:** 87–94.

JARVIK, J. and D. BOTSTEIN, 1975 Conditional-lethal mutations that suppress genetic defects in morphogenesis by altering structural proteins. Proc. Natl. Acad. Sci. USA **72:** 2738–2742.

KIM, P. S. and R. L. BALDWIN, 1982 Specific intermediates in the folding reactions of small proteins and the mechanism of protein folding pathways. Annu. Rev. Biochem. **51:** 459–489.

KING, J., 1986 Genetic analysis of protein folding pathways. Bio/Technology **4:** 297–303.

KING, J. and M.-H. YU, 1986 Mutational Analysis of Protein Folding Pathways: The P22 Tailspike Endorhamnosidase. Methods Enzymol. **131:** 250–266.

KING, J., C. A. HAASE and M.-H. YU, 1987 Temperature sensitive mutations affecting kinetic steps in protein folding pathways. pp. 109–121. In: *Protein Engineering*, Edited by D. OXENDER and C. F. FOX. Alan R. Liss, New York.

KING, J., M.-H. YU, J. SIDDIQI and C. HAASE, 1986 Genetic identification of amino acid sequences influencing protein folding. pp. 275–291. In: *Protein Engineering. Applications in Science, Medicine and Industry*, Edited by M. INOUE AND S. SARMA. Academic Press, New York.

KREBS, H., F. X. SCHMID and R. JAENICKE, 1983 Folding of homologous proteins. The refolding of different ribonucleases is independent of sequence variations, proline content and glycosylation. J. Mol. Biol. **169:** 619–635.

MILLER, J. H., 1985 Specificity of ultraviolet light. J. Mol. Biol. **182:** 45–64.

MILLER, J. H., C. COULONDRE and P. FARABAUGH, 1978 Correlation of nonsense sites in the *lacI* gene with specific codons in the nucleotide sequence. Nature **274:** 770–775.

MILLER, J. H., D. GANEM, P. LU and A. SCHMITZ, 1977 Genetic studies of the *lac* repressor. I. Correlation of mutational sites with specific amino acid residues: Construction of a colinear gene-protein map. J. Mol. Biol. **109:** 275–298.

MILLER, J. H., C. COULONDRE, M. HOFER, U. SCHMEISSNER, H. SOMMER, A. SCHMITZ and P. LU, 1979 Genetic studies of the *lac* repressor. IX. Generation of altered proteins by the suppression of nonsense mutations. J. Mol. Biol. **131:** 191–222.

POTEETE, A. R. and J. KING, 1977 Functions of two new genes in *Salmonella* phage P22 assembly. Virology **76:** 725–739.

SADLER, J. R. and A. NOVICK, 1965 The properties of repressor and the kinetics of its action. J. Mol. Biol. **12:** 305–327.

SAUER, R. T., W. KROVATIN, A. R. POTEETE and P. B. BERGET, 1982 Phage P22 tail protein: gene and amino acid sequence. Biochemistry **21:** 5811–5815.

SCHMITZ, A., C. COULONDRE and J. H. MILLER, 1978 Genetic studies of the *lac* repressor. V. Repressors that bind operator more tightly generated by suppression and reversion of nonsense mutations. J. Mol. Biol. **123:** 431–454.

SHEA, R. G., 1977 Genetic identification of phage P22 antigens and their structural location. Thesis, MIT, Cambridge Mass.

SMITH, D. H. and J. KING, 1981 Temperature-sensitive mutants blocked in the folding or subunit assembly of the bacteriophage P22 tail spike protein. III. Inactive polypeptide chains synthesized at 39°C. J. Mol. Biol. **145:** 653–676.

SMITH, D. H., P. B. BERGET and J. KING, 1980 Temperature-sensitive mutants blocked in the folding or subunit assembly of the bacteriophage P22 tail spike protein. I. Fine-structure mapping. Genetics **96:** 331–352.

SOMMER, H., A. SCHMITZ, U. SCHMEISSNER, J. H. MILLER and H.-G. WITTMANN, 1978 Genetic studies of the *lac* repressor. VI. The B116 repressor: an altered *lac* repressor containing amino acids specified by both the *trp* and *lacI* leader regions. J. Mol. Biol. **123:** 457–470.

WALKER, G. C., 1978 Inducible reactivation and mutagenesis of UV-irradiated bacteriophage P22 in *Salmonella typhimurium* LT2 containing the plasmid pKM101. J. Bacteriol. **135:** 415–421.

WINSTON, F., D. BOTSTEIN and J. H. MILLER, 1979 Characterization of amber and ochre suppressors in *Salmonella typhimurium*. J. Bacteriol. **137:** 433–439.

YU, M.-H. and J. KING, 1984 Single amino acid substitutions influencing the folding pathway of phage P22 tail spike endorhamnosidase. Proc. Natl. Acad. Sci. USA **81:** 6584–6588.

Communicating editor: J. ROTH