JOURNAL OF BACTERIOLOGY, Sept. 1986, p. 1035–1042
0021-9193/86/091035-08$02.00/0

Vol. 167, No. 3

Mutational Analysis of Bacteriophage Lambda Lysis Gene *S*

RONALD RAAB,[1] GREGORY NEAL,[1] JANETTA GARRETT,[2] RAY GRIMAILA,[3] ROBERT FUSSELMAN,[4] AND RY YOUNG[1]*

Department of Medical Biochemistry, College of Medicine, Texas A&M University, College Station, Texas 77843[1]; Department of Biology, Hamilton College, Clinton, New York 13323[2]; Department of Cellular and Developmental Biology, Harvard University, Cambridge, Massachusetts 02138[3]; and Department of Internal Medicine, Scott and White Hospital, Temple, Texas 76508[4]

Received 10 March 1986/Accepted 4 June 1986

A plasmid carrying the bacteriophage lambda lysis genes under *lac* control was subjected to hydroxylamine mutagenesis, and mutations eliminating the host lethality of the *S* gene were selected. DNA sequence analysis revealed 48 single-base mutations which resulted in alterations within the coding sequence of the *S* gene. Thirty-three different missense alleles were generated. Most of the missense changes clustered in the first two-thirds of the molecule from the N terminus. A simple model for the disposition of the *S* protein within the inner membrane can be derived from inspection of the primary sequence. In the first 60 residues, there are two distinct stretches of predominantly hydrophobic amino acids, each region having a net neutral charge and extending for at least 20 residues. These regions resemble canonical membrane-spanning domains. In the model, the two domains span the bilayer as a pair of net neutral charge helices, and the N-terminal 10 to 12 residues extend into the periplasm. The mutational pattern is largely consistent with the model. Charge changes within the putative imbedded regions render the protein nonfunctional. Loss of glycine residues at crucial reverse-turn domains which would be required to reorient the molecule to reenter the membrane also inactivate the molecule. Finally, a number of neutral and rather subtle mutations such as Ala to Val and Met to Ile are found, mostly within the putative spanning regions. Although no obvious explanation exists for this subtle and heterogeneous class of mutations, it is noted that all of the changes result in a loss of alpha-helical character as predicted by Chou-Fasman theoretical analysis. Alternative explanations for some of these changes are also possible, including a reduction in net translation rate due to substitution of a rare codon for a common one. The model and the pattern of mutations have implications for the probable oligomerization of the *S* protein at the time of endolysin release at the end of the vegetative growth period.

The bacteriophage lambda has three genes which are directly involved in lysis: *S*, *R*, and *Rz* (16, 30, 37). The *R* gene product is a transglycosylase and is responsible for the destruction of the mechanical rigidity of the peptidoglycan (5). The *Rz* gene product is required for lysis only under conditions which are known to stabilize the outer membrane and may be involved in degradation of the covalent bonds between the outer membrane and the peptidoglycan (37). The *S* gene encodes a 107-codon polypeptide which is responsible for a lethal event at the cytoplasmic membrane, allowing the release of the *R* gene product to the periplasm for the degradation of the peptidoglycan (8, 14, 15). In our hands, only a single polypeptide, with an apparent molecular weight of about 10,000, is elaborated from the wild-type *S* gene (and not from an *S*am mutant allele) in infected cells and in maxicells (3) and is associated with the inner membrane (4). The nature of the event at the cytoplasmic membrane is unclear, but a pleiotropic effect on membrane permeability and capacity to support active transport has been unambiguously demonstrated (2, 15, 35). Several workers have inferred from the release of the *R* gene product and the other associated effects that the *S* protein forms a pore in the cytoplasmic membrane (15, 31, 35). The formation of the pore must be regulated in some way, because the action of the *S* gene in the onset of lysis occurs long after the beginning of the *S* gene expression (6, 17, 32). This delayed action can be triggered early by energy poisons and protein synthesis inhibitors (6, 14, 31). A lysis gene system with many of the same properties as that of lambda exists in T4, where the *t* and *e* genes appear to fill the same roles as *S* and *R*, respectively (20, 28, 33).

Although the function of the *S* gene product is of interest because of its apparent pore-forming ability, unusual functional regulation, and, possibly, generalized ability to insert in a bilayer, a direct molecular investigation of how the polypeptide functions is impractical. Because the polypeptide is very small, hydrophobic, and made in low quantities, it was one of the last essential gene products of bacteriophage lambda to be identified (3) and is not likely to be purified with existing technology. However, a genetic approach has a number of advantages. Because expression of the *S* gene alone is necessary and sufficient for the lethality associated with lysis, although not sufficient for lysis itself (15), it should be possible to isolate mutations in the *S* gene which render the product nonlethal. Moreover, it is likely that the *S* gene product is not an enzyme and thus probably has no active site formed from residues from dispersed regions of the primary sequence. Instead, one would predict that if much of the relatively short (107-codon) sequence is involved in membrane-imbedded structures, a large collection of missense mutations which affect *S* function might fall into clusters which define imbedded domains and define structures required for membrane insertion. The *S* cistron may thus be a useful genetic probe for the study of the structural requirements for membrane-protein interactions.

MATERIALS AND METHODS

Bacterial strains and plasmids. The host bacterial strain for all of the mutational selections, CQ21, was derived from

* Corresponding author.

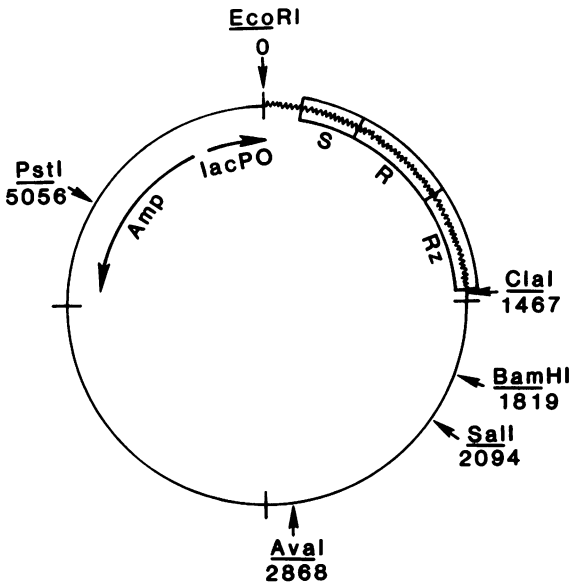

FIG. 1. Inducible lysis plasmid pRG1. The total size is 6,018 base pairs. Directions of transcription of the *bla* gene conferring Amp^r and from the *lacPO* control region are indicated. The hatched line represents λ DNA. Boxes indicate the approximate extent of the three λ lysis genes.

CSH57B of the Cold Spring Harbor collection (26) and has the genotype *ara leu lacI^{q}1 purE gal his argG rpsL xyl mtl ilv*. The plasmid pRG1 was constructed by inserting the bacteriophage lambda restriction fragment bounded by the *Eco*RI site at nucleotide 44972 and the *Cla*I site at nucleotide 46439 into the corresponding unique sites on the plasmid pBH20, which is essentially pBR322 with *lacPO* (Fig. 1) (19). The plasmid pRG1 is functionally equivalent to the previously described inducible lysis plasmid pRF26 (14), except that there is very little λ DNA present distal to the *Rz* gene in pRG1.

Media and growth conditions. Standard LB and M9 minimal liquid media and plates were used (27) with the appropriate auxotrophic supplements for growth of CQ21. Isopropyl-β-D-thiogalactopyranoside was obtained from Sigma Chemical Co., St. Louis, Mo., and was used at 1 mM. 5-Bromo-4-chloro-3-indolyl-β-D-galactopyranoside (Sigma) was used in plates at 40 μg/ml. Ampicillin was used at 50 μg/ml in LB and in other media when indicated.

Buffers and reagents. Sterile saline solution was 8 g of NaCl per liter. DNA buffer was 10 mM Tris hydrochloride–1 mM disodium EDTA (pH 7.4). HAM buffer was 0.5 M K_3PO_4–5 mM disodium EDTA, adjusted to pH 6.0 with 4 N NaOH. Hydroxylamine solution was made fresh before use by combining 0.56 ml of 4 N NaOH and 0.35 g of hydroxylamine hydrochloride (MCB Reagents) and adjusting the final volume to 5.0 ml with sterile distilled water. Reagents for dideoxy sequencing were purchased from P&L Biochemicals, Piscataway, N.J., and used according to the directions of the manufacturer. Restriction enzymes were purchased from International Biotechnologies New Haven, Conn., and Bethesda Research Laboratories, Gaithersburg, Md., and were used according to the instructions supplied.

Mutagenesis of pRG1. Mutagenesis was done by modification of the procedure of Davis et al. (10). A 0.2-ml sample of a 1-mg/ml solution of plasmid DNA purified by isopycnic density gradient centrifugation was combined with 0.4 ml of HAM buffer, 0.8 ml of hydroxylamine solution, and 0.6 ml of double-distilled sterile water and incubated for 1 h at 37°C. The DNA was then dialyzed exhaustively (1,000 volumes per change, eight or nine changes) versus DNA buffer. The DNA was then precipitated in ethanol and suspended in 150 μl of DNA buffer.

Selection of *S* mutants. The mutagenized plasmid DNA was transformed into CQ21 by the $CaCl_2$ method (23). Transformed cells were selected on LB-ampicillin plates, suspended as a slurry in 1 ml of sterile saline per plate, and plated at a 10^{-3} dilution on glycerol-minimal medium containing ampicillin, isopropyl-β-D-thiogalactopyranoside, and 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside. Previous work has shown that wild-type cells carrying the *S* gene under *lacPO* control are rapidly killed under these inducing conditions, with mutant survivors arising at a frequency of approximately 10^{-6} (15; G. S. Neal, M.S. thesis, Texas A&M University, College Station, 1984). Colonies arising after 36 h at 37°C were purified on LB-ampicillin medium and tested for the nature of the mutations conferring resistance to the inducing conditions. White colonies were assumed to have mutations affecting generalized expression from *lac* promoters (i.e., *cya*, *lacI^{sq}*, etc.) and were discarded. Plasmid minipreparations, made by the method of Holmes and Quigley (18), were made from small cultures grown from each candidate colony. This DNA was transformed into CQ21 and tested for survival on glycerol-minimal medium with ampicillin and isopropyl-β-D-thiogalactopyranoside. Plasmids revealed by this test to have lost the inducible lethality were digested with the restriction enzymes *Eco*RI and *Bam*HI and subjected to agarose gel electrophoresis. Candidate mutants which showed a restriction pattern indistinguishable from that of the parental pRG1 plasmid were checked for the inducibility of endolysin activity by a freeze-thaw procedure. Overnight cultures grown in minimal glycerol medium plus ampicillin were diluted sixfold into fresh medium containing isopropyl-β-D-thiogalactopyranoside and incubated with rolling for 2 h at 37°C. Each culture was divided into two portions, and one portion was frozen at −80°C for 30 min and thawed at 37°C for 15 min. The A_{550} was measured for the control and freeze-thawed samples.

Sequence analysis. The *Eco*RI-*Bam*HI fragment containing the lysis gene region (Fig. 1) from each candidate plasmid chosen for sequence analysis was subcloned into the dideoxy sequencing vector M13mp8 or M13mp9 (24). The

TABLE 1. Freeze-thaw lysis of *S* mutants[a]

Plasmid	Mutation	Base	Amino acid change	A_{550}		
				Before	After	Ratio
pRG1	Wild type			0.100	0.075	1.3
pBH20				0.477	0.444	1.1
pAQ1	T→A	166	Trp_{55}→END	1.100	0.200	5.2
pGN214	C→T	−313	(−35 region *lacPO*)	0.470	0.130	3.6
pGN205	G→A	52	Gly_{18}→Lys	0.498	0.076	6.6
pGN234	C→T	184	Leu_{62}→Arg	0.849	0.140	6.1
pRR13	G→A	9	Met_3→Ile	0.500	0.071	7.0
pRR30	C→T	144	Ala_{48}→Val	0.750	0.132	5.8
pRR94	G→A	67	Gly_{22}→Glu	0.397	0.070	5.7
pRR100	G→A	91	Ala_{30}→Val	0.610	0.110	5.6

[a] A_{550} readings on induced cultures carrying the wild-type and mutant *S* plasmids were taken before and after freeze-thaw (see Materials and Methods.) A ratio of greater than 5 is taken as evidence for full endolysin synthesis. The parental *lac* expression vector pBH20 is included as a nonlysing control. The low readings for pRG1 indicate that the culture was mostly lysed before freeze-thaw.

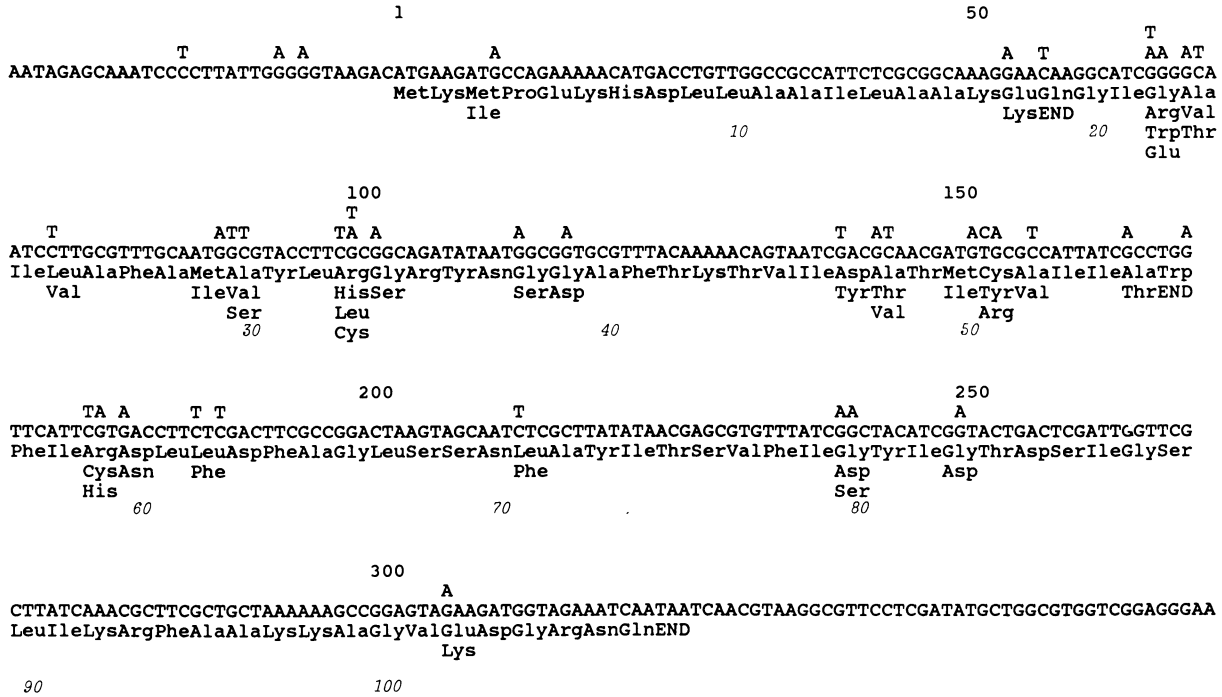

FIG. 2. Mutations leading to nonlethality of *S* in the plasmid pRG1. The nucleotide sequence is shown, numbered from the first base of the Met_1 codon, and the actual base changes are indicated above the sequence. The amino acid sequence is shown below the nucleotide sequence, and codon numbers are indicated in italics. Missense and nonsense mutations are listed beneath each codon. Not shown is a single-base change found in the *lac* promoter region of a pRG1 mutant. In this mutant the consensus −35 sequence . . .CTTTACACTTT. . . was changed to . . .TTTTACACTTT. . .

size of the *S* gene, 321 base pairs, is small enough to permit complete sequencing of a given mutant candidate on a single sequencing gel. However, three different oligonucleotide primers were used at various times for sequencing parts of the *S* gene with greater resolution. The oligonucleotide primers were synthesized on an Applied Biosystems model 380A automated DNA synthesizer. The sequences of the primers are as follows: 5′. . .ATAGAGCAAATCCCC. . .3′ (binds at positions −31 to −17), 5′. . .CC GTAGCCCCGTTAGG. . .3′ (binds at positions +58 to +73), and 5′. . .CCGCACCAGCCTCCC. . .3′ (binds at positions +352 to +366). Sequence analysis of the promoter region of some mutant plasmids was done by subcloning the *Pst*I-*Bam*HI fragment containing *lacPO* and the lysis gene region into the corresponding sites into M13mp9 and priming with a standard M13 universal (*lacZ*) primer obtained either obtained from Bethesda Research Laboratories or synthesized as described above. Chou-Fasman secondary structure predictions (7) and Kyte-Doolittle hydropathicity analysis (22) of the wild-type and mutant *S* sequences were performed with the Intelligenetics programs on the Bionet National Computer Resource for Molecular Biology (Intelligenetics, Palo Alto, Calif.).

RESULTS

Mutagenesis of S^+ plasmid. Induction of the S^+ R^+ Rz^+ plasmid pRG1 in the minimal-glycerol medium results in massive culture lysis at 1.2 to 1.5 h after the addition of isopropyl-β-D-thiogalactopyranoside (14; Neal, M.S. thesis). About 10% of the surviving cells carry mutant plasmids which, when transformed into a fresh background, are incapable of induced lethality (data not shown). These plasmids were found to have gross deletions of the *S* gene sequences easily detectable by restriction analysis and have not been studied further (J. Garrett, unpublished data). To generate a collection of point mutations in the *S* gene, the pRG1 plasmid was subjected to hydroxylamine mutagenesis under conditions determined to increase the total frequency of survivors of induction by 100-fold. The vast majority of plasmid mutants obtained by this mutagenesis were found to have no detectable deletions and were selected for further study as likely point mutants.

Characterization of hydroxylamine-induced mutations. Induction of the plasmid mutants resulted in a failure of the scheduled lysis event (data not shown). Since the *R* gene is cotranscribed with *S* and encodes a bacteriolytic activity which can be released by disruption of the membrane (31), it was possible to test the mutant plasmids for *R* gene function by freeze-thaw of induced cultures (Table 1). Approximately 20% of the mutant plasmids showed little or no lysis after freeze-thaw and were discarded as probable double mutants or gross promoter alterations (data not shown). Sequence analysis revealed that about one-third of the remaining clones were found to carry an unaltered *S* gene. Since it has already been shown that the only lethal gene among the lysis cistrons is *S*, it was presumed that these mutant plasmids carried point mutations which reduced the expression from the promoter. The promoter region of one such plasmid was sequenced and found to contain a C→T transition in the first base of the consensus −35 region of the *lac* promoter (Fig. 2). This mutant showed marginal lysis in the freeze-thaw test.

Distribution of *S* mutations. A total of 52 independently isolated plasmid mutants were found to have a single-base change in the *S* gene, with no defect on endolysin activity

TABLE 2. Missense mutations and change in empirical parameters[a]

Codon	Amino acids	Codon	No.	Codon frequency	Hydropathicity	P_a change	P_b change
3	Met Ile	ATG ATA	2	100, 1	1.9, 4.5	−0.37	0.55
18	Glu Lys	GAA AAA	2	73, 77	−3.5, −3.9	−0.35	0.37
22	Gly Arg	GGG AGG	1	33, 0.3	−0.4, −4.5	−0.41	0.18
22	Gly Trp	GGG TGG	1	33, 100	−0.4, −0.9	0.51	0.62
22	Gly Glu	GGG GAG	1	33, 27	−0.4, −3.5	0.94	−0.38
23	Ala Thr	GCA ACA	1	23, 6	1.8, −0.7	−0.59	0.36
23	Ala Val	GCA GTA	2	23, 23	1.8, 4.2	−0.36	0.87
25	Leu Val	CTT GTT	1	9, 38	3.8, 4.2	−0.15	0.40
29	Met Ile	ATG ATA	1	100, 1	1.9, 4.5	−0.37	0.55
30	Ala Val	GCG GTG	1	30, 27	1.8, 4.2	−0.36	0.87
30	Ala Ser	GCG TCG	1	30, 11	1.8, −0.8	−0.65	−0.08
33	Arg Cys	CGC TGC	1	35, 58	−4.5, 2.5	−0.28	0.26
33	Arg His	CGC CAC	1	35, 61	−4.5, −3.2	0.02	−0.06
33	Arg Leu	CGC CTC	1	35, 7	−4.5, 3.8	0.23	0.37
34	Gly Ser	GGC AGC	1	41, 22	−0.4, −0.8	0.20	0
38	Gly Ser	GGC AGC	2	41, 22	−0.4, −0.8	0.20	0
39	Gly Asp	GGT GAT	1	48, 51	−0.4, −3.5	0.44	−0.21
47	Asp Tyr	GAC TAC	1	49, 59	−3.5, −1.3	−0.32	0.93
48	Ala Thr	GCA ACA	4	23, 6	1.8, −0.7	−0.59	0.36
48	Ala Val	GCA GTA	1	23, 23	1.8, 4.2	−0.36	0.87
50	Met Ile	ATG ATA	1	100, 1	1.9, 4.5	−0.37	0.55
51	Cys Tyr	TGC TAC	1	58, 59	2.5, −1.3	0.01	0.28
51	Cys Arg	TGC CGC	1	58, 35	2.5, −4.5	0.28	−0.26
52	Ala Val	GCC GTC	1	19, 13	1.8, 4.2	−0.36	0.87
55	Ala Thr	GCC ACC	1	19, 51	1.8, −0.7	−0.59	0.36
59	Arg His	CGT CAT	2	58, 39	−4.5, −3.2	0.02	−0.06
59	Arg Cys	CGT TGT	1	58, 42	−4.5, 2.5	−0.28	0.26
60	Asp Asn	GAC AAC	2	49, 76	−3.5, −3.5	−0.34	0.35
62	Leu Phe	CTC TTC	1	7, 56	3.8, 2.8	−0.08	0.08
71	Leu Phe	CTC TTC	2	7, 56	3.8, 2.8	−0.08	0.08
80	Gly Asp	GGC GAC	2	41, 49	−0.4, −3.5	0.44	−0.21
80	Gly Ser	GGC AGC	1	41, 22	−0.4, −0.8	0.20	0
83	Gly Asp	GGT GAT	2	48, 51	−0.4, −3.5	0.44	−0.21
102	Glu Lys	GAA AAA	1	73, 77	−3.5, −3.9	−0.35	0.37

[a] For each missense mutation, the following parameters are tabulated. No. is the number of independent mutations with the identical mutation. Codon frequency is the percentage use of the codon before and after change. The first number is the frequency at which the wild-type codon is used within the possible codons for the particular amino acid. The second number is the same parameter for the altered codon. ATG is 100% since ATG is the only Met codon. Charge change is the net change in charge after substitution. Hydropathicity is the hydropathicity value (22) before and after substitution. P_a change is the net change in Chou-Fasman alpha-helix frequency after substitution (7); a negative number indicates a decrease in alpha-helix character. P_b change is the net change in Chou-Fasman beta-sheet frequency after change (7); a positive number indicates an increase in beta-sheet character.

and no detectable deletion or insertion. Four of these single-base changes were found upstream of the coding sequence, in or near the presumed Shine-Dalgarno ribosome-loading site. These mutations presumably affected only the translation of the *S* mRNA and will be described elsewhere (R. Raab, G. Neal, C. Sohaskey, and R. Young, submitted for publication). Forty-eight mutations resulted in codon changes within the *S* cistron (Fig. 2; Table 2). Twenty-five of the 107 codons were affected, and 35 *S* alleles with different predicted amino acid sequences were formed. Three of these isolates were nonsense mutations: Gln_{19}→TAA (one case) and Trp_{56}→TGA (two cases). The remaining 45 mutants (33 alleles in 23 different codons) were missense mutations. Since hydroxylamine mutagenesis primarily facilitates the transitions C→T or G→A, only 70 of the 107 *S* codons could be expected to be mutable to missense codons under our conditions (Neal, M.S. thesis). All but 5 of the 50 single-base changes were of the expected C→T or G→A type. There were also three G→T (codons 22, 30, and 47) and one C→G (codon 25) transversions, and one T→C transition (codon 51). Thus, out of 70 codons potentially convertible to missense alleles by classical hydroxylamine-induced mutation, more than one-fourth were found to be affected and converted into 28 different alleles. The *S* gene can be conveniently divided into three regions of approximately equal size, each of which has a major segment of hydropathic residues (Fig. 3). The missense mutations are found primarily in the first two regions from the amino terminus: 17 mutations in 9 codons between residues 1 and 35: 20 mutations in 10 codons between residues 36 and 69; only 7 mutations in 5 codons in the remaining residues (codons 70 to 107).

Classes of *S* missense mutations. The missense mutations isolated (plus the preexisting allele, *Sts9B*, which had been cloned previously [13]) can be grouped into three general classes: charge changes, substitutions for glycines, and neutral changes. Fourteen of these 34 different mutant alleles have an altered net charge based on the predicted primary sequence. Of the 11 glycine codons in the wild-type sequence, 6 give rise to a total of nine different missense alleles which inactivate *S* lethality under our conditions. The neutral class involves missense mutations that result in no change in charge. Fifteen of the 34 missense alleles would be predicted to have neither a charge change nor substitution for glycine. A number of these neutral changes in the first two regions of *S* are relatively conservative, resulting in little or no change in the hydropathic character of the residue. These substitutions are Met_3→Ile, Ala_{23}→Val, Ala_{23}→Thr,

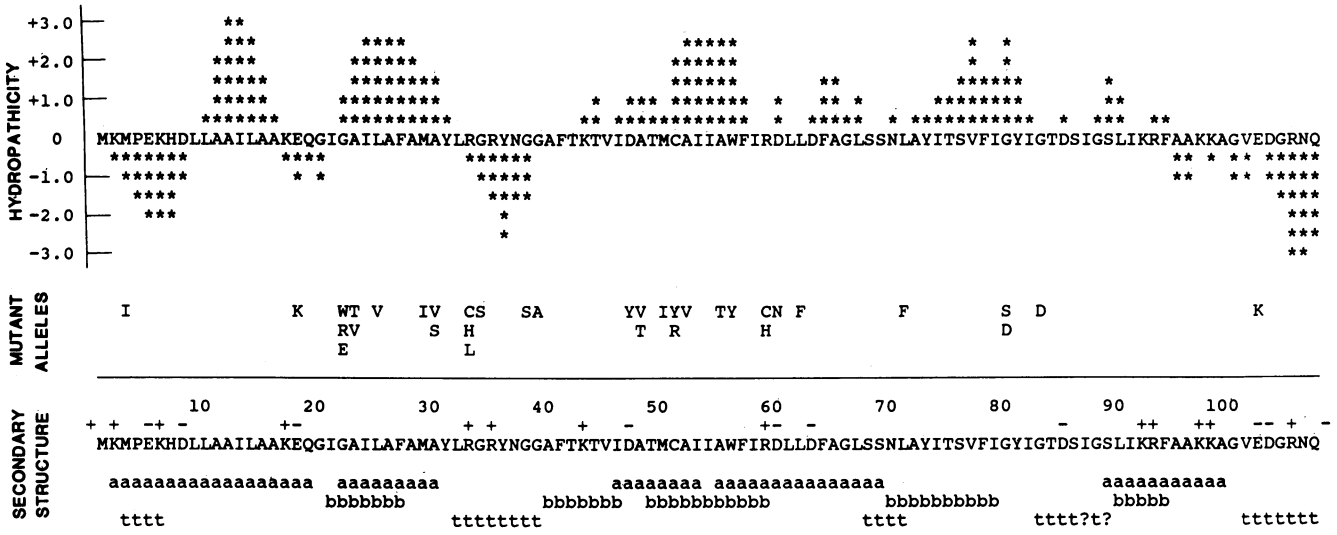

FIG. 3. Hydropathicity and Chou-Fasman analysis of the *S* gene. The hydropathicity profile for *S* as calculated by the Kyte-Doolittle algorithm with a window of six residues is shown, superimposed on the primary sequence of the *S* gene in single-letter amino acid code. Values above the median line are hydropathic; negative values are hydrophilic. Below the hydropathicity plot, the amino acid changes found in missense S^- alleles are shown. Below the mutant alleles, the *S* sequence is shown again, with charges and residue numbers. Finally, below the primary sequence, the consensus predictions made by Chou-Fasman analysis are shown as follows: a, alpha-helical region; b, beta-sheet region; t, reverse turn.

Leu_{25}→Val, Met_{29}→Ile, and Ala_{30}→Val in the region I and Ala_{48}→Val, Ala_{48}→Thr, Met_{50}→Ile, Ala_{55}→Val, and Leu_{62}→Phe in region II.

Codon usage and tRNA availability. One possible interpretation of the meaning of the neutral mutations would be that the mutant codon might be a rare codon which lowered *S* expression enough to eliminate the inducible lethality. Only in a few instances could such an interpretation be consistent with known codon frequencies (Table 2). The two Ala→Thr substitutions at residues 23 and 48 result in a low-frequency ACA codon, but a third Ala→Thr change at residue 55 also results in a defective *S* gene product even though the

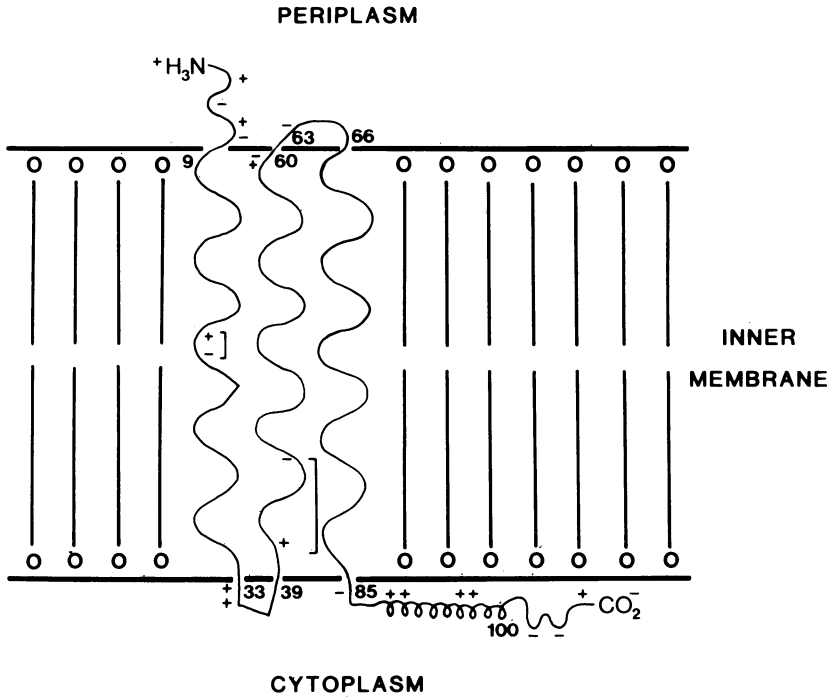

FIG. 4. Model for disposition of the *S* gene product in the inner membrane. All charges are shown in approximate location along the polypeptide chain. Numbers indicate positions of structural features discussed in the text. Brackets indicate putative salt bridges between charged side chains. Each sinusoidal wavelength represents approximately one alpha-helical turn.

substituted codon is a high-frequency ACC (21). Thus, only in the case of the three Met→Ile changes at codons 3, 29, and 50 is there a reduction in codon frequency consistent with a potential lowering of *S* translation. In these cases, the new Ile codon is ATA, found in only 1% of Ile codons in a group of 25 nonregulatory genes surveyed by Konigsberg and Godson (21). However, the wild-type *S* gene has one ATA in its 11 Ile codons, and another cistron on the same transcriptional unit, the *J* gene, has 10 ATA triplets in its 55 Ile codons (8). Based on the low intensity of labeling of pS in bacteriophage infections (3), the *S* gene is translated at low levels compared with other cistrons on the lambda late message, and its expression is unlikely to be limited by codon frequency. We infer, then, that the lesions in these missense mutants derive from a qualitative change in the primary sequence of the *S* gene product rather than from a change in its synthesis rate.

DISCUSSION

The *S* protein is thought not to have enzymatic activity but rather, in a precisely scheduled manner, to induce permeability of the cytoplasmic membrane, presumably by forming pores (31, 35). Since the *S* gene is small (107 codons), much of its primary sequence could be expected to be involved as imbedded structures within the membrane. Since pS is an inner membrane protein, it probably inserts into the bilayer posttranslationally and without interaction with the secretory apparatus of the cell (29). Are there features of its primary sequence which suggest what might drive such a spontaneous insertion? Theoretical analysis of the *S* gene sequence by Chou-Fasman and Kyte-Doolittle hydropathicity algorithms are shown in Fig. 3. There are two regions of consensus alpha-helical prediction that are also strongly hydropathic and carry no net charge, the first between codons 8 and 32 and the second between codons 39 and 62. Both of these regions contain enough residues to span the hydrophobic core of the bilayer as helices (13, 34). In fact, the first 35 codons resemble canonical membrane-spanning regions found in other polypeptides, with a stretch of 20 to 24 hydrophobic, charge-neutral residues terminated by two or more positive charges, presumably abutting the inside surface of the inner membrane (34, 36). For our working model of the *S* protein, therefore, we suggest that the first hydrophobic helix spans the membrane with the amino terminus disposed in the periplasm and a positively charged pair of residues at the cytoplasmic interface (Fig. 4). Similarly, the second hydrophobic helical region would span the membrane in the opposite orientation. Both helical spanning domains would have a salt-bridged pair of charged side chains, between Lys_{17} and Glu_{18} and between Lys_{43} and Asp_{47}. Each salt bridge can be shown by computer modeling to be easily within the rotational constraints of the side chains (data not shown). Intrahelical and interhelical salt bridges allow hydrophobic helices containing a large number of charged residues to span the membrane in rhodopsin (11). The presence of the ion pairs would confer an amphiphilic character to these imbedded helices, as found for the purple membrane protein rhodopsin, which is known to span the membrane multiple times.

This hypothetical arrangement not only takes advantage of the two putative helical regions but also suggests that the glycine-rich hydrophilic region between these domains (residues 33 through 39) may serve as a reverse-turn hinge. There is little in the rest of the molecule which suggests a particular conformation or intramembrane disposition other than the lack of charge in the 20 positions between residues 64 and 84 and the extremely hydrophilic and highly charged carboxy-terminal 23 residues (8 charged, 4 polar uncharged, plus the ionized carboxylate moiety). Intuitively, it is difficult to conceive of a spontaneous penetration of such a highly charged segment, and thus our working model redirects the *S* chain back through the membrane with undefined secondary structure between residues 64 and 84 and has the carboxy terminus disposed in the cytoplasm. This model has another feature which recommends it. Assuming that the two amino-terminal helical regions formed first and acted as a paired helix insertion domain, no net charge would be inserted either in or through the membrane (Fig. 4).

Can the collection of missense mutants be rationalized in terms of this appealingly simple but highly speculative model? Superimposing the mutational pattern on the proposed membrane and hydropathic domain structure, we find that nearly every missense change leading to a charge alteration falls within a proposed membrane-spanning region (Fig. 3). The exceptions lie at the putative C terminus of the first membrane helix, where the change of an Arg codon to either a Leu, His, or Cys inactivates the gene and, at the extreme C terminus, a Glu_{102}→Lys. The necessity of having the positively charged "anchor" for stable membrane integration of a membrane-spanning helix has been shown in other systems (1, 9), and thus the essential nature of the Arg_{33} charge is not surprising. Thus, with the exception of the Glu_{102}→Lys (net change of +2), all of the charge change mutations make sense in terms of the model.

A crucial feature of this model is that the *S* polypeptide immediately reverse its direction after the end of the first helical span and be redirected back through the membrane. There are three Gly residues, at positions 34, 38, and 39, which contribute heavily to the strong turn prediction by Chou-Fasman analysis and might be expected to be essential for the putative reverse turn. Missense mutations at all three Gly residues are found to inactivate the gene, again supporting the model. Most of the other substitutions for Gly residues result in the charge changes discussed above. A notable exception is Gly_{22} mutations, which give rise not only to charge changes (Gly_{22}→Glu, Gly_{22}→Arg) but also to the neutral missense change Gly_{22}→Trp. This suggests that the turn capacity of the Gly_{20}-Gly_{22} pair is essential to the function of *S*. In support of this conclusion, we note that there is a Pro residue in exactly the same position within the presumably analogous membrane-spanning hydrophobic amino-terminal region of the ϕX174 lysis gene *E* (36). It is tempting to speculate that an interruption or "kink" in the otherwise highly alpha-helical primary sequence of this domain is necessary for the lytic competency of these evolutionarily unrelated lysis genes (36).

The charge changes and Gly substitutions, taken together, are thus in general agreement with the simple model derived from the *S* primary sequence. Much more intriguing and difficult to interpret, however, is the relatively conservative neutral mutation class. Almost all of these are clustered in the two putative membrane-spanning helical regions (Table 2; Fig. 3). Only a few of these can be explained in terms of a significant increase in side chain polarity, e.g., Ala_{30}→Ser and, to much less of an extent, the Ala→Thr changes at codons 23 and 48. However, a number of neutral changes in these putative domains result in either no significant change or an increase in the consensus hydropathicity values, especially the Ala→Val, Leu→Val, and Met→Ile changes. This group, representing independent mutations in seven different codons, thus cannot be explained in terms of

increased polarity or charge change. There is no obvious common feature of these mutations, other than the fact that, in terms of the Chou and Fasman empirical parameters for alpha-helical and beta-sheet frequency, each change results in a decrease in alpha-helical probability and an increase in beta-sheet probability (Table 2). We would include in this group the Ala→Thr changes at codons 23 and 48 (five independent mutants) and Cys_{51}→Tyr, bringing the total of these "helix-down" (but otherwise neutral) mutations to 14 independent mutations in 8 different codons, leading to 10 different amino acid sequence variants. Also, two preexisting *S* alleles can be included in this class: the *Sts9B* allele, nonfunctional at 42°C, is Ala_{55}→Thr, and the *supF*-suppressed *Sam7*, which is slightly temperature sensitive (32), is Trp_{56}→Tyr.

Although it seems hard to believe that this uniformity is entirely coincidental, it is equally difficult to take the notion too seriously that each change actually reflects a decrease in the net propensity of a domain to assume a helical structure. We have noted that it is possible, although unlikely, that the Met→Ile changes are actually "translation-downs," due to the substitution of a rare codon for a common one. Most of the neutral, nonglycine changes are in Ala codons, and all of the changes in these codons would result in the substitution of a bulky side chain for a very small one and thus possibly distort the ability of a membrane-spanning helix to pair up smoothly with other such structures.

Although the hypothetical disposition of the *S* polypeptide within the membrane is derived from the most elementary theoretical analysis and inspection, almost all of the mutations isolated are consistent with the predictions of the model (Fig. 4). The seemingly strict requirement for the highly helical residues in the two prospective intramembrane domains and for the overall charge neutrality in the first two-thirds of the sequence suggests a simple mechanism for the insertion event, in which the first 70 residues insert into and partially through the bilayer as paired, charge neutral helices. A reasonable guess would be that this structure is sufficient for the insertion process, and that the uncharged region between residues 65 and 85 is forced into a membrane-spanning conformation which, by its intrinsic character, it might not assume. This mechanism would help explain the mutational clustering in the first two helical regions. In addition, the multiple-spanning model for *S* conformation makes it much easier to envision formation of large pores by oligomerization of *S* protein within the membrane. The two paired helices, or perhaps three if the third region between 64 and 84 is also a membrane-spanning helix, would form an imbedded monomer which could oligomerize within the bilayer to result in a large enough pore to release small proteins. We would guess that the ion pair regions of the imbedded helices would face inward toward the aqueous channel of the pore or perhaps form interaction junctions either between imbedded helices of a single monomer, as in rhodopsin (11), or between monomers. In any model of pore formation, one would expect that some of our large group of S^- missense mutants would be negatively dominant, as demonstrated for analogous mutations in the *lacY* gene (25). We have recently transferred each these *S* mutations to the bacteriophage lambda genome, which has allowed us to assess the dominant or recessive characteristics of these mutations and also to determine the effect of each change on the physiology of the infective cycle. These data will be described elsewhere (Raab et al., submitted for publication). In addition, having the mutations in the phage background will allow us to do reversion studies to look for compensatory changes within the *S* cistron and thus to identify regions of possible domain interaction. Finally, we will be able to examine the fate of the protein products of this collection of mutant *S* alleles by labeling the peptides synthesized after infection of preirradiated cells (3, 4). In our hands, this is the only practical way to detect the product of the *S* gene, which was one of the last essential gene products of bacteriophage lambda to be detected (3). With this approach, a specific molecular defect can be assigned to each missense allele, and it should be possible to determine which of these mutations affect entry into the bilayer and which affect other functions of the *S* protein (i.e., oligomerization).

ACKNOWLEDGMENTS

This research was supported by Public Health Service grant 5-RO1-GM27099-07 from the National Institutes of Health.

Donna Cleaves, Kim Neal, and Lisa Lohman were responsible for graphics preparation. The technical assistance of Sarah Plummer Raab was indispensable.

LITERATURE CITED

1. **Adams, G. A., and J. K. Rose.** 1985. Structural requirements of a membrane-spanning domain for a protein anchoring and cell surface transport. Cell **41**:1007–1015.
2. **Adhya, S., A. Sen, and S. Mitra.** 1971. The role of gene *S*, p. 743–746. *In* A. R. Hershey (ed.), The bacteriophage lambda. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
3. **Altman, E., R. K. Altman, J. M. Garrett, R. J. Grimaila, and R. Young.** 1983. *S* gene product: identification and membrane localization of a lysis control protein. J. Bacteriol. **155**: 1130–1137.
4. **Altman, E., K. Young, J. Garrett, R. Altman, and R. Young.** 1985. Subcellular localization of lethal lysis proteins of bacteriophages λ and φX174. J. Virol. **53**:1008–1011.
5. **Bienkowska-Szewczyk, K., B. Lipinska, and A. Taylor.** 1981. The *R* gene product of bacteriophage λ is the murein transglycosylase. Mol. Gen. Genet. **184**:111–114.
6. **Campbell, J. H., and B. G. Rolfe.** 1975. Evidence for a dual control of the initiation of host-cell lysis caused by phage lambda. Mol. Gen. Genet. **139**:1–8.
7. **Chou, P. Y., and G. D. Fasman.** 1978. Prediction of the secondary structure of proteins from their amino acid sequence. Adv. Enzymol. **47**:45–148.
8. **Daniels, D., J. Schroeder, W. Szybalski, R. Sanger, A. Coulson, G. Hong, D. Hill, G. Petersen, and F. Blattner.** 1983. Complete annotated lambda sequence, p. 519–661. *In* R. W. Hendrix, J. W. Roberts, F. W. Stahl, and R. A. Weisberg (ed.), Lambda II. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
9. **Davis, N. G., J. D. Boeke, and P. Model.** 1984. Fine structure of a membrane anchor domain. J. Mol. Biol. **181**:111–121.
10. **Davis, R. W., D. Botstein, and J. R. Roth.** 1980. Advanced bacterial genetics, p. 94–97. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
11. **Eisenberg, D., R. M. Weiss, and T. C. Terwilliger.** 1982. The helical hydrophobic moment: a measure of the amphiphilicity of a helix. Nature (London) **299**:371–374.
12. **Engelman, D. M., R. Henderson, A. D. McLachlan, and B. A. Wallace.** 1980. Path of the polypeptide in bacteriorhodopsin. Proc. Natl. Acad. Sci. USA **77**:2023–2027.
13. **Engelman, D. M., and T. A. Steitz.** 1981. The spontaneous insertion of proteins into and across membranes: the helical hairpin hypothesis. Cell **23**:411–422.
14. **Garrett, J. M., R. Fusselman, J. Hise, L. Chiou, D. Smith-Grillo, J. Schultz, and R. Young.** 1981. Cell lysis by induction of cloned lambda lysis genes. Mol. Gen. Genet. **182**:326–331.
15. **Garrett, J. M., and R. Young.** 1982. Lethal action of bacteriophage λ *S* gene. J. Virol. **44**:886–892.
16. **Harris, A. W., D. W. A. Mount, C. R. Fuerst, and L. Siminovitch.** 1967. Mutations in bacteriophage lambda affecting host cell lysis. Virology **32**:553–569.

17. **Herskowitz, I., and D. Hagen.** 1980. The lysis-lysogeny decision of phage lambda: explicit programming and responsiveness. Annu. Rev. Genet. **14**:399–445.
18. **Holmes, D. S., and M. Quigley.** 1981. A rapid boiling method for the preparation of bacterial plasmids. Anal. Biochem. **114**: 193–197.
19. **Itakura, K., T. Hirose, R. Crea, A. Riggs, H. L. Heyneker, F. Bolivar, and H. W. Boyer.** 1979. Expression in *Escherichia coli* of a chemically synthesized gene for the hormone somatostatin. Science **198**:1056–1063.
20. **Josslin, R.** 1970. The lysis mechanism of phage T4: mutants affecting lysis. Virology **40**:719–726.
21. **Konigsberg, W., and G. N. Godson.** 1983. Evidence for use of rare codons in the *dnaG* gene and other regulatory genes of *Escherichia coli*. Proc. Natl. Acad. Sci. USA **80**:687–691.
22. **Kyte, J., and R. F. Doolittle.** 1982. A simple method for displaying the hydropathic character of a protein. J. Mol. Biol. **157**:105–132.
23. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning: a laboratory manual, p. 250–251. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
24. **Messing, J., and J. Vieira.** 1982. A new pair of M13 vectors for selecting either DNA strand of double-digest restriction fragments. Gene **19**:269–276.
25. **Mieschendahl, M., D. Büchel, H. Bocklage, and B. Müller-Hill.** 1981. Mutations in the *lacY* gene of *Escherichia coli* define functional organization of lactose permease. Proc. Natl. Acad. Sci. USA **78**:7652–7656.
26. **Miller, J. H.** 1972. Experiments in molecular genetics, p. 13–23. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
27. **Miller, J. H.** 1972. Experiments in molecular genetics, p. 431–435. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
28. **Mukal, F., G. Streisinger, and B. Miller.** 1967. The mechanism of lysis in phage T4-infected cells. Virology **33**:398–404.
29. **Oliver, D.** 1985. Protein secretion in *Escherichia coli*. Annu. Rev. Microbiol. **39**:615–648.
30. **Reader, R. W., and L. Siminovitch.** 1971. Lysis defective mutants of bacteriophage lambda: genetics and physiology of *S* cistrons mutants. Virology **43**:607–622.
31. **Reader, R. W., and L. Siminovitch.** 1971. Lysis defective mutants of bacteriophage lambda: on the role of *S* function in lysis. Virology **43**:623–627.
32. **Rolfe, B. G., and J. H. Campbell.** 1977. Genetic and physiological control of host cell lysis by bacteriophage lambda. J. Virol. **23**:626–636.
33. **Sundar Raj, C. V., and H. C. Wu.** 1973. *Escherichia coli* mutants permissive for T4 bacteriophage with deletion in *e* gene (phage lysozyme). J. Bacteriol. **114**:656–665.
34. **von Heijne, G.** 1981. Membrane proteins: the amino acid composition of membrane-penetrating segments. Eur. J. Biochem. **120**:275–278.
35. **Wilson, D. B.** 1982. Effect of the lambda *S* gene product on properties of the *Escherichia coli* inner membrane. J. Bacteriol. **151**:1403–1410.
36. **Young, K. D., and R. Young.** 1982. Lytic action of cloned ϕX174 gene *E*. J. Virol. **44**:993–1002.
37. **Young, R., J. Way, S. Way, J. Yin, and M. Syvanen.** 1979. Transposition mutagenesis of bacteriophage lambda: a new gene affecting cell lysis. J. Mol. Biol. **132**:307–322.