Molecular Studies of the *sdaA* and *sdaB* Genes and Their Gene Products in *Escherichia coli* K-12

HONGSHENG SU

A Thesis

in

The Special Individual

Program

Presented in Partical Fulfillment of the Requirements

for the Degree of Doctor of Philosophy at

Concordia University

Montreal, Quebec, Canada

April 1991

ISBN 0-315-68774-6

CONCORDIA UNIVERSITY
Division of Graduate Studies

This is to certify that the thesis prepared

By: HONGSHENG SU

Entitled: MOLECULAR STUDIES OF sdaA AND sdaB GENES AND THEIR GENE PRODUCTS IN ESCHERICHIA COLI K - 12

and submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

complies with the regulations of the University and meets the accepted standards with respect to originality and quality.

Signed by the final examining committee:

______________________ Chair
Dr. P. Bird

______________________ External Examiner
Dr. M. Dubow

______________________ External Examiner
Dr. R. D'Ari

______________________ Examiner
Dr. J. Fraser

______________________ Examiner
Dr. R.K. Storms

______________________ Examiner
Dr. M.B. Herrington

______________________ Thesis Supervisor
Dr. E.B. Newman

Approved by ______________________
Chair of Department or Graduate Program Director

August 5 1991

____
Dean of Faculty

ABSTRACT

Molecular Studies of the *sdaA* and *sdaB* Genes and Their Gene Products in *Escherichia coli* K-12

Hongsheng Su, PH.D.

Concordia University, 1991

This work demonstrates the existence in *Escherichia coli* K-12 of two L-serine deaminating enzymes, L-serine deaminase (L-SD)#1 and L-SD#2. It demonstrates that the structure of L-SD#1 is coded by the *sdaA* gene, which has been cloned and sequenced. The entire *sdaA* gene was fused to the *lacZ* gene by mutating the stop-codon of *sdaA* and ligating in-frame to *lacZ*. The fused gene directed the formation of a large protein showing both L-SD and β-galactosidase activities. L-SD#1 has been extensively purified for the first time by use of a three-part fusion protein and some of its characteristics studied.

L-SD#2 is synthesized in wild-type cells in LB medium. A mutation in the *sdaX* gene established its expression in minimal medium. An insertion in *sdaB* abolished L-SD#2 activity in an *sdaA*::Cm^r *sdaX* strain, allowing the *sdaB* gene to be cloned by restoring growth on L-serine. The *sdaA* gene was located at 41 minutes; *sdaB* and *sdaX* both were located near 60.1 minutes and may be the same gene. Some experiments directed towards the identification of the metabolic role of L-SD are included.

ACKNOWLEDGEMENTS

I would very much like to thank Dr. E.B. Newman for the guidance and providing me the best academic support during this study. She, as supervisor, motivated and encouraged me to accept the challenge of academic work and also as a friend, give a lot of help since I came to Montreal.

I would also like to extend my thanks to Drs. R. K. Storms and M.B. Herrington, the members of committee for their helpful advice and suggestions.

Thanks also go to Drs. R. D'Ari, M. DuBow, R. Roy, and C. Cupples for their interest, advices, and help. I would like to give the special thanks to Dr. F. Lang, who helped me to do DNA sequencing in his laboratory, and then introducing this terrific method into our laboratory.

I am grateful to my colleagues in this lab, J. Garnon, E. Brown, R. Lin, Z. Shao, Y. Li, J. Moniakis, M. San Martano, and Y. Shen for friendship, help, and advice.

I also would like to thank Nicole Baur for technical help in the DNA sequencing in F. Lang's lab.

I must also thank Concordia University for giving me a J. W. McConnell Memorial Fellowship in this study, the Quebec government for foreign student fee exemption, and NSERC for financial assistance through the grant of Dr. E.B. Newman.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

INTRODUCTION

L-Serine Deaminase (L-SD) is an *Escherichia coli* K-12 enzyme which can deaminate L-serine to produce pyruvate and ammonia. The enzyme is of interest because it is produced by *E. coli* K-12 even when the cells are grown in glucose-minimal medium [Newman *et al.*, 1985a] where it would not appear to play any useful role. Pardee and Prestidge described this activity in 1955, and showed that it is induced by growth with glycine and L-leucine, but not L-serine [Pardee and Prestidge, 1955, Newman and Walker, 1982b]. Since the 1974 study of Isenberg and Newman [Isenberg and Newman, 1974], L-SD has been the focus of much of the work in Newman's laboratory.

Apart from the anomaly that L-SD is not induced by its substrate but is induced by other amino acids, synthesis of L-SD is influenced by a startling array of environmental and genetic factors. L-SD is induced by DNA damaging agents such as UV irradiation and mitomycin, by anaerobic growth, by growth at 42°C and by growth in the presence of ethanol. L-SD activity is also affected by mutations in at least 7 different genes, and is regulated by two of the *E. coli* global regulatory systems.

Despite the variety of physiological and genetic information accumulated, at the start of this work, very little was known about the molecular nature of this enzyme, or that of the gene, *sdaA*, which codes for it.

In this thesis, I present a study of molecular aspects of L-serine deamination in *E. coli*. I show that it involves not one but two L-serine deaminases coded by 2 different genes, *sdaA* and *sdaB*. The *sdaA* gene has been studied in detail, cloned, sequenced and shown to be the structural gene for L-SD#1. The *sdaB* gene has been identified, and cloned. Aspects of the regulation of both genes have been studied.

SdaA: I isolated a mutation in what proved to be a new gene *sdaA* by λplacMu9 insertion. The gene was cloned and sequenced and found to code for a 483-amino acid protein. The gene was shown to code for L-SD#1. This was done by fusing the whole *sdaA* gene to the *lacZ* gene, using site-directed mutagenesis to change the *sdaA* stop-codon into an *E.co*RI restriction site. The fact that the resultant *sdaA-lacZ* fusion protein showed L-SD activity proved that the *sdaA* gene is indeed the structural gene. I also constructed a three-part fusion using a DNA sequence coding for a collagenase-sensitive peptide [Germino and Bastia, 1984], inserted between the *sdaA* C-terminal and the *lacZ* N-terminal. This was used to purify the fusion protein, and liberate L-SD by digestion with collagenase.

SdaB: I have demonstrated that there is a second L-SD activity in *E. coli*, L-SD#2, which is only synthesized in complex media like Luria broth (LB medium). The enzyme L-SD#2 was shown to have biochemical characteristics similar to those of L-SD#1, such as a broad pH range and a requirement for high substrate concentrations. The enzyme was inactive in extracts, but could be activated in extracts with iron and dithiothreitol (DTT). A mutation which abolished the synthesis of L-SD#2 was isolated by λplacMu9 insertion. A gene which can complement this mutation was cloned from the mutant in which L-SD#2 was expressed in minimal medium. The DNA fragment containing the *sdaB* gene showed homology with a fragment carrying the *sdaA* gene. These two newly isolated mutations were mapped, and a location near 41.0 and 60.1 minute of *E. coli* chromosome map determined for *sdaA* and *sdaB*, respectively.

Studies on regulation: I took advantage of the *lacZ* insert in *sdaA* to study the regulation of *sdaA* gene expression with more sensitivity than had previously been possible. Using both the *sdaA-lacZ* fusion strain and plasmid, I confirmed that β-galactosidase synthesis was regulated by the *ssd* gene product, by glycine and L-leucine, by anaerobic growth, and by UV irradiation but was not altered in *recA* mutants. Regulation of the *sdaB* gene was studied in less detail. However it was shown to be subject to minor regulation by the *lrp* gene product.

Possible function of L-SD: I also used both *sdaA* and *sdaB* mutants to study the function of L-SD. It was shown that L-SD is involved in detoxification of L-serine, and that strains which overproduce this enzyme activity can use L-serine as carbon, energy and nitrogen source. However, that may not be the real function of L-SD.

PART 1. MUTATIONS AFFECTING L-SD ACTIVITY

It has proved quite simple to isolate mutant strains with more or less L-SD activity than the parent strain. The selections are based on the fact that the usual *E. coli* K-12 strain cannot grow with L-serine as sole carbon source but can use a combination of L-serine, glycine and L-leucine as carbon source [Pardee and Prestidge, 1955, Newman and Walker, 1982c]. It has been possible then to isolate mutants with more L-SD by selecting for growth with L-serine as carbon source [Newman *et al.*, 1982b, Lin *et al.*,1990] and to isolate mutants with less L-SD by screening for strains which cannot grow with L-serine, glycine and L-leucine [Newman *et al.*, 1985a, Newman *et al.*, 1985b].

Including the new genes defined in this work, 7 genes have been shown to affect L-serine deaminase activity. The five known prior to this work are reviewed in the following paragraphs.

1-1. Mutants Showing Decreased L-SD Activity

1-1a. Post-translational activation mutants

Mutants with decreased L-SD activity were isolated by killing with ampicillin any cells able to grow with L-serine, glycine and L-leucine. Cells unable to grow with these amino acids, called SGL$^-$ mutants, were then tested for L-SD activity in the whole cell assay. Those without whole cell L-SD activity were retained for further study.

It was a considerable surprise to find that mutations in 3 different genes resulted in a SGL$^-$ phenotype as judged by the whole-cell assay but showed a great deal of activity in crude extracts, treated with iron and dithiothreitol [Newman *et al.*, 1985a, Newman *et al.*, 1985b]. It was suggested that in all 3 mutants, the *sdaA* gene product is in fact synthesized, but it is in an inactive form and must be activated before it can deaminate L-serine. The 3 activation genes defined by mutants isolated so far are *sda191*, *sda128* and *sda84*.

The mechanism by which L-SD is activated *in vivo*, and the nature of the activation gene products is unknown. Among the possible mechanisms are the following:

i. The structural gene codes for an inactive protein. *Sda191,128* and *84* code for all or part of a system that transforms the protein into an active form.

ii. L-SD is inactive in the cell because it lacks a cofactor which the cell cannot make. The 3 activation genes would then code for enzymes of cofactor biosynthesis.

iii. The cell normally makes an inhibitor of L-SD function. This inhibitor is broken down by the activation gene products.

In all cases, iron and DTT would substitute for or bypass the effect of the activation genes. The 3 mutants all require thiamine for growth. The reason for this is also not known.

1-1b. Mutants deficient in L-SD both *in vivo* and *in vitro*

One would expect that some mutations in the structural gene would totally remove L-SD activity, whether in whole cells or in extracts, since the coding sequence for the protein would no longer exist. The first such putative structural gene mutation affecting L-SD, in strain MEW15, was isolated by MudX insertion [Newman *et al.*, 1985a]. The strain gave very low L-SD activity *in vivo* and *in vitro* with or without iron and DTT. The fact that β-galactosidase synthesis from the fused *lacZ* was regulated by the same factors as L-SD (glycine and L-leucine, UV irradiation and growth in LB medium) strongly supported the idea that the insertion was in the structural gene coding for L-SD.

Strain MEW15 would have been an appropriate strain for these studies, except that- for unknown reasons- it was not susceptible to most genetic manipulations. After long attempts to study MEW15, it became clear that it was necessary to isolate a new structural gene mutation. The fact that strain MEW15 still had L-SD activity in LB medium, however, was the first indication of the existence of a second L-SD gene which is only expressed during growth in complex medium [Newman *et al.*, 1985a].

The *sdaA* mutant described in this work was isolated in order to identify the structural gene for L-SD#1 and proved to be affected in the structural gene. The *sdaA* mutant had no activity *in vivo* or *in vitro*, with or without iron and DTT.

1-2. Mutants with Increased L-SD Activity

Mutants with high L-SD activity were isolated by direct selection on plates with L-serine as carbon source. Mutations in two genes, known as *ssd* [Newman *et al.*, 1981, Newman *et al.*, 1982b] and *lrp* [Lin *et al.*, 1990], led to very high L-SD activity. Both of these genes are involved in the global regulatory circuits of *E. coli*, suggesting that whatever its function may be, the regulation of L-SD is important in *E. coli* metabolism. In a third class, *gos*, slower growth on serine plates was not associated with an increase in L-SD activity [Brown *et al.*, 1990].

1-2a. Study of *ssd* mutants

The *ssd* gene was mapped near 87 min on the *E. coli* chromosomal map [Morris and Newman, 1980]. Mutants in this gene were isolated by growth with L-serine as sole carbon source. They showed very high L-SD activity as was expected from this selection. In addition, the mutations were exceedingly pleiotropic, affecting several areas of cell metabolism. Thus, the mutants showed a lowered efficiency of glucose utilization, inability to grow on succinate or related compounds, inability to grow anaerobically, resistance to the antibiotic kanamycin, fluoride sensitivity, and a deficiency in proline and arginine uptake [Newman *et al.*, 1981].

Several mutants showing a phenotype similar to that of *ssd* have been isolated in different laboratories. These include *ecfB* [Plate, 1976], *cpx* [Silverman, 1985], and *eup* [Plate and Suit, 1981], all located in the same area of the *E. coli* chromosomal map and sharing at least some of the phenotypic characteristics. Silverman [Rainwater and Silverman, 1990] suggested that *cpx* and *ssd* may be the same gene in view of the facts that they all map in the same area, and that the *cpxA* mutation isolated in their laboratory also affected properties associated with the *ssd/ecfB/eup* mutant. The results from this laboratory showing that a *cpxA* plasmid can complement the *ssd* mutation further confirm the suggestion. The molecular characteristics of the *cpxA* gene showed that it coded for an inner membrane protein, and it has been suggested [Weber and Silverman, 1988] that this protein is the sensor for one of the global regulatory mechanisms of *E. coli*, one which would regulate L-SD activity and a number of other cell functions.

1-2b. Studies on the *lrp* locus

Whereas the *ssd* gene codes, perhaps, for a membrane sensor protein of a global regulatory system, the *lrp* gene codes for a DNA-binding protein which controls expression of a wide variety of genes, including *sdaA*, all of which are regulated according to the availability of exogenous L-leucine.

This regulatory mechanism was elucidated by Lin and Newman [Lin *et al.*, 1990] in a study of another mutant which uses L-serine as carbon source. This mutant, known first as an *rbl* mutant, and now renamed *lrp* (leucine responsive protein) was isolated by λTn10 insertion and selection for growth with L-serine [Lin *et al.*, 1990]. The gene affected was mapped near 20 min. Deficiency in *lrp* function in an insertion mutant increased L-SD activity seven-fold, and decreased expression of *serA*, coding for phosphoglycerate dehydrogenase, the first enzyme of the L-serine biosynthesis pathway. Activities of two other enzymes, threonine dehydrogenase and acetolactate synthase III, were also altered in the *lrp* mutant.

Similar mutations, *ihb* [Ricca *et al.*, 1989,], *oppI* [Andrews *et al.*, 1986a, Andrews *et al.*, 1986b] and *livR* [Quay *et al.*, 1977], each thought to affect 1 or 2 operons, were described in other laboratories, and are now all considered to be mutations in *lrp*. Each was identified as a regulator of an L-leucine-sensitive operon. The *ihb* gene codes for a protein that binds to the regulatory region in the *ilvH* operon and activates gene expression, in the absence of L-leucine but not in its presence [Ricca *et al.*, 1989]. The *livR* gene codes for a regulator of two high-affinity L-leucine transport systems, which are repressed in the presence of L-leucine and derepressed in the *livR* mutant [Quay *et al.*, 1977]. Similarly, the *oppI* gene product regulates expression of the *oppABCD* operon, which encodes an oligopeptide permease in *E. coli*. L-Leucine can stimulate transcription of this operon [Andrews *et al.*, 1986a, Andrews *et al.*, 1986b].

1-3. Mutants That Grow with L-Serine without Increased L-SD Activity-GOS Mutants

Brown and Newman [Brown *et al.*, 1990] isolated an as yet little understood class of mutants which grow with L-serine as carbon source at 28°C, but do not show high L-SD activity. They also showed that a wild-type strain carrying a *metC* clone could grow with L-serine, presumably using a side reaction of cystathionase to deaminate L-serine. However the GOS mutants did not map in *metC*, nor did they affect cystathionase activity.

PART 2. BIOCHEMICAL CHARACTERISTICS OF L-SD

It would clearly be useful to parallel the genetic study of L-serine deamination by a study of the enzymes responsible for this function. However the *in vitro* study of L-SD has proved to be very difficult because of the apparent extreme instability of this enzyme [Newman and Kapoor, 1980]. When the phosphate buffer used for the original assays [Newman and Kapoor, 1980] was replaced by glycylglycine buffer, stability was greatly improved [Newman *et al.*, 1985a]. Moreover the level of activity measured could be greatly increased by incubating extracts with iron and DTT [Newman *et al.*, 1985a]. This made possible the demonstration, referred to earlier, that some mutants of *E. coli* made an inactive form of the enzyme and could not activate it.

Even in the parent strain, most of the L-SD activity requires incubation with iron and DTT. Some iron and DTT independent activity exists in fresh extracts, but it is quickly lost on freezing. The role of iron and DTT has been studied, without a great deal of success [Newman *et al.*, 1990, Puneker and Lardy, 1987]. Some evidence for a free radical activation mechanism was described. In any case, whatever the *in vitro* activation mechanism may be, it cannot be the same as the *in vivo* one, and its product may or may not be similar to the active form made *in vivo*.

The unactivated form of L-SD was partially purified by chromatography on a Pharmacia FPLC using a Superose-12 column with glycylglycine buffer. When all column fractions were treated with iron and DTT, L-SD activity was found only in the 12th and 13th ml fractions eluted, corresponding to a molecular weight between 40-50 kd [Newman *et al.*, 1990]. The L-SD activity from these fractions required both iron and DTT to activate. No activity was seen in the absence of either or both [Newman *et al.*, 1990].

Purification of L-SD in either its inactive or activated form has been attempted in the Newman laboratory prior to this work and was not successful. The main reason for that is the instability of both forms of the enzyme. In particular, attempts were made to rechromatograph the enzyme from fraction 12 and 13 on the same column, and most of the activity was lost [Newman *et al.*, 1990].

Because of these difficulties, there is very little detailed information about the biochemistry of L-SD, its pH requirements, the substrates it uses and their optimal concentrations. It was therefore a major goal of this work to fuse the structural gene to the *lacZ* gene, so that the fusion protein could be isolated by its β-galactosidase activity, and the L-SD recovered from that protein.

PART 3. POSSIBLE METABOLIC FUNCTION OF L-SD

The synthesis of L-serine requires energy (see below) and so it is odd that cells would both synthesize it and degrade it. No parallel to this is found in the metabolism of the other amino acids except in so far as some amino acids are biosynthetic precursors of others.

Why then does *E.coli* make an L-serine deaminating enzyme when it is grown in the absence of L-serine? L-Serine is known to be toxic to *E. coli* and other organisms [Cosloy and McFall, 1970, Hama *et al.*, 1990, Uzan and Danchin, 1978]. L-SD might function to carefully regulate the L-serine pool so that it is high enough to feed into the many biosynthetic pathways it serves, but not high enough to be toxic.

Further, L-SD converts L-serine to two major metabolites of the cell, pyruvate and ammonia. It may therefore function in the use of L-serine as carbon source and/or as nitrogen source. These aspects are discussed in the following sections.

3-1. L-Serine Toxicity

L-Serine can inhibit growth of *E. coli*. This is probably due to induction of the L-SD which would decrease the intracellular level of L-serine. T. Tsuchiya and his co-workers [Hama *et al.*, 1990] showed that L-serine inhibits the homoserine dehydrogenase, involved in L-isoleucine biosynthesis [Umbarger, 1987]. However, this inhibition can be released by homoserine, L-threonine and α-ketobutyric acid, the intermediates in the biosynthesis of L-isoleucine after homoserine dehydrogenase reaction. The end product, L-isoleucine, also can relieve this L-serine growth inhibition.

In this laboratory, all the strains are derivatives of MEW1 [Newman *et al.*, 1985b], which carries an *ilvA* deletion. Because of this, they all lack threonine deaminase (TD), the first enzyme involved in L-isoleucine biosynthesis [Umbarger, 1987]. Therefore, the cells grown on minimal medium need a supply of L-isoleucine, which also detoxifies L-serine. In the presence of L-isoleucine, L-serine should not be toxic [Hama *et al.*, 1990] unless there are other pathways and enzymes which are inhibited by L-serine in our strains.

3-2. Use of L-Serine as Carbon and Nitrogen Source

E. coli is thought to depend on L-SD for its ability to use L-serine as a carbon source. Cells with sufficient L-SD can convert L-serine to pyruvate and use that as carbon and energy source. If this is true, one has to assume that wild-type cells simply do not have enough L-SD activity. However mutants, at the *ssd* and *lrp* loci, had much more activity and were indeed selected by their ability to grow on L-serine as sole carbon source.

If all this is true, a mutant devoid of L-SD should be unable to use L-serine as a source of either carbon or nitrogen. This was already demonstrated for the activation mutants. However it was possible to argue that these were pleiotropic, and that the failure to use L-serine lay in something other than the loss of L-SD activity. The isolation of a mutation in the structural gene described in this work made it possible to resolve this problem, and demonstrate definitively that L-SD#1 is required for the use of L-serine as carbon source in wild-type *E. coli.*

The formation of pyruvate from L-serine results in the concomitant release of ammonia which might be used as nitrogen source. Indeed the *ssd* and *lrp* mutants can grow with L-serine as nitrogen source even when no other source of nitrogen is provided. However, for the same reasons, there was no direct evidence that L-SD was used for this.

Indirect evidence of the role of L-SD in nitrogen metabolism did exist. Newman and her co-workers [Newman *et al.*, 1976] showed that *E. coli* used glycine as nitrogen source by converting glycine to L-serine, using the cleavage of one molecule of glycine to C1, and condensation of that C1 with a second molecule of glycine via serine hydroxymethyltransferase (SHMT) to form L-serine. Both cleavage enzymes and SHMT were needed for the use of glycine as nitrogen source, but the role of L-SD was inferred and could not be demonstrated at that time.

E. coli also can use L-threonine as nitrogen source by converting L-threonine to glycine, and then glycine to L-serine [Potter *et al.*, 1977] (see below). It was suggested that the key enzyme in this pathway also was L-SD [Potter *et al.*, 1977].

It is perhaps worth noting that the existence of the *GOS* mutants, and the fact that high copy number of *metC* gene also permits growth on L-serine suggests that there must be other ways of deriving carbon, energy and nitrogen from L-serine even if deamination via L-SD is the most efficient one [Brown *et al.*, 1990].

PART 4. BIOSYNTHESIS OF L-SERINE IN *E. coli*

Whatever the role of L-serine deamination may be, one would expect that the cell must closely integrate deamination and biosynthesis of L-serine. I therefore review here what is known about L-serine biosynthesis in microorganisms.

The major pathway of L-serine biosynthesis in *E. coli* [Stauffer, 1987] involves a 3-enzyme pathway, starting from an Embden-Meyerhof intermediate, 3-phosphoglycerate (Fig. 1). This pathway is highly integrated with the pathway of pyridoxine biosynthesis, as elucidated in recent work [Lam and Winkler, 1990]. However, alternative routes of L-serine biosynthesis have been shown in *E. coli* and in other microorganisms.

4-1. Biosynthesis of L-Serine from Phosphoglycerate

As seen in Fig. 1, three enzymes coded by the *serA, serC* and *serB* genes are involved in the biosynthesis of L-serine. The initial reaction is the oxidation of 3-phosphoglycerate to 3-phosphohydroxypyruvate, by 3-phosphoglycerate dehydrogenase, the *serA* gene products. This is followed by transamination to 3-phosphoserine by the *serA* gene product, 3-phosphoserine aminotransferase and dephosphorylation to L-serine by *serB* encoded 3-phosphoserine phosphatase.

This pathway involves the non-energy-transducing dephosphorylation of 3-phosphoserine. Were the phosphoglycerate converted to pyruvate by the Embden Meyerhof pathway, the phosphate bond energy would be used to make ATP. One might therefore expect the use of this pathway to be closely regulated.

In fact, the first enzyme of the pathway is inhibited by both L-serine and glycine [Pizer, 1963]. This feedback inhibition occurred at L-serine and glycine concentrations of $4x10^{-5}M$ and $4.8x10^{-3}M$, respectively.

Though it is not subject to straightforward controls by the product of the pathway in which it is involved, synthesis of the *serA* gene product is nonetheless extensively regulated [McKitrick and Pizer, 1980]. This might be expected from the need to regulate L-serine biosynthesis in accordance with the extent of L-serine degradation. Synthesis

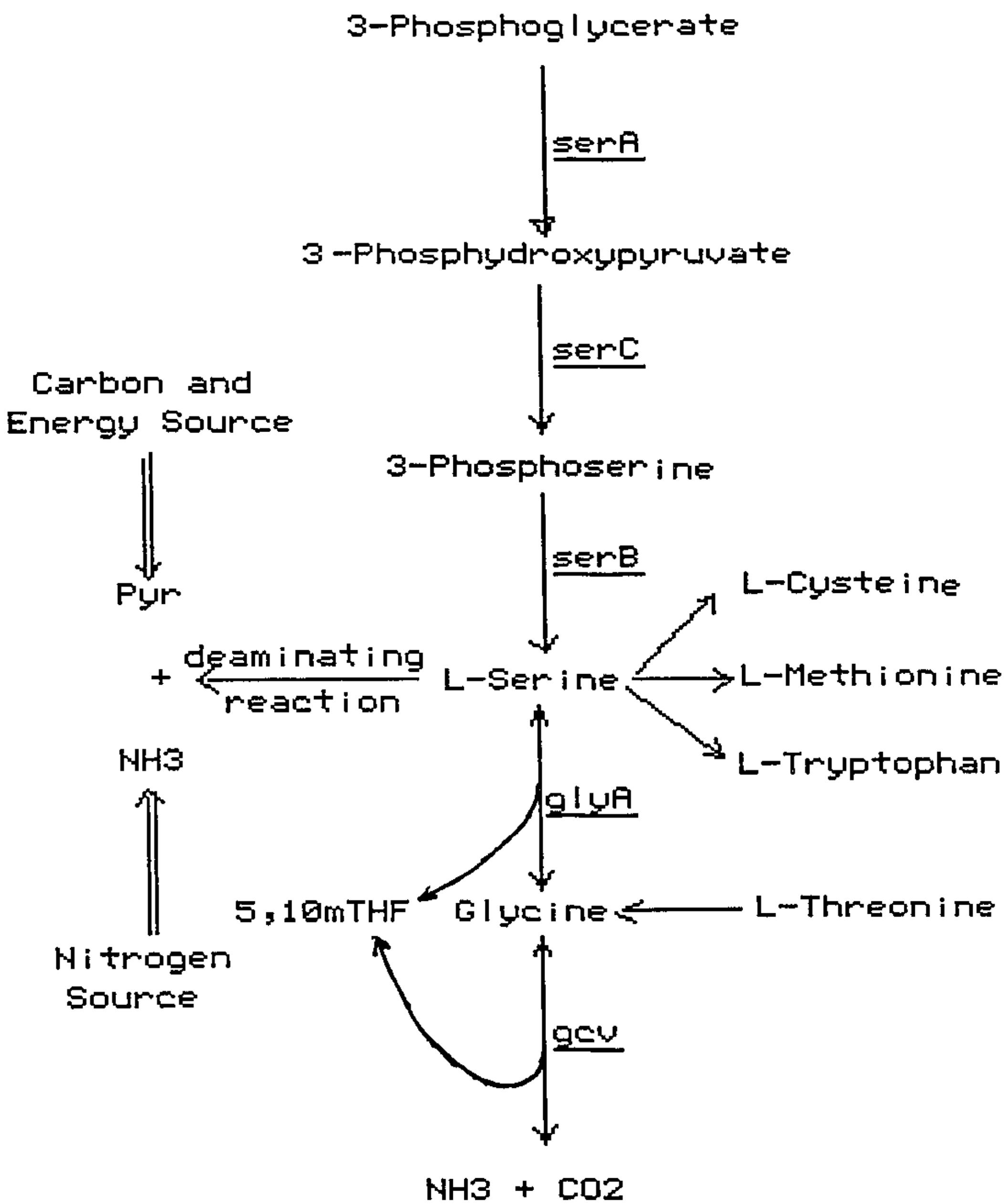

Fig. 1 The metabolic pathway of L-serine

(Major pathway is adapted from Stauffer, 1987) from *serA* varies considerably with the growth rate and carbon sources supporting growth [Pizer and Potochny, 1964].

Moreover, *serA* expression is under the control of the *lrp* gene product [Lin *et al.*, 1990]. This is organized in such a way that when the *lrp* gene product increases L-serine degradation, it decreases L-serine biosynthesis. Thus β-galactosidase synthesis from a *serA-lacZ* fusion was shown to be regulated by L-leucine and the expression of *serA* was decreased in an *lrp* mutant [Lin *et al.*, 1990].

4-2. The Relation between the L-Serine and Pyridoxine Biosynthetic Pathways

Winkler and his co-workers proposed a biosynthetic pathway for pyridoxine which parallels the L-serine biosynthetic pathway [Lam and Winkler, 1990]. In this pathway, the *pdxB* gene product converts erythronate-4-phosphate to 3-hydroxy-4-phospho-hydroxy-α-ketobutyrate. The *pdxB* gene shows homology to *serA* and the two pathways share *serC* enzyme function [Schoenlein *et al.*, 1989].

4-3. Other Pathways for the Synthesis of L-Serine

The biosynthesis from phosphoglycerate has generally been taken to the only pathway used during growth in glucose-minimal medium. This conclusion is based on the fact that mutants in which this pathway does not function require L-serine for growth [Umbarger *et al.*, 1963]. Nonetheless other potential sources of L-serine do exist and are reviewed in the following sections.

4-3a. Formation of L-serine from glycine

The production of L-serine from glycine is catalyzed by a much studied enzyme, L-serine hydroxymethyltransferase (SHMT)[Scrimgeour and Huennekens, 1962], the product of the *glyA* gene [Stauffer *et al.*, 1981]. The other substrate of this reaction is 5,10-methylenetetrahydrofolate (hTHF) which can be synthesized by adding to THF, C1 units produced from the various sources, including glycine itself by the glycine cleavage enzymes [Mudd and Cantoni, 1964].

Though these reactions can provide L-serine when exogenous glycine is provided, endogenous glycine does not replace the L-serine requirement of mutants in *serA,B* or *C*. That is, the wild-type *E. coli* does not have an alternative way to make glycine [Pizer, 1963], although some mutants do (see below).

The regulation of *glyA* is very complicated, because this reaction is a major source of glycine and C1 units. It is therefore understandable that the products of C1 metabolism, L-serine, glycine, L-methionine, purines and thymine would repress enzyme synthesis [Stauffer, 1987].

4-3b. Production of glycine from L-threonine

E. coli can derive its glycine from L-threonine via a cryptic pathway which can be established by a series of mutations as described by Fraser and Newman [Fraser and Newman, 1975]. L-Threonine can be converted to glycine by a two-step pathway, involving threonine dehydrogenase (TDH) and α-amino-β-ketobutyrate (AKB) ligase [Chan and Newman, 1981]. Then the glycine can be converted to L-serine as described above. This pathway is not used in wild-type *E. coli*, unless it is provided with exogenous L-threonine and L-leucine. The first enzyme, TDH, has been shown to be induced by L-leucine and to be regulated by *lrp* [Lin, *et al.*, 1990]. Whether the second enzyme is similarly regulated has not yet been determined.

PART 5. L-SERINE AS A BIOSYNTHETIC PRECURSOR

A large number of cellular metabolites are synthesized from L-serine (Fig. 1). Some of these use 2 or 3 of the L-serine carbons, and some use only the β carbon.

The side chain of L-tryptophan is made by condensing a complete L-serine molecule with indole glycerolphosphate [Yanofsky, 1960]. Similarly, all three carbons of cysteine are made from L-serine [Kredich and Tomkins, 1966]. Since L-cysteine is a precursor of L-methionine [Tran *et al.*, 1983], L-serine is also a direct precursor of the L-methionine carbon skeleton.

L-Serine has also been shown to be the only significant precursor of glycine in cells grown in glucose-minimal medium [Newman and Magasanik, 1963]. Each purine molecule is made from a glycine, and hence a serine molecule. Thus the L-serine biosynthetic pathway must supply enough product to meet the cells requirements for L-serine, glycine, L-tryptophan, L-cysteine, L-methionine and purines.

In cells grown in glucose minimal medium L-serine is also the source of the cell's C1 units, either directly via glycine synthesis using SHMT and forming h-THF, or indirectly by glycine cleavage [Newman and Magasanik, 1963]. The balance between these sources of C1 has been considered in detail [Newman and Magasanik, 1963].

It is clear then that the L-serine biosynthetic pathway must be extensively used during growth, and must divert a substantial amount of carbon away from the Embden Meyerhof pathway.

PART 6. OTHER ENZYMES CAPABLE OF DEAMINATING L-SERINE

This thesis is devoted to two enzymes which deaminate L-serine and are called L-serine deaminases. At least two other enzymes also deaminate L-serine, but are known as L-threonine deaminases: the biosynthetic L-threonine deaminase coded by the *ilvA* gene (*ilvA*-TD) [Calhoun *et al.*, 1973], and the biodegradative threonine deaminase coded by the *tdcB* gene (tdcB-TD) [Egan and Phillips, 1977, Goss and Datta, 1984]. The molecular characteristics of these enzymes, the genes that code for them, and the regulatory mechanisms controlling their expression have been studied in considerable detail [Goss *et al.*, 1988].

The *ilvA*-TD is known to function in the cell as a threonine deaminase because of its role catalyzing the first step in isoleucine metabolism, that is, the deamination of L-threonine to form α-ketobutyrate. Strains deficient in the *ilvA*-TD cannot deaminate L-threonine *in vivo* and require L-isoleucine for growth [Umbarger, 1987]. This enzyme is known to convert L-serine to pyruvate, at least *in vitro*, and its presence in cell extracts could confuse our L-SD assays. However this is not in fact a problem because the strains used in this laboratory all are derivatives of strain CU1008, an *ilvA* deletion mutant, and do not produce the *ilvA*-TD and require L-isoleucine for growth.

The *tdcB*-TD also deaminates L-threonine, but the enzyme is not normally synthesized in the media and growth conditions used in this work. *TdcB*-TD synthesis is normally induced under anaerobic culture conditions in tryptone-yeast extract medium lacking fermentable carbohydrates. *TdcB* gene coded for a polypeptide of 329 amino acid residues [Schweizer and Datta, 1989]. The other gene in this operon, the *tdcC* gene, seems to code for the permease for L-threonine and L-serine [Goss *et al.*, 1988 and Sumantran *et al.*, 1990].

The function of *tdcB*-TD as an anaerobic degradative enzyme made in rich medium, and the location of the gene coding for it in an operon that codes for a permease working on both amino acids, suggest that the enzyme may equally well be considered a L-serine deaminase as a L-threonine deaminase.

This conception of the *tdcB*-TD degrading both L-serine and L-threonine *in vivo* is further suggested by the fact that there is extensive peptide sequence homology between biodegradative L-threonine dehydratase, D-serine deaminase and biosynthetic L-threonine dehydratase of *E.coli* and biosynthetic L-threonine dehydrate of yeast [Datta *et al.*, 1987]. The peptide sequence of rat liver serine dehydratase [Ogawa *et al.*, 1989a], and human liver serine dehydratase also showed extensive sequence homology with the *tdcB*-TD [Ogawa *et al.*, 1989b], despite their diverse origins and metabolic significance. These mammalian L-SDs are very stable enzymes, and probably very different from the bacterial L-SD [Newman and Kapoor, 1980]. In any case, all these enzymes show relatively little discrimination between L-serine and L-threonine.

Whether the bacterial L-SD can deaminate L-threonine is not known since it has never been pure enough to test. We do however know that an increase in L-SD cannot compensate for the *ilvA* mutation, so it is unlikely that this enzyme has much activity against L-threonine.

PART 7. THE SOS AND HEAT SHOCK RESPONSES OF *E. coli*

L-SD activity in *E. coli* is induced by DNA-damaging agents and by heat shock [Newman *et al.*, 1982a]. Synthesis of many proteins is induced by DNA damage, these syntheses being coordinated by a global regulatory mechanism known as the SOS response [Walker, 1987]. Similarly, induction by an increase in temperature is coordinated by the heat shock response [Neidhardt and VanBogelen, 1987]. Since L-SD synthesis may be regulated by either of these, or by both, they are reviewed in the following sections.

7-1. The SOS Response

Exposure of *E. coli* cells to agents or conditions that damage DNA or interfere with DNA replication results in the induction of a diverse set of physiological events called the SOS response. The SOS response coordinates the expression of the more than 17 genes of the SOS regulatory network [Little and Mount, 1982], all of which are involved in DNA repair and mutagenesis.

The SOS response is organized by a complex interaction between the *recA* and *lexA* gene products. In the uninduced condition, the LexA protein binds to the upstream sequence of a group of damage-inducible (*din*) genes that share a similar sequence of which the consensus sequence is taCTGTatata-a-aCAGta. Binding of the *lexA* gene product prevents or reduces expression of the *din* genes. When DNA damage occurs, the *recA* gene product, a protease, is activated, resulting in the proteolytic cleavage of the *lexA* gene product, relieving repression and thus turning on the synthesis of products of the *din* genes, including *recA* itself [Walker, 1987, Walker, 1984].

The *recA* protease also acts on the repressor molecules which maintain lysogeny of the lysogenic phages, such as the λcI product. The proteolytic cleavage of these repressors initiates a cascade of inductions, both of prophage and bacterial genes. The prophage may then be able to enter lytic growth and escape its damaged host, and the bacteria can synthesize the molecules necessary for the repair of the damage. It has also been shown that the λ repressor is a much poorer substrate for the *recA* protease than *lexA*, so that λ prophage induction may not occur unless the cell is severely damaged by UV or other inducing treatments [Little *et al.*, 1981].

7-2. The Heat Shock Response

Using two-dimensional gels, it has been possible to define a set of proteins which increase in amount when cells are shifted from low to high temperature, e.g. from 37°C to 42°C [Herendeen *et al.*, 1979]. These are the heat shock proteins. Because L-SD activity in *E. coli* is induced by growth in high temperature, and the mechanism of this induction is unclear, I review here the heat shock response as a possible regulator of L-SD synthesis.

The heat shock regulon coordinates synthesis of at least 17 proteins in *E. coli* [Neidhardt and VanBogelen, 1987]. Expression of the genes coding for these proteins is regulated by an alternative RNA polymerase sigma factor, σ32, coded by *htpR* gene [Neidhardt and VanBogelen, 1981]. The *htpR* gene product is not produced in normal growth conditions, but is induced by sudden increases in temperature, i.e. heat shock. This sigma 32 protein binds to a consensus sequence (T tC CcCTTGAA-13-15 bp-CCCATtTA) upstream of the heat shock genes, and increases their expression.

One would expect to find such a sequence upstream of *sdaA*, if it is controlled as part of this regulon. However not all heat shock-induced genes show such a sequence. Clark and Neidhardt [Clark and Neidhardt, 1990] recently reported that the *lysU* gene, coding for lysyl-tRNA synthetase, forms part of the heat-induced system as judged by the fact that two-dimensional gels demonstrate an increase in the *lysU* gene product after temperature increase [Hirshfield *et al.*, 1981] in the parent strain but not in a mutant deficient in sigma 32 [Neidhardt and VanBogelen, 1981]. Nonetheless, on sequencing the *lysU* gene, they could not find the consensus sequence for sigma 32 binding site. They suggest that there might be another heat shock regulation system other than the *htpR*-controlled one and that this may be involved in induction of *lysU* after a temperature shift.

The existence of heat shock factors, other than sigma factor σ^{32}, was also suggested by Gross and her co-worker [Zhou *et al.*, 1988]. They constructed loss-of-function insertion and deletion mutations in the *htpR* gene and found that though there was no transcription from most heat-shock promoters, several heat shock proteins were still produced in the mutants. Synthesis of one of these proteins, DnaK protein, was abolished at low temperature and could be detected after a shift to high temperature. They suggested that there are additional mechanisms controlling the synthesis of some heat shock proteins.

Another interesting aspect of the heat shock response is that the proteins of the heat shock regulon of *E. coli* are induced when cells are grown in the presence of ethanol and that this response of the regulon is also controlled by the product of the *htpR* gene. VanBogelen *et al* [VanBogelen *et al.*, 1987] have showed that addition of 4% ethanol to the cells growing at 28°C had an effect similar to a shift to 42°C. Growth slowed, and HTP protein was induced. With 10% ethanol the induction was even greater and was comparable to that produced by a shift from 28°C to 50°C. Heat and ethanol apparently affect a common cellular process or component which, acting through the *htpR* gene product, induced the same set of polypeptides even under conditions where net growth and expression of most cellular genes were inhibited. The fact that ethanol also induces L-SD increased our interest in this regulon.

MATERIALS AND METHODS

PART 1. STRAINS, BACTERIOPHAGES, AND PLASMIDS

The strains, bacteriophages, and plasmids used in this study are listed in Table 1.

PART 2. CULTURES, MEDIA, AND GROWTH TESTS

2-1. Minimal Medium:

The minimal medium used, neutralized to pH 7, has been described elsewhere [Newman *et al.*, 1985a]. Because all derivatives of strain MEW1 carry a deletion in *ilvA* and therefore require L-isoleucine for growth, L-isoleucine and L-valine were added to all media at 50 μg/ml each.

2-2. SGL Medium:

Medium with a combination of L-serine, glycine and L-leucine as the only carbon source other than L-isoleucine and L-valine is called SGL medium. L-Serine, glycine and L-leucine were usually provided at 2,000, 300, and 300 μg/ml, respectively (unless otherwise noted).

2-3. Determination of Doubling Times

The doubling times of cultures were calculated from growth curves determined by

Table 1. Strains, Bacteriophage and Plasmids

Strain, phage or plasmid	Genotype and relevant Characteristics	Source or reference
E. coli K-12		
CU1008	*E. coli* K-12 *ilvA*	L.S. Williams
MEW1	*lacZ* derivative of strain CU1008	Newman *et al.*,1985b
MEW21	MEW1::λplacMu9 SGL⁻ Kan^r isolated by insertion of λplacMu into MEW1	This study
MEW22	MEW1::λplacMu9 SGL⁻ Kan^r isolated by transduction from MEW21 to MEW1 and selecting for Kan^r	This study
MEW22 Kan^s	SGL⁻ Kan^s UV survivor of MEW22	J. Garnon
MEW26	MEW1 *lrp*::Tn10	Lin *et al.*, 1990
MEW28	MEW1 *sdaA*::Cm^r	This study
MEW49	MEW28 *sdaA*::Cm^r	This study
MEW50	MEW28 *sdaX*; SGL⁻utilizing	This study
MEW51	MEW28 *sdaX*; *sdaB*::λplacMu9	This study
MEW52	MEW50 *lrp*::Tn10	This study
MEW53	MEW51 *lrp*::Tn10	This study
MEW55	MEW51 *sdaX*; *sdaB*::Tn10	This study

Cont.

MEW56	MEW128 *sdaA*::Cm^r by transduction from MEW28 into MEW128	This study
MEW57	MEW191C *sdaA*::Cm^r by transduction from MEW28 into MEW191C	This study
MEW58	MEW84 *sdaA*::Cm^r by transduction from MEW28 into MEW84	This study
MEW59	MEW1 *recA*::Tn10 by transduction from NK6042 into MEW1	This study
MEW60	MEW28 *recA*::Tn10 by transduction from NK6042 into MEW28	This study
MEW83	MEW1::λplacMu9 SGL- Kan^r isolated by insertion of λ placMu9 into MEW1	This study
MEW84	MEW1::λplacMu9 SGL- Kan^r isolated by transduction from MEW83 into MEW1 and selecting for Kan^r	This study
MEW128	Mutant SGL$^-$, unable to grow with serine, glycine, L-leucine; isolated by penicillin selection from strain CU1008	Newman *et al.*,1985b
MEW191	SGL$^-$ derivative by Mu::dX insertion from CAG5050 into MEW1	Newman *et al.*, 1985b
MEW191C	MEW191 SGL$^-$ cured of MudX	Dumont, 1985

Cont.

KEC9	*ssd*	Newman *et al.*, 1982b
A401	HfrC *polA1*	Russel and Holmgrem, 1988
NK6042R	NK6042 *recA*::Tn10	Basso, J.
MH2923	F+ Mu *cts* 62 *h*p1-1 *araD* (Mud5005)	Groisman and Casadaban,1986
XL1	*recA*⁻ (*recA1*, *lac*⁻ *endA1*, *gyrA96*, *thi*, *hsdR17 supE44*, *relA1*, {F' *proAB*, *lacI*^{Q} *lacZ* M15, Tn10}	Stratagene Co.
CAG5051	HfrH *nadA*::Tn10	Singer *et al.*, 1989
CAG5052	KL227 *btuB*3139::Tn10	Singer *et al.*, 1989
CAG5053	KL208 *zbc*-280::Tn10	Singer *et al.*, 1989
CAG5054	KL96 *trpB*::83::Tn10	Singer *et al.*, 1989
CAG5055	KL16 *zed*-3069::Tn10	Singer *et al.*, 1989
CAG8209	KL228 *zgh*-3057::Tn10	Singer *et al.*, 1989
CAG8160	KL14 *thi*-39::Tn10	Singer *et al.*, 1989
CAG18486	MG1655 *eda*-51::Tn10 40.75 min	Singer *et al.*, 1989
CAG12156	MG1655 *uvrC*279::Tn10 42.25 min.	Singer *et al.*, 1989
CAG12173	MG1655 *cysC*95::Tn10 59.25 min.	Singer *et al.*, 1989
CAG12079	MG1655 *fuc*-3072::Tn10 60.25 min.	Singer *et al.*, 1989
Phages		
M13K07		Vieira and Messing, 1987
λplacMu9	λ placMu1 Km^{r}	Bremer *et al.*, 1985

Cont.

λplacMu507	λ cI ts857 *Sam*7 Mu*A*+*B*+(Helper phage)	Bremer *et al.*, 1985
λTn10	λcts *Sam*S53	Wood, 1981
λ10B5	λKohara phage456	Kohara *et al.*, 1987
λ8C5	λKohara phage457	Kohara *et al.*, 1987
λ9A12	λKohara phage458	Kohara *et al.*, 1987
Plasmids		
pMES1	Plasmid able to complement *sdaA* isolated from strain MEW22 infected with MH2923 lysate	This study
pMES2	pBR322 with a 6-kb *Pst*I insert from pMES1	This study
pMES3	Bluescript (-) vector with 2.6 kb *Pst*I-*Sal*I fragment from pMES2	This study
pMES4	Bluescript (+) vector with 2.6 kb *Pst*I-*Sal*I fragment from pMES2	This study
pMES22	pBR322 with the 2.6 kb *Xho*I-*Bam*HI insert from pMES3	This study
pMEZS22	pBR322 carrying in fram fusion of transcated *sdaA-lacZ* gene	This study
pMES23	pMES22 with a 1.4-kb *cat* gene inserted at the *Hpa*I site of *sdaA* gene	This study

Cont.

pMES25	pMES4 with a new *Eco*RI site in the stop codon of the *sdaA* gene, constructed by site-directed mutagenesis	This study
pMES26	Bluescript(+) containing 2 kb *SalI*-*Eco*RI fragment	This study
pMES27	The inframe fusion of *sdaA* gene at the *Eco*RI site to the *Sma*I site in pMC1871	This study
pMES28	The plasmid pSD100 containing the *sdaA* gene inframe fusion	This study
pMES40	Plasmid conferring ability to grow on SGL medium; isolated from strain MEW55 infected with miniMu MEW50 lysate	This study
pMES41	pBR322 with a 8-kb *Pst*I insert from pMES40	This study
pBR322		Bolivar *et al.*, 1977
Bluescript	KS+ and KS-	Stratagene Co.
pMC1871	*lacZ* carried on pBR322	R.K. Storms
pACYC184		Chang and Cohen *et al.*,1978
pSD100		Moskaluk and Bastia, 1988

measuring turbidity with sidearm flasks fitted for a Klett colorimeter and using a red (420) filter. To do this, overnight cultures were subcultured, allowed to grow to exponential phase and diluted into the medium in which growth was tested. Turbidity was determined every 30 minutes.

2-4. Determination of Nutritional Requirements

To determine whether L-serine and/or glycine could be used as nitrogen source, precultures were grown with a reduced amount of ammonium sulfate (500 μg/ml) and then diluted 4,000 times into test medium so as to minimize the possibility of carryover of ammonium sulfate. Growth was followed as indicated above.

2-5. Other Additions to the Medium

Antibiotics were used at the following concentrations, in μg/ml: Ampicillin (Amp) 100, tetracycline (Tet) 25, kanamycin (Kan) 80, chloramphenicol (Chl) 40 and streptomycin (Str) 100.

PART 3. ENZYME ASSAYS

3-1. L-Serine Deaminase

L-SD was assayed as described previously in toluene-treated whole cells and in crude extracts [Newman *et al.*, 1985a], using 35 and 40 min incubations respectively. One unit of L-SD as measured in the whole cell assay is defined as the amount of enzyme which catalyzed the formation of 1 μmol of pyruvate in 35 min. in whole cell assay conditions. One unit of L-SD as measured in extracts is defined as the amount of enzyme which produced 1 μmol pyruvate per minute per mg of protein.

Assays of L-SD in LB-grown cells present the problem that results could be confused by deamination of L-serine due to another known enzyme, the biodegradative L-threonine deaminase [Goss and Datta, 1984]. To avoid this, LB cultures were grown with increased (0.5%) glucose so as to repress formation of that enzyme.

3-2. β-Galactosidase

β-Galactosidase was assayed by the method of Miller [Miller, 1972]. One unit of β-galactosidase as expressed in Miller units is the amount of enzyme which produces 1 mμ-mole/ml *o*-nitrophenol/min. in standard assay conditions at 28°C, pH 7. Specific activity is expressed here as units/mg protein or units/μl cell extracts.

3-3. Protein Assay

Protein assay was carried out by the method of Lowry assay [Lowry *et al.*, 1951] and coomassie blue protein assay [Stoscheck,1990] as indicated in each experiment.

PART 4. STRAIN CONSTRUCTIONS

4-1. Use of Mini-Mu Constructs to Isolate L-SD-deficient Strains Carrying Protein Fusions

Principle: λplacMu9(kan^r) [Bremer *et al.*, 1985] was inserted into the *E. coli* genome and kanamycin-resistant cells unable to grow in SGL were selected.

Detail: After infection of strains with λplacMu9 and λpMu507, as described by Weinstock and his co-workers [Bremer *et al.*, 1984, Bremer *et al.*, 1985], cells were subcultured in glucose-minimal medium and then incubated in SGL medium for 2 to 3 h, after which ampicillin was added and the incubation was continued for 3 h to kill cells able to grow in this medium. Cultures were then plated on LB containing kanamycin (80 μg/ml). The resulting colonies were then screened for their ability to grow in SGL medium. SGL$^-$ kan^r colonies were selected, including strain MEW22.

4-2. Isolation of a Kanamycin-sensitive Derivative of Strain MEW22

Principle: The fact that the insert in strain MEW22 conferred kanamycin-resistance interfered with plans for further constructs. The following protocol allowed isolation of a strain in which the SGL$^-$ character could be maintained, and the antibiotic resistance lost.

Detail: To isolate an SGL$^-$ kanamycin-sensitive derivative of strain MEW22, the strain was irradiated with UV and plated on LB. Colonies were replicated on LB-kanamycin plates. Antibiotic-sensitive colonies unable to grow in SGL medium were chosen for study.

4-3. Construction of an MEW1 Derivative Carrying a Chromosomal *sdaA*::Cm^r Null Mutation

Principle: A plasmid carrying an ***sdaA*::Cm^r** null mutation was transformed into a strain in which it cannot multiply. The strain in which the mutated gene was integrated in the genome was selected by chloramphenicol-resistance-all according to the method of Winans *et al* [Winans *et al.*, 1985].

Detail: Plasmid pMES23 (carrying an ***sdaA*::Cm^r** the construction of which is described later) was transformed into strain A401 [Russel and Holmgrem, 1988] and chloramphenicol-resistant colonies selected. Chloramphenicol-resistance was then transduced from such a colony into strain MEW1. The fact that the insert was indeed in *sdaA* was confirmed by showing that the strain created, strain MEW28, was both chloramphenicol-resistant and unable to grow on SGL plates.

4-4. Construction of a Kanamycin-sensitive Derivative of Strain MEW51

Principle: An *sdaB* insert conferring kanamycin-resistance was replaced by one conferring tetracycline-resistance from λTn10 [Wood, 1981], in order to allow further constructions requiring a kanamycin-sensitive host.

Detail: Strain MEW51 *sdaA*::Cm^r *sdaB*::λplacMu9 was infected with λTn10, and tetracycline-resistant colonies selected. 40 tet^r colonies were screened for kanamycin-

Fig. 2 Construction of strain MEW28 *sdaA*::Cm^r

A: To insert a chloramphenicol resistance gene into the coding region of the chromosomal *sdaA* gene, plasmid pMES22, carrying *sdaA* on a 2.6-kb fragemnt inserted into pBR322, was digested with *Hpa*I at base 1456, i.e. inside the *sdaA* coding sequence. Digestion of pACYC184 with *Hae*II produced a 1.4-kb fragemnt carrying the *cat* gene. This fragment was blunted, filled with the Klenow fragment and dNTP and ligated into the pMES22 *Hpa*I site. The resultant plasmid was transformed into MEW22, selecting Amp and Tet resistance. The plasmids from the antibiotic-resistant colonies were isolated and their size and digestion pattern checked. One of these plasmids, 8-kb in size, did not produce any L-SD activity, and was named plasmid pMES23. B: This plasmid was transformed into strain A401 and chloramphenicol-resistant colonies were selected. C: Chloramphenicol resistance was transduced from such a colony into strain MEW1. A chloramphenicol resistant SGL$^-$ strain named MEW28 was used in this study.

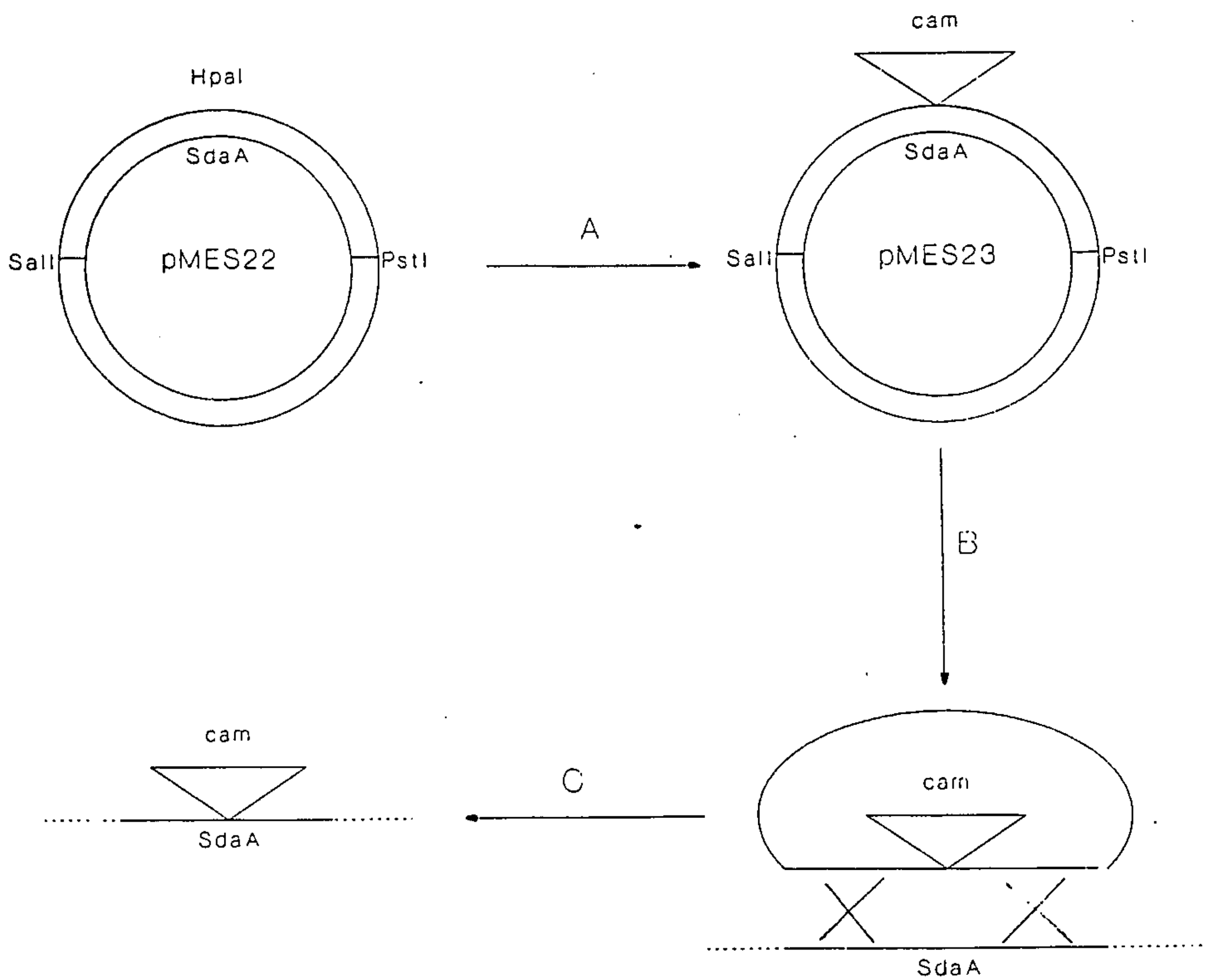

Fig. 2 Construction of strain MEW28 *sdaA*::Cm^r

sensitive SGL⁻ derivatives at 28°C and 20 of them were SGL⁻,kan^s^. One of these was named MEW55.

4-5. Construction of *lrp* Strains

Lrp strains were constructed by transduction from strain MEW26 *lrp*::Tn10 [Lin *et al.,* 1990] by selecting for tetracycline resistance.

4-6. Construction of *sdaA*::Cm^r Strains

SdaA::Cm^r strains were constructed by transduction from strain MEW28 (described in Part 2, 4-3) by selecting for chloramphenicol-resistance.

4-7. Construction of *recA*::Tn10 Strains

RecA::Tn10 strains were constructed by transduction from strain NK6042 *recA*::Tn10 (obtained from J. Basso) by selecting for tetracycline-resistance.

PART 5. CLONING OF *SdaA* AND *SdaB* FROM A MU REPLICON

The *in vivo* cloning system used here was developed by Groisman and his co-workers [Groisman *et al.*, 1984, Groisman and Casadaban, 1986], and the experiments carried out exactly as prescribed in their protocol, and detailed below.

5-1. Cloning of *SdaA* and Construction of pMES22

Principle: A miniMu replicon carrying sdaA was selected by complementing an *sdaA* deficient strain. The clone was digested with restriction enzymes, and subclones suitable for sequencing were selected.

Detail: The Mu replicon was produced by incubating strain MH2923 *sdaA* kan^r at 42°C to induce transposition and bacteriophage replication. The lysate thus formed was used to transduce the ability to grow in SGL medium in the presence of kanamycin into MEW22 Kan^s, a *sdaA* mutant kanamycin-sensitive strain. Plasmid pMES1 was isolated from a clone growing in SGL medium, digested with *Pst*I and subcloned into the pBR322 *Pst*I site, forming pMES2 (Fig. 3). Next, pMES2 was digested with *Pst*I and *Sal*I and subcloned into Bluescript KS+ and KS- vectors, forming pMES3(+) and pMES4(-). Then pMES4 was digested with *Xho*I and *Bam*HI and subcloned into pBR322 digested with *Sal*I and *Bam*HI, forming pMES22. SGL$^+$(NSIV$^+$) clones were selected at each step.

5-2. Cloning of *SdaB* from MEW51 Mucts Mini-Mu Lysogens

Principle: The method of Groisman and Casadaban uses a particular set of *E. coli* strains as host. In order to clone genes from another host, e.g. to clone the mutated form of a gene, using the same method, the mini-Mu replicon and Muts phage must be transferred into that host.

Detail: MEW50 (*sdaA*::Cm^r *sdaX*) Mucts-miniMu lysogens were constructed by infecting MEW50 with MH2923 lysate and selecting lysogens. A lysate of the MEW50 Mucts-miniMu lysogen was used to transfect MEW55(*sdaA*::Cm^r *sdaX* *sdaB*::λTn10) and SGL$^+$ kanamycin-resistant colonies were selected. Plasmid pMES40 isolated from one such colony, was digested with *Pst*I and subcloned into pBR322 *Pst*I site, forming pMES41.

Fig. 3 Cloning of *sdaA* gene

A Mu lysate was produced from strain MH2923 and transfected to strain MEW22 Kan^s, selecting SGL^+ colonies. Plasmid pMES1 was isolated from one of these colonies. Plasmid pMES2 was produced by subcloning a 6-kb *Pst*I fragment from plasmid pMES1 into pBR322. A 2.6-kb *Sal*I-*Pst*I fragment from pMES2 was subcloned into Bluescript (+) and (-) *Sal*I, *Pst*I sites, generating the plasmids pMES3 and pMES4. Then a 2.6-kb *Xho*I-*Bam*HI fragment from plasmid pMES3 was inserted into *Sal*I, *Bam*HI sites of pBR322, producing plasmid pMES22.

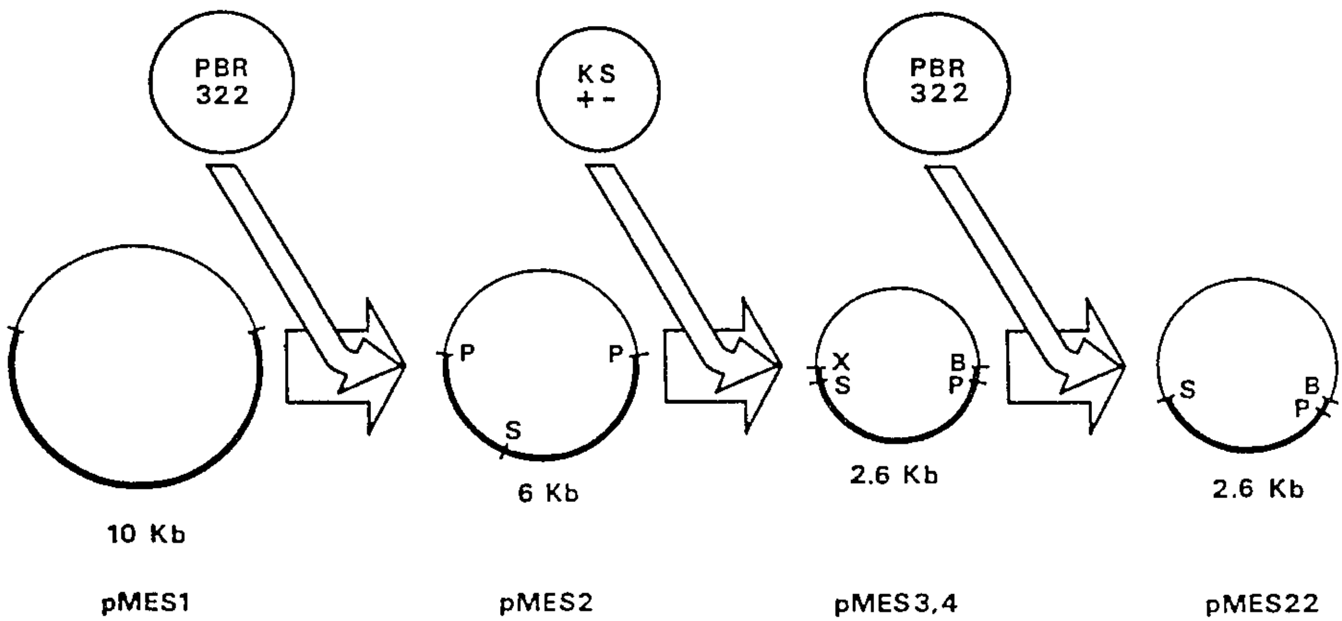

Fig. 3 Cloning of *sdaA* gene

Abbreviations: B, *Bam*HI; P, *Pst*I; S, *Sal*I; X, *Xho*I. The *E. coli* DNA inserted is represented by the heavy line, and the value under each circle (in kilobases) describes the size of the insert.

PART 6. HYBRIDIZATION

Hybridization was performed by the method of Southern [Maniatis *et al.*, 1982] with minor modifications, or using gels dried according to the method of R.K.Storms. The DNA fragments, isolated from agarose gels, were randomly oligo-labelled with ^{32}P-dATP by the procedure of the supplier (Boeringher-Mannheim).

PART 7. DNA SEQUENCING

Principle: Deletions were made in a 2.6 kb *Sal*I *Pst*I fragment from the *sdaA* clone, using the Erase-a-Base system. Single-stranded DNA was isolated from plasmids of appropriate size and the DNA sequenced with the Sequenase kit, as obtained from the United States Biochemical Corporation, P.O. Box 2240, Cleveland, Ohio 44122).

7-1. Production of Deletions with the Erase-A-Base System

Deletion experiments were carried out as described in the Erase-a-Base system purchased from Promega Biotec, Madison, Wisconsin. Plasmids pMES3 and pMES4 were double digested with *Xho*I, *Kpn*I and *Bam*HI, *Sac*I respectively. Four µg of digested DNA from each plasmid were incubated with *Exo*III nuclease for 10 different time periods (30 second intervals from 0 to 4 minutes; 5 and 6 minutes). The DNA was incubated with *S*I nuclease and then with Klenow enzyme and dNTP to generate blunt ends, and ligated with T4 DNA ligase. The resulting plasmids of mixed size from each deletion incubation were transformed into strain XL1. Plasmids were isolated from 3-5 transformants from each incubation period, and their size determined on agarose gels.

Plasmids were chosen so as to provide a series of sizes of insert DNA, and used for the isolation of single-stranded DNA.

7-2. Isolation of Single-Stranded DNA

The strains with plasmids of suitable size were used for making single stranded DNA, using the method of F. Lang (personal communication). The 100 ul of exponential phase cells carrying the appropriate plasmids were infected with 1 ul of helper phage M13KO7 (titre of 10^{13} pfu/ml) and incubated for 1 hr at 37°C with aeration by shaking. Then 2ml of 2X YT containing kanamycin (70 ug/ml) were added and the culture incubated for another 14-20 hours.

Then 1.5 ml of this culture was centrifuged in an Eppendorf centrifuge for 5 minutes. To 1.2 ml of the supernatant were added 300 μl of a solution of 20% polyethyleneglycol in ammonium acetate 3.5M pH7.5. This mixture was incubated at room temperature for 10 min, and centrifuged again. The resultant pellet was dissolved in 300 ul TES buffer (100 mM Tris pH7.5, 5 mM EDTA and 0.1% SDS). After addition of 100 μg proteinaseK, the mixture was incubated at 37°C for at least 1 hour to allow digestion of protein. The DNA was isolated by phenol and chloroform extraction followed by ethanol precipitation. The precipitated ssDNA was resuspended in 20μl TE buffer.

7-3. DNA Sequencing

DNA sequences were determined by the dideoxy-chain termination method of Sanger *et al.* [Sanger *et al.*, 1977], following the protocol indicated in the Sequenase kit from

United States Biochemical Corporation, Cleveland, Ohio. Both strands were sequenced by this method. Because part of the sequence was relatively G+C rich and contained numerous strong, secondary structures, the sequence of the plasmids deleted from pMES4 was further verified using reactions in which the dGTP was replaced by dITP.

7-4. Sequencing Gel

The DNA sequencing gel system used was developed by F. Lang [Lang and Burger, 1990].

PART 8. SITE-DIRECTED MUTAGENESIS

Principle: In order to obtain a DNA sequence coding for a fusion protein with L-SD at the N-terminal, it was necessary to alter the stop codon of *sdaA*. This could be done by changing TAA (stop codon) TAC to GAATTC, creating an *Eco*RI site at the same time. The oligonucleotide used was chosen so as to avoid complications due to secondary structure downstream of the stop codon.

Detail: The experiment was carried out using the Muta-Gene Phagemid *in vitro* Mutagenesis kit from Bio-Rad laboratories, and an oligonucleotide was synthesized and phosphorylated. The ssDNA from plasmid pMES4 was annealed with oligo-AAGAA*TTC*GTCACACTG, which corresponds to bases 1998-2016 of the *sdaA* sequence. A* and C* indicate the new bases which were introduced into the termination codon of *sdaA*. Plasmids from the resulting colonies were isolated, digested with *Eco*RI and analyzed on agarose gels. The presence of a new *Eco*RI fragment in digests of one of these plasmids, named pMES25, indicated that the stop codon had indeed been mutated as shown in Fig. 5.

Plasmid pMES26 was constructed from the mutated plasmid as detailed below. The appropriate region of pMES26 was sequenced to verify that the procedure had resulted in the intended sequence changes.

PART 9. PLASMID CONSTRUCTIONS

9-1. Construction of pMEZS22

To construct an in-frame fusion of *sdaA* to *lacZ* at the unique *Hpa*I site on plasmid pMES22, I digested pMC1871, a plasmid carrying *lacZ*, with *Sma*I and *Pst*I and isolated the 3.1-kb DNA band corresponding to *lacZ*. I then made a total digest of pMES22 with *Hpa*I and a partial digest with *Pst*I. I isolated the *Pst*I-*Hpa*I band corresponding to DNA from the *Hpa*I cut at base pair 1456 (see Fig. 11) to the *Pst*I site indicated in Fig. 3, a total of 5.4 kb. The 3.1 and 5.4 kb bands were ligated with T4 ligase (Fig. 4). The construction was verified by synthesizing the oligonucleotide TCT GCC CTG CGC CGG (corresponding to bases 1365-1380) and sequencing across the fusion junction.

9-2. Construction of pMES23

The plasmid pMES23 was generated by cloning a blunted 1.4 kb *Hae*II fragment containing the Cm^r gene, derived from pACYC184 [Chang and Cohen, 1978], into the

Fig. 4 Construction of plasmid pMEZS22

A: Plasmid pMES22 was digested with *Hpa*I totally and *Pst*I partially and a 5.4-kb DNA fragment, the size expected for DNA from *Hpa*I-*Sal*I-*Pst*I-*Pst*I as shown, was isolated from an agarose gel. B: Plasmid pMC1871 was digested with *Pst*I and *Sma*I and a 3.1-kb DNA fragment carrying the *lacZ* gene was also isolated from an agarose gel. C: These two DNA fragments were ligated and transformed into MEW1. Plasmids were isolated from resulting colonies which showed β-galactosidase activity. Their size (8.5-kb) and restriction digest pattern were checked. The junction of the *Sma*I/*Hpa*I site was further checked by DNA sequencing.

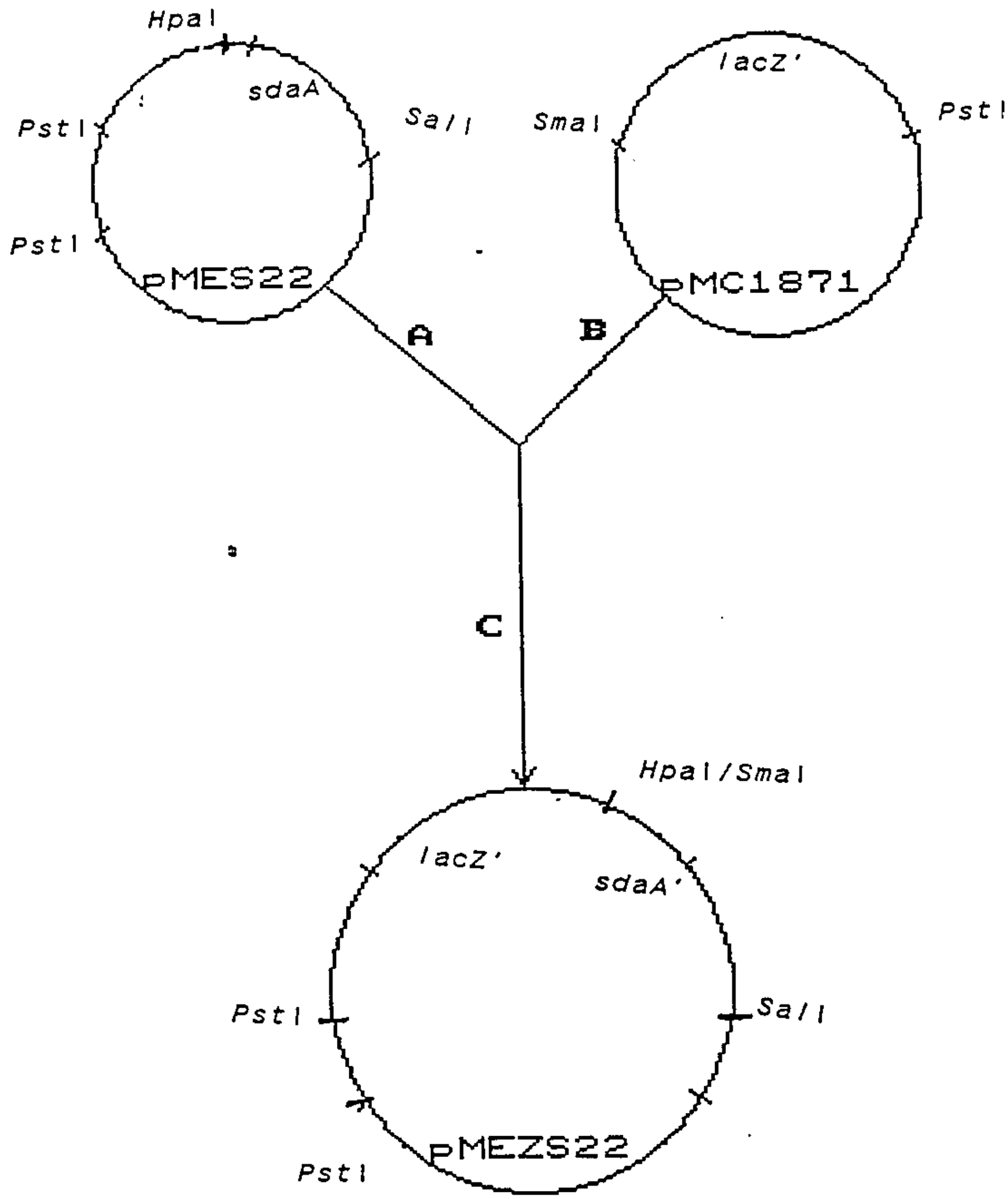

Fig. 4 Construction of plasmid pMEZS22

*Hpa*I site at base 1456 of the *sdaA* gene carried on pMES22 in a 2.6 kb fragment inserted into the *Bam*HI and *Sal*I site of pBR322 (Fig. 2).

9-3. Construction of *SdaA-LacZ* and *SdaA*-collagen-*LacZ* Fusion Plasmids

Plasmid pMES4 was mutated by site-direct mutagenesis as described in part 8 forming plasmid pMES25 (Fig. 5). The 2 kb *Sal*I-*Eco*RI fragment containing the *sdaA* gene from pMES25 was subcloned into the BS(+) *Sal*I, *Eco*RI site forming plasmid pMES26 (Fig. 6). This plasmid was fused to plasmid pMC1871 at a *Sma*I site as shown in Fig. 7, forming plasmid pMES27.

The *sdaA*-collagen-*lacZ* fusion plasmid pMES28 was constructed by isolating the 2 kb *Sal*I-*Bam*HI *SdaA* gene DNA fragment from plasmid pMES27, and a 0.3 kb *Sal*I-*Bam*HI fragment from pBR322 as a linker, ligating with pDS100 [Moskaluk and Bastia, 1988] which was digested with *Bam*HI. The *lacZ* gene from this plasmid remains under the control of λcI857. This construction is also shown in Fig. 7.

PART 10. PROTEIN PURIFICATION

10-1. Purification of the 2-Part Fusion Protein (2PF)

Strain MEW28 carrying plasmid pMES27 was grown in minimal medium with glycine and L-leucine at 37°C and harvested in late exponential-phase and resuspended in TMN (20mM TrisHCl, 10mM $MgCl_2$, 10mM β-mercaptoethanol and 1.6 M NaCl pH7.4) buffer [Ullmann, 1984]. The cell suspension was sonicated, and debris removed

Fig. 5 Scheme for generating an *Eco*RI site in the stop-codon of the *sdaA* gene

The 20 bp oligonucleotide shown in this Fig. was synthesized. 200 ng of this oligonucleotide were phosphorylated, and annealed to single stranded DNA produced by strain XL1 carrying plasmid pMES4 and the complementary strand was synthesized, all following the protocol in the instruction manual from BIO-RAD. The resultant mixture was transformed into XL1. Plasmids were isolated from 48 colonies and digested with *Eco*RI. Two of them which showed a new *Eco*RI site were further checked by double digestion with *Eco*RI and *Sal*I, which produced the expected 2-kb DNA fragment. One of these plasmids was named pMES25. The stop codon (TAA) and new *Eco*RI site (GAA TTC) generated are underlined.

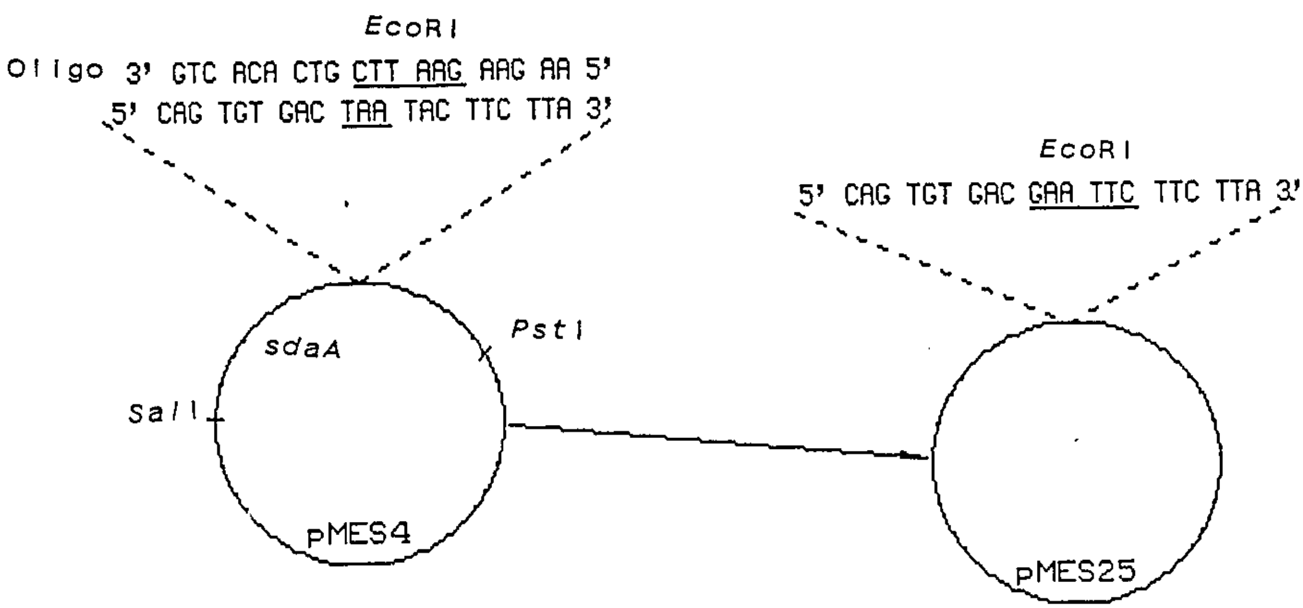

Fig. 5 Scheme for generating an *Eco*RI site in the stop-codon of *sdaA* gene

Fig. 6 Construction of plasmid pMES27

A: A 2-kb *Sal*I *Eco*RI fragment from plasmid pMES25 was isolated from an agarose gel and inserted into the *Eco*RI, *Sal*I sites of plasmid Bluescript(+) (3.0-kb), forming plasmid pMES26 (5.0-kb). The new *Eco*RI site was confirmed by DNA sequencing in the junction region and upstream of the *Eco*RI site in this plasmid. B: Plasmid pMES26 and pMC1871 (7.4-kb) were both digested with *Sma*I, ligated and transformed into strain MEW28. The plasmid size (12.4-kb) and restriction pattern of the plasmids isolated from these transformants were checked. One of these plasmids, 12.4-kb in size, allowed β-galactosidase activity synthesis, and was named plasmid pMES27. The reading-frame of both genes is indicated

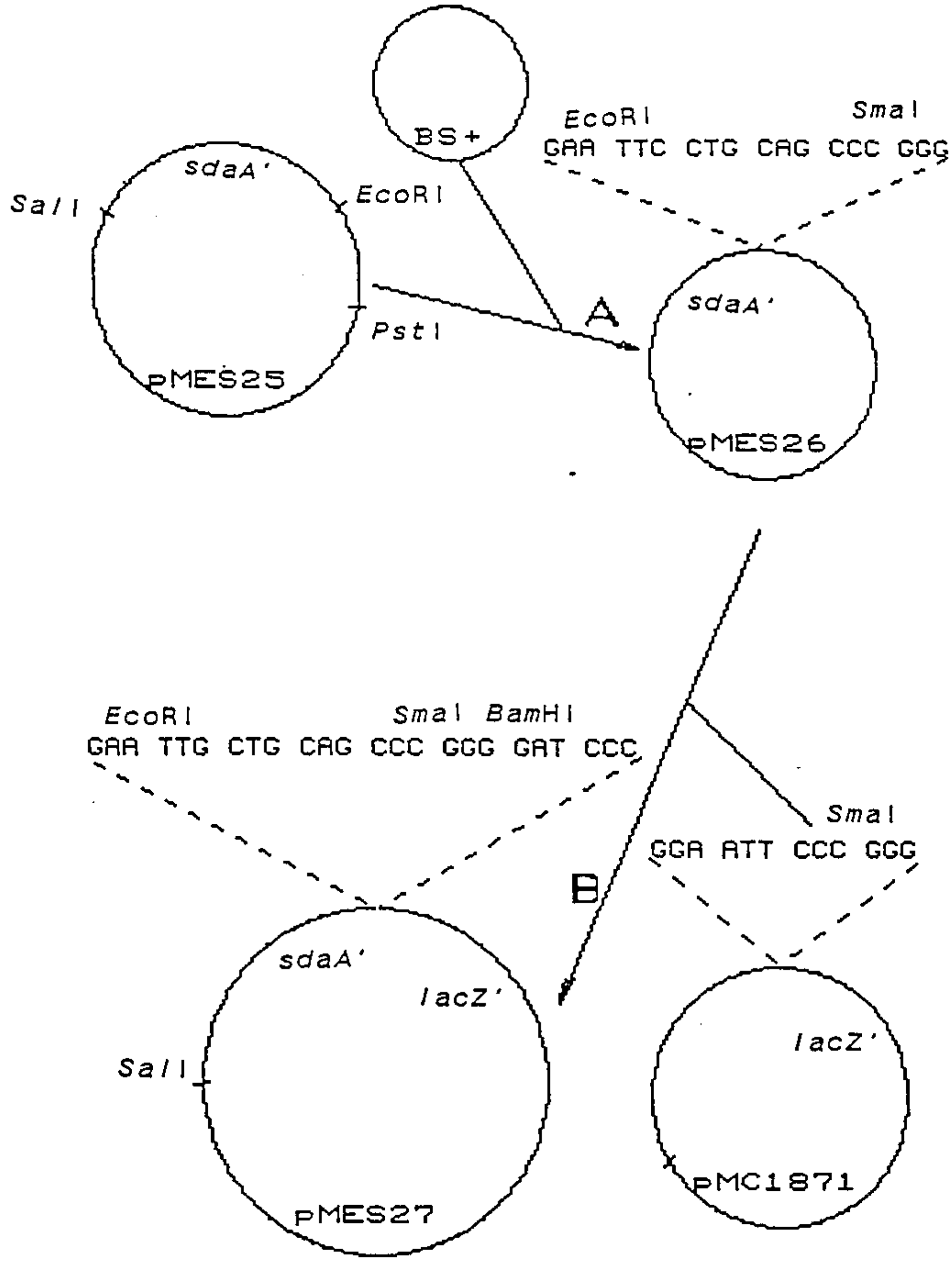

Fig. 6 Construction of plasmid pMES27

Fig. 7 Construction of plasmid pMES28

A 2-kb *Sal*I-*Bam*HI fragment from plasmid pMES27 and a 0.3 kb *Sal*I-*Bam*HI fragment from pBR322 were isolated from agarose gels and inserted into the *Bam*HI site of plasmid pSD100 (6.2-kb). The ligation mixture was transformed into strain MEW28 at 28°C and Amp resistant colonies were selected. An 8.5-kb plasmid isolated from one of these colonies showed the expected restriction site and was named pMES28. The reading-frame of these genes is indicated.

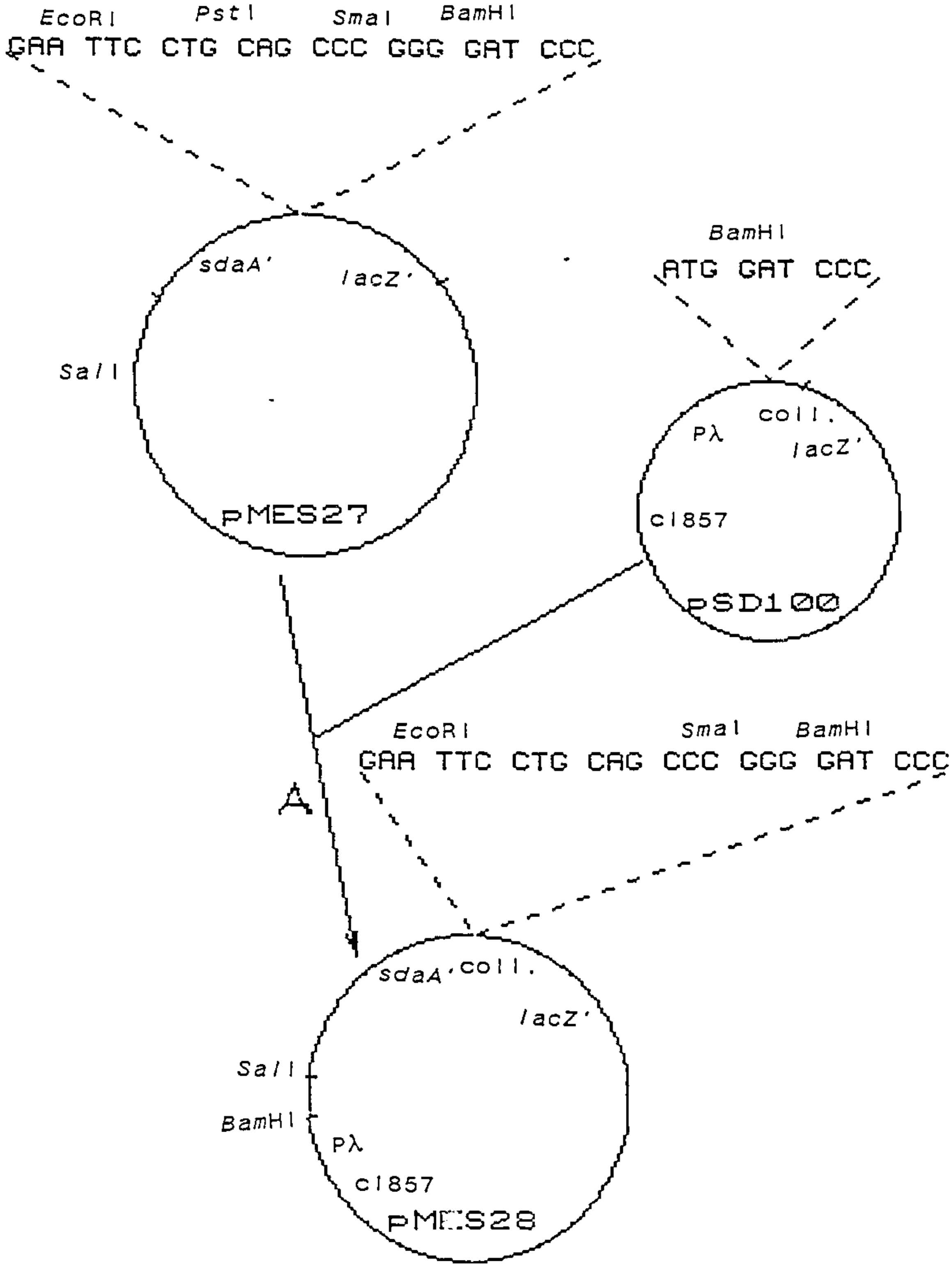

Fig. 7 Construction of plasmid pMES28

by centrifuging at 12,000 rpm. About 10^6 u of β-galactosidase were loaded onto a p-Aminobenzyl 1-thio-β-D-galactopyranoside-agarose column (Sigma A4010). The column was then washed with 100 times the bed volume of TMN buffer and eluted with TMN buffer containing 20% lactose. The eluted protein was concentrated in an Ultracent-30 Cartridge from Bio-Rad Lab.

10-2. Purification of 3-Part Fusion Protein (3PF)

Strain MEW28 carrying plasmid pMES28 was grown in minimal medium at 28°C. The cell culture in early exponential-phase was incubated at 42°C for 20 min and then at 37°C for 2-3 h. The extract was made and the protein was isolated as described for the 2-part fusion.

10-3. Purification of the *sdaA* Gene Product

The 3PF obtained as in 10-2 was treated with collagenase, purchased from Boehringer Mannheim. The treated preparation was loaded onto a superose 12 column of the FPLC and eluted with glycyl-glycine buffer (0.05M pH 8.0) at 0.5ml/min. The fractions thus collected were assayed for L-SD and β-galactosidase activities.

10-4. SDS-page Gels

SDS-page gels were prepared according to the instructions provided with the Bio-Rad Mini PROTEAN II dual slab cell.

PART 11. OTHER GENETIC METHODS

11-1. Plasmid Isolation

The plasmid "miniprep" isolations were made according to the methods of Maniatis *et al* [Maniatis *et al.*, 1982]. The large amount of plasmid used for producing for deletions and sequence determination was isolated by a method from F. Lang's lab(personal communication). 500 ml culture with plasmid was harvested by centrifugation and resuspended in 90 ml of Triton-mix buffer (0.4% Triton-100, 100 mM Tris pH8.0, 100 mM EDTA) on a 1 litre beaker, then 4 ml of 10mg/ml lysozyme was added. The beaker with cell supernatant was put on bioling water-bath for 3 minutes. After that, the supernatant was poured into centrifuge tubes and centrifugated at 15000 rpm for 15 minutes. That DNA solution was then treated with NaOAc (0.3 M) for precipitating proteins, LiCl (3.3 M) for precipitating RNA, and RNase, proteinaseK, and phenol and chloroform extraction. Finally, plasmid DNA was precipitated with ethanol and desalted with sephadex G-50 column.

11-2. Chromosomal DNA and λ DNA Isolation

Chromosomal DNA isolation was carried out as described by Silhavy *et al* [Silhavy *et al.*, 1984] and λ DNA was isolated according to the methods of Maniatis *et al* [Maniatis *et al.*, 1982].

11-3. Transformation

Transformations were performed according to Maniatis *et al* [Maniatis *et al.*, 1982].

11-4. Transduction and Conjugation for Mapping

Initial mapping of MEW28 *sdaA*::Cm^r, MEW50 *sdaA*::Cm^r *sdaX* and MEW51 *sdaA*::Cm^r *sdaX* *sdaB*::λplacMu was carried out by conjugation with Hfr strains from Singer *et al* [Singer *et al.*, 1989]. Spontaneously-occurring streptomycin-resistant derivatives of these three strains were used as recipients. The protocol used was from Miller [Miller, 1972].

More accurate mapping was done by P1-mediated transduction also according to the method described by Miller [Miller, 1972]) using Singer' mapping kit, a collection of strains carrying Tn10 insertions at convenient locations (27-42 min. for *sdaA* and 42-65 min. for *sdaX* and *sdaB*) [Singer *et al.*, 1989].

RESULTS

PART 1. ISOLATION AND CHARACTERIZATION OF NEW MUTANTS DEFICIENT IN L-SD ACTIVITY

As the first step in studying the structure of L-SD, I wished to isolate an insertion mutation in the structural gene for L-SD#1. As described in this section, this search led to the definition of 3 new genetic loci by the isolation of new mutants deficient in L-SD as assayed in whole cells. Two of the genes, *sdaA* and *sdaB* were located near 41.0 and 60.25 min of *E. coli* map [Bachmann, 1990], and studied in considerable detail. Strain MEW84 defines a new locus affecting the L-SD activation system.

1-1. Isolation of SGL$^-$ Mutants by λplacMu Insertion

Strains unable to grow with L-serine, glycine and L-leucine (SGL$^-$) were isolated by λplacMu insertion. As described in the Materials and Methods section, this was done with a double selection: strains which had acquired the λplacMu9 insertion were selected by kanamycin resistance, and then cells able to grow on SGL were killed with ampicillin. After the ampicillin selection, a high proportion of the colonies which grew on LB-kanamycin plates showed a SGL$^-$ phenotype. About 70 of these kanamycin resistant, SGL$^-$ strains were assayed for L-SD and β-galactosidase activity.

One of these SGL$^-$ clones, strain MEW21, showed low L-SD when grown in minimal medium with or without the inducers, glycine and L-leucine. If the *lacZ* gene of the λplacMu was inserted in this strain in the L-SD structural gene under the control of the L-SD promoter, this should be indicated by the fact that β-galactosidase was induced by the factors that normally induce L-SD. In fact, β-galactosidase activity was 50 units in uninduced cells, and 350 in cells grown with glycine and L-leucine.

The gene mutated in strain MEW21 by λplacMu insertion was named *sdaA*. To study it further and avoid the presence of a second insert in the strain, the λplacMu9 insert responsible for the SGL⁻ phenotype was transferred to strain MEW1 by transducing with phage P1 grown on MEW21, and selecting kanamycin-resistance. A resulting kanamycin-resistant SGL⁻ transductant, strain MEW22 *sdaA*::λplacMu9, was used for further experiments.

In the same experiment, I isolated another strain SGL⁻ strain carrying a λplacMu insertion and showing altered L-SD synthesis. This strain, MEW83, showed low L-SD when grown in minimal medium with or without inducers, and even when grown in LB medium. It showed the same high β-galactosidase in all these conditions.

It seemed possible that this insert might affect a gene of the L-SD activation system, either one of the previously defined genes (*sda128, sda191*) or a new one. To allow further study, the insert was transferred to strain MEW1 creating strain MEW84, using the transduction strategy used to create strain MEW22.

1-2. Further Characterization of Strains MEW22 and MEW84: Effect of L-SD Inducers on *lacZ* Transcription

If strain MEW22 carried an insertion in the structural gene, β-galactosidase activity in strain MEW22 should parallel L-SD activity in strain MEW1. To test this in greater detail, L-SD and β-galactosidase activities were assayed in cells grown under conditions known to affect L-SD activity (Table 2).

In fact, β-galactosidase was induced sevenfold by glycine and L-leucine, which closely parallels the 6-fold induction of L-SD in the parent strain. Growth in LB also showed parallel induction (13- and 12-fold, respectively). Nonetheless, other inducers of L-SD did not induce β-galactosidase. It was not induced by UV irradiation, and was actually decreased by growth at 42°C. I concluded nonetheless that the parallel induction that was seen was sufficient to make it likely that this strain carried an insert in the structural gene for L-SD#1.

In a similar experiment, L-SD and β-galactosidase activity from *lacZ* inserted in strain MEW84 was assayed, and is compared with similar data for a known activation mutant, strain MEW191 (Table 3). The two mutants showed similar regulation of β-galactosidase and of the low amount of L-SD that could be measured. It seemed possible then that strain MEW84 might also code for a product involved in post-translational regulation of L-SD.

1-3. *In vitro* Activation of *sdaA* and MEW84 Mutants

L-SD appears to be synthesized in an inactive form and to be activated by an as yet unknown enzyme reaction [Newman *et al.*, 1985a, Newman *et al.*, 1985b]. Activation can be mimicked by *in vitro* incubation with iron and dithiothreitol [Newman *et al.*, 1985a]. Two SGL$^-$ mutants, MEW128 and MEW191, described earlier, were deficient in the

Table 2. The effects of changes in growth conditions on synthesis of β-galactosidase from the *sdaA* promoter in strain MEW22[a]

Expt.	Incubation growth condition	L-SD activity of MEW1[b]	L-SD activity of MEW22	β-galactosidase activity of MEW22
1	37°C	19	1	50
2	Glycine and L-leucine	110	7	350
3	UV irradiation	130	2	55
4	42°C	34	8	20
5	Anaerobic growth	ND[c]	4	175
6	Luria Broth	230	65	675

a). Cells were grown in the conditions listed below, subcultured, and assayed in exponential phase for L-SD and β-galactosidase activity, at 37°C, except for experiment 4 at 42°C. In experiment 2, glycine and L-leucine (300 ug/ml each) were added. In experiment 3, exponential phase cells were irradiated with UV for 1 minute and harvested after 1 hour's further incubation. For experiment 5 cells were grown in filled, closed containers (120x16mm). For experiment 6 cells were grown in Luria Broth with 0.5%

glucose. L-SD activity was assayed in whole cells and is expressed as μmoles pyruvate synthesized by 0.1 ml of a 100 K.U. suspension of cells in 35 minutes. β-galactosidase is reported as Miller units[Miller, 1972] and numbers presented are given to the nearest 25 units, except those less than 50 which are given to the nearest 5 units. All the data represent the average of 3-5 experiments.

b). The data for strain MEW1 is taken from reference [Newman *et al* 1985a].

C). ND. Not determined

Table 3. β-Galactosidase and L-SD activities of strain MEW84[a]

Expt.	Incubation growth conditions	L-SD activity of strains MEW191[b]	MEW84	β-Galactosidase of MEW191[b]	MEW84
1	37°C	5	1	461	850
2	Glycine and L-leucine	5	2	490	925
3	LB	12	3	676	1360

a). The growth and assay conditions are the same as those in Table 2 and enzyme activity is reported in the same units. The data for MEW84 represent the average of 3 experiments.

b). The data for strain MEW191 are taken from reference [Newman *et al* 1985b].

whole-cell assay but showed activity in extracts with iron and dithiothreitol [Newman *et al.*, 1985a].

If strain MEW22 carried an insert in the structural gene for L-SD, Fe and DTT could not activate the non-existent structural gene product. The converse is of course true for an activation mutant. I therefore made extracts of the *sdaA* mutant, of strain MEW84 and of the parent strain MEW1, and assayed each of them for L-SD activity in the presence of iron and dithiothreitol (Fig. 8). The results showed that L-SD activity from *sdaA* was very low, (10-fold lower than that of the parent strain). On the other hand, strain MEW84 showed almost 70% of the activity seen in the parent strain, all being grown in glucose-minimal medium with glycine and L-leucine. This excludes the possibility that the *sdaA* mutant is deficient in the activation system for L-SD, and is consistent with the conclusion that *sdaA* codes for either the structural gene for L-SD or a regulatory factor.

The *in vitro* L-SD study suggested that strain MEW84 might carry a mutation in an activation gene. Using the Singer kit, it was roughly mapped by conjugation in order to see whether it could be differentiated by map location from the two known activation mutants [Dumont, 1985]. It was highly linked to a Tn10 element at 42 min (26/30 exconjugants with strain CAG5055) but not at all linked to Tn10 elements at 15 or 89 min (strains CAG5051 and CAG5052). Since MEW128 and MEW191 have been located near 15-17 and 86 min respectively, I conclude that the insertion in strain MEW84 defines a new locus involved in L-SD activation system. Because I was concerned with L-SD structure, I did not characterize this strain further.

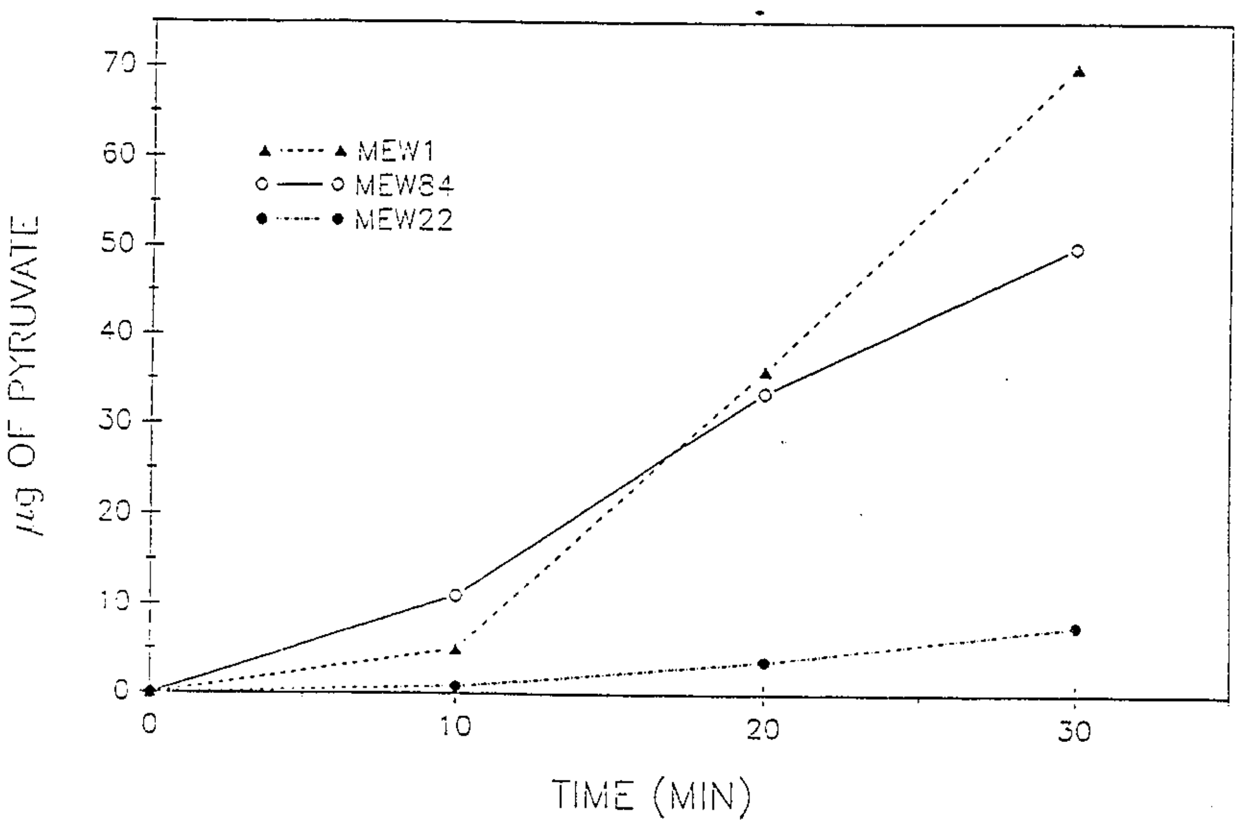

Fig. 8 *In vitro* activation of strain MEW22 and MEW84

Samples (40 μl) of crude extracts (approximately 7.5 mg of protein per ml) of strain MEW1, MEW84 and MEW22 grown on minimal medium with glycine and L-leucine as inducer were incubated with $1.5X10^{-2}$ M iron and 0.15 M DTT for the times indicated, and the amount of pyruvate formed was determinde.

PART 2. CLONING AND SEQUENCING OF THE *sdaA* GENE

2-1. Cloning of *sdaA*

In order to determine the sequence of the *sdaA* gene, it was cloned using the following strategy. Using the *in vivo* cloning system of Groisman and Casadaban [Groisman *et al.*, 1984, Groisman *et al.*, 1986], I isolated a plasmid with a Mu replicon and about 10 kb of *E. coli* DNA complementing the *sdaA* mutation. From this plasmid, smaller fragments were subcloned first into pBR322 and then into Bluescript KS+ and KS- for sequencing. The smallest fragment obtained, 2.6 kb, was then subcloned into pBR322, producing plasmid pMES22, which was used for most studies. The details of this procedure are presented in the Materials and Methods section and shown in Fig. 3.

2-2. L-SD Activity of Strains Carrying the Cloned *sdaA* Gene

The plasmids isolated all complemented the SGL$^-$ phenotype and so might be expected to restore L-SD activity to an *sdaA* mutant. To test this, the plasmids carrying 6 and 2.6 kb of chromosomal DNA (pMES2 and pMES22 respectively) were transformed into both the SGL$^-$ strain, MEW22, and its SGL$^+$ parent, MEW1.

Strains carrying either of these high copy number plasmids showed large increases in L-SD activity (Table 4) and both could grow on L-serine without inducers. The parent strain, MEW1, made 14-fold more L-SD when carrying the plasmid than without it. This large increase in L-SD production was seen with both plasmids and both hosts. L-SD synthesis from either plasmid was inducible by glycine and L-leucine, suggesting that

Table 4. L-SD activity in plasmid-carrying strains[a]

		L-SD activity[a] in:		β-galactosidase activity in:	
Host	Plasmid	Cells	grown	in glucose minimal	medium
		-	+	-	+
MEW22	none	1	7	50	350
MEW1	none	19[b]	110[b]	ND[d]	ND
MEW22	pMES22	350	1,100	45	212
MEW1	pMES22	271	1,270	ND	ND
MEW22	pMES2	203	850	50	262
MEW1	pMES2	186	930	ND	ND

a). Cells were grown in glucose minimal medium with (+) and without (-)glycine and L-leucine(300 μg/ml), subcultured and assayed, and the results reported, all as in Table 2. Antibiotic was added to plasmid-carrying cultures during both overnight growth and subculture as listed in material and methods. The data represent the average of 3 experiments.

b) These data are taken from reference [Newman *et al.*, 1985a]

c) ND. Not determined.

both plasmids contain the regulatory information needed for this induction. Furthermore, L-SD was synthesized to about the same level from both plasmids, suggesting that little if any regulatory information was lost in cutting from 6 to 2.6 kb.

On the other hand, β-galactosidase encoded by the *lacZ* gene, presumably under the control of the chromosomal L-SD promoter, was produced at the same level whether the plasmid was present or not. Its synthesis was also induced by glycine and L-leucine but to a lesser extent when the cell carried a plasmid.

It seems then that the *sdaA* clone did not alter transcription from the chromosomal *sdaA* promoter in cells grown in glucose-minimal medium. The lower β-galactosidase synthesized from the chromosomal *sdaA* promoter may indicate that the plasmid and chromosomal *sdaA* promoters shared the same regulatory sequence and compete for inducers.

2-3. Hybridization of pMES22 to Chromosomal DNA with Insertion in *sdaA*

It is clear that pMES22 complements an L-SD deficiency caused by the mutation in strain MEW22. However this does not of itself prove that pMES22 carries the same gene that is mutated in strain MEW22. It would be expected that the *sdaA* gene lies in one contiguous stretch of DNA in the parent strain but in two pieces separated by λplacMu in the *sdaA* mutant. If the 2.6 kb of chromosomal DNA in pMES22 actually carried the *sdaA* gene, it should hybridize to a 2.6-kb *Sal*I-*Pst*I fragment of chromosomal DNA from the parent strain. However, this band should be missing the *sdaA* mutant and replaced by one or two bands of different sizes.

To test this, a 2.6 kb *Bam*HI-*Sal*I fragment of pMES22 DNA was used to probe *Pst*I-*Sal*I digests of chromosomal DNA from the two strains (Fig. 9). The pMES22-derived probe hybridized to one major band and two minor bands in the digest of the parent strain. The major band (2.6 kb) was replaced in the mutant by two other bands (3.1 and 5.7 kb), indicating that the insertion in strain MEW22 *sdaA*::λplacMu was indeed in the same gene that was cloned on pMES22.

I conclude that the gene mutated in strain MEW22 was in fact cloned on pMES22. The existence of minor bands in the hybridization experiment (Fig. 9) probably indicates that other DNA sequences of *E.coli* have some homology with this sequence, since the pBR322 plasmid itself did not hybridize to these sites, and therefore contamination by pBR322 DNA could not account for these bands (data not shown). The existence of a second L-SD enzyme was suggested earlier because a strain carrying a mutation in the (then) putative structural gene, strain MEW15, still showed L-SD activity in LB medium [Newman *et al.*, 1985a].

2-4. DNA Sequence of the *sdaA* Gene

In order to sequence the 2.6-kb *Sal*I-*Pst*I fragment, it was cloned into bluescript + and - producing plasmids pMES3 and pMES4. Both of these plasmids carry the *sdaA* gene and complement the *sdaA* mutation i.e. strain MEW22 carrying either of these plasmids can grow on NSIV medium. Nine different deletion plasmids each from pMES3 and pMES4 were used for the sequencing reaction. The reaction from each single strand of appropriate length was read for up to 500 bp until a sequence was found which

Fig. 9 Hybridization of chromosomal DNA of mutant MEW22

Chromosomal DNA from strain MEW1 and MEW22 were isolated, digested with *Pst*I and *Sal*I, and electrophoresed on a 1% agarose gel. Then the DNA bands were transferred to a nitrocellulose filter and hybridized with ^{32}P-labeled *Sal*I-*Bam*HI fragment derived from plasmid pMES22. Lane a: MEW22 *sdaA*::λplacMu9. Two hybridization signals, 3.7 and 5.1 kb bands, can be seen. Lane b: MEW1 *sdaA*$^{+}$. One band, 2.6-kb in size, is seen. The molecular weight markers are indicated on the left of the diagram.

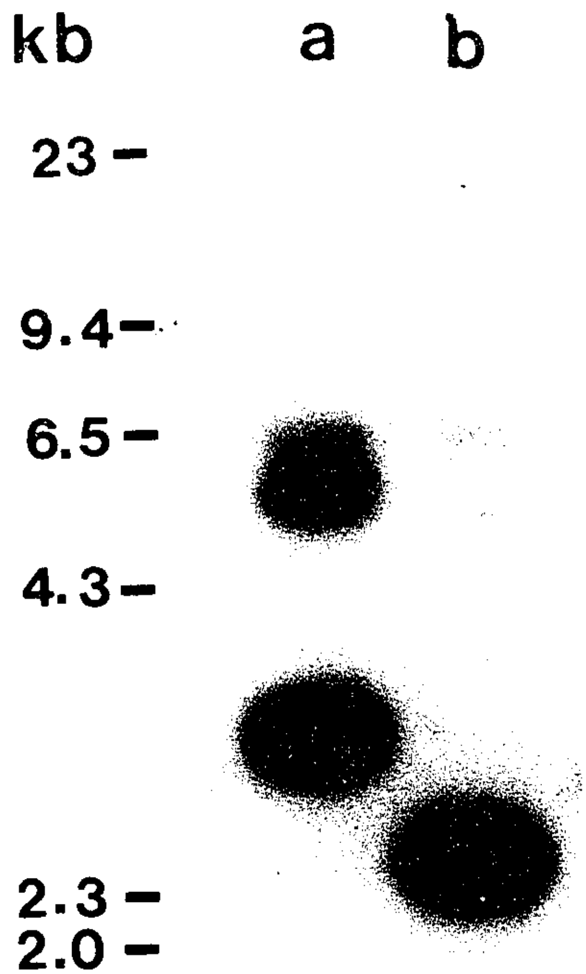

Fig. 9 Hybridization of chromosomal DNA of mutant MEW22

recurred in the next set of reactions. In this way, the sequence could be determined for the whole 2.6-kb sequence in each direction. The set of plasmids covering the entire sequence from pMES3 was also sequenced with reactions in which the dGTP was replaced by dITP in order to avoid mistakes which might be caused by the strong secondary structure in the sequence.

The sequence determined was 2,605 base pairs long, with a G+C content of 52.4% (Fig.10). It contained three possible open reading frames (ORFs) in the direction from *Sal*I to *Pst*I (Fig. 11). The first possible protein-coding region started with the *Sal*I site and ended at nucleotide 461. The second ORF was completely contained within the sequenced fragment, from an AUG codon at nucleotide 663 to position 2009, coding for 448 amino acids. A third possible ORF ran from position 2140 through to the *Pst*I site at the end of the fragment. The longest ORF on the other strand is too short to be the coding sequence.

An analysis of codon usage in the three possible protein-coding regions indicated that only the complete sequence from 663 to 2009 has a codon bias characteristic of *E. coli* protein-coding genes [Gouy and Gauthier, 1982], avoiding AUA, CUA, CGA, CCC, AGA, and GGA codons. Other features which suggest that this sequence represents a functional gene are 1)the fact that a possible ribosome-binding site (GUGAU), which complements the 3' end of 16S rRNA (AUCAU) is located 13 base pairs upstream of the AUG codon at 663 (Fig. 10) and 2)the location 11 base pairs downstream of the UAA stop codon of a possible RNA secondary structure characteristic of *E. coli* transcription termination signals (Fig. 12).

```
1     GTCGACCCCAACCCGGCTTGTTGCTGACTCAGCGTTCGATTCATCTGCGTAAACACCCTGGACAAGTGGCATTCCCTGGAGGTGCAGTCGATGACACCGA
101   CGCATCAGCTATCGCCGCCCGCTGCGCCAAGCTGAAGAAGAGCTCGCTATACCGCCTTCCGCCGTTGAAGTTATCGGCGTGCTGCCGCCCCGTCGATAGC
201   GTCACTGGCTACCAGGTAACCCCAGTGGTCGGCATTATCCCCCCCGATCTGCCGTATCGCGCCAGTGAAGATGAAGTCTCGGCGGTGTTTGAAATGCCGC
301   TCGCCCAGGCATTACATCTGGGTCGTTATCACCCTTTAGATATCTACCGCCGTGGTGATTCACATCGGGTATGGCTGTCCTGGTACGAACAGTATTTTGT
401   ATCGGGAATGACCGCAGGCATAATTCGTGAGCTGGCGCTGCAAATTGGTGTGAAACCCTGACTATACTTATCTTTACATCTACAAAACACTACTTGAGAC
501   AATCATCGCAATATTAGTTAAATCGCGGTTTTTGATTAGTTTAATTCATGTGAATAGTTAAGCCAGTCGCCGCGTTCCCTCTTACACTATGCGCTGTTAT

                                                                            M   F   K   V   C   I   G   P   S
601   TAGTTCGTTACTGGAAGTCCAGTCACCTTGTCAGGAGTATTATCGTGATTAGTCTATTCGAC  ATG TTT AAG GTG GGG ATT GGT CCC TCA

       S   S   H   T   V   G   P   M   K   A   G   K   Q   F   V   D   D   L   V   E   K   G   L   L   D
690   TCT TCC CAT ACC GTA GGG CCT ATG AAG GCA GGT AAA CAG TTC GTC GAT GAT CTG GTC GAA AAA GGC TTA CTG GAT

       S   V   T   R   V   A   V   D   V   Y   G   S   L   S   L   T   G   K   G   H   H   T   D   I   A
765   AGC GTT ACT CGC GTT GCC GTG GAC GTT TAT GGT TCA CTG TCG CTG ACG GGT AAA GGC CAC CAC ACC GAT ATC GCC

       I   I   M   G   L   A   G   N   E   P   A   T   V   D   I   D   S   I   P   G   F   I   R   D   V
840   ATT ATT ATG GGT CTT GCA GGT AAC GAA CCT CCC ACC GTG GAT ATC GAC AGT ATT CCC GGT TTT ATT CGC GAC GTA

       E   E   R   E   R   L   L   L   A   Q   G   R   H   E   V   D   F   P   R   D   N   G   M   R   F
915   GAA GAG CGC GAA CGT CTG CTG CTG GCA CAG GGA CGG CAT GAA GTG GAT TTC CCG CGC GAC AAC GGG ATG CGT TTT

       H   N   G   N   L   P   L   H   E   N   G   M   Q   I   H   A   Y   N   G   D   E   V   V   Y   S
990   CAT AAC GGC AAC CTG CCG CTG CAT GAA AAC GGT ATG CAA ATC CAC GCC TAT AAC GGC GAT GAA GTC GTC TAC AGC

       K   T   Y   Y   S   I   G   G   G   F   I   V   D   E   E   H   F   G   Q   D   A   A   N   E   V
1065  AAA ACT TAT TAT TCC ATC GGC GGC GGT TTT ATC GTC GAT GAA GAA CAC TTT GGT CAG GAT GCT GCC AAC GAA GTA

       S   V   P   Y   P   F   K   S   A   T   E   L   L   A   Y   C   N   E   T   G   Y   S   L   S   G
1140  AGC GTG CCG TAT CCG TTC AAA TCT GCC ACC GAA CTG CTC GCG TAC TGT AAT GAA ACC GGC TAT TCG CTG TCT GGT

       L   A   M   Q   N   E   L   A   L   H   S   K   K   E   I   D   E   Y   F   A   H   V   W   Q   T
1215  CTC GCT ATG CAG AAC GAA CTG GCG CTG CAC AGC AAG AAA GAG ATC GAC GAG TAT TTC GCG CAT GTC TGG CAA ACC

       M   Q   A   C   T   D   R   G   M   N   T   E   G   V   L   P   G   P   L   R   V   P   R   R   A
1290  ATG CAG GCA TGT ATC GAT CGC GGG ATG AAC ACC GAA GGT GTA CTG CCA GGC CCG CTG CGC GTG CCA CGT CGT GCG

       S   A   L   R   R   M   L   V   S   S   D   K   L   S   N   D   P   M   N   V   I   D   W   V   N
1365  TCT GCC CTG CGC CGG ATG CTG GTT TCC AGC GAT AAA CTG TCT AAC GAT CCG ATG AAT GTC ATT GAC TGG GTA AAC

       M   F   A   L   A   V   N   E   E   N   A   A   G   G   R   V   V   T   A   P   T   N   G   A   C
1440  ATG TTT GCG CTG GCA GTT AAC GAA GAA AAC GCC GCC GGT GGT CGT GTG GTA ACT GCG CCA ACC AAC GGT GCC TGC

       G   I   V   P   A   V   L   A   Y   Y   D   H   F   I   E   S   V   S   P   D   I   Y   T   R   Y
1515  GGT ATC GTT CCG GCA GTG CTG GCT TAC TAT GAC CAC TTT ATT GAA TCG GTC AGC CCG GAC ATC TAT ACC CGT TAC

       F   M   A   A   G   A   I   G   A   L   Y   K   M   N   A   S   I   S   G   A   E   V   G   C   Q
1590  TTT ATG GCA GCG GGC GCG ATT GGT GCA TTG TAT AAA ATG AAC GCC TCT ATT TCC GGT GCG GAA GTT GGT TGC CAG

       G   E   V   G   V   A   C   S   M   A   A   A   G   L   A   E   L   L   G   G   S   P   E   Q   V
1665  GGC GAA GTG GGT GTT GCC TGT TCA ATG GCT GCT GCG GGT CTT GCA GAA CTG CTG GGC GGT AGC CCG GAA CAG GTT

       C   V   A   A   E   I   G   M   E   H   N   L   G   L   T   C   D   P   V   A   G   Q   V   Q   V
1740  TGC GTG GCG GCG GAA ATT GGC ATG GAA CAC AAC CTT GGT TTA ACC TGC GAC CCG GTT GCA GGG CAG GTT CAG GTG

       P   C   I   E   R   N   A   I   A   S   V   K   A   I   N   A   A   R   M   A   L   R   R   T   S
1815  CCG TGC ATT GAG CGT AAT GCC ATT GCC TCT GTG AAG GCG ATT AAC GCC GCG CGG ATG GCT CTG CGC CGC ACC AGT

       A   P   R   V   S   L   D   K   V   I   E   T   M   Y   E   T   G   K   D   M   N   A   K   Y   R
1890  GCA CCG CGC GTC TCG CTG GAT AAG GTC ATC GAA ACG ATG TAC GAA ACC GGT AAG GAC ATG AAC GCC AAA TAC CGC

       E   T   S   R   G   G   L   A   I   K   V   Q   C   D   *
1965  GAA ACC TCA CGC GGT GGT CTG GCA ATC AAA GTC CAG TGT GAC TAA TACTTCTTACTCGCCCATCTGCAACGGATGGGCGAATTTA
2050  TACCCGCTTTCTCGTCTGCTGTAATATTCCCCACTACACTTCCACTGTTGCGTCAGGCGTTTGTCGCCATACGCTTACAGGGTGGCCCGCATGCAAAAAG
2150  CACAACGGATCATTAAAACCTATCGTCGTAATCGAATGATTGTTTGTACGATTTGCGCACTCGTTACGCTCGCTTCGACCCTGAGCGTGCGATTTATTTC
2250  ACAGCGTAACTTAAATCAACAACGGGTAGTACAATTCGCCAATCACGCTGTAGAGCAATTAGATAAAGTACTGCTTCCCCTACAGGCAGGTAGCGAAGTC
2350  TTGCTTCCGCTGATTGGTCTGCCCTGCTCTGTCGCCCATTTGCCATTACGTAAACAGGCGGCAAAACTCCAAACTGTGCGATCCATTGGCCTGGTGCAAG
2450  ACGGCACACTTTATTGCTCCAGCATTTTTGGTTATCGCAATGTGCCCGTCGTGGACATTCTGGCTGAACTTCCTGCACCGCAACCACTTTTACGCCTGAC
2550  GATCGACCGTCCCCTGATTAAAGGCAGTCCGGTTTTGATTCAATGGACGCCTGCAG
```

Fig. 10 DNA sequence of the *E. coli sdaA* region.

2,605 bp of *sdaA* gene coding region was shown, as well as the amino acids for *sdaA* gene product. A possible ribosome-binding site is underlined.

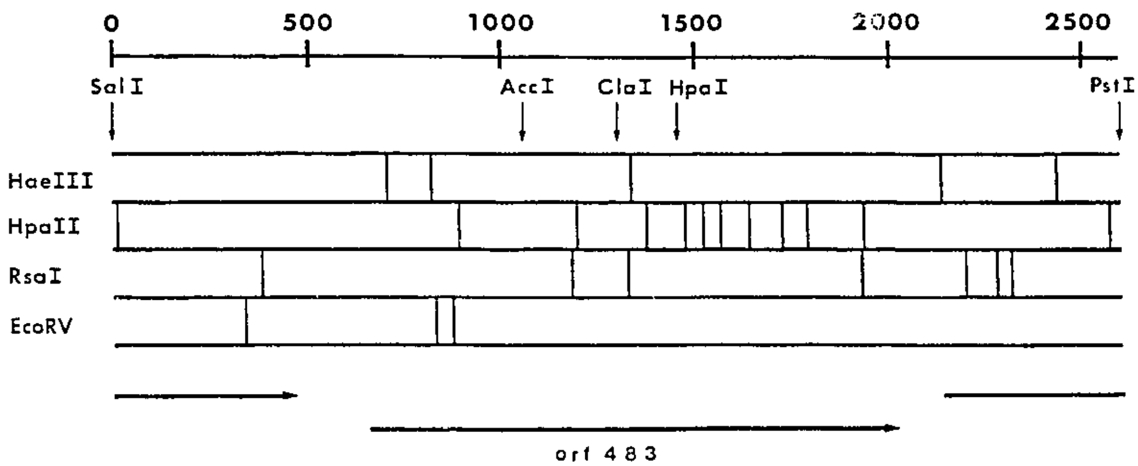

Fig. 11 Physical map of the *Sal*I-*Pst*I DNA fragment making up the *sdaA* coding sequence

A selected number of restriction sites is shown, allowing the precise assignment of the possible protein-coding regions and the possible ORF of the *sdaA* gene. The longest possible ORF which is completely included in the *Sal*I-*Pst*I fragment is ORF 483 (483 amino acid residues). Two further truncated ORF's encoded on the same DNA strand are also indicated.

```
                                    A
                                 C     A
                                  G-C
                                  U.G
                                  C-G
                                  U-A
                                  A-U
                                  C-G
                                  C-G
                                  C-G
                                  G-C
                                  C-G
5' ... GAC UAA UACUUCUUACU-AAUUUAUACCC ... 3'

---------->stop codon
    ORF
```

Fig. 12 A possible transcription termination signal of the *sdaA* gene

A transcription termination signal at position 11 downstream from the AUU stop codon is indicated.

I have postulated that the ORF begins with an AUG codon at 663. However it is also possible that it starts with a GUG codon at 645. Seven bases upstream of this codon, there is a much better Shine-Dalgarno sequence (AGGAG) [Stormo *et al.*, 1982] at position 633. On the other hand, very few E. coli coding sequences use a GUG start codon. Only about 8% of genes sequenced start with this codon. Therefore, the exact translation start of the *sdaA* gene is not certain, and should be confirmed by determining the N-terminal amino acid sequence of the *sdaA* gene product.

Whichever the actual start codon may be, since this long ORF was the only complete protein-coding sequence and showed all the criteria for an expressed ORF, I believe this to be the sequence which codes for the *sdaA* gene product.

2-5. Use of a *lacZ* Fusion to Locate the Position and Orientation of the ORF Corresponding to *sdaA*

As described in the preceding section, analysis of both strands of the *sdaA* clone indicated that only the ORF from nucleotide 663 to 2009 was likely to correspond to *sdaA*. To confirm this, I constructed an inframe fusion of a truncated *sdaA* to *lacZ*. To do this, I inserted the *lacZ* DNA sequence from pMC1871 into the unique *Hpa*I site of our clone, as described in the methods section. This produced a plasmid pMEZS22, carrying the first 1,458 bases of the putative ORF fused in frame to *lacZ*. I determined that this construction actually produced the correct inframe fusion to *lacZ* by sequencing across the fusion junction in plasmid pMEZS22. The sequence determined was AT GTC ATT GAC TGG GTA AAC ATG TTT GCG CTG GCA GTT/GGG GAT CCC GTC GTT

TTA CAA CGT CGT GAC TGG G, which corresponds to the last 38 bases prior to the *Hpa*I site of *sdaA* joined in frame to the first 34 bases of the *lacZ* of pMC1871.

If *lacZ* were in fact inserted in the appropriate orientation in the reading frame of the *sdaA* gene, it should be controlled by the same factors as regulate *sdaA* itself. I tested this in two ways. First, strain MEW1 (pMEZS22) was grown in glucose-minimal medium with and without glycine and L-leucine and its β-galactosidase activity determined. Cultures grown without inducers made 800 U of β-galactosidase, indicating that the reading frame was indeed functional and thus an active β-galactosidase could be made (Table 5). Cultures grown with glycine and L-leucine made four times more β-galactosidase. This is similar to the induction of L-SD in strain MEW1 with plasmid pMES22 (Table 4). It is clear, then, that *lacZ* was inserted in a reading frame whose promoter was regulated by L-SD inducers. Second, I further confirmed this by transforming the fusion plasmid into an *ssd* mutant. This mutant synthesized L-SD from its chromosomal gene at a much higher level (5- to 20-fold higher) than the wild type did [Newman *et al.*, 1985b]. One would therefore expect it to overexpress the gene carried on the plasmid. In agreement with this expectation, the β-galactosidase level synthesized from the fusion plasmid was much higher in the *ssd* mutant than in the parent strain (Table 5). The fact that the *ssd* mutation had an effect on the *sdaA* promoter resulting in increased synthesis of L-SD, also supported the idea that the *sdaA* gene codes for L-SD.

Table 5. β-Galactosidase activity synthesized from an *sdaA-lacZ* fusion plasmid pMEZS22[a]

Host	Plasmid	β-Galactosidase activity: glucose-minimal medium	
		without inducers	with inducers
MEW1	None	5	5
MEW1	pMEZS22	800	3,100
MEW22	None	5	ND[b]
MEW22	pMEZS22	8000	ND

a). Experiments were carried out as in Table 2. Antibiotic was added to plasmid-carrying cultures. The data represent the average of 3 experiments.

b). ND, Not determined.

PART 3. DEMONSTRATION *sdaA* GENE CODES FOR L-SD#1

A good deal of evidence suggested that the *sdaA* gene codes for the structure of L-SD#1. The *sdaA* gene codes for a 448 amino-acid protein with a caculated molecular weight around 48 kd, which is almost the same size as is seen for L-SD activity by gel filtration. Moreover, plasmids carrying the *sdaA* gene all produced very high L-SD activity. β-Galactosidase synthesis regulated by the *sdaA* promoter was induced by glycine and L-leucine, anaerobic growth, *ssd* mutation and LB medium, the conditions previously shown to induce L-SD synthesis.

All of this evidence strongly supports the idea that the *sdaA* gene carries the structural information for synthesizing L-SD#1. However they do not decide this issue definitively. The experiments described in this section demonstrate that the *sdaA* gene codes for the actual structure of L-SD#1.

The experimental approach used was to fuse an intact *sdaA* gene to an intact *lacZ* gene, isolate the fusion protein by its β-galactosidase activity, and determine whether the fusion protein also carried L-SD activity. As is detailed below, these experiments showed directly that *sdaA* does in fact code for L-SD#1.

3-1. Conversion of the Stop Codon of the *sdaA* Gene to an *Eco*RI Site

The first step in the construction of the fusion protein required the formation at the end of the *sdaA* gene of a restriction site which could also be found in a linker region of a *lacZ* clone. Inspection of the *sdaA* stop codon TGT GA<u>C</u> T<u>A</u>A TAC TTC showed that it could be mutated to an *Eco*RI site (GAATTC) by changing only the underlined bases and forming TGT GAA TTC TAC TTC. This also changes the stop codon to one coding for an amino acid.

To do this, I used site-directed mutagenesis. With the following oligonucleotide as primer (AA GAA GAA* TTC* GTC ACA CTG) which corresponds to base pairs #1998-2016 of the original sequence (Fig. 5), it was possible to change the T at #2005 and the A at # 2007 to A and C respectively as described in Materials and Methods.

The new plasmid with an *Eco*RI site (GAATTC) in the stop codon of *sdaA* was named pMES25. The 2kb *Sal*I-*Eco*RI DNA fragment from pMES25 was inserted into BS+ plasmid *Sal*I and *Eco*RI site producing plasmid pMES26 as described in Materials and Methods (Fig. 6). I confirmed that the sequence upstream of and including the *E.co*RI site was correct by sequencing, and showed that the intended changes- and no others-had been made.

3-2. Fusion of the Whole *sdaA* Gene to the *lacZ* Gene

A plasmid coding for a fused gene producing a single protein with both L-SD and β-galactosidase activity, known as the L-SD#1-β-galactosidase or two-part fusion (2PF) was constructed by inserting *sdaA* cut from the pMES26 *Sma*I site into the plasmid pMC1871 *Sma*I site. This generated plasmid pMES27 (see Fig. 7 in Part 2).

I confirmed that L-SD and β-galactosidase were both coded by this plasmid by transforming it into strain MEW28 (*sdaA*::Cm^r; construction of this strain is in the Materials and Methods and is also discussed later). Strain MEW28 carrying this plasmid showed β-galactosidase activity in minimal medium and this was slightly induced by glycine and L-leucine (Table 6). The L-SD activity from this two-part fusion plasmid was low in the whole cell assay, but induction by glycine and L-leucine was demonstrated.

If *sdaA* gene really codes for L-SD#1, one might expect the strain carrying pMES27 to have the same L-SD as a strain carrying pMES22 (Table 5). The fact that it did not is probably due to the altered conformation of L-SD in the fusion protein. The β-galactosidase molecule is a tetramer of which each monomer is almost twice big as the *sdaA* gene product. In this large fusion, the presence of the β-galactosidase molecule may interfere with either *in vivo* activation or L-SD activity.

If the construct interferes with *in vivo* activation- as by making activation sites inaccessible, it still might be possible to demonstrate L-SD activity on *in vitro* assay of the extracts with iron and dithiothreitol. I, therefore, made extracts from strains MEW28, MEW28 with pMEZS22 and the same host with pMES27 and assayed each of them for L-SD activity in the presence of iron and dithiothretiol (Fig. 13). The strain carrying pMES27 showed a great deal of L-SD activity. Neither the parent strain itself, nor the strain carrying the fusion of a part of the *sdaA* gene to *lacZ* (pMEZS22) showed L-SD activity.

The preceding results indicate that the strain carrying the *sdaA-lacZ* fusion gene in plasmid pMES27 can make an inactive form of L-SD fused to β-galactosidase and this can be activated by iron and dithiothreitol like the inactivated form of L-SD itself. However, this evidence for the existence of an activatable form of L-SD in these cells

Table 6. L-SD and β-galactosidase activity synthesized from plasmids pMES27 and pMES28[a]

			L-SD activity in:		β-Galactosidase activity in:	
Expt.	Host	Plasmid	Cells grown in glucose- minimal medium			
			-	+	-	+
1	MEW28	None	2[b]	2[b]	ND[c]	ND
2	MEW28	pMES27	13	23	2,400	3,500
3	MEW28	pMES28[d]	3	ND	12,800	ND

a). Experiments 1 and 2 were carried out as in Table 4, with the addition of antibiotic to plasmid-carrying cells. For expt. 3, cells were grown at 28 C and otherwise treated similarly.

b). These data are taken from Table 12.

c). ND. Not determined.

d). The construction of this plasmid is shown in Materials and Methods and in a later part of the text.

Fig. 13 *In vitro* activation of the fusion proteins

Strain MEW28 carrying either plasmid pMES27 (which codes for the 2PF), or plasmid pMEZS22 (which codes for the *sdaA'-lacZ* fusion referred to here as H-*lacZ*), and also without plasmid, was grown with glycine and L-leucine at 37°C. Strain MEW28 carrying plasmid pMES28 (which codes for the 3PF) were grown at 28°C and subjected to heat shock as described in Table 6. Extracts were made as in Fig. 8. The amount of protein was determined by Lowry assay and aliquots containing 25, 1.25, 0.42 and 25 μg of protein from MEW28 with pMEZS22 (H-*lacZ*), pMES27 (2PF), pMES28 (3PF), and MEW28 itself respectively were assayed as a function of time for both β-galactosidase and L-SD activity using iron and DTT for the L-SD assay as in Fig. 8. 10 μg of BSA were added to the assay tubes to avoid problems due to the low protein concentrations in these assays. The term "Control" in the L-SD figure represents results for both strain MEW28 and strain MEW28 pMEZS22.

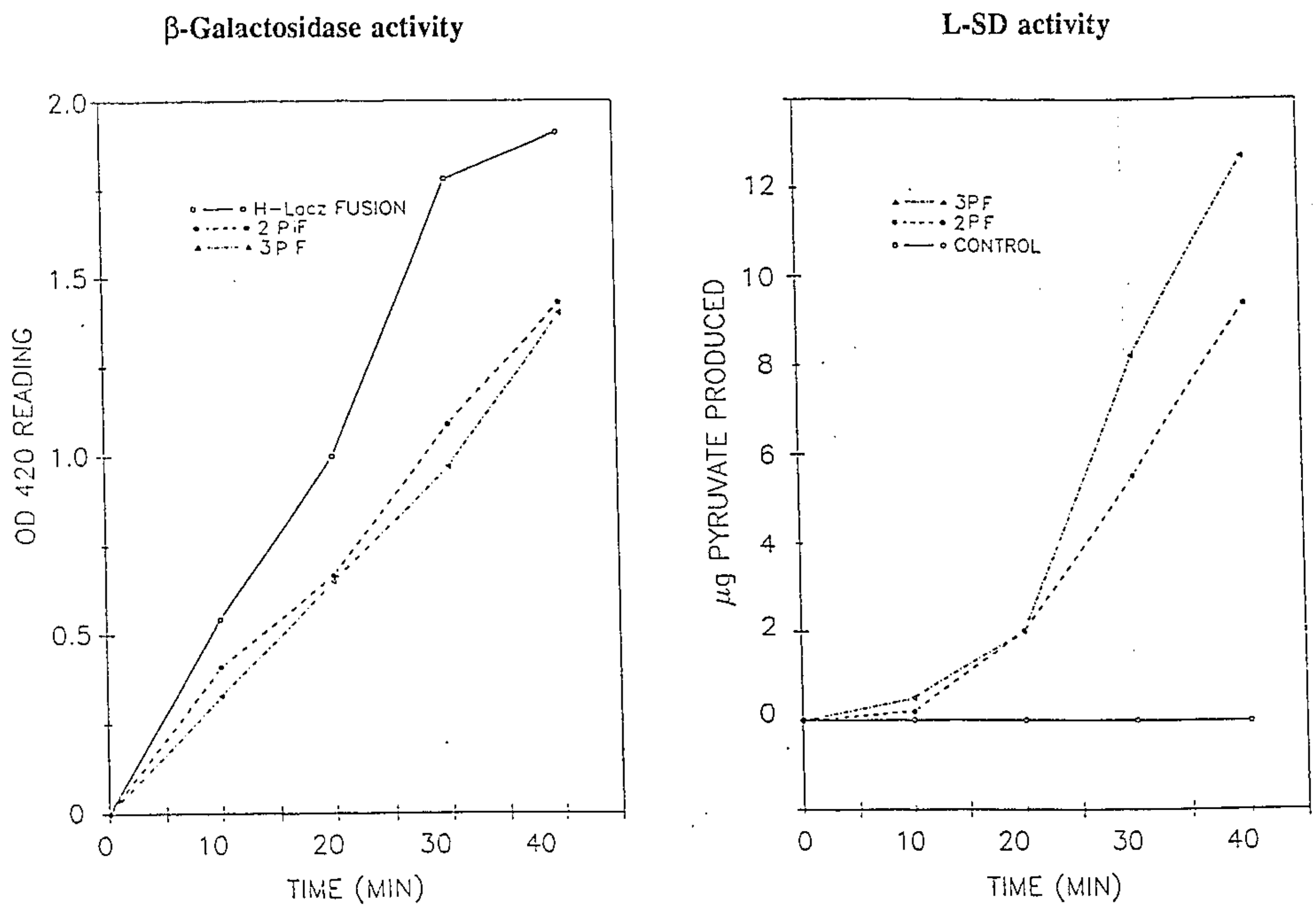

Fig. 13 *In vitro* activation of the fusion proteins

does not itself prove that the *sdaA* gene codes for the L-SD enzyme structure, even though this seems a probable explanation of the data.

3-3. Purification of 2-Part Fusion Protein

A. Ullmann had developed a one-step purification method for hybrid proteins exhibiting β-galactosidase activity, based on affinity chromatography in the presence of high salt concentration [Ullmann, 1984]. Since the *sdaA-lacZ* encoded fusion protein showed β-galactosidase activity, it might be purified by this method.

The method of A. Ullmann calls for elution by high pH sodium borate buffer. However those conditions might have a deleterious effect on L-SD activity, so I tried to elute with a high concentration of the enzyme substrate, i.e. 20% D-lactose in TMN buffer. Using this method, about 90% of β-galactosidase activity of the crude extract could be recovered from an affinity column.

The data shown in Table 7 gives the specific activities of the fusion protein in crude extracts and after eluting with 20% lactose in TMN buffer. As can be seen, a 40-fold of purification of the 2-part fusion protein was achieved. The SDS-gel profile of the purified protein are shown in Fig. 15. The preparations were quite pure and were suitable for making antibody, determining the amino acid sequence, and biochemical study.

The L-SD activity before and after the purification on the β-galactosidase affinity column is also shown in Table 7. It is clear that the L-SD was purified with the purification of fusion protein. This is the definitive proof that *sdaA* codes for the structure of L-SD. To make this even clearer, the 2-part fusion protein purified from the

Table 7. Partially purified hybrid proteins[a]

Host	Plasmid	β-Galactosidase activity[b] Crude extract	Purified protein	L-SD activity[b] Crude extract	Purified proteins
MEW28	None	5	ND[c]	4	ND
MEW28	pMES27[d]	10,000	290,000	3,600	48,000
MEW28	pMES28[d]	35,000	250,000	117,000	243,000

a). Cells were grown as shown in Table 6, harvested, and extracts made as in Fig. 8. 20 ml of 2PF or 10 ml of 3PF crude extracts were loaded on a 2 ml p-aminobenzyl 1-thio-β-D-galactopyranoside agarose column which was washed with 500 ml of TMN buffer at 4°C. The fusion proteins were then eluted with 20 ml of 20% D-lactose in TMN buffer and collected at 0.5ml/tube until most of proteins were eluted. The protein concentrations were determined by coomassie blue protein assay.

b). The β-Galactosidase activity is expressed as:

OD420 readingx380/time(minutes)x mg. proteins.

The eluted protein was diluted 100 times and 10 μl used in a 1 ml β-galactosidase assay. This avoided inhibition by the high concentration of D-lactose used to elute from the affinity column.

The L-SD activity is expressed as μMol pyruvate produced/minute/mg protein.

c). ND. Not determined.

d). pMES27 produced the 2PF and pMES28 produced the 3PF.

affinity column was subjected to gel filtration on a FPLC superose 12 column, and the resultant fractions assayed for β-galactosidase and L-SD activity. If *sdaA* codes for L-SD#1, one would expect that β-galactosidase and L-SD activities would elute together, at a higher MW than either proteins chromatographed alone. As shown in Fig. 14, the fusion protein harboring β-galactosidase and L-SD activity both peaked at fraction 20.

I conclude from the results presented in this section of the thesis that *sdaA* does code for L-SD#1, the L-SD activity produced in glucose-minimal medium.

3-4. Construction of a 3-Part Fusion to Facilitate Purification of the *sdaA* Gene Product

Bastia and his co-workers developed a method which permits rapid purification of the product of almost any fused gene [Germino and Bastia, 1984]. This is done by making the appropriate genetic fusion into a plasmid which is constructed such that site-specific proteolysis will liberate the desired gene product.

In their plasmid, the target cistron DNA is fused to a marker cistron, *lacZ*, just as was done in the construction of the two part fusion. However in this case, the junction is made via a piece of DNA that codes for a linker peptide, a DNA fragment encoding 60 amino acids from the triple helical region of chicken proα-2 collagen [Fuller and Boedtker,1981].

The peptide encoded contains a site sensitive to collagenase. This permits the tripartite hybrid protein, purified by affinity column as before, to be digested with a purified microbial collagenase to cleave the linker peptide. Then any chromatography

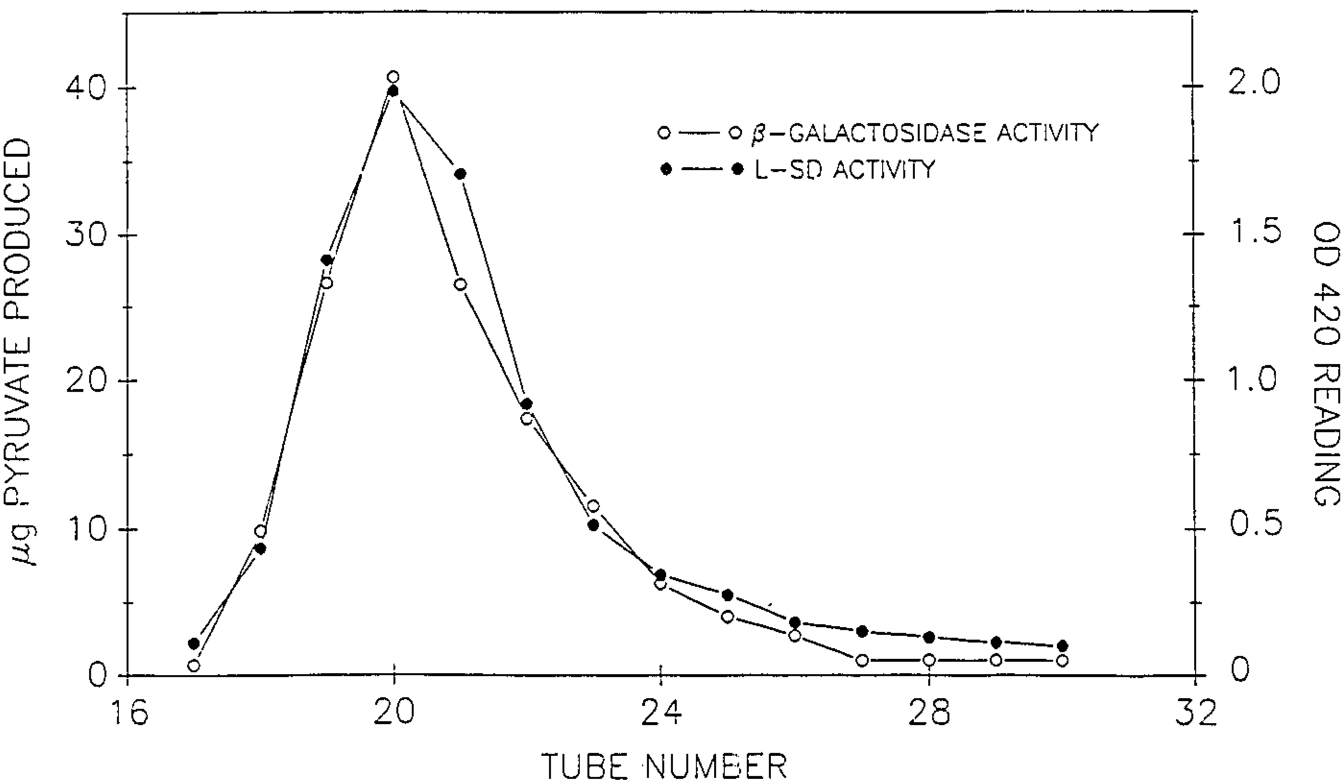

Fig. 14 FPLC analysis of the two-part fusion protein

30 μg of 2PF was loaded onto a FPLC superose 12 column and eluted at 0.5ml/min collecting 0.5ml fractions. The fractions were then assayed for β-galactosidase and L-SD activity. One experiment typical of several performed is shown.

method, in this case FPLC, allows the rapid isolation of the target protein from the marker protein.

To use this system for isolating the *sdaA* gene product, I first verified that there is no consensus cleavage site for collagenase (-Pro-X-Gly-Pro-Y-) within the *sdaA* protein. Fortunately, no such peptide sequence was found in the amino acid sequence translated from the *sdaA* gene. It was therefore possible to construct a three-part fusion gene as described in Materials and Methods. Using the same methods as for the two-part fusion, the 3-part fusion protein was purified by affinity column and showed high purity in the protein gel (Fig. 15). In the crude extract, I found a β-galactosidase level of 35,000 units/mg protein whereas after affinity chromatography, I found 250,000 units/mg protein, a purification of about 7-fold (Table 7).

The cleavage of this protein by collagenase is demonstrated in the SDS-PAGE gel in Fig. 15 lane. F. In lane F, one can see the high MW 3 part fusion protein, which has disappeared in lane E where it is replaced by a β-galactosidase band of the same size as the β-galactosidase marker protein, and a smaller molecular weight band around 48 kDa, as estimated by comparison with the 45 kDa marker protein. This lower molecular weight band is likely to be the *sdaA* gene product released by collagenase from the fusion protein.

The β-galactosidase was further separated from the *sdaA* gene product by passing the collagenase-treated 3PF over a superose 12 column of the FPLC (Fig. 16). L-SD activity was detected in fractions considerably later than those containing β-galactosidase in the high molecular weight range. A very small amount of the intact three-part fusion was

Fig. 15 SDS-page gel of fusion proteins

The fusion proteins were prepared as in Table 6. A 7.5% of SDS-page gel was used following the manual of BIO-RAD, and then stained with Coosmassie Blue. Lane A: β-galactosidase (1 μg); Lane B: crude extract (5 μg) from MEW28 with pMES27; Lane C: affinity purified 2PF (1 μg); Lane D: crude extract (5 μg) from MEW28 with pMES28; Lane E: affinity purified 3PF (1 μg); Lane F: 3PF digested with collagenase as in Fig. 15 (1 μg). Lane G: ovalbumin (5 μg). The molecular weight markers are indicated on the right.

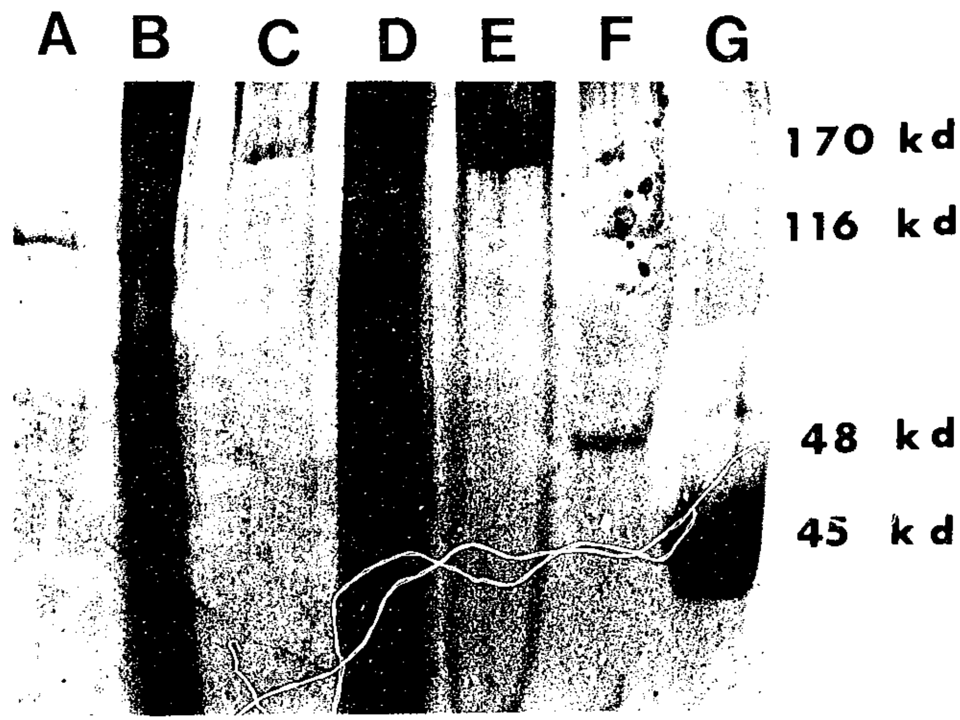

Fig. 15 SDS-page gel of fusion proteins

detected in the high molecular weight fraction. However, most of the β-galactosidase activity was found in the fractions corresponding to the MW of free β-galactosidase, and those fractions did not show L-SD activity. It seems therefore that the collagenase treatment cleaved most of the three-part fusion molecules.

It is clear that the *sdaA* gene product, L-SD#1, could be purified by this simple method. The L-SD activity so purified required the addition of iron and DTT for activation, just as the L-SD in crude extracts does. The *sdaA* gene product purified in this way can be used for biochemical study, as indicated below.

PART 5. BIOCHEMICAL STUDY OF L-SD#1

In this part of the results section, some biochemical aspects of the fusion protein are described, and the activation characteristics of L-SD and the fusion protein are considered.

5-1. Biochemical Study of the *sdaA-lacZ* Fusion Protein

L-SD has proved very difficult to study because of its instability, whether activated or inactivated [Newman *et al.*, 1985a]. This is also true for semipurified enzyme in the various fractions eluted from Superose 12. If the activity of the fusion protein were more stable, this might provide a new way to study L-SD activity.

The fusion protein indeed proved to be fairly stable. When incubated at 4°C for 24 hours, a preparation lost about half of its original activity (Table 8). This is rather more stable than L-SD in crude extracts, though not remarkable. In any case, it was possible to do a variety of studies with this protein.

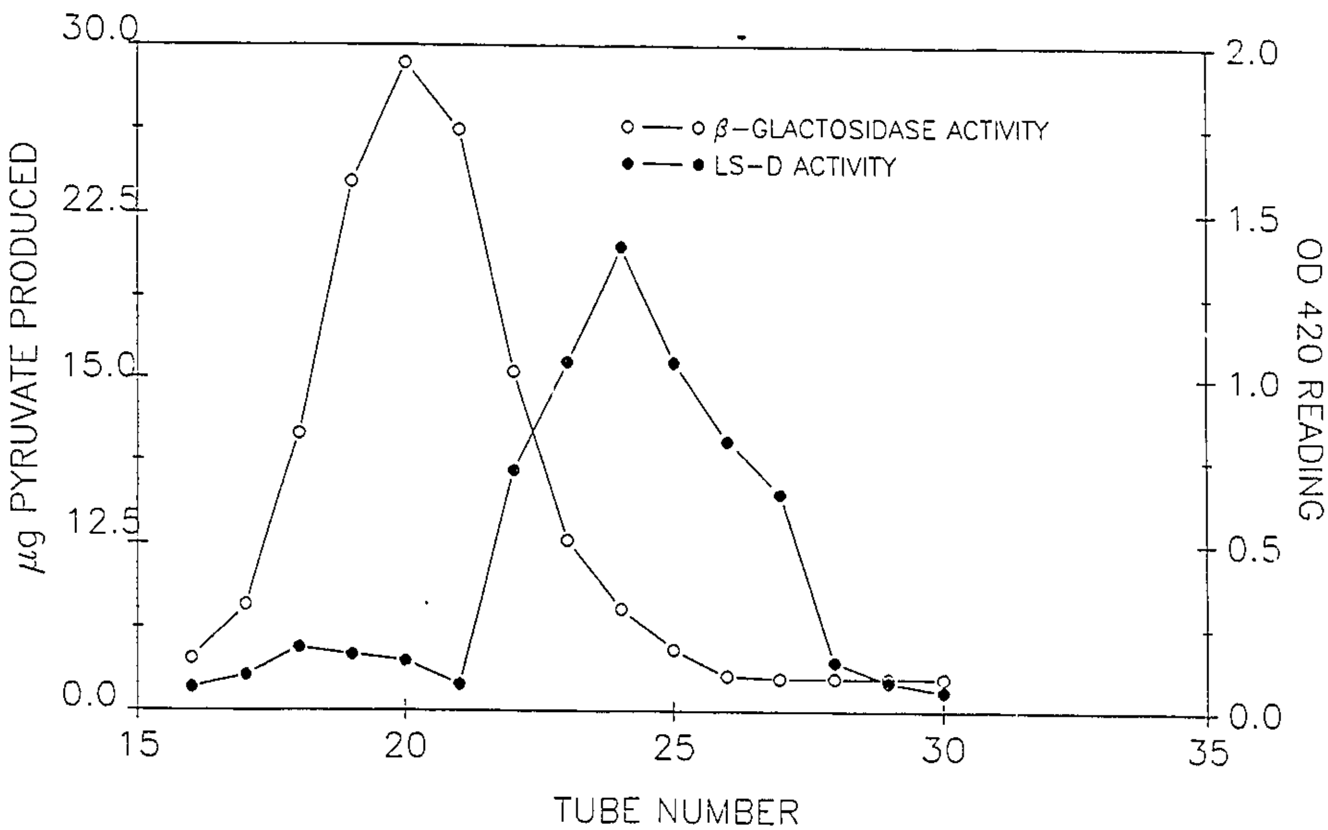

Fig. 16 FPLC analysis of the three-part fusion digested with collagenase

30 μg of 3PF was digested with 0.15U of collagenase at 30°C for 30 min. The digestion mixture was loaded and assayed as indicated in Fig. 14.

Table 8 An estimate of the stability of fusion proteins

Fusion protein	β-galactosidase activity[a]		L-SD activity[b]	
	0 hour	24 hour later[c]	0 hour	24 hour later
2PF	1,200	840	6,980	4,750
3PF	266	190	330	230

a). β-Galactosidase activity is expressed in units/μl.

b). L-SD activity is expressed as μMol pyruvate-produced/μl in 40 minutes reaction.

c). The eluted proteins were stored in 10% glycerol at -70°C. For this experiment, aliquots were thawed, assayed immediately and after incubation in the refrigerator(4°C) for 24 hours.

Various L-threonine deaminating enzymes [Umbarger, 1973] have fairly broad substrate specificity and are able to deaminate L-serine too. Because L-SD has never been sufficiently purified, it was never possible to determine its substrate specificity, and differentiate it totally from related enzymes.

To approach this question, I tested the ability of the 2PF to produce keto acid from L-threonine, glycine, L-leucine, L-aspartic acid, L-methionine, L-alanine, D-alanine and D-serine, using the assay conditions for L-SD. No deamination was seen for any of these putative substrates. This indicates that either the L-SD activation system, or its deaminating ability is very specific in its substrate requirements. This question cannot be answered better at this stage in our understanding of the enzyme.

5-2. Activation Study of L-SD and 2PF

L-SD is synthesized in an inactivated form and requires the action of at least 3 gene products to activate it [Newman *et al.*, 1985a, Newman *et al.*, 1985b]. Strains MEW128, MEW191 and MEW84 each carry mutations in a different one of these 3 genes, and each mutation causes production of an inactive form of L-SD from the chromosomal gene. However the enzyme formed from a high copy number plasmid carrying the *sdaA* gene did produce considerable L-SD activity 40-60 units (Table 9) but much less than the strain without these mutation (MEW28 with pMES22 produce 360 units).

The fact that strain with the 2PF plasmid produces much more activity in activation-competent cells shows that the activation system also works on the 2PF protein *in vivo*.

Table 9. The effect of defects in the L-SD activation system on expression of L-SD activity from plasmid-encoded genes[a]

			L-SD activity		β-Gal. activity
			Grown in glucose-minimal medium		
Expt.	Host	Plasmid	-	+	+
1	MEW28(*sdaA*::Cm^r)	None	2	2	ND[b]
2	MEW28	pMES22(*sdaA*)	360	ND	ND
3	MEW28	pMES27(2PF)	ND	13	3,500
4	MEW56(*sdaA*::Cm^r *sda128*)	None	1	1	ND
5	MEW56	pMES22	46	ND	ND
6	MEW56	pMES27	ND	1	2,300
7	MEW57(*sdaA*::Cm^r *sda191C*)	None	1	1	ND
8	MEW57	pMES22	57	ND	ND
9	MEW57	pMES27	ND	1	2,600
10	MEW58(*sdaA*::Cm^r *sda84*)	None	1	1	ND
11	MEW58	pMES22	50	ND	ND
12	MEW58	pMES27	ND	1	ND

a). Experiments were carried out as Table 2, Antibiotic was added to cultures of strains carrying plasmids. The data shown are the average of three experiments. Cells were grown in minimal medium with (+) and without (-)glycine and L-leucine.

b). ND. Not determined.

Thus the 2PF gene product can function to a limited extent in the activation system defective mutants, but is still subject to activation.

PART 6. AN EXAMINATION OF THE POSSIBILITY THAT THE *sdaA* GENE IS REGULATED BY THE SOS SYSTEM

The L-SD activity of strain MEW1 was induced by DNA-damaging agents and increased temperature, as described by Newman *et al* [Newman *et al.*, 1982a]. Induction has also been seen by ethanol shock (Newman laboratory, unpublished results). One would expect then that β-galactosidase from an *sdaA::lacZ* fusion would also be induced by these agents, but it was not (Table 2). This could be due to the fact that the DNA-damaging agents also cause induction of the λplacMu, and this may affect synthesis of β-galactosidase.

To avoid the complications due to phage induction, I tested the effect of both UV-irradiation and heat shock on synthesis of β-galactosidase from the fusion of *sdaA-lacZ* in plasmid pMEZS22 (Table 10).

This experiment suggests that the *sdaA* gene on the plasmid was induced by UV-irradiation, but not by heat shock. Whether the plasmid was carried by strain MEW1 or strain MEW28, the β-galactosidase level was increased about 4-fold after irradiation. The effect of heat shock was tested only in strain MEW28 and no significant change in enzyme activity was seen (Table 10).

It is clear from the preceding that while L-SD is clearly induced by heat shock and ethanol in some way or another [VanBogelen *et al.*, 1987], a transcriptional effect could

Table 10 The effect of UV irradiation and high temperature on synthesis from the *sdaA* promoter[a]

Expt.	Growth condition	MEW1(pMEZS22)	MEW28(pMEZS22)
1	Glucose	800[b]	390
2	UV irradiation	3,500	1,400
3	42°C growth	ND[c]	440

a). Experiments were carried out as in Table 2, using antibiotic for strain carrying plasmid. These data are the average of three experiments.

b). This data was taken from Table 5.

c). ND: not determined.

not be demonstrated. What other mechanism might be responsible for the induction of L-SD by an increase in temperature? If heat shock affected the L-SD activation system, e.g. by increasing the efficiency of converting an inactivated form of L-SD to an active L-SD, this would not be indicated by a β-galactosidase assay. To test this possibility, I made extracts from strain MEW1 grown at 37°C with and without UV-irradiation, and cells grown at 42°C, and assayed L-SD (Table 11).

Table 2 shows that L-SD as judged by the whole cell assay was induced 7-fold by UV-irradiation, and 2-fold by growth at 42°C. The results for crude extracts parallel this closely, showing a 5-fold induction by UV irradiation and 2-fold by growth at 42°C. This indicates that heat shock regulation affects enzyme synthesis.

The fact that L-SD activity is induced by DNA-damaging agents does not necessarily mean that the induction is controlled by the SOS system. Indeed the fact that an SOS consensus sequence for *lexA* protein binding site could not be found suggests that it is not regulated in this way. Newman *et al* [Newman *et al.*, 1982a] showed that the L-SD activity of a *lexA* mutant was twice as high as that of the parent strain. To confirm that further, I constructed strains MEW1 and MEW28 each carrying a *recA* mutation and assayed the L-SD activity and β-galactosidase synthesized from the *sdaA* promoter (Table 12).

The *recA* mutation causes a defective SOS response because no *recA* product is made and so the *lexA* protein is not cleaved [Walker, 1984]. This mutation decreased considerably the effect on *sdaA* expression. In strain MEW1 *recA* L-SD induction was about 2-fold, i.e. less than the 8-fold increase seen in strain MEW1 itself. A similar

Table 11 The effect of UV irradiation and high temperature on synthesis of the inactivated form of L-SD[a]

		L-SD assay in	
Expt.	Growth condition	Whole cells[b]	Extracts
1	Minimal medium	19	28
2	UV irradiation	130	151
3	42°C	34	51

a). Cells were grown, as in Table 2, and extracts prepared as in Fig. 8. The data are the average of three experiments.

b). The data was taken from reference [Newman *et al.*, 1985a].

Table 12 Effect of a *recA* mutation on expression from the *sdaA* promoter[a]

Expt.	Strain	Relevant Genotype	L-SD activity A-37°C	L-SD activity B-UV	β-Galactosidase activity C-37°C	β-Galactosidase activity D-UV
1	MEW1	Reference strain	19[b]	130[b]	ND[c]	ND
2	MEW59	MEW1 *recA*::Tn10	17	30	ND	ND
3	MEW59	with pMEZS22	13	26	290	350
4	MEW28	MEW1 *sdaA*::Cm^r	2[d]	2[d]	ND	ND
5	MEW60	MEW28 *recA*::Tn10	2	5	ND	ND
6	MEW60	with pMEZS22	2	5	260	290

a). Cells were grown in glucose minimal medium with appropriate antibiotics at 37°C as described in Table 2, footnote a. The data are the average of at least three determinations. Cells used for assay in columns B and D were treated with UV irradiation as in Table 2, Expt.4. Assays were performed and expressed as in Table 2.

b). These data are taken from reference [Newman *et al.,* 1985a].

c). ND. Not determined.

d). These data are taken from Table 13.

decrease in induction of β-galactosidase synthesized from the *sdaA* promoter after UV irradiation was also caused by *recA*. Together with the earlier finding that the *lexA* mutation also affected L-SD regulation, this all suggests that the *sdaA* promoter is under the regulation of the *recA-lexA* SOS system.

PART 7. CONSTRUCTION OF A STRAIN CARRYING AN *sdaA* NULL MUTATION AND DEMONSTRATION OF A SECOND L-SERINE DEAMINASE

In this part of the results, I describe how the *sdaA* null mutation referred to earlier in this work was made, and use it to demonstrate the existence of a second L-serine deaminating enzyme, L-SD#2, in *E. coli*.

7-1. Construction Strain of MEW28 Carrying a Stable Null Mutation in *sdaA*

To determine whether a second L-serine-deaminating activity, L-SD#2, might exist, I wished to completely eliminate L-SD#1, by making a stable null mutation in the gene coding for L-SD#1, i.e. *sdaA*, and then determine whether the strain could show any further L-SD activity, and under what conditions.

To do that, I transformed plasmid pMES23, carrying a chloramphenicol resistance cassette inserted into the *Hpa*I site of the *sdaA* coding region into a DNA-polymerase I-defective strain A401, and selected chloramphenicol-resistant (Cm^r) transformants. Since pMES23, a *colE*1-derived plasmid, cannot replicate in a DNA polymerase-deficient strain [Russel and Holmgrem, 1988], any Cm^r colonies should have the plasmid integrated into the chromosome. Four such Cm^r colonies were also resistant to ampicillin, suggesting that pMES23 had integrated into the chromosome as expected [Russel and Holmgrem, 1988, Winans *et al.*, 1985]. I then transduced from A401 *sdaA*::pMES23 into MEW1, selecting Cm^r colonies, and screened for ampicillin-sensitive strains without L-SD activity (SGL$^-$). One such Cm^r ampicillin-sensitive SGL$^-$ strain, NEW28, was used for further study.

If the gene disrupted is in fact *sdaA*, the 2.6 kb *Pst*I-*Sal*I fragment hybridizing to an *sdaA* probe would be missing in a digest of MEW28 DNA, and would be replaced by a 4.0 kb fragment. The change in the size of the *Pst*I-*Sal*I fragment seen in Fig. 17, thus verifies that mutant MEW28 carried a Cm^r cassette, inserted into the *Hpa*I site of the chromosomal *sdaA* gene.

7-2. L-SD Activity in a Strain Carrying an *sdaA* Null Mutation

It is clear that strain MEW28 does not make L-SD in minimal medium, even when grown in conditions which induce L-SD in the parent strain (Table13 lines 1-4). Similar results were previously reported for a strain carrying λplacMu inserted in *sdaA* (Table 2). This was not due to a failure to activate L-SD since extracts of cells grown in these media also had no significant L-SD activity (Fig. 8 curve MEW22, Fig. 13 curve control).

Nonetheless strain MEW28 showed a great deal of L-SD activity when grown in LB, (Table 13 line 5). Since MEW28 carries a null mutation in *sdaA*, and *sdaA* is the structural gene for L-SD#1, this activity must come from a second L-SD gene, coding for a new enzyme, L-SD#2.

Fig. 17 Hybridization of chromosomal DNA of mutant MEW28

Chromosomal DNA from strain MEW22, MEW28, and MEW1 were isolated and digested with *Pst*I and *Sal*I, and then were electrophoresed on a 1% agarose gel. The gel was dried and hybridized with the ^{32}P-oligo-random-labelled *Sal*I-*Bam*HI fragment derived from pMES22, as suggested by R.K. Storms. Lane A: MEW22 *sdaA*::λplacMu9, showing hybridization signals at 3.1 and 5.7 kb bands as in Fig. 9; Lane B: MEW28 *sdaA*::Cm^r, showing a 4.0-kb signal, corresponding to the 2.6-kb fragment carrying *sdaA* with 1.4-kb *cat* gene; Lane C: MEW1 *sdaA*$^+$, with the main hybridization signal at the expected size of 2.6-kb.

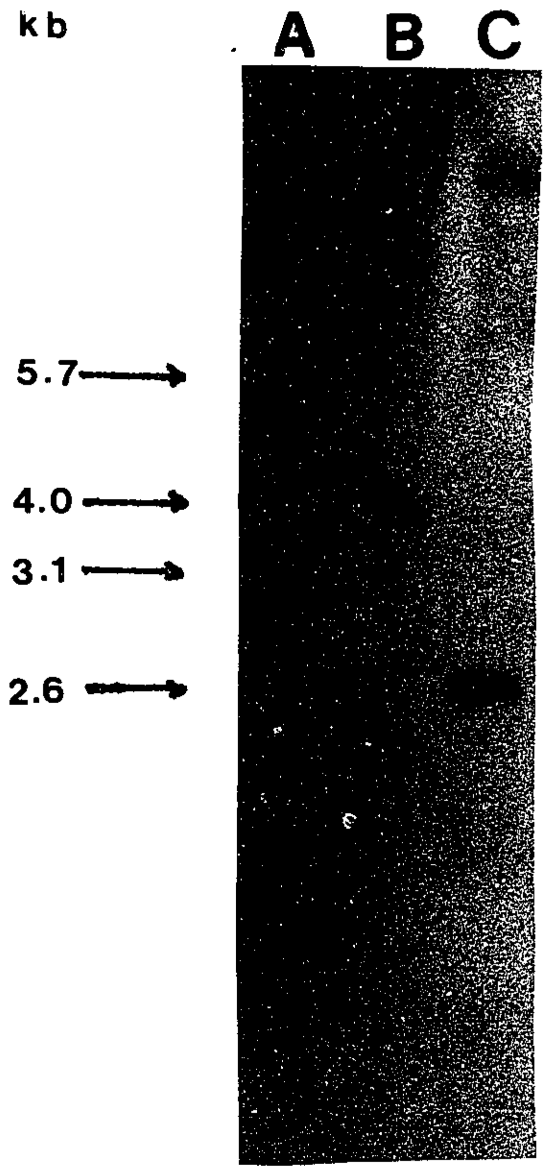

Fig. 17 Hybridization of chromosomal DNA of mutant MEW28

Table 13 L-SD activity of a strain carrying a null mutation in *sdaA*[a]

Expt.	Growth conditions	L-SD activity in strains		
		MEW28	MEW22[b]	MEW1[c]
1	37°C	2	1	9
2	Glycine+ L-leucine	2	7	110
3	UV irradiation	2	2	130
4	42°C growth	2	8	34
5	LB	29	65	230

a). Experiments were performed and expressed as described in Table 2. The data for strain MEW28 are the average of three experiments.

b). These data were taken from Table 2.

c). These data were taken from reference [Newman *et al*, 1985a].

PART 8. CHARACTERIZATION OF L-SD#2

Studying L-SD#2 was made difficult by the fact that this enzyme was not expressed in minimal medium, but only in LB. A deficiency in L-SD#1 and #2 both would not hinder growth in LB, and so I did not have growth conditions in which a loss of L-SD#2 would deter growth.

If the regulation of L-SD#2 were changed so as to permit its synthesis in minimal medium, I could look for mutants deficient in L-SD#2, and clone the gene coding for it by the same system as used for L-SD#1. The use of such a strategy is described in the next section.

8-1. Mutants with Altered Regulation of L-SD#2 Synthesis

The gene coding for L-SD#2 is clearly not expressed in minimal medium (Table 13). Therefore, strain MEW28 cannot synthesize either L-SD enzyme in minimal medium, and therefore cannot grow in medium with L-serine, glycine and L-leucine (SGL-medium) as carbon sources. In order to establish L-SD#2 synthesis in minimal medium, I isolated SGL-using derivatives of MEW28 *sdaA*::Cm^r by plating UV-irradiated cells on SGL medium containing chloramphenicol. I transduced from SGL-using strains into MEW28 to avoid other possible mutations, saving one transductant for further study as MEW50 (*sdaA*::Cm^r *sdaX*).

Strain MEW50 did show L-SD activity, albeit low, when grown in glucose-minimal medium (Table 14). This was induced further when the strain was grown with inducers (Table 14 line 2), and the cells grew correspondingly well in SGL-minimal medium.

Table 14 Synthesis of L-SD as affected by the *sdaX* and *sdaB* mutations[a]

	Growth	L-SD activity		β-Galactosidase activity
Expt[b].	condition	MEW50	MEW51	MEW51
1	37°C	6	1	45
2	Glycine+L-leucine	24	1	100
3	UV irradiation	8	1	40
4	42°C growth	8	1	20
5	LB medium	95	2	375

a). This experiment was performed, and the results expressed, as in Table 2.

b). The data in experiments 1, 2 and 5 are the average of three determinations. The data of experiments 3 and 4 are the average of two determinations.

However strain MEW50 could not use L-serine as sole carbon source without glycine and L-leucine, presumably because its level of L-SD was too low [Newman *et al.*, 1982b].

The function of the *sdaX* gene is not known. The mutation in *sdaA* clearly resulted in a change in the regulation of synthesis of L-SD#2. Whereas its parent has no significant L-SD activity when grown in glucose-minimal medium, the *sdaX* mutant showed slight activity in that medium, and considerable activity with inducers, glycine and L-leucine. When grown in LB, the *sdaX* mutant showed considerable L-SD activity, 95 units (Table 14 line 5), or about 2 times more than its parent MEW28 (Table 13 line5), but much less than the strain from which all our strains are derived, MEW1 (Table 13, 230 units). Thus the *sdaX* mutation changes the basal level of synthesis of L-SD#2 but leaves it subject to some of its usual inducers. One might suppose that the *sdaX* mutation is in the structural gene for L-SD#2 particularly in its regulatory region, or in a regulatory gene separate from it.

8-2. Map Position of *sdaA* and *sdaX*

Since the *sdaX* mutation affects L-SD#2, it might be distinguishable from *sdaA* by its map position. I therefore mapped *sdaA* and *sdaX*, using the Hfr and transduction mapping kit constructed by Singer *et al* [Singer *et al.*, 1989].

Conjugation results showed that the SGL$^-$ phenotype associated with *sdaA*::Cm^r was linked 14% with a Tn10 at 27 min in strain 5054 and 44% with a Tn10 at 42 min in strain 5055. Similarly the SGL$^+$ phenotype associated with *sdaX*, was linked 32% with the Tn10 at 42 min, and 64% with a Tn10 at 65 min in strain 8209. Closer mapping was done by P1 transduction using strains carrying Tn10 inserts in the areas identified by conjugation. This showed that *sdaA*::Cm^r was 40% cotransducible with an insert at 40.75 minutes in strain 18486, and 74% cotransducible with an insert at 41.25 minutes in strain 12608. However the SGL-using ability of strain MEW50 was 53% cotransducible with the insert at 60.5 minutes in strain 12079. It seems then that *sdaA* and *sdaX* are different genes, located 20 minutes apart.

PART 9. ISOLATION OF MUTANTS DEFICIENT IN L-SD#2 ACTIVITY AND THEIR USE IN STUDIES OF REGULATION OF ITS SYNTHESIS

If the L-SD activity in mutant MEW50 is synthesized from a second gene different from *sdaA*, it should be possible to isolate L-SD#2-deficient mutants simply by isolation of SGL^- mutants in strain MEW50. In this part of the results section, isolation of strains carrying insertion mutations in this second gene, *sdaB*, will be described.

9-1. Isolation and Characterization of Strains Carrying Insertions in *sdaB* Gene

The *sdaX* mutation of strain MEW50 could be in the structural gene coding for L-SD#2 or in a regulatory gene influencing its synthesis. Inactivating either of these should result in a loss of L-SD#2 activity. That is, using strain MEW50 in which only *sdaB* is intact, it should be possible to isolate SGL^- insertion mutants which prevent synthesis of an active form of L-SD#2.

MEW50 derivatives carrying λplacMu insertions were isolated by infecting with λplacMu, and then using first an ampicillin selection against strains able to grow on SGL medium and next, a selection on LB plates with kanamycin. The kanamycin-resistant colonies were then screened with respect to their ability to grow on SGL medium. Of 20 cultures from individual kanamycin-resistant colonies grown in LB, two showed no L-SD activity. Those two colonies were also deficient in L-SD activity when grown in glucose-minimal medium with or without glycine and L-leucine. From these strains, which I presumed to be *sdaA*::Cm^r *sdaX* *sdaB*::λplacMu, I transduced one of those inserts in *sdaB* into MEW50, selecting for kanamycin-resistance due to the λplacMu insert. One such transductant, strain MEW51, was studied further.

The gene into which the λplacMu is inserted is defined as *sdaB*. Strain MEW51 must carry three mutations, *sdaA*::cm^r, *sdaX* and *sdaB*::λplacMu. The two insertions result in the strains being unable to make L-SD either in minimal medium or in LB medium (Table 14 line 5). I conclude that the λplacMu insert prevented synthesis of L-SD#2, and that the double mutant, MEW51, grown in LB or minimal medium, had no other enzyme which could deaminate L-serine in our assay conditions.

9-2. β-Galactosidase Synthesized from the *sdaB* Promoter

If strain MEW51 carried an insertion in the *sdaB* gene, and if *sdaB* were the structural gene for L-SD#2, β-galactosidase activities measured in strain MEW51 should give a pattern of regulation similar to that of L-SD activity in strain MEW50. The results of such an experiment are listed in Table 14.

In fact, β-galactosidase activity in strain MEW51 was induced by glycine and L-leucine, as was L-SD activity in MWE50. It was induced to a much higher level during growth in LB, as expected for L-SD#2. It mimics the L-SD#2 synthesized from *sdaA*::Cm^r *sdaX* mutation. That suggested that the λplacMu insertion is in the structural gene of L-SD#2.

9-3. Mapping of the *sdaB* Mutation.

One of the possible interpretations of the *sdaX* mutation might be that it alters the promoter region of the *sdaB* gene. If that is the case, one might expect that the *sdaB* insertion mutant would be in the same map position as the *sdaX*.

By doing conjugation and P1 transduction experiments, again using the Singer kit [Singer *et al.*, 1989], I showed 38% linkage of the *sdaB*::λplacMu insertion with the insert at 60.5 minutes in strain 12079. This is essentially the same as the 53% linkage between the *sdaX* mutation and this insert. This difference in linkage may not be significant, or it may reflect the large size of the λplacMu insertion. These results are consistent with the *sdaX* mutation being located in the promoter region of *sdaB* gene.

9-4. Regulation of L-SD#2 Synthesis by the *lrp* Gene

A pleiotropic regulator of *E. coli* metabolism, the *rbl* gene, since renamed *lrp*, was described recently [Lin *et al.*, 1990]. The *lrp* gene product represses synthesis from *sdaA*, and consequently L-SD#1 activity is greatly induced in the *lrp* mutant [Lin *et al.*, 1990]. To test whether L-SD#2 synthesis is also regulated by *lrp*, I transduced *lrp* into the strains indicated in Table 15 , and measured activity of L-SD and β-galactosidase.

Table 15 **The effect of the *lrp* mutation on expression from the *sdaB* promoter**[a]

Expt.	Strains	Relevant Genotype	L-SD activity 37°C	with inducers[b]	LB	β-Galactosidase activity 37°C	with inducers	LB
1[d]	MEW28	*sdaA*::Cm[r]	2	2	29	ND[c]	ND	ND
2	MEW49	MEW28 *lrp*	2	4	89	ND	ND	ND
3[d]	MEW50	MEW28 *sdaX*	6	24	95	ND	ND	ND
4	MEW52	MEW50 *lrp*	22	17	89	ND	ND	ND
5[d]	MEW51	MEW50 *sdaX* *sdaB*::λplacMu	1	1	2	45	100	375
6	MEW53	MEW50 *lrp*	3	1	1	50	50	675

a). Activities were assay as in Table 2. The data are the average of three determinations.

b). Inducers are glycine and L-leucine which were added into the growth medium.

c). ND. Not determined.

d). The data in experiment 1 was taken from Table 13; experiment 3 and 5, from Table 14.

Table 15 shows that the *lrp* mutation does alter synthesis of L-SD#2 though most effects are rather small. In the *sdaA*::Cm^r mutant, the *lrp* mutation caused a considerable increase in L-SD activity expressed in LB (Table 15). When the *sdaX* mutation was introduced, the *lrp* mutation had a different effect, inducing increased synthesis in minimal medium, decreasing synthesis in the presence of inducers (Table 15 line 3,4) but showing no further effect on LB-grown cells. One might expect that β-galactosidase of strain MEW53 would be affected in the same way, and the reduction in the presence of inducers was in fact seen (line 5 vs line 6). However no induction was seen in minimal medium, whereas the enzyme activity was almost doubled in the *lrp* strain grown in LB medium.

It seems clear that the *lrp* gene affects synthesis of L-SD#2. However these results cannot be explained in detail.

PART 10. STUDY OF BIOCHEMICAL CHARACTERISTICS L-SD#2

In this part of the results section, I described some aspects of L-SD#2, such as enzyme activity at different pHs, and *in vitro* activation.

10-1. Activity Parameters of L-SD#2

The characteristics of L-SD#1 have been described to some extent. To see whether L-SD#2 has similar aspects, or whether it is very different, I made a preliminary characterization of L-SD#2, using the whole cell assay in conditions established for L-SD#1 [Isenberg and Newman, 1974].

In fact, these two enzymes proved to be extremely similar in their assay requirements. The L-SD#2, like L-SD#1, showed a broad pH optimum, with very little difference between pH 7 and pH 9 (Table 16). Neither showed activity at low pH, (below pH 6), and both did show enzyme activity up to pH 10 at least. Both enzymes also showed a high substrate requirement (Fig. 18) and the assay was linear in both cases for 40 min at 37°C (Fig. 19).

10-2. Activation of L-SD#2 by Iron and Dithiothreitol

The assay of L-SD#1 in extracts requires addition of iron and dithiothreitol(DTT). That this is also true for L-SD#2 is seen in Fig. 20. Extracts of cells expected to show L-SD#2 activity as judged in the whole cells showed no activity without iron and DTT. When these were added, the rate of the reaction increased gradually, corresponding to a relatively slow activation of L-SD#1 reported earlier [Newman *et al.*, 1985a,].

That this activity was due to the gene affected by the *sdaB* mutation was shown by assaying an extract of the *sdaA*::Cm^r *sdaX* *sdaB*::λplacMu triple mutant, MEW51 (Fig. 20). Whereas a strain with a functional *sdaB* gene showed L-SD activity when incubated with iron and DTT, the strain with an insertion in *sdaB* showed no activity. This indicates that once *sdaA* and *sdaB* are both rendered inactive, the extracts contain no other L-serine deaminating activity that can be activated by iron and DTT. Similarly, in an extract of strain MEW28 grown in glucose minimal medium, this being an *sdaA*::Cm^r strain in which L-SD#2 was not established by the *sdaX* mutation, incubation with iron and DTT did not produce L-SD activity (Fig. 20).

Table 16 L-SD activity in different pH buffers[a]

pH	A: L-SD#1		B: L-SD#2	
	Phosphate buffers	Tris buffers	Phosphate buffers	Tris buffers
6.0	-[b]	ND[b]	0.5	ND
6.5	+[b]	ND	10	ND
6.8	+[b]	ND	10	ND
7.1	+[b]	ND	11	ND
7.3	+[b]	ND	12	ND
7.5	+[b]	ND	12	ND
7.8	4.6	3.1	14	13
8.1	4.3	3.4	14	11
8.4	ND	3.4	ND	15
8.7	ND	3.1	ND	14
9.0	ND	3.4	ND	11
9.3	ND	4.2	ND	17
9.6	ND	4.4	ND	15
9.9	ND	4.5	ND	18

a). Cells (MEW1 in glucose-minimal medium for L-SD#1 and MEW50 in LB medium for L-SD#2) were grown and subcultured at 37°C and resuspendend to 100 K.U. in buffer at pH indicated. 50 mM of phosphate buffer was made by mixing 50 mM dibasic and monobasic potassium phosphate to the pH indicated. 50 mM Tris buffer at the pH indicated was made by adding NaOH to Tris solution as required. The data in this table are the μg of pyruvate produced by a 0.3 ml suspension in 35 min.

b). Data were taken from reference [Isenberg and Newman, 1974] and + and - indicate pyruvate produced or not.

c). ND. Not determined.

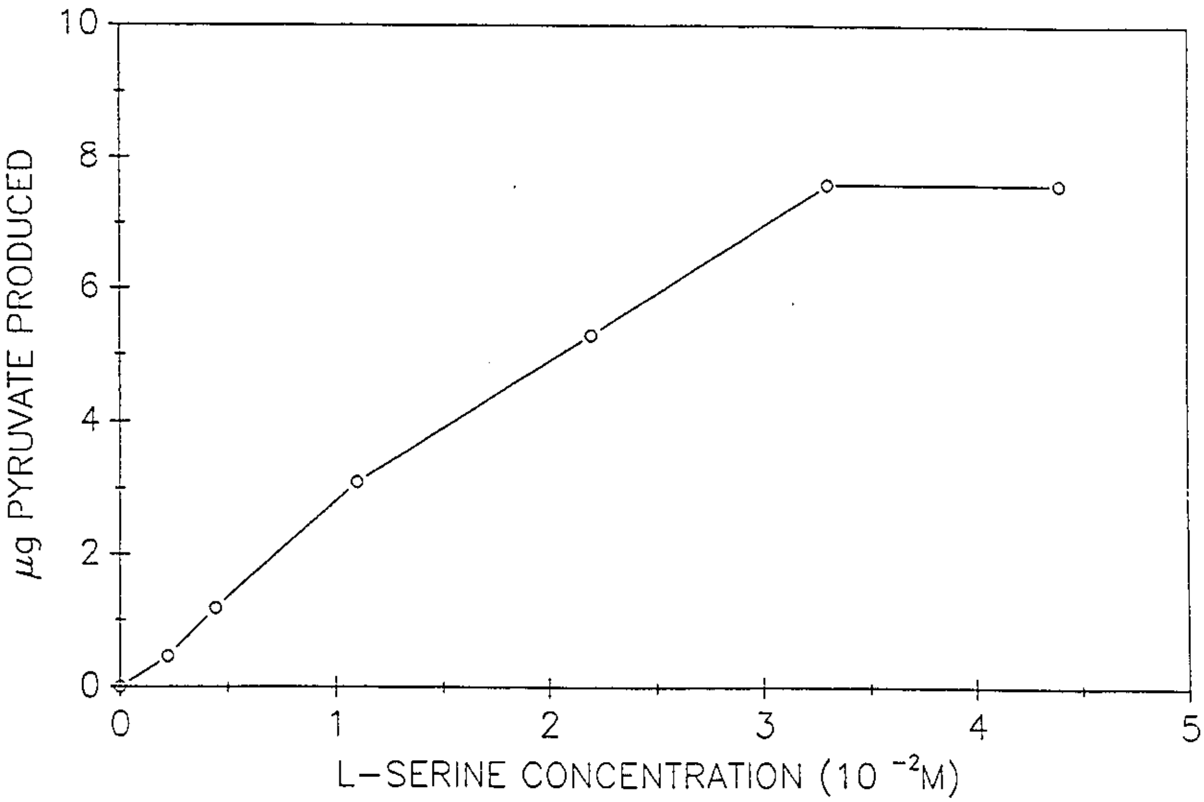

Fig. 18 The effect of L-serine concentration on L-SD#2 activity

Cells of strain MEW28 were grown in LB and assayed as indicated in Table 13, using different concentrations of L-serine as noted. Data are expressed as an average of three experiments.

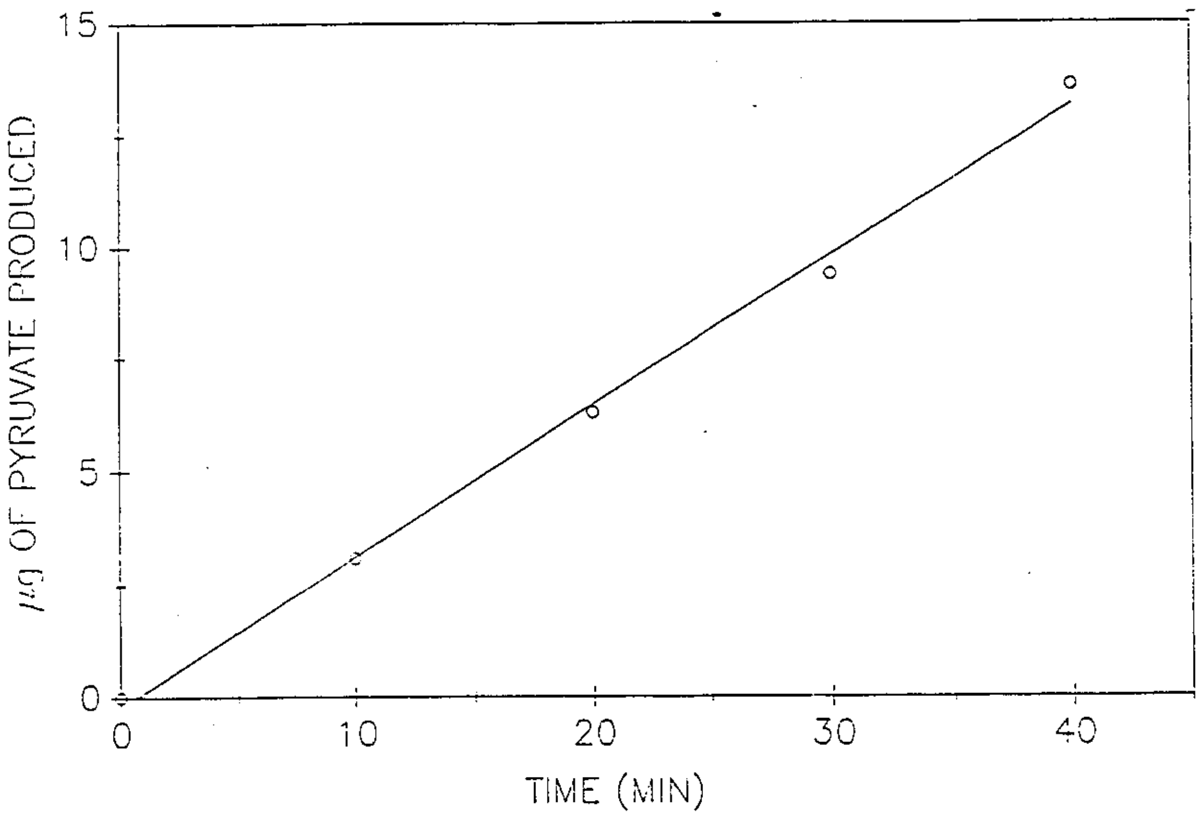

Fig. 19 The activity of L-SD#2 as a function of time

Cells of strain MEW28 were grown and assayed, and results expressed, as in Fig. 18, using different incubation times as indicated on the figure. Data are expressed as an average of three experiments.

It is clear that L-SD#2 is very similar to L-SD#1 in its various substrate parameters, and in its requirement for activation, though it is made from a different coding sequence.

10-3. Activation of L-SD#2 *in vivo*

L-SD#1 is made in the cell in an inactive form, and is activated by a system coded by at least three genes identified by the mutations [Newman *et al.*, 1985a,b] and this study. These activation-deficient mutants are unable to make an L-SD#1 enzyme which can function *in vivo*, and are therefore physiologically SGL$^-$. However they do make a protein which can be activated *in vitro* by iron and DTT, and deaminates L-serine after such treatment.

The fact that L-SD#2 was activated by the same *in vitro* treatment as L-SD#1 suggested that L-SD#2 might also be activated by the same *in vivo* system. If this were so, one might expect a strain carrying both the *sdaA*::Cm^r and the mutation from the activation-deficient mutants to have no L-SD activity *in vivo* when grown in LB medium. Such a set of strains were constructed by transducing *sdaA*::Cm^r from MEW28 into MEW128, MEW191C and MEW84, selecting antibiotic resistance, and used the transductant under the names MEW56, MEW57 and MEW58, respectively.

In the whole cell assay, strain MEW28 grown in LB showed 29 units of L-SD activity (Table 13). Strain MEW56, MEW57 and MEW58, which carried in addition the activation-deficient mutants showed very low activity. It is clear therefore that the mutation in these activation-deficient mutation has a profound effect on both enzymes.

While both L-SD#1 and L-SD#2 required *in vivo* activation, they differed in their *in* *vitro* response to iron and DTT in activation-deficient mutations. L-SD#1 could be activated readily in extracts of strains MEW128, MEW191 [Newman *et al.*, 1985a] and MEW84 (Fig. 8), from which the conclusion was made that L-SD#1 is made in an inactive form. However no activation of L-SD#2 could be demonstrated using the same conditions (Fig. 20. MEW56, similar for MEW57 and MEW58. Data not shown). This may be a difference in the activation conditions needed for L-SD#2, when synthesized in activation-deficient strains. However it may also be that *sdaB* is not transcribed in activation deficient strains and that no inactive L-SD#2 is present. I can not tell at this time whether L-SD#2 is made in an inactive form in these cells, or whether it is not made at all.

PART 11. CLONING OF *sdaB* GENE

It was possible to clone the *sdaB* gene using the mini-Mu replicon produced in a host carrying an *sdaX* mutation by transfecting strain MEW51 *sdaB*::λTn10 *sdaX* and selecting cells which could grow on SGL medium. An 8 kb *Pst*I fragment which hybridized to *sdaB* and showed high L-SD activity was cloned from the strain MEW50 with miniMu-Muts replican. It also hybridized to Kohara phage λ457 [Kohara *et al.*, 1987] which also located the *sdaB* gene at 60.1 min of *E. coli* map.

11-1. Cloning of the *sdaB* Gene from Strain MEW50

Two mutations were isolated in this part of study. A mutation in *sdaX* established synthesis of L-SD#2 in minimal medium. The second mutation, an insert in *sdaB*,

Fig. 20 *In vitro* activation of L-SD#2

Pyruvate produced by 35μg of protein in extracts from LB-grown cells was determined as in Fig. 8 using the following strains: MEW28, MEW50 *sdaA*::Cm^r *sdaX*; MEW56 *sdaA*::Cm^r, deficient in L-SD activation; and MEW51 *sdaA*::Cm^r *sdaX* *sdaB*::λplacMu. Values are corrected for nonenzymatic deamination of L-serine by iron and DTT. Corresponding whole-cell L-SD assays gave 95, 29, 7, and 2 units, respectively. Parallel experiments in the absence of iron and DTT were done in all cases and showed no activity.

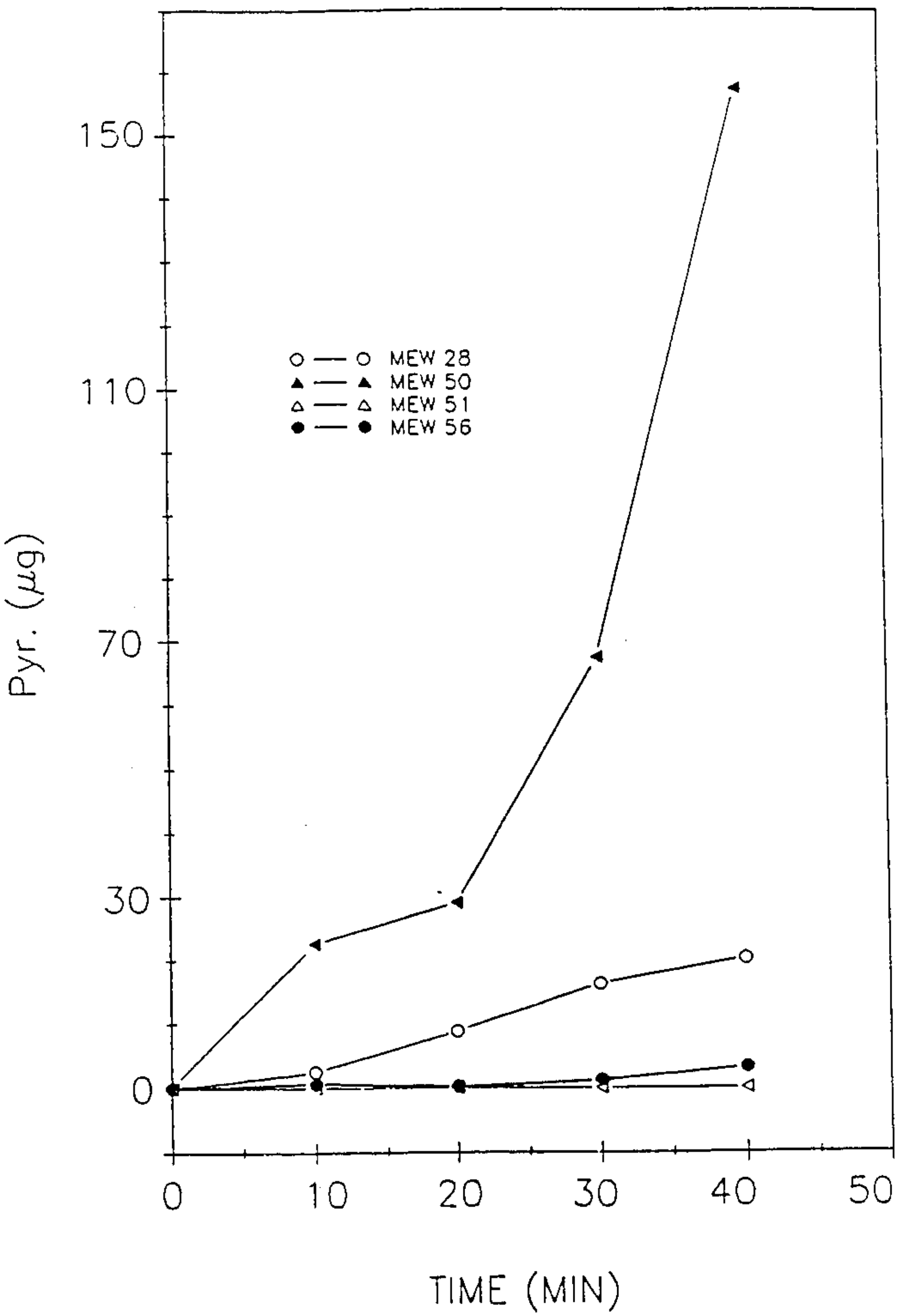

Fig. 20 *In vitro* activation of L-SD#2

destroyed the ability to synthesize L-SD#2. These two mutations could both be in the same gene, or they could be in different genes. In either case, a plasmid carrying a functional *sdaB* gene with the mutation in *sdaX* loci should confer the ability to grow on SGL upon strain MEW51 *sdaA*::Cm^r *sdaX* *sdaB*::λplacMu.

Since the strain MEW51 carried a λplacMu9 insertion specifying kanamycin resistance, it could not be used as a recipient strain for the cloning. Therefore, the λplacMu9 was replaced by a λTn10 insertion by infecting with λTn10 and selecting tetracycline-resistance, taking advantage of the possibility of homological recombination between the two λ phages. The resultant strain MEW55, *sdaB*::λTn10 remained SGL^- and lost kanamycin resistant characteristics and could be used for cloning.

The *sdaB* gene from MEW50 *sdaA*::Cm^r *sdaX* was cloned by using the mini-Mu *in vivo* cloning system. Strain MEW50 was made lysogenic for Mu cts, and a mini-Mu replicon carrying kanamycin resistance was transformed into it. Then a 42°C lysate of that strain was used to transfect into strain MEW55 *sdaX* *sdaB*::λTn10 the ability to grow on SGL in the presence of kanamycin.

In this way, approximately 15 kb of chromosomal DNA was cloned, restoring the SGL^+ (also $NSIV^+$) phenotype to strain MEW51. Since the donor strain carried a null mutation in *sdaA*, and both donor and recipient carried the mutation in *sdaX*, the gene cloned must be *sdaB*, or some other gene which complements an *sdaB* mutation.

From this plasmid, an 8 kb *Pst*I fragment was subcloned into the pBR322 *Pst*I site, forming plasmid pMES41. Plasmid pMES41 produced very high L-SD activity, in strains carrying either the wild-type *sdaX* gene (Table 17 lines 1 and 2) or the mutated one

Table 17 Synthesis of L-SD from Plasmid pMES41[a]

Strain	Relevant Genotype	Plasmid	L-SD activity 37°C	with inducers[b]	LB medium
MEW28[c]	*sdaA*::Cm[r]	None	2	2	29
MEW28		pMES41	119	129	234
MEW51[d]	*sdaA*::Cm[r] *sdaX* *sdaB*::λplacMu	None	1	1	2
MEW51		pMES41	118	150	330

a). Expressed and assayed as in Table 2. The values given are the average of three experiments. Antibiotic was added to the culture of strains with plasmid pMES41.

b). Inducers are glycine and L-leucine.

c). The data in this column was taken from Table 13.

d). The data in this column was taken from Table 14.

(Table 17 lines 3,4). This suggests that the 8 kb fragment carried either the coding gene for L-SD#2, or a gene which activates expression of an otherwise inactive L-SD#2 coding gene.

11-2. Hybridization Studies with pMES41

To characterize the cloned *sdaB* gene in pMES41, the 8 kb *Pst*I fragment of pMES41 was hybridized to *Pst*I digests of chromosomal DNA from three strains: MEW28 *sdaA*::Cm^r, MEW51 with inserts in both *sdaA* and *sdaB*, and parent strain MEW1 (Fig. 21).

The 8 kb fragment hybridized to a single 8 kb band in strain MEW1, and MEW28 indicating that the gene carried on the plasmid is intact in both those strains. In strain MEW51, the 8 kb band was replaced by bands at 6 and 10 kb. Thus the disruption of *sdaB* in strain MEW51 is in a gene carried on pMES41, i.e. the plasmid carries the *sdaB* gene. The fact that the *sdaB* gene DNA fragment did not hybridize to *sdaA* insertion mutation MEW28 further confirms that *sdaA* and *sdaB* gene were located at different sites on the chromosome.

It can also be seen from Fig. 21 that the 8 kb fragment hybridizes weakly to the 2.6 kb *Pst*I-*Sal*I fragment shown to carry the *sdaA* gene. This is also consistent with the existence of the minor hybridizing band in the experiments with the *sdaA* probe (Fig. 9). Since *sdaA* is known to code for L-SD#1, the two hybridization experiments, and the fact that both clones show high L-SD activity suggest that *sdaB* codes for L-SD#2.

The Kohara phage λ457 [Kohara *et al.*, 1987] contains an 8 kb *Pst*I fragment which

Fig. 21 Hybridization of chromosomal and plasmid DNA with an 8-kb *sdaB* gene probe

DNA samples from the sources noted were digested with *Pst*I, except for Lane A for which both *Pst*I and *Sal*I were used. Digests were electrophoresed in an 0.7% agarose gel and hybridized as described in Fig. 17 with a ^{32}P random-oligo-labelled 8-kb *Pst*I fragment derived from pMEW41. Lane A: plasmid pMES22 digested with *Sal*I and *Pst*I; Lane B: plasmid pMES41 (The same amount of DNA was used in Lanes A and B). Lanes C-E: Chromosomal DNA digests from MEW28 *sdaA*::Cm^r(Lane C) and MEW1, the parent strain (Lane E), both showed a single hybridization signal at 8-kb, as did DNA from plasmid pMES41 (Lane B). However, with DNA from MEW51 *sdaA*::Cm^r *sdaX* *sdaB*::λplacMu, this 8-kb band was replaced by two bands at 6 and 10-kb.

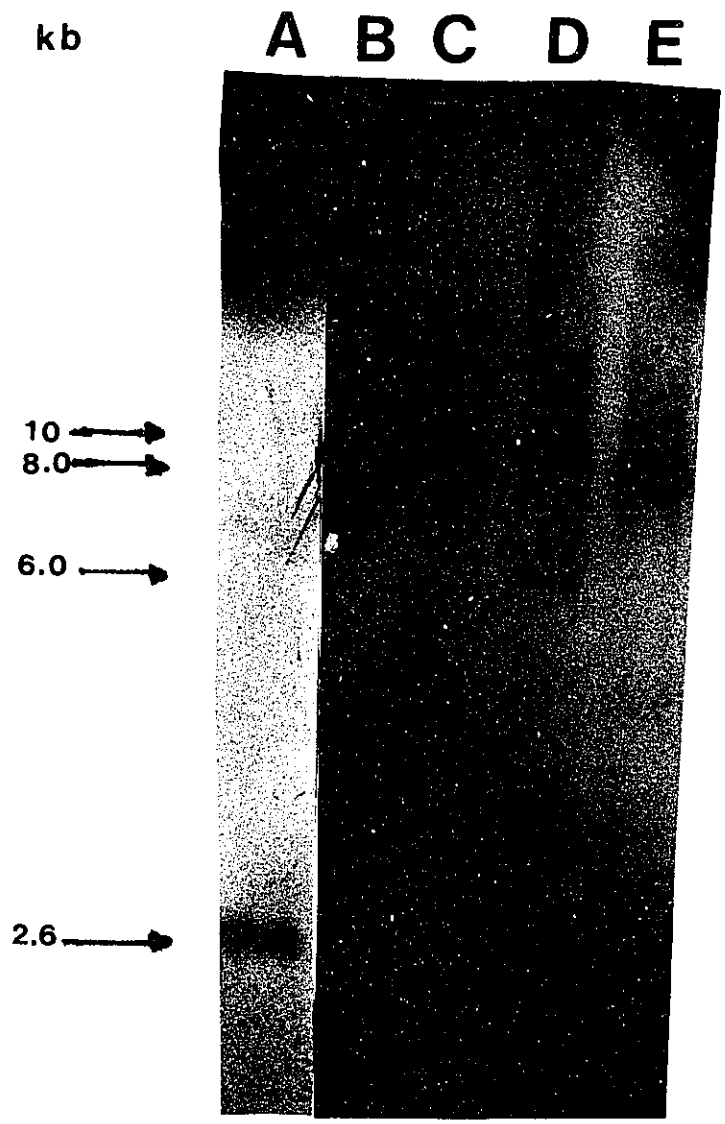

Fig. 21 Hybridization of chromosomal and plasmid DNA with an 8-kb *sdaB* gene

hybridized with the *sdaB* gene, as was shown in Fig. 22. One would expect then that phage 456 would contain hybridizing material, since it carries part of the 8 kb *Pst*I fragment. However I could not demonstrate hybridization with phage λ456 and λ458. The results with phage λ457, taken by themselves, suggest a map position of 60.1 min for the *sdaB* and agreed with the position of *sdaB* insertion determined by transduction. This is also consistent with *sdaB* and *sdaX* affecting the same gene.

PART 12. POSSIBLE METABOLIC FUNCTION OF L-SD

The metabolic function of L-SD remains unclear. Some possible functions of these enzymes may include detoxification of L-serine and use of L-serine as the carbon, energy and nitrogen source.

12-1. Detoxification of L-Serine

That L-serine inhibits growth of *E. coli* was described elsewhere [Cosloy and McFall, 1970, Uzan and Danchin, 1978]. This growth inhibition was due to L-serine inhibiting the homoserine dehydrogenase used in L-isoleucine biosynthesis [Hama *et al.*, 1990]. To investigate whether L-serine is toxic even to the cells grown with L-isoleucine and L-valine, I measured the doubling time of the strains with and without mutations in *sdaA* and *sdaB*, and *sdaA* and *sdaB* mutants with or without the plasmids (Table 18). It was shown that MEW1, *sdaA*, and MEW50, *sdaX sdaB*, grown with L-isoleucine and L-valine could tolerate L-serine at 2 mg/ml, but a much higher concentration of L-serine (10 mg/ml) slowed the growth rate {apparent doubling time (a.d.t.) 89, and 72 min,

Fig. 22 Hybridization of Kohara phage λ456, 457 and 458 with an 8-kb *sdaB* probe

λ DNA from 456, 457 and 458 was isolated and plasmid pMES41 digested with *Pst*I, electrophoresed on 0.7% agrose gel. The hybridization was carried out as in Fig. 21. Lane M: λ *Hin*III markers. Lane A: plasmid pMES41 showing 8-kb hybridization signal; Lane B: λ456; Lane C:λ458 (no hybridization is seen in these lines); Lane D and E: λ457, showing 8-kb hybridization signal. Lane M indicates the position of bands from a λ *Hind*III digest.

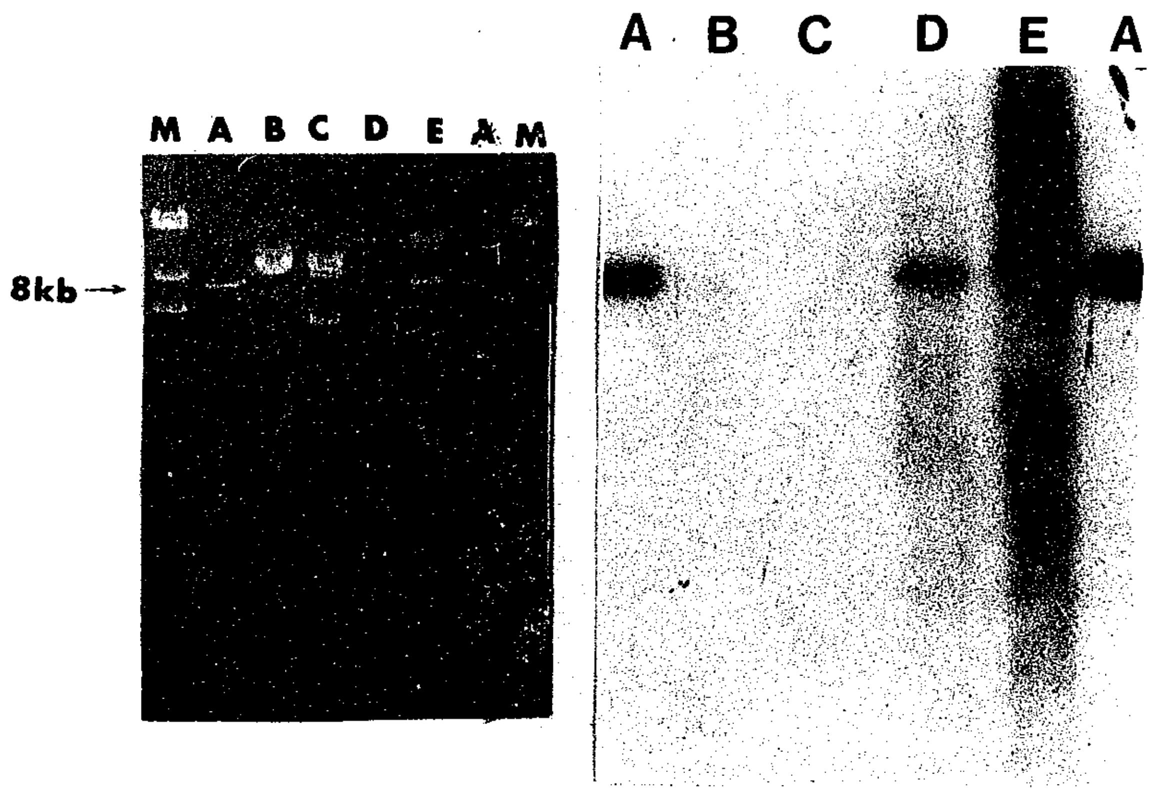

Fig. 22 Hybridization of Kohara phage λ456, 457 and 458 with an 8-kb *sdaB* probe

respectively}. This inhibition could be released by adding glycine and L-leucine, probably due to the induction of L-SD which degraded the L-serine.

The strain with mutations in *sdaA*, or and *sdaB*, showed less resistance to L-serine. Mutants MEW28 (*sdaA*::Cm^r) and MEW51 (*sdaX sdaB*::λplacMu), grew much more slowly in minimal medium with 2 mg/ml of L-serine. Their doubling time was the same as the MEW1 in 10 mg/ml of L-serine. This inhibition could be slightly relleved by adding glycine and L-leucine, from 96 min to 76 min for MEW28 and 94 to 71 min for MEW50. In medium with 10 mg/ml L-serine, these cells grew very slowly (about 170-180 min) and this slow growth could not be relieved by glycine and L-leucine. However, the L-SD activity in these mutations was the major factor in detoxification of L-serine. That was demonstrated by the fact that strains carrying the *sdaA* or *sdaB* plasmid could grow in high L-serine much faster, even faster than the reference strain (48 Vs 62 for MEW28 and 57 Vs 67 for MEW51). It is clear then that the level of L-SD is a major factor in determining the cell's ability to grow in the presence of L-serine.

12-2. Using L-Serine as Carbon and Nitrogen Source

Previous work in this laboratory suggested that the ability of cells to grow with L-serine as carbon and energy source was related to L-SD#1, as is ability to use L-serine as nitrogen source [Isenberg and Newman, 1974]. Because of the lack of a structural gene mutation, the suggestion could not be confirmed [Isenberg and Newman, 1974]. The strains constructed here allowed this confirmation. It is clear that the strains using L-serine as carbon and energy source rely on L-SD activity. Both L-SD#1 and L-SD#2

Table 18 Doubling times of strains grown in the presence of L-serine[a]

		Apparent doubling time (a.d.t.)[b]				
Strains	**Plasmid**	**A-37°C**	**B-S(2)**	**C-S(2)GL**	**D-S(10)**	**E-S(10)GL**
MEW1	None	58	61	ND	89	64
MEW28	None	62	96	76	176	168
MEW28	pMES22	ND[c]	ND	ND	48	ND
MEW50	None	63	60	ND	72	64
MEW51	None	67	94	71	176	178
MEW51	pMES41	ND	ND	ND	57	ND

a). Cells were grown overnight in glucose-minimal medium, subcultured to exponential phase, and then subcultured in the conditions noted into side-arm flasks, equipped for determining turbidity with a Klett colorimeter. Turbidity was noted at 30 minutes intervals using a 460 filter.

b). The cells were grown in glucose minimal medium with the following additions: A, none; B, L-serine 2 mg/ml; C, as B with glycine and L-leucine; D, L-serine 10 mg/ml; and E as D with glycine and L-leucine. Glycine and L-leucine were added at 300 µg/ml each. Antibiotic was added to cultures of strains carrying plasmid pMES22 and pMES41. The apparent doubling time was calculated from semi-log plots of turbidity as a function of time of incubation. The data presented is the average of three determinations.

c). ND. Not determined.

can provide the pyruvate to permit growth on L-serine. Similarly, the mutations MEW128, MEW191 and MEW84 all prohibit growth on L-serine because L-SD activity in these mutants is low, even thought the inactivated form of L-SD exists in the cells.

E. coli can use both glycine and L-serine as nitrogen source [Newman *et al.*, 1976]. The pathway of converting glycine to L-serine and using ammonia produced by the activity of L-SD was suggested [Newman *et al.*, 1976]. To test whether the *sdaA*, and *sdaB* gene products were involved in using L-serine and glycine as nitrogen source, the strains were grown in glucose-minimal medium with a limited amount of ammonium sulfate, the usual nitrogen source, and subcultured into medium with glycine or L-serine without ammonium sulfate. The turbidity of cultures was determined at intervals (Table 19) up to 48 hour. In fact, a strain with *sdaA* coding L-SD#1 could use glycine and L-serine as nitrogen source, as could the strain with a mutated *sdaA* and wild type *sdaB* established by the *sdaX* mutation. *SdaA* or/and *sdaB sdaX* insertion mutations abolished the ability to use L-serine as nitrogen source, and slowed down the rate of growth on glycine as nitrogen source. This suggests that L-SD is used in providing the ammonia from both L-serine and glycine. However, glycine may be used as nitrogen source by other minor routes.

Since high copy number plasmids carrying the *sdaA* or *sdaB* gene can provide enough pyruvate to support growth on L-serine as carbon source, one would expect these plasmids also support growth on L-serine and glycine as nitrogen source. Indeed, the *sdaA* mutant, with either plasmid can use L-serine as nitrogen source much faster than the parent strain MEW1. Similarly, growth with the plasmid permitted cells to grow with

Table 19 Test of using glycine and L-serine as nitrogen and carbon source[a]

Strain and plasmid	Growth condition[b] Glu+NS	Glu-N+G	Glu-N+GL	Glu-N+S	+S	+G
MEW1	++++	++++	++++	++++	ND[c]	ND
MEW28	++++	++	+++	+	ND	ND
MEW28(pMES22)	++++	++++	ND	++++	++++	-
MEW51	++++	++	+++	+	ND	ND
MEW51(pMES41)	++++	++++	ND	++++	ND	-

a). The experiments were carried out as described in Materials and Methods. The data in this table is the average of at least three times measurements. ++++: represent the growth over K.U. 100 within 36 hours; +++: represent the growth over 100 within 48 hours; ++: represent the growth OD around 60 within 48 hours and +: represent the growth OD near 20 reading in a red filter.

b). Growth condition was at 37°C in the medium: +Glu+NS: glucose with Ammonium sulfate and 500 μg L-serine/ml; +Glu+G: glucose with glycine at concentration of 500 μg/ml; +Glu+GL: glucose with glycine and L-leucine at concentration of 500μg/ml and 30μg/ml respectively; +Glu+S: glucose with L-serine at concentration of 500μg/ml; +S: L-serine along at concentration of 10mg/ml; and +G: glycine alone at concentration of 10mg/ml.

glycine as sole nitrogen source as fast as the wild type with L-leucine as inducer. This suggests that in these growth conditions, the growth rate depends on ammonia which is provided by L-SD deaminating L-serine.

The wild type strain of *E. coli* could not use L-serine as sole carbon source but can use it as nitrogen source. The difference between these two metabolic roles is probably the relative amounts of pyruvate and ammonia needed to support cell growth. Then if this is true, one would expect that cells carrying *sdaA* or *sdaB* on plasmids could use L-serine as sole carbon, energy and nitrogen source. In fact, both strain MEW28 carrying pMES22 *sdaA* and MEW51 carrying pMES41 *sdaB* were able to grow on L-serine as sole carbon, energy and nitrogen source as expected. Also, strains carrying pMES22 grew much faster than the strain carrying pMES41. That may be due to a difference in L-SD level or to a difference between L-SD#1 and L-SD#2. Neither of these strains could grow with glycine as sole carbon, energy and nitrogen source.

DISCUSSION

The work in this thesis has been devoted to an understanding of L-serine deamination in *E. coli* K-12. This work has proceeded along three main lines. The first project was the detailed study of the molecular characteristics of the *sdaA* gene, and the use of techniques of molecular genetics to isolate the L-SD gene product. The second project was the proof that a second L-SD enzyme exists in *E. coli* and the cloning of the gene that codes for it. A third less detailed set of experiments deals with the regulation and possible metabolic role of these enzymes.

PART 1. THE STRATEGY OF THIS STUDY

The experimental plan used for these projects is summarized in the first section of the discussion (1-1., 1-2., and 1-3.). Then the experiments are considered in detail, and questions arising from them discussed.

1-1. Studies on the Gene *sdaA* which Codes for L-SD#1

E. coli is known to make an enzyme activity called L-serine deaminase, which can convert L-serine to pyruvate both *in vivo* and *in vitro*. A strain harbouring this activity can grow with L-serine as carbon source, if glycine and L-leucine are also supplied. It should therefore be possible to isolate mutants termed SGL$^-$ which are unable to produce this activity, and unable to grow on this medium. At the time this work began, several SGL⁻ strains were known. However all of the well-characterized mutant strains were defective in enzyme activation, but not in enzyme synthesis. No mutation in the structural gene had been isolated. I therefore began this work by trying to isolate a mutation that would define the structural gene for L-SD, using insertion mutation via the transposable element, λplacMu9. If such an element were inserted in the gene coding for L-SD, one would expect that L-SD activity would disappear, whether assayed *in vitro* or *in vivo*. Of the many insertion strains isolated, one did show such a loss of L-SD activity both *in vivo* and *in vitro*. Because β-galactosidase was induced in that strain by glycine and L-leucine, just as is L-SD in the parent strain, it seemed likely that the insertion was located in the gene, *sdaA*, coding for L-SD. Since this work also showed that a second L-SD is coded by another gene, *sdaB*, the product of *sdaA* is now known as L-SD#1.

The fact that an *sdaA* clone carried on a high copy number plasmid led to very high L-SD activity was consistent with the idea of its being the structural gene for L-SD#1. Moreover, the sequence determined for the *sdaA* gene shows it to code for a 48,000 molecular weight protein, which is about the size suggested for L-SD by gel filtration [Newman *et al.*, 1990].

The direct proof that *sdaA* codes for L-SD#1 was made by experiments which show that a fusion protein, made by fusing the intact *sdaA* gene in frame to *lacZ*, carried both L-SD and β-galactosidase activity. L-SD could be purified by affinity of the fusion protein to an anti-β-galactosidase column, and the two activities comigrated during gel filtration.

Since fusion of *sdaA* to *lacZ* added the L-SD protein to the β-galactosidase molecule, it is obvious that *sdaA* must code for L-SD#1. This also allowed a simple purification of L-SD. Using the same sort of fusion, but this time with an intervening sequence coding for a collagenase-sensitive polypeptide, I was able to purify the *sdaA* gene product, L-SD. This is the first time L-SD has been extensively purified in *E. coli* (earlier reports seem to refer to a different enzyme [Alfoldi *et al.*, 1968, Alfoldi and Rasko, 1969]).

1-2. Definition of a Second L-Serine Deaminase, L-SD#2, and Genes *sdaB* and *sdaX* Affecting Its Synthesis

It is clear then that *sdaA* codes for L-SD#1, and a strain with an insert in *sdaA* showed no L-SD activity when grown in glucose minimal medium. However such a strain (*sdaA*::λplacMu or *sdaA*::Cm^r) did show L-SD activity when grown in LB medium. This suggested that there might be a second L-SD enzyme in *E. coli*, a suggestion that had been made previously [Newman *et al.*, 1985a].

Various enzymes have been reported to deaminate L-serine as a side-reaction, and it has been shown that *E. coli* can use this side-reaction of cystathionase as a way of obtaining carbon and energy from L-serine. I was therefore interested to determine whether *E. coli* actually has a second true L-serine deaminase, which is made in Luria broth, but not in minimal medium.

If such an L-SD#2 existed, it should be possible to isolate a mutant in which the regulation of its synthesis was altered to permit its synthesis in minimal medium. I therefore selected from a strain carrying *sdaA*::Cm^r, a derivative able to grow on SGL medium, with chloramphenicol, ascribing this to a change in *sdaX* affecting its regulation. Assuming that the *sdaX* strain had altered the regulation of L-SD#2 synthesis, and that this enzyme functions to deaminate L-serine, I could use the same protocol to clone *sdaB* as I did for *sdaA*.

I therefore mutated *sdaB* by λplacMu insertion, cloned the *sdaB* from the *sdaX* strain by miniMu cloning, and showed that the clone hybridized with the gene interrupted by λplacMu, i.e. that I had cloned the *sdaB* loci. I mapped both *sdaX*, and *sdaB* near 60.1 minutes, distinguishing them both from *sdaA*, mapped near 41 min. An 8 kb *Pst*I fragment from the *sdaB* clone hybridized to a Kohara phage carrying DNA from the 60.1 minute region of *E.coli*.

The preceding evidence clearly suggests that the *sdaB* and *sdaX* mutations define one or two newly discovered genes. Some evidence suggests that *sdaB* codes for the structure of L-SD#2. The *sdaB* gene cloned on a high copy number plasmid produced high L-SD activity. An insertion in *sdaB* abolished L-SD activity, as assayed both *in vivo* and *in vitro*. This attribution of *sdaB* as the structural gene for L-SD#2 was made even more likely by the fact that the *sdaB* gene could hybridize to the *sdaA* gene. However, I did not make a definitive proof that *sdaB* codes for L-SD#2.

1-3. Preliminary Studies on the Metabolic Role of L-Serine Deamination

The fact that L-SD converts L-serine to pyruvate and ammonia suggests that the metabolic role of L-SD may be in detoxification of L-serine, or in the use of L-serine as carbon and nitrogen source. As shown in this dissertation, the *sdaA* gene is regulated by the SOS regulation system and that may suggest that this enzyme is in some way involved in DNA repair, though it is not obvious what role it might play. Since L-SD#2 is made normally in rich media only, it is likely to play a degradative and gluconeogenic role.

PART 2. DETAIL OF THE STUDY OF THE *sdaA* GENE

2-1. Isolation and Characterization of a Mutation in *sdaA*, the Structural Gene for L-SD#1

The first intent of this work was to isolate a mutation in the gene coding for L-SD. Such a mutation should abolish L-SD activity as assayed both *in vivo* and *in vitro*, unlike the L-SD activation mutants which lack *in vivo* activity, but show activity *in vitro*. Several L-SD deficient strains had been isolated prior to the start of this work. One of them, strain MEW15, might have carried a mutation in the structural gene, however, the strain was refractory to standard genetic manipulations and could not be studied. I therefore decided to isolate more L-SD deficient strains.

I decided to use insertion mutagenesis with λplacMu to isolate such a mutation because this would put the *lacZ* gene under the control of the promoter of the gene of interest-providing a powerful tool to study regulation of gene expression. Previous work in this laboratory took advantage of Mud insertions [Casadaban and Chou, 1984 and Casadaban and Cohen, 1979]. However I decided to use λplacMu which carried the two ends of the transposable element bacteriophage Mu and can act as a transposable element if provided with helper phages carrying the Mu *A*, *B* genes that are necessary for the Mu transposition function. The advantage of λplacMu compared to Mud(Ap *lacZ*) [Casadaban and Chou, 1984] and MudX [Baker *et al.*, 1983] is that this λplacMu has the high transposition efficiency and very high insertion rate, about 10^{-5}, seen in Mu derivatives, but the insertions are more stable [Bremer *et al.*, 1984, and Bremer *et al.*, 1985]. Unlike Mu phage, once inserted, it rarely transposes during growth [Silhavy *et al.*, 1984]. Although Mud phages have the advantage of being small (around 39 kb), the larger size of λplacMu (48 kb) does not seriously affect transduction of the mutated gene to other hosts.

As one would expect, the chance of making an insertion in any particular gene is very low. I also did not know whether an insertion in the structural gene for L-SD would be lethal to the cell, or result in some growth requirement. However, the fact that several L-SD deficient activation mutants did grow without a supplement, and the lack of any obvious theoretical relation of L-SD activity to biosynthesis, both suggested that the isolation of structural gene insertion mutants should be possible. To increase the chance of isolating the particular mutation, I grew the population containing insertions in NIVSGL medium with ampicillin, to lyse the cells that grow on SGL medium, and thus enrich the population in cells which cannot grow on SGL medium. In this way, I was able to get over 100 colonies which were unable to grow on SGL medium.

Wild type *E.coli* can grow on L-serine with glycine and L-leucine, probably using pyruvate as carbon source via gluconeogenesis [Wang and Waygood, 1962]. Any insertion mutations in enzymes used to assimilate pyruvate would then also appear among the strains unable to grow on L-serine with glycine and L-leucine. To distinguish this kind of mutation, I tested ability of the SGL$^-$ strains to grow on L-alanine as carbon source. L-Alanine is readily used by *E. coli*, probably by making pyruvate via alanine dehydrogenase [Gottschalk,1985]. That is, apart from the first step in each pathway, L-alanine is used by the same path as L-serine. About 30% of the SGL$^-$ insertion clones were unable to grow on L-alanine, and were discarded. The remaining SGL$^-$ mutants were assayed for L-SD activity. Two of these strains, MEW21 and MEW83, which showed low L-SD activity in glucose minimal medium, with and without glycine and L-leucine, were saved for further study.

It is surprising that not all SGL$^-$ strains showed low L-SD. Indeed, most of the SGL$^-$ clones showed normal L-SD activity and a few gave quite high L-SD activity. This may indicate that deamination of L-serine to pyruvate or to any other compound the cell can use as carbon source is a multi-step procedure of which L-SD is only one reaction. However the fact that purified L-SD produces pyruvate makes this less likely. Careful study of such mutants would be needed to resolve this question.

Construction of *lacZ* gene fusions in this way has the advantage that the regulation of β-galactosidase synthesis from the structural gene promoter would mimic the regulation of L-SD. However, the λ insertion mutants made with λplacMu may have more than one insertion, and one would then risk measuring activity from *lacZ* inserted in some other gene. To avoid that, the insertions from the two strains, MEW21 and

MEW83 were transduced into MEW1 by selecting kanamycin resistance which is carried on λplacMu9. The resulting strains, MEW22 and MEW84, show the same phenotype as donor strains, such as SGL⁻; low L-SD in glucose minimal medium and with glycine and L-leucine; and similar levels of β-galactosidase as the donor strains. The fact that the L-SD deficiency followed as an unselected phenotype of kan^R indicates that the SGL⁻ characteristics are a consequence of the λplacMu insertions.

Strain MEW22 *sdaA*::λplacMu showed the physiological characteristics of a structural gene mutant: a loss of L-SD activity, and a regulation of β-galactosidase synthesis by the usual regulators of L-SD. {The β-galactosidase synthesized from the *sdaA* promoter showed 7-fold induction by glycine and L-leucine compared with 6-fold of L-SD induction in wild type strain(Table 2)}. These same characteristics were seen later in the work in a strain *sdaA*::Cm^r carrying a null mutation constructed *in vitro*. Both strains were devoid of L-SD activity whether assayed *in vivo* or *in vitro*.

The other SGL⁻ mutant studied here, strain MEW84, showed low L-SD when growing in minimal medium, with glycine and L-leucine, and LB medium. However when tested *in vitro* with activators, it showed considerable L-SD. It was therefore classified with the activation mutants, and not studied further in this work.

2-2. Cloning of the *sdaA* Gene

In order to make a molecular study of the *sdaA* gene, it was necessary to clone it. If the insertion in the *sdaA* locus indeed prevented the strain from growing on SGL medium, it should be possible to clone the *sdaA* gene by restoring to strain MEW22 the ability to grow on SGL medium. Indeed, *sdaA* carried on a high copy number plasmid might direct the synthesis of enough L-SD to permit growth in medium with L-serine alone, even without the glycine and L-leucine usually added as L-SD inducers like the *ssd* [Newman *et al.*, 1982b] and *lrp* [Lin *et al.*, 1990] mutants.

Traditional cloning was done by *in vitro* methods: isolating DNA, digesting it with suitable restriction endonucleases, ligating it to an appropriately cut vector, introducing it into a cell by transformation or transduction and selecting for complementation or screening in some other way. However a new method for *in vivo* cloning in *E. coli* was developed by Casadaban and his co-workers [Groisman and Casadaban, 1986 and Groisman *et al.*, 1984]. This was done by incorporating a plasmid origin of replication in a mini-Mu element. In the donor cells carrying this element, and also made lysogenic for Mu*cts*, after Mu induction, both the mini-Mu replicon and the Mucts lysogen can transpose hundreds of times as they replicate. In some instances, the small mini-Mu replicon may transpose itself into positions on both sides of a particular gene or a mini-Mu replicon on one side and Mu*cts* phage on the other. When the total length of DNA from the start of one Mu*cts* or mini-Mu through the gene to the end of the other mini-Mu happens to be of packageable size, that entire unit may be packaged in a Mu phage coat. When such a packaged structure is introduced by phage infection into a cell lysogenic for Mu, recombination can occur between the Mu sequences generating autonomously replicating plasmids, each carrying Mu or *E. coli* DNA. The inserted *E. coli* DNA can be up to 31 kb in the case of the construct used here, Mud5005. This reduces the preliminary cloning steps to a simple lysis of a donor strain carrying Mucts and a mini-Mu replicon, the use of the resulting lysate to transduce *E. coli* DNA into a recipient strain and the subsequent selection or screening for the complemented phenotype.

Many cases are known in which more than one gene can complement a given mutation. The fact that a clone complements a mutation, does not mean that the cloned gene is identical to the mutated gene. However the interest in this work was to clone the same gene that had been mutated. To confirm that the DNA I cloned is the same as the gene which is disrupted by λplacMu insertion, a 2.6 kb *Pst*I-*Sal*I fragment from the clone was hybridized with DNA from the insertion mutant MEW22 cut with the same enzymes (Fig. 9). It was clear that this 2.6 kb fragment of wild type strain MEW1 is replaced by two bands-3.1 and 5.1 kb in mutated strain MEW22.

That 2.6 kb *Pst*I *Sal*I DNA fragment, which had been obtained as a mini-Mu replicon, was then subcloned to pBR332. The plasmid containing this DNA fragment directs the synthesis of a very large amount of L-SD activity (14 fold higher L-SD activity than that coded by a single chromosomal gene), which is consistent with the idea that *sdaA* codes for L-SD#1, but does not prove it.

2-3. The Nucleotide Sequence of the *sdaA* Gene

Determining the DNA sequence of the 2.6 kb *Pst*I *Sal*I fragment would be expected to provide information about the molecular structure of both the *sdaA* gene, and its product L-SD. Several methods were developed for DNA sequencing at the time I started this project. I made use of the bluescript plasmid system developed by the

Stratagene Inc. in order to simplify commonly used cloning and sequencing procedures.

Most single-strand producing systems are derived from the filamentous phage (Ff', f1, fd, and M13) [Zinder and Boeke, 1982]. This is also true of the bluescript phagemid which contains a 454 nucleotide F1 phage intergenic region (M13 related) [Stratagene Inc.]. This sequence permits bluescript phagemids to be rescued from F'-bearing bacteria as single-stranded DNA packaged in phage heads. DNA can be isolated from these heads and used for DNA sequencing.

Many people use the M13 phage for producing single strands and doing sequence. However the bluescript phagemid has two great advantages- one that it can replicate both as a plasmid and as a single strand DNA phage, and the other that it can tolerate relatively large inserts of DNA, up to 10 kb. The bluescript system is arranged such that when the polylinker is cut, *Exo*III exonuclease deletes nucleotides in the direction of the inserted DNA, which facilitates sequencing [Stratagene Inc.].

Thus, I used a sequencing strategy involving inserting DNA into bluescript phagemid vectors, deleting into DNA by *Exo*III exonuclease, choosing deleted plasmids of suitable size to make single strands and sequencing these by dideoxy chain termination reactions.

A rapid, high resolution DNA sequencing gel system was developed by B. F. Lang and G. Burger [Lang and Burger, 1990]. This homemade system has several advantages as compared to other widely used systems. Among these are its relatively low cost, its use of commercially available glass plates; its use of a very thin (0.2mm thick acrylamide gel) resulting in sharper separation and clearer bands, and the ability to do *in situ* dry fixation of the gel matrix to the glass support without previous covalent binding of the gel.

By using the phagemid and gel system, I sequenced by the dideoxy-chain termination method the 2.6 kb DNA fragment which complements an *sdaA* mutation. This 2.6kb DNA had only one complete open reading frame of significant length, and that specified a protein of molecular weight around 48,000. I originally thought that the translation start site was at an ATG codon at base 663, which produced a protein of 448 amino acid residues. However, as suggested to me by W. Epstein (University of Chicago), there is another start codon GTG at base 645 of our clone, with a perfect Shine Dalgarno sequence (AGGAG) 7 bases upstream of it. This would code for a 454-amino acid *sdaA* product. Since this work was finished, an Edman degradation of the N-terminal of the *sdaA-lacZ* fusion protein showed that the actual translation start site for *sdaA* gene is in fact the GTG codon at base 645 (Moniakis, personal communication).

Transcription of an *E. coli* gene is usually ended at a strong terminator region, which consists of a stem-loop secondary structure followed by an AT-rich region [Ryan and Chamberlin, 1983]. Such a sequence is found on the *sdaA* clone, which strongly suggests that this is a single-gene operon. Thus, this TAA TACTTCTTAC<u>TCGCCCATC</u>TGCAACG<u>GATGGCGA</u>ATTTATA sequence may act as a transcription termination signal for *sdaA* gene.

It was important to confirm that the open reading frame (ORF) which I was studying was in fact the correct one. To do this, I constructed an in-frame fusion of *sdaA* to *lacZ*. From the determined sequence, by computer analysis, I picked a restriction site midway through this ORF, such that a cut there would make possible construction of such an in-frame fusion. Were this not the correct ORF, this procedure would not put the β-galactosidase under the control of *sdaA* promoter. Because the fusion could be made as predicted from this ORF and this sequence, and because β-galactosidase was synthesized from this fusion, and its synthesis was regulated by L-SD inducers, glycine and L-leucine, and the *ssd* locus, I conclude that this ORF in fact codes for the *sdaA* gene product and that its orientation and reading frame deduced by computer analysis are correct.

I conclude then that the *sdaA* gene codes for a protein of molecular weight 48,000 containing 454 amino acids. The coding region starts from a GTG codon at base 654 and contains a Shine Dalgarno sequence 7 bp upstream of GTG and a secondary structure resembling a *rho* binding site downstream from the last codon.

2-4. Analysis of the *sdaA* Sequence: Coding Region

The NBRF protein data bank and GenBank (Release 59.0) were searched for homologies both on the nucleotide and on the protein level using the Fasta and Tfasta programs of B. Pearson, University of Virginia, Charlottesville. At the time this work was completed (1989), there were no significant homologies between the *sdaA* gene and any sequence in either data bank.

The failure to find homologous sequences in the nucleotide and peptide banks indicates that the *sdaA* gene product, L-SD#1, differs in enzyme structure, active group(s) and thus probably reaction mechanism, from other deaminases. There is a great deal of homology between other deaminases: threonine deaminase [Umbarger, 1973], both biosynthetic [Lawther *et al.*, 1987] and biodegradative [Datta *et al.*, 1987] and D-serine deaminase [Marceau *et al.*, 1988] (all in *E. coli*), L-threonine deaminase in yeast [Kielland-Brandit *et al.*, 1984] and L-serine dehydratase in human and rat liver [Ogawa *et al.*, 1989a, Ogawa *et al.*, 1989b]. All of this group can deaminate both serine and threonine, and show significant homology in the functional sites and similar reaction mechanisms. However, the two enzymes described in this work resemble each other in sequence and in reaction parameters, and have no similarity with the other group.

The *sdaA* and *sdaB* genes are very similar, as judged by hybridization studies of *sdaB* to *sdaA* (Fig. 21), and by the extra bands of chromosomal DNA to which *sdaA* hybridized (Fig. 9). This was confirmed by the DNA sequence of *sdaB* gene which was obtained after this work was finished, and shows that the two nucleotide sequences shares 70% homology in the coding region and about 60% in the translated amino acid sequence. As shown here, the gene products are similar in both requiring activation by iron and DTT, and both showing little pH specificity and a requirement for a high substrate concentration. All this may indicate that the L-SD enzymes in *E. coli* have a novel structure and/or mechanism of action.

L-Serine deaminating enzymes in other bacteria are very similar to those of *E. coli* in their general characteristics [Newman and Kapoor, 1980]. Thus the *Corynebacterium* enzyme also requires activation with iron and DTT, and the *Klebsiella* *aerogenes* enzyme is induced by glycine [Vining and Magasanik, 1981]. In a more recent search of the Genbank (February 1991), I found a considerable homology between the areas upstream of *sdaA* in *E. coli* and downstream of *pabB* in both *Klebsiella aerogenes* and *Salmonella typhimurium* [Goncharoff and Nichols, 1988]. Dr. Nichols of the University of Illinois at Chicago made available to us some additional sequences at the downstream end of his clones, and it is obvious that the clone carries the N-terminal region of a gene with extremely high homology to *sdaA*. It seems likely then that this is a highly conserved enzyme across a number of bacterial species.

2-5. Analysis of the *sdaA* Sequence: Upstream Region

Expression of the cloned gene is inducible by glycine and L-leucine, and to the same extent as L-SD in the parent. This suggests that the clone carries the signals for induction by glycine and L-leucine, presumably upstream of the GTG at position 645. Therefore the upstream region was examined to try to locate possible regulatory sequences.

L-SD is one of a group of genes induced by L-leucine, a group now known as the leucine regulon [Lin *et al.*, 1990]. The expression of these genes is governed by their response to L-leucine and the *lrp* gene product, the leucine-responsive protein. Some genes are induced by L-leucine and *lrp*; others are repressed.

R. Lin has shown that the lrp protein can bind to the upstream sequence of *sdaA*, *serA* and *lysU* (personal communication), and it is also known to bind to *ilvIH* [Ricca *et al.*, 1989]. However, so far no consensus sequences have been located in the upstream regions of these genes. The *sdaA* upstream sequence was compared with that of another L-leucine regulated gene, *kbl*, which is also regulated by *rbl*, and two regions of considerable nucleotide homology were found. One region includes the putative SD sequence of both genes(*kbl* CCCTGTCTGGAGAAT; *sdaA* CCTTGTCAGGAGTAT). This represents a match of 12 of 15 positions. The second region overlaps the proposed -35 region of the *kbl* gene and occupies a similar position in the *sdaA* gene(*kbl* CGCGTTATCTCGT; *sdaA* CGCGTTCCCTCTT; a match of 10 of 13 positions). These two blocks could be involved in regulation of expression of the two genes-either as *lrp* binding sites or by some other mechanism.

L-SD#1 is also regulated by temperature of growth, being induced by an increase in temperature [Newman *et al.*, 1982a]. This induction may be under the control of the *htpR* gene, as judged by the fact that L-SD induction at 42°C is abolished by the *htpR* mutation (Newman,unpublished data). However no clear induction of β-galactosidase could be demonstrated.

It is not clear then whether L-SD is really regulated by the *htpR* system. The *htpR* mutant on which the analysis depends is a fragile strain, and low L-SD levels in that strain can be due to physiological problems of the strain. Nonetheless, I searched for the known binding site of the *htpR* gene product, sigma 32 [Neidhardt and VanBogelen, 1981], a -35 consensus sequence, TNtCNCcCTTGAA [Cowing *et al.*, 1985]. The fact that I could not find this strengthens the possibility that *sdaA* is not part of the *htpR* regulon.

L-SD#1 is also known to be induced by DNA damaging agents, such as UV irradiation. Such effects are generally mediated by the *lexA* repressor, which turns off/down expression of certain genes except as a response to DNA damage. A mutation decreasing lexA binding would lead to constitutive synthesis from genes of this regulon.

Preliminary evidence indicates that the *sdaA* gene might be controlled as part of the *lexA-recA* regulon [Newman *et al.*, 1982b]. Thus L-SD activity, was induced twofold in a strain carrying a *lexA* mutation compared with its parent strain. Similarly, in a *recA* mutant, the usual four-fold induction of the *sdaA-lacZ* fusion was abolished. However, the consensus sequence for the lexA protein binding site, taCTGTatata-a-aCAGta [Walker, 1984], was not found in the upstream sequence of the *sdaA* gene, and it is not clear whether *sdaA* is part of this regulon either.

In summary, although the *sdaA* gene may be regulated as part of several regulons, and certainly forms part of the *lrp* regulon, no consensus sequence has yet been identified in its upstream region. This could mean that *sdaA* is not regulated by *lexA* and *htpR*, or it may be that the regulatory sequence is less obvious in this gene. Binding sites for lrp protein have indeed been demonstrated by gel retardation and DNaseI protection (Lin, personal communication). Nonetheless, no consensus sequence has emerged, even from that work.

2-6. Map Position of *sdaA*

The insertion mutation in *sdaA* was mapped first by conjugation, and then by P1 transduction. The *sdaA* locus was located between 40.75 and 41.25 minutes of the latest *E. coli* map [Bachman, 1990]. This region was searched for any possible known genes which may be related to *sdaA* gene or L-SD, but no obvious candidate was recognized. This indicates that the *sdaA* gene has been demonstrated and located here for the first time.

The recognition of the homology between *E. coli* and *Klebsiella aerogenes* referred to above has led to the identification of two genes just upstream of *sdaA*. Nichols (personal communication) had sequenced, in *E. coli*, the entire region from *pabB*, and half of another folic acid related gene to the *Sal*I site, which was the site where sequencing of the *sdaA* gene started. Connecting the two sequences at this *Sal*I site indicates that there are about 200 bp between the second gene and the start of the coding region for *sdaA*. This is also consistent with *sdaA* being located at 41 minutes because the *pabB* gene of *E. coli* also mapped in this region [Goncharoff and Nichols, 1984].

Another way to map genes accurately is to hybridize to the appropriate λ Kohara phage. In this case, it is impossible to do that because the phage from 40.0-41.3 minutes region is missing from the Kohara mapping kit [Kohara *et al.*, 1987].

Prior to this work, it was thought that strain MEW15, an SGL$^-$ mutant isolated by Mu::dX insertion, might carry its insert in the structural gene now called *sdaA* [Newman *et al.*, 1985a]. This strain was exceedingly difficult to work with, but its mutation was tentatively mapped near 41 minutes (J. Garnon unpublished data, this laboratory). Using the same rationale as was used to show that the 2.6 kb *Sal*I *Pst*I

DNA fragment containing *sdaA* gene hybridized to an interrupted gene in strain MEW22, I also showed that it hybridized to an interrupted gene in strain MEW15. It seems then that strain MEW15 carries an insertion in *sdaA*.

PART 3. STUDIES ON THE *sdaA* GENE PRODUCT

3-1. Use of the *sdaA* Sequence to Produce Protein Fusions

Though the preceding evidence is suggestive, it does not prove that the *sdaA* gene codes for the actual L-SD enzyme. In previous work, using standard biochemical techniques, this laboratory was not able to make a substantial purification of L-SD. The construction of a fusion between *sdaA* and *lacZ* would advance work both by proving the gene to be structural, and by providing a quick and convenient method to purify large amounts of L-SD.

The principle used was to fuse the two genes in such a way that the product of *sdaA* would be covalently linked to β-galactosidase. The β-galactosidase could then be purified by affinity chromatography using commercially available substrates. If that molecule had L-SD activity attached to it, this would conclusively demonstrate that *sdaA* codes for L-SD.

The first fusion made fused only the N-terminal half of the *sdaA* gene to *lacZ* in plasmid pMEZS22 and it did not show any L-SD activity. Since I know now that *sdaA* does code for L-SD, it is clear that the latter part of the *sdaA* gene is necessary to obtain a functioning L-SD molecule.

To make a complete fusion, I designed an *in vitro* mutagenesis protocol that would convert the stop codon of *sdaA* into a site cut by the endonuclease *Eco*RI. This could then be fused to any *Eco*RI site which would allow in-frame fusion to *lacZ*. This could be done by changing only 2 nucleotides at the *sdaA* terminus. This was done by synthesizing the oligonucleotide AA GAA GAA TTC GTC ACA CTG, which spans the termination codon sequence TAA and directs it to be changed to the restriction endonuclease *Eco*RI site GAATTC.

This construction does not put *lacZ* exactly in-frame if I fused to the *lacZ* in plasmid pMC1871 (Fig. 6). It is true that there is an *Eco*RI site on the polylinker upstream of the pMC1871 *lacZ* gene. However, joining those two *Eco*RI cuts would put the *lacZ* one base out of frame.

I circumvented this by first inserting the newly-created *sdaA* *Eco*RI into the bluescript polylinker *Eco*RI site by subcloning the mutated plasmid pMES25 into Bluescript(+) *Sal*I and *Eco*RI sites. When that polylinker and the *lacZ* polylinker were both cut with *Sma*I, they could be religated with the *lacZ* in frame. The position of the *Eco*RI site and the sequence upstream of the *Eco*RI site on the *sdaA* clone were also confirmed by DNA sequencing in this plasmid.

The site directed mutagenesis procedure is a powerful tool for introducing mutations and redesigning sequences in genes of interest. However the low rate of mutation obtained necessitates the development of a strong selection for mutated plasmids. One way used is to synthesize single-strand DNA containing uracil in place of thymine using a strain which is a *dut- ung-* double mutant developed by Kunkel

[Kunkel, 1985]. In that strain *dut* which codes for dUTPase will inactivated, resulting in high intracellular levels of dUTP. The *ung* mutation inactivates uracil N-glycosylase, which allows the incorporated uracil to remain in the DNA. This uracil-containing template strand will be degraded in wild type strains with uracil N-glycosylase function. This method allows the production of more than 60% mutated plasmids instead of 5-10% in other methods. The mutagenesis kit purchased from Bio-Rad is designed for this kind of experiment. Unfortunately, I could not obtain uracil-containing single-stranded DNA from either Bluescript or pTZ plasmids (provided by Bio-Rad in the kit) carrying all or part of the *sdaA* gene. I therefore could not take advantage of this system and obtained about 5% mutated plasmids (2 of 48 clones) using normally constituted ssDNA as template. However, since the mutated plasmid contained a new *Eco*RI site, it was easy to identify the desired plasmid.

The strain carrying the fusion of the entire *sdaA* gene to *lacZ* did not show L-SD activity as judged by the whole cell assay but did show a huge amount of L-SD activity in cell extracts incubated with iron and DTT. That seems to indicate that the physiological activating system cannot deal with L-SD attached to β-galactosidase. This may be due to the large size of the β-galactosidase, a tetramer of which each monomer is about 3x the size of an L-SD monomer. This structure may interfere sterically with the activation site of the *sdaA* gene product.

3-2. Purification of Fusion Proteins

The proof that *sdaA* codes for L-SD requires the purification of the fusion protein.

Several methods have been used for purifying β-galactosidase fusion proteins, one of which, developed by R. Young, is based on immobilized anti-β-galactosidase antibody [Struck *et al.*, 1985]. Several companies make columns with anti-β-galactosidase antibody. However as Dr. Young informed us, it is not so simple to elute the fusion protein from the antibody column. Some methods require a fairly complex purification [Fowler and Zabin, 1983]. One of the most efficient, rapid methods, was developed by Agnes Ullmann [A. Ullman, 1984]. This uses binding in high salt and elution at high pH in borate buffer.

The method I used was a simple one. Since L-SD has proved unstable in the past, I did not want to use 100 mM borate as an eluant, as suggested by Ullman. Now β-galactosidase binds to the column (p-Aminobenzyl 1-Thio-β-D-Galactosidase-Agarose, from Sigma-A-0414) by its affinity to the substrate analogue. It might then be possible to elute it by affinity to its substrate, D-lactose, particularly if this was done in a high salt buffer. In fact, I was able to elute both fusion proteins (2PF and 3PF) with TMN buffer containing 20% lactose. This worked much better with the 2PF, of which I could elute more than 90% of the fusion protein loaded on the column. However I could only elute 10% of the 3PF. The two fusions may have different affinity for the column or for D-lactose. In any case, a better elution method for the 3PF will be required. However this one was sufficient for the experiments needed for this work.

Thus, using a modification of the method developed by A. Ullman, I obtained highly purified 2PF and 3PF. Both fusion proteins showed β-galactosidase activity, but required activation with Fe and DTT to show L-SD activity. During chromatography on Superose 12 with the FPLC, β-galactosidase and L-SD activity were found in the same fraction. This provides the complete proof that *sdaA* codes for the structure of L-SD#1.

Either fusion protein could be used as an antigen for making antibody to L-SD, a procedure which has now been accomplished by J. Moniakis.

3-3. Purification of L-SD from Fusion Proteins

The preceding system results in the purification of L-SD as a fusion protein. However one would like to have purified L-SD without another protein fused to it. This can be achieved by a modification of the 2PF procedure. Indeed several modifications have been developed [Sassenfeld and Brewer, 1984, and Germino and Bastia, 1984].

One of most elegant and efficient methods was developed by D. Bastia and his co-workers [Germino and Bastia, 1984]. This method uses the same principle as the preceding for purification-namely affinity chromatography of a fused β-galactosidase molecule. However, upstream of the *lacZ* gene used is cloned a sequence of the collagen gene coding for a collagenase-sensitive polypeptide. This allows the formation of a triple fusion: L-SD through a collagen peptide to β-galactosidase. When the purified triple-fusion protein is cleaved with collagenase at the specific peptide sequence-Pro-X-Gly-Pro-Y- [Fuller and Boedtker, 1981], the resulting products may be L-SD#1 with the amino acids coded by linker sequence and collagen gene in its c-terminal and β-galactosidase itself with some amino acids from collagenase sensitive peptide in N-terminal, and these can be readily separated on FPLC. The major problem in this method is that the collagenase used to do the cleavage is often contaminated with protease and either one must purify it considerably, or one must find a satisfactory commercial preparation.

Other similar methods have been devised [Smith *et al.*, 1984 and Sassenfeld and Brewer, 1984]. One of these involves adding a sequence coding for a polyarginine polypeptide to the C-terminus of the target gene. This changes the physical characteristics of the target protein and allows a simple purification from all other proteins by ion-exchange chromatography. This C-terminal polyarginine tail can be removed by treatment with a specific exopeptidase, carboxypeptidase B [Folk, 1970].

I preferred to use Bastia's collagen vector for the purification of L-SD because I had already set up the system for purification of β-galactosidase, and fortunately, there is no collagenase sensitive sequence within the *sdaA* product. The 3PF plasmid was constructed and a very large amount of fusion protein was produced from plasmid pMES27 which carries a λ promoter upstream of the *sdaA*-collagen-*lacZ* fusion gene. This promoter was controlled by a λcIts857 gene, and thus could be induced by heat shock in this plasmid.

As shown in section 4 of the results, the purified 3PF after digestion with collagenase, showed two peaks of enzyme activity, one showing L-SD activity, and one β-galactosidase. This corresponds to the two bands shown on the SDS-page gel

(Fig. 16). However because of elution problems, the amount of protein produced in this way is very low.

PART 4. BIOCHEMICAL CONSIDERATIONS ABOUT L-SD#1

As stated earlier, L-SD#1 has no sequence homology with any of the enzymes which might be compared to it. That fact, and the fact that no pyridoxal phosphate requirement has ever been shown for this enzyme-even after purification from the 3PF, suggest that it is fundamentally different in mechanism from either *E. coli* threonine deaminase, from D-serine deaminase of *E. coli*; from L-threonine deaminase in yeast [Datta *et al.*, 1987] and from mammalian L-serine dehydrogenase [Ogawa *et al.*, 1989a and 1989b]. All of these are relatively non-specific and deaminate both threonine and serine. All also share a functional peptide sequence.

I was able to determine for the first time that L-SD#1 is specific for L-serine and does not act on L-threonine. I conclude that the *sdaA* gene product codes for an enzyme different from all previously described proteins, and that *sdaA* is a newly discovered gene, located near 41 minutes on the *E. coli* map.

PART 5. REGULATION OF EXPRESSION OF THE *sdaA* GENE

The expression of *sdaA* is regulated by a rather astonishing number of factors. The mechanism of some of this is now being investigated. Some of it is not understood at all. The current understanding is discussed in the following paragraphs.

5-1. Regulation by Glycine and L-Leucine

The expression of the *sdaA* gene was shown here to be increased by growth with glycine and L-leucine. This is due, all or in part, to regulation by the lrp protein [Lin *et al.*, 1990]. A mutant deficient in *lrp* showed increased synthesis from *sdaA* [Lin *et al.*, 1990]. R. Lin(unpublished data) has shown that the lrp protein can bind to the upstream sequence of the *sdaA* gene in the absence of L-leucine, and its binding affinity was decreased by adding L-leucine. That suggests that the lrp protein acts as a repressor which binds to the upstream region of *sdaA* and that L-leucine can release this repression. That different concentrations of L-leucine induce different amounts of L-SD has been shown *in vivo*, which is consistent with the binding experiments.

The regulation by *lrp* integrates *sdaA* expression with that of the rest of the regulon. This includes enzymes in L-serine biosynthesis, L-threonine degradation, transport of tripeptides, L-leucine, L-isoleucine biosynthesis, and a gene coding for lysyl tRNA synthetase and at least 20 further proteins as judged by 2D gels studied by R. Matthews [Lin *et al.*, 1991]. Some of these are induced and some repressed in an *lrp* null mutant. The physiological meaning of this regulation is not altogether clear, but may be related to the availability of organic material in the cell's milieu. In any case, it is clear that *E. coli* uses the external L-leucine concentration as a major regulatory signal.

On the other hand, the mechanism of glycine induction remains unclear. Glycine and L-leucine induce when provided together or separately [Newman *et al.*, 1982b]. However β-galactosidase synthesized from the *sdaA* promoter can be induced by glycine even in the *lrp* mutant [Lin *et al.*, 1990]. Moreover glycine does nc. interact with the lrp protein as judged by Lin's gel retardation experiments. This suggests that glycine effects are transduced by something other than *lrp*. However no further information as to the glycine transducer is available.

The biological function of glycine induction is also not clear. Glycine and L-serine are interconverted by serine transhydroxymethylase (STHM) coded by *glyA*. One might then think that whenever the cell has one of these, it makes a lot of the other. But if glycine cause a high intracellular L-serine concentration, and that induces L-SD, then a high serine concentration should itself induce L-SD and it does not.

The major C1 donors of the *E. coli* cell are glycine and L-serine In glucose grown cells, about 25% of C1 units are derived from glycine through the GCV cleavage system [Newman and Magasanik, 1963]. Perhaps when glycine is provided externally, the cell is signalled that an alternative C1 source is available, and the cell can afford to operate with a reduced L-serine concentration, thus reducing its toxicity problems.

5-2. Regulation through the *ssd* Gene Product

The *ssd* mutation greatly increases expression from *sdaA* [Lin *et al.*, 1990]. The mechanism of this is not clear at present. However the mutation increases both L-SD activity in the mutant itself, and β-galactosidase activity synthesized from an *sdaA* promoter either in plasmid pMEZS22 or in the chromosomal DNA insertion strain MEW22.

The two regulons, *lrp* and *ssd*, affect *sdaA* gene expression differently. Glycine and L-leucine induce *sdaA* in the *ssd* mutant but not in the *lrp* mutant by L-leucine. These two regulons also regulate a different group of enzymes, even though they clearly overlap. The *ssd* mutant has been tested with a combination of glycine and L-leucine, but not with either alone. It could therefore define a glycine regulon, but this has not been tested.

Silverman suggested that *ssd* mutation is in the same gene which is known as *cpxA* [Rainwater and Silverman, 1990]. Unpublished data from this laboratory also support this idea. If this is true, the membrane sensor protein, coded by *cpxA* [Weber and Silverman, 1987], must work through some intermediate to regulate DNA expression. This intermediate is as yet unknown.

The *cpxA* gene product is thought to be involved in the regulation of energy transduction by the *E. coli* membrane [Weber and Silverman, 1987]. It is difficult to see how L-serine deamination is involved with energy transduction, electron transport or redox potential. However, the real function of L-SD is not clear, and *sdaA* expression is known to be affected by the oxygen supply of the cell.

5-3. Regulation by DNA-damaging Agents

L-SD activity is known to be induced by DNA-damaging agents [Newman *et al.*, 1982a]. The UV-irradiation level used in those experiments was very high, and so those results may not be physiologically significant. However the enzyme was also induced by growth with sublethal levels of mitomycin and nalidixic acid, and that seems to be consistent with a physiological response.

As reviewed above, some evidence indicates that this response is governed by the SOS regulon. However, the fact that I could not find a *lexA* consensus sequence upstream of *sdaA* suggests that the *recA-lexA* system does not act directly at the *sdaA* promoter but affects it through another factor. It is also possible that the *lexA* protein can bind to other sequences in the *sdaA* upstream region. This could be tested by *in vitro* binding experiments using lexA protein. So far, no case of lexA protein binding to a sequence other than that consensus sequence has been reported.

The genes of the SOS regulon synthesize products involved in DNA repair and mutagenesis. L-SD#1 is so far only known to function as an L-serine deaminating enzyme. Moreover the *sdaA* mutant is not particularly UV sensitive. It seems unlikely then that L-SD is involved in DNA repair. However, it might be involved in regulation of the C1 pool, related perhaps to a need for DNA methylation. It might also be a clean-up enzyme induced by DNA damage and synthesis and subsequent degradation of faulty proteins.

5-4. Induction of *sdaA* at High Temperature

It is similarly clear that L-SD activity is induced by growth at 42°C, whether or not this is mediated by the *htpR* gene product. Again the expected consensus sequence was not seen, and again regulation may be indirect. Induction might even represent a posttranslational increase in enzyme activation. Were this true, L-SD activity as measured in extracts should not be induced by growth at 42°C even though whole cell activity might be higher. However L-SD after heat shock induction was higher in both assays.

The inability to show heat shock regulation of β-galactosidase synthesis from the *sdaA* promoter from both λplacMu insertions and from plasmid pMEZS22 may be due to an increase in phage induction and decrease in plasmid stability at high temperature. It might be also possible that the increase in L-SD activity is not due to the *sdaA* gene product.

Some genes known to be part of the heat shock regulon also do not show the consensus sequence for sigma 32 binding. One of these is *lysU*, which does not show the sigma 32 binding site [Clark and Neidhardt, 1990], but is regulated by *lrp*. Perhaps then heat shock acts on the *lrp* gene product and this induces *sdaA* and *lysU* both. Were this true, the other genes of the *lrp* regulon should also be heat induced, but this has not been tested.

It is also true that several heat shock genes are also regulated by the SOS regulon. It may be that at high temperature, the cell requires biodegradative enzymes, just as it might after DNA damage.

PART 6. SOME CONSIDERATIONS CONCERNING L-SD ACTIVITY

One uncommon feature of L-SD regulation is that the *sdaA* gene product is not the active L-serine deaminase itself, but a non-functional protein activated physiologically by posttranslational modification. This physiological activation is mimicked *in vitro* by incubation with iron and dithiothreitol, and this is needed to activate both L-SD#1 and L-SD#2.

6-1. Effect of Mutations Which Prevent L-SD Activation

Three mutants have been shown to be deficient in L-SD activation, two isolated earlier (MEW128 and MEW191, [Newman *et al.*, 1985a]), and one isolated in this work,(MEW 84). These all show L-SD activity when the test is made *in vitro* with Fe and DTT; however they are unable to produce an active L-SD *in vivo*, and are SGL$^-$ in phenotype. The mutations in these strains are located at three different loci in the *E. coli* chromosome.

Very little is known about the functions coded by these genes. The factors that induce L-SD have no significant effect on synthesis of β-galactosidase from *lacZ* inserted in either *sda191* or *sda84*, nor have any other regulatory factors been demonstrated. The activation system is not necessary in all cases. L-SD is made in active form from *sdaA* carried on a high copy number plasmids even when they are placed in activation-defective mutants. The much smaller quantity made from the chromosomal *sdaA* gene is not activated when the plasmid-carried gene is not present. This seems to mean that when L-SD is synthesized in large amounts it can be activated independent of the cellular activation system. It may also be possible that the large quantity of L-SD made titrates some negative regulator of L-SD function.

The nature and relationship of the three "activation" genes remains unclear. Whether all three are structural or some regulatory, whether their gene product acts directly on the inactive *sdaA* product, or whether one or more acts indirectly (e.g. to produce a cofactor)-all remain unknown.

In any case all 3 gene products are needed for enzyme activation in the *E. coli* cell.

One possible clue is that all three mutations cause a concomitant requirement for thiamine. However added thiamine does not itself allow L-SD function, and there is no indication that these genes are involved in thiamine biosynthesis. Thiamine-nonrequiring derivatives can be isolated, and these remain L-SD deficient [Newman *et al.*, 1985b]. It may be that the activation-deficient strain is deficient in some unknown step, and that this deficiency can be relieved by a cryptic enzyme which has a much higher Km for thiamine than can be satisfied by the normal thiamine pool.

Mutations in 3 loci make the cell activation-deficient. Are there any other loci in which mutations would produce the same phenotype? The fact that the first three mutants mapped were all at different loci suggests strongly that there may be other as yet undiscovered loci.

What sort of activation procedure might this be? One possibility is suggested by the fact that activation of the Lactobacillus enzyme, histidine decarboxylase, was shown to involve serinolysis at a serine-serine bond [Recsei *et al.*, 1983]. The existence of a serine-serine bond at amino acids 243 and 244 of the *sdaA* reading frame suggested that the *sdaA* gene product might be activated by a similar mechanism. However, the *sdaB* gene, which is so homologous to *sdaA*, showing 70% nucleotide homology and 60% amino acids homology with *sdaA* gene and its predicted product, does not have a serine-serine bond in the corresponding position (Shao, personal communication) and it is therefore not likely that activation is in fact serinolytic. On the other hand, it is not obligatory that both L-SD be activated by the same *in vivo* mechanism, even if both are activated by the same *in vitro* method. Then L-SD#1 might be activated by serinolysis, and L-SD#2 by some other mechanism.

I designed an oligonucleotide to mutagenize one of the L-serines in the *sdaA* serine-serine bond into an alanine residue. However, since as I mentioned earlier, I could not produce U-containing ssDNA from the plasmid carrying the *sdaA* gene, or even part of this gene, I could not use this high-yield system to isolate the mutated plasmid. I did test a large number of plasmids but was unable to find one in which the serine-serine sequence was altered.

The fact that iron and dithiothreitol cut protein molecules [Kim *et al.*, 1985] suggests the possibility of a proteolytic activation. A mixture of iron and DTT can catalyze a large number of reactions, and which one is important here is unknown, nor is it known whether the *in vivo* process proceeds by the same mechanism as the *in vitro*. It is striking that iron and DTT can activate L-SD even when it is attached to the fusion proteins. Cutting of an L-SD bond might then not only activate, but also could remove it from an environment where it was sterically hindered.

6-2. *SdaA* on A High Copy Number Plasmid Suppresses the Activation Mutation

As discussed earlier, the three activation-defective mutants do not show L-SD activity from the chromosomal *sdaA* gene in the whole cell assay, but do produce it from *sdaA* on high copy number plasmids (Table 9). The three mutants show no significant L-SD when grown in glucose-minimal medium (1 unit) as compared to 19 units in the parent strain, MEW1. Plasmid pMES22 in strain MEW28 with its functional activation system makes 360 units of enzyme, but it still produces quite a lot of L-SD activity (40-60u) in the activation-defective mutants (Table 9). Thus production of active L-SD is very much reduced in the activation mutants, but the cells do contain active enzyme, at about 50% of the level of glycine-leucine induced parental cells. The cells have only 10-20% of the usual plasmid enzyme activity. However extracts of the activation-deficient strains incubated with iron and DTT also show less activity so that the activation problem may be smaller than it seems.

Could the L-SD levels in the plasmid-carrying strains be due to a selection for revertants in the host strain? This is not likely because the host strains without plasmid do not revert at a high rate, especially when an antibiotic selection is maintained. MEW128 revertants are more readily selected. However all transformations were done with cultures of an *sdaA*::Cm^r MEW128 strain which was checked for MEW128 function prior to use.

The activation-deficient strains are certainly not totally deficient in transcription of the *sdaA* gene since extracts of these strains all have an inactive form of L-SD (which is activated with iron and DTT). This is shown even more clearly by the fact that the *sda128* mutation did not decrease the β-galactosidase synthesized from the *sdaA-lacZ* fusion gene in plasmid pMES27 (Table 9).

The L-SD in plasmid-carrying derivatives of activation-deficient strains might arise from an intermolecular "self-activation" of L-SD. If this were true, one might expect that a similar proportion of self-activated molecules would be seen when the chromosomal *sdaA* gene is induced. However it may be that self-activation requires a concentration of L-SD molecules level higher than is made in MEW84 with glycine and

L-leucine. Self-activation might also correspond to the action of an incomplete activating system, since each of the activation mutants has several components of the system, and is missing only one.

The explanation of the synthesis of active L-SD in the activation-deficient plasmid-carrying strains is unclear. However the possibility that this depends on a high concentration of inactive L-SD molecules could be tested by constructing a series of *sdaA* gene expression plasmids which would be expected to make different levels of L-SD activity. Perhaps these will show some definite threshold level below which the inactive L-SD molecules cannot be activated.

6-3. The *sdaA-lacZ* Fusion Protein Is not Activated Efficiently *In Vivo*

The fusion protein of L-SD with β-galactosidase is not activated efficiently by the physiological activating system, though L-SD activity can be produced by incubation with iron and DTT. This can be concluded from a comparison with strain MEW1 which produces 19 units in the whole cell assay and 11 units in the extract [Newman *et al.*, 1985a]. Strain MEW28 with fusion plasmid pMES27 when grown in glucose-minimal medium with glycine and L-leucine makes 23 units as judged in the whole cell assay, and 3,600 units as judged in crude extracts. By analogy with MEW1, 3600 units in the crude extract would correspond to 1800 in the whole cell. This suggests that the physiological system has serious problems in dealing with the fused L-SD molecule.

The physiological system does act on the fusion protein to some extent, as judged by the fact that even the 23 units made from the fusion plasmid in the parent strain are not made in the activation mutants. The plasmid gene is transcribed, since the β-galactosidase level synthesized from the fusion plasmid is almost the same in the wild type and the activation mutant. It seems then that the large β-galactosidase molecule, about three times the size of the *sdaA* gene product, and existing as a tetramer, can block most of the sites needed for L-SD activation from interaction with the activators.

6-4. Instability of L-SD

Newman *et al* had shown that L-SD activity was very unstable in the intact living cell. Cultures treated with a protein synthesis inhibitor lost activity rapidly [Beeraj *et al.*, 1978]. That L-SD can exist in an inactive form was not known at that time, and extracts were not tested. Furthermore in maxicells carrying a high copy number *sdaA* plasmid, cells show no L-SD in the whole cell assay but show very high enzyme activity in extracts (data not included). This was seen in cells treated with UV and D-cycloserine. These two experiments show that instability of L-SD may be caused by degradation of the protein, but also by an inactivation that is reversible at least *in vivo*. Both experiments also suggest that the presence of the activation system enzymes is not sufficient to drive the reaction towards L-SD activation. This suggests that activation is turned off and on in some way. However, while the data shows that L-SD can be converted *in vivo* to an inactive form, and that this form can be activated *in vitro*, it does not show that inactivation is reversible *in vivo*.

L-SD is also unstable in frozen extracts [Newman and Kapoor, 1980], and even in a few hours on ice (in phosphate buffer). In these cases, however, activity is also not recovered on the usual incubation with Fe and DTT, and so inactivation may involve a different mechanism.

PART 7. STUDIES ON THE SECOND L-SERINE DEAMINATING ENZYME, L-SD#2

The existence of a second L-SD was indicated by the fact that mutants carrying insertions or a constructed null mutation in *sdaA* still showed L-SD activity when they grew on LB medium. In this section, I will discuss experiments which demonstrated the existence of this second enzyme.

7-1. Demonstration of the Existence of L-SD#2

In this study, I showed that *sdaA* codes the structure of L-SD#1. Clearly then a strain with a null mutation in *sdaA* could not make L-SD#1 and if it showed any activity, it must be coded from another gene. The fact that strain MEW22 carrying a λplacMu insertion MEW22 or the constructed null mutant strain MEW28, both showed the L-SD activity when grown in LB medium (Table 2 and Table 12), then demonstrates the existence of a second L-serine deaminating enzyme activity, L-SD#2.

This does not in itself prove that this second enzyme has L-serine deamination as its primary function. Two threonine deaminases are known to deaminate threonine [Umbarger, 1973]. However neither is made in this *ilvA* strain grown in LB. The biodegradative TD is made only in the absence of glucose and oxygen, so the assays for L-SD#2 were made in LB with excess (0.5%) glucose. L-Serine was deaminated by these LB-glucose grown cells, and this is clearly not due to the biodegradative TD.

There are other enzymes which might deaminate L-serine as a side-reaction apart from their main function. However, since the hybridization of the *sdaA* probe to chromosomal DNA gave strong evidence that a homologous gene exists, it seemed more fruitful to find it by molecular means, rather than to systematically exclude other possibilities.

L-SD#2 is synthesized in rich medium, and to some extent in minimal medium with yeast extract or tryptone added (data not included). It is not made in minimal medium supplemented with 0.1% casamino acids, but some L-SD#2 is made in minimal medium with casamino acids and the five bases found in nucleic acid. This suggests that there may be specific factors in LB inducing L-SD#2 synthesis, but no further attempt to define them was made.

7-2. Genetic Study of L-SD#2

The further study of L-SD#2 was made difficult by the fact that it was only made in LB, since this made direct selection of the mutant, or the complemented strain impossible. However, the first step in cloning the gene is to obtain a mutant deficient in L-SD#2.

I tried first to make λplacMu insertions in strain MEW28 and screen those strains which showed induction of β-galactosidase by LB medium but not in glucose minimal medium. 20 of these strains were isolated, but none of them proved to be deficient in this second L-SD activity.

I decided to alter the regulation of L-SD#2 so it would be expressed in SGL medium. From that point I could use the same methodology on L-SD#1 as on L-SD#2. I started with the strain which carried a null mutation in *sdaA* loci, and isolated strains which grew on SGL medium. From that point, it was only necessary to isolate a mutation in the L-SD#2 gene by isolating an SGL$^-$ derivative of the new SGL$^+$ strain, and showing that it did not make L-SD even in LB. Then the cloning of this L-SD#2 gene could be achieved by restoring the ability of the SGL$^-$ derivative to grow on SGL medium. However, the cloned gene obtained by this method would not be subject to the same regulation as the native gene.

Several methods have been described for constructing null mutations in *E. coli*. One of them is based on the fact that a colEI plasmid cannot replicate in a strain deficient in DNA polymerase I [Gutterson and Koshland, 1983]. If a colE1 plasmid specifying antibiotic resistance is transformed into a *polA*$^-$ host, and antibiotic resistant cells selected, they can arise only by integrating the plasmid-carried gene into the chromosome. This usually happens by integrating the whole plasmid into the chromosomal locus which has homology with some plasmid-carried gene.

Other methods include use of special plasmid with a temperature-sensitive origin of replication [Matsuyama and Mizushima, 1985], or of thermoinducible λ phage lacking the normal *att* site [Joyce and Grindley, 1984], or of a *recB, C, sbcB* [Jason and Schimmel, 1984]. A method based on plasmid segregation was also described [Kiel *et al.*, 1987].

In this work, I inserted a chloramphenicol-resistance cassette by an *in vitro* construction into the *sdaA* gene cloned on the colE1-type plasmid, pBR322. This was forced into the chromosome of the *polA* deficient strain, with the integration of the entire plasmid. This was done by selecting the chloramphenicol resistance of the construct gene, but the ampicillin resistance coded by pBR322 was also transferred to the *polA*$^-$ strain. I then used P1 to transduce chloramphenicol resistance to strain MEW1. By selecting strains which had transduced only chloramphenicol resistance, but not ampicillin resistance, I picked a strain which had lost most of the plasmid but retained the null mutation in *sdaA*. This *sdaA*::Cm^r strain MEW28 showed all the characteristics expected of a strain with a mutated *sdaA* gene. This was also confirmed further by hybridization results (Fig. 17). The strain made in this way is now used as the standard *sdaA*-deficient strain for work in our laboratory.

Altered expression of L-SD#2 was achieved by mutating strain MEW28 and selecting colonies which could grow on SGL medium. The strains which can grow on SGL medium also were assayed for L-SD activity since strains can grow on SGL for reasons other than a change in L-SD#2 regulation. Thus overexpression of *metC* allows growth on L-serine [Brown *et al.*, 1990]. So do the as yet uncharacterized GOS mutations [Brown *et al.*, 1990]. From the *sdaA*::Cm^r strain, I isolated an SGL$^+$ strain which showed L-SD activity on assay, and called the altered gene *sdaX*.

It was not clear at this stage whether the ability to grow on SGL was in fact due to increased synthesis of L-SD#2. It could also be due to establishment of synthesis of any other enzyme which can deaminate L-serine, and indeed even to an altered expression of the gene *tdc* coding for the biodegradative TD. However, by mapping this *sdaX* mutation, it might be possible to differentiate it from, or show its identity with, various other genes. In fact the location of *sdaX* around 60.25 minutes of *E. coli* map excluded the possibility of its identity with the *tdc* gene (near 68 minutes) [Schweizer and Datta, 1990]. No gene which seemed to be relevant to this area of metabolism was located near 60.25. This suggests that the *sdaX* mutation is likely to be a regulatory mutant, and could be in an independent regulatory gene, or in the promoter of the gene coding for L-SD#2.

It was then possible to isolate mutants of a strain carrying the *sdaA*::Cm^r and *sdaX* mutations using λplacMu9 insertion, screen for inability to grow on SGL medium, and demonstrate L-SD deficiency in all media tested, including LB.

The fact that the mutant had no activity in LB strongly suggested that the mutation affected L-SD#2. L-SD#2 is also activated by iron and DTT, so I checked that there was no activity in extracts incubated with iron and DTT, i.e. that the SGL$^-$ mutation was not in the activation system. The gene in which the insert leading to SGL$^-$ deficiency in the *sdaX* strain is located is named *sdaB* and may be the coding gene for L-SD#2, or a regulatory gene needed for its synthesis. This mutation was mapped by P1 transduction near 60.25 minutes which is consistent with *sdaB* and *sdaX* being the same gene. In any case, both are far from *sdaA* which mapped near 41.0 minutes.

7-3. Regulation of Expression of the *sdaB* Gene

The *sdaB* gene in the wild-type strain is expressed only in LB. The alteration in *sdaX* strain which could be in the promoter region of the gene coding for *sdaB*, established synthesis of L-SD#2 in minimal medium. Synthesis of L-SD#2 was also influenced by a mutation in the *lrp* gene, though less so than synthesis of L-SD#1. It may be then that the lrp protein, which binds to the upstream region of *sdaA* (Lin, personal communication), binds similarly to *sdaB*. This could be confirmed by gel retardation experiments using lrp protein and the DNA upstream of *sdaB*. By sequencing the *sdaB* upstream region from the *sdaX* mutant, and from the wild-type cell, one might also determine whether the *sdaX* mutation is in the *sdaB* promoter, and begin a study of promoter structure.

The enzymes, L-SD#1 and #2, both require activation by cellular enzymes, and are not made in activation-deficient mutants *in vivo*. However, L-SD#2 in extracts from activation-defective strains is not activated by iron and DTT. This suggests that even the inactive form is not made in activation-deficient strains, or that activation of L-SD#2 requires a further step which is not catalyzed by iron and DTT.

7-4. Cloning of the *sdaB* Gene

The *sdaA* gene was cloned from wild-type *E. coli* using a mini-Mu replicon. One could clone *sdaB* by the same system only if one converted the *sdaX* strain into a mini-Mu and Mu*cts* lysogenic donor, since one could only clone an allele that was expressed in minimal medium. That is, I constructed a mini-Mu Muts donor with an *sdaX* mutation, and also used a recipient with an *sdaX* so that *sdaB* would be expressed even if the donor *sdaX* mutation were not transferred. Then I selected transfectants which could grow on SGL medium, perhaps by complementing the *sdaB* mutation.

I wanted to use the mini-Mu d5005 replicon for *sdaB* gene cloning. However both it and the *sdaB* insertion strain were resistant to kanamycin. I tried to remove the kan^r gene of λplacMu9 from *sdaB* by UV irradiation but could not. I also tried to use strain MEW28 *sdaA*::Cm^r as a recipient strain for *sdaB* cloning from *sdaA*::Cm^r *sdaX* but this was also unsuccessful. This is surprising because one would expect that a plasmid carrying *sdaB sdaX* could complement *sdaA*::Cm^r, and in fact, the *sdaB* plasmid which I finally obtained did permit such a strain to produce high levels of L-SD and to grow on SGL. It seems then that the first cloning of the *sdaB* gene is only successful in a host with a nonfunctional *sdaB* gene. The attempt to clone in *sdaA*::Cm^r was made 3 times, no *sdaB* clone was found. Similarly, attempts to complement the SGL$^-$ character of MEW128 never produced an *sdaA* clone although the *sdaA* clone can complement the MEW128 SGL$^-$ character. I do not know how this phenomenon may be explained.

The way I finally used to isolate a kanamycin sensitive strain was by converting the λplacMu9 insertion into a λTn10 insertion by homologous recombination between λ phages. In that way, strain MEW55, *sdaA*::Cm^r *sdaX sdaB*::λTn10, was obtained. It is not sure whether this was really due to recombination between these λ phages, but the resulting strain was kanamycin-sensitive, and SGL$^-$, and showed no L-SD activity when tested at 28°C in LB medium. This allowed *sdaB* to be cloned in a mini-Mu replicon. 8 kb of a *Pst*I fragment from that mini-Mu was inserted into the *Pst*I site of pBR322. Both clones directed the synthesis of high levels of L-SD, whether the host strain was *sdaX* mutant or wild type.

That *sdaB* was the gene cloned was confirmed by hybridization. The 8 kb *Pst*I fragment from the plasmid hybridized to two different bands (6 and 10 kb) in a *Pst*I digest of chromosomal DNA from strain MEW50, *sdaA*::Cm^r *sdaX* *sdaB*::λplacMu, proving that the disrupted gene had in fact been cloned. It is not known whether this 8 kb fragment carries the *sdaB* gene with its own promoter or whether *sdaB* is part of an operon. If *sdaB* gene were in an operon, the insertion of λ phage might not be in the *sdaB* locus.

This fragment also hybridized to Kohora phage λ457 which contains DNA from the 60.1 minute region of *E. coli* [Kohora *et al.*, 1987]. That also indicates that the gene cloned is the same gene that carries the insertion and maps near 60.1 minutes. This is also confirmed by the fact, determined after this work, that the DNA sequence downstream of *sdaB* (Shao, personal communication) correspond to those of the *fucose* operon [Lu and Lin, 1989] which also maps in this region. The fact that *sdaX* also mapped in this region, and was transduced with *sdaB*, suggest that the *sdaX* mutation is in the promoter region of *sdaB* gene.

The 8 kb *Pst*I fragment containing *sdaB* gene showed some homology to *sdaA* DNA, but rather little. The DNA sequence of an open reading frame on a 4.2 kb clone carrying the *sdaB* gene also shows a significant homology with *sdaA* in the coding region (Shao, personal communication). The existence of this homology may indicate that *sdaB* is also a structural gene, if it may be assumed that two structural genes will show more homology than a structural and a regulatory gene. The fact that cells carrying a cloned *sdaB* gene produced very high L-SD activity may also suggest that this is indeed the structural gene. All other evidence, such as the regulation of L-SD#2 being similar to that of β-galactosidase synthesized from the *sdaB* promoter is consistent with this. However, other hypotheses cannot be formally excluded.

7-5. How Unusual Is the Existence of Two Genes Coding for the Same Enzyme Activity

L-SD#1 and L-SD#2 are extremely similar in physical parameters, high Km for substrate, wide pH range, and a requirement for iron and dithiothreitol for the *in vitro* assay. The peptide sequence coded by the *sdaB* gene shows 60% homology with that coded by *sdaA* gene (Shao, personal communication). A detailed comparison may give some idea as to which peptide sequences are important for the structure and function of this enzyme.

The existence of two or more genes coding for the same enzyme activity is not unusual. In some cases, the cell synthesizes two or more such enzymes simultaneously as is the case for the three aromatic biosynthesis isoenzymes, which catalyze the D-arabino heptulonsonate-7-phosphate synthetase reaction [Doy and Brown, 1965]. Another similar case is the isozymes acetohydroxy acid synthase I, II and III which are involved in the synthesis of α-acetolactate and α-aceto-α-hydroxybutyrate, reactions of the L-isoleucine and L-valine biosynthesis pathway [Umbarger, 1987].

In other cases, the cell regulates the enzymes so that they are synthesized in different growth conditions. *E. coli* makes two S-adenosylmethionine synthetases. One of them, coded by *metK*, is made in minimal medium [Greene *et al.*, 1973]; the other, coded by *metX*, is synthesized in LB [Satishchandran, *et al.*, 1990]. More startling is the case of the 2 lysyl-tRNA synthetases which share 95% homology but are coded by two genes, *lysU* and *lysS* and are synthesized under different growth conditions [Clark and Neidhardt, 1990].

PART 8. THOUGHTS ON THE POSSIBLE FUNCTION OF L-SD

L-SD can deaminate L-serine to ammonia and pyruvate. This suggests that the enzyme may be involved in degrading L-serine to avoid toxicity, or to provide carbon and energy, or to provide nitrogen. This section of the thesis discusses some preliminary thoughts about how the enzyme is used.

8-1. Using L-Serine as Carbon and Energy Source

Wild type *E. coli* does not grow with L-serine as the carbon and energy source but does grow in SGL medium. The straightforward explanation of this would be that the wild type strain does not make enough L-SD to permit to growth. Then the inducers, glycine and L-leucine in SGL medium would induce higher L-SD, which would permit the strain to grow. The mutations at *lrp* and *ssd* loci which allow the strains to grow on L-serine alone also increase the L-SD levels greatly. However, both mutations are pleiotropic, and alter cell metabolism extensively [Newman *et al.*, 1981, and Lin *et al.*, 1990]. It is not clear, therefore, that the higher level of L-SD is the (only) change responsible for permitting growth in L-serine.

The strain with an *sdaA* mutation can not grow in L-serine even with glycine and

L-leucine. The plasmid pMES22 carrying the *sdaA* gene increased L-SD production 14 times and gave the wild type, MEW22, and MEW28 the ability to grow with L-serine without inducer. This certainly strengthens the hypothesis that high L-serine deaminase is sufficient for growth on L-serine.

The same pattern is seen when L-SD#2 is the only deaminase made. When L-SD#2 synthesis was established in SGL, the mutant (MEW50 *sdaA*::Cm^r *sdaX*) could grow on L-serine with glycine and L-leucine but not without them. However, the plasmid pMES41 carrying a wild type *sdaB* with a *sdaX* mutation allowed any of strains MEW1, MEW22, MEW28 and MEW50 to grow on L-serine alone. It seems then that a high level of L-SD may be sufficient to confer the ability to grow with L-serine.

If this is true, one could estimate how much L-SD activity is needed for growth on L-serine. The strain MEW1 that has 19 units of L-SD#1 activity while grown on glucose-minimal medium could not grow on L-serine; when inducers provided 110 units [Newman *et al.*, 1985a], the strain grew. L-SD#2 allowed growth with 24 units on glucose-minimal medium with inducers but not with 6 unites without inducers.

The ability of cells to grow with L-serine is usually tested on plates with L-serine, L-isoleucine and L-valine. This tests the minimal requirement for growth on L-serine within 48-72 hours. However the actual growth rate on L-serine might also vary with L-SD activity. In fact the *ssd* mutant grew much faster than the *lrp* mutant on NSIV medium, and also had much higher L-SD. Even more striking were the strains which carried the *sdaA* plasmid and grew faster than any other strain tested.

It is surprising that strain MEW50 *sdaX* with 24 units of L-SD#2 grew on NIVSGL in 48 hours, but a strain with 19 units of L-SD#1 did not grow on NSIV. With such similar amounts of enzyme, one would expect some detectable growth in 48 hours, unless the two L-SDs are more different than earlier considerations indicated. Besides, L-serine would be toxic to the cells but then can be relieved by glycine and L-leucine which can let cell tolerate higher L-serine concentrations. In that case, the fact that the strain MEW1 with 19 units of L-SD activity can not grow on L-serine may be due to L-serine toxicity and not to the rate of L-serine deamination. However this comparison would be more satisfactory if it could be done with cells grown on the same medium.

Cells plated on L-serine may depend on some level of internal pyruvate to start and/or maintain growth. If some threshold level is needed, then 19 units of L-SD might not produce it, but 24 units might.

L-Serine is the precursor for synthesis of other amino acids and the major donor of C1 units. A cell making the huge level of L-SD specified by the high copy number plasmids would be expected to have problems due to depletion of the L-serine pool. However, in this study, I have shown that strains with *sdaA* or *sdaB* plasmids grow as well as the wild type cell. This may be due to the fact both L-SD#1 and L-SD#2 show a high Km, so that L-serine is never depleted to a concentration below that Km. Moreover, the other enzymes, such as STMH, show a low Km for L-serine and that will allow the L-serine required for C1 production, and other uses of L-serine, to be met before L-serine is extensively deaminated.

8-2. Using L-Serine as Nitrogen Source

E. coli can use L-serine, as well as glycine, as nitrogen source. Using glycine as nitrogen source involves combining glycine and C1 to form L-serine and using this L-serine to provide ammonia. Newman and her coworkers suggested this pathway [Newman *et al.*, 1976]. However the lack of a mutation in the L-SD structural gene made the proof that glycine is used in this way impossible.

The results from this study show clearly that both L-SD#1 and L-SD#2 can provide nitrogen from L-serine. That was demonstrated by the fact that the strain with a mutation in *sdaA* (without L-SD#1) could not use L-serine as nitrogen source; when L-SD#2 synthesis was established, the ability to derive nitrogen from L-serine was restored. Then inactivating *sdaB* again made the strain unable to use L-serine as nitrogen source. A plasmid with either of the two genes, pMES22 or pMES41, restored the use of L-serine as nitrogen source, and at a much faster growth rate.

The pathway suggested for deriving nitrogen from glycine involved its conversion to L-serine, which would be subsequently deaminated. This was supported by the fact that the rate of growth with glycine as nitrogen source was increased by also adding L-leucine. Since L-leucine alone is not used as nitrogen source, this was explained as being due to L-SD induction.

A strain with neither L-SD in fact grew only very slowly with glycine as nitrogen, suggesting that this really is the major route for use of glycine. However, the fact that strains without L-SD activity (both *sdaA*-MEW28 and *sdaB*-MEW50) grow slightly on glycine may suggest that there is another, though minor, route for deriving nitrogen from glycine.

Even without L-SD, the cell should be able to metabolize glycine via glycine cleavage (GCV) reactions, which produces ammonia. Then, why would they not use this route to get nitrogen from glycine? The GCV reactions use THF as C1 carrier. It may be that in the absence of L-SD, L-serine accumulates, inhibiting the conversion of glycine and C1-THF to L-serine. Then if the cell has no further way to take the C1-unit from THF, further use of the GCV would be impossible. If the cell can use this route to some extent, this would account for the slow growth rate on glycine as nitrogen source.

The wild type *E. coli* uses either L-serine or glycine as nitrogen source very slowly. The limiting step is probably deamination by L-SD. That is suggested by the fact that growth with glycine is faster than growth with L-serine; and this growth rate can be speeded up by adding a small amount of L-leucine. Furthermore, the high L-SD activity produced from plasmids also can increase the rate of use of either L-serine or glycine as nitrogen source.

The results from this study show that a strain with a high copy number plasmid which can produce a lot of L-SD activity can grow with L-serine as sole carbon, energy and nitrogen source. This is quite interesting because using L-serine as carbon, energy and nitrogen source needs a balance between these functions. In other words, the cell growing on L-serine alone may produce more ammonia or pyruvate than it needs. In this case, the cell might accumulate one of the end products of L-SD, and that apparently would not inhibit growth.

Using glycine as sole carbon, energy and nitrogen source might be different. Wild type *E. coli* can use glycine as nitrogen but not as carbon and energy source. The reason for that might be the speed of synthesis of L-serine, the direct source of carbon and energy source.

8-3. Detoxification of Intracellular L-Serine

L-Serine can inhibit *E. coli* growth [Cosloy and McFall, 1970, Isenberg and Newman, 1974, Uzan and Danchin, 1978]. This was shown to be due to its inhibition of homoserine dehydrogenase [Hama *et al.*, 1990], an enzyme in the isoleucine biosynthesis pathway. This inhibition can be relieved by adding L-isoleucine. In that work, L-serine was added at 1mM , or 0.1 mg/ml, and this was counteracted by adding L-isoleucine. At a much higher L-serine concentration, the L-serine toxicity could not be reversed by L-isoleucine as shown in this study. Because all strains used in this work are derived from MEW1 which carries an *ilvA* deletion, isoleucine (and L-valine) are added to all media. Thus the toxicity described here is not due to an effect on L-isoleucine biosynthesis.

The actual level of L-serine *E. coli* supplied with L-isoleucine and L-valine can tolerate depends on the amount of L-SD it is synthesizing. Wild type *E. coli* can tolerate as much as 2 mg/ml L-serine, and up to at least 10mg/ml, if inducers are provided. The growth rate of a strain without L-SD activity was decreased by even 2 mg/ml and totally inhibited by 10mg/ml. The high L-SD activity from high copy number *sdaA* or *sdaB* plasmids relieved this inhibition. Thus L-SD is responsible for the detoxification of L-serine.

L-Serine is known to inhibit homoserine dehydrogenase [Hama *et al.*, 1990]. Inhibition can be reversed by adding L-threonine and/or L-isoleucine, or in the experiments shown here, by inducing high levels of L-SD. Both L-SD#1 and L-SD#2 show a very high Km. That means that these enzymes probably only function when the cell's L-serine concentration is high. It may be that the metabolic function of L-SD is to maintain the cell's L-serine in a tolerable range.

8-4. Other Possible Functions of L-SD

L-SD clearly functions in the use of L-serine as carbon, energy and nitrogen source, and in the detoxification of L-serine. But if this is its main/only role, it is difficult to understand why the regulation of L-SD is so complex. *SdaA* can be induced by glycine and L-leucine; by DNA damaging agents; by growth at 42°C, by ethanol; by anaerobic growth; by growth in LB and is also regulated by products of two other genes, *ssd* and *lrp*. L-SD activity is also regulated by a post-translational regulation involving the products of at least three genes. Then either all other degrading enzymes must have similarly complex regulation, or there must be a dimension of L-SD metabolic function that I do not understand at this time.

The fact that not only one but two L-SD's exist in *E. coli* and that an *sdaA* gene analogue functions in both *Klebsiella aerogenes* (Nichols, personal communication), and *Salmonella typhimurium* (Shao, personal communication) excludes the possibility that L-SD is an enzyme metabolically useless. Usually, *E. coli* genes are very efficiently organized. This is seen with the D-serine deaminase operon which is expressed only when D-serine is to be degraded [McFall, 1987].

Again, L-SD#2 is only made in rich medium, but a lot of L-SD#1 is also produced in LB. Why make a second L-SD when the first one is being made? This may suggest that L-SD#2 also has some other function that we do not know.

It is clear that a strain without L-SD activity can grow well in minimal medium. L-SD is not essential for its metabolism. The fact that L-serine is the major C1 donor, and that L-SD is induced by DNA-damaging conditions, might suggest that the cell has to regulate its L-serine pool particularly carefully so as to provide sufficient C-1 units for methylation of DNA, and that L-SD is thus involved to some extent in DNA repair.

PART 9. SUMMARY

In this thesis, two genes *sdaA* and *sdaB* are characterized. *SdaA* was shown to code for L-SD#1, the L-SD produced in glucose-minimal medium. *SdaB* may be the structural gene for L-SD#2, the L-SD activity synthesized only in LB medium, and in any case influences its synthesis. The regulation and possible function of these enzymes were also studied.

An insertion mutation in *sdaA* was isolated by λplacMu insertion. The mutation showed no L-SD activity either in the whole cell assay or the extract assay with Fe and DTT. The sequence of this gene indicates that it codes for a 48,000 dalton protein. *lacZ* gene fusions to the *sdaA* gene both in chromosomal DNA by λplacMu insertion and in plasmid by construction showed that this gene was regulated by glycine and

L-leucine, by the *ssd* gene product, by UV irradiation, and by anaerobic growth, i.e. that regulation is the same as described for L-SD previously. The *sdaA* gene was located at 41 minutes on the *E. coli* gene map.

The fusion of the entire *sdaA* gene to the *lacZ* gene indicated that the *sdaA* codes for L-SD#1 activity. That was demonstrated by the fact that the fusion protein carried both L-SD and β-galactosidase activities as judged after purification on an affinity column and gel filtration.

The fusion protein is quite stable. The enzyme deaminated L-serine, but no other amino acids tested formed an α-keto acid in the conditions used for L-SD assay.

The existence of a second L-SD activity was demonstrated by the fact that a strain carrying a mutation in the *sdaA* locus can produce L-SD activity while growing in LB medium. L-SD#2 is very similar to L-SD#1. The mutation which increases the expression of this gene produces enough L-SD to permit the cell to grow on L-serine.

A mutation in *sdaB*, possibly the structural gene for L-SD#2, was also isolated by λplacMu insertion. The strain carrying this mutation produced no L-SD activity as judged either by whole cell or extract assay. The *sdaB* clone showed very high L-SD activity and hybridization experiments located the *sdaB* gene at 60.1 minutes on the *E. coli* gene map.

The known metabolic functions of L-SD include converting L-serine to pyruvate and ammonia, which would allow the cell to use L-serine as carbon and nitrogen source. Deaminating L-serine also can serve to detoxify L-serine in *E. coli*.

REFERENCES

Alfoldi, L., I. Rasko, and E. Kerekes. 1968. L-Serine deaminase of *Escherichic coli*. J. Bacteriol. *96*:1512-1518.

Alfoldi, L., and I. Rasko. 1969. L-Serine deaminating enzymes in *Escherichia coli* crude extracts. FEBS Lett. *6*:73-76.

Andrews, J.C., T.C. Blevins, and S.A. Short. 1986a. Regulation of peptide transport in *Escherichia coli*: induction of the *trp*-linked operon encoding the oligopeptide permease. J. Bacteriol. *165*:428-433.

Andrews, J.C., T.C. Blevins, and S.A. Short. 1986b. *Opp-lac* operon fusions and transcriptional regulation of the *Escherichia coli trp*-linked oligopeptide permease. J. Bacteriol. *165*:434-442.

Aronson. D., P.D. Ravnikar, and R.L. Somerville. 1988. The sequence of the 2-amino-3-ketobutyrate coenzyme A ligase(*kbl*) gene of *E. coli*. Nucleic. Acids Res. *16*:3586.

Bachmann, B.J., 1990. Linkage map of *Escherichia coli* K-12, edition 8. Microbiol. Rev. *54*:130-197.

Baker, T.A., M.M. Howe, and C.A. Gross. 1983. MudX, a derivative of MudI (*lac* Ap^r) which makes stable *lacZ* fusions at high temperature. J. Bacteriol. *156*:970-974.

Beeraj, R.D., J.F. Morris and E.B. Newman. 1978. Factors determining the *in vivo* stability of L-serine deaminase activity in *E. coli* K-12. Can. J. Microbiol. *24*:1607-1613.

Bolivar, F., R.L. Rodriguez, P.J. Greene, M.C. Betlach, H.L. Heyneker, H.W. Boyer, J.H. Cross, and S. Falkow. 1977. Construction and characterization of new cloning vehicles. II. A multi-purpose cloning system. Gene *2*:95-113.

Bremer, E., T.J. Silhavy, J.M. Wiesemann, and G.M. Wienstock. 1984. λplacMu: A transposable derivative of phage λ for creating *lacZ* protein fusions in a single step. J. Bacteriol. *158*:1084-1093.

Bremer, E., T.J. Silhavy, and G.M. Weinstock. 1985. Transposable λplacMu bacteriophages for creating *lacZ* operon fusions and kanamycin resistance insertions in *Escherichia coli*. J. Bacteriol. *162*:1092-1099.

Brown, E.A., R. D'Ari, and E.B. Newman. 1990. A relationship between L-serine degradation and methionine biosynthesis in *Escherichia coli* K-12. J. Gen. Microbiol. *136*:1017-1023.

Calhoun, D.H., R.A. Rimerman, and G.W. Hatfield. 1973. Threonine deaminase from *Escherichia coli*. I. Purification and properties. J. Biol. Chem. *248*:3511-2516.

Casadaban, M.J., and J. Chou. 1984. *In vivo* formation of gene fusions encoding hybrid β-galactosidase protein in one step with a transposable Mu-*lac* transducing phage. Proc. Natl. Acad. Sci. U.S.A. *81*:535-539.

Casadaban, M.J., and S.N. Cohen. 1979. Lactose genes fused to exogenous promotors in one step using a Mu-*lac* bacteriophage: *in vivo* probe for transcriptional control sequences. Proc. Natl. Acad. Sci. U.S.A. *76*:4530-4533.

Chan, T.T.K., and E.B. Newman. 1981. Threonine as a carbon source for *Escherichia coli*. J. Bacteriol. *145*:1150-1153.

Chang A.C.Y. and S.N. Cohen. 1978. Construction and characterization of amplifiable multicopy DNA cloning vehicles derived from the p15A cryptic miniplasmid. J. Bacteriol. *134*:1141-1156.

Clark, R.L. and F.C. Neidhardt. 1990. Roles of the two lysyl-tRNA synthetases of *Escherichia coli*: Analysis of nucleotide sequence and mutant behaviour. J. Bacteriol. *172*:3237-3243.

Cosloy, S.D. and E. McFall. 1970. L-Serine-sensitive mutants of *Escherichia coli* K-12. J. Bacteriol. *103*:840-841.

Cowing, D.W., J.C.A. Bardwell, E.A.Craig, C. Woolford, R.W. Hendrix, and C.A. Gross. 1985. Consensus sequence for *Escherichia coli* heat shock promotors. Proc. Natl. Acad. Sci. U.S.A. *82*:2679-2683.

Datta, P., T.J. Goss, J.R. Omnaas, and R.V.Patil. 1987. Covalent structure of biodegradative threonine dehydratase of *Escherichia coli*: homology with other dehydratases. Proc. Natl. Acad. Sci. USA *84*:393-397.

Doy, C.H., and K.D. Brown. 1965. Control of aromatic biosynthesis: the multiplicity of 7-phospho-2-oxo-3-deoxy-D-arabino-heptulosonate D-erythrose-4-phosphate-lyase (pyruvate phosphorylating) in *Escherichia coli* W. Biochim. Biophys. Acta *104*:377-389.

Dumont, D. 1985. The map position of the *sdaA* and *sdaB* mutations and the cloning of the *ssd* gene in *Escherichia coli* K-12. Thesis of the Master of Science. (Concordia University).

Egan, R.M. and A.T. Phillips. 1977. Requirements for induction of the biodegradative threonine dehydratase in *Escherichia coli*. J. Bacteriol. *132*:370-376.

Folk, J.E. 1970. Carboxypeptidase B. Methods in Enzymol. *19*:504-508.

Fowler, A.V., and I. Zabin. 1983. Purification, structure, and properties of hybrid β-galactosidase proteins. J. Biol. Chem. *258*:14354-14358.

Fraser, J., and E.B. Newman. 1975. Derivation of glycine from threonine in *Escherichia coli* K-12 mutants. J. Bacteriol. *122*:810-817.

Fuller, F., and H. Boedtker. 1981. Sequence determination and analysis of the 3'-region of chicken pro-α1(I) and pro-α2(I) collagen message ribonucleic acids including the carboxyl-terminal propeptide sequences. Biochem. *20*:916-1006.

Germino, J., and D. Bastia. 1984. Rapid purification of a cloned gene product by genetic fusion and site-specific proteolysis. Proc. Natl. Acad. Sci. USA. *81*:4692-4696.

Goncharoff, P., and B.P. Nichols. 1988. Evolution of Aminobenzoate Synthases: Nucleotide sequence of *Salmonella typhimurium* and *Klebsiella aerogenes pabB*. Mol. Biol. Evol. *5*:531-548.

Goncharoff, P., and B.P. Nichols. 1984. Nucleotide sequence of *Escherichia coli pabB* indicates a common evolutionary origin of p-aminobenzoate synthetase and anthranilate synthetase. J. Bacteriol. *159*:57-62.

Goss, T.J. and P. Datta. 1984. *Escherichia coli* K-12 mutation that inactivates biodegradative threonine dehydrogenase by transposon Tn*5* insertion. J. Bacteriol. *158*:826-831.

Goss, T.J., H.P. Schweizer, and P. Datta. 1988. Molecular characterization of the *tdc* operon of *Escherichia coli* K-12. J. Bacteriol. *170*:5352-5359.

Gottschalk, G. (Ed) 2nd. 1986. Bacterial Metabolism. Springer-Verlag. New York Inc.

Gouy, M., and C. Gauthier. 1982. Codon usage in bacteria: correlation with gene expressivity. Nucleic. Acids Res. *10*:7055-7074.

Greene, R.C., J.V.S. Hunter, and E.H. Coch. 1973. Properties of *metK* mutants of *Escherichia coli*. J. Bacteriol. *115*:57-67.

Groisman, E.A., and M.J. Casadaban. 1986. Mini-Mu bacteriophage with plasmid replicons for *in vivo* cloning and *lacZ* gene fusing. J. Bacteriol. *168*:357-364.

Groisman, E.A., B.A. Castilho, and M.J. Casadaban. 1984. *In vivo* DNA cloning and adjacent gene fusing with a mini-Mu-*lac* bacteriophage containing a plasmid replicon. Proc. Natl. Acad. Sci. USA *81*:1480-1483.

Gutterson N.I., and D.J. Koshland. 1983. Replacement and amplification of bacterial genes with sequences altered *in vitro*. Proc. Natl. Acad. Sci. U.S.A. *80*:4894-4898.

Hama, H., Y. Sumita, Y. Katutagi, M. Tsuda, and T. Tsuchiya. 1990. Target of serine inhibition in *Escherichia coli*. Biochem. Biophys. Res. Comm. *168*:1211-1216.

Herendeen, S.L., R.A. VanBogelen, and F.C. Neidhardt. 1979. Levels of major proteins of *Escherichia coli* during growth at different temperatures. J. Bacteriol. *139*:185-194.

Hirshfield, I.N., P.L.Block, R.A. Van Bogelen, and F.C. Nedihardt. 1981. Multiple forms of lysyl-transfer ribonucleic acid synthetase in *Escherichia coli*. J. Bacteriol. *146*:345-351.

Isenberg, S., and E.B. Newman. 1974. Studies on L-serine deaminase in *Escherichia coli* K-12. J. Bacteriol. *118*:53-58.

Jason, M., and P. Schimmel. 1984. Deletion of an essential gene in *Escherichia coli* by site-specific recombination with linear DNA fragments. J. Bacteriol. *159*:783-786.

Joyce, C.M., and N.D.F. Grindley. 1984. Method for determining whether a gene of *Escherichia coli* is essential: application to the *polA* gene. J. Bacteriol. *158*:636-643.

Kiel, J.A.K.W., J.P.M.J. Vossen, and G. Venema. 1987. A general method for the construction of *Escherichia coli* mutants by homologous recombination and plasmid segregation. Mol. Gen. Genet. *207*:294-301.

Kielland-Brandit, M.C., S. Holmberg, J.G.L. Petersen, and T. Nilsson-Tillgren. 1984. Nucleotide sequence of the gene for threonine deaminase (*ilv*1) of *Saccharomyces cerevisiae*. Carlsberg Res. Commun. *49*:567-575.

Kim, K., S.G. Rhee, and E.R. Stadtman, 1985. Nonenzymatic cleavage of proteins by reactive oxygen species generated by dithiothreitol and iron. J. Biol. Chem. *260*:15394-15387.

Kohara, Y., K. Akiyama, and K. Isono. 1987. The physical map of the whole *Escherichia coli* chromosome: application of a new strategy for rapid analysis and sorting of a large genomic library. Cell *50*:495-508.

Kredich, N.M., and G.M. Tomkins. 1966. The enzymatic synthesis of L-cysteine in *Escherichia coli* and *Salmonella typhimurium*. J. Biol. Chem. *241*:4955-4965.

Kunkel. 1985. Rapid and efficient site-specific mutagenesis without phenotypic selection. Proc. Natl. Acad. Sci. USA *82*:488-492.

Lam, H.M., and M.E. Winkler. 1990. Metabolic relationships between pyridoxine (Vitamin B_6) and serine biosynthesis in *Escherichia coli* K-12. J. Bacteriol. *172*:6518-6528.

Lang, B.F., and G. Burger. 1990. A rapid, high resolution DNA sequencing gel system. Anal. Biochem. *188*:176-180.

Lawther, R.P., R.C. Wek, J.M. Lopes,R. Pereira, B.E. Taillon, and G.W. Hatfield. 1987. The complete nucleotide sequence of the *ilvGMEDA* operon of *Escherichia coli* K-12. Nucleic Acids Res. *15*:2137-2155.

Lin, Rongtuan, R. D'Ari, and E.B. Newman. 1990. The leucine regulon of *Escherichia coli*: a mutation in *rblA* alters expression of L-leucine-dependent metabolic operons. J. Bacteriol. *172*:4529-4535.

Lin, R., B. Ernsting, I.N. Hirshfield, R.G. Matthews, F.C. Neidhardt, R.L. Clark, and E.B. Newman. 1991. The *lrp* gene product regulates expression of *lysU* in *E. coli* K-12. Submitted.

Little, J.W., and D.W. Mount. 1982. The SOS regulatory system of *Escherichia coli*. Cell *29*:11-22.

Little, J.W., D.W. Mount, and C.R. Yanisch-Perron. 1981. Purified lexA protein is a repressor of the *recA* and *lexA* genes. Proc. Natl. Acad. Sci. USA *78*:4199-4203.

Lowry, O.H., J. Rosebrough, A.L. Farr, and R. J. Randall. 1951. Protein measurement with Folin phenol reagent. J. Biol. Chem. *193*:265-275.

Lu, Z., and E.C.C. Lin. 1989. The nucleotide sequence of *Escherichia coli* genes for L-fucose dissimilation. Nucleic. Acids Res. *17*:4883-4884.

Maniatis, T., E.F.Fritsch, and J. Sambrook. 1982. In Molecular Cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.

Marceau, M., M. McFall, S.D. Lewis, and J.A. Shafer. 1988. D-Serine dehydratase from *Escherichia coli*: DNA sequence and identification of catalytically inactive glycine to aspartic acid variants. J. Biol. Chem. *263*:16926-16933.

Matsuyama S.I., and S. Mizushima. 1985. Construction and characterization of a deletion mutant lacking *micF*, a proposed regulatory gene for *ompF* synthesis in *Escherichia coli.* J. Bacteriol. *162*:1196-1202.

McFall, E. 1987. The D-serine deaminase operon. pp.1520-1526. In *Escherichia coli* and *Salmonella typhimurium* Cellular and Molecular Biology. Ed. F. C. Neidhardt. American Society for Microbiology, Washington.

McKitrick,J. C., and L.I. Pizer. 1980. Regulation of phosphoglycerate dehydrogenase levels and effect on serine synthesis in *Escherichia coli* K-12. J. Bacteriol. *141*:235-245.

Miller, J.H. (ed). 1972. Experiments in molecular genetics. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.

Morris,J.F., and E.B. Newman. 1980. Map location of the *ssd* mutation in *Escherichia coli* K-12. J. Bacteriol. *143*:1504-1505.

Moskaluk, C., and D. Bastia. 1988. DNA bending is induced in an enhancer by the DNA-binding domain of the bovine papillomavirus E2 protein. Proc. Natl. Acad. Sci. USA *81*:4692-4696.

Mudd, S.H., and G.L. Cantoni. 1964. Biological transmethylation, methyl-group neogenesis and other "one-carbon" metabolic reactions dependent upon tetrahydrofolic acid, p. 1-47. In M. Florkin and E.H. Stotz(ed.), Comprehensive biochemistry, Vol. 15. Elsevier, Amsterdam.

Neidhardt, F.C., and R.A. VanBogelen. 1981. Positive regulatory gene for temperature-controlled proteins in *E. coli*. Biochem. Biophys. Res. Commun. *100*:894-900.

Neidhardt, F.C., and R.A. VanBogelen. 1987. Heat shock response. In *Escherichia coli* and *Salmonella typhimurium* Cellular and Molecular Biology. Ed. Frederick C. Neidhardt. American Society for Microbiology, Washington.

Newman, E.B., D. Ahmad, and C. Walker. 1982a. L-Serine deaminase activity is induced by exposure of *Escherichia coli* K-12 to DNA-damaging agents. J. Bacteriol. *152*:702-705.

Newman, E.B., G. Batist, J. Fraser, S. Isenberg, P. Weyman, and V. Kapoor. 1976. The use of glycine as nitrogen source by *Escherichia coli* K-12. Biochim. Biophys. Acta *421*:97-105.

Newman, E.B., D. Dumont, and C. Walker. 1985a. *In vitro* and *in vivo* activation of L-serine deaminase in *Escherichia coli* K-12. J. Bacteriol. *162*:1270-1275.

Newman, E.B., and V. Kapoor. 1980. *In vitro* studies on L-serine deaminase activity of *Escherichia coli* K-12. Can. J. Biochem. *58*:1292-1297.

Newman, E.B., and B. Magasanik. 1963. The relation of serine-glycine metabolism to the formation of single carbon units. Biochim. Biophys. Acta *78*:437-448.

Newman, E.B., Malik, N., and C. Walker. 1982b. L-Serine degradation in *Escherichia coli* K-12: Directly isolated *ssd* mutants and their intragenic revertants.
J. Bacteriol. *150*:710-715.

Newman, E.B., B. Miller, L.D. Colebrook, and C. Walker. 1985b. A mutation in *Escherichia coli* K-12 results in a requirement for thiamine and a decrease in L-serine deaminase activity. J. Bacteriol. *161*:272-276.

Newman, E.B., J.F. Morris, C. Walker, and V. Kapoor. 1981. A mutation affecting L-serine and energy metabolism in *Escherichia coli* K-12. Mol. Gen. Genet. *182*:143-147.

Newman, E.B., and C. Walker. 1982. L-Serine degradation in *Escherichia coli* K-12: a combination of L-serine, glycine and leucine used as a source of carbon. J. Bacteriol. *151*:777-782.

Newman, E.B., C. Walker, and K. Ziegler-Skylakakis. 1990. A possible mechanism for the *in vitro* activation of L-serine deaminase activity in *Escherichia coli* K-12. Biochem. Cell Biol. *68*:123-128.

Ogawa, H., T. Gomi, K. Konishi, T. Data, H. Nakashima, K. Nose, Y. Matsuda, C. Peraino, H.C. Pitot, and M. Fujioka. 1989. Human liver serine dehydratase: cDNA cloning and sequence homology with hydroxyamino acid dehydratases from other sources. J. Biol. Chem. *264*:15818-15823.

Ogawa, H., K. Konishi, and M. Fujioka. 1989. The peptide sequences near the bound pyridoxal phosphate are conserved in serine dehydratase from rat liver, and threonine dehydratases from yeast and *Escherichia coli*. Biochim. Biophys. Acta *996*:139-141.

Pardee, A.B., and L.S. Prestidge. 1955. Induced formation of serine and threonine deaminase by *Escherichia coli*. J. Bacteriol. *70*:677-674.

Pizer, L.I. 1963. The pathway and control of serine biosynthesis in *Escherichia coli*. J. Biol. Chem. *238*:3934-3944.

Pizer, L.I., and M.L. Potochny. 1964. Nutritional and regulatory aspects of serine metabolism in *Escherichia coli*. J. Bacteriol. *88*:611-619.

Plate, C. 1976. Mutant of *Escherichia coli* defective in response to colicin K and in active transport. Proc. Natl. Acad. Sci. USA *125*:467-474.

Plate, C., and J. Suit. 1981. The *eup* genetics locus of *Escherichia coli* K-12 and its role in H^+/solute symport. J. Biol. Chem. *256*:12974-12980.

Potter, R. V. Kapoor, and E.B. Newman. 1977. Role of threonine dehydrogenase in *Escherichia coli* threonine degradation. J. Bacteriol. *132*:385-391.

Punekar, N.S., and H.A. Lardy. 1987. Phosphoenolpyruvate carboxykinase ferroactivator 1: mechanism of action and identity with glutathione peroxidase. J. Biol. Chem. *262*:6714-6719.

Quay, S.C., T.E. Dick, and D.L. Oxender. 1977. Role of transport systems in amino acid metabolism: leucine toxicity and the branched-chain amino acid transport systems. J. Bacteriol. *129*:1257-1265.

Rainwater, S., and P.M. Silverman. 1990. The cpx proteins of *Escherichia coli* K-12: evidence that *cpxA, ecfB, ssd,* and *eup* mutations all identify the same gene. J. Bacteriol. *172*:2456-2461.

Recsei, P.A., Q.K. Huynh, and E.E. Snell. 1983. Conversion of prohistidine decarboxylase to histidine decarboxylase: peptide chain cleavage by nonhydrolytic serinolysis. Proc. Natl. Acad. Sci. USA *80*:973-983.

Ricca, E., D.A. Aker, and J.M. Calvo. 1989. A protein that binds to the regulatory region of the *Escherichia coli ilvIH* operon. J. Bacteriol. *171*:1658-1664.

Russel, M., and A. Holmgrem. 1988. Construction and characterization of glutaredoxin-negative mutants of *Escherichia coli*. Proc. Natl. Acad. Sci. USA. *85*:990-994.

Ryan, T., and M. Chamberlin. 1983. Transcription analyses with heteroduplex *trp* attenuator templates indicated that the transcript stem and loop structure serves as the termination signal. J. Biol. Chem. *258*:4690-4693.

Sanger, F., S. Nicklen, and A.R. Coulson. 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA. *74*:5463-5467.

Sassenfeld, H.M., and S.J. Brewer. 1984. A polypeptide fusion designed for the purification of recombinant proteins. Bio/technology 2:76-81.

Satishchandran, C., J.C. Taylor, and G.D. Markham. 1990. Novel *Escherichia coli* K-12 mutants impaired in S-adenosylmethionine synthesis. J. Bacteriol. *172*:4489-4496.

Schoenlein, P.V., B.B. Roa, and M.E. Winkler. 1989. Divergent transcription of *pdxB* and homology between the *pdxB* and *serA* gene products in *Escherichia coli* K-12. J. Bacteriol. *171*:6084-6092.

Schweizer, H.P., and P. Datta. 1989. The complete nucleotide sequence of the *tdc* region of *Escherichia coli*. Nucleic. Acids Res. *17*:3994.

Schweizer, H.P., and P. Datta. 1990. Physical map location of the *tdc* operon of *Escherichia coli*. J. Bacteriol. *172*:2825.

Scrimgeour,K.G., and F.M. Huennekens. 1962. Serine hydroxymethylase. Methods in Enzymol. *5*:794-809.

Silhavy, T.J., M.L. Berman, and L.W. Enquist. 1984. Experiments with gene fusions. Cold Spring Harbor laboratory. Cold Spring Harbor. N.Y.

Silverman, P.M. 1985. Host cell-plasmid interactions in the expression of DNA donor activity by F^+ strains of *Escherichia coli* K-12. Bioessays 2:254-259.

Singer, M., T.A. Baker, J.Schnitzler, S.M. Deichel, M. Goel, W. Dove, K.J. Jaack, A.D. Grossman, J.W. Frickson, and C.A. Gross. 1989. A collection of strains containing genetically linked alternating antibiotic resistance for genetic mapping of *Escherichia coli*. Microbiol. Rev. *53*:1-24.

Smith, J.C., R.B. Derbyshire, E.Cook, L. Dunthorne, J. Viney, S.J. Brewer, H.M. Sassenfeld and L.D. Bell. 1984. Chemical synthesis and cloning of a poly(arginine)-coding gene fragment designed to aid polypeptide purification. Gene *32*:321-327.

Southern, E.M. 1975. Detection of specific sequences among DNA fragments separated by gel electrophoresis. J.Mol. Biol. *98*:503-517.

Stauffer, G.V. 1987. Biosynthesis of serine and glycine. In *Escherichia coli* and *Salmonella typhimurium* Cellular and Molecular Biology. Ed. Frederick C. Neidhardt. American Society for Microbiology, Washington.

Stauffer, G.V., M.D. plamann, and L.T. Stauffer. 1981. Construction and expression of hybrid plasmids containing the *Escherichia coli glyA* gene. Gene *14*:63-72.

Stormo, G.D., Schneider, T.D., and Gold, L. 1982. Characterization of translational initiation sites in *E. coli*. Nucleic Acids Res. *10*:2971-2996.

Stoscheck, C.M. 1990. Quantitation of protein. In Guide to Protein Purification. Ed. Muray P. Deutscher. Methods in Enzymology. *182*:50-68.

Stratagene Inc. Bluescript DNA sequencing system. Instruction Manual.

Struck, D.K., D. Maratea, and Ry Young. 1985. Purification of hybrid β-galactosidase proteins encoded by X174 *E lacZ* and *Escherichia coli prlA lacZ*: A general method for the isolation of *lacZ* Fusion polypeptides produced in low amounts. J. Mol. Appl. Gen. *3*:18-25.

Sumantran, V.N., H.P. Schweizer, and P. Datta. 1990. A novel membrane-associated threonine permease encoded by the *tdcC* gene of *Escherichia coli*. J. Bacteriol. *172*:4288-4294.

Tran, V.S., E. Schaeffer, O. Bertrand, R. Mariuzza, and P. Ferrara. 1983. Purification, molecular weight and N-terminal sequence of cystathionine-r-synthase of *Escherichia coli*. J. Biol. Chem. *258*:14872-14873.

Ullmann, A. 1984. A one-step purification method for the purification of enzymatically active hybrid β-galactosidase proteins. Gene. *32*:27-31.

Umbarger, H.E. 1973. Threonine deaminases. Advan. Enzymol. *37*:349-395.

Umbarger, H.E., M.A. Umbarger, and P.L.M. Siu. 1963. Biosynthesis of serine in *Escherichia coli* and *Salmonella typhimurium*. J. Bacteriol. *85*:1431-1439.

Umbarger, H.E. 1987. Biosynthesis of the branched-chain amino acids. In *Escherichia coli* and *Salmonella typhimurium* Cellular and Molecular Biology. Ed. Frederick C. Neidhardt. American Society for Microbiology, Washington.

Uzan, M., and A. Danchin. 1978. Correlation between the serine sensitivity and the derepressibility of the *ilv* genes in *Escherichia coli relA*- mutants. Mol. Gen. Gent. *165*:21-30.

VanBogelen, R.A., P.M. Kelley, and F.C. Neidhardt. 1987. Differential induction of heat-shock, SOS, and oxidation-stress regulons and accumulation of nucleotides in *Escherichia coli*. J. Bacteriol. *169*:26-32.

Vining, L.C., and B. Magasanik. 1981. Serine ultilization by *Klebsiella aerogenes*. J. Bacteriol. *146*:647-655.

Vieira, J., and J. Messing. 1987. Production of single-stranded plasmid DNA. Methods in Enzymology. *153*:3-11.

Walker, G.C. 1984. Mutagenesis and inducible responses to DNA damage in *Escherichia coli*. Microbiol. Rev. *48*:60-93.

Walker, G.C. 1987. The SOS response of *Escherichia* pp.1346-1357. In *Escherichia coli* and *Salmonella typhimurium* Cellular and Molecular Biology. Ed. Frederick C. Neidhardt. American Society for Microbiology, Washington.

Wang, D., and E. Waygood. 1962. Carbon metabolism of C^{14} labeled amino acids in wheat leaves. I. A pathway of glyoxylate-serine metabolism. Plant. Physiol. *37*:826-832.

Weber, R.F., and P.M. Silverman. 1988. The Cpx proteins of *Escherichia coli* K-12: structure of the CpxA polypeptide as an inner membrane component. J. Mol. Biol. *203*:467-478.

Winans, S.C., S.J. Elledge, J.H. Krueger, and G.C. Walker. 1985. Site-directed insertion and deletion mutagenesis with cloned fragments in *Escherichia coli*. J. Bacteriol. *161*:1219-1221.

Wood, J.M. 1981. Genetics of L-proline utilization in *Escherichia coli*. J. Bacteriol. *146*:895-901.

Yanofsky, C. 1960. The tryptophan synthetase system. Bacteriol. Rev. *24*:221-245.

Zhou, Y.N., N. Kusukawa, J.W. Erickson, C.A. Gross, and T. Yura. 1988. Isolation and characterization of *Escherichia coli* mutants that lack the heat shock sigma factor 32. J. Bacteriol. *170*:3640-3649.

Zinder, N.D., and J.D. Boeke. 1982. The filamentous phage (Ff) as vectors for recombinant DNA-a review. Gene *19*:1-10.