INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each original is also photographed in one exposure and is included in reduced form at the back of the book.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.
Characterization of the *Salmonella typhimurium* *opdA* gene, encoding oligopeptidase A: Nucleotide sequence; identity with the *Escherichia coli prIC* gene; and its role in bacteriophage P22 development

Conlin, Christopher Arthur, Ph.D.

Case Western Reserve University (Health Sciences), 1992

Copyright ©1991 by Conlin, Christopher Arthur. All rights reserved.
CHARACTERIZATION OF THE SALMONELLA TYPHIMURIUM opdA GENE, ENCODING OLIGOPEPTIDASE A: NUCLEOTIDE SEQUENCE; IDENTITY WITH THE ESCHERICHIA COLI prIC GENE; AND ITS ROLE IN BACTERIOPHAGE P22 DEVELOPMENT

by

CHRISTOPHER ARTHUR CONLIN

Submitted in partial fulfillment of the requirements for the Degree of Doctor of Philosophy

Thesis Advisor: Prof. Charles G. Miller

Department of Microbiology and Molecular Biology

CASE WESTERN RESERVE UNIVERSITY

January, 1992
CASE WESTERN RESERVE UNIVERSITY
GRADUATE STUDIES

We hereby approve the thesis of

Christopher Arthur Conlin

candidate for the Doctor of Philosophy degree.*

Signed:

[Signatures]

Date Nov. 8, 1991

*We also certify that written approval has been obtained for any proprietary material contained therein.
I grant to Case Western Reserve University the right to use this work, irrespective of any copyright, for the University's own purposes without cost to the University or to its students, agents and employees. I further agree that the University may reproduce and provide single copies of the work, in any format other than in or from microforms, to the public for the cost of reproduction.

Christopher Arthur Conlin
CHARACTERIZATION OF THE *SALMONELLA TYPHIMURIUM* *opdA* GENE,
ENCODING OLIGOPEPTIDASE A: NUCLEOTIDE SEQUENCE; IDENTITY WITH
*ESCHERICHIA COLI* *prlC* GENE; AND ITS ROLE IN BACTERIOPHAGE P22
DEVELOPMENT

Abstract

by

Christopher Arthur Conlin

The *Salmonella typhimurium* *opdA* gene, encoding oligopeptidase A, was mapped to 76 map units on the *S. typhimurium* chromosome. Plasmids containing *opdA* were isolated and the nucleotide sequence of the gene was determined. The *opdA* sequence predicted a 680 amino acid protein containing a metalloprotease Zn$^{2+}$ binding site. It also showed considerable sequence similarity to *S. typhimurium* dipeptidyl carboxypeptidase and to the hormone processing metallopeptidase from rat, EP 24.15. A possible relationship between *S. typhimurium* *opdA* and *E. coli* *prlC*, the locus of suppressors of *lamB* signal sequence mutations, was suggested by their similar map positions. It was found that plasmids carrying the *prlC* gene complemented *opdA* mutations. When the *E. coli* *prlC* gene was sequenced, it was found to be 85% identical to the *S. typhimurium* *opdA* gene. It had been previously reported that *opdA* mutant strains fail to support the normal growth of bacteriophage P22. The studies reported here show that ten to twenty-fold fewer viable phage are produced from infection of an *opdA* mutant host compared to a wild type host. Furthermore, it was found that the *opdA* gene
product was required for the previously uncharacterized proteolytic removal of 20 amino acids from the amino terminus of the P22 gene 7 gene product, a protein required for injection of the phage DNA into the host cell. P22 gene 7 mutants were isolated which functioned effectively without being processed. These mutations were sequenced and found to increase the positive charge of the unprocessed amino terminus of the gene 7 protein, suggesting that the presence of a basic amino terminus was crucial for the proper functioning of the gene 7 protein.
To the memory of my mother, Edith Jean Conlin,

whose commitment to education, curiosity, persistence and love

have been a continuing inspiration to me.
Acknowledgements

I would like to thank my parents, Daniel and Edith, for their love, support, and interest during my perhaps too extensive educational career.

I am grateful for the work and friendship of those whose expertise and advice have contributed to this work, especially: Anne Kukral, for her enthusiasm and skills as a bacterial geneticist; Mark Snavely, for his friendship and help with sequencing and maxicells; Judy Miller, for her skills and knowledge of protein purification and analysis; Tom Silhavy, for his help with prlC; Chris Bazinet, for his knowledge and advice about phage P22; and Greg Hannon, for his friendship and many fruitful discussions.

I especially thank Mary-Jane Lombardo for her editorial skills, for many helpful discussions, for her tolerance, and most of all for her friendship.

I am most grateful to Charles Miller. I could not have hoped for a better advisor. The breadth and depth of his knowledge, his curiosity and enthusiasm, his skills as a teacher, his compassion, patience, support, and friendship have been invaluable. Thank you Charles.

I am grateful for my children; Rebecca, Abigail, and Daniel. Without the patience, love, and support of the most wonderful children I know, this work would have been impossible.

Finally, I thank my wife, Patty. If we can survive this, we can survive anything. Her love, support, understanding, and tolerance made my work possible. After the last six years she deserves canonization at least. But, not being pope, all I can offer is my heart. Thank you Patty, I love you.
Table of Contents

Abstract

Dedication

Acknowledgements

Table of Contents

List of Figures

List of Tables

Chapter 1. General Introduction

Chapter 2. Cloning, Mapping, and Nucleotide Sequence of \textit{opdA} the Gene for Oligopeptidase A from \textit{Salmonella typhimurium}

Introduction

Materials and Methods

Results

Discussion

Chapter 3. \textit{opdA}, Encoding Oligopeptidase A, and \textit{priC}, Encoding Suppressors of Signal Sequence Mutations, are the Same Gene

Introduction

Materials and Methods

Results

Discussion

Chapter 4. Role of the Metalloendopeptidase Oligopeptidase A in Phage P22 Development

Introduction
Table of Contents cont.

Materials and Methods 93

Results 110

Discussion 141

Appendix. The Effect of *Salmonella ion* mutations on Protein Degradation in Multiply Peptidase Deficient Strains 149

Introduction 150

Materials and Methods 152

Results and Discussion 156

References. 167
List of Figures

2.1. Restriction Endonuclease Maps of \textit{opdA}^{+} Plasmids and Site of Insertions in pCM127 36

2.2. SDS-Polyacrylamide Gel Electrophoresis of Extracts of Strains Containing \textit{opdA}^{+} Plasmids 38

2.3. Gene Products Expressed by \textit{opdA} Plasmids and \textit{opdA} Plasmids Containing Insertions 41

2.4. Southern Hybridization to Show Replacement of Chromosomal \textit{opdA} by \textit{opdA10}:Mu dJ 43

2.5. Map Positions of \textit{opdA} on the \textit{S. typhimurium} Chromosome 47

2.6. \textit{opdA} Sequencing Strategy 50

2.7. Nucleotide Sequence of \textit{opdA} Region and Predicted Amino Acid Sequence of Oligopeptidase A 52

2.8. Amino Acid Sequence Comparisons of Oligopeptidase A, Dipeptidyl Carboxypeptidase, and Rat Metalloendopeptidase EC 3.4.24.15 56

2.9. Alignment of Deduced Amino Acid Sequences of Oligopeptidase A, Dipeptidyl Carboxypeptidase, and Rat Metalloendopeptidase EC 3.4.24.15 58

3.1. Location of \textit{Tn1000} Insertions in pTC100 74

3.2. \textit{prlC} Sequencing Strategy 77

3.3. \textit{prlC} and \textit{opdA} Sequences and Translation 79

3.4. Southern Hybridization of \textit{Salmonella} and \textit{E. coli} Chromosomal DNA with \textit{prlC} and \textit{opdA} Probes 84
List of Figures cont.

4.1. Plasmids Containing P22 DNA 101

4.2. Restriction Endonuclease Digestion of DNA from opdA suppressing P22 Phage Mutants 113

4.3. P22 Plaques on opdA\textsuperscript{+} and opdA\textsuperscript{-} hosts 115

4.4. Mapping the opdA Suppressing Mutation in Mutant P22 4a by Marker Rescue 119

4.5. P22 Sequencing Strategy 122

4.6. Nucleotide Sequence and Predicted Open Reading Frames of P22 Genes I4 and 7 124

4.7. Amino Terminal Region of Gene 7 from Mutant Phage 127

4.8. SDS-Polyacrylamide Gel Electrophoresis of Proteins from P22 Particles Grown on opdA\textsuperscript{+} and opdA\textsuperscript{-} Hosts 130

4.9. Expression and Degradation of Wild Type gp7 from pCM187 in opdA\textsuperscript{+} and opdA\textsuperscript{-} Hosts 133

4.10. Processing of gp7 136

4.11. Degradation of gp7 is lon Independent 138

4.12. Mutant gp7 from P22 4a Although not Processed is Still Rapidly Degraded in an opdA Independent Manner 140

4.13. Hydrophobicity and Hydrophobic Moment of Processed Wild Type gp7 146

A.1. Combined Effect of lon\textsuperscript{-} and Peptidase Deficiency on Puromycyl Fragment Degradation 158

A.2. Linkage Map of Salmonella lon Mutations 160
List of Figures cont.

A.3. Three Point Cross to Determine the Relative Order of lon, zba-883::Tn10, and apeB 162

A.4. Effect of lon on the Rapid Accumulation of Acid Soluble Peptides 165
List of Tables

1.1. Peptide Bond Hydrolyzing Enzymes in E. coli and S. typhimurium 3
2.1. Bacterial Strains and Plasmids 23
2.2. Primers Used to Sequence opdA 31
2.3. Relative Efficiency of Transduction of asd\(^+\) and xylB\(^+\) by Induced opdA10::Mu dP22 Lysates 44
3.1. Bacterial Strains and Plasmids 68
3.2. AcAla\(_4\) Hydrolysis by Soluble Extracts of Strains Containing priC or opdA Plasmids 71
3.3. Comparison of opdA and priC Promoter Regions to Known Heat Shock Promoters 82
4.1. Bacterial and Phage Strains and Plasmids 94
4.2. Primers Used to Sequence Phage P22 109
4.3. Marker Rescue of opdA-Independent Mutations by Plasmids Containing Mutant Phage P22 DNA 120
A.1. Bacterial Strains Used 153
Chapter 1

General Introduction
At least 34 peptide bond hydrolyzing enzymes have been identified in *Salmonella typhimurium* or *Escherichia coli*. Although general roles in cellular metabolism have been described for these enzymes, and specific roles have been assigned to many of them, for some no function is known. In this introduction I will first briefly review the cellular functions which involve peptide bond hydrolyzing enzymes. Second, I will discuss some of the strategies used to identify these enzymes and the genes which encode them. Finally, I will present the rationale for characterizing in detail one of these enzymes, namely oligopeptidase A, the product of the *S. typhimurium* gene *opdA*.

**Catabolic functions of peptide bond hydrolysis in *E. coli* and *S. typhimurium*.**

*E. coli* and *Salmonella* are able to use exogenous peptides as sources of amino acids, nitrogen, and carbon. Peptides must first be transported across the cell membrane. To this end, these bacteria have developed a complex peptide transport apparatus composed of at least three separate transport systems: a dipeptide permease, a tripeptide permease, and an oligopeptide permease (Gibson, Price, and Higgins, 1984). After transporting peptides, the cell must degrade them to their constituent amino acids. Both *E. coli* and *S. typhimurium* contain a variety of cytoplasmic enzymes which are active against peptide substrates (see Table 1.1) (Miller and MacKinnon, 1974; Miller and Schwartz, 1978). In addition to the cytoplasmic peptidases, these bacteria possess at least two periplasmic peptide hydrolases: protease III (Cheng and Zipser, 1979) encoded by *prt* (Cheng *et al.*, 1979), and the *degP* gene product (Strauch *et al.*, 1989). They also contain an outer membrane endopeptidase, protease VII, encoded by *ompT* (Sugimura and Nishihara, 1988; Sugimura and Higashi, 1988), and at least three membrane bound proteases: protease IV (the *sppA* gene product) (Ichihara *et al.*, 1986), protease V (Pacaud, 1982), and protease VI (Palmer and St. John, 1987). It is possible that these cell envelope
<table>
<thead>
<tr>
<th>Enzyme</th>
<th>Substrate</th>
<th>Gene, Locus</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Soluble Protease</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Do</td>
<td>casein, globin</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Re</td>
<td>casein, globin, oxidized glutamine synthetase</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mi</td>
<td>casein, globin</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fa</td>
<td>casein, globin</td>
<td></td>
<td></td>
</tr>
<tr>
<td>So</td>
<td>casein, globin, lipoprotein signal peptide, oxidized glutamine synthetase</td>
<td></td>
<td></td>
</tr>
<tr>
<td>La</td>
<td>casein, globin, AN, SulA, RcsA, LacZX90</td>
<td>lon (degT, capR)</td>
<td>yes</td>
</tr>
<tr>
<td>Ci</td>
<td>insulin</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Oligopeptidase A</td>
<td>Z-AAL-pNA, lipoprotein signal peptide, minimally pentapeptides</td>
<td>opdA</td>
<td>yes</td>
</tr>
<tr>
<td><strong>Protease I</strong></td>
<td><strong>NAPNE</strong></td>
<td>apeA</td>
<td></td>
</tr>
<tr>
<td><strong>Protease II</strong></td>
<td><strong>BAEE</strong></td>
<td>tip</td>
<td>yes</td>
</tr>
<tr>
<td><strong>ISP-L-Ecc</strong></td>
<td><strong>Z-AAL-pNA</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Multi-component ATP dependent protease (Ti)</strong></td>
<td>casein, globin</td>
<td>clpA and clpP</td>
<td>yes</td>
</tr>
<tr>
<td><strong>HflA</strong></td>
<td><strong>λcII protein</strong></td>
<td>hflA</td>
<td>yes</td>
</tr>
<tr>
<td><strong>Periplasmic Protease</strong></td>
<td><strong>insulin, α complementing fragment of β-galactosidase</strong></td>
<td><em>prt</em></td>
<td>yes</td>
</tr>
<tr>
<td><strong>DegP</strong></td>
<td></td>
<td><em>degP (htrA)</em></td>
<td>yes</td>
</tr>
<tr>
<td><strong>PBP3 C-terminal processing enzyme</strong></td>
<td><strong>PBP 3</strong></td>
<td><em>prc</em></td>
<td>yes</td>
</tr>
<tr>
<td><strong>alkaline phosphatase isozyme converting enzyme</strong></td>
<td><strong>alkaline phosphatase</strong></td>
<td><em>iap</em></td>
<td>yes</td>
</tr>
<tr>
<td><strong>Membrane bound Protease</strong></td>
<td><strong>lipoprotein signal peptide, Z-Val-ONP</strong></td>
<td><em>sppA</em></td>
<td>yes</td>
</tr>
<tr>
<td><strong>Protease IV</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Protease V</strong></td>
<td>casein, Z-Phe-ONP</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Protease VI</strong></td>
<td><em>E. coli</em> membrane proteins</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Protease VII</strong></td>
<td>paired basic residues</td>
<td><em>ompT</em></td>
<td>yes</td>
</tr>
<tr>
<td><strong>Signal Peptidase I</strong></td>
<td>precursors of secreted proteins</td>
<td><em>lep</em></td>
<td>yes</td>
</tr>
<tr>
<td>Enzyme</td>
<td>Substrate</td>
<td>Locus</td>
<td>Sequence</td>
</tr>
<tr>
<td>------------------------</td>
<td>----------------------------------</td>
<td>-------</td>
<td>----------</td>
</tr>
<tr>
<td>Signal Peptidase II</td>
<td>prolipoprotein</td>
<td>lsp</td>
<td>yes</td>
</tr>
<tr>
<td>Peptidases</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dipeptidase</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Peptidase D</td>
<td>broad specificity, dipeptides</td>
<td>pepD</td>
<td>yes</td>
</tr>
<tr>
<td>Peptidase E</td>
<td>Asp-X</td>
<td>pepE</td>
<td></td>
</tr>
<tr>
<td>Peptidase G</td>
<td>Gly-Gly</td>
<td>pepG</td>
<td></td>
</tr>
<tr>
<td>Peptidase Q</td>
<td>X-Pro</td>
<td>pepQ</td>
<td>yes</td>
</tr>
<tr>
<td>Aminotripeptidase</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Peptidase T</td>
<td>broad specificity, tripeptides</td>
<td>pepT</td>
<td>yes</td>
</tr>
<tr>
<td>Aminopeptidase</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Peptidase A</td>
<td>broad specificity</td>
<td>pepA (serB)</td>
<td>yes</td>
</tr>
<tr>
<td>Peptidase B</td>
<td>broad specificity</td>
<td>pepB</td>
<td></td>
</tr>
<tr>
<td>Peptidase M</td>
<td>Met-X-Y</td>
<td>pepM</td>
<td>yes</td>
</tr>
<tr>
<td>Peptidase N</td>
<td>broad specificity, naphthylamides</td>
<td>pepN</td>
<td>yes</td>
</tr>
<tr>
<td>Peptidase P</td>
<td>X-Pro-Y</td>
<td>pepP</td>
<td>yes</td>
</tr>
<tr>
<td>Carboxypeptidase</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dipeptidyl carboxypeptidase</td>
<td>minimally tetrapeptides</td>
<td>dcp</td>
<td>yes</td>
</tr>
</tbody>
</table>

Abbreviations: Z-AAL-pNA, benzylxycarbonyl Ala Ala Leu p-nitroanilide; NAPNE, N-acetyl Phe β-naphthyl ester; BAEE, N-benzoyl Arg ethyl ester; Z-Val ONP, benzylxycarbonyl Val p-nitrophenol ester; Z-Phe ONP, benzylxycarbonyl Phe p-nitrophenol ester. See Miller (1987), and text for references.
enzymes are involved in the utilization of extracellular peptides.

Another catabolic function of peptidases and proteases is the turnover of cellular proteins. Most proteins in *E. coli* and *Salmonella* have long half-lives, longer than the generation time of a cell (reviewed in Miller, 1987). However, under conditions of nutrient starvation, for example limiting carbon, nitrogen, phosphate, amino acids, the extent of degradation of normally stable proteins increases. *In vivo* experiments with pulse labelled cells have shown that about 20-30% of the cellular proteins are degraded to acid soluble peptides during carbon starvation (Pine, 1973; Yen *et al.*, 1980). Under nitrogen starvation conditions as much as 50% of the proteins are degraded (Miller, 1987). Experiments with strains lacking many of the broad specificity peptidases have shown that these peptidases are responsible for the final steps in this starvation induced proteolysis (Yen *et al.*, 1980). However, the proteases involved in the earlier steps of this degradation have not been identified.

Although most proteins are stable in these organisms, experiments by Yen *et al.* (1980a) suggest that at least 20% of the peptide bonds formed are rapidly degraded. When a wild type strain is given a pulse of radiolabelled amino acid, it is rapidly and stably incorporated into acid precipitable proteins. However, when a multiply peptidase deficient strain is given a pulse of radiolabelled amino acid, it is initially incorporated into acid precipitable protein, but within about 2 minutes, 20% becomes and remains acid soluble (Yen *et al.*, 1980a). The interpretation given these findings was that the 20% which remains acid soluble in the multiply peptidase deficient strain represents rapidly turned over peptides, which require the peptidases to break them down to amino acids so they can be reincorporated into stable proteins. Although it is not known what this rapidly turned over fraction is comprised of, there are several likely candidates, including
signal peptides.

Many proteins destined for the cell membrane or the periplasm are synthesized with a 20 to 40 amino acid N-terminal signal peptide which is rapidly removed during the course of protein localization. These signal peptides do not accumulate in the cell but are themselves degraded (Hussain et al., 1982). It has been shown that the major activity which degrades the lipoprotein signal peptide is the membrane bound protease, protease IV, the sppA gene product (Ichihara et al., 1984; Ichihara et al., 1986). Novak, Ray, and Dev (1986) have shown in vitro that there are also two cytoplasmic lipoprotein signal peptide hydrolyzing activities. Oligopeptidase A, the opdA gene product, accounts for 90% of this cytoplasmic activity, while protease So accounts for the remainder of the in vitro activity.

Another potential source of rapidly degraded polypeptides is the premature termination of translation. Manley (1978) has shown that in the case of β-galactosidase, 30% of the translation initiations terminate prematurely. Most of the resulting peptide fragments are unstable. Similarly, many fragments produced by genes containing nonsense mutations are unstable (Lin and Zabin, 1972). The degradation of one of these, the product of the ochre mutation lacZ90, has been extensively studied. LacZ90 has a half life of 7.5 minutes in wild type, exponentially growing cells (Goldschmidt, 1970). This half life is increased to 60 minutes in cells containing lon mutations (Bukhari and Zipser, 1973). The lon gene encodes an ATP-dependent protease (Chung and Goldberg, 1981). Kowit and Goldberg (1977), McKnight and Fried (1981; 1983) and Wang and Fried (1987) have shown that the LacZ90 fragment first undergoes an endoproteolytic cleavage which generates a large, relatively stable intermediate. However, subsequent steps in its degradation are not known.
Treatment of cells with sublethal amounts of puromycin leads to the production of protein fragments which are analogous to premature termination fragments, and nonsense mutation fragments and consequently rapidly degraded (Goldberg, 1972). Puromycin is incorporated into the growing peptide chain because it resembles a charged amino acyl tRNA, however after its incorporation, growth of the nascent chain stops. Maurizi et al. (1981) showed that insertional inactivation of the lon gene decreases the rate of degradation of puromycyl fragments by half. Miller and Green (1981) have shown that multiply peptidase deficient strains are similarly defective in puromycyl fragment degradation. Presumably the peptidases act at the final stages of the degradation of these fragments, and the lon gene product acts early in this process.

Other rapidly degraded proteins include misfolded proteins. Canavanyl proteins have been used to study the degradation of such misfolded proteins. Cells can incorporate canavanine into proteins in the place of arginine. Like termination and puromycyl fragments, such canavanine containing proteins are rapidly degraded, and lon mutations also decreased the rate if their degradation by half (Maurizi et al., 1985). Presumably the canavanyl proteins are improperly folded and thus more susceptible to proteolysis. Other potential sources of misfolded proteins include proteins which are lacking substrates, cofactors, or the other subunits of multimeric enzymes, and proteins damaged by heat, age, or oxygen radicals (see Rivett, 1990). One example of a damaged protein which is subject to proteolysis is glutamine synthetase. It has been shown that oxidized glutamine synthetase is proteolytically unstable (Levine et al., 1981). It has also been shown that two enzymes are able to degrade the oxidized form specifically, namely protease So, and protease Re (Roseman and Levine, 1987; Lee et al., 1988; Park et al., 1988). The removal of heat damaged proteins is presumably one function of the heat-shock response.
It has been shown that heat shock mutants (*htpR, dnaK, dnaJ, and grpE*) are defective in proteolysis (Straus et al., 1988) and at least two proteases are encoded by heat shock genes, *lon* (Goff et al., 1984), and *clpP* (Kroh and Simon, 1990).

**Proteolytic maturation in *E. coli* and *S. typhimurium***. Another broad area of activity in which peptide bond hydrolyzing enzymes play central roles is the maturation of proteins by proteolytic cleavage. One of these central processes is the removal of the amino terminal methionine from many cellular proteins. All protein synthesis in *E. coli* begins with formyl-methionine, however very few proteins have been isolated which contain the formyl group, and many lack the N-terminal methionine (see Miller, 1987). Adams (1968) showed that the formyl group is removed first, by a deformylase.

Subsequently, depending on which amino acid is second, the methionine may be removed by a specific methionine aminopeptidase. This enzyme, and its gene, *pepM*, have been identified in both *E. coli* (Ben-Bassat et al., 1987) and *S. typhimurium* (Miller et al., 1987; Movva et al., 1990). Consistent with its proposed central role in the cell, *pepM* has been shown to be a vital gene (Miller et al., 1989; Chang et al., 1989).

As mentioned earlier, many proteins exported from the cytoplasm contain a short amino terminal peptide which is removed during export. Two enzymes, signal peptidase I, the *lep* gene product (Wolfe, Zwizinski, and Wickner, 1983), and signal peptidase II, the *lsp* gene product (Tokunaga et al., 1985), have been identified which remove these signal peptides from their precursor proteins. Signal peptidase II is specific for glyceride modified proteins, such as *E. coli* lipoprotein, a major component of the outer membrane (Tokunaga et al., 1982). Signal peptidase I is responsible for the removal of signal peptides from other proteins (Wolfe, Wickner, and Goodman, 1983). Both *lep* and presumably the *lsp* genes are essential (Date, 1983).
Both N-terminal methionine removal and signal peptide processing are relatively general processes, with one peptidolytic enzyme acting on many substrates. Although not common in *E. coli* or *Salmonella*, at least three more specific post-translational proteolytic processing reactions have been described. The most dramatic of these is the processing of penicillin amidase in *E. coli*. This periplasmic heterodimeric enzyme is synthesized as an inactive 90kD polyprotein. After removal of the 26 amino acid signal peptide, this precursor is cleaved into three fragments: a 24kD amino terminal fragment which becomes the α subunit, a 62kD carboxyl terminal fragment which becomes the β subunit, and a 54 amino acid connecting peptide (Schumacher *et al.*, 1986). The enzymes responsible for this processing are not known. A second case of post-translational processing has been inferred from the observation that alkaline phosphatase is present as isozymes with differing electrophoretic mobility, probably the result of proteolytic processing. A gene, *iap*, has been identified which may play a role in generating isozymes. An *iap* mutant lacks isozymic variants (Nakata *et al.*, 1978; Ishino *et al.* 1987). A third example is the removal of 28 amino acids from the C-terminus of penicillin binding protein 3, the *fisI* gene product (Nagasawa *et al.*, 1989). The enzyme responsible for this modification is a membrane associated protease, encoded by the *prc* gene (Hara *et al.*, 1991). As with the isozymes of alkaline phosphatase, the importance of this processing is not clear. The unprocessed penicillin binding protein III is still functional (Nagasawa *et al.*, 1989).

In contrast to the apparent rarity of processing of bacterial proteins, post-translational processing of viral structural proteins is very common. Perhaps because of the necessary genetic economy of viruses, or because of the environment in which they reproduce, most bacteriophage, and many eukaryotic viruses, have structural proteins
which undergo some proteolytic processing. A major exception has been the lambdoid 
Salmonella phage P22 (Botstein et al., 1973; Casjens and King, 1974). Whereas at least 
four phage λ structural proteins (gpB, C, Nu3, and H) are processed during phage growth 
(Georgopoulos et al., 1983; Katsara, 1983), there were no known processing steps in 
phage P22 development before the work reported in this dissertation. The phage λ 
processed proteins include the head scaffolding protein. Both P22 and λ phage heads are 
assembled on a protein scaffold, gpNu3 for phage λ, gp8 for phage P22. In λ 
development, gpNu3 is degraded after the prohead has assembled onto it (Katsura, 1983). 
In P22, gp8 exits the nascent prohead intact and is used to assemble other phage heads 
(King and Casjens, 1974). The enzymes responsible for the proteolytic processing of 
phage λ structural proteins are not known.

Regulation by proteolysis. Although regulation by proteolysis does not appear to 
be common in E. coli and Salmonella, several examples have been described. Regulation 
by Lon protease has been especially well characterized (for recent review see Gottesman 
1989). lon mutations confer increased sensitivity to ultraviolet radiation as a result of 
failure to degrade the normally short lived SulA protein, an inhibitor of cell division 
(Mizusawa and Gottesman, 1983). Also, lon" mutants form mucoid colonies which is 
apparently the result of the failure of these strains to degrade the RcsA protein. RcsA is a 
normally short lived positive activator of colanic acid capsular polysaccharide synthesis 
(Torres-Cabassa and Gottesman, 1987). Lon protease has also been shown to degrade the 
phage λ regulatory protein gpN, both in vivo and in vitro (Gottesman et al., 1981; 
Maurizi, 1987).

Proteolysis also plays a role in the lysis/lysogeny decision that must be made by 
iinfecting phage λ. Increased stability of cII protein leads to increased frequency of
lysogeny, conversely, low levels of cII protein lead to increased frequency of lysis. Host mutations at two loci, hflA and hflB, affect the stability of cII protein (see Wulf and Rosenberg, 1983). The hflA locus has been cloned and shown to include three genes which encode the subunits of the HflA enzyme. This enzyme has been purified and shown to degrade λ CII protein in vitro (Cheng et al., 1988). Another well studied regulatory system involves the RecA protein. However, experiments by Little (1984) suggest that the RecA protein is not itself a protease, rather it appears to activate the autoproteolytic activity of its target proteins, eg. LexA and λ repressor.

From the foregoing discussion it is clear that peptide hydrolyzing enzymes are involved in a variety of important and complex cellular processes. In order to elucidate these processes it is necessary to identify the proteolytic enzymes involved and the specific roles they play. This enterprise requires a combination of genetic and biochemical techniques, such that in vivo effects can be correlated with in vitro activities. Table 1.1 lists the known peptide bond hydrolyzing enzymes in E. coli and S. typhimurium. The next section will discuss some of the genetic and biochemical strategies which have been used to identify these enzymes.

Strategies used to identify peptide bond hydrolyzing enzymes and the genes that encode them. General protease and peptidase substrates have been used in surveys of peptide bond hydrolyzing enzymes. Goldberg and colleagues (Goldberg et al., 1983; Sreedhara Swamy and Goldberg, 1981) used radiolabelled casein, globin, and insulin to purify eight soluble endoproteases from E. coli. Miller and coworkers (Miller and MacKinnon, 1974; Miller and Schwartz, 1978) identified four broad specificity peptidases in extracts of E. coli and S. typhimurium by using small peptides as substrates. Peptide hydrolyzing activities were identified after non-denaturing gel electrophoresis of cell
extracts. After incubation of gels with a peptide solution in an agar overlay, bands of hydrolyzing activity were detected by coupling the release of free amino acids to the oxidation of o-dianisidine through amino acid oxidase (Lewis and Harris, 1967). As these and other studies show, *E. coli* and *Salmonella* are replete with peptide bond hydrolyzing enzymes, many of which have overlapping substrate specificities (Table 1.1). This characteristic has complicated the task of identifying the functions of specific enzymes. Simply inactivating the gene which encodes for a particular peptidase or protease often has no noticeable effect on cell growth or viability. (Obvious exceptions are pepM and lon).

In spite of the overlapping substrate specificity of these enzymes, relatively specific substrates have been identified for some of them. Some of these substrates are chromogenic peptides, which are composed of a chromophore, such as *p*-nitroaniline or \( \beta \)-naphthyl amine, joined to an amino acid in an amide or ester linkage. Several enzymes and their chromogenic substrates are: peptidase N and Ala \( \beta \)-naphthyl amide (Miller and MacKinnon, 1974), ISP-L-Eco and benzylxycarbonyl Ala Ala Leu *p*-nitroanilide (Z-AAL-pNA) (Strongin *et al.*, 1979), protease IV and benzylxycarbonyl Val \( \alpha \)-naphthyl ester (Ichihara *et al.*, 1986), and protease V and benzylxycarbonyl Phe nitrophenyl ester (Pacaud, 1982). One of the limitations of the use of chromogenic peptides is that the peptidase must be able to attack the amino acid/chromophore bond to release the chromophore. Fortunately, the utility of these substrates can be expanded by coupling the hydrolysis of internal peptide bonds to the release of chromophore through the use of an aminopeptidase. For example, the cleavage of Z-AAL-pNA between Ala and Leu by oligopeptidase A generates Leu-pNA, which is then a substrate for aminopeptidase N cleavage, releasing the chromophore, *p*-nitroaniline.
The release of free amino acids as a result of the hydrolysis of peptides can be detected by coupling the oxidation of the amino acid released to the oxidation of o-dianisidine through the enzyme amino acid oxidase (Lewis and Harris, 1967). For example, peptidase E has been detected by its ability to hydrolyze the dipeptide, Asp-Leu (Carter and Miller, 1984). Alternatively, if the peptide substrate is blocked at the amino terminus, ninhydrin can be used to detect the free amino group which is exposed after the peptide is cleaved. Both dipeptidyl carboxypeptidase and oligopeptidase A have been identified by their ability to hydrolyze N-acetyl Ala₄ (AcAla₄) and release ninhydrin reactive products (Vimr and Miller, 1983). Other specific substrates which have been used to detect peptide hydrolyzing enzymes include the α complementing fragment of β-galactosidase (protease III) (Cheng and Zipser, 1979), radiolabelled lipoprotein signal peptide (oligopeptidase A and protease So (Novak et al., 1986)), and radiolabelled E. coli membrane proteins (protease VI) (Palmer and St. John, 1987).

These and other assays have been used to gain genetic access to the peptidases and proteases responsible for particular substrate hydrolysis. For example, pepN mutations were identified in colonies which failed to stain with Ala β-naphthyl amide (Miller and MacKinnon, 1974). pepE mutations were identified in a micro assay screen for failure of cell extracts to hydrolyze Asp-Leu (Carter and Miller, 1984). pepT mutants were isolated in similar micro assay screens for strains which failed to hydrolyze Met-Ala-Ser or Met-Gly-Gly (Strauch and Miller, 1983). Mutations in prt, which encodes protease III, were identified in a micro assay screen for the failure of cell extracts to degrade α complementing fragments of β-galactosidase (Cheng et al., 1979). Similarly, opdA and dcp mutations were identified by screening for failure to degrade N-Acetyl Ala₃ (Vimr and Miller, 1983) and AcAla₄ (Vimr et al., 1983).
In addition to mutations which inactivate peptidase genes, genetic access has been gained to some of these enzymes by isolating plasmids carrying the genes which encode them. Strains carrying such plasmids often overproduce the enzyme simply as a result of increased gene dosage. Consequently plasmid libraries can be screened for clones which confer increased levels of substrate hydrolyzing activity. This strategy has been used to isolate plasmids carrying spp4 (protease IV) using benzylloxycarbonyl Val α-naphthyl ester (Ichihara et al., 1986), and opd4 (oligopeptidase A) using Z-AAL-pNA.

In addition to activity assays, other more physiological screens and selections have been used to isolate peptidase and protease mutants and clones. The selection for degT mutations is a classic example of screening for loss of a particular physiological proteolytic function (Bukhari and Zipser, 1973). Small N-terminal fragments of β-galactosidase, α donors, are able to combine with C-terminal fragments of β-galactosidase, α acceptors, to produce active β-galactosidase. However, in the cell most of the α donors are short lived. In order to isolate mutations which slowed the degradation of these fragments, Bukhari and Zipser (1973) selected for β-galactosidase activity in strains containing stable α acceptor and a labile α donor. The mutations which arose from this selection were originally called degT but later were shown to be at the same locus as previously reported lon mutations (Gottesman and Zipser, 1978).

Other physiological screens have allowed isolation of peptidase mutants by identifying strains which fail to use particular peptides as sources of amino acids or nitrogen. As strains acquire an increasing number of peptidase mutations, the spectrum of peptides which they can use becomes more limited. This property has been used to screen for mutants defective in the ability to use particular peptides. For example, pepD mutants were identified in a screen for pepN pepA pepB strains which were unable to
use the dipeptide Leu-Leu as a leucine source (Miller and MacKinnon, 1974). Similarly \(dcp\) mutants were isolated as strains unable to use AcAla\(_3\) as a sole nitrogen source (Vimr and Miller, 1983), and \(opdA\) mutants were isolated as \(dcp^{-}\) strains which failed to use AcAla\(_4\) as a sole nitrogen source (Vimr et al., 1983).

Other strategies to identify peptidase and protease genes and clones have relied on the suppression of mutations in one peptidase gene by alterations in a different peptidase gene. Two general mechanisms can cause one peptidase to suppress mutations in another, namely hyper-expression and altered specificity. Hyper-expression can result from regulatory mutations which increase the expression of a chromosomal gene. For example, \(pepM100\) an allele of the gene for peptidase M, was identified as a mutation which allowed a multiply peptidase deficient strain to use Met-Gly-Gly as a methionine source. Because Met-Gly-Gly is a poor substrate for peptidase M, peptidase deficient strains with the wild type \(pepM\) allele can not use it. However, the \(pepM100\) mutation leads to a 20 fold increase in peptidase M levels, by altering the -35 region of the promoter. The resulting increased amount of peptidase M hydrolyses enough Met-Gly-Gly to provide sufficient methionine for growth (Miller et al., 1987).

Plasmids carrying the structural genes for peptidases have been identified by pseudo-suppression. Both \(pepT\) (Miller et al., 1991) and \(opdA\) carrying plasmids have been identified in screens for plasmids which confer the ability to use Met-Gly-Gly as a methionine source on a multiply deficient strain. Presumably, these phenotypes also are the result of increased levels of peptidase T and oligopeptidase A because of the increased gene dosage provided by the plasmids.

Although one would expect that the inactivation of a peptidase or protease could be suppressed by altering the specificity of a different one, no clear examples have been
reported. One possible example of suppression due to altered specificity, discussed in a later chapter, is the suppression of signal sequence mutations by mutations in *prtC*.

Reverse genetics has also been used to isolate and characterize clones of at least one peptidase gene, *pepM* (Movva et al., 1990). Reverse genetics generally entails purifying an enzyme, making a probe to detect the gene which encodes it, obtaining and inactivating clones of that gene, and returning the inactivated gene to the chromosome. The probe is typically an antibody raised against the purified protein, or as in the case of *pepM*, an oligonucleotide based on the amino acid sequence of the purified protein. Although not extensively exploited in studies of *E. coli* and *Salmonella* peptidases and proteases, these techniques offer the promise of identifying the genes for any enzymes which can be purified sufficiently to obtain amino acid sequence.

**Rationale for studying *opdA***, the gene for oligopeptidase A. Oligopeptidase A was first identified by Vimr and Miller (1983) as one of two activities in *S. typhimurium* which could degrade AcAla4. (Dipeptidyl carboxypeptidase is the other (Vimr and Miller 1983)). Mutations were isolated which prevented a *dcp* strain from using AcAla4 as a nitrogen source and eliminated the AcAla4 hydrolyzing activity in a cell extract. The locus of these mutations was called *optA* (Vimr et al., 1983). (The mnemonic has recently been changed to *opdA* to prevent confusion with an unrelated *E.coli* locus of the same name). Plasmids carrying *opdA* were initially, and intentionally, isolated as plasmids which allowed a *dcp* *opdA* strain to use AcAla4 as a nitrogen source (T. Golde, unpublished results). Curiously, *opdA* plasmids were also isolated as plasmids which allowed a multiply peptidase deficient strain to use Met-Gly-Gly as a methionine source. *opdA* carrying plasmids were also unexpectedly identified in a screen for plasmids which increased the hydrolysis of the chromogenic substrate Z-AAL-pNA.
Novak, Ray, and Dev (1986) have shown that oligopeptidase A is the major soluble activity able to degrade the lipoprotein signal peptide in vitro. Their initial studies and those of Vímr, Green, and Miller (1983) had suggested that oligopeptidase A was a carboxypeptidase. However, subsequent studies by Novak and Dev (1989) demonstrated that it was an endopeptidase with a penchant for substrates at least five amino acids long and containing either Ala or Gly on one side or the other of the scissile bond. These results made oligopeptidase A one of the relatively few endoproteases for which both the gene and a specific, potentially physiological, substrate were known.

prlC mutations are suppressors of a class of LamB signal sequence mutations (Emr et al., 1981; Thrun and Silhavy, 1987). LamB is an outer membrane protein required for the transport of maltodextrins and is the receptor for bacteriophage λ. Consequently, strains of E. coli carrying mutations in the lamB signal sequence which prevent the proper localization of the protein, are unable to use maltodextrins as a carbon source and are resistant to phage λ. In an effort to identify genes whose products acted in the export of LamB, mutations were isolated by Emr et al. (1981) which allowed strains carrying the signal sequence deletion mutation, lamBs60, to use maltodextrins. Most of these suppressors of lamBs60 were at the prlA locus, however one prlC1 was not. Subsequently, Thrun and Silhavy (1987) isolated more prlC alleles, and characterized them in detail. The strongest prlC suppressors appear to restore normal processing to the mutant precursor LamB protein, however very little LamB is exported to the outer membrane (Thrun and Silhavy, 1989; Thrun, dissertation, 1988). Thrun and Silhavy (1987) mapped prlC to 71 map units on the E. coli chromosome. When my studies were begun, opdA had been approximately mapped to the same region of the Salmonella chromosome (Vímr and Miller, 1983). The similar map positions and the in vitro signal
sequence hydrolyzing activity of OpdA (Novak and Dev, 1989), suggested a possible relationship between prlC and opdA.

While characterizing opdA mutants, Vimr noticed that phage P22 made very small plaques on the opdA\(^{-}\) host (Vimr, dissertation 1981). He also showed that this small plaque phenotype correlated with a decreased burst of infectious phage particles from an opdA\(^{-}\) host. This observation was interesting for two reasons. First, it was an example of a specific physiological effect resulting from a single peptidase or protease mutation. Second, there were no known proteolytic processing steps in phage P22 development (Botstein et al., 1973; Casjens and King, 1974).

In light of this intriguing constellation of results and observations, and in hopes of clarifying the physiological role of this enzyme, further characterization of opdA was undertaken. The cloning, mapping and nucleotide sequence of opdA are presented in chapter 2. Studies which characterize the relationship between opdA and prlC are discussed in chapter 3. Finally, studies which elucidate the role played by oligopeptidase A in bacteriophage P22 development are presented in chapter 4.
Chapter 2

Cloning, Mapping, and Nucleotide Sequence of \textit{opdA}, the Gene for Oligopeptidase A

from \textit{Salmonella typhimurium}
INTRODUCTION

Oligopeptidase A (OpdA) was originally identified by Vimr and Miller (1983) as one of two activities in extracts of *Salmonella typhimurium* which hydrolyzed N-acetyl-L-Ala₄ (AcAla₄). Mutations in the gene encoding oligopeptidase A (opdA formerly optA) were isolated from a dipeptidyl carboxypeptidase deficient (dcp) strain by screening for mutants unable to use AcAla₄ as a nitrogen source (Vimr *et al*., 1983). Dipeptidyl carboxypeptidase will also hydrolyze AcAla₄ and allow it to be used as a sole nitrogen source (Vimr and Miller, 1983). These opdA mutations were shown to map between 75 and 78 minutes on the *S. typhimurium* chromosome and to have only a small effect on bulk protein degradation and no effect on cell growth (Vimr *et al*., 1983). Surprisingly, strains with opdA mutations were unable to support normal development of phage P22. P22 formed tiny plaques on lawns of opdA strains. This phenomenon was correlated with a 10-fold reduction in burst size in these hosts (Vimr, dissertation, 1981).

Recent work by Novak and Dev (1989) and Novak *et al*. (1986) showed that OpdA is an endoprotease and that it is the major soluble activity in *Escherichia coli* able to hydrolyze the lipoprotein signal peptide *in vitro*. They suggested that OpdA may play a specific role in the degradation of signal peptides after they are released from precursor forms of secreted proteins.

Because opdA codes for an endoprotease which appears to play specific roles in both signal peptide degradation and P22 development, I have begun to characterize the locus and its product in more detail. This chapter presents the cloning and nucleotide sequence of the opdA gene, a more precise localization of the gene's position on the
S. typhimurium chromosome, and the purification and N-terminal amino acid sequence of OpdA.
MATERIALS AND METHODS

Bacteria, phage, plasmids, and media. Strains of *S. typhimurium* were derived from LT-2 and are listed in Table 2.1. *E. coli* K-12 strain JM109 (Yanisch-Perron et al., 1985) was used to propagate plasmids for purification of plasmid DNA. *E. coli* K-12 CSR603, obtained from Robert Hogg, was used to produce maxicells. Phage P22(HT12/4 int-3) and P1(clr100) were used as generalized transducing phage. P22(H5), a clear plaque mutant, was used to characterize plaque size. All plasmids were derivatives of pBR328 or pBR322 and are listed in Table 2.1. Bacteria were routinely grown in LB (Maniatis et al., 1982). E medium (Vogel and Bonner, 1956) was used as the minimal medium and supplemented with appropriate amino acids at 0.4mM and glucose at 4g/liter. The medium described by Gutnick et al. (1969), here referred to as NN, was used to test for use of peptides as sole nitrogen sources as described (Vimr and Miller, 1983). Antibiotics were added to rich media at the following concentrations in micrograms per ml: tetracycline, 25; ampicillin, 50 for solid media, 100 for liquid; kanamycin, 50; chloramphenicol, 20.

Development of benzylxycarbonyl Ala Ala Leu p-nitroanilide assay.

Benzylxycarbonyl Ala Ala Leu p-nitroanilide (Z-AAL-pNA) is a chromogenic subtilisin substrate used by Strongin et al. (1979) to assay the *E. coli* protease, ISP-L-Eco. Subtilisin directly releases p-nitroaniline from this substrate by cleavage between Leu and p-nitroaniline. Whereas Z-AAL-pNA is colorless, free p-nitroaniline is yellow, with a maximum absorbance at 410nm and an $\epsilon_{410}$ of 8900 (Strongin et al. 1979). To determine if there were activities in *Salmonella* which could hydrolyze this substrate, Z-AAL-pNA was incubated with crude cellular extracts (final concentration, 1.5mg protein/ml)
<table>
<thead>
<tr>
<th>Strain or plasmid</th>
<th>Genotype or description</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>S. typhimurium</strong></td>
<td></td>
</tr>
<tr>
<td>TN1201</td>
<td>dcp-1 zcf-845::Tn10 zhg-848::Tn5 opdA1</td>
</tr>
<tr>
<td>TN2490</td>
<td>TN1201/pJG70</td>
</tr>
<tr>
<td>TN3201</td>
<td>TN1201/pJG13</td>
</tr>
<tr>
<td>TN3202</td>
<td>TN1201/pJG73</td>
</tr>
<tr>
<td>TN3203</td>
<td>TN1201/pCM127</td>
</tr>
<tr>
<td>TN3204</td>
<td>TN1201/pCM128</td>
</tr>
<tr>
<td>TN1246</td>
<td><em>leuBCD485 pepN90 pepA16 pepB11 supQ302(DproAB pepD) pepP1 pepQ1</em></td>
</tr>
<tr>
<td>TN3058</td>
<td>TN1246/pBR328</td>
</tr>
<tr>
<td>TN3059</td>
<td>TN1246/pCM127</td>
</tr>
<tr>
<td>TN3080</td>
<td>TN1246/pCM128</td>
</tr>
<tr>
<td>TN3097</td>
<td>TN1246/pCM144</td>
</tr>
<tr>
<td>TN1379</td>
<td><em>leuBCD485</em></td>
</tr>
<tr>
<td>TN1727</td>
<td><em>leuBCD485 pepN90 pepA16 pepB11 supQ302(DproAB pepD) pepP1 pepQ1 pepT7 zhg-848::Tn5 opdA1 zxx-845::Tn10 dcp-1</em></td>
</tr>
<tr>
<td>TN3031</td>
<td>TN1727/pJG13</td>
</tr>
<tr>
<td>TN3032</td>
<td>TN1727/pJG70</td>
</tr>
<tr>
<td>TN3103</td>
<td>TN1727/pCM127</td>
</tr>
<tr>
<td>TN3104</td>
<td>TN1727/pCM128</td>
</tr>
<tr>
<td>TN3205</td>
<td>TN1727/pJG73</td>
</tr>
<tr>
<td>TN2183</td>
<td><em>leuBCD485 metA15 pepN90 pepA16 pepB11 pepP1 pepQ1 pepT7::Mu d1X</em></td>
</tr>
<tr>
<td>TN2625</td>
<td><em>leuBCD485 metA15 pepN90 pepA16 pepB11 pepP1 pepQ1 pepT7::Mu d1X opdA1 zxx-848::Tn5</em></td>
</tr>
<tr>
<td>TN2674</td>
<td>TN2625/pJG13</td>
</tr>
<tr>
<td>Strain</td>
<td>Description</td>
</tr>
<tr>
<td>--------</td>
<td>----------------------------------</td>
</tr>
<tr>
<td>TN2677</td>
<td>TN2625/pJG70</td>
</tr>
<tr>
<td>TN3206</td>
<td>TN2625/pCM127</td>
</tr>
<tr>
<td>TN3207</td>
<td>TN2625/pCM128</td>
</tr>
<tr>
<td>TN2373</td>
<td><em>ara-9 polA2</em></td>
</tr>
<tr>
<td>TR6612</td>
<td></td>
</tr>
<tr>
<td>TN3099</td>
<td>TN2373/pCM144</td>
</tr>
<tr>
<td>TN2658</td>
<td>*purC/F 42 finP301 (Lac^+)</td>
</tr>
<tr>
<td>TN3244</td>
<td>TN2658/pCM127γδ101</td>
</tr>
<tr>
<td>TN3247</td>
<td>TN2658/pCM127γδ104</td>
</tr>
<tr>
<td>TN3253</td>
<td>TN2658/pCM127γδ110</td>
</tr>
<tr>
<td>TN3254</td>
<td>TN2658/pCM127γδ111</td>
</tr>
<tr>
<td>TN3257</td>
<td>TN2658/pCM127γδ127</td>
</tr>
<tr>
<td>TN3258</td>
<td>TN2658/pCM127γδ130</td>
</tr>
<tr>
<td>TN2529</td>
<td><em>leuBCD485 metA15 pepN90 pepA16 pepB11 pepP1 pepQ1 pepT7::Mu d1X pepD3</em></td>
</tr>
<tr>
<td>TN2773</td>
<td><em>leuBCD485 metA15 pepA16 pepB11 pepP1 pepQ1 pepT7::Mu d1X pepD3 (PepN^+)</em></td>
</tr>
<tr>
<td>TN3101</td>
<td><em>leuBCD485 opdA10::Mu dJ</em></td>
</tr>
<tr>
<td>TN3192</td>
<td>(Fels-2) leuA414 r^ m^ opdA10::Mu dP</td>
</tr>
<tr>
<td>TN3193</td>
<td>(Fels-2) leuA414 r^ m^ opdA10::Mu dQ</td>
</tr>
<tr>
<td>TN3220</td>
<td>(Fels-2) leuA414 r^ m^ opdA10::Mu dP zhg-3073::Tn10 dTc</td>
</tr>
<tr>
<td>TN3221</td>
<td>(Fels-2) leuA414 r^ m^ opdA10::Mu dQ zhg-3073::Tn10 dTc</td>
</tr>
<tr>
<td>TN3227</td>
<td><em>leuBCD485 zhg-1635::Tn10 dCm</em></td>
</tr>
<tr>
<td>TN3262</td>
<td><em>asd-1</em></td>
</tr>
<tr>
<td>TN3263</td>
<td><em>glpD105</em></td>
</tr>
<tr>
<td>TN2540</td>
<td><em>metE551 metA22 ilv-452 trpB2 hisC527 galE496 xyl-404 rpsL120 flaA66 hsdR6 hsdSA29 H1-b nml H2-e,n,x</em> (Fels-2)^</td>
</tr>
<tr>
<td>DB4926</td>
<td></td>
</tr>
<tr>
<td>TN3737</td>
<td>as TN2540 with <em>opdA10::Mu dJ</em></td>
</tr>
<tr>
<td>AK3073</td>
<td>as TN2540 with zhg-3073::Tn10 dTc</td>
</tr>
<tr>
<td>AK3082</td>
<td>as TN2540 with zhg-3082::Tn10 dTc</td>
</tr>
</tbody>
</table>
AK3108  as TN2540 with zhg-3108::Tn10 dTc
AK3216  as TN2540 with zhg-3216::Tn10 dTc
AK3223  as TN2540 with zhg-3223::Tn10 dTc
xylB26
his-644

Plasmids

pJG13  7kb Sau3A fragment of TN1246 containing opdA in BamHI site of pBR328
pJG70  6.9kb Sau3A fragment of TN1246 containing opdA in BamHI site of pBR328
JG73   7.3kb Sau3A fragment of TN2294 containing opdA in BamHI site of pBR328
pCM127 8.3kb Sau3A fragment of TN1246 containing opdA in BamHI site of pBR328
pCM128 7.1kb Sau3A fragment of TN1246 containing opdA in BamHI site of pBR328
pCM144 as pCM127 but containing the opdA10::Mu dJ allele
overnight at 37°C in the following reaction mix: 42mM Tris-HCl pH8.6, 0.8mM CaCl₂, 56μg/ml Z-AAL-pNA in dimethyl formamide. It was found that extracts of TN1379 (pep⁺) released free p-nitroaniline, whereas the peptidase deficient strains, TN1246 and TN2529 did not. Since all of the peptidases missing in TN1246 and TN2529 require a substrate with a free amino terminus, this suggests that two activities, an endoprotease and an aminopeptidase, are necessary for the release of free p-nitroaniline. Since only peptidase N hydrolyses p-nitroanilides, Z-AAL-pNA was incubated with an extract of a pepN⁺ strain, TN2773, and free p-nitroaniline was released. This demonstrated that the addition of peptidase N was sufficient to complete the hydrolysis. It was noticed however, that although no free p-nitroaniline was released when Z-AAL-pNA was incubated with a peptidase deficient extract, such incubation converted Z-AAL-pNA to a material that could not be hydrolyzed by subtilisin. These results suggested that there was an activity in these extracts which could attack Z-AAL-pNA, but did not attack the Leu p-nitroaniline bond. To identify the products of this reaction, Z-AAL-pNA was incubated overnight as above with an extract made on TN2529 (pep⁻) and the products analyzed by HPLC. The product of this reaction coeluted with Leu p-nitroanilide suggesting that the protease responsible cleaved the bond between Ala and Leu, and that an aminopeptidase which could attack Leu p-nitroanilide, such as peptidase N, was required for the subsequent release of free p-nitroaniline. Based on these observations the assay and screen described below were developed.

Construction and screening of S. typhimurium plasmid libraries. Plasmid libraries containing 8-12 kb fragments of S. typhimurium chromosomal DNA in pBR328 were made by P. Hmiel and J. Glasbrook as previously described (Hmiel et al., 1986). Plasmids carrying opdA were obtained from the following three different screens.
1) Complementation of the AcAla₄ utilization defect. Plasmids were selected which allowed strain TN1201 (opdAl dcp-l) to use AcAla₄ (1.5 mM) as a nitrogen source on NN glucose agar. Increased hydrolysis of AcAla₄ was confirmed by HPLC analysis of products produced when cell extracts were incubated with AcAla₄ for 30 min at 37°C (Miller et al., 1987; Vimr, dissertation, 1981). 2) Utilization of Met-Gly-Gly as a methionine source. Plasmids were selected which allowed the multiply peptidase deficient strain TN2183 (pepN pepA pepB pepD pepP pepQ pepT dcp opdA) to use Met-Gly-Gly as a methionine source (Miller et al., 1987). 3) Hydrolysis of Z-AAL-pNA. A library of plasmids containing DNA from strain TN1246 (pepN pepA pepB pepD pepP pepQ) was screened for plasmids which conferred elevated levels of Z-AAL-pNA hydrolysis on strain TN1246. Oligopeptidase A (OpdA) cleaves Z-AAL-pNA between Ala and Leu and does not release the chromophore, p-nitroaniline. Since the product produced by OpdA cleavage contains a free amino group, it can be hydrolyzed by the aminopeptidase, peptidase N, to yield free chromophore. This coupled assay was used to detect clones that overproduced OpdA by screening microcultures grown in the wells of plastic depression plates basically as described by Carter and Miller (1984). Plasmid containing strains were transferred from agar plates to wells of plastic depression plates each containing 0.2 ml LB chloramphenicol. Cells were grown overnight at 37°C and replica cultures for storage were inoculated onto LB chloramphenicol plates. Cells in the depression plate were pelleted by centrifugation and washed with 0.2 ml 50 mM Tris-HCl (pH 7.6). The pellet was resuspended in 0.03 ml lysis solution (lysozyme, 1 mg per ml in 50 mM Tris-HCl, pH 7.6) and subjected to three cycles of freezing (-70°C) and thawing (37°C). To each well was added 0.15 ml of a solution containing 50 mM Tris-HCl (pH 7.6) and 0.5 mM cobalt chloride. Then 0.017 ml of Z-AAL-pNA (0.5 mg per ml in
dimethyl formamide) was added and the plates were incubated at 37°C for 1 hour. To each well was added 0.005ml partially purified *S. typhimurium* aminopeptidase N (4.5 mg protein/ml, 20 AU/min/mg). Aminopeptidase N activity was determined using Ala β-naphthylamide hydrolysis assayed spectrophotometrically as described by Lee *et al.* (1971). Plates were incubated at 37°C for 10 minutes. To enhance detection of free *p*-nitroaniline the reaction product was diazotized and coupled to N-(1-naphthyl) ethylenediamine to produce an intense purple color (Ohlsson *et al.*, 1986). Positive isolates were purified from the master plate. To confirm increased Z-AAL-pNA hydrolyzing activity extracts of these isolates were assayed spectrophotometrically. Cell extracts were added to a reaction mixture containing 0.07mM Z-AAL-pNA, 44mM N-methyl-diethanolamine-HCl (pH 8.0), 0.44mM CoCl₂, and aminopeptidase N (0.8AU/min/ml). The reactions were incubated at room temperature and release of *p*-nitroaniline was followed at 410nm. Extracts were diluted so that the assay was linear with added protein.

**Temperature shift induction of opdA10::Mu dJ.** An overnight culture of TN3101 grown in LB medium at 30°C was diluted 1:100 into fresh LB medium or E medium with 0.1% casamino acids, and grown at 30°C until it reached an OD₆₀₀ of 0.3. Then half the culture was transferred to a 42°C water bath and half was maintained at 30°C. Samples were taken every 15 minutes for 90 minutes and assayed for β-galactosidase activity as described by Miller (1972).

**Recombinant DNA techniques.** Restriction enzymes were obtained from New England Biolabs and BRL and used according to manufacturer’s instructions. T4 DNA ligase was from BRL. Fragments for subcloning were isolated by gel electrophoresis in low melting temperature agarose (GTG Sea Plaque agarose, FMC Bioproducts) and
ligated in agarose according to the manufacturer’s protocol. Other DNA manipulations were by standard techniques (Maniatis et al., 1982).

**Genetic techniques.** Transductions using phage P22 (Roth, 1970) or phage P1 (Heiman and Miller, 1978) were carried out as described. Mu dJ (Mud1734 (Kanr) lac operon fusion (Castillo et al., 1984)) insertions into plasmid pCM127 were obtained by the procedure of Hmiel et al. (1989). Tn1000 (γδ) insertions in pCM127 were produced by the procedure of Guyer (1983). Insertion opdA10::Mu dJ(Kanr), originally present on pCM144, was returned to the chromosome by transducing TN1379 to kanamycin resistance on MacConkey lactose agar with a P22 lysate grown on TN3097 (which contains pCM144). When carried on the plasmid the opdA10::Mu dJ insertion produced enough β-galactosidase to make the colony red on MacConkey lactose agar. A strain carrying opdA10::Mu dJ in single copy in the chromosome however, is white on MacConkey agar. To detect chromosomal replacement white kanamycinr transductants on MacConkey kanamycin agar were picked and checked for loss of plasmid encoded antibiotic resistance markers (ampicillin and chloramphenicol resistance). Replacement of the chromosomal opdA gene by the opdA::Mu dJ insertion was confirmed by Southern hybridization (Southern, 1975). Integration of entire plasmids into the chromosome of a polA strain (Whitfield and Levine, 1973) was carried out as described previously (Hmiel et al., 1989). The map position of the integration was determined by P22 transduction using linked transposon zhg-860::Tn5 (Vimr et al., 1984). Tn10 dTc (Tn10 dTcΔ16Δ17, tetracycline-resistant transposition defective Tn10 (Way et al., 1984)) insertions near opdA10::Mu dJ were identified in the set of insertions constructed by Kukral et al. (1987) by screening for Tn10 dTc insertions which simultaneously conferred tetracycline resistance and caused loss of β-galactosidase activity on LB tetracycline X-gal agar plates.
Tn10 dTc insertions near asd were identified in the same collection by screening for insertions which simultaneously conferred tetracycline resistance and ability to grow without added diaminopimelic acid on strain TN3262 (asd'). To isolate a Tn10 dCm (chloramphenicol', transposition defective Tn10 (Elliot and Roth, 1988)) insertion near opdA a population of random Tn10 dCm insertion derivatives of TN1201 (Tn5 near opdAI) was prepared (Elliot and Roth, 1988). A P22 transducing lysate of this population was used to transduce TN1379 to both chloramphenicol and kanamycin resistance simultaneously. Mu dP22 mapping was performed as described by Youderian et al. (Youderian et al., 1989). Maxicell analysis of plasmid encoded proteins was performed essentially as described by Sancar et al. (1979) with the modifications of Snavely et al. (1989).

DNA sequencing. All nucleotide sequence was obtained using the dideoxy method (Sanger et al., 1977), alkaline denatured double stranded templates (Chen and Seeburg, 1985), and Sequenase (United States Biochemicals) according to the manufacturer's instructions. Primers for sequencing from the ends of the Tn1000 insertions were those described by Liu et al. (1987), primers for pBK322 sequence were from New England Biolabs, and other oligonucleotide primers were synthesized on an Applied Biosystems nucleic acid synthesizer. Primers used are listed in Table 2.2. Sequence comparisons were performed using the University of Wisconsin Genetics Computing Group package of programs (Devereux et al., 1984).

Purification of OpdA and N-terminal amino acid sequencing. S. typhimurium OpdA was purified from a strain TN3080 containing plasmid pCM128. OpdA from E. coli has been previously purified from a single chromosomal copy using a different procedure (Novak et al., 1986). Ten liters of cells grown overnight at 37°C in LB
Table 2.2. Primers Used to Sequence *opdA*

<table>
<thead>
<tr>
<th>Primer</th>
<th>Sequence</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>NEB 1219</td>
<td>ATGCGTGCCGCGTGA</td>
<td><em>BamHI</em> site of pBR322, counterclockwise</td>
</tr>
<tr>
<td>NEB 1223</td>
<td>CACTATCGACTACGCCATCA</td>
<td><em>BamHI</em> site of pBR322, clockwise</td>
</tr>
<tr>
<td>CC1</td>
<td>TCAATAAGTTATACCAT</td>
<td>Tn<em>1000</em>, γ end</td>
</tr>
<tr>
<td>CC2</td>
<td>GAATTATCTCTTTAACG</td>
<td>Tn<em>1000</em>, δ end</td>
</tr>
<tr>
<td>OPT-1</td>
<td>TAAAACGTAAGCCCGGAT</td>
<td>439-456 in <em>opdA</em> sequence</td>
</tr>
<tr>
<td>OPT-2</td>
<td>ACGTTGATGTCTGGCAC</td>
<td>1750-1766 in <em>opdA</em> sequence</td>
</tr>
<tr>
<td>OPT-3</td>
<td>TTCGTCATACTGGTTCAA</td>
<td>complement of 1799-1783 in <em>opdA</em> sequence</td>
</tr>
<tr>
<td>OPT-4</td>
<td>CCGGTCTCACGGTTGAA</td>
<td>complement of 2518-2501 in <em>opdA</em> sequence</td>
</tr>
<tr>
<td>OPT-5</td>
<td>TTCAACCGTGAGACCCGG</td>
<td>2502-2518 in <em>opdA</em> sequence</td>
</tr>
<tr>
<td>OPT-7</td>
<td>GTTGCGCAGCTCCAGATG</td>
<td>complement of 2789-2772 in <em>opdA</em> sequence</td>
</tr>
</tbody>
</table>
chloramphenicol broth were harvested by centrifugation, lysed by sonication, and the extract cleared by centrifugation at 200,000 x g for 2 hours. The cleared extract was applied to a DES2 anion exchange column (Whatman) in 10mM Tris-HCl (pH 7.9), 5mM MgCl₂, 0.1M NaCl, and eluted with a NaCl gradient (0.1-0.35M). Active fractions detected using the Z-AAL-pNA micro-assay were pooled and concentrated by ultrafiltration using an Amicon YM-30 membrane. Concentrated fractions were applied to an Ultragel AcA34 gel filtration column. Active fractions were again pooled and concentrated with an Amicon Centriprep-10. A chloramphenicol caproate-agarose affinity column (Sigma) was used to remove a major contaminant, chloramphenicol acetyltransferase encoded by the vector (pBR328) (Zaidenzaig and Shaw, 1976). OpdA was adsorbed in 10mM Tris-HCl (pH 7.6) and eluted with 0.2M NaCl. The peak fractions were concentrated, applied to a FPLC MonoQ column (Pharmacia), and eluted with a NaCl gradient. Peak fractions were again concentrated, reapplied to the MonoQ column, and eluted with a NaCl gradient. The most active fractions were applied to a FPLC Sepharose 12 column (Pharmacia). The purity of the final product as estimated by Coomasie Blue staining of SDS polyacrylamide gel electrophoresis (Laemmli, 1970) was greater than 95%. Final purification was 3000 fold compared to levels from a single chromosomal copy. Amino acid sequence of the N-terminal region was determined using an Applied Biosystems Model 470A protein sequencer.
RESULTS

Isolation of plasmids carrying opdA gene. The gene for opdA was cloned using three different selections. It was cloned directly from a pBR328 library of Sau3a partial digest fragments from a multiply peptidase deficient strain, TN1246, by screening for clones that allowed a Δcp' opdA' strain to use AcAla₄ as a sole nitrogen source. Plasmid pJG70 was identified in this screen. It was integrated into the chromosome of a polA mutant, and the linkage of the plasmid-encoded chloramphenicol resistance to a transposon insertion (zgh-860::Tn5) near opdA was tested by P22 transduction. The cotransduction frequency of pJG70 and zgh-860::Tn5 was 51%, suggesting that pJG70 carried insert DNA homologous to the opdA region. As part of another project, similar libraries were screened for plasmids which allowed a multiply peptidase deficient strain (TN2183) to use Met-Gly-Gly as a methionine source. Plasmids pJG13 and pJG73 were identified in this screen. When integrated into the chromosome, plasmids pJG13 and pJG73 were 47% and 44% cotransducible with zgh-860::Tn5 respectively, suggesting that these plasmids also contained insert DNA with homology to the opdA region. In still another line of investigation, a similar library was screened for plasmids that led to overproduction of an activity that hydrolyzed the chromogenic endoprotease substrate Z-AAL-pNA. The assay used in this screen detected activities that cleaved Z-AAL-pNA between Ala and Leu. Plasmids pCM127 and pCM128 obtained from this screen also integrated into the chromosome by homology with the opdA region. The integrated plasmids showed 45% (pCM127) and 38% (pCM128) cotransduction with zgh-860::Tn5.

Although isolated using different screens, all these plasmids (pJG13, pJG70, pJG73, pCM127, pCM128) contained inserts from the same region of the chromosome and conferred the same set of phenotypes: AcAla₄ utilization, Met-Gly-Gly utilization, and
elevated Z-AAL-pNA hydrolysis. Assays of AcAla₄ hydrolyzing activity in extracts of strains carrying these plasmids showed levels approximately 100-fold higher than wild type. In addition, all of these plasmids conferred the ability to support normal growth of phage P22 on an opdA⁺ strain. Phage P22 forms small plaques on an opdA⁺ host (50), on plasmid carrying opdA⁺ hosts it formed large plaques. Restriction mapping showed that these plasmids shared a common 4kb EcoRI-KpnI fragment (Fig. 2.1). Additionally, a Mu dJ insertion in pCM127 (see below) simultaneously eliminated its ability to complement the AcAla₄ utilization defect, the Z-AAL-pNA hydrolysis defect and the phage P22 small plaque phenotype. These results argue that these plasmids carry the opdA gene and that this gene is responsible for the phenotypes conferred by these plasmids.

Analysis of insertions in the cloned opdA gene. In order to define the location of the opdA gene in the cloned DNA, insertions of both the Mu dJ mini-MuΔ element and Tn1000 (γδ) were generated in pCM127. Strains containing plasmids carrying these insertions were screened for loss of their ability to rapidly hydrolyze the chromogenic substrate Z-AAL-pNA. Plasmids containing such insertions were restriction mapped. All insertions that led to loss of Z-AAL-pNA hydrolysis were localized to a 2kb region of pCM127, a region shared by all of the plasmids (Fig. 2.1).

Maxicell analysis of pCM127 and pCM128 showed that both plasmids encoded a 66kD protein (Fig. 2.3). In addition, this 66kD protein could be seen as a major Coomassie blue stained protein after SDS-PAGE of crude soluble protein extracts of strains carrying each of the opdA plasmids and was not present in an extract of a strain containing only the pBR328 vector (Fig. 2.2).

Maxicell analysis showed that the insertion opdA10::Mu dJ in plasmid pCM127
FIGURE 2.1. Restriction endonuclease maps of opdA\textsuperscript{+} plasmids and sites of insertions in pCM127. Horizontal lines represent chromosomal DNA carried on the plasmid. Restriction endonuclease sites are indicated as follows: E, \textit{EcoRI}; B, \textit{BglII}; P, \textit{PvuII}; K, \textit{KpnI}; S, \textit{SalI}. The enlarged region of pCM127 shows the locations and allele numbers of insertions in or near opdA. Open symbols indicate insertions which reduced Z-AAL-pNA hydrolyzing activity. Insertion 101, represented by a closed circle, did not. \textit{Tn1000} insertions are represented by circles, the \textit{Mu dJ} insertion is represented by a square. Plasmid pJG70 was originally isolated by T. Golde, pJG73 by P. Martello, and pJG13 by K. Strauch.
FIGURE 2.2. SDS-polyacrylamide gel electrophoresis of extracts of strains containing opdA⁺ plasmids. Plasmid carrying derivatives of TN1727 were grown to late log phase in LB medium with chloramphenicol. Cells were pelleted by centrifugation, then broken by sonication. The sonicate was cleared by a 90 min. centrifugation in SS34 rotor at 15,000 rpm. Proteins from equivalent amounts of each extract were separated on 10% polyacrylamide SDS slab gels with Tris-glycine buffer (Laemli, 1970). Proteins were visualized by staining with Coomassie Blue R250. Lanes are as follows: S, molecular weight standards, values are given in kD; 1, TN1727 without plasmid; 2, TN1727/pJG13; 3, TN1727/pJG70; 4, TN1727/pCM127; 5, TN1727/pCM128; 6, TN1727/pJG73; 7, TN1727/pBR328. Open arrow indicates the OpdA band present in the plasmid extracts. Filled arrow at 50kD indicates band present only in pCM127 carrying strain, and filled arrow at 28kD indicates band present in pCM127 and pJG70 carrying strains.
eliminated the 66kD band (Fig. 2.3B) as did insertions opdA104::Tn1000 and opdA130::Tn1000 (Fig. 2.3A). These insertions also eliminated Z-AAL-pNA hydrolyzing activity. Insertion opdA101::Tn1000 however, did not affect either Z-AAL-pNA hydrolyzing activity or the 66kD band. The presence of a new 23kD band in opdA104::Tn1000 and a new 50kD band in opdA130::Tn1000 indicated that opdA is transcribed in the direction from opdA104::Tn1000 to opdA130::Tn1000. This is the same direction as lac transcription in opdA10::Mu dJ. This was later confirmed by sequencing from the end of opdA10::Mu dJ.

Mapping opdA on the Salmonella chromosome. The results of conjugation crosses had suggested that opdA is located between asd and xylB (Vimr et al., 1983). To confirm this independently and to map the chromosomal location of opdA more accurately, I first returned insertion opdA10::Mu dJ to the chromosome by P22 transduction. I confirmed the replacement of the chromosomal locus by testing for the opdA+ phenotypes and by Southern hybridization (Fig. 2.4). The chromosomal opdA10::Mu dJ was exchanged with the Mu dP22 elements (Mu dP and Mu dQ) of Youderian et al. (1989). These elements form locked-in P22 prophage which when induced cannot excise, but still make phage particles and preferentially package chromosomal DNA from the pac site. Mu dP packages in the same direction as lacZ is transcribed in Mu dJ, Mu dQ packages in the opposite direction. Mu dP and Mu dQ transducing lysates were used in crosses with strains carrying asd and xylB markers. The numbers of Asd+ and Xyl+ transductants were determined and normalized to the number of His+ transductants (Table 2.3). Because Mu dP transduced asd much more frequently than xylB and Mu dQ transduced xylB more frequently than asd, opdA must lie between asd and xylB. Additionally, because Mu dP packages in the same direction as lac is
FIGURE 2.3. Gene products expressed by opdA plasmids and opdA plasmids containing insertions. Gene products of plasmids in E. coli strain CSR603, were expressed in maxicells labelled with $[^{35}\text{S}]$methionine. Proteins were separated on SDS-10% polyacrylamide gels and visualized by autoradiography. Open arrowheads indicate opdA gene product. Filled arrowheads indicate fragments generated by Tn1000 insertions in pCM127. Panel A: lane 1, pBR322; lane 2, pCM127; lane 3, pCM127 opdA10::Mu dJ. Panel B: lane 1, pCM127; lane 2, pCM128; lane 3, pCM127 opdA101::Tn1000, lane 4, pCM127 opdA104::Tn1000; lane 5, pCM127 opdA130::Tn1000, lane 6, pBR328. Numbers between the panels indicate molecular weight standards in kD. Note that proteins of ~50kD and 26kD are produced by pCM127 but not by pCM128. These proteins must be coded by the region of pCM127 to the left of opdA (Fig. 2.1) that is not carried by pCM128.
FIGURE 2.4. Southern hybridization to show replacement of chromosomal \textit{opdA} by \textit{opdAI0::Mu dJ}. Chromosomal DNA from TN1379 and TN3101 (\textit{opdAI0::Mu dJ}) was digested with \textit{EcoRI} and separated on 0.6\% agarose gel. Fragments were transferred to charged nylon membrane using a vacuum transfer apparatus. The 700bp \textit{PvuII} fragment of pCM127, which spanned the \textit{Mu dJ} insertion site, was radiolabelled with $[^{32}\text{P}]$-dCTP (Feinberg and Vogelstein, 1983; 1984), and hybridized to the membrane overnight at $50^\circ$C. After washing, the membrane was exposed to X-ray film.
**TABLE 2.3. Relative efficiency of transduction of \( \text{asd}^+ \) and \( \text{xyI}B^+ \) by induced \( \text{opdA10}:\text{Mu dP22} \) lysates**

<table>
<thead>
<tr>
<th>Selected marker</th>
<th>Donor lysate(^a)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>TN3192(Mu dP)</td>
</tr>
<tr>
<td>( \text{asd}^+ )</td>
<td>146.2(^b)</td>
</tr>
<tr>
<td></td>
<td>(4824)(^c)</td>
</tr>
<tr>
<td>( \text{xyI}B^+ )</td>
<td>3.3</td>
</tr>
<tr>
<td></td>
<td>(110)</td>
</tr>
<tr>
<td>( \text{his}^+ )</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>(33)</td>
</tr>
</tbody>
</table>

\(^a\)TN3192 and TN3220 are independent replacements of \( \text{opdA10}:\text{Mu dJ} \) by Mu dP. TN3193 and TN3221 are independent replacements of \( \text{opdA10}:\text{Mu dJ} \) by Mu dQ.

\(^b\)Efficiency with which the selected marker was transduced by the Mu dP22 lysate relative to the efficiency with which \( \text{his}^+ \) was transduced.

\(^c\)Number of transductants per \( \mu l \) of lysate
transcribed in Mu dJ, which is also the same direction as opdA transcription, opdA must be transcribed toward asd.

Although opdA is between asd and xylB, it is unlinked by P22 transduction to either of them. Eight Tn10 dTc insertions linked to opdA by P22 transduction were isolated by screening a collection of random Tn10 dTc insertions (Kukral et al., 1987). A Tn10 dCm linked to opdA by P22 transduction was isolated from a random Tn10 dCm population made on strain TN1201. None of these elements was linked to asd or to xylB by P22 transduction. Three Tn10 dTc insertions P22 cotransducible with asd were isolated from a collection of random Tn10 dTc insertions. None of these was cotransducible with opdA using P22. However, two of these insertions, zhs-3223::Tn10 dTc and zhs-3216::Tn10 dTc, were cotransducible with opdA::Mu dJ using phage P1. P1 transducing particles carry approximately twice as much DNA as those formed by P22, and so are able to cotransduce markers too distant to be carried on a single P22 transducing fragment. These results clearly place opdA at 76 map units on the Salmonella chromosome (Fig. 2.5).

Purification and N-terminal sequence of OpdA. Salmonella typhimurium OpdA was purified to near homogeneity from a strain containing plasmid pCM128. Though the yield was low (1.4%), a 3000 fold purification compared to the level of activity in an extract of a strain containing a single chromosomal copy was achieved. The purest fraction obtained was subjected to N-terminal amino acid sequencing and the first 18 amino acids were determined to be: (Met) Thr Asn Pro Leu Leu Thr Ser Phe (?) Leu Pro Pro Phe Ser Ala Ile Lys. Sequencing indicated that two species were present which differed only by the presence of an N-terminal methionine. Apparently the level of overproduction exceeds the cell’s capacity for N-terminal methionine removal. Purified
FIGURE 2.5. Map position of opdA on the S. typhimurium chromosome.
Linkage data above the line was obtained with phage P1, below the line with phage P22 as indicated. Arrows point from the selected marker to the unselected marker. Distances above the arrows are in kb's calculated from co-transduction frequencies using the formula in ref. 45. Mapped genes are shown with their positions in map units. The opdA allele used in all crosses was opdA10::Mu dJ. Transposon insertions are indicated by their allele numbers. Insertion 1635 is a Tn10 dCm, all others are Tn10 dTc. The arrow above opdA indicates direction of transcription of the gene.
OpdA was stimulated by 0.01mM Co$^{2+}$ and to a lesser extent by 0.1mM Mn$^{2+}$. It was completely inhibited by the following: 0.1mM Zn$^{2+}$, 1mM Ni$^{2+}$, 0.1mM Cu$^{2+}$, and 1mM EDTA. It hydrolyzed AcAla$_4$ and Z-AAL-pNA, but not insulin A or B chain. Curiously, even though opdA containing plasmids allowed a peptidase deficient strain to use Met-Gly-Gly as a Met source, purified OpdA failed to hydrolyze this peptide. The rate of hydrolysis of Met-Gly-Gly was at least $10^5$-fold less than that of AcAla$_4$.

Nucleotide sequence of the opdA gene. The nucleotide sequence for the gene opdA was determined using dideoxy sequencing techniques on double stranded plasmid template DNA. Oligonucleotide primers were used to sequence from both ends of Tn1000 insertions and from the L end of Mu dJ. To sequence a difficult region, a 700bp PvuII fragment from pCM127 was cloned into the BamHI site of pBR328 and sequenced using primers homologous to the pBR328 sequence around the BamHI site. Additionally, five primers were synthesized which allowed the sequencing of regions which were inaccessible using the Tn1000 insertions. As shown in figure 2.6, both strands were sequenced throughout with the exception of 34 nucleotides on the 5' end and 89 on the 3' end. The complete sequence of opdA is shown in figure 2.7.

The sequence contains a 680 amino acid open reading frame (orf) starting at nucleotide 603 with an N-terminal sequence identical to that determined by directly sequencing OpdA and with codon usage typical of other S. typhimurium genes (Wada et al., 1991). This orf is preceded by a possible ribosome binding site. The orf predicts a 77kD protein with a pI of 5.0. The calculated pI is consistent with the behavior of OpdA on anion exchange columns: elution from a DE52 column at pH 7.9 and 0.2M NaCl, and from a Mono-Q column at pH 8.2 and 0.3M NaCl. Both the purified protein and the protein observed in maxicells and crude extracts display a mobility on SDS-PAGE.
FIGURE 2.6. opdA sequencing strategy. Arrows indicate the extent of sequence (5' to 3') derived from individual priming sites. Circles indicate Tn1000 insertions in pCM127 which were used as priming sites. Allele numbers are indicated above the circles. All insertions, except 101, eliminated opdA complementing activity. The square indicates the opdA10::Mu dJ insertion in pCM127, also used as a priming site.
FIGURE 2.7. Nucleotide sequence of opdA region and predicted amino acid sequence of oligopeptidase A. Doubly underlined DNA sequence indicates potential $\sigma^{32}$ promoter (Cowling et al., 1985). Singly underlined DNA sequence indicates putative ribosome binding site. Underlined amino acid sequence indicates metalloprotease Zn$^{2+}$ binding site motif (Jongeneel et al., 1989). Doubly underlined amino acid sequence indicates sequence determined directly from purified OpdA protein.
indicating a molecular weight of 66,000.

There is no strong consensus sigma\(^{70}\) -10 or -35 region upstream of the start of translation, which is consistent with the low level of expression assayed from the chromosomal Mu dJ insertion (37 units of \(\beta\)-galactosidase). However there is a near consensus sigma\(^{32}\) heat-shock promoter at nucleotide 504 to 540 (see Table 3.3, chapter 3) (Cowing \textit{et al.}, 1985). To determine if this is an active heat-shock promoter, \(\beta\)-galactosidase activity of a culture of TN3101 (\textit{opdA10::Mu dJ}) was assayed at 15 minute intervals after shifting the culture from 30°C to 42°C. No increase in \(\beta\)-galactosidase levels was observed after the temperature shift (37 units ±5). This procedure has been used to analyze the induction of other heat shock genes (Yano \textit{et al.}, 1987).

Examination of the sequence revealed a site [amino acids 467-475; (L)(F)HE(F)(G)\(\gamma\)H(G)(L)] with considerable similarity to a proposed metalloproteinase Zn\(^{2+}\) binding site motif: (uncharged)-(uncharged)-H-E-(uncharged)-(uncharged)-H-(uncharged)-(hydrophobic) (Jongeneel \textit{et al.}, 1989). This is consistent with the observed inhibition of OpdA by EDTA.

The predicted amino acid sequence of OpdA shows considerable similarity to the predicted amino acid sequence of dipeptidyl carboxypeptidase (Dcp) (Hamilton, M.S. thesis, 1991). Both orf's are 680 amino acids long and display 33% amino acid identity over their entire lengths. In addition, there are two regions of >50% identity. One region is centered around the putative Zn\(^{2+}\) binding site and extends from residue 440 to 512. The other is between amino acids 590 and 677, the C-terminal eighth of the proteins. These similar regions are graphically displayed in Figure 2.8.

A search of GenBank using the FASTA program (Pearson and Lipman, 1988)
revealed significant similarity of both OpdA and Dcp to rat metalloendopeptidase EC 3.4.24.15 (EP 24.15) (Pierotti et al., 1990). EP 24.15 is a zinc metalloendopeptidase found in brain, testes, and pituitary, and appears to play a role in peptide hormone processing and degradation (Lasdun et al., 1989; Orlowski et al., 1989). OpdA and EP 24.15 showed 28% amino acid identity in a 565 residue overlap, while Dcp and EP 24.15 displayed 24% amino acid identity in that region. The regions of similarity are graphically represented in figure 2.8. Figure 2.9 shows the multiple alignment of OpdA, Dcp and EP 24.15. The region around the putative zinc binding site is very similar in all three. Whereas orf's for OpdA and Dcp are both 680 amino acids long, EP 24.15 is only 645 and lacks the extended region of similarity shared by the C-terminal eighth of OpdA and Dcp. Except for the minimal Zn$^{2+}$ binding site motif (HEXXH), no similarity was found to the other E. coli zinc metallopeptidase (aminopeptidase N), or to any other metalloproteases in the data base.
FIGURE 2.8. Amino acid sequence comparisons of oligopeptidase A (OpdA), dipeptidyl carboxypeptidase (Dcp), and rat metalloendopeptidase EC 3.4.24.15 (EP 24.15). Comparisons were performed using the program Compare in the UWCG package. The comparison matrix was based on PAM250, the window size was 30, and the stringency was 15. Numbers indicate amino acid residue.
FIGURE 2.9. Alignment of deduced amino acid sequences of oligopeptidase A (OpdA), dipeptidyl carboxypeptidase (Dcp), and rat metalloendopeptidase EC 3.4.24.15 (EP 24.15). Alignment was performed with the Pileup program in the UWGCG package using a comparison matrix based on PAM250, a gap penalty of 4, and a gap extension penalty of 0.1. Amino acid residues common to OpdA and Dcp are shown above the Dcp sequence. Residues common among all three are shown below the EP 24.15 sequence. Conservative amino acid replacements, represented by an asterisk, are as follows: D or E; L or I or V; K or R; F or Y or W; S or T. Dcp sequence is from Hamilton (1991), EP 24.15 sequence is from Pierotti et al. (1990).
DISCUSSION

Clones carrying the \textit{opdA} gene have arisen from three different screens. The first, intended for the isolation of \textit{opdA}, involved complementing the AcAla\textsubscript{4} utilization defect of an \textit{opdA} mutant strain. The second, intended to identify new endoprotease activities, involved screening libraries for plasmids conferring elevated Z-AAL-pNA hydrolyzing activity. When this work was begun, studies with extracts of \textit{opdA}\textsuperscript{+} and \textit{opdA}\textsuperscript{−} strains and purified protein had suggested that oligopeptidase A (OpdA) was a carboxypeptidase not an endopeptidase (Novak \textit{et al.}, 1989; Virmr, dissertation, 1981). Thus, the chromogenic substrate Z-AAL-pNA, which is blocked at both the amino and carboxyl termini, was not expected to be an OpdA substrate. Subsequently, Novak and Dev (1989) showed that purified OpdA from \textit{E. coli} cleaves lipoprotein signal peptide \textit{in vitro} at several internal sites, usually on one side or the other of an Ala or Gly residue, demonstrating that OpdA is an endoprotease. It is therefore not surprising that OpdA cleaves Z-AAL-pNA between Ala and Leu. The third screen was intended to identify clones carrying \textit{pepM}, the gene encoding the activity responsible for N-terminal methionine removal (Miller \textit{et al.}, 1987). This screen also produced plasmids which were eventually shown to carry \textit{opdA}. Since purified OpdA does not hydrolyze Met-Gly-Gly to free methionine, it is not clear how plasmids carrying \textit{opdA} allow a peptidase deficient strain to use this peptide a source of methionine. One possibility is that overproduction of OpdA leads to the degradation of peptides which are inhibitors of the activity which does hydrolyze Met-Gly-Gly. Peptides are known to accumulate in peptidase deficient strains (Yen \textit{et al.}, 1980, 1980a). Another possibility is that the \textit{in vitro} conditions used to assay OpdA may not mimic those in which it acts in the cell. For example, it has been shown
that striking differences in the specificities of another bacterial peptidase can be generated by changing the metal ion present in the assay (Hayman et al., 1974).

Novak et al. (1986) have shown that OpdA is the major soluble activity able to cleave lipoprotein signal peptide in vitro. Elimination of opdA, however, has no effect on cell growth. Apparently there are multiple signal peptide degrading activities in vivo. Just as both *S. typhimurium* and *E. coli* have three aminopeptidases with overlapping broad specificities (Miller and MacKinnon, 1974; Miller and Schwartz, 1978), there are at least three activities which can degrade cleaved lipoprotein signal peptides in vitro. In addition to OpdA, both protease IV, a membrane bound signal peptide degrading activity encoded by *sppA* (Ichihara et al., 1986), and protease So, a soluble enzyme (Chung and Goldberg, 1983), have been shown to attack this signal peptide in vitro (Novak et al., 1986; Novak and Dev, 1989). The So activity is weak, however, and under the conditions used, OpdA comprises about 90% of the soluble lipoprotein signal peptide degrading activity. It is possible that in vivo the broad specificity aminopeptidases, peptidases N, A, and B, and dipeptidyl carboxypeptidase (Dcp), may be involved in signal peptide degradation as well. These enzymes have been shown to play a role in protein degradation (Miller and Green, 1981; Yen et al., 1980).

Inspection of the predicted amino acid sequence of *opdA* revealed a thermolysin-like metalloproteinase Zn$^{2+}$ binding site. This site has been found in a number of Zn$^{2+}$ metalloproteins from a variety of sources including; human collagenase, crayfish digestive protease, *Bacillus* spp. thermolysin and neutral protease, and *E. coli* peptidase N (Jongeneel et al., 1989). OpdA shows no extended amino acid similarity to any of these enzymes. Comparison of *opdA* sequence with the sequence for *dcp* showed considerable similarity between the two. Although both enzymes will hydrolyze AcAla$_4$, 
the sequence similarity is surprising as Dcp is a true carboxypeptidase which removes dipeptides from the C-termini of polypeptides (Yaron et al., 1972) and OpdA is an endoprotease. Such sequence similarity is not typical of E. coli and Salmonella peptidases. Although at least 14 E. coli or S. typhimurium peptidolytic enzymes have been sequenced, only opdA and dcp, and pepP and pepQ (Nakahigashi and Inokuchi, 1990) have been shown to be similar to each other. In addition to their similarity to each other, OpdA and Dcp share considerable similarity with the rat metallopeptidase EP 24.15. There are several other examples of a mammalian peptidase exhibiting sequence similarity to an E. coli or S. typhimurium peptidase. Peptidase Q (pepQ) shows 29% amino acid identity to human prolidase (Nakahigashi and Inokuchi, 1990). Aminopeptidase A (pepA) from E. coli shows 35% amino acid identity to bovine lens leucine aminopeptidase (Stirling et al., 1989). Protease III (prt) from E. coli shows 26% identity to human insulin degrading enzyme (AFFHOLTER et al., 1989), and E. coli aminopeptidase N (pepN) shows 21% amino acid identity in a 361 residue overlap with human aminopeptidase N (Olsen et al., 1988).

In addition to their sequence similarity, OpdA and EP 24.15 have similar specificities and possibly similar functions. EP 24.15 cleaves peptides on the carboxyl side of hydrophobic amino acids, including Gly and Ala (Orlowski et al., 1989). It has an extended substrate binding site able to contain at least 5 amino acids. Additionally, EP 24.15 will not cleave polypeptides longer than about 20 amino acids (Orlowski et al., 1988, 1989). Based on the available evidence, OpdA appears to prefer Gly or Ala on either side of the scissile bond, and it requires substrates at least 5 amino acids long with 2 to 3 amino acids on either side of the scissile bond (Novak and Dev, 1989; VIMR, dissertation, 1981). OpdA will only hydrolyze the lipoprotein signal peptide after it has been removed from the precursor protein, suggesting that OpdA, like EP 24.15, will not
attack larger polypeptides (Novak and Dev, 1989).

OpdA and EP 24.15 both appear to be involved in post-translational processing of polypeptides and degradation of bioactive peptides. EP 24.15 is found in peptide hormone producing tissues such as brain, pituitary, and testes (Orlowski et al. 1989). It hydrolyses several enkephalin containing peptides in vitro releasing the active peptide hormone (Orlowski et al. 1989). It is also the major activity from these tissues able to degrade gonadotropin releasing hormone (GnRH) in vitro (Orlowski et al. 1989).

Additionally, studies with a specific inhibitor of EP 24.15 have shown that EP 24.15 is active in the degradation of GnRH in vivo (Lasdun et al., 1989). As described in chapter 4, OpdA is required for the post-translational removal of 20 amino acids from the N-terminus of a phage P22 protein. As shown in chapter 3, opdA is the Salmonella homolog of the E. coli gene prIC, a site of suppressors of LamB signal sequence mutations. This result, along with the observation that OpdA degrades free lipoprotein signal peptide, suggests that OpdA may play a role in signal peptide processing or degradation in vivo.

The degree of sequence similarity among OpdA, Dcp, and EP 24.15 suggests they represent a previously unidentified sub-class of zinc metallopeptidases. The regions shared by all three could serve as a profile to identify other members of this sub-class. The similarity between OpdA and Dcp suggests that they may have arisen from a duplication of an earlier peptidase gene. More data would be required to determine if the branching which ultimately led to EP 24.15 occurred before or after this duplication.

However, based on the greater similarity of OpdA than Dcp to EP 24.15, and the similar specificity and perhaps similar function of OpdA and EP 24.15, one might speculate that the branching which led to EP 24.15 occurred after the duplication which led to Dcp.
The apparent conservation of both functional and structural elements between OpdA and EP 24.15, and protease III and human insulin-degrading enzyme (Affholter et al., 1988), suggests that functional similarity with their counterparts in higher organisms may be a common theme among Salmonella and E. coli peptidases. This may provide insights into the largely uncharacterized roles of these enzymes in vivo.

The opdA sequence contains a near consensus sigma\(^{32}\) promoter at position 504. The consensus \(\sigma^{32}\) promoter is: TNNtCNCCCTTGAA 13-15 bp spacer

\[
\text{CCCAT}tTaNNNNNtca \quad \text{(Cowin et al., 1985), opdA has: gtcCatCaTTGAA 15 bp spacer CCCCATTtctagtctca. Several } \sigma^{32} \text{ dependent heat shock proteases have been identified,}
\]

including Lon (Goff et al., 1984) and ClpP (Kroh and Simon, 1990). Our failure to observe an increase in \(\beta\)-galactosidase levels from the \(\text{opdA}^{10}:\text{Mu dJ}\) fusion after heat shock does not rule out the possibility that \(\text{opdA}\) is a heat shock gene but might be the result of the transient nature of the response, or suggest that the cell maintains a relatively constant level of OpdA whether heat stressed or not. Obviously additional experiments are required to examine the induction of \(\text{opdA}\) mRNA after heat shock, or the \textit{in vitro} transcription of \(\text{opdA}\) by \(\sigma^{32}\) containing RNA polymerase.
Chapter 3

*OpdA*, Encoding Oligopeptidase A, and *prlC*, Encoding Suppressors of Signal Sequence

Mutations, are the Same Gene
INTRODUCTION

One strategy used to identify components of the cell's protein secretion apparatus has been to isolate mutations which allow export of proteins with defective signal sequences. One class of suppressor mutations isolated in this way affects the prIC locus (Emr et al. 1981). The first prIC suppressor was selected as a mutation which allowed a strain carrying a 36 bp deletion (lamBss60) in the lamB signal sequence to grow on maltodextrins. LamB is an outer membrane protein that serves as the phage lambda receptor and is required for uptake of maltodextrins. Mutations at prIC are rare among lamBss60 revertants - only one allele (prIC1) was obtained in the original selection which produced mainly prL1A mutations (Emr et al. 1981). More recently, Thrun and Silhavy (1987) have used localized mutagenesis to isolate additional prIC alleles. One of the strongest prIC suppressor mutations, prIC8, obtained in this way, has been characterized in detail. Although strains carrying prIC8 efficiently remove the mutant lamBss60 signal sequence, very little of the LamB protein is exported to the outer membrane (Thrun and Silhavy, 1989). Thrun and Silhavy (1987) have cloned wildtype, suppressor, and amber alleles of prIC and shown that prIC is dispensable for growth. The prIC locus has been mapped to approximately 71 map units on the E. coli chromosome (Thrun, dissertation 1988).

The gene opdA was first identified by Vimr, Green, and Miller (1983) as the locus of mutations which prevented a peptidase deficient Salmonella typhimurium strain from using the peptide N-acetyl-L-alanyl-L-alanyl-L-alanyl-L-alanine (AcAla₄) as a sole nitrogen source. As described in the previous chapter, the opdA gene has been mapped to 76 minutes on the S. typhimurium chromosome by cotransduction with transposon insertions near asd. It has been cloned and sequenced and shown to encode a 680 amino
acid protein. Oligopeptidase A, the product of the opdA gene, is a metalloproteinase which has been purified from both E. coli and S. typhimurium (Novak et al., 1986). It has also been shown that oligopeptidase A is the major soluble activity in E. coli which is able to hydrolyze free lipoprotein signal peptide in vitro (Novak et al., 1986).

Several facts suggested the possibility that prlC and opdA might be related. prlC and opdA map in the same region of the chromosome. The molecular weights of their gene products are similar (68kD and 66kD respectively, determined by denaturing polyacrylamide gel electrophoresis) (Thrun, dissertation, 1988). In addition, the results of Novak and Dev (1989) suggest that oligopeptidase A can recognize a signal sequence although it can cleave it only after release from the precursor protein. The results in this chapter show that prlC and opdA are indeed the same gene.
MATERIALS AND METHODS

Bacterial strains and plasmids. Bacterial strains and plasmids used in this study are shown in Table 3.1. Except for *E. coli* strain DH5 (Hanahan, 1983), all strains used were derivatives of *S. typhimurium* LT2. Bacteria were routinely grown in LB medium (Maniatis *et al.*, 1982). To determine the ability to use AcAla₄ as a nitrogen source, the medium of Gutnick *et al.* (1969) (here called NN) was used. Plasmids and other markers were transferred between *S. typhimurium* strains using the generalized transducing phage P22 (HT12/4 int:103) (Roth, 1970). Plasmids pTA108, pTC100 (*prlC*⁺), pTC101 (*prlCI*), and pTC108 (*prlC8*) are described in Thrun and Silhavy (1987). Plasmid pTC231 is a pTC101 derivative carrying the *prlC* gene with both the *prlCI* allele and the *prlC31* amber mutation. This amber mutation eliminates the suppressor activity of the *prlCI*. Plasmid pTC169 consists of a 2.7kb fragment of pTC101, which contains the *prlCI* gene, in the vector pTA108 (Thrun, dissertation, 1988). Plasmid pCM191 carries a 2.9kb BamHI fragment of pTC169 containing *prlCI* in the BamHI site of pBR322.

Complementation of *opdA*⁻ phenotypes by *prlC* plasmids. Phage P22 forms pinpoint plaques on *opdA*⁻ strains (Vimr, dissertation, 1981). Complementation of this small plaque phenotype was tested by plating P22H5, a clear plaque mutant, on plasmid containing strains. An *opdA*⁻ *dcp*⁻ strain is unable to use AcAla₄ as a sole nitrogen source (Vimr *et al.*, 1983). The ability of plasmids to allow strain TN1201 (*opdA*⁻ *dcp*⁻) to use AcAla₄ as a sole nitrogen source was tested by plating plasmid containing strains on NN glucose medium containing 1.5mM AcAla₄. AcAla₄ hydrolysis by cell extracts was determined by HPLC analysis of reaction products as described above.

Benzylxycarbonyl Ala Ala Leu p-nitroanilide (Z-AAL-pNA) hydrolysis by cell extracts was determined as described above.
TABLE 3.1. Bacterial Strains and Plasmids

<table>
<thead>
<tr>
<th>Strain or plasmid</th>
<th>Genotype or description</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>S. typhimurium</strong></td>
<td></td>
</tr>
<tr>
<td>TN1727</td>
<td>leuBCD485 pepN90 pepAI6 pepB11 supQ302(ΔproAB pepD) pepPl pepQ1 pepTI dcp-1 zcf845::Tn10 zgh848::Tn5 opdA1</td>
</tr>
<tr>
<td>TN3031</td>
<td>TN1727/pJG13</td>
</tr>
<tr>
<td>TN3032</td>
<td>TN1727/pJG70</td>
</tr>
<tr>
<td>TN3033</td>
<td>TN1727/pTC100</td>
</tr>
<tr>
<td>TN3034</td>
<td>TN1727/pTC108</td>
</tr>
<tr>
<td>TN3035</td>
<td>TN1727/pTC231</td>
</tr>
<tr>
<td>TN3067</td>
<td>leuA414(Adj) hisC527(Adj) supJ60 zgh-848::Tn5 opdA1</td>
</tr>
<tr>
<td>TN3068</td>
<td>TN3067/pJG13</td>
</tr>
<tr>
<td>TN3070</td>
<td>TN3067/pTC100</td>
</tr>
<tr>
<td>TN3071</td>
<td>TN3067/pTC108</td>
</tr>
<tr>
<td>TN3072</td>
<td>TN3067/pTC231</td>
</tr>
<tr>
<td>TN1201</td>
<td>dcp-1 zcf845::Tn10 zgh-848::Tn5 opdA1</td>
</tr>
<tr>
<td>TN1292</td>
<td>dcp-1 zcf845::Tn10 zgh-848::Tn5</td>
</tr>
<tr>
<td>TN3393</td>
<td>purC7 recA1 srl-202::Tn10 / F'42 finP301</td>
</tr>
<tr>
<td>TN3392</td>
<td>TN3393/pTC100</td>
</tr>
<tr>
<td><strong>Plasmids</strong></td>
<td></td>
</tr>
<tr>
<td>pTC100</td>
<td>prlC+ in pSC101 based vector pTA108</td>
</tr>
<tr>
<td>pTC101</td>
<td>prlC1 in pTA108</td>
</tr>
<tr>
<td>pTC108</td>
<td>prlC8 in pTA108</td>
</tr>
<tr>
<td>pTC231</td>
<td>prlC1 prlC31(Adj) in pTA108</td>
</tr>
<tr>
<td>pTC169</td>
<td>2.7 kb prlC1 containing fragment from pTC101 in pTA108</td>
</tr>
<tr>
<td>pCM191</td>
<td>2.9 kb prlC1 containing fragment from pTC169 in pBR322</td>
</tr>
<tr>
<td>pJG13</td>
<td>opdA+ in pBR328</td>
</tr>
<tr>
<td>pJG70</td>
<td>opdA+ in pBR328</td>
</tr>
</tbody>
</table>

Tn1000 (γ8) insertions in pTC100. Tn1000 insertions in pTC100 were isolated using the method of Guyer (1983). Briefly, plasmid pTC100 was moved into strain TN3393 to make strain TN3392. Strain TN3393 contains F'42 which carries at least one Tn1000 insertion. Four mls each of exponentially growing cultures of TN3392 and TN1201 were combined and incubated at 37°C without shaking for one hour. These cells were vortexed vigorously, pelleted by centrifugation, and resuspended in 0.5 ml of LB broth. Samples were plated on LB agar containing both ampicillin (50μg per ml) and kanamycin (50μg per ml) and incubated overnight at 37°C. Individual colonies were patched to another LB ampicillin kanamycin agar plate and replica plated to NN medium containing 1.5mM AcAla₄ as sole nitrogen source. Strains which failed to use AcAla₄ were single colony isolated. Plasmid DNA was isolated from these strains and the locations of the Tn1000 insertions were determined by restriction mapping.

DNA sequencing and manipulations. DNA sequence was determined using the dideoxy nucleotide chain termination method (Sanger et al., 1977), double stranded plasmid templates (Chen and Seeburg, 1985) and Sequenase (USB) according to the manufacturer's directions. Template DNA was isolated using Qiagen columns (Qiagen Inc.) according to the manufacturer's directions. Primers for sequencing from the Tn1000 insertions were those described by Liu et al (1987). In addition, because of the high degree of homology between prlC and opdA, oligonucleotide primers which had been synthesized to sequence opdA were also used to sequence prlC. Restriction enzymes and T4 DNA ligase were purchased from BRL (Life Technologies Inc.) and used according to manufacturer's directions. Other manipulations were performed using standard techniques (Maniatis et al. 1982).
RESULTS

AcAla₄ and Z-AAL-pNA hydrolysis by soluble extracts of strains containing prlC plasmids. Oligopeptidase A, the product of the opdA gene is one of two activities in the cell which hydrolyze AcAla₄ (Vimr and Miller, 1983). Extracts of strains containing opdA plasmids have elevated AcAla₄ hydrolyzing activity. To test whether strains containing prlC plasmids have elevated levels of AcAla₄ hydrolyzing activity, the prlC plasmids pTC100, pTC108, and pTC231 were moved into TN1727 (opdA dcp pepN pepA pepB pepD pepT pepP pepQ). The specific activity of AcAla₄ hydrolyzing activity in the extracts was determined by J. Miller (Table 3.2). Extracts of the strains carrying pTC100 (prlC⁺) and pTC108 (prlC8) had specific activities much higher than the parent strain without any plasmid. The activity in the extracts of strains carrying pTC100 and pTC108 was lower than that in the strain carrying pJG13 (opdA⁺). This is expected since the prlC plasmids are derivatives of pTA108 which has a significantly lower copy number than pBR328 from which pJG13 was derived. Strain TN3035, containing pTC231 (prlCl prlC31Am), had a specific activity equivalent to the parent, indicating that the amber mutation which prevents the suppression of the lamb8s60 signal sequence mutation (Thrun, dissertation, 1988), also eliminated AcAla₄ hydrolyzing activity.

Extracts of peptidase deficient strains carrying prlC⁺ (pTC100) or opdA⁺ (pJG70) plasmids (strains TN3033 and TN3032 respectively) were also assayed for Z-AAL-pNA hydrolyzing activity. Z-AAL-pNA is a chromogenic peptide substrate of OpdA and was used to identify plasmids carrying the opdA gene (chapter 2). The specific activity of TN1727, the peptidase deficient host strain without any plasmid, was <0.002 μmol min⁻¹ mg⁻¹. The activity of TN3032 (pJG70) and TN3033 (pTC100) were 0.05 and 0.01 μmol min⁻¹ mg⁻¹ respectively. As described above, pTC100 is a low copy plasmid whereas
TABLE 3.2. AcAla$_4$ hydrolysis by soluble extracts of strains containing *prIC* or *opdA* plasmids

<table>
<thead>
<tr>
<th>Strain</th>
<th>Parent</th>
<th>Plasmid</th>
<th>Specific activity $\mu$M min$^{-1}$ mg$^{-1}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>TN3031</td>
<td>TN1727</td>
<td>pJG13 (opdA$^+$)</td>
<td>1.3</td>
</tr>
<tr>
<td>TN3033</td>
<td>TN1727</td>
<td>pTC100 (prICwt)</td>
<td>0.94</td>
</tr>
<tr>
<td>TN3034</td>
<td>TN1727</td>
<td>pTC108 (prIC8)</td>
<td>0.44</td>
</tr>
<tr>
<td>TN3035</td>
<td>TN1727</td>
<td>pTC231 (prICam)</td>
<td>0.002</td>
</tr>
<tr>
<td>TN1292 (opdA$^+$)</td>
<td>none</td>
<td></td>
<td>0.02</td>
</tr>
<tr>
<td>TN1727 (opdA$^-$)</td>
<td>none</td>
<td></td>
<td>0.001</td>
</tr>
</tbody>
</table>
pJG70 is a medium copy plasmid. It is not clear why the difference between the Z-AAL-pNA hydrolyzing activities of the plasmid containing strains is greater than the difference between AcAla₄ hydrolyzing activities. This inconsistency in relative activities may reflect a slight difference in substrate specificity between OpdA and PrlC.

Complementation of the opdA P22 small plaque defect by prlC plasmids. One characteristic of opdA mutations is that strains carrying these mutations fail to support the normal growth of phage P22. P22 makes small plaques on an opdA⁻ strain (Vimr, dissertation, 1981). This small plaque phenotype is complemented by opdA⁺ plasmids. In experiments performed by A. Kukral (unpublished results), the prlC plasmids pTC100 and pTC108 also complemented the small plaque phenotype, whereas pTC231 (prlCam) did not. When pTC231 (prlCam) was tested for complementation of the opdA small plaque phenotype in a strain carrying an amber suppressor (TN3070), normal plaques were formed. To confirm that pTC231 still carried the amber mutation, it was moved from TN3070 back into TN1727 (which contains no suppressor) where it once again failed to complement. Thus, the prlC plasmids complemented not only the Z-AAL-pNA and AcAla₄ defect, but also the P22 growth defect characteristic of opdA⁻ strains. Similarly, the prlC amber mutation which eliminated LamB signal sequence suppressor activity, also eliminated the ability of the prlC plasmid to complement the opdA peptide hydrolyzing defect and the P22 growth defect.

The opdA complementing activity and prlC are coincident on pTC100. Tn1000 insertions in pTC100 were isolated. Of 191 insertions characterized, 30 failed to complement the opdA⁻ AcAla₄ utilization defect. Sixteen of these 30 insertions were localized by restriction enzyme analysis, and found to lie in a 1.4kb region of pTC100 (Fig. 3.1). This 1.4kb region was within the 2.7kb region which had been shown (Thrun,
FIGURE 3.1. Location of Tn1000 insertions in pTC100. Tn1000 insertions were generated in pTC100 as described in the text, and were localized on the plasmid by restriction mapping and sequencing out from the Tn1000 ends. The insertions indicated by open circles, eliminated the ability of pTC100 to complement the AcAla\_4 utilization defect of the opdA\_ strain, TN1201. The thin line represents *E. coli* chromosomal DNA on pTC100, the thick lines at the ends indicate vector DNA. The hatched bar above the line indicates the region to which prlC had been previously localized by Thrun (dissertation, 1988). Restriction enzyme sites are: A, Av\_I; B, Bgl\_II; H, Hpa\_I; P, Pst\_I; and S, Sal\_I.
dissertation, 1988) by deletion analysis to contain \( prIC \). Thus the \( \text{Tn}1000 \) insertions which eliminated the \( opdA \) complementing activity of pTC100 mapped to the same region of the plasmid as \( prIC \). A 2.9kb fragment of pTC169 containing the 2.7kb region previously shown to encode \( prIC \) and 0.2 kb of pTA108 vector sequence, was cloned into pBR322. This plasmid, pCM191, complemented the \( opdA^+ \) small plaque defect in TN1201 and suppressed the \( lamB_{ts60} \) signal sequence mutation (A. Flowers, unpublished data). This result indicated that this same region was not only required for both \( opdA \) complementation and \( lamB_{ts60} \) suppression, but was also sufficient for both. This result also demonstrated that the difficulty cloning \( prIC \) on a medium copy plasmid (Thrun and Silhavy, 1987), not seen in cloning \( opdA \), was the result of other sequence around \( prIC \), and not the result of \( prIC \) itself.

The nucleotide sequence of \( prIC \) is very similar to that of \( opdA \). The wild type \( prIC \) gene from pTC100 was sequenced using primers to the \( \text{Tn}1000 \) insertions which eliminated \( opdA \) complementation, and primers originally synthesized for sequencing \( Salmonella typhimurium \) \( opdA \). The sequencing strategy is shown in figure 3.2, and the complete sequence is shown in figure 3.3. This sequence contained an open reading frame (orf) (positions 337 to 2377) encoding a 680 amino acid protein. This orf is preceded by a possible ribosome binding site. The region from nucleotide 243 to nucleotide 2578 showed 84% identity to \( opdA \) sequence. The predicted amino acid sequence of the 680 amino acid orf showed 95% identity to the predicted amino acid sequence of oligopeptidase A (\( opdA \)) (Fig. 3.3). Nucleotide sequence similarity extended 3' to the open reading frame to the end of the \( opdA \) sequence. Sequence similarity also extended 5' to the coding region until immediately after a putative \( \sigma^{32} \) heat shock promoter which is present in both the \( opdA \) and the \( prIC \) sequences (Fig. 3.3). Table 3.3
FIGURE 3.2. *priC* sequencing strategy. Arrows indicate extent (5' to 3') of nucleotide sequence obtained from individual priming sites. Small circles indicate the positions of Tn1000 insertions in pTC100.
FIGURE 3.3. *prlC* and *opdA* sequences and translation. The *prlC* and *opdA* nucleotide and predicted amino acid sequences are shown aligned with each other. Only bases or amino acids which differ between the two are indicated. The overlined *prlC* nucleotide sequence indicates the putative $\sigma^{32}$ promoter. The underlined *prlC* nucleotide sequence indicates the region which is duplicated in *prlCI*. 
shows the similarity between this 5' region of both prlC and opdA and known σ32 promoters.

The difference between the opdA and prlC sequences 5' to the putative σ32 promoter may explain the difficulty in subcloning a fragment carrying prlC and over 3kb of 5' DNA into pBR322 (Thrun and Silhavy, 1987). S. typhimurium opdA could be readily cloned into pBR vectors, and the 50 fold overproduction of oligopeptidase A that resulted did not affect cell growth. Southern hybridization of an internal opdA probe to E. coli and Salmonella chromosomal DNA identified a unique fragment in both. However, hybridization with a prlC probe which contained ~500 bp 5' of the coding region, identified the same band in E. coli but an additional band in Salmonella (Fig. 3.4). This result demonstrated that both E. coli and S. typhimurium contain the sequence upstream from E. coli prlC, but in different locations on their chromosomes. This is consistent with the different chromosomal map positions of opdA and prlC as determined by cotransduction with chromosomal markers.

A search of the GenBank sequence data base using the BLAST program (Altschul et al., 1990) indicated that the sequence upstream of prlC (bases 1 to 133 in figure) was 67% identical to an uncharacterized orf from Haemophilus influenzae (GenBank accession no. M62809). This orf is immediately downstream from the Haemophilus homolog of the E. coli gene mcrA which codes for penicillin-binding protein 1A (Tomb et al., in press). The mcrA gene maps at 75 map units on the E. coli chromosome.

Sequencing of the prlCI allele from pCM191 identified a 21 bp duplication at position 1422 in figure 3.3. This in-frame duplication results in the duplication of 7 amino acids in a region highly conserved among the related metallopeptidases prlC, opdA, dcp, and rat metallopeptidase EP 24.15. This is consistent with Thrun's result which
Table 3.3. Comparison of opdA and prlC promoter regions to known heat shock promoters

<table>
<thead>
<tr>
<th>Gene</th>
<th>-35 region</th>
<th>spacer</th>
<th>-10 region</th>
<th>+1</th>
</tr>
</thead>
<tbody>
<tr>
<td>groE</td>
<td>TTTCCCCCTTGAAG</td>
<td>GGGCGGAAGGCAAT</td>
<td>CCCATTTCTCTGG</td>
<td>TCAC</td>
</tr>
<tr>
<td>dnaKp1</td>
<td>TCTCCCCCTTGAAG</td>
<td>GACGTGGTTTACG</td>
<td>CCCATTTTGTAG</td>
<td>TCAC</td>
</tr>
<tr>
<td>dnaKp2</td>
<td>TTGGCCAGTTTGAAG</td>
<td>ACCAGACGTGTCG</td>
<td>CCCCTATTACAGAC</td>
<td>TCAC</td>
</tr>
<tr>
<td>htpGp1</td>
<td>GCTTCTCGTTGGAAG</td>
<td>ATATTTCTCCTTGT</td>
<td>CCCATCTCTCCT</td>
<td>ACACT</td>
</tr>
<tr>
<td>rpoDpHs</td>
<td>TGCCACCCCTTGAAG</td>
<td>AAACCTGTGATGTCG</td>
<td>GACGATATAAGG</td>
<td>ATRAT</td>
</tr>
<tr>
<td>htpX</td>
<td>ATCCAGACTTGAAG</td>
<td>AAATGGCTGGGTG</td>
<td>ACCCATACGATGTCG</td>
<td>GTAT</td>
</tr>
<tr>
<td>grpE</td>
<td>TGTTCCCTTGAAG</td>
<td>ACCCTGAAACTGAT</td>
<td>CCCATAATAGC</td>
<td>GAAG</td>
</tr>
<tr>
<td>topAa</td>
<td>ACAAGGGGTCTCA</td>
<td>ATCCGCAAGACAGCA</td>
<td>GTCGATATCGGTA</td>
<td>ACTC</td>
</tr>
<tr>
<td>htrC</td>
<td>AAACGTGACCTGAA</td>
<td>CTAATTGATGCAAATC</td>
<td>CCGGAAAGATGCG</td>
<td>ATAC</td>
</tr>
<tr>
<td>clpB</td>
<td>CAATAACCTTGAAG</td>
<td>TAATTGAGGAGTGAG</td>
<td>CTCATATTATCT</td>
<td>CAG</td>
</tr>
<tr>
<td>clpPa</td>
<td>TGTTATGCTTGAAG</td>
<td>TATGGTGATGCGGT</td>
<td>ACCATAACACAGG</td>
<td>ACTAG</td>
</tr>
<tr>
<td>lon</td>
<td>TCTCGGCGTTGAAAG</td>
<td>TGTGGGGAATACAT</td>
<td>CCCATGACTGAGCG</td>
<td>TACG</td>
</tr>
<tr>
<td>opdAa</td>
<td>GTCCATCATTGAAAG</td>
<td>ATTCCTTACTACTC</td>
<td>CCCATTCTAGATC</td>
<td>TCGAT</td>
</tr>
<tr>
<td>prlCa</td>
<td>GTTTCTCATTTGAAAG</td>
<td>ATTCCTACACTTTA</td>
<td>CCCATCGCTACACA</td>
<td>CAT</td>
</tr>
</tbody>
</table>

Consensus**<sup>a</sup>** t---c---Gcctgaa 14-15bp spacer ccccctacatc+a-t c


**a** Transcription start site has not been determined for opdA, prlC or clpP.

**b** Consensus sequence was derived from the above sequences. Frequencies of a given base at a given position are indicated as follows: >50%, lowercase; ≥75%, uppercase; and >90%, underlined.
FIGURE 3.4. Southern hybridization of *Salmonella* and *E. coli* chromosomal DNA with *priC* and *opdA* probes. Chromosomal DNA from *E. coli* K12 and *S. typhimurium* TN1379, was digested with *EcoRI*. Fragments were separated by agarose gel electrophoresis, then transferred to charged nylon membrane using a vacuum transfer apparatus (Hoeffer Scientific). The *opdA* probe was a 700bp *PvuII* fragment completely internal to the *opdA* coding region. The *priC* probe was a 1.7kb *AvaI/HpaI* fragment from pTC100 which included about 500 bp upstream of the *priC* coding region. Probes were labelled using random primers and [32P]dCTP (Feinberg and Vogelstein, 1983, 1984), hybridized to the membranes at 50°C overnight, then washed at 50°C in 15mM NaCl, 1.5mM sodium citrate, and 0.1% SDS. Autoradiograms were established on Kodak X-Omat AR film. The arrow indicates the band which hybridized to the *priC* probe, but not to the *opdA* one.
\begin{align*}
&\text{pCM127} \\
&\text{pTC100} \\
&\text{opdA probe} \\
&\text{prlC probe} \\
\end{align*}
shows by maxicell analysis that the PrlC protein made by strains carrying pTC101 (prlCI) is slightly larger than that made by strains carrying pTC100 (prlC+) (Thrun, dissertation, 1988). From this identification of the prlCI mutation and the mapping results of Thrun and Silhavy (1987) it can be deduced that the 8 prlC alleles which have been characterized, are either in the same region as prlCI or 3' from it.
DISCUSSION

It has been shown that the plasmids carrying the wild-type and suppressor alleles of \textit{prlC}, originally isolated by Thrun and Silhavy (1987), complemented an \textit{opdA} mutation and that a plasmid containing a \textit{prlC(am)} allele was unable to complement except in an amber suppressor background. Furthermore, Tn\textit{1000} insertions which eliminated the ability of \textit{pTC100} (\textit{pribC}+) to complement the \textit{opdA} \textit{AcAlaQ} utilization defect, mapped to the region of the plasmid previously shown to contain \textit{pribC}. I have sequenced this region and found that the \textit{pribC} gene from \textit{E. coli} is the same as the \textit{opdA} gene in \textit{S. typhimurium} (84\% nucleotide identity, 95\% amino acid identity). Thus it is clear that the \textit{pribC} mutations which suppress \textit{lamB} signal sequence mutations, are mutations in the metalloendoprotease oligopeptidase A.

A number of previous observations are relevant in considering how mutations that affect a protease might suppress defective localization caused by signal sequence mutations. Each of the eight \textit{pribC} suppressors characterized are dominant to the wild type allele (Thrun, dissertation, 1988, Emr \textit{et al.}, 1981; Thrun and Silhavy, 1987). The \textit{pribC} mutations are weak suppressors of the \textit{lamB} signal sequence mutant phenotype (Emr \textit{et al.}, 1981; Thrun and Silhavy, 1987). They are not effective as suppressors of all \textit{lamB} signal sequence mutations, but are specific for a subset of signal sequence mutations, namely those which involve changes in the hydrophobic core of the signal sequence (Emr and Bassford, 1982; Thrun and Silhavy, 1989). The rate and extent of \textit{in vivo} processing (conversion of a precursor form to a stable, smaller product form) of the mutant signal sequence protein, LamB\textit{S78}, by strains containing the two best characterized \textit{pribC} alleles, \textit{pribC1} and \textit{pribC8}, is almost normal. Very little LamB is however, exported to the outer membrane (Emr and Bassford, 1982; Thrun and Silhavy, 1989). In the case of \textit{pribC1}, the
processed forms of both LamBs78 and LamBs71, migrate more slowly than the correctly processed forms, suggesting that the processing in a prlCI strain may not be accurate (Emr and Bassford, 1982). In fact, the accuracy of the processing has not been determined for any of the prlC strains. The low level of correct localization of the LamB protein that does occur is dependent on secB and secA (Thrun and Silhavy, 1989). The processed LamB which remains in the cytoplasm is proteolytically unstable (Thrun and Silhavy, 1989). Finally, as has been shown here, prlC mutations are mutations in the metallopeptidase, oligopeptidase A.

One possible suppression mechanism is that the mutant PrlC proteases directly cleave the LamB signal sequence mutant precursor and that localization is secondary to this cleavage. This would suggest that the prlC mutations lead to altered substrate specificity of the protease so it can recognize and cleave the mutant precursor. In support of this hypothesis, evidence by Emr and Bassford (1982) suggests that although both LamBs71 and LamBs78 are processed in a prlCI strain, this processing may not produce the same product as that formed from the wild type LamB precursor. That the prlCI mutation and the seven other prlC mutations, have been localized to a region of the protein highly conserved among opdA, dcp, and rat metallopeptidase EP 24.15. suggests that they are in a region crucial to the function of these proteases. This region includes the thermolysin-like Zn²⁺ binding site, which is almost certainly part of the active site. Such an activity of the mutant PrlC protease is consistent with the apparent role for opdA in the N-terminal processing of a vital phage P22 protein (see chapter 4). Challenging this hypothesis that prlC suppressors process signal sequence mutants directly, is the proposal by Thrun and Silhavy (1989), that the processing in a prlC8 strain is dependent on leader peptidase, encoded by lep. Their conclusion is based on the observation that
dinitrophenol inhibits LamBs78 processing in a prlC8 strain, and on the results of experiments with lep(Am) mutants (Thrun and Silhavy, 1989).

The small amount of LamBs78 protein that does end up in the outer membrane of a prlC8 strain is transported in a secB/secA dependent manner (Thrun and Silhavy, 1989). Perhaps some small portion of the processed LamB is exported without the need for a cleavable signal sequence. This would be analogous to the case with lamBs60 protein. The lamBs60 mutation is a 36 base pair deletion in the signal sequence which was used to select for the prlC and prlA suppressor mutations. In the presence of these suppressors the lamBs60 protein is exported without being processed (Emr and Bassford, 1982). An alternative and exotic possibility is that at a low frequency, the prlC suppressor protein transpeptidates a good signal peptide onto the processed LamB protein, and the resulting chimera is exported to the outer membrane. Although this is speculative, oligopeptidase A, the opdA/prlC gene product, is able to recognize and hydrolyze free lipoprotein signal peptide (Novak and Dev, 1989), and it has been observed in vitro that purified oligopeptidase A appears to catalyze some sort of transpeptidation reaction involving the peptide Met-Met-Ala (J. Miller, unpublished observations).

If the PrlC protein does not directly process the mutant signal peptide it could indirectly facilitate processing by leader peptidase or some other protease. This could occur if the mutant PrlC protein, recognizes the mutant signal sequence, binds to the LamB protein, and prevents it from prematurely folding into a refractory conformation, playing a role similar to that proposed for secB (Kumamoto, 1989). The fact that prlC mediated export of the mutant LamB proteins is still secB dependent does not necessarily preclude an anti-folding mechanism of prlC suppression. This hypothesis is consistent with the proposed lep dependence of the processing which was questioned above.
However, it should be remembered that although almost all of the LamB protein is processed only a very small amount is actually exported. Obviously other hypotheses for \textit{prlC} suppression are possible, and more studies will be required to elucidate the actual mechanism.

Another question which is raised by the results reported here, is whether wild type PrlC/OpdA protein normally functions in protein export. Neither \textit{prlC(Am)} nor \textit{opdA} insertion mutations are lethal or even have a noticeable effect on cell growth as might be expected if they were necessary for as crucial an activity as protein localization. If PrlC/OpdA normally function in protein export, redundancy or non-essential function would make them dispensable. On the other hand, the \textit{prlC} suppressor phenotype could result from the recruitment of PrlC/OpdA into a protein export role unrelated to its normal function. Once again, more experimentation is needed to ascertain the normal role of \textit{prlC/opdA} in cellular metabolism.
Chapter 4

Role of the Metalloendopeptidase Oligopeptidase A in Phage P22 Development
INTRODUCTION

Oligopeptidase A (OpdA) was originally identified by Vimr and Miller (1983) as one of two activities in extracts of *Salmonella typhimurium* which hydrolyzed N-acetyl-L-Ala<sub>4</sub>. Oligopeptidase A has been purified from both *Escherichia coli* (Novak, Ray, and Dev, 1986) and *S. typhimurium*. Recent work by Novak, Ray, and Dev (1986) and Novak and Dev (1989) has shown that oligopeptidase A is the major soluble activity able to hydrolyze the lipoprotein signal peptide *in vitro*. Inhibition studies, and limited substrate specificity studies have shown that oligopeptidase A is a metalloendopeptidase (Vimr *et al.*, 1983; Novak *et al.*, 1986; Novak and Dev, 1989).

As described earlier, the *opdA* gene has been cloned and sequenced and shown to encode a 77kD protein. The predicted OpdA amino acid sequence includes a metalloprotease Zn<sup>2+</sup> binding site. Additionally, *opdA* is the same as *prtC*, the site of certain signal sequence suppressor mutations in *E. coli*, which proteolytically process LamB proteins with altered signal peptides.

Mutations were originally isolated in the gene encoding oligopeptidase A, *opdA*, by screening for mutants unable to use AcAla<sub>4</sub> as a sole nitrogen source. During the characterization of these *opdA* mutations, it was noticed that phage P22 made much smaller plaques on an *opdA*<sup>-</sup> host (Vimr, dissertation 1981). This is surprising because, unlike many phage, including the related phage λ, there are no known proteolytic processing steps in P22 development (Botstein *et al.*, 1973; Casjens and King, 1974). In this chapter I will show that OpdA is required for a previously unreported proteolytic processing step in P22 development, namely the removal of 20 amino acids from the N-terminus of the P22 gene 7 protein (gp7). I will describe the isolation and characterization of gene 7 mutants which function in the absence of OpdA, and present
the sequence of these gene 7 mutations and of the gene 7 and gene 14 region of wild type P22, and comment on what this sequence may indicate about the role of gp7 in phage P22 DNA injection.
MATERIALS AND METHODS

Bacterial and phage strains. All _S. typhimurium_ strains used were derived from LT2 and are shown in Table 4.1. _E. coli_ strain DH5 (Hanahan, 1983) was used to propagate plasmids for sequencing. Phage P22 strains used are also shown in Table 4.1. Bacteria were routinely grown in LB medium. LB with galactose, 0.05%, was used to propagate phage P22 on _galE_ strains. E medium (Vogel and Bonner, 1956) was used as the minimal medium and supplemented with appropriate amino acids at 0.4mM and glucose at 4g/liter. Antibiotics were used at the following concentrations: ampicillin 100μg/ml in liquid media, 50μg/ml in solid; chloramphenicol 25μg/ml; tetracycline 25μg/ml; and kanamycin 50μg/ml.

Propagation and purification of phage particles. Small scale lysates of P22 phage were made essentially as described by Roth (1970). Briefly, 0.25 ml of an overnight culture of host bacteria was added to 10^7 p.f.u. phage in 10 ml LB medium. Cultures were incubated at 37°C overnight in screw cap test tubes on a rotating wheel. After incubation, 0.5 ml CHCl₃ was added to each tube, vortexed, and incubated for at least 15 minutes more. The cell debris was removed by centrifugation at 2000 x g for 10 minutes, and the supernatant was transferred to a sterile tube and stored over a few drops of CHCl₃ at 4°C. For large scale purification of all phage except _φ3_ phage, host strains were grown in liquid culture to OD₆₀₀ 0.5, and then infected with phage at an m.o.i. of 1 and incubated for about 3h, before the addition of CHCl₃ to kill and lyse any remaining cells. DNase and RNase were added to a final concentration of 1μg/ml each and the solution incubated at room temperature for 30 min. NaCl was then added to 1M, and the solution incubated on ice for 1h. Cell debris was removed by centrifugation at 11,000 x g for 10 min. at 4°C. The phage were precipitated from the supernatant by the addition of
### Table 4.1. Bacterial and Phage Strains and Plasmids

<table>
<thead>
<tr>
<th>Strain</th>
<th>Genotype or Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>S. typhimurium</strong></td>
<td></td>
</tr>
<tr>
<td>TN1379</td>
<td>leuBCD485</td>
</tr>
<tr>
<td>TN2684</td>
<td>lon101 zba883::Tn10</td>
</tr>
<tr>
<td>TN3101</td>
<td>leuBCD485 opdA10::Mu dJ</td>
</tr>
<tr>
<td>TN3491</td>
<td>TN3101/pCM172</td>
</tr>
<tr>
<td>TN3625</td>
<td>TN1379/pCM187</td>
</tr>
<tr>
<td>TN3627</td>
<td>TN3101/pCM187</td>
</tr>
<tr>
<td>TN3629</td>
<td>TN1379/pCM185</td>
</tr>
<tr>
<td>TN3630</td>
<td>TN3101/pCM185</td>
</tr>
<tr>
<td>TN3647</td>
<td>leuBCD485 trp::[T7 gene 1 lacI2 spc']</td>
</tr>
<tr>
<td>TN3684</td>
<td>TN3647/pT7-5</td>
</tr>
<tr>
<td>TN3685</td>
<td>TN3647/pCM193</td>
</tr>
<tr>
<td>TN3803</td>
<td>TN3101/pDH3(mutD5)</td>
</tr>
<tr>
<td>TN3804</td>
<td>purC7 recA1 srl-202::Tn10/F'42,finP301/pCM172</td>
</tr>
<tr>
<td>DB7155</td>
<td>leuA414(Am) hisC527(Am) supE20 (Winston et al., 1979)</td>
</tr>
<tr>
<td>JR501 (TN3428)</td>
<td>hsdSA29 hsdSB121 hsdL6 metA22 metE551 trpC2 ilv-452 fla-66 rpsL120 xyl-404 galE719 nml H1-b H2-e,n,x (cured of Fels2)</td>
</tr>
<tr>
<td><strong>Phage</strong></td>
<td></td>
</tr>
<tr>
<td>P22 wild type (from R. Maurer)</td>
<td></td>
</tr>
<tr>
<td>P22 c2H5</td>
<td></td>
</tr>
<tr>
<td>P22 HT12/4 int103</td>
<td></td>
</tr>
<tr>
<td>P22 c2H5 13amH101</td>
<td></td>
</tr>
<tr>
<td>P22 14amH600 c2-5 steA44 (=MS523, from M. Susskind)</td>
<td></td>
</tr>
<tr>
<td>P22 14amH611 c2-5 steA44 (=MS679, from M. Susskind)</td>
<td></td>
</tr>
<tr>
<td>P22 7amH1035 cl-7 h21 (=MS1029, from M. Susskind)</td>
<td></td>
</tr>
<tr>
<td>P22 7amH1375 cl-7 (from A. Poteete)</td>
<td></td>
</tr>
<tr>
<td>P22 20amH1030 cl-7 h21 (=MS1212, from M. Susskind)</td>
<td></td>
</tr>
<tr>
<td>P22 7-1a c2H5 (=P22 1a)</td>
<td></td>
</tr>
</tbody>
</table>
P22 7-3a c2H5 (=P22 3a)
P22 7-4a c2H5 (=P22 4a)

**Plasmids**

<table>
<thead>
<tr>
<th>Plasmid</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>pBR328</td>
<td>(Soberon <em>et al.</em>, 1980)</td>
</tr>
<tr>
<td>pBR322</td>
<td>(Bolivar <em>et al.</em>, 1977)</td>
</tr>
<tr>
<td>pT7-5</td>
<td>derivative of pT7-1 (Tabor and Richardson, 1985)</td>
</tr>
<tr>
<td>pSE380</td>
<td><em>lacUV5</em> promoter, <em>lacI</em>, superpolylinker in pBR322 (Brosius, 1989)</td>
</tr>
<tr>
<td>pDH3</td>
<td><em>mutD5</em> on pBR322. From R. Maurer.</td>
</tr>
<tr>
<td>pCM146</td>
<td>9.2 kb <em>EcoRI</em> fragment of P22 4a in <em>EcoRI</em> pBR328</td>
</tr>
<tr>
<td>pCM147</td>
<td>5.6 kb <em>EcoRI/SalI</em> fragment of P22 4a in <em>EcoRI/SalI</em> pBR328</td>
</tr>
<tr>
<td>pCM148</td>
<td>3.2 kb <em>EcoRV/SalI</em> fragment of P22 4a in <em>SalI/PvuII</em> pBR328</td>
</tr>
<tr>
<td>pCM149</td>
<td>1.5 kb <em>EcoRV/AvaI</em> fragment of P22 4a in <em>AvaI/PvuII</em> pBR328</td>
</tr>
<tr>
<td>pCM150</td>
<td>1 kb <em>AvaI/BstEII</em> fragment of P22 4a with <em>BamHI</em> linkers added, in <em>BamHI</em> pBR322</td>
</tr>
<tr>
<td>pCM151</td>
<td>0.5 kb <em>EcoRV/BstEII</em>(blunt) fragment of P22 4a in <em>AvaI/PvuII</em> pBR328</td>
</tr>
<tr>
<td>pCM152</td>
<td>1.7 kb <em>SalI/HpaI</em> fragment of P22 4a in <em>SalI/PvuII</em> pBR328</td>
</tr>
<tr>
<td>pCM153</td>
<td>3.4 kb <em>EcoRI/SalI</em> fragment of P22 4a in <em>SalI/EcoRI</em> pBR328</td>
</tr>
<tr>
<td>pCM154</td>
<td>5.6 kb <em>EcoRI/SalI</em> fragment of P22 1a in <em>EcoRI/SalI</em> pBR328</td>
</tr>
<tr>
<td>pCM155</td>
<td>5.6 kb <em>EcoRI/SalI</em> fragment of P22 3a in <em>EcoRI/SalI</em> pBR328</td>
</tr>
<tr>
<td>pCM156</td>
<td>5.6 kb <em>EcoRI/SalI</em> fragment of P22 H5 in <em>EcoRI/SalI</em> pBR328</td>
</tr>
<tr>
<td>pCM158</td>
<td>3 kb <em>SalI/EcoRV</em>(HindIII linker) fragment of P22 4a in <em>SalI/HindIII</em> pSE380</td>
</tr>
<tr>
<td>pCM162</td>
<td>3.2 kb <em>EcoRV/SalI</em> fragment of P22 1a in <em>SalI/PvuII</em> pBR328</td>
</tr>
<tr>
<td>pCM163</td>
<td>3.2 kb <em>EcoRV/SalI</em> fragment of P22 3a in <em>SalI/PvuII</em> pBR328</td>
</tr>
<tr>
<td>pCM164</td>
<td>3.2 kb <em>EcoRV/SalI</em> fragment of P22 H5 in <em>SalI/PvuII</em> pBR328</td>
</tr>
<tr>
<td>pCM166</td>
<td>1.5 kb <em>EcoRV/AvaI</em> fragment of P22 1a in <em>AvaI/PvuII</em> pBR328</td>
</tr>
<tr>
<td>pCM167</td>
<td>1.5 kb <em>EcoRV/AvaI</em> fragment of P22 3a in <em>AvaI/PvuII</em> pBR328</td>
</tr>
<tr>
<td>pCM168</td>
<td>1.5 kb <em>EcoRV/AvaI</em> fragment of P22 H5 in <em>AvaI/PvuII</em> pBR328</td>
</tr>
</tbody>
</table>
pCM169 2.8 kb EcoRV/Clal(blunt) fragment of P22 1a in PvuII/SalI(blunt) pBR328
pCM170 2.8 kb EcoRV/Clal(blunt) fragment of P22 3a in PvuII/SalI(blunt) pBR328
pCM171 2.8 kb EcoRV/Clal(blunt) fragment of P22 H5 in PvuII/SalI(blunt) pBR328
pCM172 2.8 kb EcoRV(HinDIII linker)/Clal(blunt) fragment of P22 4a in Ncol(blunt)/HinDIII pSE380
pCM185 1.5 kb EcoRV(HinDIII linker)/Aval(blunt) fragment of P22 4a in Ncol(blunt)/HinDIII pSE380
pCM186 Ncol digested pSE380, made flush with S1 nuclease and Klenow, religated.
pCM187 1kb Aval(BamHI linker)/BsrEII(BamHI linker) fragment of P22 H5 in BamHI pCM186(=pSE380 NcolF) correct orientation for transcription of gene 7 from lacUV5 promoter
pCM188 1kb Aval(BamHI linker)/BsrEII(BamHI linker) fragment of P22 H5 in BamHI pCM186(=pSE380 NcolF) opposite orientation to pCM187
pCM193 EcoRI/HinDIII fragment from pCM187 in EcoRI/HinDIII pT7-5
pCM172γ67 Tn1000 insertion in pCM172, eliminates complementation of small plaque defect
pCM172γ610 Tn1000 insertion in pCM172, eliminates complementation of small plaque defect
pCM172γ611 Tn1000 insertion in pCM172, eliminates complementation of small plaque defect
pCM172γ617 Tn1000 insertion in pCM172, eliminates complementation of small plaque defect
pCM172γ619 Tn1000 insertion in pCM172, eliminates complementation of small plaque defect
polyethylene glycol (PEG 8000) to 10%. Phage were allowed to precipitate overnight at 4°C, pelleted at 11,000 x g for 15 min., and resuspended in 10mL TM (50mM Tris-HCl, pH 7.8, 10mM MgCl₂) (Sambrook et al., 1989). For I3 phage, cells were infected at an m.o.i. of ~10. Infected cells were incubated for 2 hours, after which they were pelleted by centrifugation, resuspended in 0.01 volume TM, and lysed by the addition of CHCl₃. Cell debris was removed from this concentrated phage suspension by centrifugation at 11,000 x g for 15 min. Concentrated phage were further purified by isopycnic CsCl gradient centrifugation. 0.75 g of CsCl was added to each ml of phage suspension, which was then centrifuged in 70Ti rotor for 24 to 48 hours at 150,000 x g at 4°C. The opalescent band of phage particles was removed by puncturing the side of the centrifuge tube. CsCl was removed by passing the phage suspension over a small Sephadex G-10 column (PD10, Pharmacia) previously equilibrated with TM. The resulting suspension was titered on TN1379 (sup⁰) or DB7155 (sup⁺) and the efficiency of plating (PFU/OD₂₆₀) was determined.

Isolation of phage mutants. An overnight culture of strain TN3803 (leu-485 opdA10::Mu dJ/pDH3(mudD5)) was diluted 1:100 into 50 mL LB + ampicillin and incubated with shaking at 37°C for 2 hours. P22 H5 (10⁸ p.f.u.) was added and the cultures were incubated overnight. CHCl₃ (0.5ml) was added to kill and lyse remaining bacteria. Lysates were spun in polypropylene tubes at 2000 x g for 10 min to pellet debris. The supernatant was titered on TN3101 (opdA10::Mu dJ). Large plaques were picked and repurified twice by streaking on a lawn of TN3101 or by suspending a single plaque in T2 buffer (0.03mM K₂SO₄, 0.07mM NaCl, 0.01mM KH₂PO₄, 0.02mM Na₂HPO₄, 1mM MgSO₄, 0.1mM CaCl₂, 0.001% gelatin), and plating dilutions for isolated plaques. Single plaques were used to make lysates on TN3101 which were then
stored at 4°C over CHCl₃. Only one mutant was saved from each original lysate.

Subcloning phage DNA. To purify phage DNA, phage particles were concentrated by centrifugation, 15000 rpm for 90 min. at 4°C in SS34 rotor, and resuspended in 1ml TM. This suspension was extracted with an equal volume of phenol (equilibrated with Tris-HCl pH7.8), then repeatedly extracted with an equal volume of phenol:CHCl₃:isoamyl alcohol (25:24:1, v:v:v) until no precipitate was seen at the interface. After a final extraction with CHCl₃:isoamylalcohol (24:1), the aqueous layer was removed and the DNA was precipitated by the addition of 0.1 volumes 3M sodium acetate, 2 volumes 95% ethanol and incubated at -20°C. DNA was pelleted at 12000 x g for 30 min., and washed with 0°C 70% ethanol. P22 DNA was digested with restriction endonucleases according to manufacturer’s recommendations (BRL). DNA fragments were purified using Gene Clean (Bio 101, Inc.), according to the manufacturer’s directions, after being separated on agarose gels. Fragments were ligated to cleaved pBR328 or pBR322 using standard protocols (Sambrook et al., 1989). Deletions were made in plasmids containing P22 DNA, by digesting with appropriate restriction endonucleases, blunting the ends with S1 nuclease and the Klenow fragment of DNA polymerase if necessary (Sambrook et al., 1989), and religating the plasmid. Ligation reactions (1 to 2 μl) were electroporated into S. typhimurium strain JR501 using a BioRad electroporation device and the manufacturer’s recommendations for electroporation of E. coli. Plasmids pCM150 and pCM179 were constructed by first digesting pCM149 or pCM156 with AvaI and BsrEII, blunting the ends of the DNA with S1 nuclease and the Klenow fragment of DNA polymerase, and adding BamHI linkers (NEB). The 1kb fragment of P22 DNA with BamHI cohesive ends was gel-purified, and ligated into pBR322, which had been digested with BamHI and dephosphorylated with calf intestinal
alkaline phosphatase (USB). The construction of plasmids used in this study is shown in figure 4.1.

**Marker rescue by plasmids containing P22 DNA.** To detect marker rescue by plasmids containing mutant phage DNA, strain JR501 containing the recombinant plasmid to be tested was grown overnight in LB+galactose. A small scale lysate was made on this host using P22 H5. This phage lysate was titered on both TN3101 (opdA\(^+\)) and TN1379 (opdA\(^-\)). The presence of large plaques on TN3101 indicated the presence of recombinant phage in the lysate. The number of large plaques on TN3101 divided by the number of plaques on TN1379 determined the frequency of recombination. Marker rescue of P22 strains containing amber mutations by plasmids containing phage DNA was performed by moving the plasmid into the supE strain, DB7155. This permissive host was infected with the mutant phage to be tested, and the resulting lysate was titered on both DB7155 (supE) and TN1379 (sup\(^B\)).

**Expression of P22 genes from plasmid borne trc and T7 promoters.** Two expression vectors were used in these studies, pSE380 (Brosius, 1989), and pT7-5, a derivative of pT7-1 (Tabor and Richardson, 1985). Plasmid pSE380 contains the following elements: the hybrid trc promoter, which is repressible by lac repressor; a polylinker following the trc promoter; the lacI gene, which encodes the lac repressor; the gene for \(\beta\)-lactamase; and the pBR322 origin of replication (Brosius, 1989). Plasmids pCM158, pCM172, and pCM185 contain DNA from P22 mutant 4a under control of the trc promoter, as described in figure 4.1. Plasmids pCM187 and pCM188 contain the 1kb Aval/BstEII fragment from P22 H5 behind the trc promoter. The P22 DNA in pCM187 is in the correct orientation for expression from the trc promoter, the phage DNA in pCM188 is in the opposite orientation.
FIGURE 4.1. Plasmids containing P22 DNA. Plasmids pCM146, pCM147, pCM148, pCM149, and pCM150 carry DNA from mutant P22 4a. pCM146 carries a 9.3kb EcoRI fragment in the EcoRI site of pBR328. pCM146 was digested with SalI and religated to make pCM147. pCM147 was digested with PvuII and EcoRV and religated to make pCM149. pCM149 was digested with Avel and BstEII, DNA ends were made flush with S1 nuclease and Klenow fragment of DNA polymerase I, and BamHI linkers were added. The 1kb fragment was purified and ligated into BamHI digested pBR322 to make pCM150.
Complementation of the small plaque phenotype by pCM172 was tested by diluting an overnight culture of TN3491 (opdA10::Mu dJ/pCM172) 1:100 into LB and incubating it at 37°C with shaking for 1-2 hours. The culture was made 1mM in isopropyl β-D-thiogalactopyranoside (IPTG) to induce the trc promoter. After an hour at 37°C, 0.5ml of this culture was combined with P22 H5 (200 p.f.u.) and 2.5 ml LB soft agar (0.75%agar) and plated on LB agar containing ampicillin and IPTG (1mM). These plates were incubated overnight at 37°C and then examined for large plaques.

Transposon Tn1000 insertions which eliminated the ability of pCM172 to complement the small plaque phenotype were isolated by the method of Guyer (1983). TN3804 (purC7 recA1 srl-202::Tn10/F’42 finP301/pCM172) was mixed with TN3101 and transconjugants were selected on LB+ampicillin+tetracycline. Transconjugants were purified and tested for the ability of the plasmid in them to complement the small plaque phenotype. Of the 32 transconjugants tested, 5 failed to complement the small plaque phenotype. The locations of the 5 Tn1000 insertions which prevented complementation, were determined by restriction mapping and later confirmed by nucleotide sequencing.

Pulse chase experiments were performed to visualize the proteins expressed from pSE380 based plasmids. An overnight culture of the plasmid containing strain was diluted 1:50 into minimal glucose medium with leucine and ampicillin. The culture was incubated at 37°C until it reached OD_{600} of 0.8 then IPTG was added to 1mM. After further incubation, to allow induction of the trc promoter, a mixture of ^{14}C-amino acids (55.5mCi/matom carbon) was added to a final concentration of 10 or 30μCi/ml. After either 1 minute or 5 minutes, casamino acids were added to 1%. Samples (0.2ml) were taken at times after the addition of casamino acids and put on ice. Cells were pelleted, resuspended in 20 or 50μl SDS-sample buffer (62.5mM Tris-HCl, pH 6.8, 2% SDS,
10% glycerol, 5% 2-mercaptoethanol), and heated to 95°C for 3-5 minutes. If necessary, samples were briefly sonicated to reduce viscosity before being loaded onto SDS-polyacrylamide gels for electrophoresis. SDS-polyacrylamide gel electrophoresis was performed using 12% polyacrylamide slab gels and the Tris-glycine buffer system of Laemmli (1970). After electrophoresis, gels were stained with Coomassie Blue, destained, and dried. Bands of radiolabelled protein were detected by exposure to Kodak X-Omat AR film.

The 1kb AvaI/BsrEII fragment from P22 H5 (see figure 4.1) was also expressed from the phage T7 promoter on plasmid pT7-5. The EcoRI/HindIII fragment from pCM187, which contains the 1kb AvaI/BsrEII fragment from P22 H5, was cloned behind the phage T7 promoter in pT7-5. This promoter is specific for T7 RNA polymerase and is not recognized by the bacterial RNA polymerase. This construction was moved into TN3647 to make TN3685. The trp operon of TN3647 contains an insertion composed of the following; the gene for T7 RNA polymerase under control of the lacUV5 promoter, the lacI gene, and a gene for spectinomycin resistance. (TN3647 was constructed by J. Miller). An overnight culture of TN3685 was diluted 1:50 into minimal glucose medium with leucine, tryptophan, and ampicillin, and incubated at 37°C until it reached an OD₆₀₀ of 0.5. Expression of the T7 RNA polymerase was induced by adding IPTG to 1mM. After 30 min. at 37°C, rifampicin (20mg/ml in dimethyl sulfoxide) was added to 0.4mg/ml in order to inhibit bacterial RNA polymerase and reduce labeling of host proteins. After 30 min. at 37°C to allow pre-existing bacterial messages to be degraded, ¹⁴C-amino acids (55.5mCi/matom carbon) were added to 3μCi/ml. After one minute, casamino acids were added to 1% final concentration. Samples were removed at various times after the pulse of radioactivity, pelleted, resuspended in 50μl SDS sample buffer,
boiled for 3 minutes, and electrophoresed through 10% polyacrylamide gels and Tris-tricine buffer (Schägger and von Jagow, 1987). After staining with Coomassie Blue, gels were dried and radiolabelled proteins were detected by autoradiography.

**DNA sequencing.** All DNA sequencing was performed using the dideoxy nucleotide chain termination method (Sanger *et al.*, 1977), and Sequenase (USB). Double-stranded plasmid templates prepared as described by Chen and Seeburg (1985). The primers used are shown in Table 4.2. Sequence from the *AvaI/BsrEII* fragments of pCM150 and pCM179 was obtained using primers to the vector. Based on this sequence, oligonucleotide primers were synthesized in order to obtain the rest of the sequence, and to sequence this region from other phage strains. Sequence was also obtained by priming from the ends of the Tn1000 insertions in pCM172 (Liu *et al.*, 1987). This sequence was used to design oligonucleotide primers which were used to obtain the sequence of the neighboring gene *I4* region. The sequence of the *I4* amber mutations was obtained by first amplifying the region by symmetric polymerase chain reaction (PCR). The template was phage DNA from strains MS523 (*I4am* H600) and MS679 (*I4am* H611) was prepared as described above. Primers used were P22-5 and Oligo-14. The symmetric PCR reaction included 1-10 ng template DNA, 10 pmol each primer, 16 μl 1.25 mM dNTP mix, 10 μl 10X PCR reaction buffer, and 0.5 μl Amplitaq polymerase (Perkin Elmer Cetus) in a 100 μl reaction. DNA was denatured at 95°C for 1 min., hybridized at 60°C for 1 min., and extended at 72°C for 2 min. for 20 cycles. To generate single stranded sequencing template, 2 μl of the symmetric reaction was used as the template in a reaction as above except only one primer was added and the number of cycles was increased to 30. The single stranded product was purified for sequencing by precipitation with 2M ammonium acetate and 1 volume of isopropanol.
Table 4.2. Primers used to sequence phage P22

<table>
<thead>
<tr>
<th>Name</th>
<th>Sequence</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>CC1</td>
<td>TCAATAAGTATACCAT</td>
<td>Tn1000 gamma end&lt;sup&gt;3,4&lt;/sup&gt;</td>
</tr>
<tr>
<td>CC2</td>
<td>GAATTATCTCTTAAACG</td>
<td>Tn1000 delta end&lt;sup&gt;3,4&lt;/sup&gt;</td>
</tr>
<tr>
<td>NEB1223</td>
<td>CACTATCGACTACCGATCA</td>
<td>BamHI site of pBR322 (clockwise)&lt;sup&gt;2&lt;/sup&gt;</td>
</tr>
<tr>
<td>NEB1219</td>
<td>ATGGCTCCGCGGCCTAGA</td>
<td>BamHI site of pBR322 (counterclockwise)&lt;sup&gt;2&lt;/sup&gt;</td>
</tr>
<tr>
<td>P22-1</td>
<td>TGAATAACGGGAGATAA</td>
<td>position 52-71&lt;sup&gt;5&lt;/sup&gt;</td>
</tr>
<tr>
<td>P22-2</td>
<td>TGGCAAGCAACCTGCTTG</td>
<td>position 1560-1544&lt;sup&gt;5&lt;/sup&gt;</td>
</tr>
<tr>
<td>P22-4</td>
<td>ACCAATACCACCAAGGA</td>
<td>position 1279-1262&lt;sup&gt;5&lt;/sup&gt;</td>
</tr>
<tr>
<td>P22-5</td>
<td>TCGTGCTGCTATGTTG</td>
<td>position 897-880&lt;sup&gt;5&lt;/sup&gt;</td>
</tr>
<tr>
<td>P22gp14-1</td>
<td>CGTACATGGTCTGTCGG</td>
<td>position 284-302&lt;sup&gt;6&lt;/sup&gt;</td>
</tr>
<tr>
<td>P22gp14-2</td>
<td>GCCGTGGCAGTCAAAGGT</td>
<td>position 354-336&lt;sup&gt;6&lt;/sup&gt;</td>
</tr>
<tr>
<td>Oligo-14</td>
<td>AGCCGAATCTGAAAATCCAGGC</td>
<td>position 52-71&lt;sup&gt;6&lt;/sup&gt;</td>
</tr>
</tbody>
</table>

<sup>1</sup>Numbers refer to the position of the primer on the sequence in figure 4.6.
<sup>2</sup>Oligonucleotides purchased from New England BioLabs.
<sup>3</sup>Oligonucleotides as described in Liu et al. 1987.
<sup>4</sup>Synthesized by Multidisciplinary Laboratory, School of Medicine, Case Western Reserve University, Cleveland, OH.
<sup>5</sup>Synthesized by Oligos Etc. Inc., Guilford, CT.
<sup>6</sup>Synthesized by University of Illinois Biotechnology Center Genetic Engineering Facility, Urbana, IL.
RESULTS

The *opdA* gene was first identified as the locus of mutations which eliminated the ability of strains to hydrolyze N-acetyl L-Ala\(_4\) and to use it as a sole nitrogen source (Vimr *et al.*, 1983). During the characterization of these mutations Vimr noticed that bacteriophage P22 makes small plaques on an *opdA*\(^-\) host (Vimr, dissertation, 1981). He showed that two independent *opdA* mutations, *opdA1* and *opdA2*, had the same effect on P22 plaque size. Additionally strains carrying *opdA3*, a leaky mutation which decreased but did not eliminate N-acetyl Ala\(_4\) hydrolysis, yielded plaques which were intermediate in size between those on *opdA1* and *opdA*\(^+\). He tested the *opdA*\(^-\) plaque phenotypes of several phage including P22, A3, L, KB1, 9NA, and ES18.h1. Of these only P22 and the closely related phage L and A3 made smaller plaques on an *opdA*\(^-\) host than on an *opdA*\(^+\) host (Vimr, dissertation, 1981). These results, and my observation that both *opdA*\(^+\) and *opdA*\(^-\) strains had the same generation times in LB, 24 minutes at 37°C, demonstrate that the small P22 plaques are not the result of a general growth defect, but are the result of a specific effect on the growth of P22 and closely related phage.

When moving *opdA* mutations by P22 transduction, it was noticed that the number of transductants from a lysate made on an *opdA*\(^-\) strain was much lower than that typically obtained from a lysate made on an *opdA*\(^+\) strain. However the defect in phage production was not in the ability of an *opdA*\(^-\) host to be infected. The number of transductants obtained upon transducing an *opdA*\(^-\) strain by P22 HT12/4 *int-3* transducing lysates was the same as that obtained from an *opdA*\(^+\) strain. P22 H5 was also able to infect an *opdA*\(^-\) host with the same efficiency as it did an *opdA*\(^+\) host. However, infection by P22 H5 of the *opdA*\(^-\) host, TN3101, resulted in a burst of phage that was only 5-10% that of an *opdA*\(^+\) host, TN1379; 38 and 450 PFU/infectious center.
respectively.

Phage P22 lysates were made on opdA+ and opdA− hosts. The phage particles were purified by isopycnic CsCl gradient centrifugation and the efficiency of plating (e.o.p.) was measured as PFU/OD_{260}. The e.o.p. of phage grown on an opdA− host was about 5% that of phage grown on an opdA+ host, 2.3 \times 10^{10} and 6.5 \times 10^{11} PFU/OD_{260} respectively. In other words, about 95% of the phage produced from an opdA− infection were defective. Thus the small plaques appeared to be the result of a reduced yield of viable phage after infection of an opdA− host.

Isolation of opdA independent phage. In an attempt to define the P22 function which requires opdA for full activity, phage mutations were sought which led to larger plaques on an opdA− host. To enrich for such mutants, P22 H5 lysates were made on TN3101 containing plasmid pDH3. TN3101 contains a non-reverting opdA10::Mu dJ insertion in opdA, and pDH3 contains the E. coli mutD5 allele to enhance mutation frequency. Three independent phage mutants were isolated which made large plaques on TN3101. Because S. typhimurium LT2, from which the host strain was derived, has been shown to release several different kinds of phage following P22 infection (Young et al., 1966) it was confirmed that the opdA independent mutant phage were indeed P22. First, the mutant phage made clear plaques, as did the parent phage, P22 H5. Second, the fragments generated when DNA from these phage was digested with three endonucleases, BamHI, EcoRI, and HindIII, were the same sizes as those from P22 H5 (Fig. 4.2). Thus we were confident these mutant phage were P22.

Figure 4.3 shows plaques made by the mutant phage on lawns of opdA+ and opdA− bacteria. Although all three mutants made larger plaques on an opdA+ lawn than did wild type P22 H5, the plaque phenotypes of the three mutants were different. Of the
FIGURE 4.2. Restriction endonuclease digestion of DNA from opdA suppressing P22 phage mutants. Phage DNA was digested with HinDIII, BamHI, and EcoRI, separated by agarose gel electrophoresis, and stained with ethidium bromide. Lanes marked W are wild type P22, and those marked H5 are P22 H5, the clear plaque parent of the mutant phage. Lanes 1, 3, and 4 are opdA suppressing mutants P22 1a, P22 3a, P22 4a respectively. Lanes marked S are λ DNA digested with HinDIII to generate size markers.
FIGURE 4.3. P22 plaques on opdA⁺ and opdA⁻ hosts. 0.05 ml of overnight cultures of strains TN1379 (opdA⁺) and TN3101 (opdA⁻) were combined with 2.5 ml of molten LB soft agar and 200 to 300 pfu's of phage. This was poured over an LB agar plate, and incubated at 37°C overnight. Plaques are shown approximately actual size.
three, mutant 4a made the largest plaques on an opdA+ lawn. Plaques made by mutant 4a on an opdA+ lawn were similar in size to those made on an opdA- lawn. Plaques made by mutant 3a on an opdA+ lawn were generally slightly smaller than those on an opdA- lawn, although this effect was variable. Plaques of mutant 1a on an opdA- lawn were significantly smaller than on an opdA+ lawn (although still much larger than those made by the parent). Mutant phage were purified from TN1379 (opdA+) and TN3101 (opdA-) by isopycnic CsCl gradient centrifugation and the efficiency of plating (PFU/OD260) was measured for each mutant on an opdA+ host (TN1379). The relative efficiencies of plating, opdA+/opdA+, of P22 H5 and the mutants were as follows: P22 H5, 0.05; mutant 4a, 1.3 mutant; mutant 3a, 0.91; and mutant 1a, 0.27. These results are generally consistent with the plaque sizes. They also demonstrate that the mutations in these phage allow them to produce a higher proportion of viable particles in the absence of opdA than are produced by the parent phage P22 H5. Furthermore, they demonstrate that these mutants were different from each other.

Mapping the phage mutations. A marker rescue strategy was adopted to identify the P22 gene or genes carrying the opdA-independent mutations. Restriction fragments of DNA from the mutant phage were subcloned into pBR322 or pBR328. Lysates, using P22 H5, were made on permissive opdA+ strains carrying these plasmids, and these lysates were screened for the presence of recombinant phage by plating on an opdA- lawn and looking for large plaques. Initially, an 8kb EcoRI fragment which contained most of the P22 structural genes (genes 8, 5, 4, 10, 26, 14, 7, 20, and 16), was cloned and found to contain the mutation. Progressively smaller fragments were screened until a 1kb A\text{val}/BstEII fragment, in pCM150, was identified which contained the mutation from mutant 4a (Fig. 4.4). Mutations from mutants 1a and 3a were mapped to a
slightly larger 1.5kb *Aval/EcoRV* fragment which included the 1kb *Aval/BstEII* fragment (pCM166 and pCM167 respectively) (see Table 4.3). Both the *Aval/EcoRV* fragment and the *Aval/BstEII* fragment from P22 H5 were also cloned, resulting in plasmids pCM168 and pCM179 respectively. Based on the restriction map of P22 (Casjens *et al.*, 1983; Chisholm *et al.*, 1980), the 1kb *Aval/BstEII* fragment probably contained all or part of genes 14, 7, or 20. To further identify the genes, or parts of genes, present on this fragment the ability of pCM179 to rescue P22 strains with amber mutations in genes 14, 7, and 20 was tested. Plasmid pCM179 was able to rescue gene 7 amber mutation H1035, and not gene 20am H1030 or 14am H600. A 1.5kb *Aval/EcoRV* fragment in pCM168, which extended 500 bp 3' of the fragment in pCM179, was able to rescue the gene 20 amber mutation. This showed that the 1kb fragment on pCM179 contained the region of gene 7 which included the amber mutation H1035, and did not contain the regions of genes 14 and 20 represented by mutations H600 and H1030 respectively.

**Sequence of P22 genes 7 and 14.** This region of phage P22 had not been sequenced. In order to characterize the *opdA* independent mutations further we first sequenced the 1kb *Aval/BstEII* fragment from P22 H5, then 566 bp 5' to the *Aval* site were also sequenced. The sequencing strategy is shown in figure 4.5, and the nucleotide sequence and the translations of predicted open reading frames (orfs) are shown in figure 4.6. The *Aval/BstEII* region contained a 229 amino acid orf, positions 611 to 1297 in figure 4.6. The molecular weight of the protein predicted by this orf was 23405Da. This suggests that this orf is gp7 (mw 22kD) (Youderian and Susskind, 1980) not gp14 (mw 15.5kD) (Youderian and Susskind, 1980). The gene 7 orf was confirmed by sequencing two gene 7 amber mutations, H1035 (Poteete and King, 1977) and H1375 (Youderian and Susskind, 1980). The sequence differences in the amber mutations are
FIGURE 4.4. Mapping the opdA suppressing mutation in mutant P22 4a by marker rescue. P22 H5 lysates were made on strains of JR501 containing plasmids carrying various regions of P22 4a DNA. These lysates were titered on both TN1379 (opdA\(^+\)) and TN3101 (opdA\(^-\)). The number of large plaques on TN3101 compared to the total number of plaques on TN1379 determined the recombination frequency. The approximate positions of P22 genes are indicated by the shaded bars. Below them is a restriction map of the relevant region of the P22 genome. Restriction sites are indicated as follows: RI, EcoRI; S, SalI; A, AvaI; H, HpaI; B, BstEII; and RV, EcoRV. The P22 DNA carried on each plasmid is indicated by the open bars.
Plasmid | Recombination frequency
---|---
pCM146 | 5.6%
pCM153 | <0.002%
pCM147 | 1.6%
pCM148 | 0.19%
pCM149 | 0.04%
pCM152 | <0.003%
pCM150 | 0.025%
pCM151 | <0.0008%
Table 4.3. Marker Rescue of \( opdA^- \) Independent Mutations by Plasmids Containing Mutant Phage P22 DNA

<table>
<thead>
<tr>
<th>Phage DNA on Plasmid</th>
<th>Plasmid</th>
<th>Recombination Frequency, %</th>
</tr>
</thead>
<tbody>
<tr>
<td>P22 4.a</td>
<td>pCM146</td>
<td>5.6</td>
</tr>
<tr>
<td></td>
<td>pCM147</td>
<td>1.6</td>
</tr>
<tr>
<td></td>
<td>pCM153</td>
<td>&lt;0.002</td>
</tr>
<tr>
<td></td>
<td>pCM148</td>
<td>0.19</td>
</tr>
<tr>
<td></td>
<td>pCM152</td>
<td>&lt;0.003</td>
</tr>
<tr>
<td></td>
<td>pCM149</td>
<td>0.04</td>
</tr>
<tr>
<td></td>
<td>pCM151</td>
<td>&lt;0.0008</td>
</tr>
<tr>
<td></td>
<td>pCM150</td>
<td>0.025</td>
</tr>
<tr>
<td>P22 1.a</td>
<td>pCM154</td>
<td>0.22</td>
</tr>
<tr>
<td></td>
<td>pCM162</td>
<td>1.7</td>
</tr>
<tr>
<td></td>
<td>pCM169</td>
<td>0.14</td>
</tr>
<tr>
<td></td>
<td>pCM166</td>
<td>0.04</td>
</tr>
<tr>
<td>P22 3.a</td>
<td>pCM155</td>
<td>1.03</td>
</tr>
<tr>
<td></td>
<td>pCM163</td>
<td>0.9</td>
</tr>
<tr>
<td></td>
<td>pCM170</td>
<td>0.19</td>
</tr>
<tr>
<td></td>
<td>pCM167</td>
<td>0.03</td>
</tr>
<tr>
<td>P22 H5</td>
<td>pCM156</td>
<td>&lt;0.003</td>
</tr>
<tr>
<td></td>
<td>pCM164</td>
<td>&lt;0.003</td>
</tr>
<tr>
<td></td>
<td>pCM171</td>
<td>&lt;0.003</td>
</tr>
<tr>
<td></td>
<td>pCM168</td>
<td>&lt;0.005</td>
</tr>
<tr>
<td></td>
<td>pCM179</td>
<td>&lt;0.005</td>
</tr>
</tbody>
</table>

Plasmids were constructed which carried DNA from \( opdA^- \) independent phage mutants or wild type phage P22H5 (see Fig. 4.1). \( opdA^+ \) strains containing these plasmids were infected with P22 H5. The resulting lysates were titered on both \( opdA^+ \) (TN1379) and \( opdA^- \) (TN3101) host strains. The percentage of total plaques which were large on TN3101 is given as percent recombination.
FIGURE 4.5. P22 sequencing strategy. Arrows indicate extent of sequence (5' to 3') obtained from each priming site. Open circles indicate Tn1000 insertions in pCM172 which eliminated the ability of this plasmid to complement the small plaque phenotype in TN3101.
FIGURE 4.6. Nucleotide sequence and predicted open reading frames. Gene 14 orf extends from 153 to 605. Gene 7 orf extends from 611 to 1297. Gene 20 orf begins at 1311 and continues beyond the available sequence. Amber mutations sequenced are indicated above the nucleotide sequence. The underlined amino acid sequence indicates the region of gp7 removed during processing.
1 ATG AAA GGC GTA TGT CCG ATG CAA TTC GCC CTC ACC CAT ACT TAC AAC GAG GAC ACG GCT ATA GCC
   Asn Lys Gly Val Phe Asp Ala Asp Leu Thr Phe Ala Val Ser Asp Thr Tyr Gin Ser Glu Ile Gin Ile Ala Ala

79 ATG GGT CTA ATT ACT GAG CGG CGC ACT AAG GCT TTG GAA GAC GCC TGG CTC GCA CAT GGG TTG ATT CAT TA ATG ATT
   Asn Ala Leu Ile Thr Glu Arg Arg Ala Asp Leu Ala Leu Arg Arg Ala His Gly Leu Ile Asp Met Ile 2

159 ACC TTC ACT CCA ACA GGC ACC ACG CTC GTA GAA ATG GCC GCG CAC CCC GAC ATT GCC GCA ACC AAC
   Thr Phe Thr Pro Arg Asn Ile Asp Thr Glu Met Val Gly Asn Ala Pro Asp Ile Ala Gly Ser Asn Asn

237 GCT GAC GCA TAC GAC TAC ARG CCT GAG TGT CTC GTC TTT GAA GTA GAC GTA CAT GCT GAC TTC ATG TAC GAC
   Gly Asp Gly Tyr Asp Tyr Tyr Gly Pro Glu Cys Arg Tyr Phe Gly Val Asn Val Asp Gin Cys Phe Gly Ile Val Tyr

315 TAC ACG AAC ATT AGG CGG ATC TCC TAT CTC CAC GCC ATG TAC CTG CCC CAC ATG ACC AGC AAC
   Tyr Asn Glu Ile Gin Pro Met Thr Phe Asp Cys His Ala Met Tyr Leu Pro Glu Ile Arg Gly Phe Ser Lys Glu Ile

(16am H6603, H6511) T

395 GCA CTG GCG TTC TCG CAA TAT ATT CTC ACC AAC ACC GTT CAG TGC GCT ACA TCA TTT GCT GCA CCG AAT AAA TCG
   Gly Leu Ala Phe Thr Arg Tyr Ile Leu Thr Asn Thr Val Thr Val Gin Cys Thr Ser Phe Ala Arg Lys Phe Ala

471 GAC GCT CAC AGT TAC TGC GCA ATG ATT GGC CTT AGT GCT GCA ACC ATG AAG AAA TAC TIC AAA GCC GTA GAT GAC
   Asp Gly Gin Met Tyr Cys Ala Met Ile Gin Gly Leu Arg Val Gly Thr Tyr Lys Lys Phe Gly Lys Gly Asp

549 TGG GCG TTT TAC GCC GCC ACC CGA GAA GAG TTA ACC GAA YTA CTG AAT AAC GGG AGA TAA ACG ATG TTA CAT GCA TTT ACG
   Val Thr Phe Thr Ala Ala Thr Arg Glu Ile Gin Pro Ser Tyr Pro Gin Lys Gin Lys Gin Gly Ala Ser Lys Ala

629 CTT GGC ACC AAA CTG CGC GCT GGT GAG GAA OTC TAT CCT GCT AAT CAA AAT ACC ACG CAG CCT TCC AAT ACC ATC ACG AAG
   Ser Gly Arg Arg Ala Asp Leu Glu Arg Gin Pro Ser Tyr Pro Gin Lys Gin Lys Gly Asp Ser Lys Ala

707 TAT GCA GCC GAA CGG CCA AAG TAT GCC GCA GGT CCA AAC CAG CAG TAC TCT AAC ACC ACG AAG CGG CAC AAC
   Tyr Ala Ala Glu Ala Gin Lys Tyr Ala Ala Asp Leu Gin Asn Gin Cys Asn Phe Thr Pro Leu Ala Asp Lys Tyr Ile Gin Ser Leu Gin Leu Gin

785 TTT ACT CTT CTG GCA GAT AAG AGT ATC TAT GGC GCC GCA GCC GCA ACC ATG CTT CAA GCT TTA TCG TCT CTC GTA GGT CAG GGG CAG GCC
   Tyr Phe Thr Pro Leu Ala Asp Lys Tyr Ile Gin Ser Leu Gin Leu Ser Ser Leu Gin Gin Gin Gin Leu Asn Ala

863 ATT TAC TAT AAC TCC CCA GAA CAA CAC TAC CAG CAC CTC GGC CAG CCT GCC CTA TAT CAG CAA GTC CGG GCC GCA GCC
   His Tyr Tyr Ser Ser Gin Met Ile Gin Ala Gin Ala Asp Gin Cys Gin Cys Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin

941 GGT CTC GGT GAT TCT ACA GCC ACC AGT AAC CAC CTT TCA GCC ATC GCC CCA CTT GAT CAG TCA GCT GTC GAT
   Gly Leu Leu Gin Ser Thr Thr Gin Ser Thr Gin Leu Gin Cys Thr Gin Leu Gin Gin Leu Leu Leu Gin Gin Gin Gin Leu

T (7am H17375) T

1019 CAG ATG AAT AAC TAT CAC ACC CTC GCA AAT ATT GAT GTC ATT GCT GCC CTT CAG CGG CAG CAC ACA TAT
   Gin Met Asn Asn Asp Thr Gin Asn Leu Ala Asn Gin Leu Asn Gin Asn Ala Asn Gin Thr Gin Thr

1097 GCC AAC ATG AAT AGT CAG TTA CAG CAA ACG GCC CCT GCC GCC GCA AAT GCC AAC AGA CCA TCA GCT ATG CAA
   Ala Asn Cys Met Gin Asp Thr Cys Ala Ala Asp Ala Asp Ala Arg Cys Gin Gin Gin Gin Gin

1175 TGT CTT ATT GCG GGA GGT CGC TCT GGT GTC ATT GGT CGT GCA GAT GGC CAA TTA ATT ATT ATT ATT ATT ATT ATT ATT ATT ATT ATT ATT ATT ATT ATT ATT ATT ATT ATT ATT ATT ATT ATT ATT ATT ATT ATT ATT ATT ATT ATT ATT ATT ATT ATT ATT ATT ATT ATT AT

1253 GCT GTC GCC ATC GCT GGT ATT GST GAT GCT CCT GTC TGG TTT TTAGGCGTTATCA ATG GCC ATG TCC CAG CGG
   Gly Ala Ile Gin Leu Gin Leu Gin Leu Ser Thr Met Thr Gin Gin

1332 AAT TAC TCA GCT GCT TTG GCT GCC ATT GCT GCC CAA CAC CAC GGA AAT CAG CAC CAG AAG GCA AAT ATT AAC AAC
   Asp Thr Asn Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Thr Cys Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gins

1470 GCT GCC GTG ATT GAT TAT AAC CAG GAC CAA CCG CCA ACG CTA AAC GCC TAC CAG GCT CAT GCC AAC ACA
   Ala Gly Val Ala Asp Thr Tyr Asn Gin Glu Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Leu Leu Ala Phe Gin Gin Val His Ala Asn

BotEl

1566 TGG GCT ACT GAT GCC CGG TCT
   Trp Ala Thr Gin Asp Pro Ser
shown in figure 4.6. Both alleles contained C to T transitions, at position 1019 in H1035 and at position 1073 in H1375.

The gene immediately 5' from gene 7 on the P22 chromosome is gene \textit{I4}. As further confirmation of this assignment, 566 bp upstream of the \textit{AvrI} site were sequenced. This sequence contained a 151 amino acid orf, from position 153 to 605 in figure 4.6. The predicted molecular weight of this orf was 17240, similar to the 15500 molecular weight of gp14 as determined by SDS-polyacrylamide gel electrophoresis. To confirm that this open reading frame encoded gp14, this region was sequenced from two independent P22 gene \textit{I4} amber alleles, H600 and H611. Both alleles carried the same mutation, a C to T transition at position 445 which changed the Gln codon to a UAG termination codon. These results confirmed the marker rescue results and showed that the same region of P22 chromosome which contained the \textit{opdA} independent mutations, also contained gene 7.

Sequence of \textit{opdA} independent mutations. The gene 7 region of the three \textit{opdA} independent mutants was sequenced (Fig. 4.7). Mutant 1a had a C to A transversion at position 627 and an A to G transition at position 651 which resulted in Thr6 to Lys and Glu14 to Gly amino acid changes. Mutant 4a had G to A transitions at both positions 653 and 668 resulting in Glu15 and Glu20 to Lys substitutions. Mutant 3a was an in frame deletion of 36 bases, from 650 to 688 which resulted in a deletion of the 13 amino acids from Glu14 to Gly26. Mutant 3a also had a silent T to C transition at position 826.

Proteolytic processing of gp7. Eppler \textit{et al.} (1991) have obtained N-terminal sequence of the gp7 protein isolated from P22 particles. This sequence, ?GG?GGADKSA, was not found at the beginning of the gene 7 orf, but was found beginning at residue 21, KGGKGGADKSA. This suggested that the N-terminal 20 amino
FIGURE 4.7. Amino terminal region of gene 7 from mutant phage. The nucleotide sequence, 599 to 703, and predicted amino acid sequence of wild type P22 is presented. The nucleotide changes in the mutant phage are indicated above it with the corresponding amino acid changes below. The solid line indicates the extent of the deletion in P22 3a. The amino terminus of mature gp7, as determined by Eppler et al. (1991) is shown in italics.
(3a)  
(4a)  
(1a)  
(WT)  CGGGAGATAAAC ATG TTA CAT GCA TTT ACG CTG GGC AGG AAA CTG CGC GGT GAG GAA CCT  
(WT)  Met Leu His Ala Phe Thr Leu Gly Arg Lys Leu Arg Gly Glu Glu Pro Gly Lys  
(1a)  
(4a)  
(3a)  

(3a)  
(4a)  
(1a)  
(WT)  TCT TAT CCT GAA AAA GGC GGT AAA GGT GGC GCA GAT AAA AGC GCA  
(WT)  Ser Tyr Pro Glu Lys Gly Gly Lys Gly Gly Ala Asp Lys Ser Ala  
(1a)  
(4a)  Lys  
(3a)  

---

A
G
Lys
acids are removed before gp7 is incorporated into phage particles. To determine if 
oligopeptidase A was required for this processing, P22 H5 13am particles were prepared 
from opdA⁺ and opdA⁻ hosts. Phage proteins were separated by SDS-PAGE and 
visualized by Coomasie blue staining. The results, shown in figure 4.8, clearly show that 
gp7 from the opdA⁻ particles is about 2kDa larger than that from the opdA⁺ particles. 
This is the predicted size difference between the orf beginning at position 611, and that 
beginning at 671, the N-terminus as determined by Eppler et al. These results 
demonstrate that gp7 is processed and oligopeptidase A is required for processing.

The processing of the opdA-independent phage mutants was studied in a similar 
way. Figure 4.8 also shows these results. None of the mutant proteins is cleaved in the 
opdA⁻ host. One of the mutants (1a) was processed in the opdA⁺ infection, two (mutants 
3a and 4a) were not. Protein gp7 from mutant 3a particles is smaller than the 
unprocessed wild type gp7 and larger than the processed form. This is consistent with the 
nucleotide sequence data for mutant 3a which predicts a protein that is 13 amino acids 
shorter than the wild type. None of the mutant gp7 proteins was processed in the opdA⁻ 
host. These results show that all three mutations lead to proteins that function more 
effectively without proteolytic cleavage than the uncleaved wild type protein. They do not 
alter the gene product in such a way that it can be cleaved by some other cellular 
protease.

Expression of gp7. To study the function and processing of gp7 in the absence 
of other phage proteins, plasmids were constructed which allowed the expression of 
parental (P22 H5) and mutant 4a (pCM172) gene 7 under control of the trc promoter. 
The marker rescue experiments had shown that the mutants were able to confer their 
opdA-independent plaque phenotype when they were recombined into the phage genome.
FIGURE 4.8. SDS-polyacrylamide gel electrophoresis of proteins from P22 particles grown on opdA$^+$ and opdA$^-$ hosts. Phage were grown on TN1379 (opdA$^+$) or TN3101 (opdA$^-$) and purified by CsCl$_2$ gradient centrifugation. Proteins were separated on a 12% polyacrylamide slab gel using the buffer system of Laemmli (1970) and visualized by staining with Coomassie Blue R250. T lane is trypsinogen (mw 24,000), S lane contains molecular weight markers. WT is P22 H5 13am, 1a is P22 1a, 3a is P22 3a, 4a is P22 4a. (+) indicates phage were grown in TN1379 (opdA$^+$), (-) that they were grown in TN3101 (opdA$^-$).
To determine if the mutant gene 7 could confer its opdA-independent phenotype when expressed from a plasmid and not the phage itself, pCM172 was introduced into TN3101 (opdA10::Mu dl). P22 H5 made large plaques on a lawn of TN3101/pCM172 when expression of the mutant gene 7 was induced by IPTG, but small plaques when expression was not induced. This lead to the following conclusions: first, it showed that the mutant gene 7 could be expressed in Salmonella. Second, it showed that the mutant gene 7 did not have to be expressed from the phage genome to confer its opdA-independent phenotype. Third, it showed that the opdA-independent mutation was dominant over the wild-type. To further confirm that expression of the mutant gp7 was responsible for the opdA-independent phenotype, transposon Tn1000 insertions in pCM172 were isolated. Five insertions which eliminated the plasmids ability to confer the opdA-independent phenotype were mapped on the plasmid. One was in the vector, between the promoter and the insert, the other four were located within the Aval/BstEI fragment (Fig. 4.5).

Figure 4.9 shows the expression of wild type gp7 from plasmid pCM187 in opdA− and opdA+ cells. Only the mature forms are seen, suggesting that the initial processing is rapid, half-life less than 5 minutes, and does not require any other phage proteins. Furthermore, the gp7 protein is rapidly degraded in both opdA− and opdA+ cells. The half-life of gp7 in opdA+ and opdA− cells was about 3 min. and 10 min. respectively.

To study the processing step directly, wild type gene 7 was cloned behind the phage T7 promoter (pCM193), and expressed in the S. typhimurium strain TN3685. TN3685 contains the phage T7 RNA polymerase gene under inducible control in its chromosome. After inducing the expression of the T7 polymerase and treating the cells with rifampicin to inhibit the S. typhimurium RNA polymerase, gp7 was labelled with a one min. pulse of 14C-amino acids, then chased with an excess of casamino acids. The
FIGURE 4.9. Expression and degradation of wild type gp7 from pCM187 in opdA\(^+\) (TN1379) and opdA\(^-\) (TN3101) hosts. gp7 expression was induced by adding IPTG. Proteins were labelled with a 5 minute pulse of \(^{14}\)C-amino acids, and chased with unlabelled casamino acids. Samples were removed at 0, 2, 5, 10, 20, and 40 minutes after the pulse. Cells were pelleted, then suspended in SDS sample buffer, heated to 95°C for 5 minutes. Proteins were separated on a 12% polyacrylamide slab gel using the Laemmli (1970) buffer system and detected by autoradiography.
autoradiogram in figure 4.10 shows that the rate of processing is very rapid with a half-life of less than 2 min.

The proteolytic instability of gp7 was unexpected and intriguing. Preliminary studies of gp7 degradation were performed to determine 1) whether gp7 degradation was dependent on lon protease, and 2) whether the mutation in mutant 4a stabilized gp7. The stability of many abnormal and foreign proteins is increased in lon" strains. To determine if the degradation of gp7 was dependent on lon protease, pCM187 was moved into TN2684 (lon" S. typhimurium). Protein gp7 was pulse labelled with 14C-amino acids and chased with casamino acids. Figure 4.11 shows that gp7 was rapidly degraded in the lon" strain, half life ~5 min., demonstrating that gp7 degradation was not lon dependent. To study the effect of the gene 7 mutations in mutant 4a on stability of gp7, the mutant 4a gp7 protein was expressed from the trc promoter in pCM185. Although mutant 4a gp7 is not processed, it is degraded as rapidly as the processed wild type gp7, with a half life of ~3 min. in both opdA+ and opdA- backgrounds (Fig. 4.12).
FIGURE 4.10. Processing of gp7. TN3685, contains an IPTG inducible chromosomal T7 RNA polymerase gene and wild type gene 7 under T7 polymerase control on pCM193. In order to observe processing directly, the T7 RNA polymerase was induced with IPTG, rifampicin was added to inhibit the host polymerase, then $^{14}$C-amino acids were added. After one minute, an excess of unlabelled casamino acids was added. Samples were removed to ice at 0, 1, 2, 3, 5, and 7 minutes after the pulse. Cells were pelleted, resuspended in SDS sample buffer and heated to 95°C for 5 minutes. Proteins were separated on 10% polyacrylamide Tris-tricine gels (Schägger and von Jagow, 1987) and detected by autoradiography.
FIGURE 4.11. Degradation of gp7 is lon independent. As figure 4.9, except the host strain for pCM187 was TN2684 (lon-10I).
FIGURE 4.12. Mutant gp7 from P22 4a although not processed is still rapidly degraded in an $opdA$ independent manner. As figure 4.9 except the plasmid was pCM185, which carries gene 7 from P22 4a and also a fragment of gene 20 in pSE380. The host strains were TN1379 ($opdA^+$) and TN3101 ($opdA^*$). The asterisk indicates the gp20 fragment expressed from this plasmid.
DISCUSSION

The major conclusions of these studies are the following: 1. OpdA is required for the normal growth of phage P22. 2. Phage P22 gp7 undergoes a previously unrecognized post-translational processing in which a 20 amino acid peptide is removed from the N-terminus of the protein. 3. This processing is dependent on the host metalloendopeptidase OpdA, and occurs in the absence of other phage proteins. 4. Mutations which reduce the requirement for this processing are located in the N-terminal portion of gp7. 5. Phage protein gp7 is rapidly degraded when expressed in the absence of phage infection. This degradation occurs whether gp7 is processed or not and is independent of the lon protease.

The only host mutations previously reported to affect the growth of phage P22 are those of Joshi et al. (1982). These mutations were selected as mutations in S. typhimurium which increased resistance to the antibiotic thiolutin. It was subsequently found that strains carrying these mutations failed to support the normal growth of P22 at elevated temperature (40°C). Phage grown on one class of these mutants appeared morphologically normal by electron microscopy, however they failed to inject their DNA into the host. We have not tested the thiolutin sensitivity of our opdA⁻ strains, and the thiolutin resistant mutations have not been mapped in S. typhimurium. Other thiolutin resistance mutations have been mapped in E. coli to 11 and 85 map units (Sivasubramanian and Jayaram, 1980). However, opdA maps at 76 map units on the S. typhimurium chromosome and prlC, the E. coli analogue of opdA, maps at 71 map units on the E. coli chromosome (Thrun and Silhavy, 1987). Consequently, although the relationship between the thiolutin resistance mutations and opdA is unknown, the map positions of the E. coli mutations suggest they affect different genes.
Although proteolytic processing is a common aspect of phage development, in fact of virus development in general, previous studies have reported no proteolytic processing of P22 proteins during morphogenesis (Botstein et al., 1973, Casjens and King, 1974). Whereas some of the phage lambda minor coat protein, gpC, is cleaved and joined to gpE, the major coat protein (Hendrix and Casjens, 1974a), the P22 major coat protein, gp5, undergoes no proteolytic modifications (Casjens and King, 1974). The lambda scaffolding protein, gpNu3, is degraded (Ray and Murialdo, 1975). The scaffolding protein of P22, gp8, differs from that of phage lambda in that it is not degraded but rather recycled (Casjens and King, 1974). The processing of gp7 reported here may not have been detected in earlier studies for several reasons. First, the processing is rapid and only detectable in pulse chase experiments with a short pulse and early time points. The studies by Botstein et al. (1973) used 5 minute pulses and the earliest time point was at 6 minutes. Second, gp7 is similar in size to several early phage proteins, and in pulse chase experiments of phage infected, irradiated host cells these early phage proteins might have masked the processing of gp7. Finally, both the low level of expression of gp7 during infection relative to the major late proteins, gp5 and gp8, and the small size of gp7 (22kD) might have made detection of gp7 processing difficult.

The proteolytic processing of gp7 is both independent of other phage proteins and dependent on the host protein oligopeptidase A. The simplest explanation for this is that the metalloendopeptidase, oligopeptidase A, is directly responsible for the proteolytic processing of gp7 by cleaving the bond between Glu20 and Lys21, two charged residues. However, limited in vitro specificity studies, mainly with the lipoprotein signal peptide as substrate, have indicated that oligopeptidase A prefers an Ala or Gly on either side of the scissile bond (Novak and Dev, 1989; Vimr et al., 1983). It should be noted that as there
is no Glu-Lys site or even two adjacent charged residues in the lipoprotein signal peptide, studies with this substrate provide no information about the specificity of oligopeptidase A for such a site. Even if oligopeptidase A fails to attack the Glu-Lys bond *in vitro*, it is possible that this does not accurately reflect its substrate specificity *in vivo*.

It has also been reported that oligopeptidase A will not attack the lipoprotein signal peptide while it is attached to the precursor protein (Novak and Dev, 1989). This suggests that oligopeptidase A prefers small polypeptide substrates, and that it would be unlikely to attack the unprocessed gp7. Once again however, these are *in vitro* studies, and the *in vivo* situation may be different. Alternatively, if oligopeptidase A is not directly responsible for the removal of the N-terminal peptide, it may be indirectly necessary for gp7 processing, possibly by activating another protease, or by interacting with the unprocessed gp7 to keep it susceptible to processing.

The phage proteins gp7, gp16, and gp20, are required for ejection of the phage DNA from the phage particle (gp16) and injection of that DNA into the host (Poteete and King, 1977; Botstein *et al.*, 1973; Hoffman and Levine, 1975; Israel, 1977). Although these proteins are incorporated into proheads early in development, before the DNA is packaged, phage particles lacking any one of these proteins are morphologically indistinguishable from wild type particles (Poteete and King, 1977; Botstein *et al.*, 1973). Inactivation studies using DNA intercalating agents have shown that in the mature phage particle, gp7, gp16, and gp20 are in close association with the phage DNA (Bryant and King, 1984). They are all ejected from the phage head during DNA ejection (Israel, 1977), but whether they remain associated with the DNA, and what their subsequent fates are, is unknown. It has been shown that the ejection of gp7 is dependent upon gp16. The ejection of gp16 however, is not dependent on either gp7 or gp20 (Israel, 1977).
Furthermore, a 16+ phage can complement a 16- phage if both infect the same cell at the same time (Hoffman and Levine, 1975). These results suggest that gp16 is released from the phage head first after adsorption and subsequently acts to facilitate the release of gp7 and the phage DNA.

Although no significant sequence similarity was found between gene 7 and any phage lambda genes, or any other genes in Genbank, it is interesting to note that the phage lambda tail protein gpH, which has been implicated in phage DNA injection, is also proteolytically processed (Hendrix and Casjens, 1974). This suggests that processing may be a common occurrence among proteins involved in phage DNA injection. It has been suggested that a possible function of gpH processing might be to provide energy for DNA ejection by forcing gpH to assume a stressed conformation (Hendrix and Casjens, 1974). However, gpH and gp7 are very different proteins. In addition to the lack of primary sequence similarity, gpH is almost 4 times larger than gp7 (90kD and 24kD respectively). Furthermore, mature gpH is 11kD smaller than its precursor, and mature gp7 is only 2kD smaller than its precursor (Hendrix and Casjens, 1974).

For phage DNA to enter a Gram negative bacteria such as *Salmonella*, the negatively charged, very large (28.6 MDa) DNA molecule must traverse two phospholipid bilayers and the periplasmic space. How this is done is unknown. Phage P22 adsorbs to the lipopolysaccharide (LPS) by its tail spike. The tail spike protein is an active endorhamnosidase and digests the polysaccharide side chains of the LPS (Iwashita and Kanegusaki, 1973). Presumably this allows access to an unidentified secondary receptor in the outer membrane. It is possible that by interacting with this outer membrane receptor, the phage gains access to the periplasm. Subsequently, the DNA must be released and pass through the periplasm and the inner membrane. Therefore, possible
roles for injection proteins include releasing the DNA from the phage particle, directing it across the periplasm, protecting it from degradation, and getting it through the inner membrane. It is not known if P22 pirates a cellular transport system to gain entry to the cytoplasm, or if the injection proteins themselves are sufficient to allow the phage DNA to enter the cytoplasm.

There are several elements of the gene 7 sequence which might provide clues to the function of this protein. The mature N-terminus has a large number of basic residues, 5 lysines in 19 amino acids, net charge of +3 (predicted pI 9.5). When projected onto a helical wheel (Schiffer and Edmondson, 1967), the basic residues concentrate on the same face of the helix. The opdA independent mutations all make the unprocessed N-terminus more basic than the wild type unprocessed N-terminus. Wild type unprocessed N-terminus has a predicted pI of 7.01, whereas the predicted pI’s of the first twenty amino acids in the unprocessed N-termini of the mutant phage 1a, 3a, 4a, are 9.61, 10.36, and 10.26 respectively. This suggests that a basic N-terminus may be important to gp7 function.

The C-terminal half of gp7, on the other hand, is rather hydrophobic (Fig. 4.13). There are two especially hydrophobic regions at the very C-terminus, 189 to 200, and 212 to 229, which are predicted, by the method of Eisenberg et al. (1984), to be potential membrane spanning regions. These regions are delimited by charged residues.

These findings suggest two possible functions for gp7. One, the basic N-terminus of gp7 binds to the phage DNA and the hydrophobic C-terminus inserts into the inner membrane of the host cell, perhaps directing the DNA to its entry site. Two, the function of the positively charged N-terminus may not be to bind DNA, but to direct the protein to the negatively charged phospholipid head groups of the membrane. Subsequently, the
FIGURE 4.13. Hydrophobicity and hydrophobic moment of processed wild type gp7. The top panel show the hydrophobicity of gp7 as predicted by the method of Eisenberg et al. (1984). The bottom panel shows the hydrophobic moment plot. The middle lines show the structural predictions: SUR, surface seeking protein; GLO, globular protein; TM, transmembrane protein.
Hydropobicity and Hydrophobic Moment of Processed gp7

[Graph showing data with axes labeled phi and pho and markers for SUR, GLO, and TM]
hydrophobic C-terminus would insert itself into the membrane. This would be similar to
the insertion of bacteriophage M13 procoat protein into the cytoplasmic membrane
(Galluser and Kuhn, 1990).

The unprocessed N-terminus contains 6 charged residues in 20 amino acids.
However there is no net charge (predicted pI 7.01). A possible function for this
hydrophilic N-terminal peptide would be to keep the relatively hydrophobic gp7 protein
soluble and in an assembly competent conformation. It might also prevent the basic
mature N-terminus from prematurely binding DNA, or directing the protein to the
membrane. Removal of the N-terminal peptide in an oprBA dependent manner, would
consequently allow the protein to fold into its active conformation, and expose the mature
basic N-terminus. Obviously, further experiments are required to elucidate the function
and mechanism of action of gp7.
Appendix

The Effect of *Salmonella ion* Mutations on Protein Degradation in Multiply Peptidase Deficient Strains
INTRODUCTION

Salmonella typhimurium and Escherichia coli contain a large number of enzymes which hydrolyze peptide bonds. One approach to characterizing these activities and to elucidating their roles in cellular metabolism is to study the effect of mutations which eliminate these activities. This approach has been used successfully to characterize the many peptidase activities present in *S. typhimurium* and *E. coli* (Miller, 1987). These peptidases have been shown to function late in the process of protein degradation, namely they are required to convert small peptides to free amino acids (Yen *et al.*, 1980; Miller and Green, 1981).

A recurring theme in these studies has been the apparent overdetermination of peptidase activities. The initial goal of my project was to extend this genetic approach to the study of enzymes which would attack larger protein substrates and so presumably act earlier in the protein degradation pathway. One well characterized protease, known to be required for the degradation of a variety of unstable proteins, is the Lon protease (for recent review see Gottesman, 1989). Lon protease has been shown to play a role in the degradation of abnormal proteins, such as the incomplete protein fragments made in the presence of puromycin, or proteins which have incorporated the amino acid canavanine, or protein fragments such as LacZΔ90, which are the result of nonsense mutations. Lon protease also function in the degradation of unstable regulatory proteins such as the *sulA* gene product, the *rcsA* gene product, a regulator of capsular polysaccharide production, and the phage lambda N protein. Although *lon* mutations have been known in *E. coli* for many years, it was only recently that they were isolated in *Salmonella* (Downs *et al.*, 1986). As part of my original project to gain genetic access to new endoproteases, I decided to first get Lon protease under control. This section describes the isolation of *lon*
mutations in *S. typhimurium* using the procedure of Downs et al. (1986) and the
colorization of these mutations, particularly, the precise genetic mapping of the *lon*
gene in *Salmonella*, and the effect of the *lon* mutation on protein degradation in a multiply
peptidase deficient strain.
MATERIALS AND METHODS

Bacterial strains and growth conditions. All strains were derivatives of *S. typhimurium* LT2 and are shown in Table A.1. Bacteria were routinely grown in LB medium nutrient broth with NaCl (Downs et al. 1986). Minimal medium was the E medium of Vogel and Bonner (1956) supplemented with glucose at 4g/liter. Amino acids were added at 0.4mM when necessary. Tetracycline and nitrofurantoin were included at 25μg/ml and 3μg/ml respectively.

Genetic techniques. Phage P22 HT12/4 int-3 was used for transducing genetic markers between bacterial strains as described by Roth (1970). Because very little thiamine is required for growth of Thr⁻ strains, transduction of thiI15 to Thr⁺ was performed by first selecting growth on minimal medium then patching individual colonies to another minimal plate.

N-acetyl Leu α-naphthyl ester (NALNE) stain for *apeB*. To detect the presence of *apeB* mutants, colonies were stained with NALNE as described by Heiman and Miller (1978). To 1.5 ml LB soft agar was added 0.4 ml NALNE (7.5mg/ml in DMSO). This solution was poured over colonies on an LB agar plate and allowed to incubate at room temperature for 45 min. Over this was poured 10 ml of Fast Garnet GBC solution (1mg/ml in 0.1 M Tris-HCl, pH 7.5). After 15-30 min. at room temperature, Ape⁺ colonies became dark purple, while Ape⁻ colonies were only slightly purple.

Isolation of *lon* mutants. *S. typhimurium* *lon* mutants were isolated as described by Downs et al. (1986). Briefly, an overnight culture of TN1379 was diluted 1:100 into nutrient broth containing 75μg/ml chlorpromazine. Fifteen separate cultures were incubated at 37°C. After 72 hours incubation, 8 of the 15 cultures were turbid. These were diluted, and plated for isolated colonies on LB agar. After overnight incubation at
<table>
<thead>
<tr>
<th>Strain</th>
<th>Genotype</th>
</tr>
</thead>
<tbody>
<tr>
<td>LT2</td>
<td>Wild type</td>
</tr>
<tr>
<td>TN677</td>
<td>( \text{apeB21 ~ apeA18 ~ argF88} )</td>
</tr>
<tr>
<td>TN698</td>
<td>( \text{thiI15 ~ galE} )</td>
</tr>
<tr>
<td>TN1246</td>
<td>( \text{leuBCD485 ~ pepB11 ~ pepN90 ~ pepA16 ~ supQ302 ~ pepP1 ~ pepQ1} )</td>
</tr>
<tr>
<td>TN1379</td>
<td>( \text{leuBCD485} )</td>
</tr>
<tr>
<td>TN2634</td>
<td>( \text{recA' ~ srl-::Tn10 ~ leuBCD485 ~ pepB11 ~ pepN90 ~ pepA16 ~ supQ302 ~ pepP1 ~ pepQ1} )</td>
</tr>
<tr>
<td>TN2681</td>
<td>( \text{leuBCD485 ~ lon-101} )</td>
</tr>
<tr>
<td>TN2682</td>
<td>( \text{leuBCD485 ~ zba-883::Tn10} )</td>
</tr>
<tr>
<td>TN2684</td>
<td>( \text{lon-101 ~ zba-883::Tn10} )</td>
</tr>
<tr>
<td>TN2686</td>
<td>( \text{leuBCD485 ~ lon-102} )</td>
</tr>
<tr>
<td>TN2688</td>
<td>( \text{lon-102 ~ zba-883::Tn10} )</td>
</tr>
<tr>
<td>TN2696</td>
<td>( \text{lon-101 ~ zba-883::Tn10 ~ leuBCD485 ~ pepB11 ~ pepN90 ~ pepA16 ~ supQ302 ~ pepP1 ~ pepQ1} )</td>
</tr>
</tbody>
</table>
37°C, 2 of the 8 cultures contained mucoid colonies. These mucoid colonies were purified and tested further.

Testing for UV sensitivity. Cultures were screened for sensitivity to UV light by streaking an overnight culture of the strain to be tested onto LB agar. This streak was exposed to increasing amounts of UV radiation from a germicidal lamp; 0, 100, and 200 ergs mm⁻². The plates were then covered with aluminum foil and incubated at 37°C overnight. UV irradiation at 200 ergs mm⁻² had little to no effect on growth of LT2, the lon strains were moderately inhibited by 200 ergs mm⁻², and TN2634 (recA) was completely inhibited by 100 ergs mm⁻².

UV killing curves were determined for LT2, TN2634 (recA), and lon strains. Overnight cultures were diluted to ~1000 cfu/ml, spread on LB agar, and exposed to UV radiation at the following levels: LT2, 0, 100, 200, 300 ergs mm⁻²; lon strains, 0, 50, 80, 100 ergs mm⁻²; TN2634, 0, 6, 12, 18 ergs mm⁻². After overnight incubation in the dark, colonies were counted.

Puromycin fragment degradation. The procedure was essentially that of Yen et al. (1980). Overnight cultures in minimal medium were diluted 1:100 in minimal medium, and incubated at 37°C until OD₆₀₀ reached 0.15. Puromycin was added to 100μg/ml and incubation continued for 10 min. ¹⁴C-Leu (330mCi/mmol) was added to 0.1μCi/ml, and cultures were incubated 5 minutes. The cells were pelleted by centrifugation, washed 3 times with cold Leu (2mM), resuspended in minimal glucose medium with Leu, and incubated at 37°C. 0.9ml samples were withdrawn at various times, and made 5% in trichloroacetic acid (TCA). 0.4ml of this was added to liquid scintillant to determine the total amount of radioactivity in the sample. The rest was held on ice for at least 1 hour, after which precipitated protein was pelleted in a
microcentrifuge at 12,000 x g, 4°C for 30 min. The radioactivity in 0.4ml of the supernatant was counted in liquid scintillant to determine acid soluble counts.

**Rapid accumulation of acid soluble peptides.** The procedure was essentially that of Yen et al. (1980a). Overnight cultures were diluted 1:50 in minimal medium with proline and grown at 37°C to OD$_{600}$ 0.2. $^{14}$C-Leu (330mCi/mmol) was added to 0.1 μCi/ml. 0.9ml samples were removed at 1, 3, 5, 15, 30, and 45 minutes after the pulse of radioactivity. These were made 5% in TCA and held on ice for at least 1 hour. To determine total number of counts, 0.4 ml of this was removed. The rest was centrifuged at 12,000 x g, 4°C, for 30 min. to pellet precipitated protein. To determine acid soluble counts, the amount of radioactivity in 0.4 ml of supernatant was measured.
RESULTS AND DISCUSSION

Isolation of *S. typhimurium* lon mutants. Two independent *lon* mutations were isolated in TN1379 using the procedure of Downs *et al.* (1986). These were given the allele numbers *lon-101* and *lon-102*. These mutations conferred the following *lon* phenotypes: excess capsular polysaccharide production (Trisler and Gottesman, 1984), sensitivity to UV radiation which is intermediate between wild type and *recA* (Howard-Flanders *et al.*., 1963), and reduced ability to degrade puromycin generated protein fragments to acid soluble peptides (Fig. A.1) (Gottesman and Zipser, 1978).

Mapping *lon* on the *S. typhimurium* chromosome. The *lon* gene maps at 10 map units on the *E. coli* chromosome (Bachman and Low, 1980). In order to map my *S. typhimurium* lon mutations, they were checked for linkage to *zba-883::Tn10*, a transposon which is 48% cotransducible with *apeB*. Heiman and Miller (1978) have mapped *apeB* to between *thiI* and *apt*. Transposon *zba-883::Tn10* was 35% cotransducible with *lon-101* and 33% cotransducible with *lon-102*. The cotransduction frequencies of nearby markers were as follows: *thii15* and *apeB2I*, 11%; *thii15* and *zba-883::Tn10*, 6%; *thii15* and *lon-101*, 1%. As shown in figure A.2, the two point crosses place *zba-883::Tn10* between *thiI* and *lon*. The three point cross described in figure A.3 was performed in order to confirm the order of *zba-883::Tn10*, *apeB*, and *lon*. First, consistent with the results of the 2 point cross, *zba-883::Tn10* is not between *apeB* and *lon*. The linkage of the Tn10 to *lon* increases in Tet* ApeB+* transductants compared to Tet* ApeB* transductants. If the Tn10 had been between *apeB* and *lon*, the linkage of *lon* to the Tn10 in the Tet*, ApeB+* transductants would decrease. The fact that the rare class in the 3 point cross is Tet*, ApeB*, Lon*, proves that the order is Tn10, *apeB*, *lon* (order 1, Fig. A.3). This is the order shown in figure A.2.
FIGURE A.1. Combined effect of lon and peptidase deficiency on puromycyl fragment degradation. Strains TN2684 (lon<sup>-</sup>, Pep<sup>+</sup>), TN1246 (lon<sup>+</sup>, Pep<sup>-</sup>), and TN2696 (lon<sup>-</sup>, Pep<sup>-</sup>) were transduced to Leu<sup>+</sup>. Overnight cultures of these strains and LT2 (lon<sup>+</sup>, Pep<sup>+</sup>) were diluted into minimal glucose medium and incubated to OD<sub>600</sub> of 0.15. Puromycin was added to 100 µg/ml. After 10 min. at 37°C, ¹⁴C-Leu was added to 0.1 µCi/ml. After 5 min. more, cells were pelleted, washed 3 times in cold unlabelled Leu (2mM), and resuspended in minimal glucose medium with Leu at 37°C. At 0, 20, 45, 60, 90, 120, 150, and 180 minutes after pulse, 0.9 ml samples were removed and made 5% in TCA. 0.4 ml of this was counted to determine total radioactivity, the rest was incubated on ice for >1h, then pelleted at 15,000 × g, for 30 min. The radioactive material in 0.4 ml of the supernatant was determined.
FIGURE A.2. Linkage map of *S. typhimurium* *lon* mutations. All transduction were performed with the general transducing phage P22 HT12/4 *int-103*. Arrows point toward the unselected marker. Cotransduction frequencies were converted to physical distance using the formula in Sanderson and Roth (1988). *apeB* was scored using NALNE staining, *lon* was scored using nitrofurantoin sensitivity (Downs *et al.*, 1986).
FIGURE A.3. Three point cross to determine the relative order of lon,
zba-883::Tn10, and apeB. A P22 transducing lysate on TN2682 (Lon\(^-\), ApeB\(^+\), Tet\(^f\))
was used to transduce TN677 (Lon\(^+\), ApeB\(^-\), Tet\(^f\)) to Tet\(^f\). ApeB was scored by NALNE
staining, and Lon by nitrofurantoin sensitivity. The rare class for order 1 is: Tet\(^f\), ApeB\(^-\),
Lon\(^-\). The rare class for order 2 is: Tet\(^f\), ApeB\(^+\), Lon\(^-\). The rarest class among the
transductants was Tet\(^f\), ApeB\(^-\), Lon\(^-\), establishing order 1 as the correct order.
<table>
<thead>
<tr>
<th>Donor</th>
<th>Recipient</th>
<th>Selected marker</th>
<th>Number</th>
<th>ApeB</th>
<th>Lon</th>
<th>Number</th>
<th>Per cent</th>
</tr>
</thead>
<tbody>
<tr>
<td>TN2682</td>
<td>TN677</td>
<td>Tet-R</td>
<td>37</td>
<td>+</td>
<td>-</td>
<td>10</td>
<td>27</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>4</td>
<td>11</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0</td>
<td>&lt;3</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>23</td>
<td>62</td>
</tr>
</tbody>
</table>
Combined effect of *lon* and *pepN, pepA, pepB, pepD* on puromycin fragment degradation. Strains containing mutations at the *lon* locus have a reduced ability to degrade incomplete proteins containing puromycin. Strains which are deficient in peptidases N, A, B, and D also are deficient in degrading puromycin protein fragments (Yen *et al.*, 1980). As shown in figure A.1, there is little further reduction in puromycin fragment degradation when *lon* and peptidase deficiency are combined. This suggests that Lon protease and the peptidases are on the same pathway of puromycin fragment degradation because if they were on different pathways, one would expect their effects to be additive. This suggests that Lon protease acts early in the pathway, attacking larger acid insoluble fragments which are then degraded to acid soluble fragments by the peptidases.

Combined effect of *lon* and peptidase deficiency on the rapid accumulation of acid soluble peptides. Wild type Pep*+* strains rapidly and completely incorporate a pulse of labeled amino acid into large peptides and proteins which are and remain acid precipitable (Fig. A.4, LT2). However, Yen *et al.* (1980a) observed that in a multiply peptidase deficient strain, such as TN1246 in figure A.4, about 20% of a pulse of labeled amino acid remains acid soluble. Sampling over short time courses has shown that the label appears to be initially incorporated into acid precipitable material, but it appears that material is rapidly degraded to acid soluble material which is stable in the peptidase deficient strain. The interpretation given to this result is that at least 20% of the peptide bonds formed are rapidly degraded and the amino acids reincorporated into stable protein. This is a striking finding and a major cellular activity. To determine if Lon protease had a role in this rapid recycling of amino acids, a multiply peptidase deficient *lon* strain was tested. It can be seen from figure A.4, that the lack of Lon protease has little effect on
FIGURE A.4. Effect of lon⁻ on the rapid accumulation of acid soluble peptides.

Overnight cultures of LT2 and Leu⁺ derivatives of TN1246 (Pep⁻, lon⁺), TN2696 (Pep⁻, lon⁻) and TN2684 (Pep⁺, lon⁻) were diluted into minimal glucose medium and incubated at 37°C until OD₆₀₀ of 0.2. ¹⁴C-Leu was added to 0.1 μCi/ml and 0.9 ml samples removed at 1, 2, 4, 5, 10, 20, 30, and 45 minutes after the pulse. Samples were made 5% in TCA and incubated on ice at least 1 hour. 0.4 ml was removed to determine total radioactivity, the rest was pelleted and 0.4ml of the supernatant was removed to determine acid soluble radioactivity.
the accumulation of acid soluble material. Apparently, Lon protease does not play as major a role in the degradation which leads to the acid soluble material, as it does in puromycyl fragment degradation. If it did, one would expect to see less material become acid soluble in the lon\(^-\) peptidase deficient strain than in the lon\(^+\) peptidase deficient strain. These results suggest that Lon protease is not crucial to this pathway which breaks down 20% of the peptide bonds formed in the cell. Of course, Lon protease could still play a role in this degradation pathway, but it would not be a unique one.

In conclusion, this section has described the isolation and verification of the lon-101 and lon-102 alleles. It describes the precise mapping of these alleles with respect to the known markers thiI and apeB. It also suggests that Lon protease and peptidases N, A, B, and D are on the same pathway for degrading puromycyl fragments, but that Lon protease is not crucial to the degradation which leads to the rapid turnover seen as the accumulation of acid soluble material in a multiply peptidase deficient strain.


