THE UNIVERSITY OF CALGARY

Analysis of Intervening Sequences in the 23S Ribosomal RNA Genes of *Salmonella* species

by

Kanti Pabbaraju

A THESIS SUBMITTED TO THE FACULTY OF GRADUATE STUDIES IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF MASTER OF SCIENCE

DEPARTMENT OF BIOLOGICAL SCIENCES

CALGARY, ALBERTA

JANUARY, 2000

© KANTI PABBARAJU

National Library of Canada
Bibliothèque nationale du Canada

Acquisitions and Bibliographic Services
Acquisitions et services bibliographiques

395 Wellington Street
Ottawa ON K1A 0N4
Canada

395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file Votre référence

Our file Notre référence

The author has granted a non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of this thesis in microform, paper or electronic formats.

L'auteur a accordé une licence non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de cette thèse sous la forme de microfiche/film, de reproduction sur papier ou sur format électronique.

The author retains ownership of the copyright in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

L'auteur conserve la propriété du droit d'auteur qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

0-612-49650-3

Canada

ABSTRACT

Eighty six strains representing eight subgenera of the genus *Salmonella* were tested to investigate ribosomal RNA (rRNA) fragmentation which occurs as a result of intervening sequence (IVS) excision, and the pattern of transfer of IVSs in salmonellae.

Fragmentation of the 23S rRNA was restricted to conserved cleavage sites at helix-25 and helix-45 locations. Random cleavage at sites where IVSs could not be detected was not seen. Uncleaved IVSs were not detected in any case, indicating a strong selection for the removal of IVSs. Selection for maintenance of the primary stem, which is the RNase III cleavage site, was suggested by the high rate of compensatory mutations in this region. The pattern of possession and of sequence identity of IVSs was diverse and it could not be related to the multi-locus enzyme electrophoresis (MLEE) pattern of the various strains, indicating that the IVSs are frequently transferred between strains by horizontal transfer.

ACKNOWLEDGEMENTS

My gratitude to Dr. K. E. Sanderson for allowing me to borrow from his wisdom, experience and resources. His patience and support have contributed beyond words, in the completion of my thesis.

A special thanks to Dr. Michael Surette, Dr. Steven Zimmerly and Dr. Michael Hynes for all their advice and words of wisdom.

I am grateful to Dr. S. -L. Liu for his encouraging and supportive words. I would also like to thank him for helping me in obtaining the genomic DNA fragments by pulsed-field gel electrophoresis.

I would like to thank all the members of my lab: Wayne Miller for his assistance in analyzing some of the strains, Ivy Ng, Loes Pronk, Renee Brost and Sonali Majumdar for all the memories I shall cherish.

DEDICATION

To my parents who are my inspiration

To Srinivas and Nymisha for their love and support

TABLE OF CONTENTS

Approval page i

Abstract iii

Acknowledgements iv

Dedication v

Table of Contents vi

List of Tables x

List of Figures xi

Abbreviations xiii

CHAPTER 1: INTRODUCTION 1

CHAPTER 2: LITERATURE REVIEW 4

2.1 Prokaryotic ribosomal RNA processing 4

2.1.1 Organization and processing of the *rrn* operon 4

2.1.2 Enzymes involved in rRNA processing 4

2.1.3 Role of RNase III in rRNA processing 8

2.2 Intervening Sequences 9

2.2.1 Diversity of distribution and general properties 9

2.2.2 Intervening sequences in *Salmonella* 13

2.2.3 Intervening sequences and introns 13

2.2.4 Horizontal transfer of intervening sequences 15

2.2.5 Speculated functions of intervening sequences 16

2.3 Ribosomal RNA in the stationary phase 17

2.3.1 Bacterial responses to starvation in the stationary phase 17

2.3.2 Effect of starvation on ribosomes and rRNA 19

2.4 Classification of *Salmonella* 20

2.4.1 The genus *Salmonella* 20

2.4.2 Construction of SARB and SARC sets of strains 21

2.5 Research objectives 22

CHAPTER 3: MATERIALS AND METHODS 23

3.1 Bacterial strains and cultivation conditions 23

3.2 Enzymes and Chemicals 28

3.3 Primers 28

3.4 RNA isolation 29

3.5 Denaturing RNA electrophoresis 29

3.6 Northern blotting 30

3.7 Polymerase Chain Reaction and Agarose gel electrophoresis 30

3.8 Sequencing strategies 31

3.8.1 Preparation of template from strains in which only one of the seven *rrl* genes has an insertion in helix-25 or helix-45 31

3.8.2 Preparation of template from strains with more than one insertion in helix-25 and/or helix-45 32

3.8.2.1 Preparation of genomic DNA embedded in agarose and enzymatic digestion 32

3.8.2.2 Pulsed Field Gel Electrophoresis....32
3.8.3 Amplification of IVSs from individual *rrl* genes....33
3.8.4 Purification of PCR products....34
3.9 Cell lysis, RNA isolation and gel electrophoresis....34
3.10 Viable cell counts....35
3.11 Computer analysis....35
CHAPTER 4: RESULTS- Analysis of IVSs in the SARB and SARC sets....36
4.1 Diversity of distribution of IVSs in the SARB and SARC sets on the basis of PCR analysis....36
4.2 Diversity of distribution of IVSs in the SARB and SARC strains by RNA analysis ..42
4.3 Analysis of IVSs in the 72 strains from the SARB set and 16 strains from the SARC set of *Salmonella*....49
4.4 Analysis of nucleotide sequences of helix-25 IVSs....55
4.5 Analysis of nucleotide sequences of helix-45 IVSs....63
4.6 Sequence alignment studies....69
4.7 Secondary structures of IVSs....80
CHAPTER 5: RESULTS- Role of IVSs in rRNA degradation....88
5.1 Growth curves and viability counts of *Salmonella* spp. and *E. coli*....88
5.2 Degradation of rRNA during the different phases of growth....93
5.3 Stationary phase lysis of cells....97
5.4 Degradation of rRNA after extraction....100

CHAPTER 6: DISCUSSION 103

6.1 Occurrence of IVSs at the conserved helix-25 and helix-45 locations in *Salmonella* .103

6.2 Ribosomal RNA fragmentation is always a result of IVS excision 104

6.3 Selective pressure for IVS removal 105

6.3.1 IVSs are not retained in the mature 23S rRNA 105

6.3.2 Compensatory mutations in IVSs 106

6.4 Transfer of IVSs 107

6.4.1 Transfer of IVSs in salmonellae involves both vertical and horizontal transfer .107

6.4.2 Horizontal transfer of IVSs in *Enterobacteriaceae* 109

6.4.3 Entry and spread of IVSs in *Salmonella* 111

6.5 Function of IVSs and their role in the breakdown of rRNA in *Salmonella* and *E.coli* 112

6.6 Future studies 115

CHAPTER 7: LITERATURE CITED 117

LIST OF TABLES

Table 3.1: Strains of *Salmonella* from the *Salmonella* Reference Collection B (SARB)..... 24

Table 3.2: Strains of *Salmonella* from the *Salmonella* Reference Collection C (SARC)..... 27

Table 4.1: Comparison of the helix-25 IVSs of *Salmonella*.....61

Table 4.2: Comparison of helix-25 IVSs of *Salmonella, Proteus* and *Providencia*.....62

Table 4.3: Comparison of helix-45 IVSs of *Salmonella*.....65

Table 4.4: Comparison of helix-45 IVSs of *Salmonella*, *Proteus*, *Providencia* and *Yersinia*.....66

LIST OF FIGURES

Figure 2.1: Structure and cleavage sites of the rRNA primary transcript.....6

Figure 2.2: 23S rRNA structural model.....11

Figure 4.1: Diagrammatic representation of the PCR strategy.....38

Figure 4.2: A. PCR amplicons from selected SARB strains.....40

B. Interpretation of the number of *rrl* genes with IVSs in each strain.....40

Figure 4.3: Schematic representation of 23S rRNA fragmentation on excision of IVSs.45

Figure 4.4: A. 23S rRNA fragmentation from selected SARB strains.....47

B. Interpretation of rRNA fragmentation pattern.....47

Figure 4.5: Pattern of possession of IVSs for the SARB strains.....51

Figure 4.6: Pattern of possession of IVSs for the SARC strains.....53

Figure 4.7: *rrn* operons in the *Salmonella* genome.....57

Figure 4.8: Pulsed-field gel electrophoresis of I-CeuI digested genomic DNA from strains of *Salmonella* from the SARB strains.....59

Figure 4.9: Distribution of helix-25 and helix-45 IVS families on the SARB strains.....67

Figure4.10: Sequence alignments for helix-25 IVSs from individual *rrl* genes for *Salmonella*, *Proteus*, and *Providencia*.....71

Figure 4.11: Sequence alignments for helix-45 IVSs from individual *rrl* genes for *Salmonella*, *Proteus*, *Providencia*, and *Yersinia*.....75

Figure 4.12: Proposed secondary structures for the helix-25 and helix-45 tetraloops in *E. coli* and *S. typhimurium*.....82

Figure 4.13: Proposed secondary structures for representatives of helix-25 IVS families from *Salmonella* and *Proteus* 84

Figure 4.14: Proposed secondary structures for representatives of helix-45 IVS families from *Salmonella*, *Proteus* and *Yersinia* 86

Figure 5.1: Growth curves for *E. coli*, *S. agona*, *S. typhimurium* and SARC 10 89

Figure 5.2: Total viable counts for *E. coli*, *S. agona*, *S. typhimurium* and SARC 10 91

Figure 5.3: rRNA from *E. coli* and *Salmonella* over a 25 hour period 95

Figure 5.4: Lysis of stationary phase cells using various methods 98

Figure 5.5: Action of degradative enzymes on stationary phase rRNA 101

ABBREVIATIONS

BLAST	Basic Local Alignment Search Tool
bp	Base pair (s)
DMSO	Dimethylsulfoxide
dNTP	Deoxynucleotide triphosphate
DTT	Dithiothreitol
EDTA	Ethylenediamine tetraacetic acid
IVS	Intervening Sequence
L	Liter
LB	Luria-Bertani medium
mg	Milligram
ml	Milliliter
MLEE	Multi-Locus Enzyme Electrophoresis
mM	Millimolar
µl	Microliter
µg	Microgram
OD	Optical density
ORF	Open Reading Frame
PCR	Polymerase Chain Reaction
PFGE	Pulsed-Field Gel Electrophoresis

PMSF	Phenylmethylsulfonyl fluoride
PNPase	Polynucleotide phosphorylase
RNase III	Ribonuclease III
rrl	Gene for Ribosomal RNA large
rRNA	Ribosomal Ribonucleic Acid
RNP	Ribonucleoprotein
S	Svedberg unit
SDS	Sodium dodecyl sulfate
SARB	*Salmonella* Reference Collection B
SARC	*Salmonella* Reference Collection C
SGSC	Salmonella Genetic Stock Centre
SSC	Sodium chloride and Sodium citrate
TBE	Tris-HCl-Boric acid-EDTA
tRNA	Transfer RNA
V/ cm	Volts per centimeter

CHAPTER 1: INTRODUCTION

The presence of fragmented 23S ribosomal RNA (rRNA) was first discovered in *Salmonella typhimurium* (Winkler, 1979). This fragmentation was later shown to be a result of excision of transcribed sequences, called Intervening Sequences (IVSs) from the rRNA. These IVSs have been reported at base pair 550 and bp 1170 in the 23S rRNA genes of *Salmonella* species (Burgin et al., 1990) (Mattatall and Sanderson, 1996). The two sites correspond to helix-25 and helix-45 in the proposed secondary structure for the *E. coli rrl* (ribosomal RNA large) gene (Noller, 1984). The tetraloops predicted at these positions in *E. coli* are replaced at the corresponding positions in *Salmonella* by an extended stem-loop structure of the IVS. The transcribed IVSs are cleaved out by ribonuclease III (RNase III), an enzyme involved in the 23S rRNA maturation (Burgin et al., 1990).

IVSs have not been reported in *E. coli* strains, which raises the question whether IVSs entered the salmonellae after the divergence of *Salmonella* from *Escherichia* or if *Escherichia* lost its IVSs. The presence of IVSs in the stable rRNA genes, without any proven function, is surprising. The occurrence of IVSs may have implications regarding the evolution of rRNA transcription units. IVSs are distributed sporadically among the different bacterial genera. They have been identified in the *Enterobacteriaceae* (Burgin et al., 1990) (Skurnik and Toivanen, 1991) (Miller et al., 2000), *Rhizobiaceae* (Selenska-Pobell and Evguenieva-Hackenberg, 1995), *Rickettsiaceae* (Afseth et al., 1995), *Leptospiraceae* (Ralph and McClelland, 1993), and the genera *Campylobacter* (Konkel et al., 1994) (Linton et al., 1994b) (Trust et al., 1994) and *Helicobacter* (Hurtado et al.,

1997; Linton et al., 1994a). It has also been shown that the distribution of IVSs among the several copies of *rrl* genes is heterogeneous. This raises questions about how IVSs are distributed to different cells and how they are spread to different copies of the rRNA genes. IVSs with similar nucleotide sequence have been seen in *Salmonella* and *Yersinia* (Skurnik and Toivanen, 1991); *Salmonella* and *Proteus*; and *Salmonella* and *Providencia* (Miller et al., 2000) in a background of low chromosomal homology. This suggests that IVSs may be recent evolutionary additions as a result of horizontal transfer by plasmid-mediated conjugation or phage-mediated transduction.

Like introns, IVSs are cleaved out from the transcribed RNA, thus they disrupt the normal continuity of the 23S rRNA gene. IVSs have terminal inverted repeats flanking the central region, which is a characteristic of a mobile genetic element. In this regard IVSs resemble introns, however unlike introns, after the removal of IVSs the exons are not religated. In addition most of the IVSs do not have known coding open reading frames (ORFs) or terminal sequences which facilitate the splicing reactions. ORFs coding for putative proteins of 121-133 amino acids are present in an IVS from *Leptospira*, however the protein it codes for has not been detected in the cell (Ralph and McClelland, 1993). The ability of the fragmented rRNA to assemble into an active conformation, presumably by RNA-protein interactions, appears to obviate the need for exon ligation (Belfort et al., 1995).

The function of IVSs in the cell is still unclear. A RNase III-deficient strain with the IVSs intact in the rRNA is viable, but it has a reduced growth rate, presumably because RNase III is also involved in other rRNA processing reactions (Mattatall and Sanderson, 1998). It has also been shown that the transfer of a plasmid with an IVS-containing *rrl*

gene from *Salmonella* into *E. coli* leads to presence of fragmented 23S rRNA in *E. coli*, but the cell maintains wild type growth rates (Gregory et al., 1996). One function attributed to IVSs is evasion of bacteriocins that target the rRNA. Due to fragmentation of the rRNA, as a result of IVS excision, the rRNA becomes insensitive to the action of these bacteriocins (Skurnik and Toivanen, 1991). It has also been shown that rRNA in *Salmonella* is degraded as a regulatory response in the stationary phase; the presence of fragmented rRNA is postulated to confer a selective advantage by increasing the rate of rRNA degradation (Hsu et al., 1994).

IVSs are intriguing genetic elements present in the stable rRNA genes without any proven function and therefore they merit further study. It has not yet been determined if rRNA fragmentation at sites other than bp 550 and bp 1770 is seen either as a result of IVS excision or random fragmentation. It would be interesting to determine if the cells ensure the removal of all IVSs indicating that presence of IVSs in the mature rRNA is selected against. I report experiments that answer these questions and also investigate the speculated function of IVSs in rRNA degradation in the stationary phase. Speculations on the transfer and spread of IVSs in the *Enterobacteriaceae* will be made based on the experimental results.

CHAPTER 2: LITERATURE REVIEW

2.1 Prokaryotic ribosomal RNA processing.

2.1.1 Organization and processing of the *rrn* operon.

Seven copies of the *rrn* operons are present around the chromosome of *Escherichia coli* and *Salmonella typhimurium* (Lindahl and Zengel, 1986). They are all organized in a similar way: promoter, 16S rRNA, tRNA, 23S rRNA, 5S rRNA and terminator. The first intergenic region separates the 16S and 23S rRNA genes, and may contain one or two tRNAs. The second intergenic region separates the 23S and 5S rRNAs. Beyond the 5S towards the 3' end of the 30S rRNA distal tRNAs may be present (Morgan et al., 1980) (Fig.2.1). Ribosomal RNA is synthesized as a long, primary transcript that must be extensively processed to generate the mature, functional species. In *Escherichia coli*, it is known that the initial 30S precursor is synthesized and subjected to primary processing. Primary processing involves cleavage by the endonuclease RNase III to generate precursors to the 16S, 23S, and 5S. The product of the primary processing event then acts as the substrate for secondary processing. This may be an RNP particle (King et al., 1986). The overall strategy of processing and the site of action of the primary processing enzymes are shown in Fig 2.1.

2.1.2 Enzymes involved in rRNA processing.

It has been established that three enzymes are involved in the primary processing of the nascent rRNA transcript. They are RNase III, RNase P and RNase E (Gegenheimer

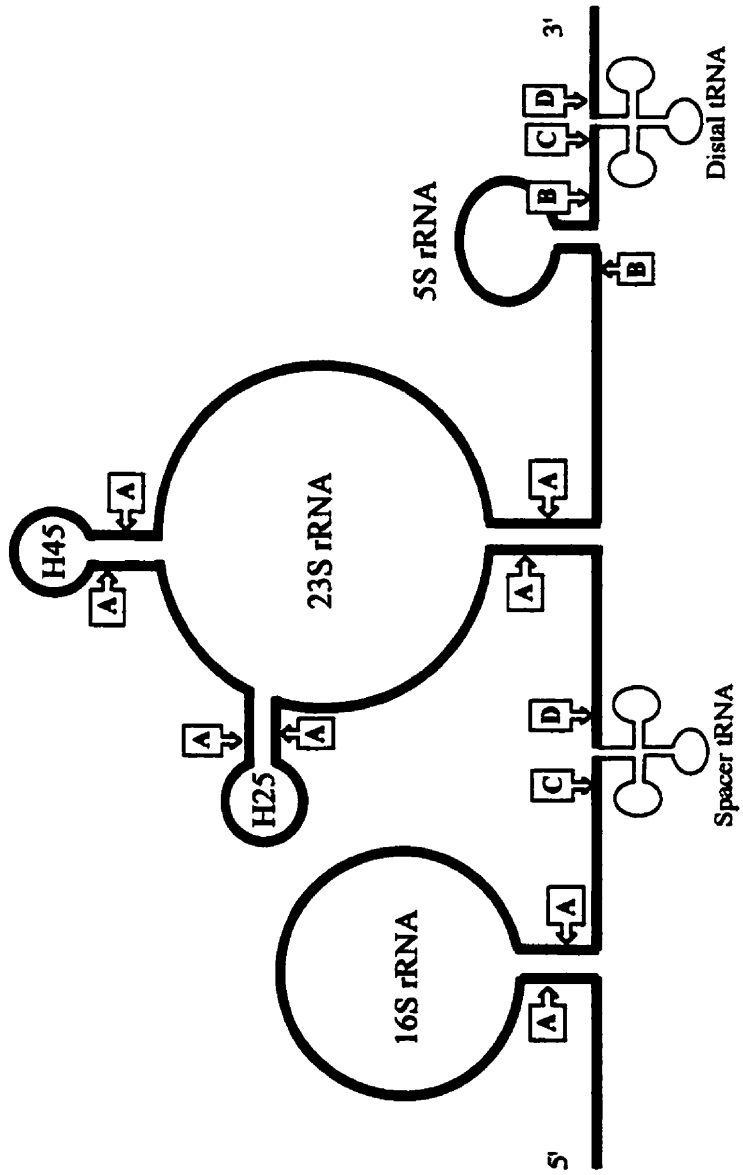
and Apirion, 1981). RNase III introduces a double cleavage in each of the two stems that produce the 16S and 23S rRNAs (Ginsburg and Steitz, 1975) (Gegenheimer and Apirion, 1980). RNase P introduces cleavages at the 5' end of all the tRNA molecules in the transcript (Altman et al., 1975), and RNase E introduces two cleavages, one before and one after the 5S rRNA. These RNA processing events occur on the nascent rRNA transcript in association with ribosomal proteins; however, the above mentioned enzymes do not require any additional proteins for efficient processing (Apirion and Miczak, 1993).

Precursor 16S (p16) and 23S (p23) rRNA are then subjected to secondary processing events, which use an RNP particle as a substrate (Hayes and Vasseur, 1976) (Meyhack et al., 1974). RNase E and CafA (proposed to be renamed as RNase G) have been shown to be responsible for the two step, sequential maturation of p16 to mature 16S rRNA (Li et al., 1999b). RNase T participates in the shortening of the 3' end of p23 along with other exoribonucleases to form the mature 23S rRNA (Li et al., 1999a).

Figure 2.1: Structure and cleavage sites of the rRNA primary transcript.

A 30S rRNA transcript transcribed from an *rrn* operon with 16S, 23S and 5S rRNAs is shown. The spacer tRNA (of which there may be one or two) and the distal tRNA (of which there may be none, one or two) are also shown. The p16 (precursor 16S) and p23 (precursor 23S) are produced as a result of the action of RNase III and are indicated by [**A**] (Dunn, 1976) (Robertson, 1982) (Suryanarayana and Burma, 1975). The 5S rRNA is processed by RNase E to release the precursor 5S rRNA [**B**] (Szeberenyi et al., 1983). Primary processing of the spacer and distal tRNAs uses RNase P at the 5' end [**C**] and an uncharacterized 3' endonuclease at [**D**] (King et al., 1986).

The positions of helix-25 IVS (indicated as H25) at bp550 and helix-45 IVS (indicated as H45) at bp 1170 are also shown as stem-loop structures on the 23S rRNA. A staggered nick in the stem of the IVS by RNase III [**A**] is responsible for their removal after transcription.


5'
16S rRNA
A
A
H25
A
A
23S rRNA
A
A
H45
A
A
C
D
Spacer tRNA
A
A
5S rRNA
B
B
C
D
Distal tRNA
3'

2.1.3 Role of RNase III in rRNA processing.

Ribonuclease III is a double strand-specific endoribonuclease coded by the *rnc* gene (Dunn, 1976) (Robertson, 1982). It functions as a dimer and each homologous polypeptide has a molecular weight of 25,000 daltons (Dunn, 1976). It catalyzes the hydrolysis of phosphodiester bonds within the processing signals. The presence of a double stranded RNA though required is not sufficient to establish RNase III selectivity (Gegenheimer and Apirion, 1981) (Robertson, 1982) (Krinkle and Wulff, 1990) (Chelladurai et al., 1991) (Li et al., 1993)

RNase III can efficiently degrade polymeric double stranded RNAs of low sequence complexity such that the involvement of base pair sequences in determining cleavage sites is unclear. An alignment of RNase III substrates reveals the exclusion of specific Watson-Crick base pair sequences at defined positions relative to the cleavage sites. Inclusion of these "disfavored" sequences in a model substrate strongly inhibits cleavage in vitro by interfering with RNase III binding (Zhang and Nicholson, 1997). Presence of 10 to 14 base pairs of double stranded RNA facilitates RNase III cleavage (Krinkle and Wulff, 1990). Reduction in the length of the double stranded region reduces the efficiency of RNase III cleavage (Chelladurai et al., 1991).

RNase III introduces a double cleavage in each of the two stems, thus producing the 16S and 23S rRNAs (Ginsburg and Steitz, 1975) (Gegenheimer and Apirion, 1980). The study of mutants deficient in RNase III production has shown that they accumulate 30S and 25S (23S and 5S) precursor rRNA (Gegenheimer et al., 1977) (King et al., 1984) (Apirion et al., 1976). Mutants deficient in RNase III have longer generation times as compared to wild type isogenic controls (Apirion et al., 1976) (Mattatall and

Sanderson, 1998). The 16S rRNA in these cells is fully mature as the machinery that deals with the precursor produced after the RNase III cleavage can recognize larger precursors and process them to the normal mature 16S rRNA (Gegenheimer et al., 1977). RNase E, another rRNA processing enzyme, can process 25S rRNA to functional 23S rRNA rendering the ribosome functional. RNase III is thus not an essential enzyme (Gegenheimer and Apirion, 1981) as the products of processing by RNase III are not the mature rRNA but rather precursors that require further processing to complete the maturation process (King et al., 1984).

2.2 Intervening Sequences.

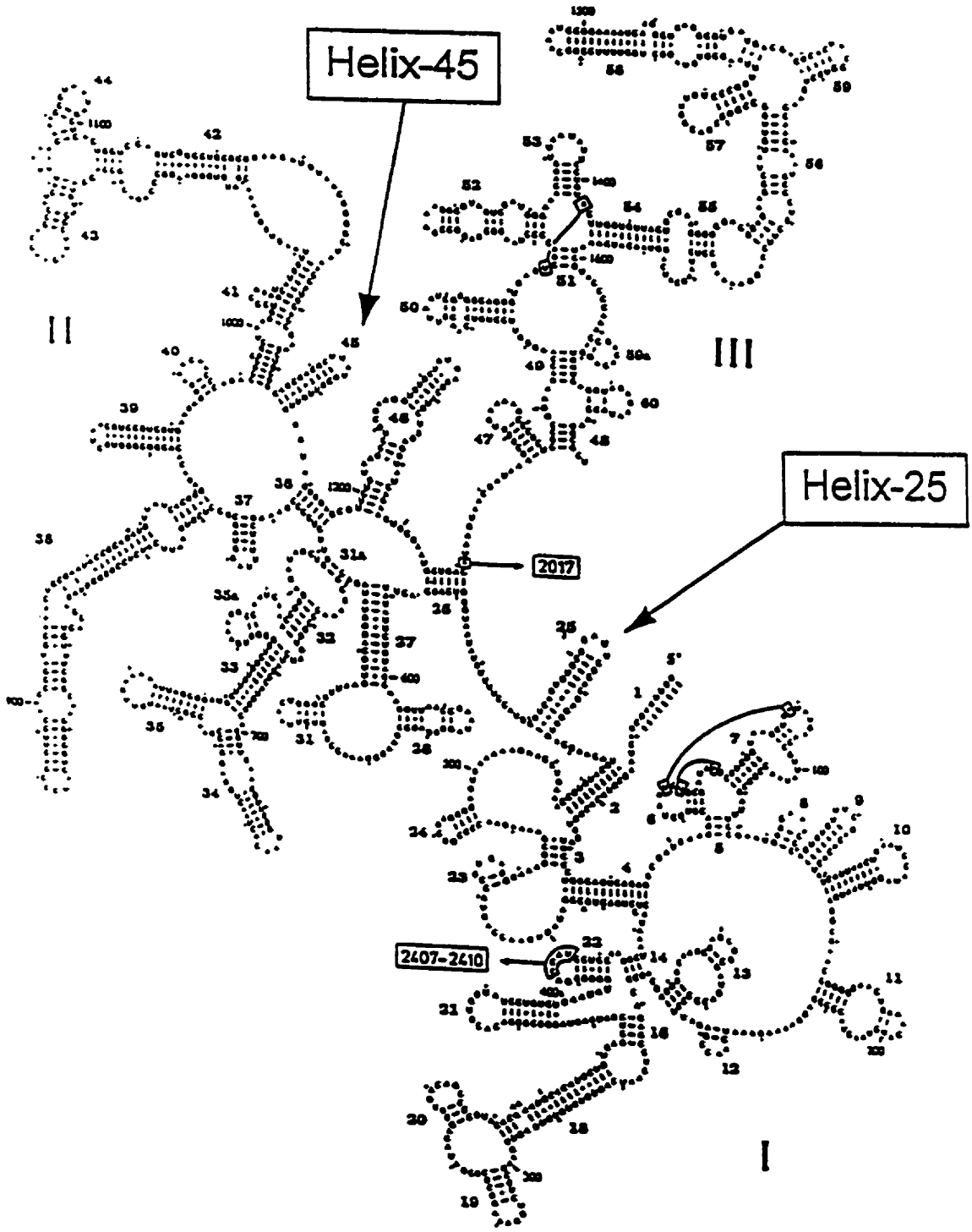
2.2.1 Diversity of distribution and general properties.

The 23S rRNA of several prokaryotes is fragmented as a result of the post-transcriptional removal of sequences called IVSs by the rRNA processing enzyme RNase III (Fig. 2.1) (Burgin et al., 1990). The IVSs replace a tetraloop on a short stem in the putative secondary structure of the 23rRNA (Fig. 2.2). These IVSs largely occur around base pairs 550 and 1170, which corresponds to helix-25 and helix-45 in the putative secondary structure of the rRNA (Noller, 1984)(Fig. 2.2). IVSs have also been found at other locations, at bp 135 and bp 400 of the *rrl* gene in Rhizobiaceae (Selenska-Pobell and Evguenieva-Hackenberg, 1995) as well as in the 16S rRNA of *Campylobacter helveticus* (Linton et al., 1994b), *Helicobacter canis* (Linton et al., 1994a) (Redburn and Patel, 1993), *Clostridium* spp. (Rainey et al., 1996), *Rhizobium tropici* (Willems and Collins, 1993), and Caedibacter caryophila (Springer et al., 1993). IVSs are unevenly

distributed among the bacterial genera, being found in *Agrobacterium* (Hsu et al., 1992), *Brucella* (Hsu et al., 1992), *Rhizobium*(Selenska-Pobell and Evguenieva-Hackenberg, 1995) (Alpha proteobacteria); *Coxiella* (Afseth et al., 1995), *Hemophilus* (song et al., 1999), *Salmonella* (Winkler, 1979) (Burgin et al., 1990) (Mattatall and Sanderson, 1996), *Yersinia* (Skurnik and Toivanen,1991), *Proteus* (Miller et al., 2000), *Providencia* (Miller et al., 2000) (Gamma proteobacteria); *Campylobacter* (Konkel et al., 1994) (Linton et al., 1994) (Trust et al., 1994), *Helicobacter* (Hurtado et al., 197) (Epsilon proteobacteria) and *Leptospira* (Ralph and McClelland, 1993) (Spirochetes). They disrupt the normal continuity of the 23S rRNA gene and the mature ribosome thus contains fragmented rRNA. Most IVSs are about 90-150 bp in length and do not contain ORFs, except in *Leptospira*; however the putative protein coded for in *Leptospira* has not been detected in the cell (Ralph and McClelland, 1994).

Figure 2.2: 23S rRNA structural model.

The 23S rRNA model of *E.coli* (Egebjerg et al., 1990) is shown. Bases 1 to 1647 with the primary domains I, II, and III and the positions of helix-25 and helix-45 tetraloops are indicated. The domains are defined on the basis of long-range double helices. The structure is based on phylogenetic comparisons and the base pairs are supported by compensating base changes (Leffers et al., 1988). Base pairs within a putative double helix for which there is no negative evidence are joined by a line or dot; the latter represent guanosine-uridine pairings. Helices have been numbered according to the *Desulfurococcus mobilis* 23S rRNA model (Leffers et al., 1987). Every 10^{th} nucleotide from the 5' end is indicated by a short line and every 50th nucleotide is indicated by a longer line. The figure was created using EDSTRUC (Egebjerg et al., 1990).


Helix-45
Helix-25
2017
2407-2410
II
III
I

2.2.2 Intervening Sequences in *Salmonella.*

The 23S rRNA of *Salmonella typhimurium* is fragmented while that of *E. coli* is present as an intact 2900bp fragment (Winkler, 1979). The basis for this fragmentation is the removal of novel sequences called IVSs by the rRNA processing enzyme RNase III (Burgin et al., 1990). The IVSs are present at base pair 550 (helix-25) and bp 1170 (helix-45) in the 23S rRNA of *S. typhimurium* and *S. arizonae.* The distribution of the number of IVSs at the helix-25 and helix-45 positions in *Salmonella* is heterogenous in the seven copies of the *rrl* genes distributed around the genome (Mattatall and Sanderson, 1996). Most of the strains of *Salmonella* studied have IVSs identical in sequence in the different copies of the *rrl* gene, including *S. typhimurium* LT2 where 6 identical helix-45 IVSs are present, however *S.typhimurium* LT2 has 2 helix-25 IVSs, which were only 56% identical (Mattatall and Sanderson, 1996).

2.2.3 Intervening sequences and introns.

Group I and Group II introns are mobile genetic elements which function as ribozymes and catalyze their own splicing (Lambowitz and Belfort, 1993) (Belfort et al., 1995). Group I introns catalyze splicing via a series of guanosine-initiated transesterification reactions. In some cases the excised Group I intron cyclizes by an additional transesterification and can be detected in the cell after cleavage (Cech and Bass, 1986) (Cech, 1990) (Saldanah et al., 1993). Group II introns splice when the 2' OH of a bulged nucleotide attacks the 5' splice site, thus generating a characteristic lariat structure (Saldanah et al., 1993) (Peebles et al., 1986) (van der Veen et al., 1986). The mobility of Group I and Group II introns is also mediated by different types of intron-

encoded proteins (Saldanah et al., 1993) (Michel et al., 1989). Group I and Group II introns are capable of "Homing" i.e. transfer from intron-containing to intronless alleles of the same gene. This proceeds via intron-encoded site-specific endonucleases in Group I introns and using intron-encoded reverse transcriptase-like proteins in Group II introns (Dujon et al., 1989) (Belfort et al., 1995). Horizontal transmission of Group I and Group II introns, although unproven, has been invoked to explain cases where closely related introns are present at different locations in different organisms or when introns at the same site are more strongly conserved than the exon sequences (Colleaux et al., 1990) (Lang, 1984) (Hardy and Clark-Walker, 1991).

IVSs are comparable to introns as they disrupt the continuity of the rRNA gene without affecting its function (Burgin et al., 1990). However, IVSs cannot be classified as true introns because the rRNA fragments are not religated after removal of IVSs. Most IVSs do not have ORFs or terminal consensus sequences to facilitate their excision or mobilization. The mechanism of excision of IVSs is different from that of introns, for it is mediated by RNase III. The recognition sites for RNase III are conserved secondary structures; terminal consensus sequences have not been identified (Robertson, 1982) (Gegenheimer et al., 1977) (Dunn, 1976) (Krinkle and Wulff, 1990) (Chelladurai et al., 1991).

2.2.4 Horizontal transfer of intervening sequences.

On the basis of the sporadic nature of the occurrence of fragmented rRNAs in bacteria and substantial sequence divergence among homologous IVSs in closely related organisms it was suggested that such fragmentation is a derived state and not a primitive

one (Burgin et al., 1990). The IVSs that occur in the 23S rRNA genes of *S. typhimurium* and *S. arizonae* in helix-25 are 94% identical and indicate common ancestry. In contrast the two helix-25 IVSs that occur in different *rrl* genes of a single strain of *S. typhimurium* LT2 are only 56% identical. This suggests that the helix-25 IVS was horizontally transferred between *S. typhimurium* and *S. arizonae* and the two helix-25 IVSs in *S. typhimurium* arose independently (Burgin et al., 1990) (Mattatall and Sanderson, 1996). Horizontal transfer is transfer of genetic material between strains and is mediated by methods like transformation, transduction and conjugation; this is unlike vertical transfer by replication from parent cells.

Comparison of nucleotide sequences of the IVSs of *Y. enterocolitica* and *Salmonella* spp. reveals that the IVSs of highly virulent *Y. enterocolitica* are more identical to *Salmonella* IVSs than to the IVSs of poorly virulent *Y. enterocolitica* (Skurnik and Toivanen, 1991). This is not consistent with the overall chromosomal DNA homology as the DNA homology (determined by DNA-DNA reassociation) between different *Y. enterocolitica* strains is 60% whereas that between *Y. enterocolitica* and salmonellae is about 20% (Bercovier et al., 1980). Similar discrepancies in nucleotide identities of helix-25 IVSs from different *rrl* genes of *Proteus vulgaris* suggest horizontal transfer of IVSs between *Proteus* and *Salmonella* (Miller et al., 2000). This indicates that the IVSs have been introduced by horizontal transmission into the bacteria late in evolution. Similar observations of recent horizontal transfer of IVSs have been made between the chromosomes of the spirochete species, *Leptospira noguchi* and *Leptospira weillii* (Ralph and McClelland, 1994). The IVSs sequenced from *Coxiella burnetti*

contain a single ORF that is 70% identical to the ORFs from *Leptospira* IVSs (Afseth et al., 1995). These similarities have been explained on the basis of horizontal transfer.

2.2.5 Speculated functions for intervening sequences.

IVSs lead to the fragmentation of the 23S rRNA; the resulting fragments are not religated and are incorporated into the mature functional ribosome. It has been suggested that the rRNA fragments are held together by RNA-protein interactions (Burgin et al., 1990). The function of IVSs in the cell is still unclear. An RNase III-deficient strain with the IVSs intact in the rRNA is viable in culture, however it has a reduced growth rate, presumably because RNase III is also involved in other rRNA processing reactions (Mattatall and Sanderson, 1998). It has also been shown that the transfer of a plasmid with an IVS-containing *rrl* gene from *Salmonella* into *E. coli* leads to presence of fragmented 23S rRNA in *E. coli*, but the cell maintains wild type growth rates (Gregory et al., 1996). Thus the removal of IVSs from the mature rRNA is not essential in mutant strains maintained in culture, but it is not clear if IVSs are ever retained in wild type strains.

The IVSs have not been proven to perform any function except the fragmentation of 23S rRNA, however their maintenance in some genera suggests the existence of a positive selective pressure. Two of the functions proposed for IVSs are as follows: 1) IVSs with conserved sequence identity are found in the enterobacterial species of *Yersinia* and *Salmonella*. It was suggested that IVSs play a role in resistance to bacteriocins elaborated by the normal flora of the gut. Bacteriocins colicin E3 and DF 13 attack and cleave the 16S rRNA, thus inhibiting protein synthesis. Skurnik and Toivanen,

(1991) speculated the presence of an unknown bacteriocin which uses the native 23S rRNA loop at the sites of IVS insertion as a target. Altering the target by IVSs would make the site resistant to the action of bacteriocins. 2) It has also been shown that rRNA in *Salmonella* is degraded as the cells enter the stationary phase. This degradation is a mechanism for the regulation of the bacterial 23S rRNA and ribosome concentration. This degradation might provide a selective advantage to the cells as the degraded rRNA and ribosomes are used as a source of nucleotides and amino acids for subsequent re-synthesis (Kaplan and Apirion, 1975a) (Kaplan and Apirion, 1975b). Experiments showing faster degradation of stationary phase rRNA containing IVSs as compared to rRNA without IVSs have been reported (Hsu et al., 1994). The presence of fragmented rRNA resulting from the excision of IVSs was postulated to confer a selective advantage by increasing the rate of rRNA degradation, helping the cell respond more efficiently to the onset of adverse conditions (Hsu et al., 1994).

2.3 Ribosomal RNA in the stationary phase.

2.3.1 Bacterial responses to starvation in the stationary phase.

Bacterial strains have a remarkable capacity for growth when nutrients are available under appropriate conditions. Similarly, they also very efficiently enter a starvation-induced program that results in a metabolically less active and more resistant state.

In the stationary phase, cells undergo physical and morphological adaptation concomitantly with the acquisition of a dormant state and resistance against physical and

chemical stresses. Morphological changes include reduction in size and becoming more spherical as a result of the induction of the *bol*A gene in *E. coli* (Loewen and Hengge-Aronis, 1996) (Lange and Hengge-Aronis, 1991). The cytoplasm becomes condensed (Reeve et al., 1984), the membrane hydrophobicity increases making it less fluid, fatty acid composition of cell membranes changes (Kjelleberg et al., 1987), and the nucleoid becomes condensed by replacement of some DNA-binding proteins (Spassky et al., 1984). The overall metabolic rate decreases to a low level. It has been suggested that bacteria have a general starvation response under conditions of starvation and stress. Subtle fluctuations in the ATP levels and proton motive force across the membrane may provide a common signal that regulatory sensors can transduce to the cell's machinery, signifying starvation (Kolter et al., 1993).

The stationary phase adaptation is accompanied by marked changes in the pattern of global gene expression, which is mainly mediated by modulation of the specificity of RNA polymerase by replacement of the promotor recognition subunit σ^{70} by σ^{S}. The stationary phase-specific σ^{S} subunit is encoded by the *rpo*S gene and was originally identified as the product of a regulatory gene *kat*F (Mulvey and Loewen, 1989) (Tanaka et al., 1993) (Tanaka et al., 1995). Other regulatory molecules controlling the induction of transcription in the stationary phase depend on σ^{32}. Many genes induced in response to glucose starvation depend on cAMP and are induced in an *rpo*S-independent manner (Schultz and Matin, 1988). Approximately 100 genes which are highly expressed in the exponentially growing *E. coli* cells are mostly turned off or markedly repressed in the stationary phase cells and a set of 50-100 genes that are repressed in the growing cells begin to be expressed upon entry into the stationary phase (Ishihama, 1997).

In summary, the changes occurring in nondifferentiating bacteria in response to nutrient deprivation and onset of stationary phase are: 1) synthesis of new proteins which make cells effective scavengers of scarce nutrients and confer on them a more resistant phenotype to escape starvation (Matin et al., 1989), 2) increase in the salvage metabolism rate (Dawes, 1976), and 3) utilization of storage polymers for survival (Kjelleberg et al., 1987).

2.3.2 Effect of starvation on ribosomes and rRNA.

Degradation of ribosomes and rRNA is a well known major response of bacteria to starvation and stress (Mandelstam and Halvorson, 1960) (Ben-Hamida and Schlessinger, 1966). This process provides nucleotides and amino acids from the degradation of rRNA and ribosomes (Kaplan and Apirion, 1975a) (Kaplan and Apirion, 1975b). In *E. coli* 20-30% of the cellular RNA is lost within the first 4 hours of starvation for phosphate, degradation then proceeds at a slower rate and 10% more is lost in the next 20 hours (Maruyama and Mizuno, 1966). Ribosomal RNA is preferentially degraded and the loss of viability of starving *E. coli* can be directly correlated with the degradation of ribosomes. It has been suggested that ribosomal degradation is initially an adaptive response to starvation but ultimately becomes suicidal (Davis et al., 1986).

With the onset of starvation, polysomes are broken down to 70S monosomes, which are degraded by RNaseI into their 30 and 50S subunits. 16S rRNA is cleaved by RNaseII and PNPase and 5S rRNA is cleaved by an unidentified endonuclease. After further nuclease and PNPase attack, the ribosomal subunits are dissociated into proteins and nucleotides; the proteins become attached to the cell membrane and the nucleotides

are used in the synthesis of new RNA (particularly mRNA) (Matin et al., 1989). A multicomponent ribonucleolytic degradosome complex formed around RNase E has been shown to be the major site of degradation of 16S and 23S rRNAs (Bessarab et al., 1998). Some of the 70S ribosomes associate with the Ribosome Modulation Factor (RMF) to form 100S ribosome dimers in the stationary phase. It has been suggested that RMF is a stationary-phase-specific inhibitor of ribosome functions and that 100S dimers are stored forms of ribosomes (Wada et al., 1990) (Wada et al., 1995).

2.4 Classification of *Salmonella*.

2.4.1 The genus *Salmonella*.

The genus *Salmonella* belongs to the large eubacterial family *Enterobacteriaceae*. It was originally classified into 2200 serotypes by use of somatic and flagellar antigens to develop the Kaufmann-White classification, but more recently all strains have been proven to belong to the same DNA reassociation group (over 70% DNA reassociation) representing two species, *S. enterica* (for almost all strains) and *S. bongori* (for subgenus V) (Crosa et al., 1973) (Le Minor, 1984). The salmonellae have been separated by electrophoretic typing into numerous clones (Selander et al., 1991). Studies on the cellular and molecular biology of *S. typhimurium* have been summarized (Neidhardt et al., 1987). The use of the name *Salmonella enterica* serovar Typhimurium rather than the name *S. typhimurium* is gaining acceptance, as it is a better indicator of relationships within this group. However, the traditional species names have been used in this report.

2.4.2 Construction of SARB and SARC sets of strains.

Serotyping of the flagellar (H-antigen) and lipopolysaccaride (O-antigen) has been traditionally used for epidemiologically categorizing isolates of *Salmonella* spp. However serotyping does not provide a basis for estimating evolutionary genetic relatedness among strains. Studies of electrophoretically-demonstrable allelic variation at multiple enzyme loci have demonstrated that isolates of the same serovar may be distantly related in chromosomal genotype and that conversely strains of different serovars may be virtually identical in overall genetic character (Beltran et al., 1991) (Selander et al., 1991). The *Salmonella* Reference Collection B (SARB) set is a collection of 72 strains representing 37 serovars of *Salmonella* subspecies. The isolates were characterized by enzyme electrophoresis for allelic variation in 25 chromosomal genes and represent 71 distinctive multilocus genotypes, i.e., electrophoretic types (Boyd et al., 1993). In the present study, the *Salmonella* Reference Collection B (SARB) set of 72 (Boyd et al., 1993) and the *Salmonella* Reference Collection C (SARC) set of 16 strains representing seven subspecies of *Salmonella* and also *S. bongori* (subspecies V) (Boyd et al., 1996) were analyzed. These sets of strains were established by MLEE in the laboratory of R. K. Selander, and are available from the *Salmonella* Genetic Stock Centre (SGSC).

2.5 Research objectives.

The objectives of the study are as follows: to determine if IVSs in *Salmonella* are always present at the previously reported helix-25 and helix-45 locations, or if they could be present at other sites; to determine if fragmentation of rRNA can occur without the

presence of IVSs; to determine if any IVSs are present which are not excised by RNase III; to study the pattern of distribution and nucleotide sequence of IVSs among 72 SARB and 16 SARC strains to determine the extent of vertical and horizontal transmission; and to determine the role of IVSs in the degradation of stationary phase rRNA.

CHAPTER 3: MATERIALS AND METHODS

3.1 Bacterial strains and cultivation conditions.

The *Salmonella* Reference Collection B (SARB) (Boyd et al., 1993) and *Salmonella* Reference Collection C (SARC) (Boyd et al., 1996) sets of strains were obtained originally from R.K.Selander and are maintained at the *Salmonella* Genetic Stock Centre (SGSC, University of Calgary, Calgary; www.ucalgary.ca/~kesander). All strains used in this study are available from the SGSC and are listed in Table 3.1 for the SARB strains and Table 3.2 for the SARC strains. Strains are stored at –70°C in 15% glycerol. Single colonies isolated from the stocks were used in all experiments. All cultures were grown in Luria Broth Base (LB broth) (tryptone 10g/l, yeast extract 5g/l, NaCl 10g/l, adjusted to pH 7.0 with 10M NaOH). 1.5% Bacto-agar (Difco) was added when solid medium was used. The ECOR set of 72 *E. coli* serotypes (Ochman and Selander, 1984) was used for a thorough analysis of all *E. coli*. The wild type strain of *Salmonella* used was *S. typhimurium* LT2 (SGSC 1412). Thirteen *S. bongori* strains (obtained from R.K. Selander), which represent subgroup V of *Salmonella* and are the most divergent *Salmonella* serovars were used (SGSC 3098, 3099, 3100, 3101, 3102, 3103, 3104, 3105, 3106, 3107, 3108, 3109 and 3110).

Table 3.1: Strains of *Salmonella* from the *Salmonella* Reference Collection B (SARB) [a].

Electro-phoretic Type (ET)	SARB No. [b]	RKS Strain No. [c]	*Salmonella* sp.
Ag1	1	1701	*S. agona*
An1	2	2403	*S. anatum*
Ba2	3	4231	*S. brandenburg*
Cs1	4	1280	*S. choleraesuis*
Cs6	5	1239	"
Cs11	6	3169	"
Cs13	7	4640	"
Dt1	8	4647	*S. decatur*
De1	9	246	*S. derby*
De13	10	241	"
De31	11	243	"
Du1	12	1518	*S. dublin*
Du3	13	4717	"
Du2	14	1550	"
Di1	15	4239	*S. duisburg*
En1	16	53	*S. enteritidis*
En2	17	761	"
En3	18	69	"
En7	19	1208	"
Em1	20	1216	*S. emek*
Ga2	21	2962	*S. gallinarum*
Ha1	22	4241	*S. haifa*
He1	23	539	*S. heidelberg*
He3	24	1391	"
Id1	25	4250	*S. indiana*
In1	26	1490	*S. infantis*
In3	27	1452	"
Mi1	28	2833	*S. miami*
Mi5	29	4381	"
Mo1	30	1762	*S. montevideo*

Mo6	31	1740	"
Mu1	32	3121	*S. muenchen*
Mu2	33	4288	"
Mu3	34	4300	"
Mu4	35	4272	"
Np8	36	2016	*S. newport*
Np11	37	1915	"
Np15	38	1956	"
Pn1	39	1793	*S. panama*
Pn2	40	1776	"
Pn12	41	1779	"
Pa1	42	4993	*S. paratyphi* A
Pb1	43	3222	*S. paratyphi* B
Pb3	44	3202	"
Pb4	45	3201	"
Pb5	46	3274	"
Pb7	47	3215	"
Pc1	48	4587	*S. paratyphi* C
Pc2	49	4594	"
Pc4	50	4620	"
Pu3	51	2266	*S. pullorum*
Pu4	52	2246	"
Re1	53	4256	*S. reading*
Ru1	54	4938	*S. rubislaw*
Sp3	55	1690	*S. saintpaul*
Sp4	56	1686	"
Sw1	57	4261	*S. schwarzengrund*
Se1	58	2866	*S. sendai*
Sf1	59	2358	*S. senftenberg*
St1	60	4264	*S. stanley*
Sv2	61	4267	*S. stanleyville*
Th1	62	1767	*S. thompson*
Tp1	63	3333	*S. typhi*
Tp2	64	3320	"
Tm1	65	284	*S. typhimurium*
Tm7	66	203	"
Tm12	67	837	"
Tm23	68	4535	"
Ts1	69	3134	*S. typhisuis*
Ts3	70	3133	"

Wi1	71	4000	*S. wien*
Wi2	72	3998	"

[a] The SARB set of strains was obtained from R.K. Selander (Boyd et al., 1993) and is circulated by the *Salmonella* Genetic Stock Centre.

[b] The SARB No. is the *Salmonella* Reference Collection B number for a set of 72 strains of *Salmonella* representing sub-genus I.

[c] The RKS No. is the R. K. Selander number assigned to stock strains.

Table 3.2: Strains of *Salmonella* from the *Salmonella* Reference Collection C (SARC) [a].

Group	SARC No. [b]	RKS No [c]	Species
I	1	s4194	*S. typhimurium*
	2	s3333	*S. typhi*
II	3	s2985	
	4	s2993	
IIIa	5	s2980	
	6	s2983	
IIIb	7	s2978	
	8	s2979	
IV	9	S3015	
	10	S3027	
V	11	S3041	*S. bongori*
	12	S3044	*S. bongori*
VI	13	s2995	
	14	S3057	
VII	15	S3013	
	16	S3014	

[a] The SARC sets of strains was obtained from R. K. Selander (Boyd et al., 1996) and is circulated by the *Salmonella* Genetic Stock Center, University of Calgary.

[b] The SARC No. is the *Salmonella* Reference Collection C number for a set of 16 strains of *Salmonella* representing sub-genus I.

[c] The RKS No. is the R. K. Selander number assigned to stock strains.

3.2 Enzymes and Chemicals.

Taq polymerase, dNTPs and DNaseI were purchased from Pharmacia. RNasin (placental RNase inhibitor) was purchased from Promega. DEPC (Di Ethyl Pyro Carbonate) and N-laurosyl sarcosine from Sigma. Endonuclease I-*Ceu*I was purchased from New England Biolabs. Proteinase K was from Roche Molecular Biochemicals. Most other chemicals including PMSF (Phenyl Methyl Sulfonyl Flouride) and agarose were purchased from Gibco-BRL.

3.3 Primers.

The region of the ribosomal DNA containing helix-25 was amplified using primers P1 and P2 and the region containing helix-45 was amplified using primers P3 and P4 (See Fig.4.1 for location of primers). Primers rrl6 and rrl7 were the nested primers used for sequencing the helix-25 region and rrl8 and rrl9 were used for sequencing the helix-45 region. All primers were synthesized by the University Core DNA Services (Health Science Centre, University of Calgary): P1 (5'gcgtcggtaaggtgatatg3'), P2 (5'gctatctcccggtttgattg3'), P3 (5'ccgatgcaaactgccaatac3'), P4 (5'ttctctacctgaccacctg3'), rrl6: (5'tactcctgactgaccgatag3'), rrl7: (5'ggctagatcaccgggtttcg3'), rrl8: (5'cctgcgcggaagatgtaacg3'), rrl9: (5'gcattcgcacttctgatacc3'). The PCR primers are located at *E. coli rrlB* base (Noller, 1984) 74-92, 786-805, 901-920 and 1616-1634, respectively and the sequencing primers are located at base 430-450, 675-695, 1119-1139, 1240-1259 respectively.

3.4 RNA isolation.

Cells from a single colony were grown with shaking in 3 ml of LB to obtain an overnight culture. 1 ml of the overnight culture was inoculated in 20 ml fresh LB and grown for 4 hours to an O.D. of about 200-250 Klett units. Cells from 10 ml of this culture were isolated by centrifugation at 5000 X g for 15 minutes at 4°C. These cells were suspended into 3 ml extraction buffer (10mM sodium acetate and 0.15 M sucrose (pH 4.8)) at 4°C, SDS was added to a final concentration of 1% and the solution was vortexed. 3 ml of phenol (sodium acetate buffered (pH 4.8)) was added and the solution was vortexed, incubated at 65°C for 5 minutes, incubated at 0°C for 5 minutes and centrifuged at 5000 x g for 15 minutes at 4°C. The phenol extraction was repeated thrice with the aqueous phase, followed by a final extraction with 3 ml of chloroform. 2.5 volumes of 100% ethanol was added and the RNA was precipitated overnight at -20°C. The RNA was centrifuged at 5000 x g for 15 min to obtain a pellet and the pellet was washed with 1ml of 75% ethanol and resuspended in 91 µl of DNase buffer (50 mM Tris-Cl, 10 mM $MgCl_2$ (pH 7.5)) with dithiothreitol to a final concentration of 5 nM. To this solution, 7 U/µl of DNaseI and 36 U/µl of RNasin were added, and the solution was incubated for 30 minutes at 37°C, then extracted with an equal volume of chloroform, precipitated with 2.5 volumes of ethanol, and stored at -70°C in DEPC (Di Ethyl Pyro Carbonate) treated ddH_2O. DEPC is an RNase inhibitor and all equipment and solutions were treated with 0.1% DEPC at 37°C overnight before use.

3.5 Denaturing RNA electrophoresis.

The RNA was quantitated spectrophotometrically at 260 nm and 10 - 20 μg of RNA was electrophoresed in a 1.5% agarose gel using glyoxal-DMSO denaturation (Sambrook et al., 1989). A final concentration of 10 nM sodium iodoacetate was added to the gel to inactivate the RNases. RNA can be denatured by heating a sample with a mixture of 1M glyoxal and 50% DMSO for 1 hour at 50°C. Glyoxylation introduces an additional ring into the guanosine residues, which in turn causes steric hindrance of guanosine-cytosine base-pairings. The denatured RNA is then electrophoresed using 10 mM sodium phosphate buffer (pH 6.8) at 5 V/cm with constant recirculation from anode to cathode.

3.6 Northern blotting.

The RNA was blotted to a Hybond-N+ membrane (Amersham) by capillary action for 18-20 hrs in 10 X SSC (Stock solution of 20 X SSC contains 3M NaCl, 300mM sodium citrate (pH 7.0)). After transfer the membrane was fixed in 200 ml of 5% acetic acid and stained using 0.04% methylene blue in 0.5 M sodium acetate (pH 5.2)(Sambrook et al., 1989). The blot was then washed several times with water, and photographed using the DIAMED photodocumentation system or scanned using Scan Jet 6100C (Hewlett-Packard).

3.7 Polymerase Chain Reaction and Agarose gel electrophoresis.

PCR reactions were carried out according to the instructions accompanying the Taq polymerase on a Techne Gene E Thermal Cycler. Templates were prepared by boiling bacterial cells (obtained on the tip of a toothpick from a single colony) in 500 μl of water

for 5 min and then rapidly cooling on ice. 2 µl of the template was used for each reaction. The thermal profile consisted of thirty cycles of 1 min denaturation (94°C), 1 min annealing (56°C), and 1 min extension (72°C). A final extension step for 10 min at 72°C was performed. All PCR products were electrophoresed in 1.5% agarose gels in 0.5 X TBE buffer (1 X TBE buffer contains 90 mM Tris-HCl, 90 mM boric acid, 2 mM EDTA, pH 8.0) at 8 V/cm in the presence of 0.5 µg/ml of ethidium bromide.

3.8 Sequencing strategies.

Templates for sequencing were prepared by PCR and purified using the Wizard PCR preps DNA purification kit (Promega). Cycle sequencing of PCR products was performed by the University Core DNA Services (Health Science Centre, University of Calgary), using Automated Applied Biosystems (ABI) Sequencing and Taq Dye Deoxy Terminator Cycle Sequencing Kit (ABI). Nested sequencing primers were used for this purpose.

3.8.1 Preparation of template from strains in which only one of the seven *rrl* genes has an insertion in helix-25 or helix-45.

PCR reaction was performed using the appropriate primers to amplify the helix with a single insertion and electrophoresed in a 2% agarose gel. The slow running single band containing the IVS was excised under UV light and the band was stored in 50 µl of ddH_2O overnight for the DNA to diffuse out of the band. This band was again PCR amplified using the water with DNA as the template,

purified and sequenced. Four to five cycles were required to achieve a pure band for sequencing.

3.8.2 Preparation of template from strains with more than one insertion in helix-25 and/or helix-45.

3.8.2.1 Preparation of genomic DNA embedded in agarose and enzymatic digestion (Liu et al., 1993): Cells grown overnight at 37°C in LB broth were harvested by centrifugation and resuspended in 0.5 ml of Cell Suspension Solution (10 mM Tris-HCl, 20 mM NaCl, 100 mM EDTA (pH 7.2)). The cells were then mixed in 0.5 ml of 1.4% agarose. The cell-agarose mixture was drawn into 1 ml syringes from which the needle had been removed. After solidifying, the agarose was cut into discs about 1 mm thick and placed in Lysing Buffer (10 mM Tris-HCl (pH 7.2), 50 mM NaCl, 100 mM EDTA, 0.2% SDS, 0.5% N-laurosyl sarcosine) and heated in a water bath at 65°C for 45-60 min. After lysing the cells, the disks were washed in Wash Solution (20 mM Tris-HCl, 50 mM EDTA (pH 8.0)) and treated with 3ml of Proteinase K Solution (100 mM EDTA, 0.2% SDS, 1%N-laurosyl sarcosine, with 0.66 mg of Proteinase K per ml). The disks were incubated at 42°C with gentle agitation for 18-24 hours. After protein digestion the disks were rinsed twice in Wash Solution and incubated in 5 ml of PMSF Solution (Phenylmethylsulfonyl Fluoride) (1 mM PMSF in Wash solution) for 2 hours with gentle shaking at room temperature to remove remaining Proteinase activity. After rinsing the disks twice in Wash Solution and in Storage Solution they were used for

endonuclease digestion. The genomic DNA embedded in agarose disks was digested by equilibriating a single disk in 2 X concentration of I-*Ceu*I NE buffer (20 mM Tris HCl, 20 mM $MgCl_2$, 2 mM DTT, pH 8.5 @ 37°C) for 15 min at room temperature. Digestion was then done using 0.4 units of I-*Ceu*I enzyme in 1X I-*Ceu*I buffer and 1X BSA for 2-3 hours at 37°C.

3.8.2.2 Pulsed Field Gel Electrophoresis (PFGE):A single disk digested with I-*Ceu*I was inserted in each of the wells of a 0.7% agarose gel and electrophoresed in 1X TBE buffer with 1μg of ethidium bromide added per ml. The electrophoresis profile for the initial separation of the I-*Ceu*I fragments consisted of pulse times of 30-60s, an angle of 120°, current of 300V and a run time of 23 hours. For the separation of fragments B and G (in cases where the band lengths of B and G was sufficiently different and allows their separation) the gel was further run for 10 hours with pulse times of 70-80s, angle of 150° and a current of 300V. The electrophoresis was performed with either a Bio-Rad CHEF Mapper, or Bio-Rad DRII PFGE apparatus.

3.8.3 Amplification of IVSs from individual rrl genes.

Individual IVSs were isolated from the I-*Ceu*I fragments and used for sequencing. The agarose band containing the I-*Ceu*I fragment was dissolved by immersing the tube containing the band in boiling water for 2-5 min. The band was diluted to a 1 in 10 and 1 in 100 concentration and the diluted DNA was used as the template for PCR amplification of the IVS from the band.

3.8.4 Purification of PCR products.

The PCR products were purified using the Wizard PCR preps DNA purification system. 200-300 µl of the aqueous phase of the PCR reaction was transferred to a microfuge tube and 100 µl of the direct purification buffer was added. To this 1 ml of the resin was added and vortexed briefly. The DNA/resin mix was pipetted into a syringe with a minicolumn at the tip. Vacuum was applied to draw the DNA/resin mix into the minicolumn, the column was washed with 2 ml of 80% isopropanol and the DNA was eluted out in 50 µl of sterile distilled water.

3.9 Cell lysis, RNA isolation and gel electrophoresis.

Cells from a single colony were grown with shaking for 8 hours and 500 µl of this culture was inoculated into 50 ml of fresh medium for growth curves, viability studies and RNA degradation experiments. Growth was monitored with a Klett-Summerson colorimeter at 540 nm every hour coinciding with the RNA extractions. 5 ml of the culture was used for RNA extraction; the cells were pelleted by centrifugation at 5000 X g for 15 min at 4°C, resuspended in 3 ml of extraction buffer (10mM sodium acetate, 0.15 M sucrose, pH 4.8), sodium dodecyl sulfate (SDS) was added to a final concentration of 1%, and the cells were incubated at 65°C for 5 min to ensure complete cell breakage. In addition to this, to test the most efficient method of cell breakage, the cells were lysed by sonication using 5-7 25s pulses, or by passing the cell suspension through a French press (2 x 9000lb/in^2; 2 x 62 MPa). 3 ml of phenol (sodium acetate buffered (pH 4.8)) was added to the lysed cells and the solution was vortexed, incubated

at 65°C for 5 min, incubated at 0°C for 5 min, and centrifuged at 5000Xg for 15 min at 4°C. The phenol extraction was repeated thrice and after the final extraction, the aqueous layer was stored at -20°C.

10 µl of the RNA extract was electrophoresed in 1% agarose gels in 0.5 X TBE buffer (1 X TBE buffer contains 90 mM Tris-HCl, 90 mM boric acid, 2 mM EDTA, pH 8.0) at 8 V/cm in the presence of 0.5 µg/ml of ethidium bromide.

3.10 Viable cell counts.

Viable cells counts were determined by making 10-fold serial dilutions in LB broth, spreading the cell suspension with a glass spreader in duplicate on LB agar plates, and incubating the plates at 37°C overnight.

3.11 Computer analysis.

Sequence alignments were done using CLUSTALX version 1.6 for Macintosh. The rRNA structure was predicted by free energy minimization using the algorithm of Zuker as implemented in the mfold program (Zuker et al., 1999). Percent nucleotide identity comparisons were done using DNASIS version 2.0 for Macintosh (Hitachi). The GenBank, EMBL, and DDBJ databases were searched for similar sequences using the BLASTn service available at the NCBI (National Center for Biotechnology Information). Advanced BLAST searches with decreased e values were done to search for IVSs substantially diverged from the *Salmonella* IVSs which could not be detected by a Basic BLAST search (Altschul et al., 1997).

CHAPTER 4: RESULTS- Analysis of IVSs in the SARB and SARC sets

4.1 Diversity of distribution of IVSs in the SARB and SARC sets on the basis of PCR analysis.

Using whole genomic DNA of the bacterial strains as template, primers P1 and P2 were used for the amplification of the region of DNA including helix-25 of the *rrl* gene (amplicon A). Primers P3 and P4 were used for the amplification of the helix-45 region in the DNA (amplicon B) (Fig. 4.1). The PCR products were subjected to gel electrophoresis to determine the number of *rrl* genes with IVS insertions (Fig.4.2). The PCR products represent the helix-25 and helix-45 regions amplified from seven copies of the *rrl* genes. The PCR amplicon is smaller (faster migrating) and is expected to be 731bp (helix-25) or 733bp (helix-45) when there is no IVS present (based on *E. coli rrl* gene numbering) (Noller, 1984). The presence of an IVS yields a PCR product that is larger (slower migrating). On the basis of relative intensities of the slow and fast migrating bands the number of *rrl* genes with IVSs out of the total seven *rrl* genes was determined; this analysis is dependent on the assumption that amplification of all *rrl* genes is equivalent.

A comparison of the PCR amplicons from *E. coli* (lanes 1 and 2 of Fig. 4.2, panel A) and *Salmonella* (lanes 3 and 4 of Fig. 4.2, panel A) shows the absence of IVSs in *E. coli* and presence in *S. typhimurium* LT2. Both amplicons A and B in *E. coli* K12 (Ec) are fast migrating, showing that *E. coli* does not have any IVSs. In *S. typhimurium* LT2 (Stm) the band intensity of the slow migrating band for amplicon A (helix-25) corresponds to 5 copies of *rrl* genes without IVSs and the band intensity of the fast

migrating band corresponds to 2 copies of *rrl* genes with IVSs. Similarly amplicon B indicates 6 *rrl* genes with IVSs in helix-45, and one without; these data are consistent with the earlier report (Mattatall and Sanderson, 1996). This reveals the presence of 2 helix-25 and 6 helix-45 IVSs. Similar analysis of fig. 4.2 panel A leads to the conclusion that Lane 13 representing the PCR amplicons from Pb1 (S. *paratyphi* B, ET1) shows that the amplified helix-25 region does not have any slow migrating (IVS containing) bands. The amplification products from all seven *rrl* genes are represented by the fast migrating (non-IVS containing) bands in lane 14 of Fig 4.2. This suggests that Pb1 has no insertions in helix-25 and has 7 insertions in helix-45. On the basis of similar comparative analysis Pb3 (*S. paratyphi* B, ET3) has one insertion in helix-25 and no insertions in helix-45 (Fig. 4.2, lanes 11 and 12). Pb4 (*S. paratyphi* B, ET4) has 2 helix-25 and 4 helix-45 IVSs (Fig 4.2, lanes 9 and 10), while Pb5 (*S. paratyphi* B, ET5) and Pb7 (*S. paratyphi* B, ET7) have 0:1 and 2:4 helix-25 and helix-45 IVSs, respectively (Fig 4.2, lanes 5, 6, 7, 8). An interpretation of the number of IVSs for which the PCR data is shown in Fig. 4.2, panel A, is shown in Fig. 4.2, panel B. All the strains in the SARB and SARC sets and the *S. bongori* strains were analyzed similarly by PCR to evaluate the number of IVSs in both helices. In some cases (e.g., Pb1) the assignment of numbers of IVSs is clear, but in others it is more difficult; to achieve this, several independent PCR reactions were done. Quantitation of PCR products on the basis of the relative intensities of the bands from the seven *rrl* genes may help in the accurate analysis of genes with and without IVSs. The amount of DNA in each band can be analyzed on the basis of band intensity, which can then be used for conversion to molar ratios of the DNA present in each band.

Figure 4.1: Diagrammatic representation of the PCR strategy.

A single rRNA operon is comprised of one 16S rRNA (*rrs* gene), one or two tRNAs, and one 23S rRNA (*rrl* gene). The *rrl* gene may include intervening sequences at H25 (helix-25 of the proposed secondary structure of the rRNA, at bp 550) and/or at H45 (helix-45, at bp 1170). Primers P1 and P2 were used for amplification of a 731 bp region containing helix-25 (amplicon A), and primers P3 and P4 for amplification of a 733 bp region containing helix-45 (amplicon B). When these amplicons contain IVSs, they result in PCR products larger by about 100-150bp.

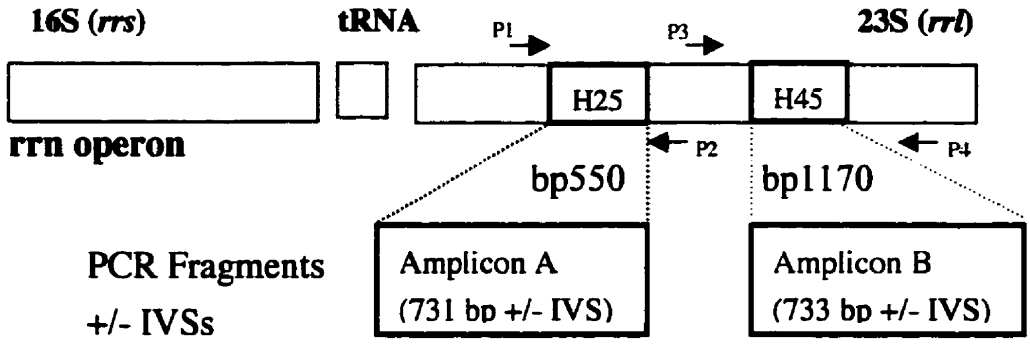

16S (rrs)
tRNA
P1
P3
23S (rrl)
H25
H45
rrn operon
P2
P4
bp550
bp1170
PCR Fragments
+/- IVSs
Amplicon A
(731 bp +/- IVS)
Amplicon B
(733 bp +/- IVS)

Figure 4.2:

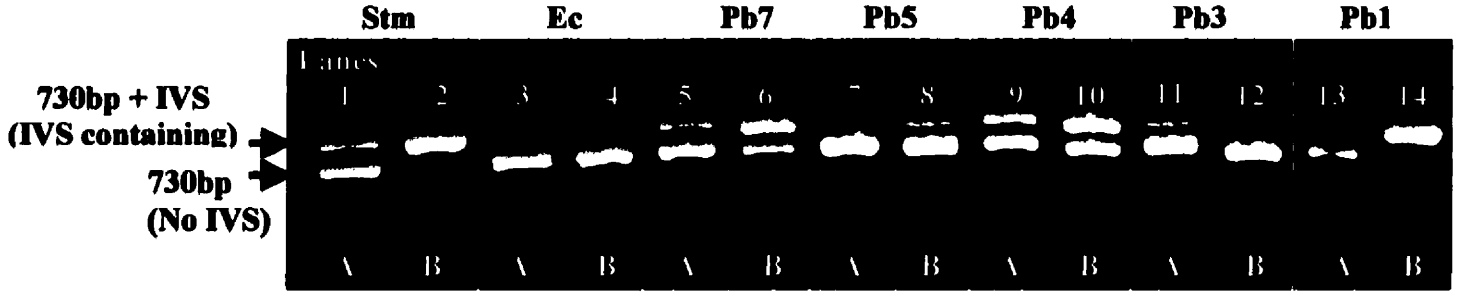
Panel A. PCR amplicons from selected SARB strains.

PCR products obtained by amplification of the helix-25 (A) and helix-45 (B) regions of the 23S rRNA using templates from the following strains: Stm, *S. typhimurium* LT2; Ec, *E. coli* K-12; Pb7, *S. paratyphi* B, ET7 (=Electrophoretic type 7); Pb5, *S. paratyphi* B, ET5; Pb4, *S. paratyphi* B ET4; Pb3, *S. paratyphi* B, ET3; Pb1, *S. paratyphi* B, ET1. "A" indicates amplicon A; "B" indicates amplicon B. Each strain has seven copies of the *rrl* gene. The lower band in each lane (approximately 730 bp) represents PCR product from an *rrl* gene which has no IVS; the upper band results from an *rrl* gene which has an IVS.

Panel B. Interpretation of the number of *rrl* genes with IVSs in each strain.

The number of IVSs in helix-25 and helix-45 for strains whose PCR products are shown in PanelA. The strains and their electrophoretic types are shown and the number of *rrl* genes out of seven containing helix-25 and helix-45 IVSs are indicated.

A



B

Strain/ Electrophoretic type	Insertions in helix-25	Insertions in helix-45
S.typhimurium LT2 (Stm)	2	6
E.coli (Ec)	0	0
S. paratyphi B ET=7 (Pb 7)	2	5
S. paratyphi B ET=7 (Pb 5)	0	1
S. paratyphi B ET=7 (Pb 4)	2	4
S. paratyphi B ET=7 (Pb 3)	1	0
S. paratyphi B ET=7 (Pb 1)	0	7

4.2 Diversity of distribution of IVSs in the SARB and SARC strains by RNA analysis.

The extent of fragmentation of the 23S rRNA was examined in order to determine the site(s) of cleavage in the rRNA, the proportion of the rRNA molecules in which cleavage has occurred, and thus to infer the number of *rrl* genes containing IVSs. Fig.4.3 shows the known sites of the IVSs previously determined in *Salmonella rrl* genes, and the sizes of the rRNA fragments which would result if cleavage occurs at these sites. Fig.4.4, panel A shows the rRNA fragmentation patterns for the same strains for which PCR data reveals the presence of IVSs in Fig.4.2, panel A.

The extent and pattern of fragmentation produced in the rRNA was studied by analyzing the different band sizes and the intensity of individual bands produced by the seven copies of the *rrl* gene. The different band lengths add up to reflect the cleavage products resulting from 7 copies of *rrl* genes after the removal of the IVSs. The band intensity resulting from seven copies of the 16S rRNA (1.5kb band) was used as a standard for comparison to the other band intensities to determine the number of copies of a band present. An intact 23S rRNA fragment is 2.9 kb in length. 23S rRNA fragments of 2.4 and 0.5 kb indicate that one *rrl* gene carries an IVS in helix-25, resulting in cleavage of the rRNA without religation. 23S rRNA fragments of 1.7 and 1.2 indicate that one *rrl* gene carries an IVS in helix-45. 23S rRNA fragments of 1.7, 0.7 and 0.5 kb indicate that one *rrl* gene carries IVSs in both helices (Fig.4.3).

E. coli K12 gives an intact 2.9kb band representing the 23S rRNA as well as an intact 1.5kb band representing the 16S rRNA (Fig. 4.4, panel A, lane 2). This shows that there is no fragmentation of the 23S rRNA, confirming earlier reports that it has no IVSs

(Burgin et al., 1990); as expected, sequencing the whole genome of *E. coli* K12 reveals no IVSs in all seven *rrl* genes (Blattner et al., 1997). In the case of Pb1 (*S. paratyphi* B, ET1) (Fig. 4.4, lane 7) the 2.9kb band is completely absent and all seven *rrl* gene fragments are present as 1.7 kb and 1.2 kb bands. The fragmentation pattern and the band lengths indicate the excision of helix-45 IVSs from all the seven *rrl* genes. Lane 6 of Fig. 4.4 representing rRNA from Pb3 (*S. paratyphi* B, ET3) shows the presence of a 2.4 kb band and a 0.5 kb band (barely visible in this picture), indicating that the fragmentation results from the excision of one helix-25 IVS. A large amount of the 2.9kb band (uncleaved) rRNA and the absence of 1.7kb, 1.2kb and 0.7kb bands indicates that the strain has no IVSs in helix-45. On the basis of similar analysis it can be interpreted that Pb5 (*S. paratyphi* B, ET5) (Fig. 4.4, lane 4) has no helix-25 IVSs and one helix-45 IVS. In the case of *S. typhimurium* LT2 (Fig. 4.4, lane 1), the absence of an intact 2.9kb band demonstrates that all the seven *rrl* genes have either a helix-25 or a helix-45 IVS or both. The presence of a 2.4kb band with one of the two 0.5kb bands suggests that one of the seven *rrl* genes has a helix-25 IVS only. The presence of five out of the six 1.7kb and five 1.2kb bands suggests that five of the seven *rrl* genes have helix-45 IVSs. One of the 1.7kb band, one 0.7kb bands, and one 0.5kb band add up to suggest that one *rrl* gene has both helix-25 and helix-45 IVSs. Pb3, Pb4 and other SARB and SARC strains were similarly analyzed. Fig. 4.4 panel B shows the interpretation of the rRNA fragments lengths to analyze the number of IVS cleaved out for the strains whose rRNA is shown in Fig. 4.4 panel A. In some cases (e.g., Pb1) the assignment of numbers of IVSs is clear, but in others it is more difficult; to achieve this, several independent northern blots were

done at different concentrations of the product to be able to visualize the different bands clearly.

Figure 4.3: Schematic representation of 23S rRNA fragmentation on excision of IVSs.

23S rRNA fragments of 2.4 and 0.5 kb indicate that one *rrl* gene carries an IVS in helix-25. 23S rRNA fragments of 1.7 and 1.2 indicate that one *rrl* gene carries an IVS in helix-45. 23S rRNA fragments of 1.7, 0.7 and 0.5 kb indicate that one *rrl* gene carries IVSs in both helices.

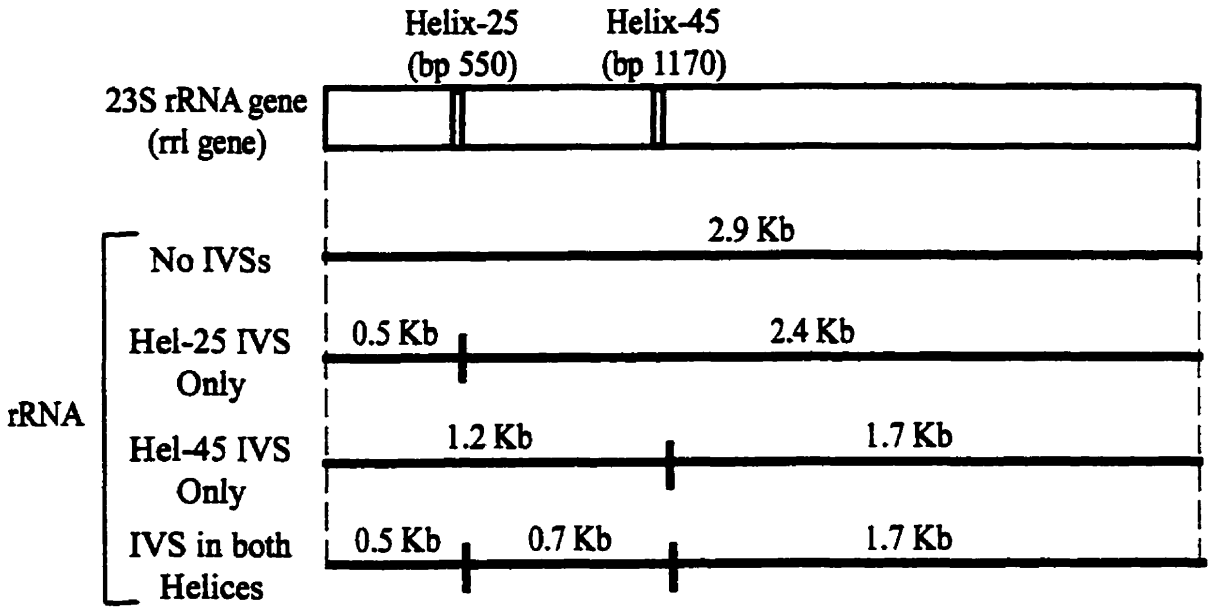

Helix-25
(bp 550)
Helix-45
(bp 1170)
23S rRNA gene
(rrl gene)
2.9 Kb
No IVSs
0.5 Kb
2.4 Kb
Hel-25 IVS
Only
rRNA
1.2 Kb
1.7 Kb
Hel-45 IVS
Only
0.5 Kb
0.7 Kb
1.7 Kb
IVS in both
Helices

Figure 4.4:

Panel A. 23S rRNA fragmentation from selected SARB strains.

Ribosomal RNA from bacterial strains, detected by electrophoresis, Northern blotting, and methylene blue staining. See Table 3.1 for listing of the strains tested. Normal rRNA sizes are indicated on the right (23S, and 16S). The sizes in kb shown on the left indicate fragmented rRNA.

Panel B. Interpretation of rRNA fragmentation pattern.

On the basis of the fragmentation pattern resulting from the excision of IVSs (if present) from the seven copies of the *rrl* genes, the number of helix-25 and helix-45 IVSs in all strains for which rRNA data is shown in panel A is represented. The conclusions column shows that the number of number of IVSs determined from the PCR data and the rRNA fragmentation pattern is in complete agreement.

A

B.

Electro-phoretic type/ Strain	RNA Data						Possession in (H25/ H45)				Conclusions			
											RNA data		PCR data	
	# 0.5 kb	# 0.7 kb	# 1.2 kb	# 1.7 kb	# 2.4 kb	# 2.9 kb	-/-	+/-	-/+	+/+	H25	H45	H25	H45
Stm	2	1	5	6	1	0	0	1	5	1	2	6	2	6
Ec	0	0	0	0	0	7	7	0	0	0	0	0	0	0
Pb 7	2	2	5	3	0	2	2	0	3	2	2	5	2	5
Pb 5	0	0	1	1	0	6	6	0	1	0	0	1	0	1
Pb 4	2	1	3	4	1	2	3	1	2	1	2	4	2	4
Pb 3	1	0	0	0	1	6	6	1	0	0	1	0	1	0
Pb 1	0	0	7	7	0	0	0	0	7	0	0	7	0	7

4.3 Analysis of IVSs in the 72 strains from the SARB set and 16 strains from the SARC set of *Salmonella*.

The SARB set is a collection of 72 strains representing 37 serovars of *Salmonella* subspecies I. The isolates were characterized by enzyme electrophoresis for allelic variation in 25 chromosomal genes and represent 71 different electrophoretic types (ETs) (Boyd et al., 1993). Estimates of genetic relationships among the ETs is indicated in the evolutionary tree in Fig. 4.5. The tree was generated by the neighbour joining method (Saitou and Nei, 1987) from a pairwise matrix of Nei's standard genetic distance for the 25 loci assayed (Saitou and Nei, 1987). The SARC set is a collection of 96 strains of *Salmonella* representing 80 electrophoretic types and eight subgenera (Fig. 4.6). 16 of these strains representing all eight subgenera were additionally analyzed for variation in the nucleotide sequence of five housekeeping genes. These 16 from the SARC set were used for the analysis of IVSs in this study.

The number of *rrl* genes containing IVSs in helix-25 and/or helix-45 regions were determined by PCR as described above (Fig. 4.1) for all the SARB and SARC strains, and the data are recorded in Fig. 4.5 and 4.6. Similarly the rRNA of all 88 strains was examined as previously shown in Figs. 4.3 and 4.4, and the pattern of fragmentation was used to determine the number, out of seven, of the *rrl* genes that produced rRNA that was cleaved at helix-25 and/or helix-45; these data are also recorded in Fig.4.5 and 4.6. The number of IVSs detected by PCR corresponded to the number of IVSs calculated from the fragmentation of rRNA (determined from Northern blots of the rRNA) in all 72 strains of the SARB set and all 16 strains of the SARC set. The strains portray extensive diversity in the possession of IVSs, leading to the following conclusions:

a) Possession of IVSs is common, but there are some strains (8 of 72 in SARB and 0 of 16 in SARC) that have no IVSs.

b) No strains are saturated with the presence of IVSs in all the helix-25 and helix-45 locations, the maximum number of IVSs present in a strain was a 7:4 distribution in helix-25 and helix-45, respectively, in the case of SARC 10 and SARC 16.

c) Closely related strains (according to MLEE) often show identical or very similar pattern of distribution of IVSs, for example Pa1 and Se1; Pc1, Pc2 and Ts1 (shown in blue in Fig 4.5).

d) In spite of this, there is no strong overall pattern of possession of IVSs on the trees in either the helix-25 or helix-45, for example Pn 1, Pn 2, and Pc 4, Ru 1 (shown in red in Fig. 4.5).

Figure 4.5: Pattern of possession of IVSs for the SARB strains.

Evolutionary tree showing relationships determined by multi-locus enzyme electrophoresis, among the 71 ETs represented by 72 SARB strains of *Salmonella* subgenus I (Boyd et al., 1993). The *Salmonella* species, electrophoretic types, and SARB numbers are listed in Table 3.1. The total numbers of IVS insertions in H25 (helix-25) and H45 (helix-45) in the seven *rrl* genes for these strains were determined through PCR and also through rRNA cleavage; the inferred numbers of IVSs, which were the same for both methods in all 72 strains, are shown. Closely related strains showing identical pattern of distribution of IVSs are shown in blue. Closely related strains showing a very different pattern of distribution of IVSs are shown in red.

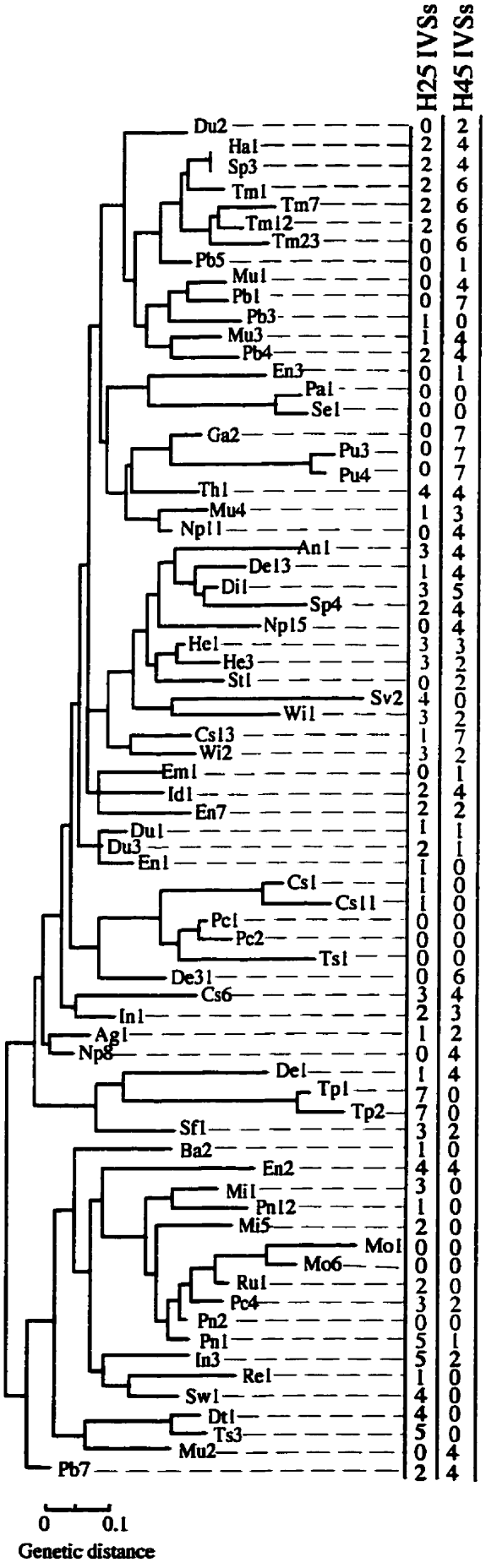
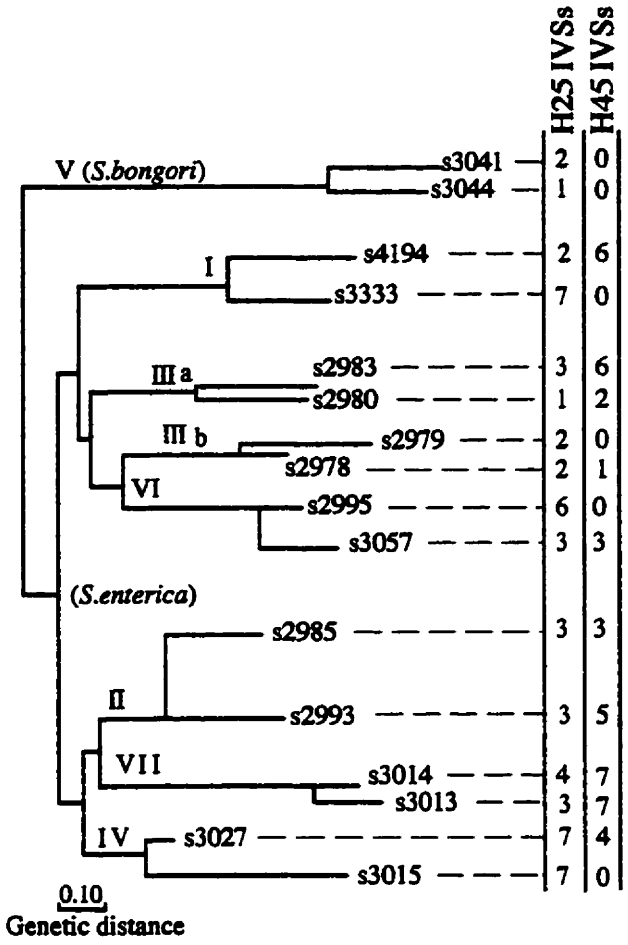

H25 IVSs
H45 IVSs
Du2 0 2
Ha1 2 4
Sp3 2 4
Tm1 2 6
Tm7 2 6
Tm12 2 6
Tm23 0 6
Pb5 0 1
Mu1 0 4
Pb1 0 7
Pb3 1 0
Mu3 1 4
Pb4 2 4
En3 0 1
Pa1 0 0
Se1 0 0
Ga2 0 7
Pu3 0 7
Pu4 0 7
Th1 4 4
Mu4 1 3
Np11 0 4
An1 3 4
De13 1 4
Di1 3 5
Sp4 2 4
Np15 0 4
He1 0 3
He3 3 2
St1 0 2
Sv2 4 0
Wi1 3 2
Cs13 1 7
Wi2 3 2
Em1 0 1
Id1 2 4
En7 2 2
Du1 1 1
Du3 2 1
En1 1 0
Cs1 1 0
Cs11 1 0
Pc1 0 0
Pc2 0 0
Ts1 0 0
De31 0 6
Cs6 3 4
In1 2 3
Ag1 1 2
Np8 0 4
De1 1 4
Tp1 7 0
Tp2 7 0
Sf1 3 2
Ba2 1 0
En2 4 4
Mi1 3 0
Pn12 1 0
Mi5 2 0
Mo1 0 0
Mo6 0 0
Ru1 2 0
Pc4 3 2
Pn2 0 0
Pn1 5 1
In3 5 2
Re1 1 0
Sw1 4 0
Dt1 4 0
Ts3 5 0
Mu2 0 4
Pb7 2 4
0
0.1
Genetic distance

Figure 4.6: Pattern of possession of IVSs for the SARC strains.

Evolutionary tree for 80 ETs of the eight subgenera of *Salmonella* based on MLEE (only the 16 SARC strains which were analyzed are shown) (Boyd et al., 1996). The SARC number and strains are listed in Table 3.2. The total number of IVSs in helix-25 (H25) and helix-45 (H45) were determined through PCR and through analysis of rRNA fragmentation patterns; the inferred number of IVSs from both methods were the same for all strains. The number of IVSs for the analyzed strains is indicated.


H25 IVSs
H45 IVSs
V (S.bongori)
s3041 2 0
s3044 1 0
I
s4194 2 6
s3333 7 0
III a
s2983 3 6
s2980 1 2
III b
s2979 2 0
s2978 2 1
VI
s2995 6 0
s3057 3 3
(S.enterica)
s2985 3 3
II
s2993 3 5
VII
s3014 4 7
s3013 3 7
IV
s3027 7 4
s3015 7 0
0.10
Genetic distance

4.4 Analysis of the nucleotide sequences in helix-25 IVSs.

To obtain IVSs from individual *rrl* genes, two approaches were used. In cases where a single IVS was present, it was isolated by cutting the agarose band and using the DNA to reamplify the IVS until a pure band was achieved (see Material and Methods, section 3.8). In cases where more than one IVS was present, PFGE fragments of I-*Ceu*I-digested genomic DNA were used as templates for PCR amplification of IVSs. The reasons to use I-*Ceu*I are as follows. I-*Ceu*I is an intron-encoded endonuclease which recognizes and cuts a 19 bp sequence that occurs only at bp 2200 in the *rrl* genes for 23S rRNA of *Salmonella* (Liu et al., 1993). The rRNA operons are oriented away from the origin of replication, which is present in the I-*Ceu*I-digested fragment A (Fig. 4.7). Hence digestion of genomic DNA with I-*Ceu*I (at bp 2200 in every *rrl* gene) results in the absence of the helix-25 and helix-45 regions of the *rrl* gene from I-*Ceu*I A, and the presence of two of these regions in I-*Ceu*I C (Fig. 4.7). As a result, IVSs amplified from I-*Ceu*I C could be present in *rrl*D or *rrl*C. Among the products of *I-Ceu*I digested genomic DNA from *Salmonella*, fragments B and G which contain the helix-25 and helix-45 regions of *rrn*G and *rrn*H respectively are approximately of the same length in several strains of *Salmonella*. This restricts the separation of bands B & G and thus it cannot be unambiguously stated from which *rrn* operon the IVS is derived. However the identities of the following helix-25 and helix-45 regions can be inferred: I-*Ceu*I D contains regions from *rrn*A, I-*Ceu*I E from *rrn*B and I-*Ceu*I F from *rrn*E. I-*Ceu*I digested DNA of 10 strains of Salmonella was separated by PFGE (Fig. 4.8); these fragments were excised and used as templates for PCR, as described in Materials and Methods.

Cycle sequencing of the purified PCR products was done by the University Core DNA Sequencing Centre from both strands using nested primers. A sample of strains in the SARB set was tested. The number of nucleotides in the IVSs ranged from 113 to 151 bases. Analysis of the sequences for percent nucleotide identity was done using DNASIS version 2 for Macintosh. IVSs having 90% or greater sequence identity were placed in the same family. The sequenced helix-25 IVSs from *Salmonella* were thus divided into three families, A to C (Table 4.1). Helix-25 IVSs from different *rrl* genes of a particular strain belonged to the same family with the sole exception of *S. typhimurium* LT2, in which the two helix-25 IVSs represent families A and B, as previously reported (Mattatall and Sanderson, 1996). The IVS from *S. enteritidis* (En1), the sole representative of family C, is 151bp, about 37 bp longer than the other helix-25 IVSs which are 113-115bp and it bears only about 57-63% nucleotide identity to the family A and B IVSs (Table 4.1).

The helix-25 IVSs from *Proteus* have been reported as belonging to 3 families (D, E and G) and those of *Providencia rettgeri* have been assigned to family F (Miller et al., 2000). The percent nucleotide identities of helix-25 IVSs from *Proteus* and *Providencia* with those from *Salmonella* are shown in Table 4.2. Family D from *Proteus* is 74% identical to family A from *Salmonella*, however it is only 50% identical to the other helix-25 families from *Proteus* (E and G). The family D IVS from *Proteus* is only 115 bp in length thus resembling the *Salmonella* families A and B, while the other helix-25 IVSs from *Proteus* are up to 185 bp in length (section 4.6). Family F from *Providencia* is 80-85% similar to families E and G from *Proteus*, but is only 49% similar to family D.

Figure 4.7: *rrn* operons in the *Salmonella* genome.

Orientation and position of each *Salmonella rrn* operon on the chromosome is shown as reported earlier (Liu et al., 1993). The positions in the *rrl* genes for 23S rRNA of helix-25 (h25) at bp 550, helix-45 (h45) at bp 1170, and the I-*Ceu*I cleavage site at bp 2200 are shown for all the operons. The approximate locations for the origin of replication (OriC) and terminus (Ter) are indicated. The letters A to G in bold indicate the I-*Ceu*I cleavage fragment and the *rrl* genes present at its boundaries is also indicated. Most I-*Ceu*I fragments should contain one helix-25 and one helix-45 region (e.g., I-*Ceu*I F should have these segments for *rrl*E, fragment E for *rrl*B, etc.). Fragment C should contain helix-25 and helix-45 segments for *rrl*C and D; fragment A should contain no helix-25 and helix-45 regions. An *rrn* operon is represented in a linear fashion at the bottom and shows the 16S, 23S and 5S rRNA in addition to h25, h45 and I-*Ceu*I site.

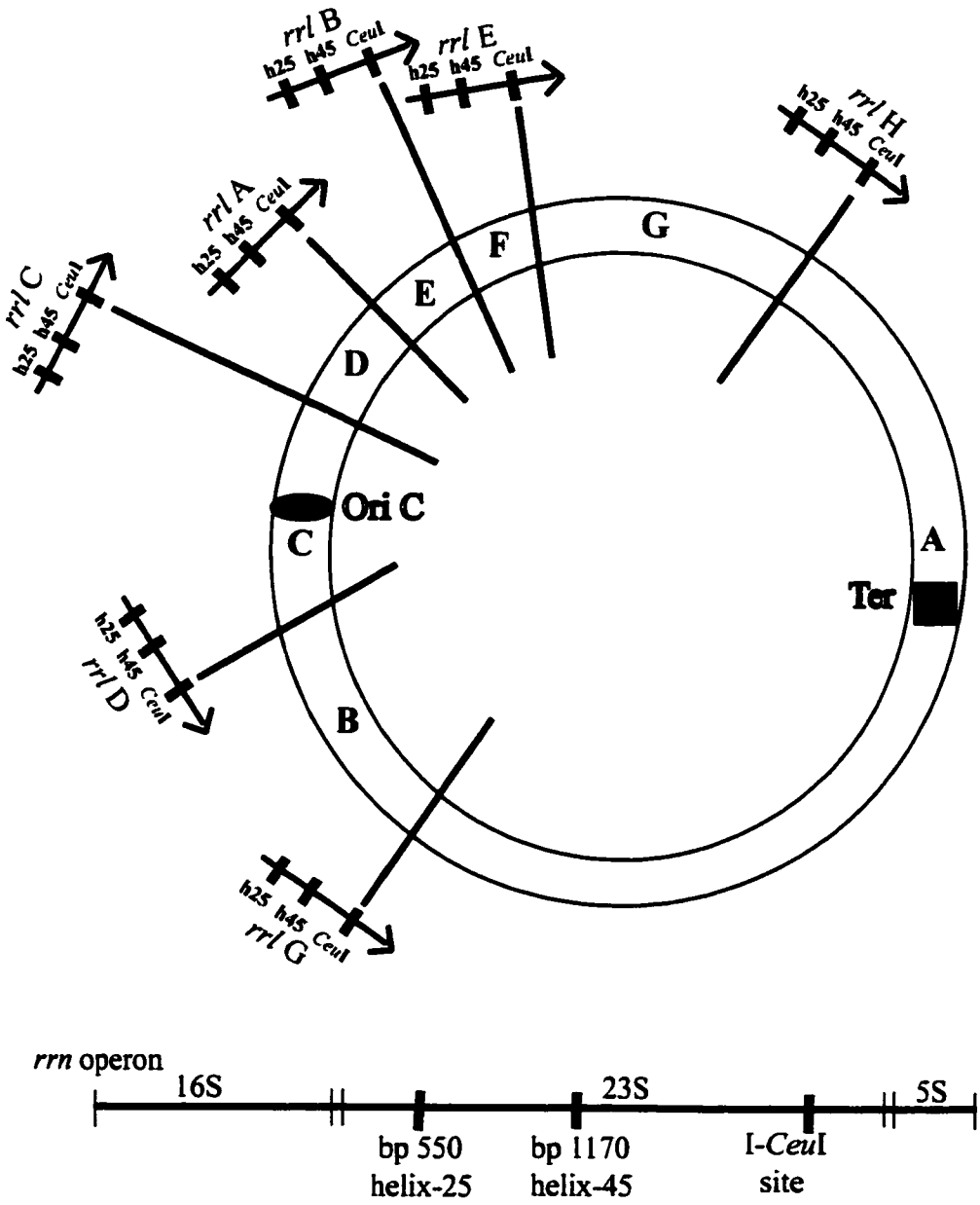

rrl B
h25 h45 CeuI
rrl E
h25 h45 CeuI
rrl H
h25 h45 CeuI
rrl A
h25 h45 CeuI
rrl C
h25 h45 CeuI
G
F
E
D
Ori C
C
A
Ter
B
rrl D
h25 h45 CeuI
rrl G
h25 h45 CeuI
rrn operon
16S
23S
5S
bp 550
helix-25
bp 1170
helix-45
I-CeuI
site

Figure 4.8: Pulsed-field gel electrophoresis of I-*Ceu*I digested genomic DNA from strains of *Salmonella* from the SARB set.

Genomic DNA from *Salmonella* strains was digested with I-*Ceu*I and separated by PFGE, using a 0.7% agarose gel containing ethidium bromide; the fragments were visualized by UV light. The I-*Ceu*I fragments are indicated. A lambda ladder was run alongside so that the fragment sizes could be determined; the size of the λ multimers is indicated on the right. The following strains of *Salmonella* were tested: Lane 1, *S. decatur* ET1; 2, *S. dublin* ET3; 3, *S. enteritidis* ET1; 4, *S. enteritidis* ET7; 5, *S. emek* ET1; 6, *S. indiana* ET1; 7, *S. miami* ET1; 8, *S. panama* ET12; 9, *S. pullorum* ET3; 10, *S. typhisuis* ET3.

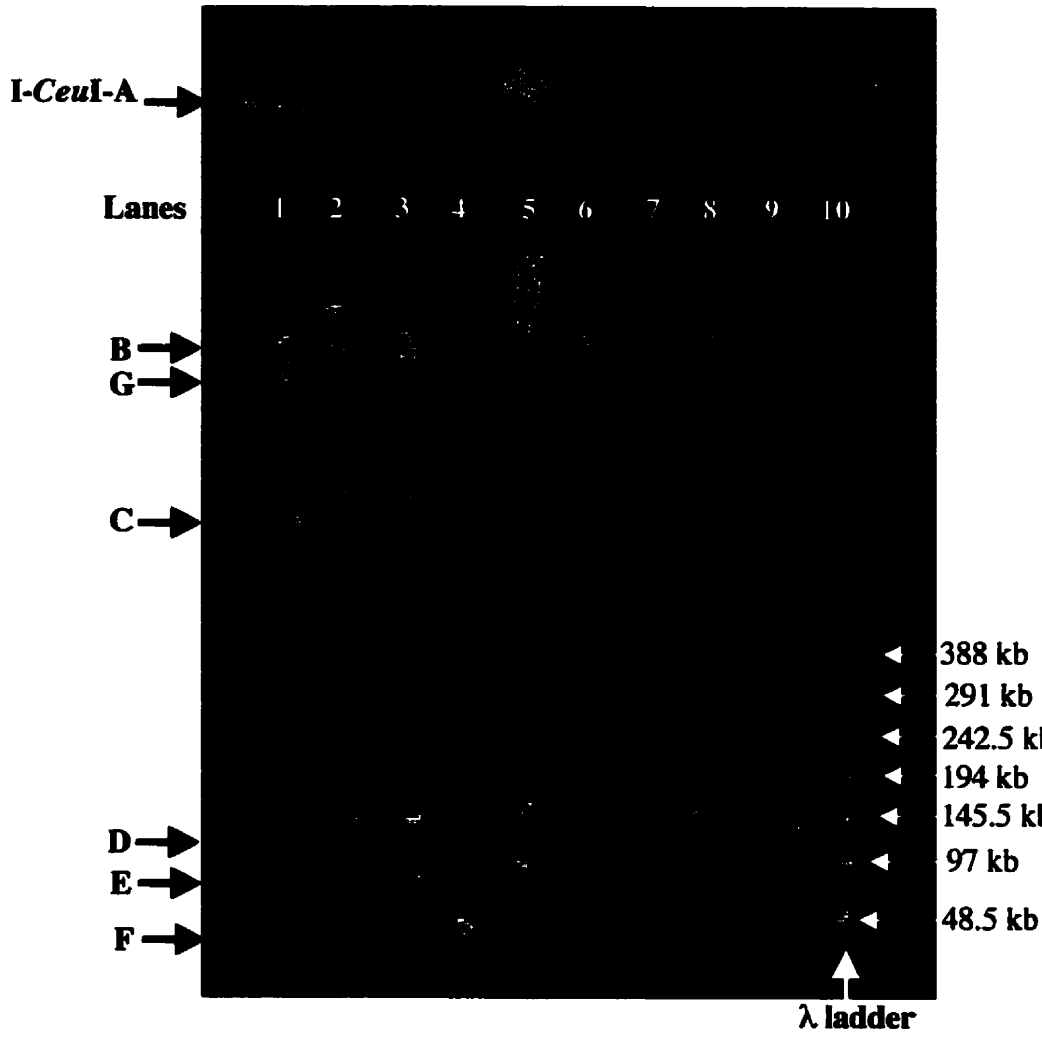

I-CeuI-A
Lanes
1 2 3 4 5 6 7 8 9 10
B
G
C
D
E
F
388 kb
291 kb
242.5 kb
194 kb
145.5 kb
97 kb
48.5 kb
λ ladder

Table4.1:Separation of the Helix-25 IVSs of *Salmonella* into families based on nucleotide identity.

% nucleotide identity [a] with Helix-25 IVS of:																	
IVS Family [b]	**A**												**B**				**C**
Strain(ET)[c]/*rrl* gene	Tm 1	Tp 1	Dt 1	Dt 1	Id 1	Mi 1	Du 3	Ts 3	Ts 3	Ba 2	De 1	Mu 4	Tm 1	Cs 1	Mu 3	En 7	En 1
Tm1(H)	100	96	98	97	96	98	97	97	98	97	96	96	65	65	66	66	58
Tp1(all)	96	100	98	98	95	98	97	98	98	96	98	97	66	66	66	66	58
Dt 1(G,A)	98	98	100	99	96	100	99	99	100	98	98	99	65	65	65	65	59
Dt 1(B,E)	97	98	99	100	95	99	98	100	99	100	97	98	65	66	66	66	58
Id 1(D/C)[d]	96	95	96	95	100	96	95	95	96	94	94	95	66	67	66	67	57
Mi 1(D/C,H)[d]	98	98	100	99	96	100	99	99	100	98	98	99	65	65	65	65	59
Du 3(H,B)	97	97	99	98	95	99	100	98	99	97	97	98	65	65	65	65	59
Ts 3(D/C,A,B)[d]	97	98	99	100	95	99	98	100	99	97	98	98	65	66	66	66	58
Ts 3(G,E)	98	98	100	99	96	100	99	99	100	98	98	99	65	65	65	65	59
Ba 2[e]	97	96	98	100	94	98	97	97	98	100	98	99	64	64	64	64	58
De 1[e]	96	98	98	97	94	98	97	98	98	98	100	99	64	65	64	65	58
Mu 4[e]	96	97	99	98	95	99	98	98	99	99	99	100	64	65	65	65	58
Tm 1(G)	65	66	65	65	66	65	65	65	65	64	64	64	100	99	96	99	65
Cs 1[e]	65	66	65	66	67	65	65	66	65	64	65	65	99	100	97	100	63
Mu 3[e]	66	66	65	66	66	65	65	66	65	64	64	65	96	97	100	97	63
En 7(A,H)	66	66	65	66	67	65	65	66	65	64	65	65	99	100	97	100	63
En 1(G/H)[d]	58	58	59	58	57	59	59	58	59	58	58	58	65	63	63	63	100

[a] DNASIS was used to align sequences for identity comparisons.

[b] The IVSs are said to belong to the same family when they are 90% or more identical.

[c] The strain studied is identified by the Electrophoretic type (ET), the species is listed in Table 3.1; e.g. , Tm1 is *S.typhimurium* ET1. The letters in brackets show the *rrl* gene inferred from their presence on specific I-*Ceu* I fragments from PFGE (see section 4.4).

[d] IVSs from *rrl* G/*rrl* H and those from *rrl* D/*rrl* C cannot be distinguished as described in section 4.4.

[e] These strains contain only one *rrl* gene with an IVS; the DNA with the gene for the one IVS-containing *rrl* gene was purified. as described in Materials and Methods under sequencing strategies.

Table 4.2: Comparison of percent nucleotide identity[a] between IVSs from helix-25 of the 23S rRNA of *Salmonella*[b], *Proteus* and *Providencia*[c].

IVS Family [d]	A [e] *Salmonella*	B [e] *Salmonella*	C [e] *Salmonella*	D [f] *Proteus*	E [f] *Proteus*	F [g] *Providencia*	G [f] *Proteus*
A (25/10) [h]	93-100	59-63	52-54	73-74	50	48-49	49
B (5/4) [h]	59-63	95-100	57-60	61	48	47	48
C (1/1) [h]	52-54	57-60	100	54	53	54	52
D (4/1) [h]	73-74	61	54	100	50	49	49-50
E (2/1) [h]	50	48	53	50	100	85	85-86
F (1/1) [h]	48-49	47	54	49	85	100	80-81
G (11/3) [h]	49	48	52	49-50	85-86	80-81	93-100

[a] Percent identity comparisons were done using DNASIS Version 2 for Macintosh.

[b] The *Salmonella* strains used for sequencing are listed in Table 1.

[c] The *Proteus* and *Providencia* families are reported in Miller et al., 2000.

[d] A family consists of IVS sequences with more than 90% nucleotide identity.

[e] Represent the different families from *Salmonella*.

[f] Represent the different families from *Proteus* (Accession numbers for the sequences are mentioned in Fig. 4.10).

[g] Represent the family from *Providencia* (Accession number for the sequence is mentioned in Fig. 4.10).

[h] The numerator of the fraction in parentheses indicates the number of sequenced IVSs, each isolated from an individual *rrl* gene and the denominator represents the number of strains from which these IVSs were sequenced.

4.5 Analysis of nucleotide sequences in helix-45 IVSs.

Nucleotide sequence analysis was done for helix-45 IVSs as described in section 4.4 for helix-25 IVSs. Sequence analysis of the helix-45 IVSs revealed the presence of two families, M and O which are 90% or more identical within the families and 80-84% identical between families (Table 4.3). Fourteen different IVS sequences from 4 different strains were analyzed in family M and 25 IVS sequences from 8 strains were analyzed in family O.

The helix-45 IVSs determined in this study in *Salmonella* were compared with previously determined IVS sequences. The helix-45 IVS from *Yersinia enterocolitica* group II (renamed family N in this study) was about 86% identical to family M and 80% identical to family O but only 60% identical to *Yersinia enterocolitica* group I IVS (renamed family P) (Table 4.4). The helix-45 IVSs from *Providencia rettgeri* are assigned to family Q and those from *Proteus vulgaris* to family R. The two families share 78% nucleotide similarity. Families Q and R are 65-73% similar to families M and O from *Salmonella.*

The sequences of helix-25 and helix-45 IVSs were determined in 22 strains from the SARB set to analyze if closely related strains possess IVSs from the same family, and to study if all IVSs from the seven *rrl* genes of a strain are similar. Presence of the same IVS family in closely related strains is suggestive of vertical transfer during clonal propagation. The presence of distantly related families of IVSs in closely related strains indicates horizontal transfer of IVSs. The IVS families are shown for the strains where they were determined (Fig. 4.9). For both helix-25 and helix-45 IVSs, closely related strains were seen to possess IVSs from the same family in most of the cases, however

closely related strains in some cases had IVSs from different families. All the *Salmonella* strains from which IVSs were sequenced showed the presence of similar sequences in the different *rrl* genes of a specific strain. The maximum sequence divergence seen between the different IVSs from a strain was 3 bases. The only exception to this is seen *in S. typhimurium* where the two helix-25 IVSs from *rrl*G (family A) and *rrl*H (family B) are only 56% identical (Mattatall and Sanderson, 1996).

Table 4.3: Separation of Helix-45 IVSs of *Salmonella* and *Yersinia* into families based on nucleotide identity.

%nucleotide identity [a] with Helix-45 IVSs of:															
IVS Family [b]	M						O								
Strain(ET)[c]/*rrl* gene	Tm1	Cs13	Cs13	Cs13	Id1	Pb5	Pb1	Ga2	Pu3	Pu4	Du3	Em1	Em1	Pn1	En7
Tm1(A)	100	100	99	98	99	98	84	84	84	84	84	83	84	84	83
Cs13(B)	100	100	99	98	99	98	84	84	84	84	84	83	84	84	83
Cs13(G,H,E)	99	99	100	99	98	97	83	83	83	83	83	82	83	83	82
Cs13(D,C,A)	98	98	99	100	97	96	82	82	82	82	82	80	82	82	80
Id1(B,D/C) [d]	99	99	98	97	100	99	84	84	84	84	84	83	84	84	83
Pb5 [e]	98	98	97	96	99	100	83	83	83	83	83	82	83	83	82
Pb1(all) [f]	84	84	83	82	84	83	100	100	100	100	100	99	100	100	99
Ga2(all) [f]	84	84	83	82	84	83	100	100	100	100	100	99	100	100	99
Pu3(all) [f]	84	84	83	82	84	83	100	100	100	100	100	99	100	100	99
Pu4(all) [f]	84	84	83	82	84	83	100	100	100	100	100	99	100	100	99
Du3(A)	84	84	83	82	84	83	100	100	100	100	100	99	100	100	99
Em1(G/H) [d]	83	83	82	80	83	82	99	99	99	99	99	100	99	99	94
Em1(B)	84	84	83	82	84	83	100	100	100	100	100	99	100	100	99
Pn1 [e]	84	84	83	82	84	83	100	100	100	100	100	99	100	100	99
En7(G)	83	83	82	80	83	82	99	99	99	99	99	94	99	99	100

[a] DNASIS was used to align sequences for identity comparisons.

[b] The IVSs are said to belong to the same family when they are 90% or more identical.

[c] The strain studied is identified by electrophoretic type (ET), the species is listed in Table 3.1; e.g., Tm1 is *S.typhimurium* ET1. The letters in brackets show the *rrl* gene inferred from their presence on specific I-*Ceu* I fragments from PFGE (see section 4.4).

[d] IVSs from *rrl* G/*rrl* H and those from *rrl* D/*rrl* C cannot be distinguished as described in section 4.4.

[e] These strains contain only one *rrl* gene with an IVS; and was purified as described in Materials and Methods under sequencing strategies.

[f] All the IVSs from the seven *rrl* genes were identical in these strains.

Table 4.4: Comparison of percent nucleotide identity[a] between IVSs from helix-45 of the 23S rRNA of *Salmonella*[b], *Yersinia*[c], *Proteus*[d] and *Providencia*[d].

IVS Family[e]	M[f] *Salmonella*	N[g] *Yersinia II*	O[f] *Salmonella*	P[h] *Yersinia I*	Q[i] *Providencia*	R[j] *Proteus*
M(14/4)[k]	96-100	85-87	80-84	61-62	65-66	68-69
N(1/1)[k]	85-87	100	78-80	60	67	66
O(25/8)[k]	80-84	78-80	94-100	65-67	69-70	70-73
P(1/1)[k]	61-62	60	65-67	100	62	64
Q(2/1)[k]	65-66	67	69-70	62	100	77-78
R(7/2)[k]	68-69	66	70-73	64	77-78	99-100

[a] Percent identity comparisons were done using DNASIS Version 2 for Macintosh.

[b] The *Salmonella* strains used for sequencing are listed in Table 1.

[c] The sequences from *Yersinia* are reported by Skurnik and Toivanen, 1991.

[d] The *Proteus* and *Providencia* families are reported by Miller et al., 2000.

[e] A family consists of IVS sequences with more than 90% nucleotide identity.

[f] Represent the different families from *Salmonella*.

[g] Represents the IVS family from *Yersinia enterocolitica* GroupII (Accession number for the sequence is mentioned in Fig. 4.11).

[h] Represents the IVS family from *Yersinia enterocolitica* GroupI (Accession number for the sequence is mentioned in Fig. 4.11).

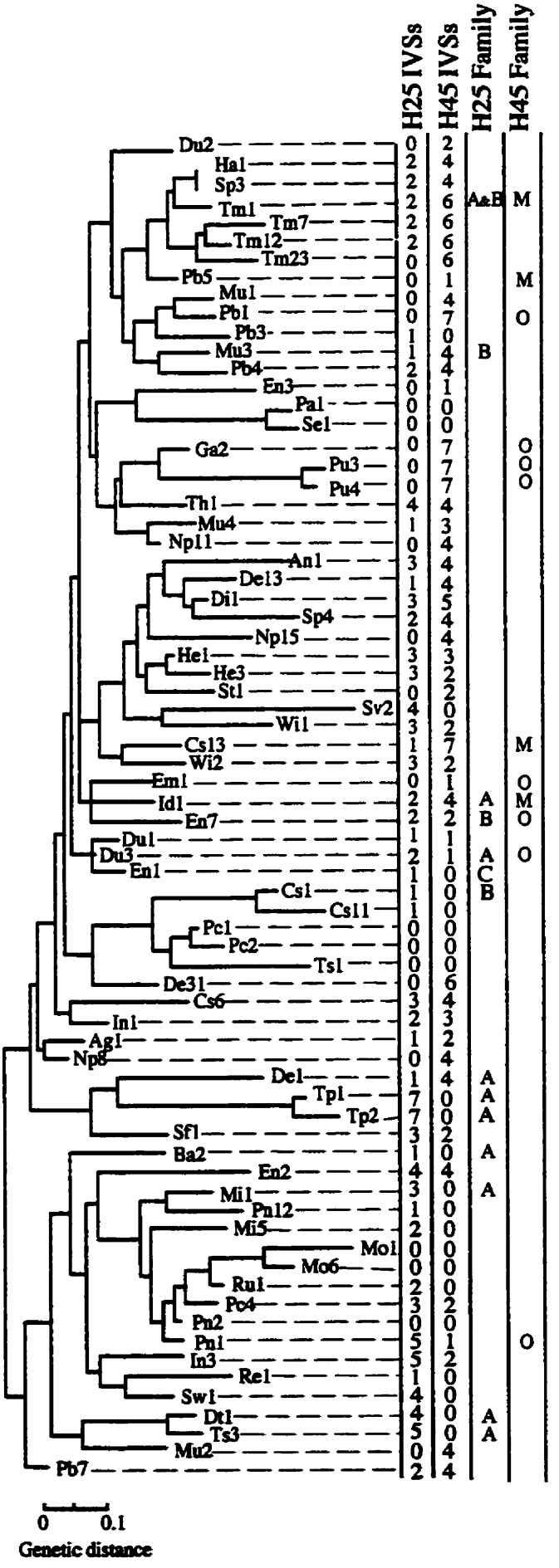
[i] Represents the family from *Providencia* (Accession number for the sequence is mentioned in Fig. 4.11).

[j] Represents the family from *Proteus* (Accession numbers for the sequences are mentioned in Fig. 4.11).

[k] The numerator of the fraction in parentheses indicates the number of sequenced IVSs, each isolated from an individual *rrl* gene and the denominator represents the number of strains from which these IVSs were sequenced.

Figure 4.9: Distribution of helix-25 and helix-45 IVS families among the SARB strains.

The relationships of 72 *Salmonella* strains of the SARB set determined by MLEE (Boyd et al., 1993) and the distribution of helix-25 (H25) and helix-45 (H45) IVSs is shown, as previously illustrated in Fig. 4.8. The *Salmonella* species, electrophoretic types and SARB numbers are listed in Table 3.1. The families of the IVSs are indicated for the strains from which IVSs were sequenced (tables 4.1, 4.2, 4.3, 4.4). The presence of the same family as well as different families in closely related strains is seen.


H25 IVSs
H45 IVSs
H25 Family
H45 Family
Du2 0 2
Ha1 2 4
Sp3 2 4
Tm1 2 6 A&B M
Tm7 2 6
Tm12 2 6
Tm23 0 6
Pb5 0 1 M
Mu1 0 4
Pb1 0 7 O
Pb3 1 0
Mu3 1 4 B
Pb4 2 4
En3 0 1
Pa1 0 0
Se1 0 0
Ga2 0 7 O
Pu3 0 7 O
Pu4 0 7 O
Th1 4 4
Mu4 1 3
Np11 0 4
An1 3 4
De13 1 4
Di1 3 5
Sp4 2 4
Np15 0 4
He1 3 3
He3 3 2
St1 0 2
Sv2 4 0
Wi1 3 2
Cs13 1 7 M
Wi2 3 2
Em1 0 1 O
Id1 2 4 A M
En7 2 2 B O
Du1 1 1
Du3 2 1 A O
En1 1 0 C
Cs1 1 0 B
Cs11 1 0
Pc1 0 0
Pc2 0 0
Ts1 0 0
De31 0 6
Cs6 3 4
In1 2 3
Ag1 1 2
Np8 0 4
De1 1 4 A
Tp1 7 0 A
Tp2 7 0 A
Sf1 3 2
Ba2 1 0 A
En2 4 4
Mi1 3 0 A
Pn12 1 0
Mi5 2 0
Mo1 0 0
Mo6 0 0
Ru1 2 0
Pc4 3 2
Pn2 0 0
Pn1 5 1 O
In3 5 2
Re1 1 0
Sw1 4 0
Dt1 4 0 A
Ts3 5 0 A
Mu2 0 4
Pb7 2 4
0 0.1
Genetic distance

4.6 Sequence alignment studies.

All sequences from the different helix-25 and helix-45 families of *Salmonella*, *Proteus*, *Providencia* and *Yersinia* were aligned using CLUSTAL X as shown in Fig 4.10 and Fig 4.11. All the sequences possess an imperfect terminal inverted repeat which forms the primary stem in the secondary structure and the rest of the sequence folds into secondary stems and loops (see section 4.7). An alignment of sequences belonging to the same family shows that the sequence conservation within a family is strong (90% or more as shown in Tables 4.1 and 4.3) and the sequence conservation between different families resides mainly in the primary stem rather than in the secondary stems and loops (Fig. 4.10 and Fig.4.11). With the exception of a few base changes in the region of the primary stem discussed in section 4.7, this region is largely conserved as seen in the alignments in Figs. 4.10 and 4.11. Indel (insertions and deletions) type of mutations are not seen in this region. However, in the secondary stems and loops, indels as well as base pair changes define the differences between the families; for example families A and B (sizes vary from 113-114 bp) differ from family C (total size 151 bp) by 8 deletions which vary from 1 to 16 bases in length (Fig 4.10a). CLUSTAL alignments of family A of *Salmonella* and family D of *Proteus* (Fig 4.10b) indicate a high degree of identity between these sequences from different genera. Family D of *Proteus* has many deletions compared with family G of *Proteus* (Fig. 4.10c). An alignment of helix-25 families of *Salmonella* (A, B, C), *Proteus* (D, E, G) and of *Providencia* (F) (Fig 4.10d) reveals that the only sequence conservation seen resides in the region of the primary stem.

Sequence similarity based on CLUSTALX alignments between the helix-45 IVSs from *Salmonella*, *Proteus*, *Providencia* and *Yersinia* is shown in Fig. 4.11. Families M

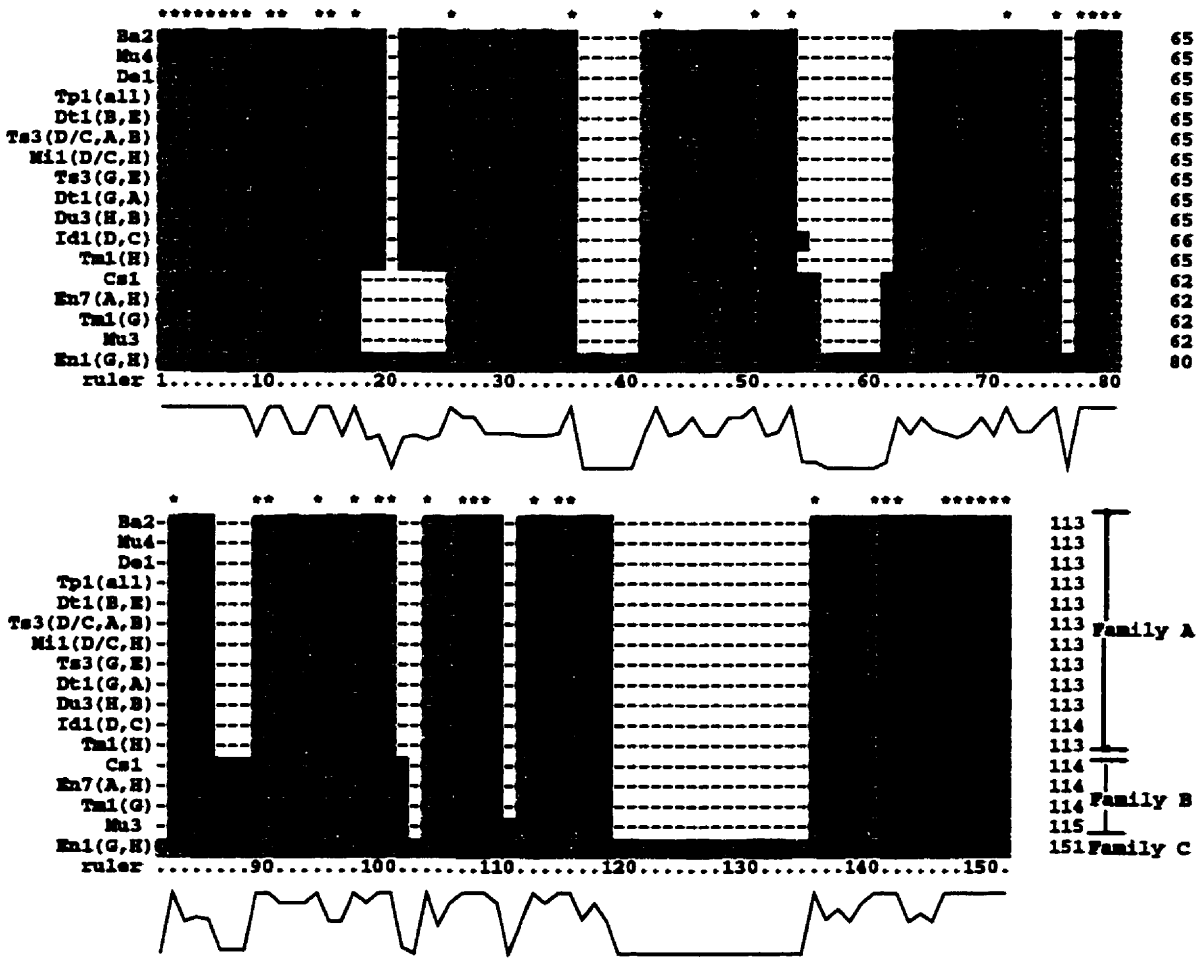
and O of *Salmonella* differ only by a few base changes and indels (Fig. 4.11a). Fig. 4.11b shows a high degree of sequence identity for helix-45 families from *Salmonella* (M) and *Yersinia* (N) and the lack of identity between IVSs of *Yersinia* GroupI (family P) and GroupII (family N) is shown in Fig. 4.11c. CLUSTAL alignments for *Salmonella* (M), *Providencia* (Q) and *Proteus* (R) (Fig. 4.11d) indicate a high degree of sequence conservation. The difference between the three families is characterized by two segments, both of which are deleted from *Salmonella* (family M), one of the deleted sequences is also absent in *Proteus*, family R. Fig. 4.11e indicates the well conserved stem region in the 6 different helix-45 families identified in the *Enterobacteriaceae*. The primary stem has been proven to be the region which is recognized and cleaved by RNase III (Burgin et al., 1990) and hence its conservation is suggestive of a positive selective pressure for IVS excision.

Figure 4.10: Sequence alignments for helix-25 IVSs from individual *rrl* genes for *Salmonella*, *Proteus*, and *Providencia*.

Sequences from individual *rrl* genes were obtained as described in Fig. 4.7 and Fig. 4.8 and aligned using CLUSTAL X. Each sequence is identified by the electrophoretic type for *Salmonella* and by the strain number for *Proteus* and *Providencia*. The letters in parentheses indicate the *rrl* gene from which the IVS was isolated for *Salmonella*. For *Proteus* and *Providencia* the I-*Ceu*I fragment is shown in parentheses. Asterisks above the sequence indicate absolutely conserved bases, and numbers at the end of each row indicate the number of bases in each IVS. The family of the IVS was determined as described in section 4.4 and is shown at the end of the sequence. a) All helix-25 IVSs from *Salmonella*; b) Family A of *Salmonella* and family D of *Proteus*; c) Families D and G of *Proteus*; d) helix-25 IVSs from representatives of family A, B and C of *Salmonella*, families D, E & G of *Proteus* and family F of *Providencia*;

The GenBank Accession numbers for the *Proteus* and *Providencia* sequences are as follows; AF176794: SA5564 (G, C, D, F); AF176790: SA5474 (D,G); AF176785: SA5461 (C, D); AF176791: SA5473 (B); these sequences were previously determined in this laboratory (Miller et al., 2000)

a.

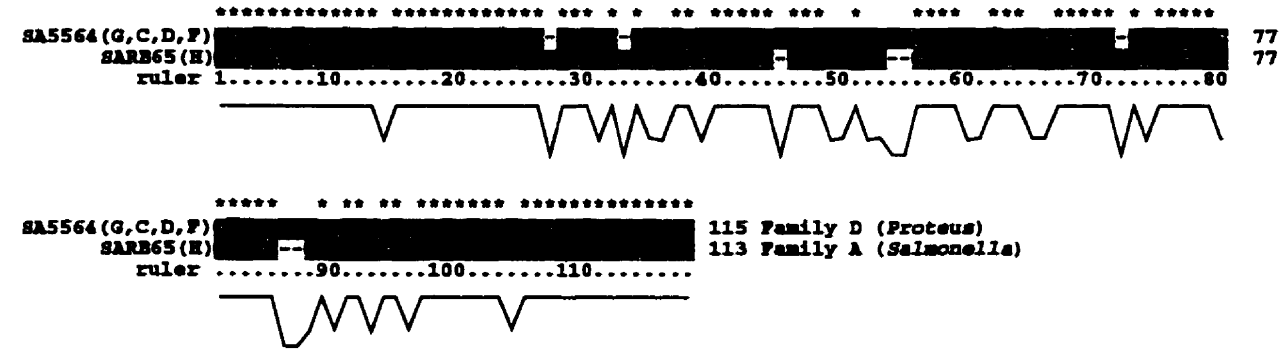


Family A *(Salmonella)*

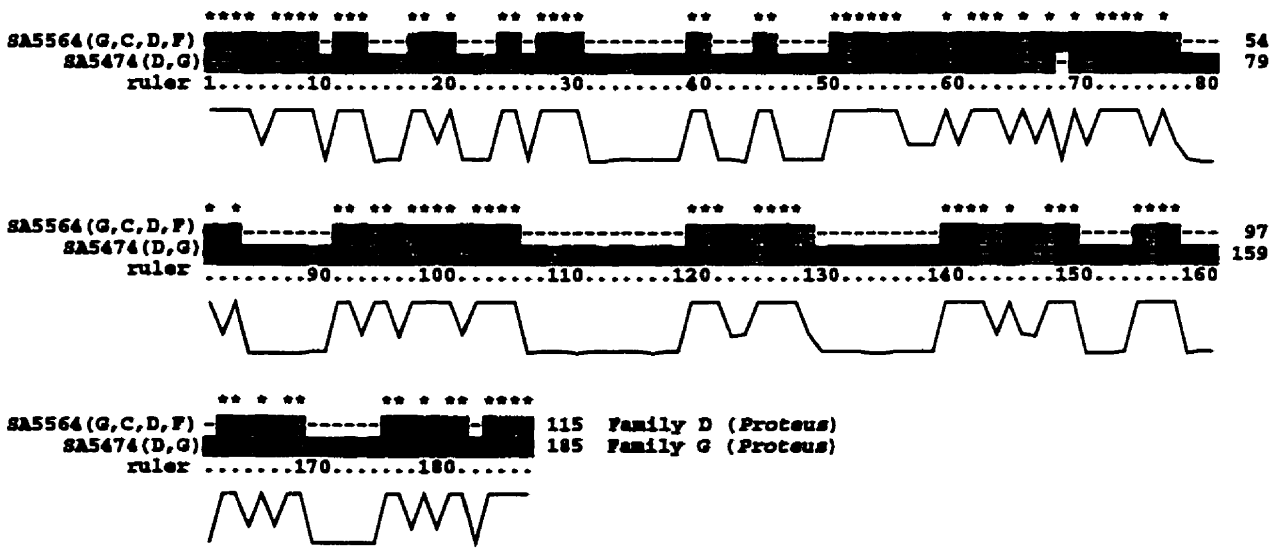
Family B *(Salmonella)*

Family C *(Salmonella)*

b.


SA5564(G,C,D,F)
SARB65(H)
ruler
77
77
115 Family D (Proteus)
113 Family A (Salmonella)

c.


SA5564(G,C,D,F)
SA5474(D,G)
ruler
54
79
97
159
115 Family D (Proteus)
185 Family G (Proteus)

d.

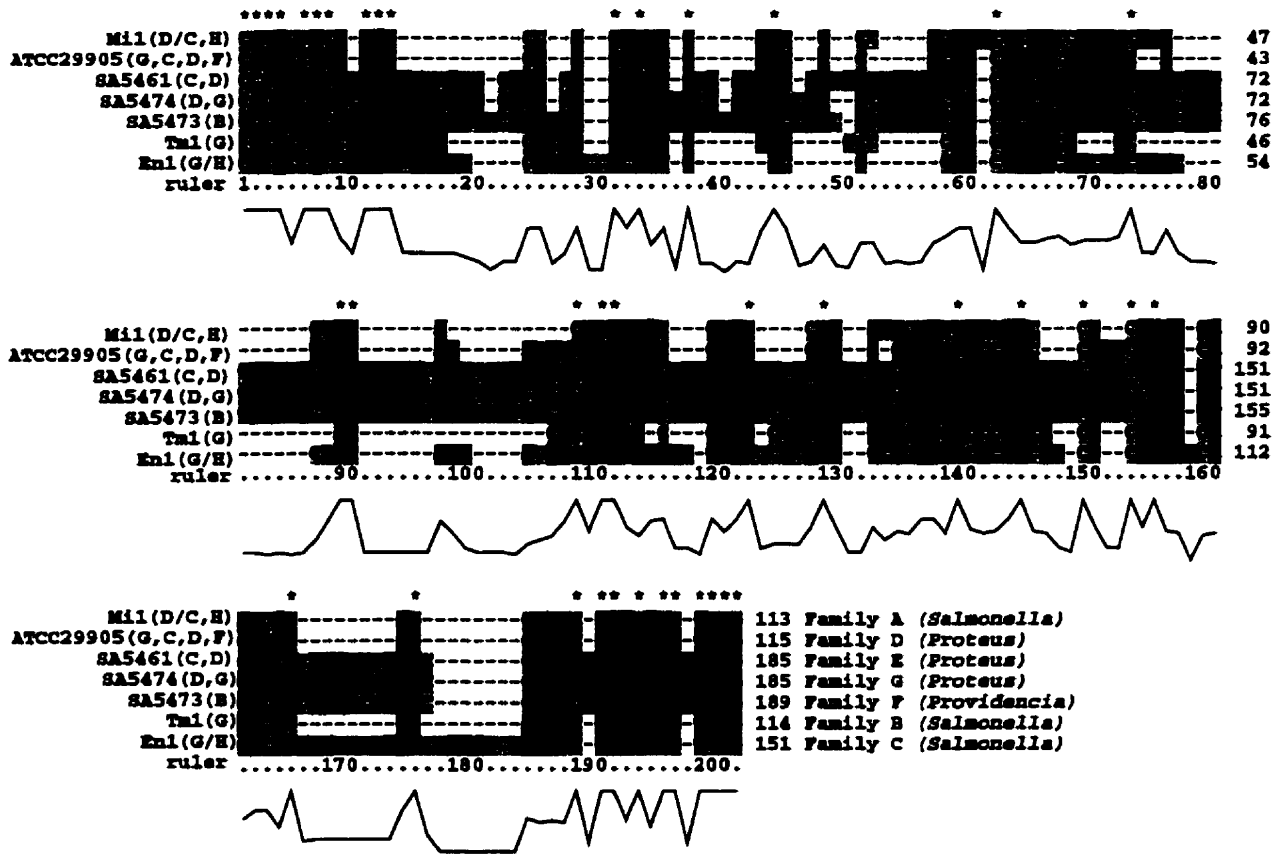
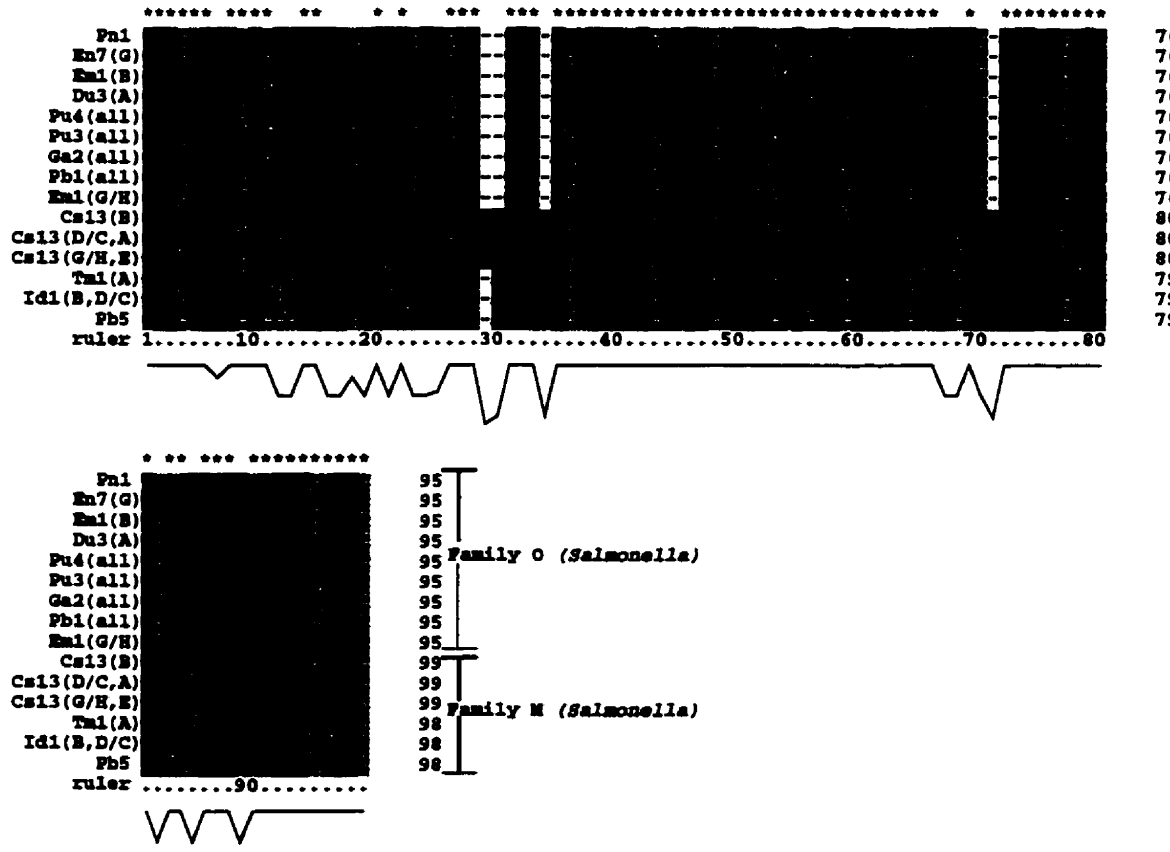

Mi1(D/C,E) 47
ATCC29905(G,C,D,F) 43
SA5461(C,D) 72
SA5474(D,G) 72
SA5473(B) 76
Tm1(G) 46
En1(G/E) 54
ruler 1.....10.....20.....30.....40.....50.....60.....70.....80
Mi1(D/C,E) 90
ATCC29905(G,C,D,F) 92
SA5461(C,D) 151
SA5474(D,G) 151
SA5473(B) 155
Tm1(G) 91
En1(G/E) 112
ruler .....90.....100.....110.....120.....130.....140.....150.....160
Mi1(D/C,E) 113 Family A (Salmonella)
ATCC29905(G,C,D,F) 115 Family D (Proteus)
SA5461(C,D) 185 Family E (Proteus)
SA5474(D,G) 185 Family G (Proteus)
SA5473(B) 189 Family F (Providencia)
Tm1(G) 114 Family B (Salmonella)
En1(G/E) 151 Family C (Salmonella)
ruler .....170.....180.....190.....200.

Figure 4.11: Sequence alignments for helix-45 IVSs from individual *rrl* genes for *Salmonella*, *Proteus*, *Providencia*, and *Yersinia*.

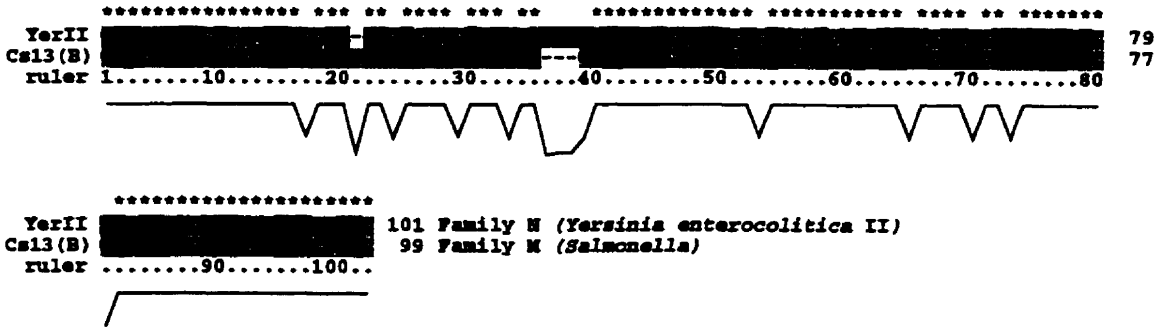
Sequences from individual *rrl* genes were obtained as described in Fig. 4.7 and Fig. 4.8 and aligned using CLUSTAL X. Each sequence is identified by the electrophoretic type for *Salmonella*, by the strain number for *Proteus* and *Providencia* and as YerGrpI & YerGrpII for *Yersinia enterocolitica*. The letters in parentheses indicate the *rrl* gene from which the IVS was isolated for *Salmonella*. For *Proteus* and *Providencia* the I-*Ceu*I fragment is shown in parentheses. Asterisks above the sequence indicate absolutely conserved bases, and numbers at the end of each row indicate number of bases in the sequence. The family of the IVS was determined as described in section 4.5 and is shown at the end of the sequence. a) All helix-45 IVSs from *Salmonella*; b) Family M of *Salmonella* and family N of *Yersinia* GrpII; c) Families N and O of *Yersinia*; d) Family M of *Salmonella*, Q of *Providencia* and family R of *Proteus* e) Helix-45 IVSs from representatives of family M and O of *Salmonella*, families N and P of *Yersinia*, Q of *Providencia* and family R of *Proteus*.

The GenBank Accession numbers for the *Proteus* and *Providencia* sequences are as follows; AF176791: SA5473(B); AF176787: SA5474(E) (Miller et al., 2000). The GenBank accession number for *Y. enterocolitica* GroupI helix-45 IVS is M35805 and Group II is M35811 (Skurnik and Toivanen, 1991).

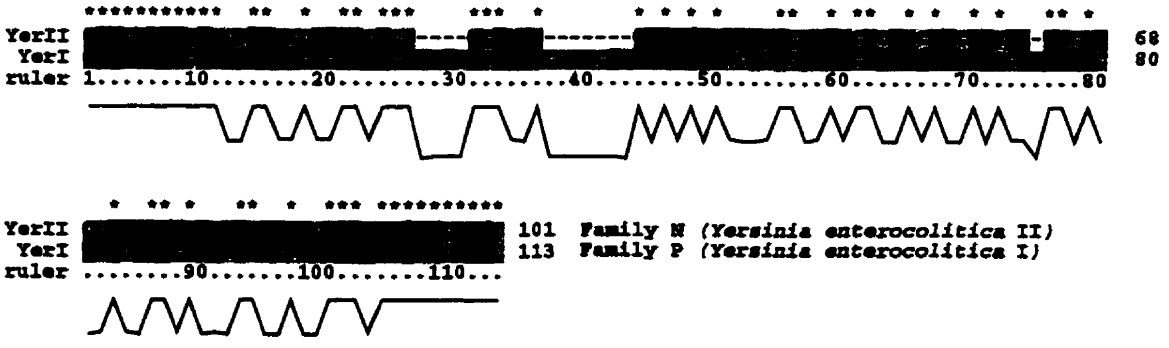
a.


Pn1
En7(G)
Em1(B)
Du3(A)
Pu4(all)
Pu3(all)
Ga2(all)
Pb1(all)
Em1(G/H)
Cs13(B)
Cs13(D/C,A)
Cs13(G/H,E)
Tm1(A)
Id1(B,D/C)
Pb5
ruler
1.....10.....20.....30.....40.....50.....60.....70.....80
76
76
76
76
76
76
76
76
76
80
80
80
79
79
79
90
95
99
98
Family O (Salmonella)
Family M (Salmonella)

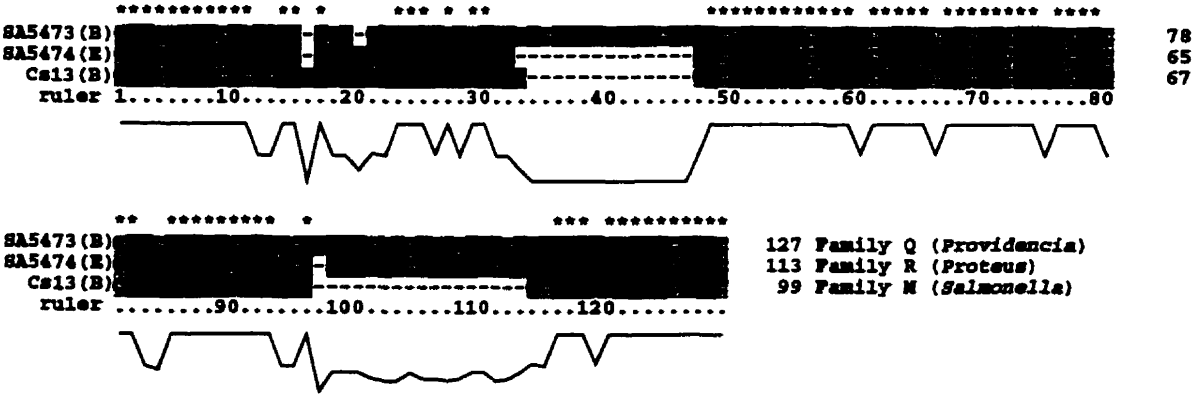
b.


YerII
Cs13(B)
ruler
79
77
101 Family N (Yersinia enterocolitica II)
99 Family M (Salmonella)

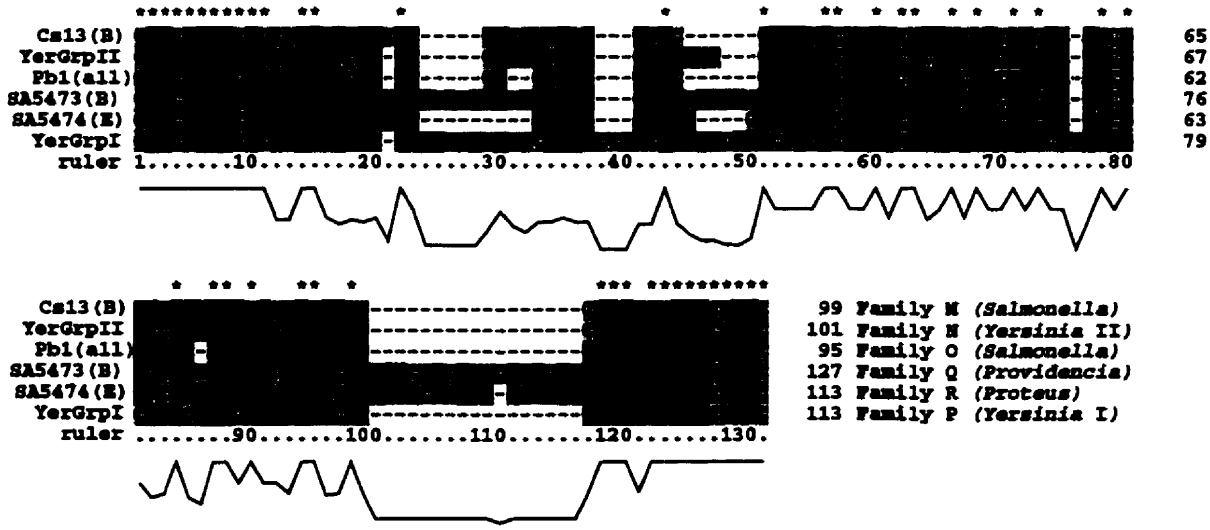
c.


YerII
YerI
ruler
68
80
101 Family N (Yersinia enterocolitica II)
113 Family P (Yersinia enterocolitica I)

d.


SA5473(B)
SA5474(E)
Cs13(B)
ruler 1.....10.....20.....30.....40.....50.....60.....70.....80
78
65
67
SA5473(B)
SA5474(E)
Cs13(B)
ruler .....90.....100.....110.....120.....
127 Family Q (Providencia)
113 Family R (Proteus)
99 Family M (Salmonella)

e.


Cs13(B)
YerGrpII
Pb1(all)
SA5473(B)
SA5474(E)
YerGrpI
ruler 1.....10.....20.....30.....40.....50.....60.....70.....80
65
67
62
76
63
79
Cs13(B)
YerGrpII
Pb1(all)
SA5473(B)
SA5474(E)
YerGrpI
ruler .....90.....100.....110.....120.....130.
99 Family M (Salmonella)
101 Family N (Yersinia II)
95 Family O (Salmonella)
127 Family Q (Providencia)
113 Family R (Proteus)
113 Family P (Yersinia I)

4.7 Secondary structures of IVSs.

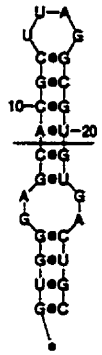
The rRNA secondary structure for all the sequences was predicted by free energy minimization using the algorithm of Zuker as implemented in the mfold program (Zuker et al., 1999). The extended stem-loop structure of the IVS replaces a tetraloop at the helix-25 and helix-45 locations from *E.coli*; a similar tetraloop is seen at these locations in the 23S rRNA of *Salmonella* when an IVS not present (fig. 4.12). Figs. 4.13 and 4.14 show the secondary structures of one representative of each of the helix-25 and helix-45 families. The RNA secondary structure reveals a primary stem composed of at least 16bp in both helix-25 and helix-45 which results from an imperfect inverted repeat at the DNA level. The sequences in the stem region are highly conserved; this conservation is also seen in the CLUSTAL alignment (Fig 4.10 for helix-25 IVSs and Fig 4.11 for helix-45 IVSs). The base pair changes between the families in the region of the primary stem are largely compensatory in nature wherein a base change at the DNA level resulting in a rRNA change which would disrupt pairing between two bases of the stem is usually associated with a corresponding change at the DNA level resulting in a change at the RNA level in the base it pairs with, restoring the Watson-Crick pairing between the bases. The difference between the primary stems of the helix-25 IVSs (Fig. 4.13) is characterized by 3 changes, 2 of which are compensatory in nature (a base pair change and a base pair deletion). One of the changes is a single base deletion leading to a pucker in the stem; changes of this type have been shown to enhance RNaseIII cleavage (Nicholson, 1999). Similarly compensatory mutations are common in helix-45 (Fig. 4.14); the difference between the primary stems of the helix-45 IVSs is characterized by

6 changes, 4 of which are compensatory base pair changes, one is a single base change and one is a multiple base change.

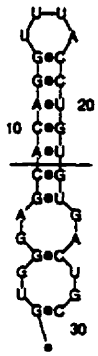
Figure 4.12: Proposed secondary structures for the helix-25 and helix-45 tetraloops from the 23S rRNA of *E. coli* and *S. typhimurium*.

The region of the helix-25 and helix-45 tetraloops at bp 550 and bp 1170 respectively is shown. The sequences for this region of the rRNA were obtained from the study of IVSs in *S. typhimurium* (Mattatall and Sanderson, 1996), and were confirmed in this study. The tetraloops are replaced by the extended stem-loop structure of the IVS when present. The free energy of folding is shown in kcal/mol. The horizontal line indicates the position where the tetraloop is replaced by the IVS.

Helix-25 tetraloops

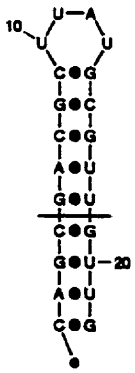


ΔG = -15.9kcal/mol
S.typhimurium

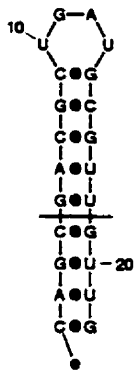


ΔG = -14.0kcal/mol
E.coli

Helix-45 tetraloops



ΔG = -12.3kcal/mol
S.typhimurium

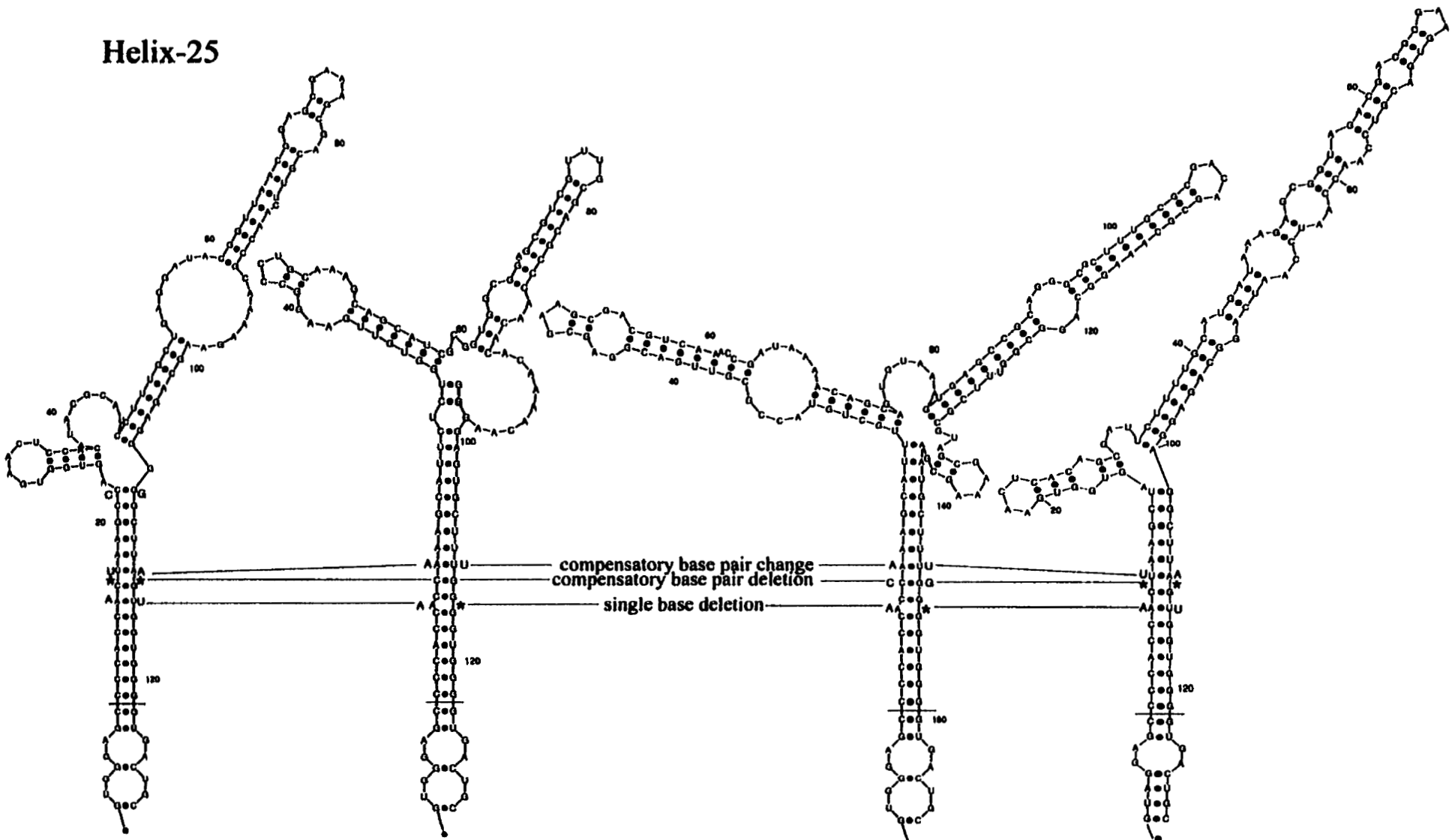


ΔG = -12.3kcal/mol
E.coli

Figure 4.13: Proposed secondary structures for representatives of helix-25 IVS families from *Salmonella* and *Proteus*.

Secondary structure predictions were made using the mfold program (Zuker et al., 1999). One representative from each of families A, B and C of *Salmonella* and family D of *Proteus* are shown. The free energy of folding is shown in kcal/mol and the length of the sequence in bases is indicated. The *rrl* gene from which the IVS was isolated is shown for IVSs from *Salmonella* and the I-*Ceu*I fragment is indicated for *Proteus*. The horizontal line on the IVS at the base of the stem indicates the position where the tetraloops shown in Fig. 4.12 are replaced by the extended stem-loop structure of the IVS. In an analysis of the base pair changes in the stem region of different families, the altered base pairs are shown and the type of mutation leading to the change are indicated to illustrate their compensatory or non-compensatory nature. All other base pairs in the region of the primary stem were conserved.

Helix-25



ΔG = -72.5kcal/mol
Family A(129 bases)
S.decatur (*rrl*G,A)

ΔG = -66.6kcal/mol
Family B (130 bases)
S.enteritidis (*rrl*A,H)

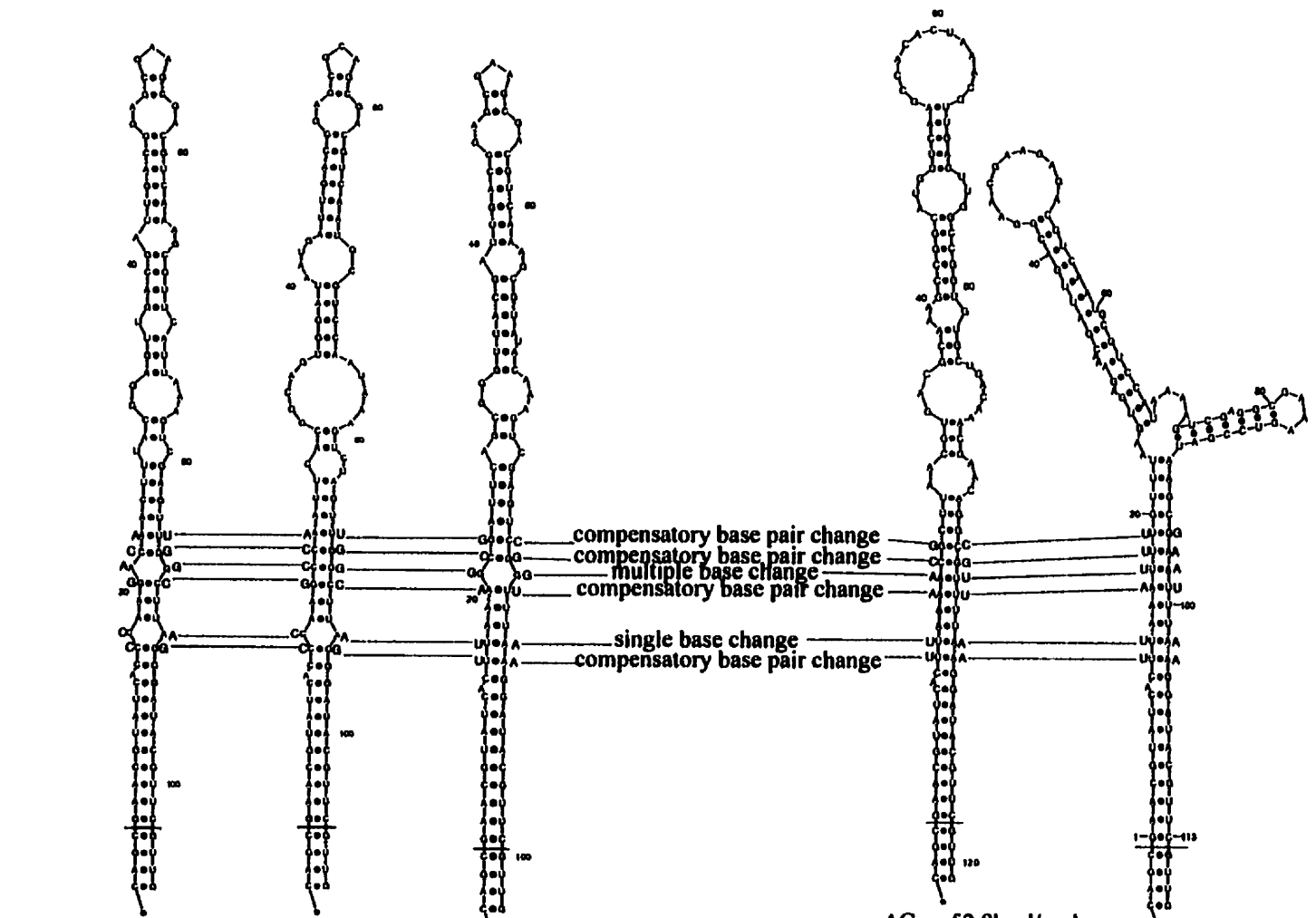
ΔG = -88.1kcal/mol
Family C (167 bases)
S.enteritidis (*rrl*G or H)

ΔG = -46.8kcal/mol
Family D (131 bases)
P.vulgaris (I-*Ceu*I G,C,D & F)

Figure 4.14: Proposed secondary structures for representatives of helix-45 IVS families from *Salmonella*, *Proteus* and *Yersinia*.

Secondary structure predictions were made using the mfold program (Zuker et al., 1999). One representative from each of families M and O of *Salmonella*, family N and P of *Yersinia* and family R of *Proteus* are shown. The free energy of folding is shown in kcal/mol and the length of the sequence in bases is indicated. The *rrl* gene from which the IVS was isolated is shown for IVSs from *Salmonella* and the I-*Ceu*I fragment is indicated for *Proteus*. A base pair comparison of the stem region was made for the IVSs shown. The altered base pairs are shown and the type of mutation leading to the change are indicated to illustrate their compensatory or non-compensatory nature. All other base pairs in the region of the primary stem were conserved. The horizontal line on the IVS at the base of the stem indicates the position where the tetraloops shown in Fig. 4.12 are replaced by the extended stem-loop structure of the IVS.

Helix-45


compensatory base pair change
compensatory base pair change
multiple base change
compensatory base pair change
single base change
compensatory base pair change
ΔG = -48.3kcal/mol
Family M (106 bases)
S.cholerasuis (rrlG,H,E)
ΔG = -51.7kcal/mol
Family N (109 bases)
Y.enterocolitica GrpII
ΔG = -51.8kcal/mol
Family O (103 bases)
S.paratyphiB (all rrl genes)
ΔG = -52.8kcal/mol
Family P (121 bases)
Y.enterocolitica GrpI
ΔG = -38.6kcal/mol
Family R (113 bases)
Proteus vulgaris (I-CeuI G)

CHAPTER 5: RESULTS- Role of IVSs in rRNA degradation

5.1 Growth curves and viability counts of *Salmonella* spp. and *E. coli*.

Cells of *E. coli*, *Salmonella agona* (SARB 1), *S. typhimurium* and SARC 10 (*Salmonella* of subspecies IV) were used in this study, the *Salmonella* strains used have a varying degree of IVS-facilitated rRNA fragmentation. 50ml of LB broth was inoculated with 0.5% inoculum and incubated at 37°C with aeration for 25 hours. The optical density of the culture was determined at 540nm using the Klett-Summerson colorimeter. The cells multiplied exponentially for 3 hours, after that the increase in optical density was slower and reached a maximum of 380-405 Klett units after about 9-11 hours of growth for the *Salmonella* spp. and *E. coli* used. This optical density was maintained for 15-17 hours of continued growth after which there was a steady but not sharp decline to 365-380 Klett units (Fig. 5.1) till 25 hours. Viable counts for the same cultures are shown in Fig.5.2 for a 25 hour period, wherein the total viable counts were determined on LB plates. For every time point 2 independent dilution series were done and 100µl of the dilution was plated on LB plates. Averages from these two dilution series is reported here. All the strains studied reached a maximum cell density of 6.2 x 10^9 CFU/ml to 9.2 x 10^9 CFU/ml after 13 hours of growth; on further growth there was a decline in the viability of the cells to 1 x 10^9 CFU/ml to 3 x 10^9 CFU/ml at 25 hours.

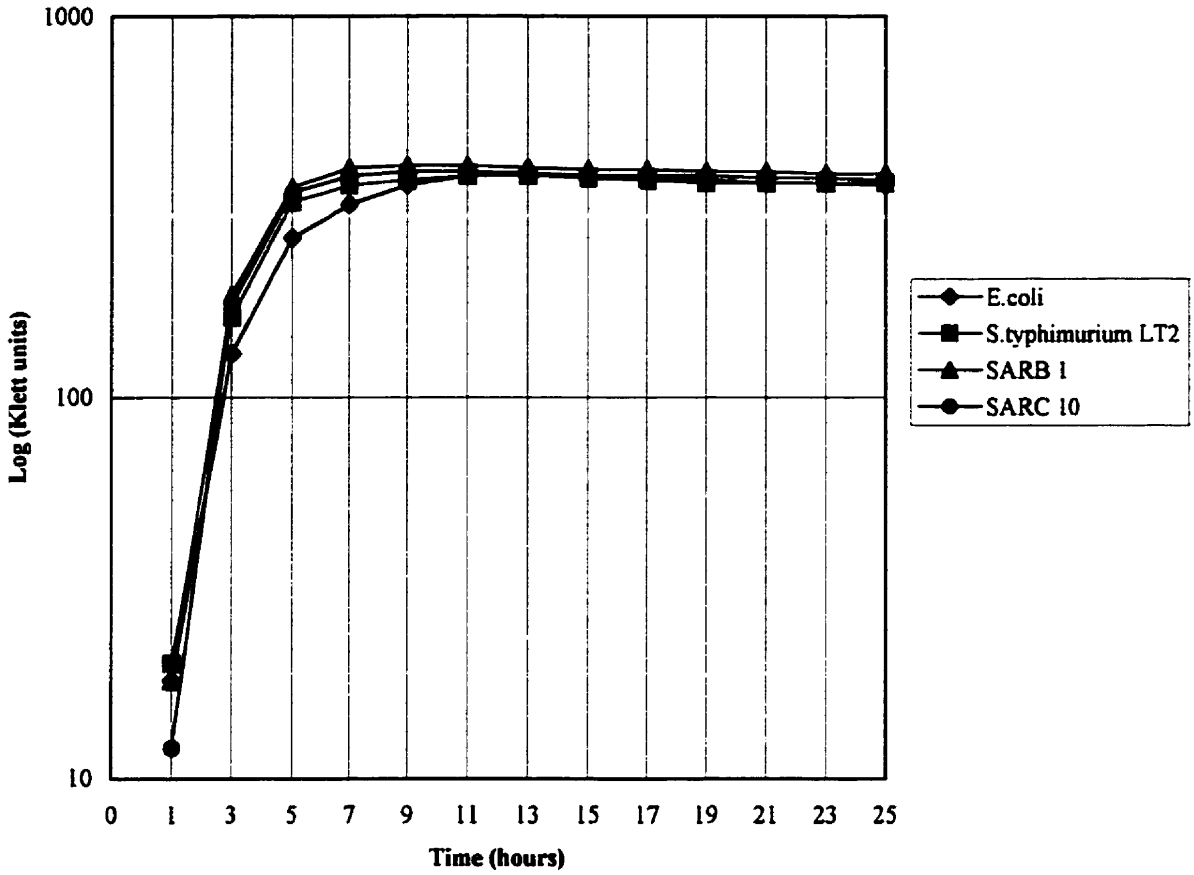
Figure 5.1: Growth curves for *E. coli*, *S. agona*, *S. typhimurium* and SARC 10.

0.5% inoculum from a 12 hour old culture obtained from a single colony isolate was inoculated into 50 ml of fresh LB broth in a Klett flask.

Panel A: Growth was monitored by measuring the O.D. every two hours at 540nm for 25 hours. The O.D units are plotted on a log scale on the Y-axis.

Panel B: To facilitate comparisons, the data recorded in panel A are shown in numbers.

A.



B.

Hours	Klett units			
	E.coli	*S.typhimurium* LT2	SARB 1	SARC 10
1	18	20	18	12
3	130	162	185	175
5	262	325	355	345
7	320	360	400	380
9	360	370	405	390
11	380	380	405	390
13	380	380	400	385
15	380	375	395	380
17	375	370	395	380
19	370	365	390	380
21	365	365	390	375
23	365	365	385	375
25	360	365	385	370

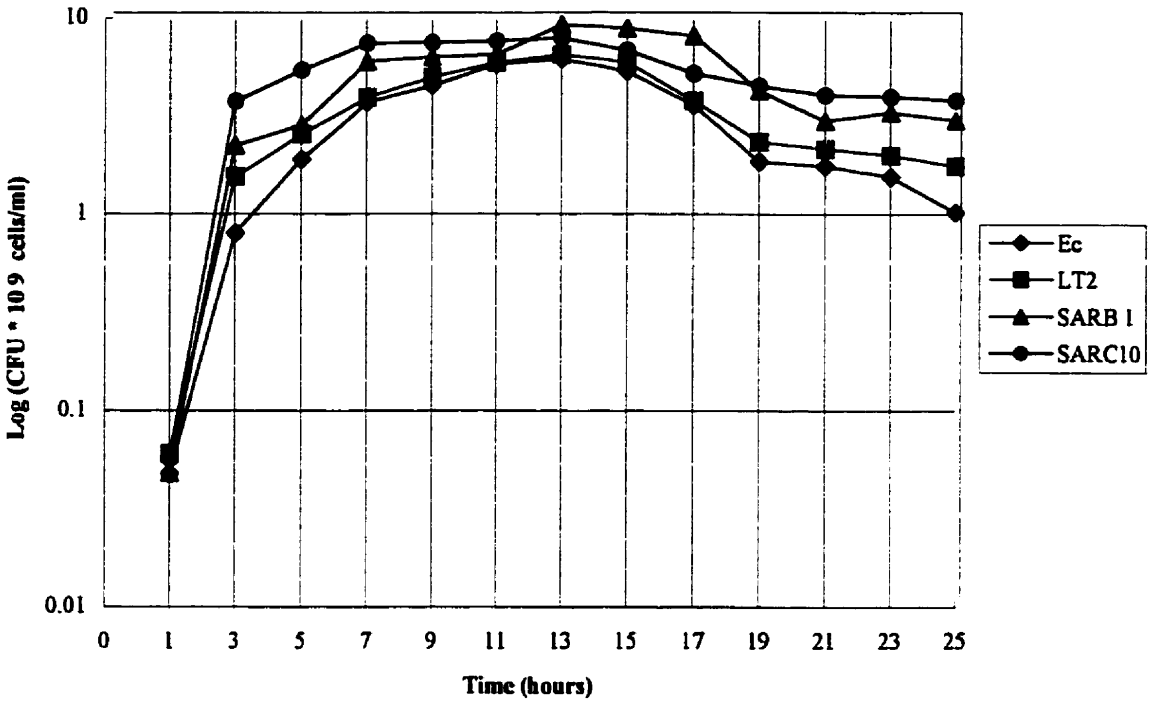
Figure 5.2: Total viable counts for *E. coli*, *S. agona*, *S. typhimurium* and SARC 10.

0.5% inoculum from a 12 hour old culture obtained from a single colony isolate was inoculated into 50 ml of fresh LB broth. Viable counts were monitored by plating appropriate dilutions in duplicate on LB plates and incubating the plates at 37°C overnight.

Panel A: The counts are represented as average CFU x 10^9 cells /ml on a log scale on the Y-axis.

Panel B: To facilitate comparisons, the data recorded in section A is shown in numbers.

A.



B.

Hours	CFU * 10^9 cells/ml			
	E. coli	*S. typhimurium* LT2	SARB 1	SARC10
1	0.0475	0.06	0.048	0.057
3	0.8	1.55	2.24	3.75
5	1.9	2.55	2.85	5.4
7	3.7	3.92	6	7.41
9	4.5	5	6.3	7.5
11	5.8	5.9	6.5	7.61
13	6.2	6.5	9.26	7.92
15	5.35	5.95	8.85	6.85
17	3.6	3.78	8.1	5.21
19	1.85	2.34	4.25	4.5
21	1.76	2.15	3	4.05
23	1.55	2	3.31	3.98
25	1.02	1.76	3.01	3.8

5.2 Degradation of rRNA during the different phases of growth.

Variation in rRNA content in strains of *Salmonella* and *E. coli* during growth in LB broth over a 25 hour period was monitored. 0.5% inoculum from a 12 hour culture was used and rRNA content was monitored every 2 hours at the same time points when the Klett readings were taken and viability counts were done. Fig. 5.3a represents the rRNA extracted from *E. coli.* The RNA was extracted by lysing cells from 5 ml of the culture using 1% SDS + heat (at 65°C for 5 min) and repeated phenol extractions. After three phenol extractions 2ml of the crude RNA extract was recovered as the aqueous phase. Each lane was loaded with 10 µl of the rRNA extract for the hours indicated on the figure. Figure 5.3b represents rRNA from *S. agona* extracted at the same time periods. *S. agona* has fragmented rRNA as a result of excision of a helix-25 IVS from one *rrl* gene and helix-45 IVSs from two *rrl* genes out of the seven *rrl* genes present in *Salmonella.* Figure 5.3c shows the rRNA extracted from *S. typhimurium* LT2 which has been shown to have 2 helix-25 and 6 helix-45 IVSs resulting in greater fragmentation of the rRNA as compared to *S. agona.* Figure 5.3d shows RNA isolated from SARC 10 (*Salmonella* Reference Collection C) representing *Salmonella* of subspecies IV; SARC 10 possesses the maximum number of IVSs detected in *Salmonella* with 7 helix-25 and 4 helix-45 insertions. The RNA was extracted from the same stages of growth for all the analyzed strains. 10 µl of the extract containing RNA isolated from 5 ml of cells from different stages in growth was used in gel electrophoresis.

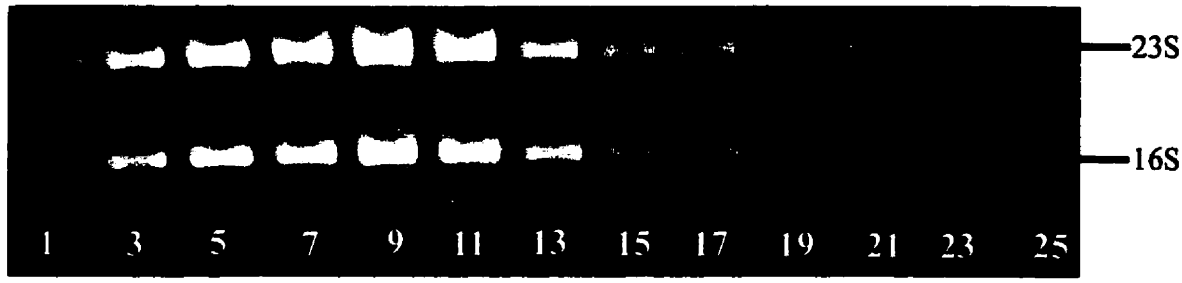
Figure 5.3 indicates that the rRNA content in *E. coli* and the *Salmonella* strains grown for one hour is low; this increases steadily, reaching the maximum level at 11 hours of growth which coincides with the late stationary phase as shown in Fig. 5.1. The

maximum number of viable counts were attained at 13 hours of hours, after which a decline in viability was seen. On further growth of these cultures at 37°C, a decline in the concentration of rRNA is observed till 19 hours of growth after which the decline was steep till 24 hours. No difference in the decline of rRNA concentration was seen when degradation of rRNA from *E. coli* was compared to that from the *Salmonella* strains which have varying degree of fragmentation of the 23S rRNA as a result of excision of IVSs. No difference was noted between the degradation of the 16S and 23S rRNA. In addition, the fragments from 23S rRNA degrade at the same rate as 16S or intact 23S rRNA.

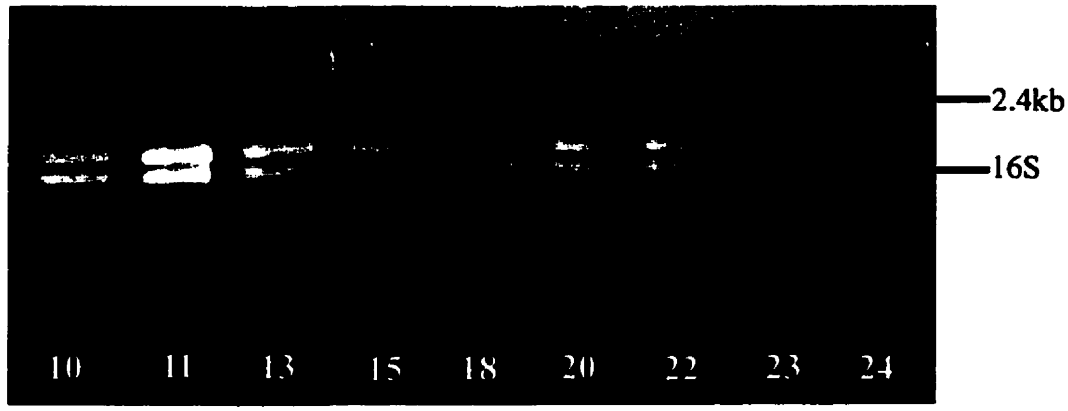
Figure 5.3: rRNA from *E. coli* and *Salmonella* over a 25 hour period.

0.5% inoculum from a 12 hour culture of a single colony isolate was inoculated in fresh LB and the rRNA content during the different phases of growth in LB broth at 37°C with shaking. The hours at which the rRNA was extracted are indicated on the figure. The intact 23S and 16S rRNA which give 2.9kb and 1.6kb bands are indicated. a) rRNA from *E.coli*, which contains no IVSs; b) rRNA from *S. typhimurium* LT2, which has two IVSs in helix-25 and 6 IVSs in helix-45 (Mattatall and Sanderson, 1996). This leads to the fragmentation of the 23S rRNA (2.9kb band). The excision of IVSs from the seven copies of rRNA yields fragments of 2.4kb, 1.7kb, 1.2kb, 0.7kb, and 0.5kb; c) rRNA from *S. agona*, which possesses two IVSs in helix-25 and one in helix-45, this leads to the fragmentation of the 23S (Fig. 4.5); d) rRNA from SARC 10, which has seven IVSs in helix-25 and four IVSs in helix-45, leading to extensive fragmentation of the 23S rRNA (Fig. 4.6).

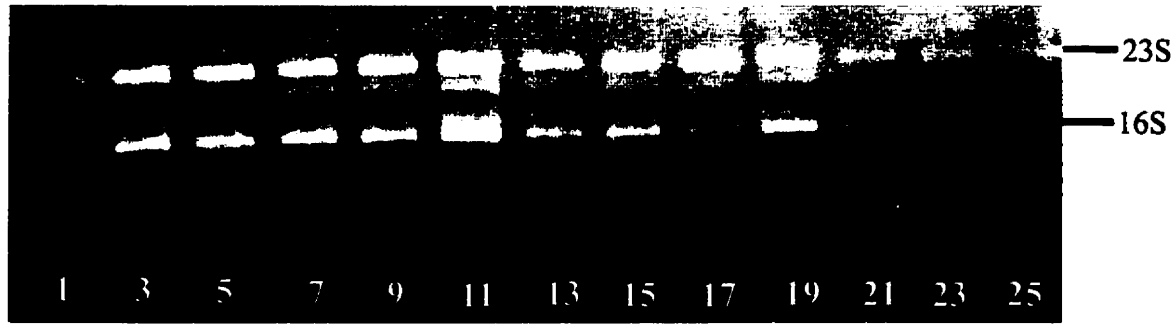
a)


23S
16S
1 3 5 7 9 11 13 15 17 19 21 23 25

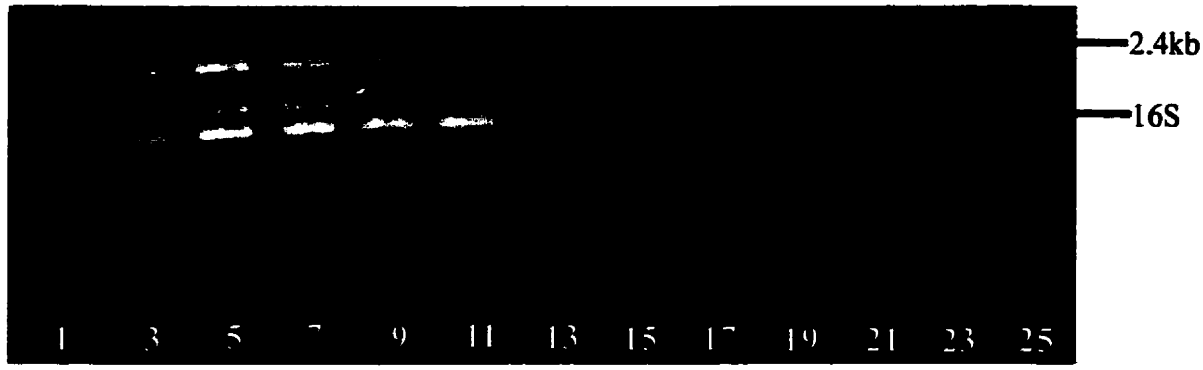
b)


2.4kb
16S
10 11 13 15 18 20 22 23 24

c)


23S
16S
1 3 5 7 9 11 13 15 17 19 21 23 25

d)


2.4kb
16S
1 3 5 7 9 11 13 15 17 19 21 23 25

5.3 Stationary phase lysis of cells.

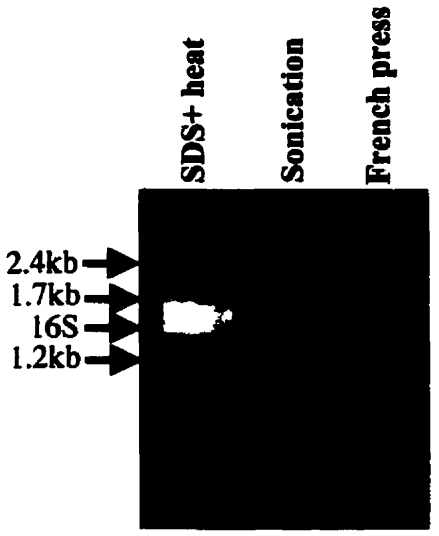
There is evidence that in some situations the breakage of stationary phase cells is difficult (Kolter et al., 1993), therefore the apparent decline in rRNA content after 13 hrs (Fig. 4.3) could be a result of failure to lyse the cells and not due to decrease in rRNA content in the stationary phase. This explanation seemed unlikely, because no CFU were detected on LB agar after extraction of RNA by SDS + heat. However, this idea was further tested by various methods of lysis, testing RNA in the extract and plating the lysed cells on LB plates to check for viable intact cells. 24 hour old cells of *S. typhimurium* LT2 were lysed by the methods shown in Fig. 5.4, i.e., by SDS + heat, by the French press and by sonication. The rRNA extracted from these cells is shown in Fig. 5.4, panel A; lane 1 represents the rRNA from cells lysed by SDS + heat, lane 2 from cells lysed by sonication and lane 3 from cells lysed using a French press. The CFU/ml was determined by dilution and plating on LB medium (Panel B) The original culture had 6×10^9 CFU/ml, rRNA extracted from cells lysed using SDS and heat treatment was maximum and no viable cells were detected on the LB plates. The cells lysed using sonication showed a lower concentration of rRNA and retained 1×10^8 CFU/ml. The rRNA extracted from cells lysed using a French press showed the lowest concentration and the number of viable cells retained was 8.3×10^8 CFU/ml. The concentration of rRNA from the cells lysed using a French press was less than expected. The reason for this could be that the recovery of cells from a French press after lysis was poor.

Figure 5.4: Lysis of stationary phase cells using various methods.

Panel A. Ribosomal RNA extracted from 5 ml of 24 hour old cells of *S. typhimurium* was lysed using the various methods indicated in the figure. The 16S rRNA and the fragmentation products of the 23S rRNA are indicated. The 0.5 and 0.7 kb bands are not visualized due to low concentration.

Panel B. The number of CFU/ml in the original culture and in the aliquot of cells lysed using the various methods is shown.

A.



B.

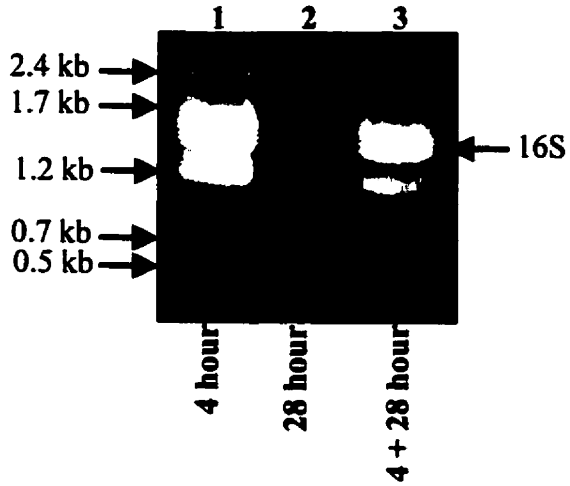
Method of cell lysis	CFU/ml
Original culture	6×10^9
SDS + heat	0
Sonication	1×10^8
French Press	8.3×10^8

5.4 Degradation of rRNA after extraction.

There is a possibility that the apparent reduction in the rRNA content in the stationary phase cells (Fig. 5.3) may be due to the action of RNA degrading enzymes which are released from stationary-phase cells which are not present in log phase cells. Ribosomal RNA extracted from stationary phase cells may also be subjected to the lytic action of enzymes that it would be subjected to in a cell extract, but would not be exposed to in an intact cell. To test if stationary phase cells contain a high concentration of such degradative enzymes, rRNA was extracted from 4 hour old cells mixed with cells from a 28 hour culture. Lane 1 of Fig. 5.5 represents 10 µl of the rRNA extract from 10ml of a 4 hour culture, lane 2 represents 10 µl of the rRNA extract from 10 ml of a 28 hour culture. Ribosomal RNA in lane 3 was extracted from a mixture of 5 ml of 4 hour and 5 ml of 28 hour old culture. No rRNA was detected from the 28 hour culture. The concentration of rRNA in lane 3 is close to 50% of the rRNA in lane 1, indicating that rRNA extracted from the four hour culture was not degraded by the action of degradative enzymes, present in the 28 hour culture. Therefore basis for rapid decline in the level of rRNA in the older cultures does not appear to be degradative enzymes in the extract.

Figure 5.5: Action of degradative enzymes on stationary phase rRNA.

Lane 1 represents rRNA from a 4 hour culture, lane 2 from a 28 hour culture and lane 3 represents rRNA from a mixture of 4 and 28 hour cultures.


1
2
3
2.4 kb
1.7 kb
1.2 kb
0.7 kb
0.5 kb
16S
4 hour
28 hour
4 + 28 hour

CHAPTER 6: DISCUSSION

The number of IVSs in the 7 *rrl* genes (detected by PCR) and the number of cleavage sites in the rRNA (detected by the fragmentation pattern of the rRNA) from 72 *Salmonella* strains of the SARB set (representing sub-genus I) and 16 strains of the SARC set (representing all 8 sub-genera of *Salmonella*) were determined. The following important conclusions were derived from the above analysis.

6.1 Occurrence of IVSs at the conserved helix-25 and helix-45 locations in *Salmonella*.

All IVSs in *Salmonella* occur in or close to helix-25 (bp 550) or helix-45 (bp 1170) regions of the 23S rRNA. The PCR and RNA fragmentation data indicate that the site of IVS insertion is indistinguishable from the expected sites. IVSs in both helix-25 and helix-45 have been previously reported in *Helicobacter* (Hurtado et al., 1997), *Haemophilus* (Xin-Ming-Song et al., 1999), *Proteus* and *Providencia* (Miller et al., 2000). IVSs around bp 1170 have been reported in *Actinobacillus* (Haraszthy et al., 1992), *Campylobacter* (Konkel et al., 1994), *Leptospira* (Ralph and McClelland, 1993), *Yersinia* (Skurnik and Toivanen, 1991), *Rhodobacter* (Konkel et al., 1994) and *Coxiella* (Afseth et al., 1995). IVSs have also been found at other locations, at bp 135 and bp 400 of the *rrl* gene in Rhizobiaceae (Selenska-Pobell and Evguenieva-Hackenberg, 1995) as well as in the 16S rRNA of *Campylobacter helveticus* (Linton et al., 1994b), *Helicobacter canis* (Linton et al., 1994a) (Redburn and Patel, 1993), *Clostridium* spp.

(Rainey et al., 1996), *Rhizobium tropici*(Willems and Collins, 1993), and *Caedibacter caryophila* (Springer et al., 1993). Thus the positions where IVSs can occur are diverse but by no means random, and in *Salmonella* all the IVSs are at the two most common sites. Insertion of IVSs and excision of these elements after transcription are thus tolerated at certain conserved locations in the rRNA suggesting that continuity of the *rrl* gene in this region of the 23S rRNA is not required for ribosome functioning. It is interesting that IVSs in *Salmonella* occur at positions in the 23S rRNA gene that are homologous to two of the positions in which expansion elements occur in eukaryotes, as reviewed by (Burgin et al., 1990). Expansion elements in eukaryotes are extremely variable in sequence and position and are present in *Drosophila* and several other eukaryotes (Lanversin and Jacq, 1989). Some have acquired functions and some of them are removed during rRNA processing; these have been called transcribed spacers or fragmentation spacers (Lanversin and Jacq, 1989).

6.2 Ribosomal RNA fragmentation is always a result of IVS excision.

Fragmentation of the rRNA in *Salmonella* is always due to the excision of an IVS present at bp 550 and/or bp 1170 in all the strains of *Salmonella* tested (Fig. 4.3 and Fig. 4.6). The fragmentation pattern of the rRNA in all 86 strains tested did not show any bands that could not be attributed to excision of IVSs. Fragmentation of the rRNA in some genera is the result of other causes; in the central region of the 23S rRNA, at sites where no IVSs were present, fragmentation was seen in several *Agrobacterium* strains as well as most strains of *Rhizobium leguminosarum* and *Rhizobium etli* as well as other *Rhizobium* spp. (Selenska-Pobell and Doring, 1998). Nonrandom fragmentation of the

rRNA, that could not be attributed to excision of IVSs was seen in *Helicobacter pylori* during conversion from bacillary to coccoid forms. Both 16S and 23S rRNA show distinct highly specific fragmentation patterns at conserved sites (Selenska-Pobell and Doring, 1998). However, in *Salmonella* our data indicate that all fragmentation of rRNA can be directly attributed to the presence of IVSs.

6.3 Selective pressure for IVS removal.

6.3.1 IVSs are not retained in the mature 23S rRNA.

The number of IVSs at the helix-25 and helix-45 locations for all strains was detected by PCR. Independently the rRNA from these strains was analyzed to determine the number of IVSs excised from the 23S rRNA to give the fragmentation pattern characteristic of the strain. The number of IVSs detected by PCR and the number of IVSs determined on the basis of analysis of the rRNA fragmentation pattern were identical for all strains suggesting that IVSs when present are always excised out from the mature 23S rRNA, presumably by RNase III. Uncleaved IVSs were not detected in any of the strains tested. RNase III-deficient mutants of *S. typhimurium*, which retain IVSs in all their ribosomes, are known to be viable in culture (Mattatall and Sanderson, 1998), yet uncleaved IVSs were not detected in 86 wild type strains of *Salmonella* in the present study (Fig. 4.5, Fig. 4.6). This indicates that these strains must have a strong selective pressure for maintenance of sequences which form a base-paired stem at the site of the IVS in the rRNA transcript (this stem is the substrate for RNase III excision of the IVS), otherwise unexcised IVSs would occasionally be detected. This suggests that the

continued presence of the IVS in the rRNA might hinder optimal functioning of the ribosome. It has been shown that the regions at the end of the 23S rRNA form a stable stem which is essential for RNase III-mediated 23S rRNA maturation (Liiv and Remme, 1998); it is seen that there is a similar conservation of the stem region of the IVS (discussed later) to ensure their removal. This suggests that though strains with intact IVSs are viable in culture (Mattatall and Sanderson, 1998), they would not survive in nature and removal of IVSs confers selective advantage.

6.3.2 Compensatory mutations in IVSs.

Comparative sequence analysis in other organisms has revealed positional covariance (also called, and here referred to as compensatory mutations). Compensatory mutations in the DNA occur when nucleotide substitutions at one position are associated with substitutions at another location, which restores canonical base-pairing of the RNA product of the DNA. The occurrence of compensatory mutations indicates that the pair of altered bases in the RNA interact through secondary structure (Gutell et al., 1994). For example, comparative analysis of 16S rRNA sequences from the ECOR set of *E. coli* strains has revealed compensatory mutations in specific regions, which confirm previously predicted secondary structure (Martinez-Murcia et al., 1999). The importance of secondary structure in the spacer region for RNase III processing (which brackets the 23S rRNA gene of *E. coli*) was proven through construction of compensatory mutations in this region (Noller, 1984). Similarly, my comparative analysis of the IVS sequences in the primary stem in *Salmonella, Proteus, Providencia* and *Yersinia* confirms the importance of secondary structure in the primary stem region as a substrate for the RNase

III. In the primary stem region of helix-25 (Fig. 4.13) and helix-45 (Fig. 4.14) the base changes are largely compensatory in nature. The high frequency of these compensatory mutations confirms that physical interactions occur at the sites predicted by the free energy minimization program (Zuker et al., 1999). The existence of these compensatory mutations strongly suggests that there is selective pressure to maintain the RNase III cleavage sites; it also confirms the importance of removal of the IVSs. This indicates that the continued presence of the IVS in the rRNA might hinder optimal functioning of the ribosome; RNase III deficient mutants of *S. typhimurium* LT2 in which all ribosomes retain IVSs are viable in culture, but may not survive in nature. Even in wild type strains which have many IVSs, uncleaved IVSs are never detected (Fig. 4.5, Fig. 4.6) even though mutation of the primary stem region is seen between the enterobacterial IVSs (Fig. 4.10d and 4.11e); these facts, and the compensatory mutations, reveal the importance of removal of the IVSs from wild types strains.

6.4 Transfer of IVSs.

6.4.1 Transfer of IVSs in salmonellae involves both vertical and horizontal transfer.

Evolutionary trees showing relationships of 72 strains of the SARB set (Boyd et al., 1993) and 16 strains of the SARC set (Boyd et al., 1996) of *Salmonella* were previously constructed, based on multi-locus enzyme electrophoresis. The number of IVSs in the seven *rrl* genes for each strain was determined. The data show that very closely related strains often show identical or very similar patterns of distribution of IVSs (Fig. 4.5 and

Fig. 4.6); for example, of eight strains with no IVSs, seven are members of three groups found at different locations on the SARB tree (Fig. 4.5) (Pa1 and Se1; Pc1, Pc2 and Ts1; Mo1 and Mo6). This indicates some degree of stability of IVSs, indicating vertical transmission. However, there is also much evidence for horizontal transfer of IVSs. IVSs are often found in strains which are unrelated according to MLEE criteria; IVSs seem randomly distributed through the various divisions of the SARB set, and are usually present in each of the eight subgenera of the SARC set (Fig. 4.6). Thus there is no strong overall pattern of possession of IVSs on the tree either in helix-25 or in helix-45. An apparent exception to this is sub-genus V (*S. bongori*) which has no helix-45 IVSs in the two strains tested (Fig.4.6), suggesting the possibility that helix-45 entered after the divergence of the other *Salmonella* from *S. bongori*, but this apparent exception is not real, for IVSs were detected by PCR in helix-45 in 2 of the 13 strains of *S. bongori* which were tested (data not shown). It was earlier reported that *E. coli* strains do not have IVSs(Burgin et al., 1990). I confirmed this; all 72 strains from the ECOR set (Ochman and Selander, 1984) were tested by PCR for the possession of IVSs at the helix-25 and helix-45 regions, and none were detected (data not shown). It can be suggested that *Salmonella* acquired IVSs, presumably due to a horizontal transfer event from an external source, after the divergence of *Salmonella* from *Escherichia*, estimated to be over 100 million years ago (Ochman and Wilson, 1987). Representatives of the different families of helix-25 (A to C) and helix-45 (M & O) are distributed rather randomly among the strains of the SARB set (Fig. 4.11). The absence of a strong overall pattern in terms of possession of IVSs and distribution of families on the SARB and SARC trees indicates that the possession of IVSs is a random phenomenon and that the process of acquisition

through horizontal transfer and spread among the seven *rrl* genes accompanied by sporadic loss of IVSs is continuous among the *Salmonella* spp. It can thus be postulated that *Salmonella* spp. acquired IVSs by genetic transfer from an external source after the divergence of *Salmonella* from *Escherichia*, presumably due to a horizontal transfer event into *Salmonella*

6.4.2 Horizontal transfer of IVSs in *Enterobacteriaceae.*

Helix-25 IVSs from family D of *Proteus vulgaris* show 74% nucleotide identity to helix-25 IVS from the *rrlH* operon of *S. typhimurium* (Mattatall and Sanderson, 1996; Miller et al., 2000). Nucleotide identity of family D of *Proteus vulgaris* to the other helix-25 families of *Proteus mirabilis and Proteus penneri* is 50% (Table 4.2). This is suggestive of a horizontal transfer event between *Salmonella* and *Proteus vulgaris*. Blocks of similar sequences are shared between the helix-45 IVSs from *Proteus* and *Providencia* (Fig. 4.11d) (Miller et al., 2000), and the difference between these families is characterized by two large insertions. It has been reported that the helix-45 IVS from *Yersinia enterocolitica* group II is more related to the helix-45 from *Salmonella typhimurium* (89% nucleotide identity) and *Salmonella enterica* (84% nucleotide identity) (Skurnik and Toivanen, 1991) than to the helix-45 IVS from *Yersinia enterocolitica* group I (60%). The conservation in the stem region of helix-25 (Fig. 4.10d) and helix-45 (Fig. 4.11e) is shown. The discrepancies in nucleotide identities of IVSs from the different enterobacterial genera in which they have been detected as compared to the overall chromosomal homology indicate extensive horizontal transfer of IVSs within the family. Advanced BLAST searches, with increased E values, against the

non-redundant nucleotide database at the National Centre for Biotechnology Information (Altschul et al., 1997) do not reveal homology between the enterobacterial IVSs and IVSs from other genera of bacteria. Thus there is evidence for horizontal transfer of IVSs between the four different enterobacterial genera, but there is no evidence for horizontal transfer between the enterobacterial genera and other bacterial groups. Yet it is striking to note, as discussed above, that even when there is no homology, the IVSs in different bacterial genera occur at relatively conserved locations, usually in helix-25 and helix-45.

The primary stem regions of families A, B, C (*Salmonella*) and D (*Proteus*) are very similar; so are placed in superfamily I. The differences between the families are characterized by changes in the secondary regions. Primary stems of families E, F and G from *Proteus* and *Providencia* resemble each other, but are not similar to superfamily I in this region, so are postulated to originate from a different ancestor and to belong to superfamily II. All the previously reported helix-45 IVSs in families Q and R (*Proteus* and *Providencia*), families N and P (*Yersinia*), and families M and O (*Salmonella*) (present study) show a well conserved primary stem and thus belong to the same superfamily III indicating origin from a common ancestor. The secondary region of most families are also related, though there has been many base pair changes and indels, but P (*Yersinia*), in spite of well conserved identity in the primary stem, is so different in the secondary region, that it appears to have an independent origin.

6.4.3 Entry and spread of IVSs in *Salmonella*.

The above observations lend support to the speculation that IVSs are transferred among the bacterial species by horizontal transfer using phage-mediated transduction, or

other methods like transformation or conjugation. Spread between the different *rrl* genes of an individual cell could be a result of reciprocal recombination (RecA mediated) or by gene conversion (non-reciprocal) (Mattatall and Sanderson, 1998). However, an alternative hypothesis rather than a continuous process of IVS-transfer between the species of *Salmonella*, is that the progenitor *Salmonella* obtained an IVS (probably by horizontal transfer) but that subsequent variation between different species was simply due to loss of IVSs from some strains.

The presence of three families of helix-25 IVSs might be explained in two ways. First, the IVSs might have entered *Salmonella* in three independent horizontal transfer events; however this seems unlikely since nucleotide identity between the different families in the region of the primary stem is almost complete. A second explanation is that there was only a single horizontal transfer event of a helix-25 IVS into *Salmonella* and this sequence underwent random mutations. Many of these mutations in the region of the primary stem were selected against or compensatory due to the functional importance of this region for excision by RNase III; however mutations in the regions of the secondary stem and loops were not selected against and led to the sequence divergence. This however should have led to IVSs diverged randomly with the similarity residing mainly in the region of the primary stem and does not explain the tight nucleotide conservation within the different families. It can be speculated that IVSs entered *Salmonella* by a horizontal transfer event and the spread among the different serovars of *Salmonella* is a result of a combination of vertical and horizontal transfer. IVSs could not successfully establish in *Escherichia* and in other enterobacterial genera in which they are absent, as these elements do not perform any function in the cell, or

these elements entered *Salmonella* after the divergence of *Salmonella* from *Escherichia* which is estimated to have been 100 million years ago (Ochman and Wilson, 1987).

6.5 Function of IVSs and their role in the breakdown of rRNA in *Salmonella* and *E. coli*.

The random distribution of IVSs in the different bacterial genera, as shown by their presence in *Salmonella* and absence in *Escherichia*, and the random distribution of the number of IVSs in the 7 *rrl* genes of the different *Salmonella* species, indicates that they are not essential elements for survival. Ribosomes with fragmented 23S rRNA are functional, but fragmentation of *Salmonella* rRNA is not necessary since 8 out of 72 *Salmonella* strains in the SARB set have no IVSs and hence no fragmentation, yet they are viable. It is hence not clear if the IVSs have been recently acquired in evolution or if they have been present in all bacteria and some may have lost them in the course of evolution.

One of the proposed functions of IVSs is that they play a role in the degradation of stationary phase rRNA (Hsu et al., 1994). The degradation of ribosomes and rRNA in the stationary phase has been previously reported as an adaptive response by cells to the onset of adverse conditions (Mandelstam and Halvorson, 1960) (Ben-Hamida and Schlessinger, 1966). This process provides nucleotides and amino acids for cell metabolism (Kaplan and Apirion, 1975a) (Kaplan and Apirion, 1975b). It has been reported that the degradation of rRNA in strains of *Salmonella* with IVSs is more rapid as compared to the rRNA degradation in *Vibrio* spp. and *E. coli*, which do not have

fragmented rRNA (Hsu et al., 1994). Ribosomal RNA in *E. coli* cells grown in LB broth was undetectable on an agarose gel after 8 hours of growth while the rRNA in *Salmonella ohio* with IVSs was shown to be undetectable after only 6 hours of growth suggesting that the degradation in *Salmonella* is faster (Hsu et al., 1994). This could be a selective advantage to the cells which helps the cells respond more efficiently to the depletion of nutrients. A less rapid degradation of 16S rRNA (without IVSs) than of 23S rRNA (with IVSs) in *Salmonella* was also noted in this study (Hsu et al., 1994).

In the present study, with a better understanding of the distribution of the number of IVSs among *Salmonella* strains of the SARB and SARC sets, the role of IVSs in the degradation of stationary phase rRNA was re-examined. The rRNA content in *E. coli* and in the *Salmonella* strains with varying degrees of IVS-facilitated fragmentation, is low after one hour (Fig. 5.3); it increases steadily, reaching the maximum level at 12 hours of growth which coincides with the late stationary phase (based on optical density as shown in Fig. 5.1) and with maximum CFU/ml (Fig. 5.2). On further growth of these cultures at 37°C, a smooth decline in the concentration of rRNA is seen till 20 hours of growth, after which the decline was steep till 24 hours. No difference in the rate of decline of rRNA concentration was seen when degradation of rRNA from *E. coli* was compared to that from the *Salmonella* strains which have varying degree of fragmentation of the 23S rRNA. No difference was noted between the degradation of the 16S and 23S rRNA. In addition, the loss of rRNA in the stationary phase was not dependent on the number of IVS insertions. The extent and pattern of rRNA degradation observed was thus in contrast to that observed by Hsu et al., 1994. It is thus concluded that IVSs do not play a role in the degradation of stationary phase rRNA.

Most experiments on the degradation of ribosomes and rRNA in *E. coli* have been conducted under conditions of carbon, nitrogen or phosphate starvation. A comparison of the degradation of ribosomes shows that degradation is more rapid in cells grown in L-broth as compared to cells grown in phosphate medium, and suggests that the catabolic enzymes for ribosome degradation are induced differently in the two media (Ramagopal, 1984). Experiments where the nucleic acid content in the cell was monitored through [^{3}H]-uridine labelling by growing *E. coli* cells in media limited for nitrogen, carbon or phosphate show that the rRNA content began to decrease at 12-21 hrs with the onset of loss of viability. The 16S and 23S rRNAs decreased 30-40% in the early stationary phase and to 10-15% following 21-29 hours (Davis et al., 1986). It has been suggested that the loss of ribosomes and rRNA begins with the onset of loss of viability and once begun, proceeds rapidly to completion (Ramagopal, 1984) (Kaplan and Apirion, 1975a) (Kaplan and Apirion, 1975b).

In summary, intervening sequences are transferred between the enterobacterial strains by horizontal transfer. The primary stem of IVSs, which is the RNaseIII cleavage site is well conserved, however, the region of the secondary stem and loops can undergo extensive sequence divergence. No wild type strains were detected with uncleaved IVSs in the mature rRNA suggesting that there is a positive selection pressure for the removal of IVSs. IVSs were detected only at the helix-25 and helix-45 locations indicating that discontinuity in other regions of the 23S rRNA is not tolerated.

6.6: Future studies.

Some interesting aspects of IVSs that can be studied are:

1. Analysis of uncleaved IVSs in RNaseIII mutants to study if they are located on the surface of the ribosome, such that they do not disrupt ribosome structure. Cells carrying uncleaved IVSs in the 23S ribosome are viable (Mattatall and Sanderson, 1996) which indicates that the IVSs must be arranged in a way that prevents them from interfering with the structure and function of the ribosome. The location of IVSs in the mature ribosome can be determined by testing the binding efficiency of bulky groups to the surface structures of the ribosome and using primer extension analysis to determine the location of the bulky groups (Egebjerg et al., 1990).

2. IVSs have been reported in the enterobacterial species of *Salmonella* (Winkler, 1979) (Burgin et al., 1990) (Mattatall and Sanderson, 1996), *Yersinia* (Skurnik and Toivanen, 1991), *Proteus* (Miller et al., 2000), and *Providencia* (Miller et al., 2000). Other enterobacterial genera tested, including *Citrobacter*, *Enterobacter*, *Klebsiella*, and *Morganella* do not possess IVSs (Miller et al., 2000). Further analysis of other genera from *Enterobacteriaceae* for the possesion of IVSs is being conducted in our laboratory. This will shed light on the extent of horizontal transfer of IVSs in *Enterobacteriaceae*.

CHAPTER 7: LITERATURE CITED

Afseth, G., Mo, Y.-Y., and Mallavia, L.P. (1995). Characterization of the 23S and 5S rRNA genes of *Coxiella burnetti* and identification of an Intervening sequence within the 23S rRNA gene. J. Bacteriol. *177*, 2946-2949.

Altman, S., Bothwell, A.L., and Stark, B.C. (1975). Processing of *E.coli* tRNA Tyr precursor RNA in vitro. Brookhaven Symp. Biol. *26*, 12-25.

Altschul, S.F., Madden, T.L., Schaffer, A.A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D.J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. *25*, 3389-3402.

Apirion, D. and Miczak, A. (1993). RNA processing in prokaryotic cells. Bioessays *15(2)*, 113-120.

Apirion, D., Watson, N., and Neil, J. (1976). Consequences of losing ribonuclease III on the *E.coli* cell. Mol. Gen. Genet. *144*, 185-190.

Belfort, M., Raeban, M.E., Coetzee, T., and Dalgaard, J.Z. (1995). Prokaryotic introns and inteins: a panoply of form and function. J. Bacteriol. *177*, 3897-3903.

Beltran, P.S., Plock, A., Smith, N.H., Whittman, T.S., Old, D.C., and Selander, R.K. (1991). Reference collection of strains of the *Salmonella typhimurium* complex from

natural populations. J. Gen. Microbiol. *137*, 601-606.

Ben-Hamida, F. and Schlessinger, D. (1966). Synthesis and breakdown of ribonucleic acid in *Escherichia coli* starving for nitrogen. Biochim Biophy. Acta *119*, 183-191.

Bercovier, H., Brenner, D.J., Ursing, J., Steigerwalt, A.G., Fanning, G.R., Alonso, J.M., Carter, G.P., and Mollaret, H.H. (1980). Characterization of *Yersinia enterocolitica* sensu stricto. Curr. Microbiol. *4*, 201-206.

Bessarab, D.A., Kaberdin, V.R., Wei, C.-L., Liou, G.-G., and Chao, S.L. (1998). RNA components of *Escherichia coli* degradosome: Evidence for rRNA decay. Proc. Natl. Acad. Sci. *95*, 3157-3161.

Blattner, F., Plunkett, G., Bloch, C., Perna, N., Burland, V., Riley, M., Collado-Vides, J., Glasner, J., Rode, C., Mayhew, G., Greyor, J., Davis, N., Kirkpatrick, H., Goeden, M., Rose, D., Mau, B., and Shao, Y. (1997). The complete genome sequence of *Escherichia coli* K-12. Science *277*, 1453-1462.

Boyd, E.F., Wang, F., Beltran, P., Plock, S.A., Nelson, K., and Selander, R.K. (1993). *Salmonella* Reference Collection B (SARB): strains of 37 serovars of subspecies I. J. Gen. Microbiol. *139*, 1125-1132.

Boyd, E.F., Wang, F.S., Whittam, S.T., and Selander, R.K. (1996). Molecular genetic relationships of the Salmonellae. Appl. Env. Microbiol. *62*, 804-808.

Burgin, A.B., Parodos, K., Lane, D.J., and Pace, N.R. (1990). The excision of intervening sequences from *Salmonella* 23S rRNA. Cell *60*, 405-414.

Cech, T.R. (1990). self splicing of groupI introns. Annu. Rev. Biochem. *59*, 543-568.

Cech, T.R. and Bass, B.L. (1986). Biological catalysis by RNA . Annu. Rev. Biochem. *55*, 599-629.

Chelladurai, B.S., Li, H., and Nicholson, A.W. (1991). A conserved sequence element in ribonuclease III processing signals is not required for accurate in vitro enzymatic cleavage. Nucleic Acids Res. *19*, 1759-1766.

Colleaux, L., Michel-Wolwertz, M.R., Matagne, R.F., and Dujon, B. (1990). The apocytochrome b gene of *Chlamydomonas smithii* contains a mobile intron related to both *Saccharomyces* and *Neurospora* introns. Mol. Gen. Genet. *223*, 288-296.

Crosa, J.H., Brenner, D.J., Ewing, W.H., and Falkow, S. (1973). Molecular relationships among the salmonellae. J. Bacteriol. *115*, 307-315.

Davis, B.D., Luger, S.M., and Tai, P.C. (1986). Role of ribosome degradation in the death of starved *Escherichia coli* cells. J. Bacteriol. *166*, 439-445.

Dawes, E.A. (1976). Endogenous metabolism and the survival of starved prokaryotes. In The Survival of Vegetative Microbes. T.R.G. Gray, J.R. Postgate,London: Cambridge

Univ. Press.:19-53

Dujon, B., Belfort, M., Butow, R.A., Jacq, C., Lemieux, C., Perlman, P.S., and Vogt, M.J. (1989). Mobile introns: definition of terms and recommended nomenclature. Gene *82*, 115-118.

Dunn, J.J. (1976). RNaseIII cleavage of single-stranded RNA. Effect of ionic strength on the fidelity of the cleavage. J. Biol. Chem. *251*, 3807-3814.

Egebjerg, J., Larsen, N., and Garrett, R.A. (1990). The Ribosome- Structure, Function and Evolution. Washington, D. C.: American Society for Microbiology, pp. 168-179.

Egebjerg, J., Larsen, N., and Garrett, R.A. (1990). Structural map of 23S rRNA. In:The ribosome. Structure, function, & evolution, edited by Hill, W.E., Dahlberg, A., Garrett, R.A.,Moore, P.B., Schlessinger, D., and Warner, J.R.Washington,D.C.American Society for Microbiology, 168-179.

Gegenheimer, P. and Apirion, D. (1980). Precursors to 16S and 23S ribosomal RNA from a ribonucleaseIII$^-$ strain of *Escherichia coli* contain intact RNaseIII processing sites. Nucleic Acids Res. *8*, 1873-1891.

Gegenheimer, P. and Apirion, D. (1981). Processing of prokaryotic ribonucleic acid. Microbiol. Revs. *45*, 502-541.

Gegenheimer, P., Watson, N., and Apirion, D. (1977). Multiple pathways for primary processing of ribosomal RNA in *Escherichia coli*. J. Biol. Chem. *252*, 3064-3073.

Ginsburg, D. and Steitz, T.A. (1975). The 30S ribosomal precursor RNA from *Escherichia coli*. A primary transcript containing 23S, 16S and 5S sequences. J. Biol. Chem. *250(14)*, 5647-5654.

Gregory, S.T., O'Connor, M., and Dahlberg, A.E. (1996). Functional *Escherichia coli* 23S rRNA containing processed and unprocessed intervening sequences from *Salmonella typhimurium*. Nucleic Acids Res. *24*, 4918-4923.

Gutell, R.R., Larsen, N., and Woese, C.R. (1994). Lessons from an evolving rRNA: 16S and 23S rRNA structures from a comparative perspective. Microbiol. Revs. *58*, 10-26.

Haraszthy, V.I., Sunday, G.J., Bobek, L.A., Motley, T.S., Preus, H., and Zambon, J.J. (1992). Identification and analysis of the gap region in the 23S ribosomal RNA from *Actinobacillus actinomycetemcomitans*. J. Dent.Res. *71*, 1561-1568.

Hardy, C.M. and Clark-Walker, G.D. (1991). Nucleotide sequence of the COX1 gene in *Kluyveromyces lactis* mitochondrial DNA : evidence for recent horizontal transfer of a groupII intron. Curr. Genet. *20*, 99-114.

Hayes, F. and Vasseur, M. (1976). Processing of the 17S *Escherichia coli* precursor RNA in the 27-S pre-ribosomal particle. Eur. J. Biochem. *61(2)*, 433-442.

Hsu, D., Shih, L.M., and Zee, Y.C. (1994). Degradation of ribosomal RNA in *Salmonella* strains: a novel mechanism to regulate the concentration of ribosomal RNA and ribosomes. J. Bacteriol. *176*, 4761-4765.

Hurtado, A., Clewley, J.P., Linton, D., Owen, R.J., and Stanley, J. (1997). Sequence similarities between large subunit ribosomal RNA gene intervening sequences from different *Helicobacter* species. Gene *194*, 69-75.

Ishihama, A. (1997). Adaptation of gene expression in stationary phase bacteria. Curr. Opin. Genet. Dev. *7*, 582-588.

Kaplan, R. and Apirion, D. (1975a). Decay of ribosomal acid in *Escherichia coli* cells starved for various nutrients. J. Biol. Chem. *250*, 3174-3178.

Kaplan, R. and Apirion, D. (1975b). The fate of ribosomes in *Escherichia coli* cells starved for a carbon source. J. Biol. Chem. *250*, 1854-1863.

King, T.C., Sirdeshmukh, R., and Schelessinger, D. (1986). Nucleolytic processing of ribonucleic acid transcripts in prokaryotes. Microbiol. Revs. *50*, 428-451.

King, T.C., Sirdeshmukh, R., and Schlessinger, D. (1984). RNaseIII cleavage is obligate for maturation but not for function of *Escherichia coli* pre 23S rRNA. Proc. Natl. Acad. Sci. *81*, 185-188.

Kjelleberg, S., Hermansson, M., Marden, P., and Jones, G.W. (1987). The transient phase between growth and nongrowth of heterotrophic bacteria, with emphasis on the marine environment. Annu. Rev. Microbiol. *41*, 25-49.

Kolter, R., Siegele, D.A., and Tormo, A. (1993). The stationary phase of the bacterial life cycle. Annu. Rev. Microbiol. *47*, 855-874.

Konkel, M.E., Marconi, R.T., Meda, D.J., and Ceiplak, W. (1994). Identification and characterization of an intervening sequence within the 23S ribosomal RNA genes of *Campylobacter jejuni*. Mol. Microbiol. *14*, 235-241.

Krinkle, L. and Wulff, D.L. (1990). The cleavage specificity of RNaseIII. Nucleic Acids Res. *18*, 4809-4815.

Lambowitz, A.M. and Belfort, M. (1993). Introns as mobile genetic elements. Annu. Rev. Biochem. *62*, 587-622.

Lang, B.F. (1984). The mitochondrial genome of the fission yeast *Schizosaccharomyces pombe*: highly homologous introns are inserted at the same position of the otherwise less conserved cox1 genes in *Schizosaccharomyces pombe* and *Aspergillus nidulans*. EMBO J. *3*, 2129-2136.

Lange, R. and Hengge-Aronis, R. (1991). Growth phase regulated expression of bolA and morphology of stationary phase *E.coli* cells are controlled by the novel sigma factor σ^{S}.

J. Bacteriol. *173*, 4474-4481.

Lanversin, G. and Jacq, B. (1989). Sequence and secondary structure of the central domain of *Drosophila* 26S rRNA: A universal model for central domain of the large rRNA containing the region in which the central break may happen. J. Mol. Evol. *28*, 403-417.

Le Minor, L. (1984). Bergey's manual of systematic bacteriology. Williams and Wilkins, Baltimore, pp. 427-458.

Leffers, H., Egebjerg, J., Andersen, A., Christensen, T., and Garrett, R.A. (1988). Domain VI of *E. coli* 23S rRNA. Structure, assembly and function. J. Mol. Biol. *204*, 507-522.

Leffers, H., Kjems, J., Ostergaard, L., Larsen, N., and Garrett, R.A. (1987). Evolutionary relationships among archaebacteria. A comparative study of the 23S rRNAs of a sulphur-dependent thermophile, an extreme halophile and a thermophilic methanogen. J. Mol. Biol. *195*, 43-61.

Li, H., Chelladurai, C.S., Zhang, K., and Nicholson, A.W. (1993). RibonucleaseIII cleavage of a bacteriophage T7 processing signal. Divalent cation specificity, and specific anion effects. Nucleic Acids Res. *21*, 1919-1925.

Li, Z., Pandit, S., and Deutscher, M.P. (1999a). Maturation of 23S ribosomal RNA

requires the exoribonuclease RNase T. RNA *5*, 139-146.

Li, Z., Pandit, S., and Deutscher, M.P. (1999b). RNase G (Caf A protein) and RNase E are both required for the 5' maturation of 16S ribosomal RNA. EMBO J. *18(10)*, 2878-2885.

Liiv, A. and Remme, J. (1998). Base-pairing of 23S rRNA ends is essential for ribosomal large subunit assembly. J. Mol. Biol. *276*, 537-545.

Lindahl, L. and Zengel, J.M. (1986). Ribosomal genes in *Escherichia coli*. Annu. Rev. Genet. *20*, 297-326.

Linton, D., Clewely, P.J., Burnens, A., Owen, J.R., and Stanley, J. (1994a). An intervening sequence (IVS) in the 16S rRNA gene of the eubacterium *Helicobacter canis*. Nucleic Acids Res. *22*, 1954-1958.

Linton, D., Dewhirst, F.E., Clewley, J.P., and Owen, R.J. (1994b). Two types of 16S ribosomal RNA genes are found in *Campylobacter helveticus*: analysis, application and characterization of the intervening sequence found in some strains. Microbiology *140*, 847-855.

Liu, S.-L., Hessel, A., and Sanderson, K.E. (1993). Genomic mapping with I-*Ceu*I, an intron-encoded endonuclease, specific for ribosomal RNA, in *Salmonella* spp., *Escherichia coli*, and other bacteria. Proc. Natl. Acad. Sci. *90*, 6874-6878.

Loewen, P.C. and Hengge-Aronis, R. (1996). The role of sigma factor (S) (KatF) in bacterial global regulation. Annu. Rev. Microbiol. *48*, 53-80.

Mandelstam, J. and Halvorson, H. (1960). Turnover of protein and nucleic acid in soluble and ribosome fractions of non growing *Escherichia coli.* Biochim. Biophy. Acta *40*, 43-49.

Martinez-Murcia, A.J., Anton, A.I., and Rodriguez-Valera, F. (1999). Patterns of sequence variation in two regions of the 16S rRNA multigene family of *Escherichia coli.* Int. J. Syst. Bacteriol. *49*, 601-610.

Maruyama, H. and Mizuno, D. (1966). The participation of ribonuclease in the degradation of *Escherichia coli* ribosomal ribonucleic acid as revealed by oligonucleotide accumulation in the phosphorus-deficient stage. Biochim. Biophy. Acta *108*, 593-604.

Matin, A., Auger, E.A., Blum, P.H., and Schultz, J.E. (1989). Genetic basis of starvation survival in nondifferentiating bacteria. Annu. Rev. Microbiol. *43*, 293-316.

Mattatall, N.R. and Sanderson, K.E. (1996). *Salmonella typhimurium* LT2 possesses three distinct 23S rRNA intervening sequences. J. Bacteriol. *178*, 2272-2278.

Mattatall, N.R. and Sanderson, K.E. (1998). RNaseIII deficient *Salmonella typhimurium* LT2 contains intervening sequences in the 23S rRNA. FEMS Microbiol. Lett. *159*, 179-185.

Meyhack, B., Mayhack, I., and Apirion, D. (1974). Processing of precursor particles containing 17S rRNA in a cell free system. FEMS Microbiol. Lett. *49(2)*, 215-219.

Michel, F., Umesono, K., and Ozeki, H. (1989). Comparative and functional anatomy of GroupII catalytic introns. Gene *82*, 5-30.

Miller, W.L., Pabbaraju, K., and Sanderson, K.E. (2000). Fragmentation of 23S ribosomal RNA in strains of *Proteus* and *Providencia* results from intervening sequences (IVSs) in the *rrn* (rRNA) genes. J. Bacteriol. In Press.

Morgan, E.A., Ikemura, T., Post, L.E., and Nomura, M. (1980). tRNA genes in *rrn* operons of *Escherichia coli*. In Transfer RNA: Biological Aspects. Cold Spring Harbor Laboratory, Cold Springs Harbor: pp. 259-266.

Mulvey, M.R. and Loewen, P.C. (1989). Nucleotide sequence of *katF* gene of *Escherichia coli* suggests KatF protein is a novel σ transcription factor. Nucleic Acids Res. *17*, 9979-9991.

Neidhardt, F.C., J. L. Ingraham, J.L., Low, K.B., Magasanik, B., Schaechter, M., and Umbarger, H.E. (1987). *Escherichia coli* and *Salmonella typhimurium*: cellular and molecular biology. Washington, D. C.: American Society for Microbiology.

Nicholson, A.W. (1999). Function, mechanism and regulation of bacterial ribonucleases. FEMS Microbiol. Rev. *23*, 371-390.

Noller, H.F. (1984). Structure of ribosomal RNA. Ann. Rev. Biochem. *53*, 119-162.

Ochman, H. and Selander, R.K. (1984). Standard reference strains of *Escherichia coli* from natural populations. J. Bacteriol. *157*, 690-693.

Ochman, H. and Wilson, A.C. (1987). Evolutionary history of enteric bacteria. In *Escherichia coli* and *Salmonella typhimurium*: cellular and molecular biology. F. Neidhardt, ed. Washington, D. C.: American Society of Microbiology, pp. 1649-1654.

Peebles, C.L., Perlman, P.S., Mcklenburg, K.L., Petrillo, M.L., and Tabor, J.H. (1986). A self splicing RNA excises an intron lariat. Cell *44*, 213-223.

Rainey, F.A., Ward-Rainey, N.L., Janssen, P.H., Hippe, H., and Stackebrandt, E. (1996). *Clostridium paradoxum* DSM 7308 contains multiple 16S ribosomal RNA genes with heterogenous intervening sequences. Microbiology *142*, 2087-2095.

Ralph, D. and McClelland, M. (1993). Intervening sequences with a conserved open reading frame in eubacterial 23S rRNA genes. Proc. Natl. Acad. Sci. *90*, 6864-6868.

Ralph, D. and McClelland, M. (1994). Phylogenetic evidence for horizontal transfer of an intervening sequence between species in a spirochete genus. J. Bacteriol. *19*, 5982-5987.

Ramagopal, S. (1984). Metabolic changes in ribosomes of *Escherichia coli* during prolonged culture in different media. Eur. J. Biochem. *140*, 353-361.

Redburn, A.C. and Patel, B.K.C. (1993). Phylogenetic analysis of *Desulfotomaculum thermobenzoicum* using polymerase chain reaction amplified 16S rRNA-specific DNA. FEMS Microbiol. Lett. *113*, 81-86.

Reeve, C.A., Amy, P.S., and Matin, A. (1984). Role of protein synthesis in the survival of carbon starved *Escherichia coli* K12. J. Bacteriol. *160*, 1041-1046.

Robertson, H.D. (1982). *Escherichia coli* ribonuclease III cleavage sites. Cell *30*, 669-672.

Saitou, N. and Nei, M. (1987). The neighbour-joining method: a new method for reconstructing phylogenetic trees. Mol. Biol. Evol. *4*, 406-425.

Saldanah, R., Mohr, G., Belfort, M., and Lambowitz, A.M. (1993). Group I and Group II introns. FASEB J. *7(1)*, 15-24.

Sambrook, J., Fritsch, E.F., and Maniatis, T. (1989). Molecular cloning, A laboratory manual. Cold Spring Harbor Laboratory Press.

Schultz, J.E. and Matin, E. (1988). Differential regulation by cyclic AMP of starvation protein synthesis in *Escherichia coli*. J. Bacteriol. *170*, 3903-3909.

Selander, R.K., Beltran, P., and Smith, N.H. (1991). Evolutionary genetics of Salmonella. Sunderland, Mass., Sinauer Associates Inc., pp. 25-57.

Selenska-Pobell, S. and Doring, H. (1998). Sequences around the fragmentation sites of the large subunit ribosomal RNA in the family *Rhizobiaceae.* Antonie van Leeuwenhoek *73*, 55-67.

Selenska-Pobell, S. and Evguenieva-Hackenberg, E. (1995). Fragmentation of the large subunit ribosomal RNA in the family Rhizobiaceae . J. Bacteriol. *177*, 6993-6998.

Skurnik, M. and Toivanen, P. (1991). Intervening sequences (IVSs) in the 23S ribosomal RNA genes of pathogenic *Yersinia enterocolitica* strains. The intervening sequences in *Y. enterocolitica* and *Salmonella typhimurium* have a common origin. Mol. Microbiol. *5*, 585-593.

Spassky, A., Rimsky, S., Garreau, H., and Buc, H. (1984). Hla, an *E.coli* DNA -binding protein which accumulates in stationary phase, strongly compacts DNA in vitro. Nucleic Acids Res. *12*, 5321-5340.

Springer, N., Ludwig, W., Amann, R., Schmidt, H.J., Goertz, H.D., and Schleifer, K.H. (1993). Occurrence of fragmented 16S rRNA in an obligate bacterial symbiont of *Paramecium caudatum.* Proc. Natl. Acad. Sci. *90*, 9892-9895.

Suryanarayana, T. and Burma, D.P. (1975). Substrate specificity of *Salmonella typhimurium* RNaseIII and the nature of the products formed. Biochim. Biophys. Acta *407*, 459-468.

Szeberenyi, J., Roy, M.K., and Apirion, D. (1983). Precursor nucleotides at the 5' end are not required for processing by RNAse E at the 3' end of 5-S rRNA. Eur. J. Biochem. *136(2)*, 321-326.

Tanaka, K., Fujita, N., Ishihama, A., and Takahashi, H. (1993). Heterogeneity of the principal sigma factor in *Escherichia coli*: the *rpo*S gene product, 38, is a principal sigma factor of RNA polymerase in stationary phase of *Escherichia coli*. Proc. Natl. Acad. Sci. USA *90*, 3511-3515.

Tanaka, K., Kusano, S., Fujita, N., Ishihama, A., and Takahashi, H. (1995). Promotor determinants for *Escherichia coli* RNA polymerase holoenzyme containing σ^{38} (the *rpo*S gene product). Nucleic Acids Res. *23*, 827-834.

Trust, T.J., Logan, S.M., Gustafson, C.E., Romaniuk, P.J., Kim, N.W., Chan, V.L., Ragan, M.A., Guerry, P., and Gutell, R.R. (1994). Phylogenetic and molecular characterization of 23S rRNA gene positions in the genus *Campylobacter* in the epsilon subdivision of Proteobacteria shows that the presence of transcribed spacers is common in *Campylobacter* sp. J. Bacteriol. *176*, 4597-4609.

van der Veen, R., Arnberg, G., van der Horst, L., Bonen, L., Tabak, H.F., and Grivell, L.A. (1986). Excised Group II introns in yeast mitochondria are lariats and can be formed by self-splicing in vitro. Cell *44*, 225-234.

Wada, A., Igarashi, K., Yoshimura, S., Aimoto, S., and Ishihama, A. (1995). Ribosome

modulation factor: stationary growth phase-specific inhibitor of ribosome functions from *Escherichia coli*. Biochem. Biophys. Res. Commun. *214*, 410-417.

Wada, A., Yamazaki, Y., Fujita, N., and Ishihama, A. (1990). Structure and probable genetic location of a "ribosome modulation factor" associated with 100S ribosomes in stationary phase *Escherichia coli* cells. Proc. Natl. Acad. Sci. *87*, 2657-2661.

Willems, A. and Collins, M.D. (1993). Phylogenetic analysis of rhizobia and agrobacteria based on 16S ribosomal RNA gene sequences. Int. J. Syst. Bacteriol. *43*, 305-313.

Winkler, M.E. (1979). Ribosomal ribonucleic acid isolated from *Salmonella typhimurium*: absence of the intact 23S species. J. Bacteriol. *139*, 842-849.

Xin-Ming-Song, A. Forsgren, and H. Janson (1999). Fragmentation heterogeneity of 23S ribosomal RNA in *Haemophilus* species. Gene *230*, 287-293.

Zhang, K. and Nicholson, A.W. (1997). Regulation of ribonucleaseIII processing by double helical sequence determinants . Proc. Natl. Acad. Sci. *94*, 13437-13441.

Zuker, M., Mathews, D.H., and Turner, D.H. (1999). Algorithms and Thermodynamics for RNA Secondary Structure Prediction : A Practical Guide in RNA Biochemistry and Biotechnology, J. Barciszewski & B.F.C. Clark, eds., NATO ASI Series.