SUBCLONING, CHARACTERIZATION AND SEQUENCING OF THE STREPTOMYCIN RESISTANCE GENE OF A MULTIPLE ANTIBIOTIC RESISTANCE TRANSPOSON FROM *Salmonella typhi*

RAMESVARI A/P PARARAJASINGAM PILLAI

A dissertation submitted to the University of Malaya, in partial fulfillment of the requirements for the Degree of Master of Biotechnology

INSTITUTE OF POSTGRADUATE STUDIES AND RESEARCH
UNIVERSITY OF MALAYA

KUALA LUMPUR

June 2000

Dedicated to,

The loving memory of my Daddy (late), ...

the wellspring of wisdom and perseverance and

to my Mummy,.....

for her devotion and her constant demonstrations

of love.

ABSTRACT

Salmonella typhi S8 is resistant to ampicillin (Ac), chloramphenicol (Cm), co-trimoxazole (Ctm), streptomycin (Sm), and tetracycline (Tc). The Ac, Cm, Ctm, and Sm resistance traits have been shown to be mediated by a multiple antibiotic resistance transposon located on a large conjugative plasmid. Transposition of the transposon encoding Ac^R (ampicillin resistance), Cm^R, Ctm^R, and Sm^R, designated Tn*X8*, was done by Wong (1999) to a recipient replicon, plasmid pUB307. The Ac^R, Cm^R, Ctm^R, and Sm^R genes were subcloned into plasmid pKan by digesting the recombinant plasmid, designated pCL8 (pUB307::Tn*X8*), with *Sal*I. This subcloning strategy generated different recombinant pKans that harbour different antibiotic resistance genes. Plasmid pCLS55 is a recombinant pKan that harbours a *Sal*I insert bearing Sm^R genes from the multiple antibiotic resistance transposon of *S. typhi* S8.

The main objective of this study was to subclone and characterizes, by DNA sequencing, the Sm^R genes derived from the transposon of *S. typhi* S8, in pCLS55. Successive subclonings into pUC19 and M13mp18 vectors, and construction of deletions localized the Sm^R genes on a 2.5 kb *Eco*RI DNA fragment. Southern hybridization, performed before DNA sequencing, confirmed that the 2.5 kb *Eco*RI Sm^R inserts carried by M13pSMR13 and M13pSMS16 originated from the plasmid in *S. typhi* S8.

A total of 2,128 nucleotides was determined based on complete DNA sequencing of the 2.5 kb *Eco*RI Sm^R insert in recombinant M13mp18. The nucleotide sequences were analyzed and the gene structure (containing *str* genes, *strA* and *strB*) within the 2.5 kb *Eco*RI Sm^R insert from pCLS55 was determined. These sequences were identical to the *strA* and *strB* genes encoding Sm resistance in plasmid RSF1010 (Scholz *et al.*, 1989) and in Tn*5393* of *Erwinia amylovora* (Chiou and Jones, 1993).

The nucleotide sequences obtained in this study indicated that there are three open reading frames (ORF) in the 2.5 kb *Eco*RI Sm^R fragment. The first ORF of 267 amino acids was designated ORF A'. The deduced amino acid sequence corresponded to a molecular mass of 29,595 Daltons and has high similarity with the amino acid sequence of Sm phosphotransferase protein A, which is also encoded by RSF1010

and Tn*5393*. This ORF translation starts at UUG (TTG in the DNA sequence) and this suggests that this gene did not originally come from members of the family *Enterobacteriaceae*.

The second ORF, designated ORF B', encodes a polypeptide of 278 amino acids, which corresponded to a molecular mass of 30,824 Daltons and has high similarity with the amino acid sequence of Sm phosphotransferase protein B.

The Sm^R genes were also observed to contain part of the transposase gene belonging to IS*26* and its variants. This was observed in the third ORF, ORF C', in the nucleotide sequences obtained. Partial DNA sequencing of the 5.5 kb *Sal*I insert in pCLS55 showed that the two ends of this insert each contained part of the transposase gene of IS*26* and its variants.

The Sm^R genes in pCLS55 appeared to be borne on a composite element, flanked by IS*26* at its ends. The presence of the composite IS*26* element within Tn*X8* supports the hypothesis that multiple antibiotic resistance transposons evolved by insertion of antibiotic resistance determinants, which are themselves transposable.

ABSTRAK

Salmonella typhi S8 adalah rintang terhadap ampisilin (Ac), kloramfenicol (Cm), ko-trimoxazole (Ctm), streptomisin (Sm), dan tetrasiklin (Tc). Ciri-ciri kerintangan terhadap Ac, Cm, Ctm, dan Sm adalah diperantarakan oleh suatu transposon pelbagai kerintangan antibiotik yang berlokasi pada suatu plasmid konjugatif yang besar. Transposisi transposon Tn*X8*, transposon yang mengekodkan Ac^R (kerintangan terhadap ampisilin), Cm^R, Ctm^R, dan Sm^R, kepada replikon penerima iaitu plasmid pUB307 telah dilakukan oleh Wong (1999).

Hasil transposisi tersebut iaitu plasmid rekombinan pUB307, yang didesignasi sebagai pCL8 (pUB307::Tn*X8*), dicerna dengan enzim pencerna *SalI* untuk mensubklonkan gen-gen Ac^R, Cm^R, Ctm^R, dan Sm^R ke dalam plasmid pKan. Strategi pensubklonan ini menjana beberapa pKan rekombinan yang mengandungi gen-gen kerintangan antibiotik yang berlainan. Plasmid pCLS55 adalah sejenis pKan rekombinan yang mengandungi gen-gen Sm^R dalam satu fragmen *SalI* bersaiz 5.5 kb daripada transposon pelbagai kerintangan *S. typhi* S8.

Objektif utama kajian ini adalah untuk mensubklon dan mencirikan, melalui penjujukan DNA, gen-gen Sm^R dalam pCLS55 yang diterbitkan dari transposon *S. typhi* S8. Pensubklonan berturutan kedalam vektor pUC19 dan sejurusnya kedalam vektor M13mp18, serta penjanaan delesi berjaya melokasikan gen-gen Sm^R pada satu fragmen *Eco*RI bersaiz 2.5 kb. Sebelum penjujukan DNA, hibridisasi Southern dilakukan dan ia menunjukkan bahawa asalan fragmen Sm^R *Eco*RI bersaiz 2.5 kb ini adalah sememangnya dari plasmid *S. typhi* S8.

Nukeotida-nukleotida berjumlah 2,128 ditentukan berdasarkan penjujukan DNA seluruhan fragmen *Eco*RI bersaiz 2.5 kb yang terkandung dalam M13mp18 rekombinan. Jujukan nukleotida yang diperolehi dianalisakan dan struktur gen (yang mengandungi gen-gen *str*, iaitu *strA* dan *strB*) dalam fragmen *Eco*RI bersaiz 2.5 kb daripada pCLS55 berjaya ditentukaan. Jujukan-jujukan ini didapati seiras dengan gen-gen *strA* dan *strB* yang mengekodkan kerintangan terhadap Sm dalam plasmid RSF1010 (Scholz *et al.*, 1989) dan dalam Tn*5393* yang termaktub dalam *Erwinia amylovora* (Chiou and Jones, 1993).

Jujukan nukelotida yang diperolehi dalam kajian ini menunjukkan bahawa ada tiga bingkai bacaan terbuka (ORF, "open reading frame") dalam fragmen Sm^R *Eco*RI bersaiz 2.5 kb. Bingkai bacaan pertama yang berdesignasi ORF A' mengandungi 267 asid amino dan sejajar dengan berat molekul 29,595 Dalton. Ia mempunyai kesamaan tinggi dengan jujukan asid amino protin A fosfotransferase Sm yang dikodkan oleh RSF1010 dan Tn*5393*. Translasi ORF ini bermula dengan UUG (TTG dalam jujukan DNA) dan mencadangkan bahawa gen *strA* ini tidak berasal daripada ahli famili *Enterobacteriaceae*.

Bingkai bacaan kedua berdesignasi ORF B' mengekodkan polipetida sebesar 278 asid amino dan sejajar dengan berat molekul 30,824 Dalton. Analisa jujukan asid amino ini sejajar dengan protin B fosfotransferase Sm yang dikodkan oleh gen *strB* RSF1010 dan Tn*5393*. Gen-gen Sm^R ini didapati juga mengandungi sebahagian gen enzim transposisi IS*26* dan variasinya. Ini diperhatikan dalam bingkai bacaan ketiga, yakni ORF C', dalam jujukan nukleotida yang diperolehi dalam kajian ini. Penjujukan DNA separa juga dilakukan terhadap fragmen *SalI* bersaiz 5.5 kb dari pCLS55. Analisa jujukan nukleotid separa yang diperolehi ini juga menunjukkan kehadiran gen enzim transposisi IS*26* dan varasinya.

Gen-gen Sm^R dalam pCLS55 didapati termaktub dalam elemen komposit kerana ia diapiti oleh IS*26* pada kedua-dua hujungnya. Kehadiran elemen komposit IS*26* dalam Tn*X8* mengukuhkan lagi hipotesis bahawa transposon pelbagai kerintangan antibiotik berevolusi melalui insersi penentu-penentu antibiotik, yang mempunyai keupayaan untuk bertransposisi secara sendiri.

ACKNOWLEDGEMENT

First, I would like to express my deepest gratitude to my supervisors, Prof. Dr. Koh Chong Lek and Assoc. Prof. Dr. Sam Choon Kook, who have been ever so supportive during my research and dissertation. Their invaluable advice, constructive criticisms and patient discussions throughout the duration of this project have helped me immensely. They have also provided and offered me enthusiasms and acuity that sharpened my resolve and thinking.

Special thanks and gratitude to Miss Lim Moo Eng for her valuable time and assistance in preparing the diagrams and photographs.

A special note of thanks to my dearest friends and colleagues Ling, May, Dr. Don Gary Benjamin, and Mr. Pathi, Wan Chee, Dr. Fong who constantly gave me the opportunities to share ideas in a large variety of settings and for their friendship.

Among my profound debts is to Wong Hann Ling, who have shared his findings with me, and whose efforts are reviewed and synthesized here.

My indebtedness and gratitude to Kanga, her (Kanga) family, Katherine, and Vinder for their hospitality, encouragement, and assistance in the process of writing this dissertation.

Thanks to staffs of the Institute of Biological sciences and IPSR, especially Mr. Gopal and Mrs. Gan, for offering their assistance to me.

My sincere thanks also to the University of Malaya for financial assistance under the Fellowship and PASCA scheme for funding my study.

A very special note of indebtedness to my family; my mummy, brother and sister, who have provided the cocoon of warmth and love, for their inspirations, encouragement, intellectual and spiritual support to persevere in very trying circumstances. If ever a cliché was true, it would be that without them, this dissertation would never have been possible.

Lastly, my utmost gratification to the LORD for the countless blessings HE bestowed upon me. I thank YOU, LORD.

ABBREVIATIONS

Most of the abbreviations used are standard. However, attention is drawn to the following:

μg	microgram
μl	microlitre
μm	micrometer
A	adenine (in DNA nucleotide sequence)
Ac	ampicillin
Ac^R	ampicillin resistant or resistance
bp	base pair
BPB	bromophenol blue
C	cytosine (in DNA nucleotide sequence)
Cm	chloramphenicol
Cm^R	chloramphenicol resistant or resistance
CsCl	cesium chloride
Ctm	cotrimoxazole
Ctm^R	cotrimoxazole resistant or resistance
dATP	deoxyadenosine 5'-triphosphate
DNA	deoxyribonucleic acid
EDTA	ethylenediaminetetraacetate
EtBr	ethidium bromide
g	gram
G	guanine (in DNA nucleotide sequence)
hr	hour
int	integron or integrase
IR	inverted repeats
IS	insertion sequence
kb	kilobase pair or kilobase
Km	kanamycin

Km^R	kanamycin resistant or resistance
LB	Luria-Bertani
M	molar
mA	milliampere
mg	milligram
MH	Mueller Hinton
min	minute
ml	millilitre
mM	millimolar
MW	molecular weight
N	Normal
ng	nanogram
nm	nanometer
°C	degree Celsius
OD	optical density
ORFS	open reading frame
ori	origin
P	promoters
p.s.i	pounds per square inch
PEG	polyethylene glycol
rec	recombination
RF	replicative form (of M13 DNA)
RHSs	recombination hot spots
rpm	revolutions per minute
SDS	sodium dodecyl sulphate
sec	second
Sm	streptomycin
Sm^R	streptomycin resistant or resistance
Smz	sulphamethoxazole
Smz^R	sulphamethoxazole resistant or resistance
SSC	sodium chloride-sodium citrate
Su	sulfonamide
Su^R	sulfonamide resistant or resistance

T	thymine (in DNA nucleotide sequence)
TBE	Tris-borate-EDTA
Tc	tetracycline
Tc^R	tetracycline resistant or resistance
TE	Tris-EDTA
TEMED	N,N,N',N' - tetramethylenediamine
TGE	transposable genetic element
Tn	transposon
Tp	trimethoprim
Tp^R	trimethoprim resistant or resistance
Tris	Tris (hydroxymethyl) methylamine
UV	ultraviolet
V	volt
W	watt

LIST OF CONTENTS

LIST OF FIGURES

LIST OF TABLES