Restriction alleviation and modification enhancement by the *ral* gene of bacteriophage λ

Gareth King

A thesis presented for the degree of Doctor of Philosophy

Institute of Cell and Molecular Biology
University of Edinburgh

June, 1995

Abstract

Bacteriophages have a variety of mechanisms to escape host-encoded restriction and modification (R-M) systems. Bacteriophage λ has a gene, *ral*, whose product alleviates restriction and enhances modification by a host encoding a type IA R-M system. The genes of the type IA R-M systems encode an hetero-oligomeric DNA methyltransferase (MTase) that together with an additional subunit forms a DNA endonuclease (ENase). Both enzymatic activities share a common subunit for target-site recognition; unmodified targets are normally a good substrate for the ENase, whereas hemimethylated targets, those that are modified on only one strand of the duplex, are a good substrate for the MTase. The MTase component of the type-IA R-M systems are the only prokaryotic MTases shown to have a strong preference for hemimethylated DNA *in vivo*, and are also the only systems in which restriction is alleviated by Ral. This thesis describes an investigation of the mechanism of Ral activity, both *in vivo* and *in vitro*.

Unmodified λ*ral*$^+$ phages are restricted by bacteria expressing type IA R-M systems, presumably restriction occurs before the infecting phage has time to express the *ral* gene. In order to examine Ral activity *in vivo*, the degree of modification of progeny phages was measured after a single round of infection on a restriction-proficient strain, and the prediction that the progeny of *ral*$^+$ phages are more likely to become modified than *ral*$^-$ phages was confirmed. This modification would be advantageous to those phages that subsequently infect bacteria with an R-M system of the same specificity. Interestingly, in some hosts progeny phages were similarly modified irrespective of their *ral* genotype. Some strains of *E. coli* contain a defective prophage, termed Rac, that is known to encode a Ral-like antirestriction function, Lar. The inability to detect Ral activity was found to correlate with the presence of the Rac prophage; presumably *lar* is expressed in a sub-population of Rac$^+$ hosts. The gene encoding *lar* was isolated by cloning DNA fragments from the Rac prophage in a plasmid vector and screening for anti-restriction activity. In this way, *lar* was localised to a 500 bp DNA fragment, and was further defined by site-directed mutagenesis in conjunction with an analysis of the phenotypes of the mutants, and the polypeptides produced. The predicted amino acid sequences of Ral and Lar are remarkably dissimilar, although a number of residues do align which may indicate important features for structure and/or function.

The *ral* gene has been overexpressed and its product purified. Methylation assays, with purified Ral and the type IA *Eco*KI MTase failed to show any change in methylation activity in the presence of Ral. There are many reasons why no effect of

Ral was detected, not least that the purification process may disrupt the polypeptides. In order to overcome these problems, similar experiments were performed using extracts of cells overexpressing *ral* and the genes encoding the *Eco*KI MTase. These extracts, like extracts from cells producing only the MTase, methylated synthetic hemimethylated substrates efficiently, but only methylated unmodified substrates poorly. This is inconsistent with the hypothesis that Ral changes the activity of *Eco*KI from a maintenance MTase to a *de novo* MTase. An alternative explanation, consistent with present results and observations, is that Ral both blocks ENase activity and stimulates MTase activity.

Declaration

The composition of this thesis and the work presented within it are, unless otherwise stated, my own. Many of the approaches to the work were devised in collaboration with my supervisor, Professor N.E. Murray.

Acknowledgements

First and foremost I would like to thank my supervisor, Noreen Murray, for her encouragement, support and ideas.

I would also like to thank all the people who have helped me with techniques or services, including A. Brown, L. Cooper, A. Cronshank, D. Dryden, J. Fothergil, G. Leadbetter, L. Powell, K. Shaw and A. Titheradge.

Finally, I'm grateful to all those people who made my stay in Edinburgh so enjoyable and my family for a place to go home to. Thanks Jane, Pete, Laurie, Damion, Vicki, Jules, Catherine, Thorsten, Noreen, Dave, Lynn, Marcus, Annette, Diane, John, Glenn, Petra, Ruth, Rex, Anne, Tracy, Tanya, Phil and all the others, too numerous to mention.

Table of Contents

Abstract i
Declaration iii
Acknowledgements iv
Table of contents v
Abbreviations viii
Conventions xi
Amino acids xii
The genetic code xiii

Chapter one: Introduction

Historical introduction 1
Type I R-M systems 2
Genetic determinants 2
Sequence comparisons 4
The enzymes 6
Reaction mechanism 11
Evolution of new specificities 12
Type II R-M systems 15
Type III R-M systems 18
Modification-dependent systems 19
Distribution of R-M systems 22
Regulation of R-M activity 23
Phage antirestriction mechanisms 25
Modification of phage DNA 26
Inactivation of R-M systems 28
Modulation of R-M activity 29
Possible roles of R-M systems 32

Chapter two: Materials and methods

Bacterial strains 36
Phages 36
Plasmids 36
Synthetic oligonucleotides 36
Enzymes, isotopes and chemicals 36
Standard solutions and buffers 43
Media 45

Manipulation of bacteria and phages 45
Long-term storage of bacterial cells 45
Preparation of plating cells 45
Preparation of λ lysates 46
Restriction alleviation tests 46
Modification enhancement tests 46
Construction of strains carrying *hsdR*Δ4 47
Preparation of λDE3 lysogens 47
P1 transduction 48
Construction of NM793(DE3) 48
Manipulation of nucleic acids 49
Large-scale preparation of plasmid DNA 49
Removal of proteins with phenol 50
Precipitation of DNA with ethanol 50
Small-scale preparation of plasmid DNA 50
Small-scale preparation of phage λ DNA 50
Cutting of DNA with type II ENases 51
Removal of 5'-phosphates 51
Generation of blunt-ended fragments 51
Ligation of DNA 52
Agarose gel electrophoresis 52
Recovery of DNA from agarose gels 52
Transformations 52
The polymerase chain reaction 53
End-labelling of oligonucleotides 54
Nucleotide sequence analysis 54
DNA-denaturing PAGE 55
Autoradiography 55
Fluorography 55
Manipulation of proteins 56
Columns and media 56
Determination of protein concentration in solution 56
Coupling of Ral to keyhole limpet hemocyanin 57
Coupling of Ral to NHS-activated Sepharose 57
Gel retardation 58
Non-denaturing PAGE 58
Labelling of proteins 58

Tris-tricine SDS-PAGE 59
Staining of proteins with coomassie 60
Western blotting 60
Analysis of DNA and protein sequences 61
Isoelectric 61
Compare and DotPlot 61
Gap 61

Chapter three: Ral activity *in vivo*

Introduction 62
Single round of infection with modified phages 64

Chapter four: The Ral protein

Overexpression of the *ral* gene of bacteriophage λ 69
Production and purification of Ral Abs 76
Purification of the *ral* gene product 78
Analysis of Ral activity *in vitro* 86
In vitro methylation 89
Gel retardation 91
In vitro methylation with cell extracts 91
Discussion 95

Chapter five: The Lar activity of bacteriophage λ*reverse*

Introduction 101
Identification of the *lar* gene 103
Identification of the *lar* gene product 118
In vitro methylation with cell extracts 121
Comparisons between the predicted aa sequences of Ral and Lar 121
Discussion 123

Chapter six: Conclusions and general discussion

Conclusions and general discussion 128

References

References 132

Appendix

Publications 156

Abbreviations

A	adenosine
aa	amino acid
Ab	antibody
AdoMet	S-adenosyl methionine
AMPS	ammonium persulphate
ATP	adenosine triphosphate
bp	base pair
C	cytosine
C-	carboxyl-
cfu	colony forming units
Ci	Curie
cm	centimetre
CM	carboxymethyl
CTP	cytosine triphosphate
ddNTP	dideoxynucleoside triphosphate
DEAE	diethylaminoethyl
dH_2O	deionised water
DMF	dimethylformamide
DMSO	dimethyl sulphoxide
DNA	deoxyribonucleic acid
dNTP	deoxynucleoside triphosphate
DTT	dithiothreitol
EDTA	diaminoethanetetra acetic acid
ENase	DNA endonuclease
eop	efficiency of plating
g	gram

g	standard acceleration due to gravity
G	guanosine
GTP	guanosine triphosphate
hmC	hydroxymethyl cytosine
HPLC	high performance (high pressure) liquid chromatography
hr	hour
IPTG	isopropyl-β-D-thiogalactopyranoside
kb	kilobase
kDa	kilodalton
KLH	keyhole limpet hemocyanin
l	litre
m	milli (10^{-3})
M	molar
MALDI-TOF	matrix-assisted, laser-desorption ionisation, time of flight
MBS	maleimidobenzoic acid-N-hydroxysuccinimide ester
mcs	multiple cloning sites
MES	2-[N-morpholino]ethanesulphonic acid
min	minute
mm	millimetre
moi	multiplicity of infection
mol	mole
MTase	DNA methyltransferase
n	nano (10^{-9})
N-	amino-
NHS	N-hydroxysuccinimide
nm	nanometre
NMR	nuclear magnetic resonance
nt	nucleotide

NTP	nucleoside triphosphate
OD	optical density
orf	open reading frame
PAGE	polyacrylamide gel electrophoresis
PBS	phosphate-buffered saline
PCR	polymerase chain reaction
PEG	polyethylene glycol
pfu	plaque forming units
pH	$-\log_{10}[H^+]$
phage	bacteriophage
PMSF	phenylmethylsulphonyl fluoride
PNK	polynucleotide kinase
PVDF	polyvinylidene difluoride
R-M	restriction and modification
RNA	ribonucleic acid
SDS	sodium dodecyl sulphate
sec	second
sri	single round of infection
TBE	Tris-borate-EDTA (buffer)
TBS	Tris-buffered saline
TE	Tris-EDTA (buffer)
TEMED	N,N,N',N'-tetramethyl ethylenediamine
TFA	trifluoroacetic acid
tir	translational initiation region
TRD	target recognition domain
Tris	2-amino-2-(hydroxymethyl)-1,2-propandiol
U	unit
uv	ultraviolet

UWGCG	University of Wisconsin Genetics Computer Group
V	volt
v/v	volume per volume
W	watt
w/v	weight per volume
Δ	deletion
ε	extinction coefficient
μ	micro (10^{-6})
μF	microfaraday
Ω	ohm
°C	degree centigrade

Conventions

1. Bacterial genes are in italics and their products start with a capital letter, e.g. the *hsdR* gene encodes HsdR. Bacteriophage genes occasionally follow different conventions, e.g. the *c*I gene.
2. The nomenclature for restriction and modification (R-M) systems is according to Smith and Nathans (1973).
3. Prophages are indicated in brackets after the host, e.g. BL21(DE3).
4. The host on which a bacteriophage was last propagated is indicated after the phage symbol, e.g. λ.C600, λ.K, and λ.0 indicate, respectively, that bacteriophage λ was grown on C600, on an *Eco*KI-modifying strain and on a strain with no R-M system.
5. Restriction and modification phenotypes are indicated with superscript negative or positive symbols, e.g. r^-m^+, restriction-deficient modification-proficient strain. Subscript letters are used to indicate the R-M system associated with the phenotype, e.g. $r_K^-m_K^+$, restriction-deficient modification-proficient for *Eco*KI.

Amino acids

Amino acids	Three-letter symbol	Symbol
Alanine	Ala	A
Arginine	Arg	R
Asparagine	Asn	N
Aspartic acid	Asp	D
Cysteine	Cys	C
Glutamine	Gln	Q
Glutamic acid	Glu	E
Glycine	Gly	G
Histidine	His	H
Isoleucine	Ile	I
Leucine	Leu	L
Lysine	Lys	K
Methionine	Met	M
Phenylalanine	Phe	F
Proline	Pro	P
Serine	Ser	S
Threonine	Thr	T
Tryptophan	Trp	W
Tyrosine	Tyr	Y
Valine	Val	V

The genetic code

2nd (across) 1st (down)	U	C	A	G	3rd (down)
	Phe	Ser	Tyr	Cys	U
U	Phe	Ser	Tyr	Cys	C
	Leu	Ser	Ochre	Opal	A
	Leu	Ser	Amber	Trp	G
	Leu	Pro	His	Arg	U
C	Leu	Pro	His	Arg	C
	Leu	Pro	Gln	Arg	A
	Leu	Pro	Gln	Arg	G
	Ile	Thr	Asn	Ser	U
A	Ile	Thr	Asn	Ser	C
	Ile	Thr	Lys	Arg	A
	Met	Thr	Lys	Arg	G
	Val	Ala	Asp	Gly	U
G	Val	Ala	Asp	Gly	C
	Val	Ala	Glu	Gly	A
	Val	Ala	Glu	Gly	G

Chapter one: Introduction

Historical introduction

Observations consistent with the activity of restriction and modification (R-M) systems were made as early as the 1930's (Craigie and Yen, 1938), although an explanation based on the non-heritable, host-controlled modification of bacteriophage (phage) wasn't offered until the 1950's (Luria and Human, 1952; Bertani and Weigle, 1953). Bertani and Weigle (1953) noted that phage lambda propagated on *E. coli* K-12 grew on the *E. coli* strains K-12 and C with the same efficiency, however, phage propagated on *E. coli* C grew only poorly on *E. coli* K-12. Furthermore, only a single round of growth on *E. coli* C was required for the phage to plate with a low efficiency on *E. coli* K-12.

An idea to account for the restricted growth of unmodified phage was proposed by S. Lederberg (1957) who suggested "an agent which .... interferes specifically with phage DNA. A possible candidate could be a DNase". Work done by Arber and colleagues in the early 1960's, identified the mechanism of host controlled modification. The fate of λ grown on *E. coli* K-12 lysogenic for phage P1 was followed in a single round of growth on *E. coli* K-12 . It was found that about the same number of progeny as were used for the infection carried the parental host specificity, λ.K (P1), and the rest were protected against restriction by *E. coli* K-12 not lysogenic for phage P1. Therefore, the host-specificity of the parent is transferred to the progeny, but is not replicated. The link between transferred host specificity and parental DNA was demonstrated using deuterium-labelled DNA; labelled phages have a higher than normal density. Only DNA containing at least one parental strand retained the specificity of the parent, λ.K (P1), and non-deuterated progeny phage conferred the host-specificity of the most recent host, λ.K. Thus, the host-controlled modification is carried on the phage DNA molecule (Arber and Dussoix, 1962). ^{32}P-labelled DNA and genetic marker rescue experiments with phage λ showed that the unmodified DNA was digested by the host cell shortly after injection of the DNA (Arber and Dussoix, 1962; Dussoix and Arber, 1962). Preliminary work also suggested that exchange of bacterial DNA might also be affected by host-specificity (Arber and Dussoix, 1962), and was proven using crosses between *E. coli* K-12 Hfr donors and F$^-$ K-12 (P1) recipients (Arber and Morse, 1965).

The observation that a methionine auxotroph of *E. coli* K-12 starved of methionine is unable to confer the host-controlled modification, led Arber (1965a) to suggest that the host-specificity was a result of methylation of particular sites on the phage DNA.

Arber (1965b) also proposed that host-specificity might be determined by a two enzyme system; one enzyme a sequence-specific DNA endonuclease (restriction endonuclease, ENase), cleaving unmodified DNA, and the other a DNA methyltransferase (MTase), altering the DNA so that it was no longer susceptible to cleavage. The DNA target site for the ENase and MTase were later shown to be the same (Vovis and Zinder, 1975).

An ENase was identified by Takano and colleagues (1966), using cell-free extracts of restrictive bacteria with purified λ DNA as a substrate. They found that unmodified DNA was efficiently degraded by the extracts, whereas modified DNA was not. This confirmed the idea of 'restriction enzymes'. On this basis, ENases were purified from cell extracts of *E. coli* K-12 and B (*Eco*KI and *Eco*BI; Meselson and Yuan, 1968; Linn and Arber, 1968; Roulland-Dussoix and Boyer, 1969), and found to have a molecular mass of 400 kDa (Meselson *et al.*, 1972). Both enzymes required ATP, S-adenosyl-L-methionine (AdoMet) and Mg^{2+} to introduce a number of double stranded scissions in either unmodified phage or bacterial DNA. A 'simpler' restriction enzyme, *Hin*dII, was purified from *Haemophilus influenzae* that only required Mg^{2+} to cleave unmodified DNA (Smith and Wilcox, 1970). It became clear there were at least two types of R-M enzymes, the simpler type-II (e.g. *Hin*dII) and the more complex type-I (e.g. *Eco*KI and *Eco*BI; Boyer, 1971). As more R-M enzymes were identified, and another group of ATP-dependent endonucleases was found (Kauc and Piekarowitz, 1978), the host-specificity systems were rearranged into three groups (Yuan, 1981).

Type I R-M systems

Genetic determinants

The observation that bacterial DNA was susceptible to host-controlled restriction and modification (Dussoix and Arber, 1962) enabled Boyer (1964) to position the *hsd* locus, for host-specificity of DNA, on the *E. coli* chromosome by Hfr mapping. He found that the restriction and modification genes were linked, and that the *hsd* loci of *E. coli* K-12 and B were allelic. P1 transduction experiments positioned the *hsd* loci close to *serB* at 98.5 minutes on the *E. coli* chromosome (Glover and Colson, 1969).

The first restriction deficient (r^-) derivatives of *E. coli* K-12 and B were isolated by Wood (1966). He found that the r^- mutants were either deficient or proficient in modifying ability, r^-m^- or r^-m^+ respectively. He proposed that a common gene product was required for both restriction and modification, which was absent in r^-m^- strains. Complementation studies by Boyer and Roulland-Dussoix (1969), using Hfr,

identified three genes involved in host-controlled restriction and modification, subsequently termed *hsdR, M* and *S*. A mutation in *hsdR* resulted in the loss of restriction ability but not modification activity (r^-m^+). Mutations in two other genes abolished both restriction and modification activity (r^-m^-). No mutants were isolated that abolished just modification activity (r^+m^-) presumably as this would lead to auto-restriction. The host-specificity genes of *E. coli* K-12 and B were shown to be sufficiently related that their gene products could complement each other. This work also demonstrated that the *hsdS* gene product was responsible for the specificity of both the restriction and modification activities of the system. Continuing this work Glover (1970) proposed that the *hsdR, M* and *S* gene products interact to form two oligomeric enzymes, HsdM and S forming a strain-specific modification enzyme, and HsdR, M and S forming a strain-specific restriction enzyme. This was supported by the isolation of a temperature sensitive mutation, $r^-m_{ts}^+$, which was mapped to *hsdM*, first by complementation studies (Hubacek and Glover, 1970) then by DNA sequence analysis (A. Daniel, unpublished).

Borck and colleagues (1976) cloned the *E. coli* K-12 *hsdM* and *hsdS* genes in phage λ. Clones carrying the genes for K-specific modification were recognised as they grew with a high efficiency on *E. coli* K-12 after propagation on a non-modifying host. One of these phages was then used to recover a clone containing the whole *hsd* region, following integration by homologous recombination and aberrant excision resulting in the formation of λ*hsdRMS* (Sain and Murray, 1980). These were identified as lysogens of *E. coli* C, a strain that does not possess an R-M system (Raleigh, 1992), that were phenotypically r^+m^+. Deletion derivatives of these phages and complementation tests showed the gene order to be *hsdR, M* then *S* (Sain and Murray, 1980).

At the beginning of the 1980's, five allelic type I R-M systems had been identified, *Eco*KI (from *E. coli* K-12), *Eco*BI (from *E. coli* B; Glover and Colson, 1969), *Eco*AI (from *E. coli* 15T$^-$; Arber and Wauters-Willems, 1970), *Sty*LTIII (from *S. enterica* serovar *typhimurium* LT2; Van Pel and Colson, 1974) and *Sty*SPI (from *S. enterica* serovar *potsdam*; Bullas and Colson, 1975). DNA hybridisation studies and immunological methods, using antibodies against *Eco*KI, indicated sequence similarity between *Eco*KI, *Eco*BI, *Sty*LTIII and *Sty*SPI, but none to *Eco*AI (Murray *et al.*, 1982). As more type-I R-M systems were identified (Price *et al.*, 1987a), it became clear that they could be further divided on the basis of complementation analysis (Arber and Linn, 1969; Glover, 1970; Price *et al.*, 1987a), DNA cross-hybridisation (Sain and Murray, 1980; Daniel *et al.*, 1988), and immunological cross reactivity (Murray *et*

al., 1982). *Eco*KI and relatives constitute the type-IA R-M systems, *Eco*AI the type-IB, and *Eco*R124I the type-IC. There is now evidence for a fourth group, type-ID, the first representative of which is *Sty*SBLI (from *S. enterica* serovar *blegdam*; A. Titheradge and N. Murray, unpublished; Barcus *et al.*, 1995). The type IA, IB and ID *hsd* genes behave as alleles, and yet some alleles are so dissimilar that their DNA fails to cross-hybridise (e.g. *Eco*KI and *Eco*AI; Murray *et al.*, 1982). Comparison of the nucleotide sequence of the *hsd* genes, has confirmed this allelic diversity. Some alleles of *hsdR* and *hsdM* encode polypeptides with very similar amino acid (aa) sequences, even when from different species e.g. 94% identity between *Eco*KI and *Sty*LTIII. However, other alleles encode remarkably dissimilar polypeptides, irrespective of the host species, and retain little evidence of a common ancestry e.g. 32% identity between *Eco*KI and *Eco*AI (Sharp *et al.*, 1992; Murray *et al.*, 1993). These genes have been termed pseudoalleles (Sharp *et al.*, 1992). This unusual allelic diversity is shared by a number of other genes that are concerned with differentiating 'self' from 'non-self', for example, the major histocompatibility complex class II locus of mammals and the self-incompatibility genes of plants. This allelic diversity indicates a strong selection for variation, possibly aided by the lateral transfer of short DNA segments between different species (Sharp *et al.*, 1992).

Sequence comparisons

Sequence comparisons, between different R-M systems and sequences in the database, have been very useful in identifying conserved regions that may serve structural or functional roles. A comparison of the predicted aa sequences of the *hsdR* genes of *Eco*KI, *Eco*AI and *Eco*R124I has identified seven regions conserved among a family of putative helicases, the so-called 'DEAD-box' proteins (Gorbalenya and Koonin, 1991; Murray *et al.*, 1993). Region I is an abridged version of the 'A' component of the NTP-binding motif (Walker *et al.*, 1982), and ATP is known to be both a cofactor and substrate in restriction by type I enzymes (Bickle *et al.*, 1978). Region II includes the characteristic aa sequence DExH, where x is either an A or C. Mutational analysis of this region indicates that H is essential for restriction by *Eco*KI *in vivo*, as either an N or D substitution at this position abolishes activity (G. King, D. Ternent, S. Bolton and N. Murray, ms in preparation). It is interesting that substituting D for H blocks restriction, as this is a substitution to the super-family consensus, DEAD. Mutational analysis of another 'DEAD-box' protein, the translational initiation factor eIF4A, demonstrated that activity was abolished if the DEAD motif was changed to DEAH (Pause and Sonenberg, 1992). Sequence comparisons, identifying all seven conserved motifs, have placed the DExH and DEAD-box proteins in two distinct

groups, and all of the putative DNA helicases contain the DExH motif (Gorbalenya and Koonin, 1991). Two changes have been noted in the original *Eco*KI sequence (Sain and Murray, 1980; Blattner *et al.*, unpublished; GeneBank acc. no. U14003) which has meant that region III, the SAT motif that was difficult to identify in the original alignments (Murray *et al.*, 1993), has been realigned. The presence of these motifs in HsdR, and the preliminary analysis of mutants in the DExH motif, would indicate that the type I R-M systems may be ATP-dependent DNA helicases.

A comparison of the predicted aa sequences of a range of methyltransferases has identified a number of conserved motifs. The sequence (D/E/S)XFXGXG, termed motif I, is found in all MTases that use AdoMet as the methyl donor (Klimisaukas *et al.*, 1989; Pósfai *et al.*, 1989). The co-crystal structure of *Hha*I MTase and AdoMet confirmed that these residues interact with AdoMet (Cheng *et al.*, 1993; Klimisaukas *et al.*, 1994). A second motif, motif II (N/D/S)PP(Y/F), is also aligned in MTases that methylate the N^6 position of adenine, like the MTases of the type-I R-M systems, or the N^4 position of cytosine (Klimisaukas *et al.*, 1989; Pósfai *et al.*, 1989; Sharp *et al.*, 1992). Mutational analysis of these two motifs in *Eco*KI confirms that motif I is required to bind AdoMet. Changes have been made to motif II which don't affect AdoMet or DNA binding, and yet abolish enzyme activity. These results imply a role for motif II in the catalysis of methyl group transfer (Willcock *et al.*, 1994). The sequence of HsdM that includes motif I and motif II has been aligned with the type II MTases, and identified a common AdoMet-binding domain (Dryden *et al.*, 1995). The crystal structures of a number of type II MTases have been determined (Klimisaukas *et al.*, 1994; Labahn *et al.*, 1994) and the tertiary structure of this domain from the type I MTase of *Eco*KI, has been modelled on that of the type II enzyme, *Hha*I. These results suggest a common evolutionary origin for all MTases (Dryden *et al.*, 1995).

The MTases of the type IA R-M systems are the only prokaryotic MTases shown to have a strong preference for hemimethylated DNA *in vivo*. Mutants of *E. coli* K-12 have been isolated that no longer show this preference for hemimethylated DNA, and efficiently methylate unmodified target sites (Kelleher *et al.*, 1991). These mutations are loosely clustered in *hsd*M, and the activity of these mutants indicate *de novo* methylation by *Eco*KI.

The three type I R-M systems *Eco*KI, *Eco*BI and *Eco*DI recognise different DNA target sequences. When the nucleotide sequence of the *hsdS* genes from these systems was determined, it was apparent that although they share very similar functions, they are remarkably dissimilar in primary structure (Gough and Murray, 1983). Gough and Murray (1983) identified two highly conserved regions in the nucleotide and predicted

amino acid sequence, flanked by variable regions that they predicted would encode the target recognition domains (TRDs). As the nucleotide sequence of more *hsdS* genes were determined, it became clear that this pattern of variable and conserved regions was common to all type I R-M systems that recognised different DNA target sequences (figure 1.1; for recent reviews see Bickle, 1993; Barcus and Murray, 1995). Argos (1985) identified a short repeat in the predicted aa sequence of *Eco*KI, *Eco*BI and *Eco*DI that overlapped the two conserved regions in HsdS. These repeats have since been found in all HsdS polypeptides (figure 1.1; Kannan *et al.*, 1989), and residues in these regions are now believed to be involved in protein:protein interactions with HsdM or R (Cooper and Dryden, 1994; Kneale, 1994).

The specificities of many of the type-I R-M enzymes have been elucidated (see table 1.1) by identifying DNA fragments containing targets, either by the ^{3}H-methyl-labelling of DNA *in vitro* (e.g. Nagaraja *et al.*, 1985a) or the restriction of recombinant phages (e.g. Gann *et al.*, 1987), and using computers to compare the sequence of the mapped fragments. The recognition sequences are asymmetric and bipartite, consisting of two short, defined sequence components separated by a non-specific spacer of fixed length. Two specific adenine residues within the recognition sequence are methylated (Kühnlein and Arber, 1972). The conserved position of these residues at 10 or 11 base pairs apart, suggest that the enzyme may bind to the DNA along one face of the double helix interacting with the adenine residues in two successive major grooves with most of the non-specific spacer sequence in the minor groove (Nagaraja *et al.*, 1985b). Methylation interference footprinting, of the *Eco*KI MTase and DNA, confirmed that the base contacts are in the major groove (Powell and Murray, 1995). UV crosslinking of an oligonucleotide containing a substitution of thymidine with the photoreactive base-analogue bromodeoxyuridine, in the AAC component of the *Eco*KI recognition sequence, also demonstrated a close contact between the recognition sequence in the major groove of the DNA and the MTase (Chen *et al.*, 1995). Proteolysis of the crosslinked complex and peptide sequencing, identified a tyrosine in the amino-variable domain of the HsdS subunit as the site of crosslinking, and this residue is conserved among type I systems that have AA in the trinucleotide component of their recognition sequence.

The enzymes

The restriction enzymes from *E. coli* K-12 and B (*Eco*KI and *Eco*BI) were purified in the mid 1960's (Meselson and Yuan, 1968; Linn and Arber, 1968; Roulland-Dussoix and Boyer, 1969). Enzymes were also purified from $r^{-}m^{+}$ and $r^{-}m^{-}$ strains of *E. coli* K-

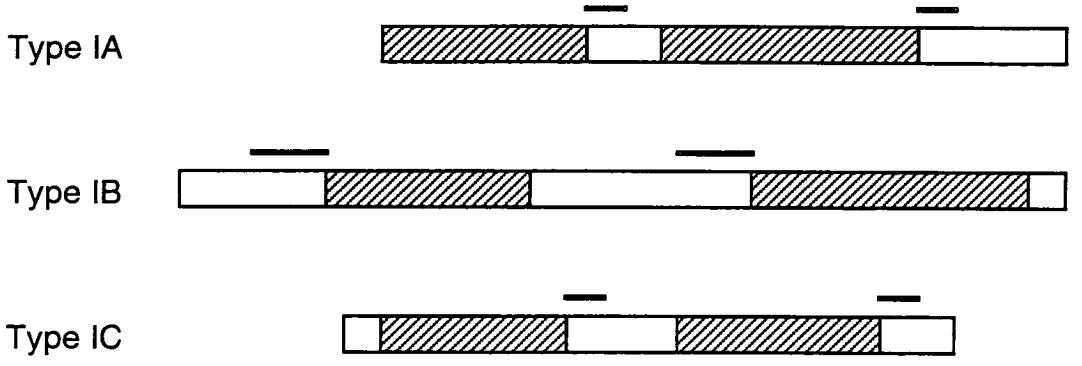


Figure 1.1. Structure of the HsdS subunits. Hatched box, variable regions; open box, conserved regions; bars, direct repeats.

Type	System	Specificity	Reference
IA	*Eco*KI	AAC (N_6) GTGC	Kan *et al.*, 1979
	*Eco*BI	TGA (N_8) TGCT	Lautenberger *et al.*, 1978
			Ravetch *et al.*, 1978
	*Eco*DI	TTA (N_7) GTCY	Nagaraja *et al.*, 1985a
	*Sty*SPI	AAC (N_6) GTRC	Nagaraja *et al.*, 1985b
	*Sty*LTIII	GAG (N_6) RTAYG	Nagaraja *et al.*, 1985b
	*Eco*R5I	new	Barcus *et al.*, 1995
	*Eco*R10I	new	Barcus *et al.*, 1995
	*Eco*R11I	same as *Eco*BI	Barcus *et al.*, 1995
	*Eco*R12I	same as *Eco*KI	Barcus *et al.*, 1995
	*Eco*R13I	new	Barcus *et al.*, 1995
	*Eco*R14I	same as *Eco*BI	Barcus *et al.*, 1995
	*Eco*R23I	new	Barcus *et al.*, 1995
	*Eco*R24I	same as *Eco*KI	Barcus *et al.*, 1995
	*Eco*R25I	same as *Eco*R10	Barcus *et al.*, 1995
	*Eco*R70I	same as *Eco*BI	Barcus *et al.*, 1995
IB	*Eco*AI	GAG (N_7) GTCA	Suri *et al.*, 1984b
	*Eco*EI	GAG (N_7) ATGC	Cowan *et al.*, 1989
	*Sty*STI	unknown	A. Titheradge and N. Murray, unpublished
	*Sty*SKI	unknown	D. Ternent, unpublished
	*Cfr*AI	GCA (N_8) GTGG	Kannan *et al.*, 1989
	*Eco*R15I	unknown	Barcus *et al.*, 1995
	*Eco*R17I	unknown	Barcus *et al.*, 1995
	*Eco*R42I	new	Barcus *et al.*, 1995
IC	*Eco*R124I	GAA (N_6) RTCG	Price *et al.*, 1987b
	*Eco*DXXI	TCA (N_7) RTTC	Piekarowicz and Goguen, 1986
	*Eco*prrI	CCA (N_7) RTGC	Tyndall *et al.*, 1994
ID	*Sty*SBLI	unknown	A. Titheradge, unpublished
	*Eco*R9I	unknown	Barcus *et al.*, 1995
	*Eco*R31I	unknown	Barcus *et al.*, 1995
	*Eco*R65I	unknown	Barcus *et al.*, 1995
	*Eco*R68I	unknown	Barcus *et al.*, 1995

Table 1.1. Type I R-M systems and their recognition sequences.

12 and B, and found to have the *in vitro* properties predicted from the phenotypes of the strains that they were purified from (Kühnlein *et al.*, 1969; Hadi and Yuan, 1974). Further work estimated the molecular mass of *Eco*KI to be 400 kDa. When denatured the enzyme could be separated into three subunits of estimated molecular mass 135 kDa (α), 62 kDa (β) and 52 kDa (γ). The molecular ratio was estimated to be $\alpha_2\beta_2\gamma_1$ (Meselson *et al.*, 1972). Similar results were found for *Eco*BI, and inconsistencies probably reflect differences in the methods used (Eskin and Linn, 1972a).

Lautenberger and Linn (1972) purified the MTase from *E. coli* B and found it to possess the two smaller subunits, β and γ, demonstrating that the large subunit, required for restriction, was dispensable for modification activity. Polypeptide analysis of λ*hsdRMS*, and deletion derivatives of this phage, correlated the 135 kDa polypeptide with *hsdR*, the 62 kDa polypeptide with *hsdM*, and the 52 kDa polypeptide with *hsdS* (Sain and Murray, 1980). The modification MTase was purified from *E. coli* K-12 lysogenic for λ*hsdMS* (Suri *et al.*, 1984a) and found to have the same properties as the *Eco*BI MTase purified by Lautenberger and Linn (1972). A number of kinetic studies were performed on the methylation properties of the restriction enzyme and purified methyltransferase. It was found that the methylation of hemimethylated pBR322, obtained by hybridising modified with unmodified DNA, was much faster with both enzymes than methylation of unmodified DNA. For the restriction enzyme the figure is about 150 times faster, and the MTase 35 times.

Large quantities of the *Eco*KI MTase have been purified (Dryden *et al.*, 1993), and used to examine the protein:protein and protein:DNA interactions in more detail. The subunit composition of the purified MTase is $HsdM_2S_1$, and a domain structure of the MTase has been derived from limited proteolysis experiments (figure 1.2; Cooper and Dryden, 1994). The purified *Eco*KI MTase has the expected preference for hemimethylated DNA substrates (Dryden *et al.*, 1993), but this preference can't be explained by differences in the binding affinity of the MTase for unmodified or hemimethylated oligonucleotides that contain the *Eco*KI target site (Powell *et al.*, 1993). However, differences are detected by the methylation interference footprinting of *Eco*KI MTase with DNA of different methylation states (Powell and Murray, 1995). In the presence of the cofactor, AdoMet, there is no interference outside of the recognition sequence with an unmodified oligonucleotide, but the interference extends into the non-specific, spacer-sequence when the target is hemimethylated. Powell and Murray (1995) conclude that whether *Eco*KI restricts or modifies is determined by conformational changes in the enzyme, dependent on the methylation state of the DNA

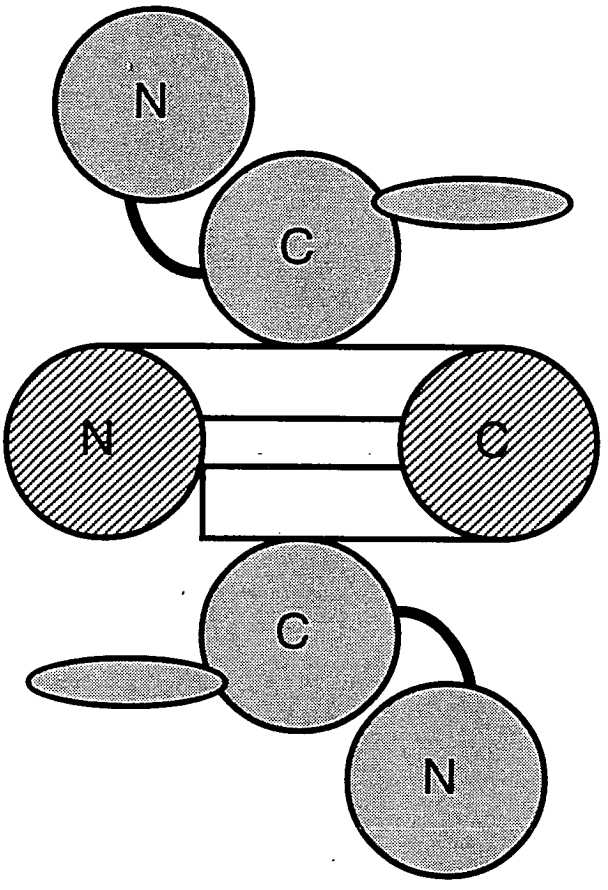


Figure 1.2. A representation of the domain structure of the *Eco*KI MTase. Hatched circles, variable regions of HsdS; open box, conserved regions of HsdS; Stippled objects, HsdM domains; N, amino-terminal domain; C, carboxyl-terminal domain. Motif I is in the N-terminal domain of HsdM, and motif II is in a short exposed loop between the N and C-terminal domains. The C-terminal domain of HsdM contacts the HsdS subunit, and there is a 'tail' that is sensitive to the methylation state of the DNA (Cooper and Dryden, 1994; Willcock *et al.*, 1994).

and the presence of AdoMet. This is detected in changes in the DNA:protein interface with the MTase and AdoMet, and supports the hypothesis that the cofactor is used to communicate the methylation state of the DNA to the enzyme (Burckhardt *et al.*, 1981b; Powell and Murray, 1995).

Reaction mechanism

Some of the features of the reaction mechanism for *Eco*KI and *Eco*BI have been elucidated *in vitro* (Burckhardt *et al.*,1981a).

1. *Eco*KI rapidly binds AdoMet.
2. *Eco*KI-AdoMet undergoes a slow transition to an activated form, *Eco*KI*, which continues to have AdoMet bound (Hadi *et al.*, 1975).
3. *Eco*KI* forms a non-specific complex with the DNA.
4. If the *Eco*KI recognition site is present on the DNA a more stable complex is formed (Yuan *et al.*, 1975).
5. Depending on the methylation state of the target sequence, one of several things can happen:
   - With a fully methylated site and ATP, the enzyme is released.
   - If the site is hemimethylated, the unmethylated adenine is rapidly methylated. This occurs some 150 times faster than the methylation of an unmodified target sequence (Suri *et al.*, 1984). ATP and Mg^{2+} stimulate methylation by the restriction enzyme (Vovis *et al.*, 1974), however, ATP does not stimulate the activity of the *Eco*KI methylase (Suri *et al.*, 1984), nor is Mg^{2+} required for DNA binding or methylation (Powell *et al.*, 1993; G. King, unpublished).
   - If the target site is unmodified, ATP induces the transition of *Eco*KI* to *Eco*KI^{+}. This form has lost AdoMet and shows a decrease in diameter, possibly due to rearrangements in the subunits of the complex (Bickle *et al.*, 1978; Burckhardt *et al.*, 1981a). The DNA is then translocated past the enzyme in an ATP-dependent reaction (Yuan *et al.*, 1980; Rosamond *et al.*, 1979), followed by the formation of a single-stranded, then double-stranded break in the DNA at a non-specific site (Murray *et al.*, 1973a; Meselson and Yuan, 1968). The enzyme does not turn over in the cleavage reaction, but it is still an active ATPase after DNA cutting is complete (Eskin and Linn, 1972b).

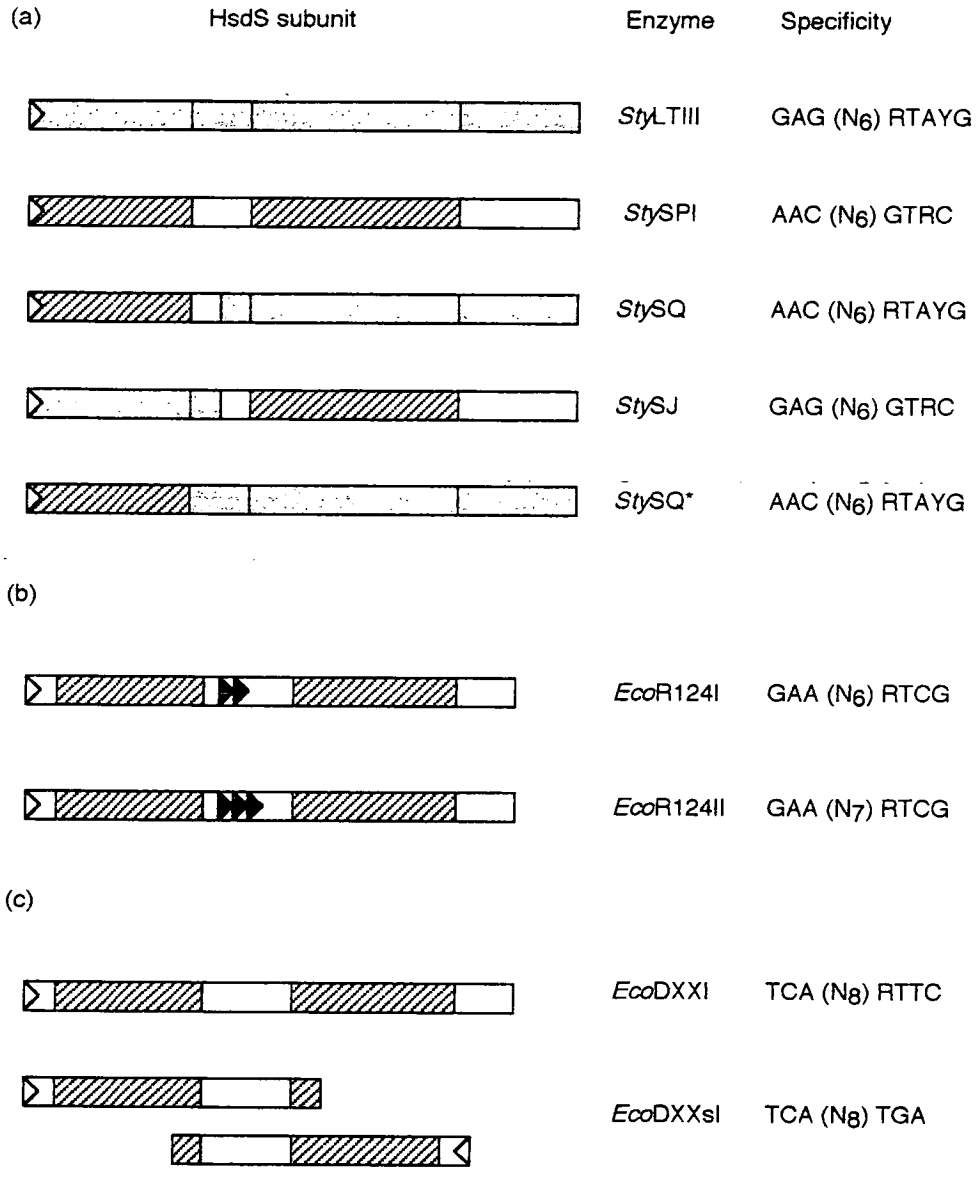
A model for the cleavage of DNA by *Eco*KI has been postulated (Burckhardt *et al.*, 1981b). The HsdS subunit binds weakly to the recognition sequence, along the minor groove. AdoMet binding to HsdM induces a conformational change in the enzyme, and the methyl groups of AdoMet would be used to probe the methylation state of the

target bases in the major groove. If the DNA was fully modified, the two HsdM subunits would not be able to enter the major grooves of the DNA, due to steric hindrance from the methylated adenines, and thus remain in an open configuration, *Eco*KI^{k}. If the DNA was hemimethylated , that is with only one methylated adenine residue in the recognition sequence, then one of the HsdM subunits would be able to bind to the major grove at the unmodified site. This would result in a partially open conformation, *Eco*KI^{h}, and methylation would occur at this site. With unmodified DNA, both HsdM subunits would be able to bind to the unmethylated sites and result in the formation of a closed conformation, *Eco*K^{o}, in which the two HsdR subunits form a functional restriction complex. In contradiction of this model, the footprinting results of Powell and Murray (1995) show that the *Eco*KI MTase binds to the major groove of the DNA. The HsdM subunits could still methylate the adenine in the major groove if the base was 'flipped' out of the DNA helix, as has been shown for the type II MTase, *Hha*I (Klimisaukas *et al.*, 1994). *In vitro* and *in vivo* results suggest that the restriction enzyme bound at the recognition sequence translocates DNA towards itself, from both directions, and when two enzymes collide they cut the DNA between them (Brammar *et al.*, 1974; Endlich and Linn, 1985; Studier and Bandyopadhyay, 1988). It has been demonstrated that the *Eco*KI MTase is able to distinguish the methylation state of the target DNA (Suri *et al.*, 1984a; Dryden *et al.*, 1993), thus it must be the interaction of the HsdM and S subunits with the DNA target sequence that determines the enzymatic action of the complex.

Evolution of new specificities

Changes in the specificity of type I R-M systems have been observed in the laboratory, both serendipitously, involving homologous recombination, unequal crossing over and transposon insertion, or by design using recombinant DNA techniques. The R-M system *Sty*SQ was the outcome of P1-mediated transduction between two strains with different type IA R-M systems (Bullas *et al.*, 1976). *Sty*SQ was shown to have arisen by homologous recombination between the *hsdS* central conserved regions of *Sty*SPI and *Sty*LTIII (figure 1.3a; Fuller-Pace *et al.*, 1984), and the sequence recognised by the recombinant system comprised one component of each parental system (Nagaraja *et al.*, 1985a; 1985b). An R-M system that recognised the other components of the parental systems, *Sty*SJ, was similarly isolated after recombination *in vivo* (Gann *et al.*, 1987). These experiments demonstrated that there were two target recognition domains, TRDs, one in each half of HsdS, that could be reassorted to generate novel DNA specificities. Cowan and colleagues (1989) further defined the target recognition domains by using site directed mutagenesis to generate a hybrid system, *Sty*SQ*,

Figure 1.3. Evolution of new specificities by rearrangements in *hsdS*. (a) Exchange of regions in *hsdS* identifies two target recognition domains (TRDs) corresponding to the two variable regions. (b) Increasing the length of the central conserved region, between the TRDs, can also increase the length of the non-specific 'spacer' sequence in the DNA target. (c) A type I R-M system may be isolated with only one kind of TRD. See text for more details. Shading indicates origin of regions of HsdS; hatched box, variable region; open box, conserved region; open arrows, amino-terminus; filled arrows, repeats; R, purine; Y, pyrimidine; N, any base.


(a)
HsdS subunit
Enzyme
Specificity
StyLTIII
GAG (N6) RTAYG
StySPI
AAC (N6) GTRC
StySQ
AAC (N6) RTAYG
StySJ
GAG (N6) GTRC
StySQ*
AAC (N6) RTAYG
(b)
EcoR124I
GAA (N6) RTCG
EcoR124II
GAA (N7) RTCG
(c)
EcoDXXI
TCA (N8) RTTC
EcoDXXsI
TCA (N8) TGA

containing only the HsdS N-terminal variable domain of *Sty*SPI and the remaining sequence of *Sty*LTIII. *Sty*SQ* recognises the same target sequence as *Sty*SQ, demonstrating that the N-terminal variable domain is all that is required to specify one half-site in the recognition sequence (figure 1.3a). By analogy the C-terminal variable domain recognises the other component (Cowan *et al.*, 1989). The same conclusions may be drawn from similar experiments that have been performed on the type IC R-M systems, *Eco*R124I and *Eco*DXXI (Gubler *et al.*,1992).

A second R-M system, *Eco*R124II, has been isolated that appears to have arisen by unequal crossing over between the central conserved regions of two *hsdS* genes from *Eco*R124I (Price *et al.*, 1989). Recombination lead to an extra 12 bp repeat in the *hsdS* central region of *Eco*R124II, compared to *Eco*R124I, and the two components of the recognition sequence are separated by an extra nucleotide (figure 1.3b). Again, similar results have been achieved with *Eco*DXXI (Gubler and Bickle, 1991).

Type I R-M systems have been isolated with only one TRD, either as a result of transposon insertion (Meister *et al.*, 1993), or mutagenesis (Abadjeva *et al.*, 1993). In both cases, the R-M system recognises a symmetrical sequence consisting of two half sites, each a single component of the recognition sequence from the parental system (figure 1.3c). These recognition sequences are consistent with two truncated HsdS subunits being required for DNA recognition, rather than the usual one.

The repeated sequences in *hsdS*, the primary structure of it's product, and the isolation of a type I R-M system with only one TRD, all indicate that *hsdS* may have arisen by the duplication of a gene encoding only one TRD (Argos, 1985; Kannan *et al.*, 1989). This duplication is most apparent in the IB family, in which repeated sequences are particularly well conserved (Kannan *et al.*, 1989).

Type II R-M systems

The type-II R-M systems are the simplest of the three classical systems. They have been defined as double-stranded endonucleases that recognise specific DNA sequences and cleave at a defined point within or close to that sequence (Roberts and Halford, 1993). The first type II ENase was purified in the late 1960's (Kelly and Smith, 1970; Smith and Wilcox, 1970), although it was later shown to be a mixture of two enzymes, *Hin*dII and *Hin*dIII (Old *et al.*, 1975). The use of type II ENases in the development of recombinant DNA technology meant that by the mid 1970's over 40 enzymes had been identified (see Roberts, 1976), and there are currently more than 2400 (Roberts and Macelis, 1993).

Most type II R-M systems consist of two separate enzymes with separate activities. The *res* gene encodes a Mg^{2+}-dependent site-specific DNA endonuclease. The *mod* gene encodes a AdoMet-dependent DNA methyltransferase that modifies the same sequence, either at the N^4 or N^5 position of cytosine, or the N^6 position of adenine. As expected, the *res* and *mod* genes have always been found linked, and efficient cloning strategies have been developed that take advantage of this (Lunnen *et al.*, 1988). The genes encoding type II R-M systems are found on both bacterial plasmids and chromosomes, and the nucleotide sequence for about 60 systems is now known (Roberts and Halford, 1993). Interestingly, no aa similarity is seen between the *res* and *mod* gene products of any cognate systems (Chandrasegaron and Smith, 1988; Roberts and Halford, 1993). This lack of similarity may be because the ENase and MTase of the systems examined recognise the DNA target in different ways; the ENase is normally a dimer and each subunit need only recognise half the recognition sequence, whereas the MTase acts as a monomer and must, therefore, recognise the complete sequence (Wilson and Murray, 1991).

The crystal structure of a number of type II ENases and MTases, either free or complexed with DNA, are now known. These include *Eco*RI (complexed with DNA; Kim *et al.*, 1990), *Eco*RV (complexed with DNA; Winkler *et al.*, 1993), *Bam*HI (Newman *et al.*, 1994), and *Pvu*II (complexed with DNA; Cheng *et al.*, 1994), and the MTases of *Hha*I (complexed with DNA; Klimasaukas *et al.*, 1994) and *Taq*I (Labahn *et al.*, 1994). *Eco*RI and *Bam*HI have little similarity in primary sequence, but a surprising amount of similarity in higher structures (Newman *et al.*, 1994). *Eco*RV shows little structural similarity to either *Eco*RI or *Bam*HI (Winkler *et al.*, 1993), but does to *Pvu*II (Athanasiadis *et al.*, 1994). The structure of *Eco*RV has been solved in three forms, free protein, complexed with non-specific DNA, and complexed with DNA containing the *Eco*RV recognition sequence. Sequence specific contacts, between *Eco*RV and its target, occur between aa in the recognition loop (R-loop) and the major grove of DNA. When the DNA target is methylated, and hence not a substrate for the ENase, this R-loop is displaced (Winkler *et al.*, 1993).

The co-crystal structure of the *Hha*I MTase, trapped in a reaction complex with an oligonucleotide containing the *Hha*I target sequence, shows a novel DNA:protein interaction (Klimisaukas, *et al.*, 1994). The cytosine to be methylated in the DNA target sequence is 'flipped out' of the DNA and into the active site of the protein, bringing all the components required for methylation together (figure 1.4; Klimisaukas *et al.*, 1994). The crystal structure of the *Taq*I MTase has also been solved, and is very similar to *Hha*I in the region of the active site (Labahn *et al.*, 1994). It seems

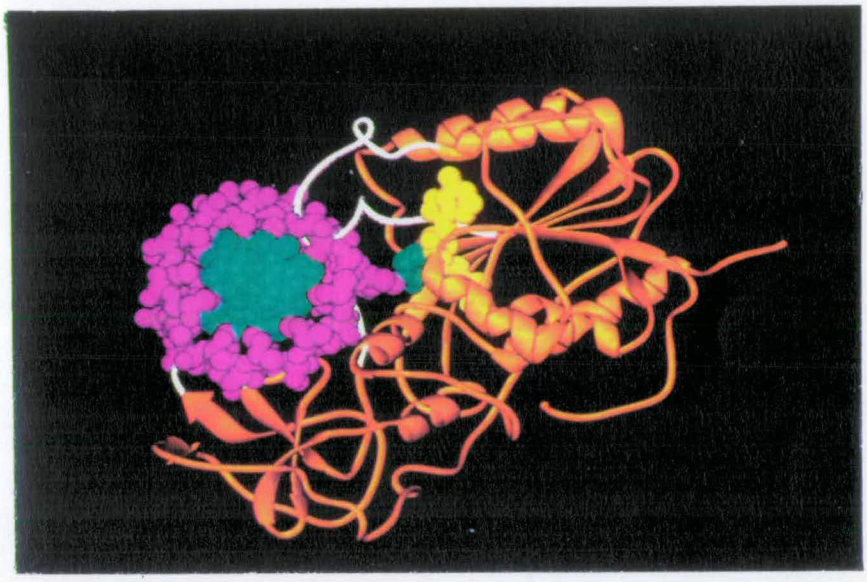


Figure 1.4. A view looking down the DNA axis of the *Hha*I MTase covalently bound to an oligonucleotide containing its recognition sequence. The protein is coloured brown, apart for the active-site loop and recognition loops which are white; the DNA backbone is magenta; bases are green; S-adenosyl homocysteine is yellow (Klimisaukas *et al.*, 1994)

likely that this method of accessing nucleotides in a DNA helix will be common to all DNA MTases, and maybe other DNA-modifying enzymes (Klimisaukas *et al.*, 1994).

There are some R-M systems currently classed as type II, which have rather different properties. The type-IIs are very similar to the type II R-M systems, except that they cleave a short distance from their recognition sequence, hence 'IIs' for 'shifted cleavage' (Szybalski *et al.*, 1991). Like the type IIs systems, *Eco*57I recognises an asymmetric target sequence and cleaves a short distance from this site. This enzyme is different, however, in that one subunit is both an ENase and MTase which has an absolute requirement for Mg^{2+} and whose activity is stimulated by AdoMet (Janulaitis *et al.*, 1992). *Eco*57I also has a cognate MTase. The R-M system *Bcg*I cleaves DNA bilaterally, either side of its recognition sequence, and consists of two subunits both of which are required to form an active ENase and MTase (Kong *et al.*, 1994). These unusual enzymes may represent new classes of R-M systems, and *Eco*57I has been tentatively termed a type IV (Janulaitis *et al.*, 1992).

Type III R-M systems

The type III R-M systems have the smallest number of members, currently only four (see Bickle, 1993). The *E. coli* phage P1 and defective prophage P15, encode very similar systems, *Eco*PI and *Eco*P15I respectively (Arber and Dussoix, 1962; Arber and Wauters-Willems, 1970). *Haemophilus influenzae* strain Rf and many *Salmonella* strains also encode type III systems (*Hin*fIII and *Sty*LTI respectively; Piekarowicz and Kalinowska, 1974; Bullas *et al.*, 1980). There is also some evidence for a system, or the remnants of one, in *Bacillus cereus* (Hegna *et al.*, 1992), and it is probable that many other examples of these systems have yet to be identified.

The type III systems are intermediate in complexity, between the type I and type II enzymes, and are encoded by two genes, *res* and *mod* (Iida *et al.*, 1983). MTase activity is encoded by the *mod* gene, which together with the *res* gene product can form an ENase. The DNA is cut 25 or 26 bp downstream from the recognition sequence, which are short (5-6 nt), and asymmetric (see Bickle, 1993). Surprisingly, only one strand of the DNA is modified, forming N^6meA. At first, it was difficult to see how hosts encoding these systems avoided auto-restriction after replication, until it became clear that a single unmodified target is not the substrate for a type III ENase (Meisel *et al.*, 1992).

The nature of the target for a type III ENase became apparent from an examination of the nucleotide sequence of the closely related phages, T3 and T7. Phage T3 is restricted by *Eco*P15I, whereas T7 is not. Phage T7 contains 36 *Eco*P15I target sites

in its genome (Dunn and Studier, 1983), however, they are all in the same orientation (Schroeder *et al.*, 1986). Phage T3 contains at least two inverted sites (Krüger *et al.*, 1990). This strongly suggested that it is the presence of two inversely oriented target sequences that elicit DNA cutting. This was further investigated, and shown to hold true, for M13 clones containing different numbers and orientations of *Eco*P15I sites (Meisel *et al.*, 1992). Thus, the *Eco*P15I recognition target for methylation is 5' CAGCAG whereas the target for restriction is two of these, in inverse orientation and separated by a variable distance i.e. 5' CAGCAG-N_x-CTGCTG. A host expressing a type III R-M system does not cut its own DNA after replication, as there is always one methylated site in an inverse-orientated target (figure 1.5; Meisel *et al.*, 1992).

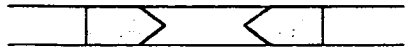
Sequence comparisons among members of the type III R-M systems have revealed similar identities to the type I systems. The predicted aa sequence of the *mod* gene product, responsible for DNA sequence specificity and methylation, contains a central variable region flanked by two conserved regions. By analogy to the type I systems, this variable region probably encodes the DNA target recognition domain (Hümbelin *et al.*, 1988). Motifs I and II, common to all adenine MTases (Klimisaukas *et al.*, 1989; Pósfai *et al.*, 1989), are found in the conserved regions (Dartois *et al.*, 1993). Alignment of the *res* gene products shows little similarity, except in a short central region which has near identity and may be involved in subunit interactions (Dartois *et al.*, 1993). Like the type I HsdR polypeptide, seven motifs characteristic of 'DEAD-box' proteins can be found in the predicted aa sequence of type III *res* gene products, indicating that they may both be ATP-dependent DNA helicases (Dartois *et al.*, 1993; Murray *et al.*, 1993). This is apparently contradicted by the observation that type III enzymes can cut DNA without hydrolysing ATP (Yuan *et al.*, 1980). Restriction by type III systems require communication between enzymes at two sites (Meisel *et al.*, 1992), which could be achieved by DNA translocation (Murray *et al.*, 1993). Type I R-M systems also translocate DNA before cutting occurs (Yuan *et al.*, 1980; Studier and Bandyopadhyay, 1988). DNA translocation by both type I and III systems may involve ATP-dependent helicase activity, which is supported by the recent reporting of ATP hydrolysis associated with restriction by a type III ENases (Saha and Rao, 1995; Meisel *et al.*, 1995).

Modification-dependent systems

The systems originally described in 1952 by Luria and Human are not classical R-M systems, as modified DNA is the target for the ENase and there is no cognate MTase. Luria and Human (1952) isolated T-even phages, T*, that were restricted by *E. coli* B.

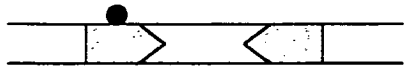
Figure 1.5. Targets for restriction and modification by the *Eco*P15I R-M system. (a) A single DNA target is modified, but two inversely-orientated unmodified sites are required for restriction. (b) Protection of host DNA after semi-conservative replication. See text for more details. Shaded arrow, *Eco*P15I recognition site (5' CAGCAG); solid circle, methylated base (Meisel *et al.*, 1992)

(a)



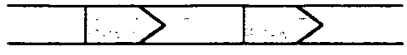
Methylation	Restriction
+	+

2 inversely oriented sites



Methylation	Restriction
-	-

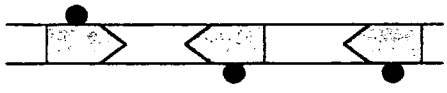
2 inversely oriented, hemimethylated sites



Methylation	Restriction
+	-

2 sites in the same orientation

(b)



Semiconservative DNA replication

An understanding of this change in host restriction activity, based on the modification of phage DNA, was made in the late 1960's (Revel, 1967; Revel and Luria, 1970). T-even phages normally contain glucosylated hydroxymethylcytosine (hmC) in place of cytosine in their DNA. The T* phages do not glucosylate hmC-containing DNA and the systems responsible for their restricted growth were termed RglA and RglB, for restricts glucose-less phages (Revel, 1967; Revel and Luria, 1970). These systems may be an evolutionary relic, as the normal glucosylation of phage DNA renders them totally resistant to restriction by RglA and RglB (Revel and Luria, 1970; Bickle and Krüger, 1993). However, interest in these systems was stimulated when it became apparent that they interfered in the cloning of certain MTases (Blumenthal *et al.*, 1985; Noyer-Weidner *et al.*, 1986; Raleigh and Wilson, 1986). Restriction by these systems was shown to be both modification-dependent and sequence-specific and they have been renamed McrA and McrB, RglA and RglB respectively, for modified cytosine restriction (Raleigh *et al.*, 1989). The region encoding McrB is now known to include two genes, *mcrB* and *mcrC* (Raleigh, 1992). A third modification-dependent system has been identified in *E. coli*, termed Mrr, for modified adenine recognition and restriction (Heitman and Model, 1987), although phages with cytosine-methylated DNA are also restricted (Kelleher and Raleigh, 1991). The *hsd*, *mrr*, and *mcrBC* genes are closely linked in both *E. coli*K-12 and *E. coli* B. This region at 99 minutes on the *E. coli* chromosome has been termed the immigration control region (Raleigh, 1992). The *mcrA* gene is part of e14, a prophage-like element on the *E. coli* K-12 chromosome (Hiom and Sedgwick, 1991).

Distribution of R-M systems

There is no efficient screen for all types of R-M systems. Methods based on the restriction of phages have been generally unsuccessful, because many natural isolates of bacteria are resistant to productive phage infection (Bullas *et al.*, 1980; Barcus, 1993). Other methods rely on similarities with known systems; type II R-M systems are detected by their ability to produce discrete DNA fragments, and type I systems have been identified using DNA probes derived from the *hsd* genes (Daniel *et al.*, 1988).

Over 2400 type II R-M systems have been identified, mainly due to their importance to molecular biologists, and efficient screens that have been developed to identify them (Roberts and Macelis, 1993). Sequence-specific DNA endonucleases are identified by incubating cell extracts with suitable DNA substrates and looking for discrete fragments separated by agarose gel electrophoresis. This activity is normally

interpreted as due to the ENase of a type II R-M system. Over 10 000 bacterial strains have been examined in this way, and approximately a quarter are thought to encode type II R-M systems (Wilson and Murray, 1991; Roberts and Halford, 1993). This figure is supported by a systematic screen of the Enterobacteriacea for type II systems, which identified sequence-specific endonucleases in 25 % of over 1000 strains tested (Janulaitis *et al.*, 1988).

R-M systems which do not cut DNA at specific points or have other requirements, like the type I and type III R-M systems, are not detected in these screens. Only four type III R-M systems are known, and these were identified by the restriction of phages (Bickle, 1993). Type I R-M systems were originally detected by phage restriction, but have recently been identified by DNA hybridisation to known *hsd* genes (Daniel *et al.*, 1988; Barcus and Murray, 1995; Barcus *et al.*, 1995). A collection representing the genotypic diversity of *E. coli* strains, the ECOR collection (Ochman and Selander, 1984), has been used to screen for type I R-M systems (Barcus *et al.*, 1995). Probes from known *hsd* genes, of type IA, IB and ID R-M systems, hybridised to the DNA of 17 strains out of 37 tested in the collection. The distribution of these systems, using a dendrogram of the ECOR collection, indicates the horizontal transfer of the *hsd* genes (Barcus *et al.*, 1995). Some of the R-M genes identified in the ECOR collection have been cloned, and five new DNA specificities identified (table 1.1; Barcus *et al.*, 1995). This brings the total of type I specificities to at least 16. The type III systems encode four different specificities, and the type II encode over 180 (Roberts and Macelis, 1993). The selective pressure imposed on a bacterial population for variation in the specificity of R-M systems, may be the challenge by phages (Levin, 1986).

Members of the ECOR collection that did not hybridise to probes of known *hsd* genes may encode other R-M systems, either members of different type I families or other types of R-M system, and the screening of strains for type II systems does not identify other types of R-M system. It is interesting to note that only one natural isolate, *E. coli* C, has been extensively studied and found not to encode an R-M system (Raleigh, 1992). However, *E. coli* C can readily acquire a system by transduction or transformation, and it seems likely that the majority of bacteria do possess an R-M system.

Regulation of R-M activity

It is essential for a bacterium to regulate R-M activity under certain conditions, for example when an R-M system is transferred between hosts the DNA of the recipient must be modified before an active ENase is produced. Many R-M systems have

separate promoters for the genes encoding the MTase and ENase (Wilson and Murray, 1991), so R-M activity could be modulated by differential gene expression. The control of gene expression may involve passive or active mechanisms. The promoter for the MTase of *Pst*I overlaps the weaker promoter for the ENase, this would allow time for the MTase to accumulate before an active ENase is formed (Walder *et al.*, 1984). Some of the genes encoding type II systems are transcriptionally regulated by repressor-like C proteins of about 100 aa with a characteristic helix-turn-helix DNA-binding motif. The first *C* gene identified was linked with the genes encoding the *Bam*HI R-M system (Nathan and Brooks, 1988), and there is now a small family of proteins that regulate the systems of *Bam*HI, *Pvu*II, *Sma*I, *Eco*RV, *Eco*72I and *Mun*I (Tao Tao *et al.*, 1991; Siksnys *et al.*, 1994). Little is known about the molecular mechanism of C protein activity, although C proteins are probably site-specific DNA-binding proteins that activate or repress expression of the *res* and *mod* genes of the R-M systems with which they are associated. C proteins are much more similar to each other, and to some other gene regulators e.g. λ repressor, than the R-M systems with which they are associated, implying that the *C* genes have been independently recruited into different R-M systems (Tao Tao *et al.*, 1991).

Type I and III R-M systems also have separate promoters for the genes encoding the MTase (P_{mod}) and the additional subunit that is required for restriction (P_{res}; Loenen *et al.*, 1987; Sharrocks and Hornby, 1991). Promoter fusion studies, with the *lacZ* operon, have found no evidence of differential gene expression with P_{res} and P_{mod} of the type I R-M system *Eco*KI (Loenen *et al.*, 1987; Prakash-Cheng *et al.*, 1991) or the type III system *Eco*PI (Sharrocks and Hornby, 1991). Experiments examining the restriction phenotypes of *E. coli* after conjugation, reveal that there is a delay in restriction activity when the *hsd* genes of *Eco*KI are transferred (Prakash-Cheng and Ryu, 1993). It is not possible, from the results of Prakash-Cheng and Ryu (1993), to conclude whether modification activity is observed immediately upon transfer, but it seems probable. Promoter fusion studies show that both P_{res} and P_{mod} are transcribed immediately upon transfer, and yet there is a delay before restriction is detected. A mutant has been isolated, termed *hsdC*, that can't act as a recipient of the *Eco*KI *hsd* genes in conjugation, presumably because restriction activity is not regulated, leading to auto-restriction (Prakash-Cheng *et al.*, 1993). Further analysis of this mutant should identify a mechanism of controlling the R-M activity of type I systems.

Restriction by *Eco*KI is reduced after cells have been irradiated with ultraviolet (uv) light (Day, 1977) or treatment with chemical agents that cause DNA damage (Efimova *et al.*, 1988a). Kelleher and Raleigh (1994) examined the effect of uv-induced DNA

damage on the restriction activity of three R-M systems in *E. coli* K-12, the type I system *Eco*KI and the modification-dependent systems McrA and McrBC. Restriction by McrA was alleviated by exposure to uv light, but this was because the genetic element that encodes it, e14, is rapidly lost (Hiom *et al.*, 1991a; Kelleher and Raleigh, 1994). The excision of e14 is probably analogous to parts of the lytic pathway of phages, where the phage repressor is cleaved by activated RecA (Roberts and Roberts, 1975). Restriction by McrBC and *Eco*KI was also alleviated in these experiments; alleviation of McrBC restriction is dependent on the SOS response, whereas alleviation of *Eco*KI is dependent on actual DNA damage as restriction was not alleviated in strains that expressed SOS-regulated genes constitutively (Kelleher and Raleigh, 1994). The mechanism of these activities, therefore, is likely to be different. Alleviation of McrBC restriction may be by the direct interaction with an SOS-induced protein. The ENase of *Eco*KI is not turned-over after cutting DNA, but remains bound to the target site. Restriction by *Eco*KI may be alleviated, therefore, by titration of the enzyme with small DNA fragments containing unmodified target sites that are produced as a by-product of the *E. coli* DNA repair machinery (Kelleher and Raleigh, 1994). This may also explain the alleviation of restriction in *E. coli* K-12 *dam* mutants (Efimova *et al.*, 1988b); *Eco*KI could be titrated by the excess of unmodified targets generated by uncoordinated DNA replication (Boye and Lϕbner-Olesen, 1990; Kelleher and Raleigh, 1994). It is not clear why restriction need be alleviated in response to DNA damage or whether, in the case of *Eco*KI at least, it is merely a saturation of the restriction ability of the cell.

Phage antirestriction mechanisms

Many phages have methods of avoiding R-M systems (see Krüger and Bickle, 1983; Bickle and Krüger, 1993), however, like the host-encoded systems, there has been little work performed on the distribution of phage antirestriction systems in nature. One study has examined the restriction of a collection of phages, by R-M systems that may be expected in the phage host-range (Korona *et al.*, 1993). Twenty four naturally occurring coliphage were tested for restriction by three type I and four type II R-M systems *in vivo*; 19 were restricted by at least one R-M system. Twenty one of the 24 phages could also be tested for cutting by the four type II ENases *in vitro*; 16 of the phages that were not restricted *in vivo* were also not cut by the ENase *in vitro*. This may indicate that the targets for these R-M systems are not present in the phage genomes (Korona *et al.*, 1993) or alternatively, that the phage DNA is modified so that the targets are not recognised as a substrate for the ENase. The remaining five phages that are not restricted *in vivo*, but are cut by the type II ENase *in vitro*, probably

encode antirestriction systems that interfere with the host's R-M system. In the well-studied laboratory phages, most antirestriction systems are directed against type I R-M systems, although phage T5 does encode a function that makes it resistant to *Eco*RI, a type II R-M system (Eskridge *et al.*, 1967). Unlike most other antirestriction systems, the antirestriction gene of T5 appears to encode an essential function as the only mutants that can be isolated sensitive to *Eco*RI are those that have additional target sites in the region of phage DNA first introduced into the cell during infection (Davison and Brunel, 1979a, 1979b; Brunel and Davison, 1979).

A number of statistical analyses have been performed on the nucleotide sequence of phages. An examination of the nucleotide sequence of a total of 12 DNA phages, T7, λ, ϕX174, G4, M13, f1, fd, IKe, ϕ29, P1, PZA and T4, shows that six-base palindromes are significantly underrepresented in all but phage T4 (Sharp, 1986; Karlin *et al.*, 1992). These sequences are not depleted in the RNA phage MS2 or viruses that infect eukaryotes, e.g. EBV, HSV, CMV or SV40, and strongly suggests a counter-selection for six-base palindromes in DNA phages. Fifty-five of the 64 possible six-base palindromes are known to be targets for the type II R-M systems (Roberts and Macelis, 1993), and it seems probable that all possibilities do exist in nature. It is possible that a phage may encounter a type II system with any of 64 six-base specificities, although there is some evidence to suggest that the range of DNA targets recognised by type II R-M system enzymes is limited within a species (Janulaitis *et al.*, 1988). The underrepresentation of six-base palindromes in the genomes of DNA phages probably results from the selection against phages containing targets for host-encoded type II R-M systems (Sharp, 1986; Karlin *et al.*, 1992). Phage T4 is not underrepresented in these sequences, but it is completely protected from restriction by most systems as it contains glucosylated hydroxymethylcytosine (hmC) in its DNA (Wyatt and Cohen, 1952).

Modification of phage DNA

Like phage T4, a number of other phages are protected from restriction by the modification of their DNA. This modification may be methylation, as with the *Bacillus* phage ϕ3T (Noyer-Weidner *et al.*, 1981), acetimidation, by phage Mu (Swinton *et al.*, 1983) or the glucosylation of hmC, as with phage T4 (Wyatt and Cohen, 1952).

DNA methylation is an obvious mechanism to avoid restriction and a wide range of phages encode MTases, including the *Bacillus* phages ϕ3T and SPR (Noyer-Weidner *et al.*, 1981), the *Shigella* phage DDVI (Nikolskaya *et al.*, 1979), phages T2 and T4 (Hattman, 1970), and the *Lactococcus* phage ϕ50 (Hill *et al.*, 1991). The MTases of

the *Bacillus* phages are similar to the MTases of type II R-M systems, but often contain more than one TRD and recognise more than one DNA sequence (Lange *et al.*, 1991). The host and phage-encoded MTases of *Bacillus* appear to be unrelated, but the MTase of the *Lactococcus* phage ϕ50 is clearly derived from the host. The *Lactococcus* self-transmissible plasmid pTR2030 encodes a type II R-M system, *Lla*I, and shares a 1.2 kb region of sequence identity with ϕ50 that encodes a functional MTase (Hill *et al.*, 1991).

The DNA of phages Mu and D108 contain N^6 acetamidoadenine instead of adenine, and the enzyme responsible is encoded by the *mom* gene (Swinton *et al.*, 1983). The *mom* gene product is a sequence-specific modifying enzyme and phages are protected from restriction by those systems whose target overlaps this sequence (Kahmann, 1984). The *mom* gene of phage Mu has been sequenced, and encodes a polypeptide of predicted molecular mass 28.3 kDa (Kahmann, 1984). Constitutive expression of *mom* is lethal to the host, and its expression is tightly regulated (Kahmann *et al.*, 1985). Transcription of the *mom* gene is dependent on the methylation of sites upstream from the promoter, by host-encoded Dam MTase (Hattmann, 1982), and the phage-encoded *com* gene product is required for translational initiation (Wulczyn *et al.*, 1989).

The presence of glucosylated hmC in the DNA of the T-even phages, and the *Shigella* phage DDVI, renders them totally resistant to most R-M systems (Wyatt and Cohen, 1952). This modification also serves to differentiate between phage and host DNA, as phage-encoded nucleases rapidly digest host DNA upon infection by T-even phages (Warner and Snustad, 1983). The DNA of a T-even phage is replicated containing hmC, instead of cytosine, and then glucosylated at this position postsynthetically (Revel, 1983). Phages may be isolated that are not glucosylated, but these are still resistant to most R-M systems due to the presence of hmC. The modification-dependent systems, McrA, McrBC and Mrr, are able to restrict the growth of these phages, although the T-even phages do encode a gene, *am*, whose product disrupts the activity of McrBC (Dharmalingam and Goldberg, 1976; Dharmalingam *et al.*, 1982). Arn is not produced sufficiently quickly to protect its own DNA, but it is able to protect superinfecting phages from restriction by McrBC. T-even phages that contain cytosine instead of the modified base, hmC, are restricted by classical R-M systems. T-even phages, therefore, appear to have a number of redundant antirestriction systems as glucosylation of the DNA renders the phage totally resistant to restriction by most R-M systems. This may indicate a close evolutionary link between the phages and

their host, and *am* and hmC may be part of an evolutionary legacy (Krüger and Bickle, 1983).

The region of DNA that contains the *hsd* genes of the type I R-M system *Eco*prrI also encodes a phage exclusion gene, *prrC* (Levitz *et al.*, 1990; Tyndall *et al.*, 1994). PrrC is activated by a T4-encoded peptide, Stp, and cleaves the host $tRNA^{Lys}$, blocking translation and phage development (Amitsur *et al.*, 1987, 1992). The only known function of T4-encoded polynucleotide kinase and RNA ligase is to repair this damage and allow phage development to continue (Amistur *et al.*, 1987). The PrrC protein and *Eco*prrI are closely associated, as they can be co-precipitated using specific antibodies, and results indicate that the *Eco*prrI complex masks the activity of PrrC (Amitsur *et al.*, 1992; Morad *et al.*, 1993). The T4 Stp peptide may be an antirestriction mechanism that dissociates the ENase complex and in so doing releases, and activates, PrrC (Snyder, 1995). The PrrC exclusion system may be an RNA-based phage restriction system, and T4 polynucleotide kinase and RNA ligase a phage antirestriction mechanism (Bickle and Krüger, 1993).

Inactivation of R-M systems

The *Bacillus* phages ϕ1rH and ϕNR2rH are not restricted by the type II R-M system *Bam*NxI (Makino *et al.*, 1979; 1980). DNA prepared from phage ϕNR2rH grown in a strain expressing *Bam*NxI, is still cut by the ENase *in vitro*, suggesting that both enzymatic activities of the R-M system are inhibited (Makino *et al.*, 1979). A polypeptide was isolated from cells infected with ϕ1rH, that specifically inhibited the *Bam*NxI ENase *in vitro*, and this inhibitory protein probably forms a complex with the ENase that blocks its activity (Makino *et al.*, 1980).

Phages T3 and T7 encode a more general antirestriction system that blocks both restriction and modification by type I R-M systems (Studier, 1975; Studier and Movva, 1976; Spoerel *et al.*, 1979; Krüger and Bickle, 1983). The T3 system is also active against type III R-M systems (Krüger *et al.*, 1982; Bickle and Krüger, 1993), but both phages T3 and T7 are efficiently restricted by type II systems (Bandyopadhyay *et al.*, 1985). The phage gene encoding this function is termed *ocr* or gene *0.3* (Studier, 1975; Krüger *et al.*, 1977), and the molecular mechanism of T7 Ocr activity has been examined in detail; Ocr binds tightly to free ENase, either to the HsdM or HsdS subunit, and prevents the complex from binding DNA. This explains why both enzymatic activities are inhibited, and the phage is neither restricted or modified. If an ENase:DNA complex is formed before Ocr is added, then nuclease activity is still inhibited, suggesting that the binding site for Ocr is accessible after

DNA binding, but before cutting occurs. The MTase activity is strongly inhibited by Ocr added at any stage of the reaction (Bandyopadhyay *et al.*, 1985). The T3 *ocr* gene product is also able to hydrolyse AdoMet, and mutants can be isolated that map to *ocr* and still retain the Ocr phenotype, but have lost the ability to hydrolyse AdoMet (Krüger *et al.*, 1977). AdoMet stimulates the ENase activity of type III R-M systems (Haberman, 1974), so the hydrolysis of AdoMet by the phage T3 *ocr* gene product explains why these phages are resistant to restriction by type III systems.

The *ocr* gene is in a region of the phage genome that is introduced into the cell at the start of injection of the phage DNA (Pao and Speyer, 1973). Moffatt and Studier (1988) followed the fate of ^{32}P-labelled DNA of infecting T7 phages in hosts with different R-M systems, and concluded that the phage DNA was not susceptible to degradation until six or seven minutes after infection. The *ocr* gene product should be made just three to four minutes after infection (Studier, 1972), and can be visualised at about this time in experiments following the synthesis of phage proteins in cells infected with T7 (Moffatt and Studier, 1988). The first *Eco*KI target site is in a region of the phage DNA that is introduced into the cell after *ocr*, but T7 phages containing a site before *ocr* are still protected from restriction by *Eco*KI, although it is not clear if a single target at this position would be a substrate for the ENase. However, the evidence does indicate that transcription and translation of the early region of the T7 genome occurs before the DNA can be degraded by host-encoded R-M systems, and the production of Ocr during this time would protect the phage DNA from cutting by type I ENases (Moffatt and Studier, 1988).

The absence of potential target sites for type II R-M systems (Sharp, 1986; Karlin *et al.*, 1992), the production of the Ocr function that blocks the activity of type I enzymes, and the underrepresentation of sites for the *E. coli* MTases Dam and Dcm (Schroeder *et al.*, 1986), have lead Krüger and colleagues (1989) to suggest that these may be methods for phage T7 to avoid methylation. Apart from R-M systems, DNA methylation in prokaryotes has been linked to DNA replication and repair, regulation of gene expression and packaging of phage DNA (see Marius, 1987), but any relevance of undermethylation to the biology of phage T7 is unknown.

Modulation of R-M activity

Phage P1 encodes both a type III R-M system, *Eco*PI, and an antirestriction system that is encoded by the genes *darA* and *darB* (Iida *et al.*, 1987; Streiff *et al.*, 1987). The Dar proteins are associated with the DNA when injected into the cell and block restriction by type I R-M systems. The *darA* gene product is sufficient to protect

against restriction by type IB family of R-M systems, but both the *darA* and *darB* gene products are required to protect against restriction by the type IA systems. Furthermore, any DNA packaged in a P1 phage head is protected from restriction, which allows efficient generalised transduction between hosts with different type I R-M systems (Iida *et al.*, 1987). The DarA and DarB proteins are associated in the P1 head and DarB does not seem to be present in the phages of a P1*darA* mutant, suggesting that DarA is required for the correct incorporation of DarB into the phage head. The Dar proteins probably bind DNA in the phage head and remain bound to the DNA after injection. Dar may alleviate restriction by blocking the DNA translocation required for cutting by the type I ENases (Iida *et al.*, 1987). The Dar proteins also stimulate methylation of unmodified DNA by the MTases of type IA R-M systems (Iida *et al.*, 1987). These enzymes are the only prokaryotic MTases shown to have a strong preference for hemimethylated DNA. DarA may be responsible for blocking restriction, and DarB for modification enhancement, but this can't be tested using P1*darA* mutants as these phages do not incorporate DarB into the phage heads. If DarA remains bound to the DNA and blocks restriction, the DNA would only be poorly methylated by the MTase of a type IA R-M system, and when DarA-mediated protection is lost, the DNA would be cut by the ENase. However *in vivo*, the MTases of other type I R-M systems do not have a strong preference for hemimethylated DNA and would methylate unmodified DNA protected from restriction by DarA. The DNA would not be cut, therefore, when DarA protection was lost and this may explain why P1*darB* mutants are only resistant to restriction by non-type IA R-M systems.

The lambdoid phages λ, λ*reverse* (λ*rev*), ϕ21 and P22 encode a very specialised antirestriction system, encoded by the genes *ral* (λ, ϕ21 and P22; Zabeau *et al.*, 1980; Franklin, 1984; Semerjian *et al.*, 1989), or *lar* (λ*rev*; Toothman, 1981). Most work has been performed on the *ral* gene of phage λ. Zabeau and colleagues (1980) discovered that unmodified λ phages can escape restriction upon coinfection with modified λ helper phages, and mapped the gene responsible between *N* and *c*III. When this region of DNA was sequenced, a short open reading frame (*orf*) of 201 bp was assigned to *ral* (Ineichen *et al.*, 1981; Sanger *et al.*, 1982). A *ral* mutant, *ral*18 (Debrouwere *et al.*, 1980a), was sequenced and shown to have two mutations in this *orf* (Loenen and Murray, 1986). Both λ*ral*$^+$ and λ*ral*$^-$ phages are restricted efficiently by *Eco*KI, however, after a single round of infection of a restriction-deficient, modification-proficient host, the progeny of the *ral*$^+$ phages are fully modified, whereas those of the *ral*$^-$ phages are only poorly modified (Zabeau *et al.*, 1980). This effect may be seen even after multiple rounds of infection (Kelleher *et al.*, 1991). Ral

is not itself a MTase, as the λ phage only carries the modification pattern of the host in which it was last propagated (Toothman, 1981).

The *ral* gene has been cloned in an expression plasmid, and the gene product visualised using both an *in vitro* transcription and translation system, and minicells transformed with the plasmid (Loenen and Murray, 1986). The only R-M systems in which restriction is alleviated, and modification enhanced by Ral, are the type IA. These are also the only systems whose MTase has a strong preference for hemimethylated DNA *in vivo*. Recombinant λ phage that over-express the genes for the *Eco*KI MTase, do not efficiently methylate their DNA unless they also contain *ral* (Loenen and Murray, 1986). An examination of the polypeptides produced by two phages, *ral*$^+$ or *ral*$^-$, encoding the *Eco*KI MTase, concluded that Ral altered the ratio of HsdM and HsdS subunits produced, and that this explained the increase in *de novo* methylation activity with Ral (Zinkevich *et al.*, 1986). However, the phages in these experiments are very different. Both phages contain the same region of DNA, encoding the MTase from *Eco*KI, but only in the *ral*$^+$ phages are the *hsd* genes transcribed from the strong λ promoter, P_L. This fragment of DNA encodes most of *hsd*R, all of *hsd*M, *hsd*S, *mcr*BC and other downstream genes. McrB and HsdS comigrate by SDS PAGE (Sain and Murray, 1980) and these two products may have been confused in this work as Zinkevich and colleagues (1986) appear to have made no attempt to confirm their identification of the polypeptides produced. A favoured model to explain Ral activity is that Ral interacts directly with either HsdM or HsdS and changes the kinetics of methylation so that unmodified DNA is efficiently methylated (Loenen and Murray, 1986).

The *ral* gene appears to have no role in the life cycle of phage λ, as it is dispensable for normal phage development (Zabeau *et al.*, 1980; Debrouwere *et al.*, 1980a, 1980b). Debrouwere and colleagues (1980b) observed that *E. coli* strains in which *ral* was constitutively expressed were phenotypically similar to strains carrying a mutation in *rho* that encodes a transcriptional termination factor. They postulated that *ral* might have an antitermination function which modulated the activity of *Eco*KI. However, no evidence of antitermination activity was observed with cloned *ral* (Loenen and Murray, 1986) and the activity observed by Debrouwere and colleagues (1980b) may be due to *Ea*10; a single-strand binding protein of unknown function that is also expressed by the defective prophage used in this work.

The *ral* gene of φ21 has been identified by the nucleotide sequence analysis of a recombinant λ phage (Franklin, 1984). The λ and φ21 *ral* genes are identical, and it may be that the gene identified as φ21 *ral* is derived from λ sequence in the

recombinant phage that was sequenced. The *ral* gene of the *Salmonella* phage P22 has also been sequenced, and shares 75% identity with λ *ral* (Semerjian *et al.*, 1989). The *ral* gene from phage P22 has been cloned in an expression plasmid, and used to transform both *E. coli* K-12 and *S. enterica* serovar *typhimurium* LT2. Restriction by the *E. coli* strain is alleviated, demonstrating that the P22 *ral* gene encodes a functional protein, but restriction by the *Salmonella* strain is unaffected (Semerjian *et al.*, 1989). *S. enterica* serovar *typhimurium* LT2 encodes three R-M systems, termed *Sty*LTI, *Sty*LTII and *Sty*LTIII; only *Sty*LTIII is a type IA R-M system (Barcus and Murray, 1995). Therefore, although the P22 *ral* gene product is probably effective against *Sty*LTIII, the two other R-M systems would still be active and restrict the test phage.

The Ral-like antirestriction function of λ*rev* is encoded by the *lar* gene (Toothman, 1981). λ*rev* arises by a recombination event between phage λ and the excised Rac prophage (Kaiser and Murray, 1979). The Rac prophage is a defective lambdoid prophage in the chromosome of many *E. coli* strains, and encodes genes that can serve the same functions as the early genes of phage λ, including integration, recombination, immunity, and replication (Low, 1973; Gottesman *et al.*, 1974; Diaz *et al.*, 1979). *E. coli sbcA* mutants express the *recE* and *recT* genes, and λ*rev* relies on these genes to provide the general recombination functions normally encoded by λ *exo* and *bet* (Kushner *et al.*, 1974; Gottesman *et al.*, 1974; Clark *et al.*, 1993). *E. coli sbcA* mutants are also phenotypically Lar^+ (Simmon and Lederberg, 1972), and the antirestriction activity of these mutants is functionally separable from the recombination functions encoded by *recET* (Toothman, 1981).

Possible roles of R-M systems

Recently, it has been proposed that the *res* and *mod* genes of the type II R-M systems may be a selfish symbiotic unit, that ensure the maintenance of the R-M system that they encode (Naito *et al.*, 1995). Naito and colleagues (1995) examined the stability of plasmid-encoded type II R-M systems in *E. coli*, and found that plasmids encoding both the *res* and *mod* genes were maintained better than vector plasmid alone or those encoding just *mod*. They postulate that once a bacterium express an R-M system, loss of the genes encoding that system is fatal. Once the genes have been lost, R-M activity would continue with residual ENase and MTase present in the cell. However, the concentration of these enzymes would decrease through successive generations until the MTase activity is not sufficient to ensure that the host DNA is completely modified, and the host DNA then becomes a substrate for the ENase, leading to cell inviability. This hypothesis cannot be extended easily to R-M systems that have common subunits

in the ENase and MTase, like the type I and III systems, but may account, at least in part, for the maintenance of those R-M systems that have separate ENases and MTases.

Traditionally, R-M systems have been described as primitive immune systems that protect against phage infection, although evidence is now accumulating that R-M systems influence the flux of genetic material between bacterial populations (for recent reviews see Arber, 1994; King and Murray, 1994; Barcus and Murray, 1995). R-M systems are certainly a barrier to phage infection, however studies with populations of bacteria and phages show that R-M systems only confer a transient advantage to the host. In the laboratory, a bacterial population that is challenged by phages quickly acquire mutations that make the host totally resistant to infection (Levin, 1988; Korona and Levin, 1993). Phage resistance was also acquired by the host in experiments where bacteria colonised an environment in which phages were already present, but a host-encoded R-M system increased the likelihood of the bacterial population becoming established (Korona and Levin, 1993). This colonisation selection may be important in the maintenance of R-M systems, but it also seems relevant that the maintenance of an R-M system affords protection to a range of phages simultaneously, unlike mutations conferring phage resistance. Mutations may also be disadvantageous as the products of the genes, changed to acquire phage resistance, are commonly involved in other processes in the cell and their disruption may compromise the fitness of the bacterium in nature. However, these mutations are rapidly acquired in laboratory studies (Korona and Levin, 1993). R-M systems are not a complete barrier to phage infection, and those phages that do escape restriction are likely to become modified. An R-M system, therefore, can only confer a transient advantage to the host as the phage becomes rapidly modified, although bacteria encode R-M systems with many different specificities. This diversity of recognition sequences may be explained by the frequency-dependent selection of rare specificities by phages (Levin, 1988). There are still a number of points which are perhaps inconsistent with R-M systems having evolved solely as a barrier to phage infection (Levin and Lenski, 1985).

1. R-M systems are a leaky barrier to phage infection; the eop of phages vary, depending on the particular phage and R-M system challenged, but it is typically reduced to between 10^{-1} and 10^{-5} by a single system.
2. Many phages encode anti-restriction mechanisms. However, one study does suggest that a sufficient proportion of naturally occurring coliphages are sensitive to R-M systems to account for the maintenance of these systems by phage-mediated selection (Korona *et al.*, 1993).

R-M systems are a barrier to the transfer of any DNA molecule, however, the consequence of DNA cutting may be different for extra-chromosomal elements and chromosomal DNA. The cutting of a phage genome leads to the restriction of phage growth, and similarly, a plasmid that is cut is not maintained. It is interesting to note that a number of self-transmissible plasmids also encode an anti-restriction function, termed Ard, that is functionally similar to the T7 antirestriction function Ocr, and the two products share regions of sequence identity (see Wilkins, 1995; Belogurov *et al.*, 1995). The fate of bacterial DNA is likely to be different, however, as fragments may be rescued by homologous recombination with the host chromosome.

Linear fragments of DNA, generated by R-M systems, are rapidly degraded in *E. coli* by the RecBCD nuclease (Simmon and Lederberg, 1972), the same enzyme that is required for the repair of DNA double-strand breaks and the promotion of homologous recombination (see Smith, 1988). Enzymes analogous to RecBCD are found in many bacterial species, and an eight nt asymmetric DNA sequence termed χ, 5' GCTGGTGG, has been identified as a hotspot for RecBCD-promoted recombination (Smith *et al.*, 1981). RecBCD is a multifunctional enzyme that has a number of enzymatic activities; a single-strand and double-strand DNA exonuclease, a single-strand DNA endonuclease, a DNA helicase and a site-specific nuclease that nicks DNA when approaching a χ site from the 3' direction. A model for RecBCD-promoted homologous recombination is that site-specific nicking of DNA at the χ site and the subsequent unwinding of DNA produces a single-stranded end that may be used to initiate RecA-dependent strand exchange (Smith, 1988). The ss DNA end that is produced would normally be degraded by RecBCD, but this activity is lost after association with the χ site (Dixon and Kowalczkowski, 1993; Kuzminov *et al.*, 1994) possibly by dissociation of the RecD subunit (Dixon *et al.*, 1994). The effect of this is to increase the survival of linear DNA that contains a χ site (Kuzminov *et al.*, 1994). Thus, it is RecBCD and its interaction with χ that determines the fate of linear DNA in the cell, and it seems relevant that χ sites occur more frequently than expected, and have an orientation bias, in the *E. coli* genome (Burland *et al.*, 1993).

The ends generated by ENases can be used to initiate recombination, although most experiments have been performed in the absence of RecBCD (for example, Eddy and Gold, 1992). One experiment, in the presence of RecBCD, demonstrates that a fragment generated by cutting with the type II ENase, *Eco*RI, can be used to initiate recombination *in vivo* (Stahl *et al.*, 1983). The overall level of recombination, however, is reduced in the presence of an R-M system. Experiments with conjugal

plasmids show that restriction, by type I and III systems, both reduce linkage and the acquisition of early genetic markers (Boyer, 1964; Pittard, 1964; Arber and Morse; 1965). Similarly, the frequency of transduction mediated by phage P1*dar*⁻ is severely reduced by the presence of a type I R-M system (Iida *et al.*, 1987). These observations seem inconsistent with the suggestion that the ends generated by ENases are recombinogenic (Endlich and Linn, 1985; Price and Bickle, 1986). Preliminary experiments, however, using the PCR to examine restriction fragment length polymorphisms, indicate that R-M systems may influence the transfer of DNA by P1-mediated transduction (McKane and Milkman, 1995). McKane and Milkman (1995) examined the recombination patterns of a number of transductants, centred around a single allele, and showed that transduction between different strains of the ECOR collection (Ochman and Selander, 1984) could result in the replacement of multiple fragments much shorter than the entrant molecule. When the DNA from two of these transductants was transduced back into the parent, the substitutions were much longer, and no multiple replacements were observed suggesting an effect due to restriction by the host (McKane and Milkman, 1995). Further experiments, with hosts of known R-M and recombination genotypes and P1 phages of known *dar* (anti-restriction) genotype, should help clarify the situation. The present results are consistent with the finding that the chromosomes of naturally occurring *E. coli* differ in short regions of DNA (Milkman and McKane-Bridges, 1993), and indicate a more general role for R-M systems in the flux of genetic material between populations.

Chapter two: Materials and methods

Bacterial strains

Bacterial strains are described in table 2.1.

Phages

Phages are described in table 2.2.

Plasmids

Plasmids are described in table 2.3.

Synthetic oligonucleotides

Synthetic oligonucleotides are described in table 2.4.

Enzymes, isotopes, and chemicals

Klenow DNA polymerase, type II ENases and T4 DNA ligase were purchased from Boehringer Mannheim Ltd.; $Vent_R$ DNA polymerase was from New England Biolabs Inc.; Lysozyme was from Sigma-Aldrich Company Ltd.; Sequenase was from Amersham International plc.; HK Phosphatase was purchased from Epicentre Technologies and Taq DNA polymerase and polynucleotide kinase (PNK) were kindly donated by D. Clark and S. Bruce, respectively.

Stabilised aqueous solutions of [^{35}S]dATPαS (400 Ci mmol^{-1}), [γ–^{32}P]dATP (3000 Ci mmol^{-1}) and S-adenosyl-L-[methyl-^{3}H]methionine ([methyl-^{3}H]AdoMet, ~75 Ci mmol^{-1}) solubilised in hydrochloric acid, were purchased from Amersham International plc., and aqueous L-[^{35}S]cysteine (600 Ci mmol^{-1}) was from DuPont Ltd. Amplify and Hyperfilm MP, purchased from Amersham International plc., were used for fluorography; ^{35}S was detected using Biomax film from IBI-Kodak Ltd., and Cronex film from DuPont Ltd. was used for other applications.

Antibiotics (chloramphenicol, tetracycline and rifampicin) were purchased from Sigma-Aldrich Company Ltd., and ampicillin (Penbritin) was from Beecham Research plc.

Dialysis tubing was purchased from Spectrum Medical Instruments Inc. or MediCell International Ltd.; 0.2 μm syringe filters were from Whatman Scientific Ltd.; 0.45 μm (HV) filters, used for the filtration of organic solvents, and Immobilon-P polyvinylidene difluoride (PVDF) membrane, used in Western blotting, were from

Strains[§]	Relevant features	Source
E. coli K-12 derivatives		
C600	$r_K^+m_K^+$ *supE*	Appleyard, 1954
5K	C600*hsdR*	Hubacek and Glover, 1970
NM710	C600*srl*::Tn*10*Δ(*rac*)	King and Murray, 1995
NM792	NM710*hsdR*Δ4	This thesis
AB1157	$r_K^+m_K^+$ Δ(*rac*)	Low, 1973
JC8679	*sbcA*23 in AB1157	Gillen *et al.*, 1977
NM519	JC8679*hsdR*Δ4	Lab collection
W3110	$r_K^+m_K^+$ *sup*°	C. Yanofsky
W3110*trpC*::Tn*10*	*trpC*::Tn*10*	C. Yanofsky
NM679	W3110Δ(*mrr-hsd-mcr*)	King and Murray, 1995
NM777[†]	W3110*zcj*3061::Tn*10*(*lar*$^-$)	King and Murray, 1995
NM778	W3110*zcj*3061::Tn*10*(*lar*$^+$)	Lab collection
NM794[†]	NM777*hsdR*Δ4	This thesis
Ymel[†]	$r_K^+m_K^+$ *supF*	Lab collection
NM526[†]	Ymel*hsdR*Δ4	King and Murray, 1995
NM787	Ymel*trpC*::Tn*10*	This thesis
E. coli B derivatives		
BL21[†]	$r_B^-m_B^-$ *sup*°*hsdS*	Studier and Moffat, 1986
NM793[†]	BL21*supF*	This thesis

Table 2.1. Bacterial strains. §, NM refers to stock numbers in N.E. Murrays laboratory; †, lysogenic for λDE3, where indicated.

Phages[§]	Relevant features	Source
P1*kc*		Lennox, 1955
DE3	λ*int*::*lacI*lac*UV5*-T7gene*1 imm*^{21}	Studier and Moffat, 1986
NM22	λ*Nam7Nam53c*I26	Lab collection
NM63	λ*c*I26	Lab collection
NM75	λ*b2red3imm*$^{\lambda}$	Lab collection
NM123	λ*c*I857*Sam7*	Lab collection
NM144	λ*h*^{82}*imm*$^{\lambda}$*c*I	Lab collection
NM220	λ*h*^{80}*imm*$^{\lambda}$*c*I	Lab collection
NM243	λ*vir*	Lab collection
NM324	λ*bio1imm*^{434}*c*I	Lab collection
NM325	λ*bio232*Δ(*int-ral*)*imm*^{434}*c*I	Lab collection
NM435	λ*Pam3c*I26	Lab collection
NM507	λ*b2imm*^{21}*c*I	Lab collection
NM508	λ*h*^{80}*imm*^{21}*c*I	Lab collection
NM891	λ*recE*	Kaiser and Murray, 1979
NM1090	λ*ral*18*c*I60	Lab collection
NM1094	λ*c*I60	Lab collection
NM1118	λ*hsdR*Δ4*c*I857Δ(*att*)	Lab collection

Table 2.2. Phages. §, NM refers to stock numbers in N.E. Murrays laboratory.

Plasmid	Relevent features	Source
pUC18	ColE1 replicon, carrying *bla*, P_{lac}, and multiple cloning sites (mcs).	Yanisch-Perron *et al.*, 1985
pUC19	As pUC18, but mcs in opposite orientation.	Yanisch-Perron *et al.*, 1985
pT7-6	ColE1 replicon, carrying P_{T7}, *bla* and mcs.	Tabor, 1990
pT7-7	Similar to pT7-6, also contains translational initiation region (tir) of T7 gene *1*.	Tabor, 1990
pJLA503	ColE1 replicon, carrying P_L, P_R, *c*I857, *bla*, tir *atpE*, and mcs.	Schauder *et al.*, 1987
pET16b	Similar to pT7-7, also contains coding sequence for polyHis and factor X cleavage.	Novagen Inc.
pJF118	ColE1 replicon, carrying *bla*, P_{tac}, and mcs.	Fürste *et al.*, 1986
pACYC184	p15A replicon carrying *cat* and *tetA*.	Chang and Cohen, 1978
pPT7 (pJFMS)	Genes encoding *Eco*KI MTase in pJF118.	Dryden *et al.*, 1993
pJH16	Genes encoding *Eco*RI MTase in pACYC184.	Heitman *et al.*, 1989
R-factor	Genes encoding RI R-M system and *bla*.	Yoshimori *et al.*, 1972
pUC*ral*	4.2 kb *Bam*HI (λ*bio*214) *ral* fragment cloned in pCQV2.	Loenen and Murray, 1986
p*ral*5	400 bp *Hae*III-*Bam*HI *ral* fragment in pUC13.	Loenen and Murray, 1986
pGK1	400 bp *Eco*RI-*Hin*dIII (pUC*ral*) fragment in pT7-6.	This thesis
pGK3	270 bp blunt-ended *ral* fragment (amplified by the PCR) in pUC18.	This thesis
pGK4	220 bp *Nde*I-*Eco*RI (pGK3) *ral*fragment in pJLA503.	This thesis
pGK5	220 bp *Nde*I-*Eco*RI (pGK3) *ral* fragment in pT7-7.	This thesis
pGK6	7.6 kb *Hin*dIII (λNM891) *lar* fragment in pUC18.	This thesis

pGK8	3.5 kb *Bam*HI (pGK7) *lar* fragment in pUC18.	This thesis
pGK9	1.4 kb *Eco*RI-*Cla*I (pGK8; blunt-ended) *lar* fragment in pUC18.	This thesis
pGK10	500 bp *Xho*I-*Eco*RI (pGK9; blunt-ended) *lar* fragment in pUC18.	This thesis
pGK11	500 bp *Hin*dIII-*Eco*RI (pGK10) *lar* fragment in pUC19.	This thesis
pGK12	500 bp *Hin*dIII-*Eco*RI (pGK11) *lar* fragment in pUC18.	This thesis
pGK15	290 bp *Hin*dIII-*Eco*RI *lar* fragment (amplified by the PCR) in pT7-6.	This thesis
pGK16	Similar to pGK15, encodes Lar K42Am.	This thesis
pGK17	Similar to pGK15, encodes Lar M1I.	This thesis
pGK18	Similar to pGK15, encodes Lar M0I.	This thesis
pGK19	Similar to pGK15, encodes Lar M14I.	This thesis
pGK20	Similar to pGK15, encodes Lar M1Am.	This thesis
pGK21	Similar to pGK15, encodes Lar M10Am.	This thesis
pGK22	Similar to pGK15, encodes Lar M14Am.	This thesis
pGK23	Similar to pGK15, encodes Lar M10Am M14I.	This thesis
pGK24	220 bp *Nde*I-*Bam*HI (pGK5) *ral* fragment in pET16b.	This thesis
pGK25	570 bp *Bgl*II-*Bam*HI (pGK5) *ral* fragment in pACYC184.	This thesis

Table 2.3. Plasmids used in this work.

Designation	Sequence (5'-3')	Use
755G	GACTATCCATATGACTACG	Amplification of *ral* using the PCR to introduce an *Nde*I target site (figure 4.3)
642C	GTTTTCCCAGTCACGAC	With 755G (-40 primer)
F799	ATAAGCTTGGGCTGCAGG	T7 reverse sequencing primer
T7 promoter	TAATACGACTCACTATAGGG	T7 promoter sequencing primer
408L	TGTCTAGATATCGGCCTAAC CACGTGGTGCGTACGAGCT CAGGCG	To create an *Eco*KI target used in MTase assays and gel retardations ('top' strand; figure 4.12)
409L	CGCCTGAGCTCGTACGCAC CACGTGGTTAGGCCGATAT CTAGACA	With 408L or 860L ('bottom' strand; figure 4.12)
860L	TGTCTAGATATCGGCCTA^{Me}A CCACGTGGTGCGTACGAGC TCAGGCG	To create a modified *Eco*KI target, used in the MTase assay ('top' strand, modified; figure 4.12)
861L	CGCCTGAGCTCGTACGC^{Me}A CCACGTGGTTAGGCCGATA TCTAGACA	With 860L or 408L ('bottom' strand, modified; figure 4.12)
G2370	GGAAGCTTCAGAGGAATAA TTCAGC	Amplification and mutagenesis of *lar* by the PCR (figure 5.3)
G2371	TACCGAGCGTAGTGTAAC	Site-directed mutagenesis (SDM; Lar K42Am; figure 5.3)
G2372	GGGAATTCTTCCTTTTCAAT AGTGG	With G2370 (figure 5.3)
G2437	GTTACACTACGCTCGGTA	With G2371
G3788	GTCATATCTGATAATTTTTC	SDM (Lar M14I; figure 5.5)
G3789	GAAAAATTATCAGATATGAC	With G3788 (figure 5.5)
G5859	GTGTAATACACCGCCAAC	SDM (Lar M1I; figure 5.5)
G5860	GTTGGCGGTGTATTACAC	With G5859 (figure 5.5)
G7230	TTTATCAGAAAAATTATGAG	SDM (Lar M10I; figure 5.5)
G7231	CTCATAATTTTTCTGATAAA	With G7230 (figure 5.5)
G8266	GTTGGCGGTGCTATACAC	SDM (Lar M1Am; figure 5.5)
G8267	GTGTATAGCACCGCCAAC	With G8266 (figure 5.5)

H1207	TTTTAGAGAAAAATTATGAG	SDM (Lar M10Am; figure 5.5)
H1208	CTCATAATTTTTCTCTAAAA	With H1207 (figure 5.5)
H1831	GTCATATCTCTAAATTTTTC	SDM (Lar M14Am; figure 5.5)
H1832	GAAAAATTTAGAGATATGAC	With H1831 (figure 5.5)

Table 2.4. Synthetic oligonucleotides. T7 promoter (sequencing) primer was purchased from Amersham International plc., all other oligonucleotides were from Oswel DNA service, Dept. Chemistry.

Millipore Corporation; 5 μm Acrodisc filters, used in-line during liquid chromatography, were from Gelman Sciences; Microsep concentrators were from Filtron Technology Corp.

Benzamidine hydrochloride, phenylmethylsulfonyl fluoride (PMSF), 2-mercaptoethanol, dithiothreitol (DTT), acrylamide, N,N'-methylene-bis-acrylamide, N,N,N',N'-tetramethylethylenediamine (TEMED), keyhole limpet hemocyanin (KLH), 3-maleimidobenzoic acid N-hydroxysuccinimide ester (MBS), ethidium bromide, brilliant blue R, bromophenol blue, 1,3-diaminopropane, piperazine, ethanolamine, acetonitrile (HPLC grade), water (HPLC grade), and anti-rabbit IgG horseradish peroxidase conjugate, were purchased from Sigma-Aldrich Company Ltd; trifluoroacetic acid (TFA; HPLC grade) was from Fisons plc.; agarose was from Flowgen Instruments Ltd.; caesium chloride was from ICN Biomedicals Inc.; 2-amino-2-(hydroxymethyl)-1,2-propandiol (Tris) and urea were from Boehringer Mannheim Ltd.; high-purity isopropyl-β-D-thiogalactopyranoside (IPTG) was from Calbiochem-Novabiochem Corp.; 40% w/v acrylamide/bis-acrylamide (19:1) and 30% w/v acrylamide/bis-acrylamide (37.5:1) solutions were from NBL Gene Sciences Ltd., and the ECL detection reagents were from Amersham International plc.

General laboratory supplies were purchased from Sigma-Aldrich Company Ltd., Fisons plc., or Merck Ltd.

Standard solutions and buffers

Ethidium bromide: 10 mg ml^{-1} in deionised water (dH_2O), stored in the dark at 4°C.

Ampicillin: stock solution 100 mg ml^{-1} in dH_2O, 100 μg ml^{-1} end concentration.

Chloramphenicol: stock solution 34 mg ml^{-1} in ethanol, 25 μg ml^{-1} end concentration.

Tetracycline: stock solution 25 mg ml^{-1} in dH_2O, 50 μg ml^{-1} end concentration.

Rifampicin: stock solution 100 mg ml^{-1} in methanol, 200 μg ml^{-1} end concentration.

Antibiotic solutions were sterilised by filtration, and stored at -20°C.

PMSF: 20 mM in ethanol, stored at 4°C.

Benzamidine hydrochloride: 10 mM in dH_2O, stored at 4°C.

Tris-EDTA (TE): Tris base (10 mM) and ethylenediaminetetraacetic acid (EDTA; 1 mM) adjusted to appropriate pH (usually 7.5) with concentrated HCl.

Tris-Borate-EDTA (TBE) 20x: Tris base (1.78 M), boric acid (1.78 M) and EDTA (0.05 M).

Phage buffer: KH_2PO_4 (3 g), Na_2PO_4 (7 g), NaCl (5 g), $MgSO_4.7H_2O$ (1 ml of 0.1 M), $CaCl_2$ (10 ml of 0.1 M), gelatine (1 ml, 1 % w/v), and dH_2O to 1 l.

M9 Salts (4x): $Na_2HPO_4.7H_2O$ (64 g l^{-1}), KH_2PO_4 (15 g l^{-1}), NaCl (2.5 g l^{-1}), NH_4Cl (5.0 g l^{-1}).

18 aa mix: DL-alanine (10 mg ml^{-1}), L-arginine.HCl (2 mg ml^{-1}), L-asparagine (10 mg ml^{-1}), L-aspartate (10 mg ml^{-1}), L-glutamine (10 mg ml^{-1}), glycine (10 mg ml^{-1}), L-histidine (2 mg ml^{-1}), L-isoleucine (2 mg ml^{-1}), L-leucine (2 mg ml^{-1}), L-lysine.HCl (10 mg ml^{-1}), L-phenylalanine (2 mg ml^{-1}), L-proline (3 mg ml^{-1}), DL-serine (10 mg ml^{-1}), DL-threonine (10 mg ml^{-1}), L-tryptophan (2 mg ml^{-1}), L-tyrosine (2 mg ml^{-1}), DL-valine (4 mg ml^{-1}), L-glutamate (10 mg ml^{-1}).

Acetate buffer (pH 5.0): 50 mM acetic acid, adjusted to pH 5.0 with 10 M NaOH.

Tris-MES (pH 6.5 or 8.0): Tris base (20 mM), 2-[N-morpholino]ethanesulphonic acid (MES, 20 mM), $MgCl_2$ (10 mM), EDTA (1 mM), adjusted to required pH with concentrated HCl or 10 M NaOH.

Piperazine buffer (pH 9.5): 20 mM piperazine, adjusted to pH 9.5 with concentrated HCl.

Diamino propane buffer (pH 10.3): 20 mM 1,3-diamino propane, adjusted to pH 10.3 with concentrated HCl.

Phosphate buffer (50 mM, pH 6.0): Na_2HPO_4 (6.0 ml, 1M), NaH_2PO_4 (44 ml, 1M), dH_2O (to 1 l).

Phosphate buffer (10 mM, pH 7.2): Na_2HPO_4 (6.8 ml, 1M), NaH_2PO_4 (3.2 ml, 1M), dH_2O (to 1 l).

Phosphate buffer (pH 11.9): Na_2HPO_4 (50 ml, 50 mM), NaOH (23 ml, 0.1 M), and dH_2O (to 100 ml).

Phosphate-buffered saline (PBS): NaCl (8 g l^{-1}), KCl (0.2 g l^{-1}), Na_2HPO_4 (1.44 g l^{-1}), KH_2PO_4 (0.24 g l^{-1}), adjusted to pH 7.5 with concentrated HCl.

Tris-buffered saline (TBS): Tris (10 mM), NaCl (150 mM) adjusted to pH 7.5 with concentrated HCl.

Cracking buffer: Tris.HCl pH 6.8 (50 mM), sodium dodecyl sulphate (SDS, 2% w/v), bromophenol blue (0.1% w/v), glycerol (10% v/v), 2-mercaptoethanol (10% v/v), DTT (100 mM)

Sterile solutions were prepared by autoclaving at 15 lb in^{-2} for 15 min, unless otherwise stated.

Media

Luria (L) Broth: Difco Bacto tryptone (10 g l^{-1}), Difco Bacto yeast extract (5 g l^{-1}), NaCl (10 g l^{-1}), pH adjusted to 7.2 with NaOH before autoclaving.

Minimal Broth: to sterile dH_2O (300 ml) was added 4x M9 Salts (100 ml), glucose (4 ml, 20%), $MgSO_4$ (1 ml, 1M), $CaCl_2$ (50 µl, 1M) and vitamin B1 (100 µl, 2 mg ml^{-1}).

Luria broth and minimal broth were converted to solid media by adding 15 g l^{-1} and 10 g l^{-1} Difco agar respectively.

Baltimore Biological Laboratory (BBL) Agar: trypticase (10 g l^{-1}), NaCl (5 g l^{-1}), Difco agar (10 g l^{-1}; top layer 6.5 g l^{-1}).

TY (2x): Difco Bacto yeast extract (10 g l^{-1}), Difco Bacto tryptone (16 g l^{-1}), NaCl (10 g l^{-1}).

SOC: L broth (100 ml) supplemented with glucose (1.8 ml, 20% w/v), $MgSO_4$ (1 ml, 1 M) and $MgCl_2$ (1 ml, 1 M).

Media was sterilised by autoclaving at 15 lb in^{-2} for 15 min.

Manipulation of bacteria and phages

Long-term storage of bacterial cells

A single colony was grown overnight in liquid media, under appropriate conditions, and used to inoculate L-agar in a small bottle ('stab'). After overnight incubation, the bottle cap was sealed and the stab stored at room temperature.

Preparation of plating cells

An overnight culture was diluted 20-fold in L-broth, grown to mid-logarithmic phase (OD_{650} ~0.6, ~5 x 10^8 cfu ml^{-1}), and the cells harvested by centrifugation (3 000 x *g*, 5 min). The pellet was resuspended in the same volume of 10 mM $MgSO_4$, as the culture the cells were harvested from, and stored at 4°C.

Preparation of λ lysates

A single plaque was picked using a toothpick and resuspended in 1 ml phage buffer containing a few drops of chloroform (10^7 pfu ml^{-1} of λ*c*I). 0.1 ml of an appropriate

dilution of the phage, to give confluent lysis (~10^5 pfu), was mixed with an equal volume of plating cells, left for 15 min to allow phage adsorbtion, and 2.5 ml molten BBL soft agar was added before the mixture was poured on fresh L-agar in a petri dish ('plate'). The plate was incubated until confluent lysis was observed (usually 7-8 hr) and placed at 4°C overnight with an overlay of L-broth (4 ml). The broth was decanted, remaining cells lysed with a few drops of chloroform, and clarified by centrifugation (3 000 x *g*, 5 min). The phage lysate was transferred to a fresh bottle and stored at 4°C.

Restriction alleviation tests

Plating cells were prepared from a restriction-proficient strain transformed with the plasmid of interest, and infected with appropriate dilutions of modified and unmodified phages. After adsorbtion (15 min), infective centres were assayed by plating in top-layer (2.5 ml) on BBL agar. Plaques were counted after overnight incubation and the titre of unmodified phages was normalised against the titre of modified phages, to correct for variation resulting from factors other than restriction.

Modification enhancement tests

Plating cells were prepared from a restriction-deficient, modification-proficient strain, transformed with the plasmid of interest (if relevant), and infected with unmodified λ phages at a multiplicity of infection (moi) of 0.1 (normally, 100 µl each in a 1.5 ml microcentrifuge tube). After adsorbtion of the phages (15 min), cells were harvested by centrifugation (11 000 x *g*, 5 min) to remove unabsorbed phages, resuspended in pre-warmed L-broth (10 ml in a 100 ml flask), and grown with aeration under appropriate conditions (normally 37°C). After the phages had undergone a single round of infection (70 min), a sample (1 ml) of the lysate was removed and remaining cells lysed by the addition of a few drops of chloroform. The degree of phage modification was estimated by measuring the efficiency of plating (eop) of progeny phages on a restricting strain versus a non-restricting strain. Modified λ phages were used as a control to ensure that the plating efficiencies of the indicator bacteria were comparable.

Construction of strains carrying hsdRΔ4

The *hsdR*Δ4 derivatives of NM710 (NM792), and NM777 (NM794) were made by transferring the deletion from an λ*hsd* phage to the bacterial chromosome (Gough and Murray, 1983). The host was infected with λNM1118 (λ*hsdR*Δ4*c*I857Δ*att*) and

grown in top-layer on BBL-agar overnight at 30°C. Lysogens were selected by streaking samples from the centre of lysis on L-agar seeded with homo-immune λ*c*I phages of different host ranges (~10^9 pfu λNM63 and λNM220 or λNM144). Cells were purified and lysogens confirmed by their sensitivity to λ*vir*, but immunity to λ*c*I phages. λNM1118 has a deletion removing the attachment site and is only able to integrate efficiently into the host chromosome by homologous recombination, therefore, lysogens probably have the phage genome integrated in the *hsd* genes.

The λNM1118 lysogens were streaked on L-agar and grown overnight at 42°C. Of the surviving cells, a proportion will retain the *hsdR*Δ4 deletion after excision of the λ genome from the host chromosome. The cured-lysogens were purified and hosts that were no longer able to restrict unmodified phages were presumed to have acquired the *hsdR*Δ4 deletion.

Preparation of λDE3 lysogens

Phage λDE3 (Studier and Moffatt, 1986) contains T7 gene *1*, encoding T7 RNA polymerase, inserted in the *int* gene of a phage λ vector (λD69; Mizusawa and Ward, 1982). In order for λDE3 to lysogenise its host, the *int* gene product has to be provided *in trans*. λDE3 was isolated from a culture of BL21(DE3), by infecting plating cells (100 µl) with the helper phage λNM75 (moi ~10; $\lambda imm^{\lambda} int^{+} xis^{+}$) and selecting for λDE3 (imm^{21}) by growth on a λimm^{λ} lysogen (NM477imm^{λ}). Phages were plaque-purified and the immunity of phage 21 confirmed by their inability to grow on BL21(DE3).

Lysogens of λDE3 were prepared by co-infecting cells with λDE3 and λNM75 (moi ~10) in top layer on BBL-agar, and selecting for lysogens with the immunity of phage 21 by streaking on L-agar seeded with two homo-immune λ*c*I phages of different host ranges (~10^9 pfu λNM507 and λNM508). Lysogens were purified and the immunity confirmed by their sensitivity to λ*vir*, but not λimm^{21} phages. Monolysogens were confirmed by their sensitivity to λ*c*I, and these imm^{21} hosts were presumed to have been lysogenised by λDE3.

P1 transduction

Transduction using P1*kc* was achieved in two stages (Miller, 1992). A P1 lysate was made on the donor strain, resulting in the packaging of host DNA in a small proportion of phage heads, and the recipient transduced by growth with the phage and selection for recombinants.

P1 lysates were prepared in a similar way to λ plate lysates. An overnight culture of the donor strain was diluted 20-fold in L-broth supplemented with $CaCl_2$ (5 mM) and grown to mid-logarithmic phase (OD_{650} ~0.6) with good aeration at 37°C. Cells were harvested (3 000 x *g*, 5 min) and resuspended in a tenth volume of L-broth containing $CaCl_2$ (5 mM). Dilutions of P1*kc* (100 µl) were used to infect an equal volume of cells and allowed to adsorb for 20 min before the mixture was plated in top layer (2.5 ml) on L-agar. Lysates were harvested after overnight incubation at 37°C by adding L-broth (4 ml) and transferring to 4°C, for a minimum of 6 hr, before decanting the phage lysate. Remaining cells were lysed by adding a few drops of chloroform, the phage lysate clarified by centrifugation and stored at 4°C.

Cells from an overnight culture of the strain to be transduced were harvested by centrifugation (3 000 x *g*, 5 min) and resuspended in an equal volume of an aqueous solution containing $MgSO_4$ (10 mM) and $CaCl_2$ (5 mM). Samples (100 µl) of the P1*kc* phage lysate (grown on the donor strain and serially diluted) were mixed with an equal volume of cells, and incubated at 37°C for 20 min to allow phage adsorbtion. Sodium citrate (200 µl; 1 M, pH 8.5) was added, to prevent re-infection of the bacteria by the P1 phages, and infected cells were spread on solid selective media. After incubation (usually at 37°C, for 24 to 36 hr), transductants were purified and phenotypes confirmed or, in the case of co-transduction, transductants were screened for the required phenotype.

Construction of NM793(DE3)

NM793(DE3) was made by a series of P1 transductions. An efficient donor of *supF*, tagged with a closely linked transposon, was generated by the P1-mediated transfer of *trpC*::Tn*10* from a *sup*° strain (W3110*trpC*::Tn*10*) to the *supF* strain, Ymel. Tetracycline-resistant transductants of Ymel were selected on L-agar supplemented with 10 µg ml^{-1} tetracycline and screened for the retention of *supF* by their ability to suppress λ*Sam*7 (λNM123), but not λ*Pam*3 (λNM435). Phage P1 grown on this host (NM787) was used to transduce BL21(DE3). Tetracycline-resistant transductants were selected for as before and screened for the presence of *supF* by their ability to suppress λ*Nam*7*Nam*53 (λNM22), a phage chosen to avoid complementation by the resident imm^{21} prophage. Phage P1 was grown on the $trpC^+$ host and used to regenerate a prototrophic strain (NM793). These were selected on M9-minimal agar and screened for the presence of the supressor mutation as described above.

Manipulation of nucleic acids

Large-scale preparation of plasmid DNA

Large amounts of purified plasmid DNA were prepared by a modified method of Clewell and Helinski (1969). Cells transformed with the required plasmid were grown overnight (250 ml, usually at 37°C) and harvested by centrifugation (8 000 x *g*, 10 min). The pellet was resuspended in lysis solution (7 ml; 25 mM Tris.HCl pH 8.0, 10 mM EDTA pH 8.0, 1% w/v glucose), cells disrupted by the addition of lysozyme (10 mg) and left on ice for 5 min. Freshly prepared alkaline-SDS (14 ml; 0.2 M NaOH, 1% w/v SDS) was added to the mixture and left for 10 min on ice before potassium acetate (3.5 ml; 3 M, pH 5.0) was added. After a further 5 min incubation on ice, the DNA solution was clarified by centrifugation (8 000 x *g*, 10 min) and passed through a funnel stuffed with glass wool to remove any remaining debris. Nucleic acid was precipitated with isopropanol (15 ml) at 4°C (15 min) and sedimented by centrifugation (8 000 x *g*, 10 min). The pellet was washed with 70% ethanol, dried under vacuum (5 min), and resuspended in dH_2O (9 ml) containing ethidium bromide (0.5 ml, 10 mg ml^{-1}). CsCl (9.5 g) was dissolved in the DNA solution, transferred to a Sorvall Ultracrimp tube (DuPont Ltd.) and a density gradient formed by centrifugation (130 000 x *g*, 48 hr).

The DNA was visualised with uv-light and two bands were normally distinguished. The upper band contains chromosomal DNA, linearised and nicked plasmid DNA, whereas the lower band should contain only supercoiled plasmid DNA and this was collected using a hypodermic needle and syringe to puncture the side of the tube. Ethidium bromide was extracted in isopropanol saturated with TE and NaCl, by vigorously mixing the two solutions and leaving the phases to separate before removing the solvent layer. The solution was diluted in two volumes of dH_2O and the DNA precipitated by adding a further two volumes of ice-cold ethanol. The CsCl, removed with the DNA from the gradient, was not usually precipitated using this method, but if it was thought necessary, excess salt was removed by dialysis. The DNA was sedimented by centrifugation (29 000 x *g*, 15 min), rinsed with 70% ethanol and dried under vacuum. The DNA was resuspended in dH_2O (500 µl), any remaining protein removed using phenol and precipitated with ethanol (see below).

Removal of proteins with phenol

Proteins were removed from solutions containing DNA by vigorously mixing equal volumes of DNA and buffered phenol solutions, separating the two phases by

centrifugation (11 000 x *g*, 3 min) and removing the aqueous, DNA-containing, layer to a fresh tube. Buffered phenol solutions were prepared from either solid or water-saturated phenol (Hastings and Kirby, 1966). Tris.HCl pH 7.8 (110 ml, 1 M), m-cresol (14 ml), 2-mercaptoethanol (500 µl), and hydroxyquinoline (0.28 g) were added to the phenol (250 g solid and 14 ml dH_2O or 241 ml water-saturated phenol) and mixed using a magnetic stirrer-bar for 30 min at 20°C. The buffered phenol solution was transferred to a polypropylene bottle and stored in the dark at 4°C

Precipitation of DNA with ethanol

DNA in solution was precipitated by the addition of a tenth volume of sodium acetate (3 M, pH 5.3), two volumes of ice-cold ethanol, and left at -20°C for 20 min. The DNA was sedimented by centrifugation (11 000 x *g*, 20 min), rinsed with 70% ethanol, dried under vacuum and resuspended in an appropriate volume of dH O or TE buffer.

Small-scale preparation of plasmid DNA

Cells transformed with the plasmid were harvested by centrifugation (11 000 x *g*) from a culture (1.5 ml) grown overnight and plasmid DNA isolated by a method based on those of Holmes and Quigley (1981) and He *et al.* (1990). The pellet was resuspended in TELT buffer (400 µl; 50 mM Tris.HCl pH 7.5, 62.5 mM EDTA pH 7.5, 2.5 M LiCl, 0.4% v/v Triton X100), and left at 20°C (2 min) after the addition of lysozyme (40 µl of 100 mg ml^{-1} in TELT). The mixture was heated to 100°C (2 min) to denature proteins, then cooled on ice (10 min) and clarified by centrifugation (11 000 x *g*, 15 min). Plasmid DNA in the supernatant was precipitated by adding isopropanol (200 µl) at 4°C (2 min) and the DNA was sedimented by centrifugation (11 000 x *g*, 20 min). The pellet was rinsed with 70% ethanol, dried under vacuum, resuspended in dH_2O or TE buffer (20 µl) and stored at -20°C.

Small-scale preparation of phage λ DNA

Phage DNA was prepared by a method devised by A. Titheradge (unpublished). A culture of an appropriate host was diluted 40-fold in L-broth (4 ml), supplemented with $MgSO_4$ (10 mM), infected with the required phage at a moi ~5 (~10^8 pfu) and grown with aeration until lysis (normally, 37°C for 4 hr). Remaining cells were lysed by the addition of chloroform (100 µl), cell debris was removed by centrifugation (3 000 x *g*, 10 min), and nucleases (5 µl, 10 mg ml^{-1} RNase and DNase) added to the supernatant in a fresh tube. The lysate was incubated at 37°C for 30 min, and the

phage precipitated by the addition of polyethylene glycol (PEG) and salt (4 ml; 20 g 6K PEG, 11.7 g NaCl, 78 ml phage buffer). The mixture was left overnight at 4°C, and the phage particles sedimented by centrifugation (13 000 x *g*, 20 min, 4°C). Care was taken to remove most of the PEG, and the phages were resuspended in phage buffer (500 µl). Chloroform was added (500 µl) to remove any remaining PEG, the mixture clarified by centrifugation (5 000 x *g*, 3 min) and the aqueous phase (containing the phage particles) transferred to a tube containing phenol (500 µl) and TE buffer (100 µl). The two phases were separated by centrifugation (11 000 x *g*, 3 min), and the aqueous phase transferred to a tube containing phenol (250 µl) and chloroform (250 µl). The aqueous phase, containing naked phage DNA, was separated as before, mixed with an equal volume of chloroform and the aqueous phase collected. To remove some of the excess salt carried over from the stages with phage particles, the DNA in solution was precipitated with ethanol twice, resuspended in TE (100 µl) and stored at -20°C.

Cutting of DNA with type II ENases

DNA was cut by type II ENases in buffers recommended and provided by the supplier. Digests were normally done with 1 to 2 µg DNA, and 5 units of enzyme in 1x buffer (10 µl), and incubated at the recommended temperature (normally 37°C) for a minimum of 1 hr.

Removal of 5'-phosphates

The phosphates were removed from the 5' end of vector DNA to aid the recovery of recombinant molecules, as these ends are not a substrate for DNA ligase. The reaction was done in the buffer supplied with the enzyme (TA buffer), generally with 1 to 2 µg DNA, $CaCl_2$ (5 mM) and 1 unit of HK phosphatase. The mixture was incubated at 30°C for 1 hr, and the enzyme inactivated by heating to 65°C for 30 min.

Generation of blunt-ended fragments

Recessed 3'-ends were filled using Klenow DNA polymerase, to generate blunt-ended fragments. Klenow polymerase (1 unit in 20 µl) and dNTPs (0.4 mM) were added to the DNA solution (usually in TE buffer), and the mixture was incubated at 20°C for 30 min.

Ligation of DNA

DNA was ligated using T4 DNA ligase (Murray *et al.*, 1979), in buffer supplied with the enzyme (66 mM Tris.HCl pH 7.5, 5 mM $MgCl_2$, 1 mM DTT, 1mM ATP). The ligation reaction was normally done in a final volume of 10 µl buffer, with 50 to 100 ng vector DNA, 0.5 units enzyme and incubated at 16°C overnight.

Agarose gel electrophoresis

DNA fragments were separated by electrophoresis through horizontal agarose gels submerged in 1x TBE containing ethidium bromide (0.8 mg ml^{-1}). The concentration of agarose varied, depending on the size of DNA fragments, but was generally between 0.7 and 1.5% w/v. Samples containing DNA were mixed with Ficol-loading dye (30% w/v; 40% w/v Ficol 400, 0.05% w/v bromophenol blue in TE buffer), applied to the gel and electrophoresis done at 5 Vcm^{-1}. The DNA was visualised over a uv-light transilluminator and photographed with a video copy processor (Mitsubishi Electronics Corp.). λ DNA (λ*Sam7c*I857) cut with *Hin*dIII, or a 123 bp ladder (Life Technologies Ltd.) were routinely used as standards.

Recovery of DNA from agarose gels

The region of the gel containing the required fragment of DNA was isolated using a scalpel and placed in dialysis tubing closed at one end. The piece of gel was covered with 1x TBE, any air removed and the tubing sealed. The dialysis tubing was submerged in 1x TBE and the DNA eluted from the agarose by electrophoresis (5 V cm^{-1}, 30 min). The DNA was released from the sides of the tubing by reversing the direction of electrophoresis (10 sec), and the DNA solution removed to a microcentrifuge tube. Ethidium bromide was extracted with dH_2O-saturated isobutanol, protein removed with phenol, the DNA precipitated with ethanol and resuspended in an appropriate volume of dH_2O or TE buffer.

Transformations

Cells were made competent by electroporation (Heery and Dunican, 1989). An overnight culture was diluted 20-fold in 2x TY media and grown to mid-logarithmic phase (OD_{650} ~0.6), the cells harvested by centrifugation (3 000 x *g*, 5 min, 4°C) and washed three times in ice-cold dH_2O. The pellet was resuspended in an equal volume of dH_2O (normally 100 µl) and samples (40 µl) added to dilute DNA solutions (usually 5 to 20 ng in 5 µl). The mixture was placed in an electroporation cuvette (0.2 cm) and pulsed (4.8 msec) with an electric field (12.5 kV cm^{-1}) using a Gene Pulser

electroporator (BioRad Laboratories Ltd.; settings, 25 μF, 2.5 kV, 200 Ω). SOC media (1 ml) was immediately added to the mixture, transferred to a small bottle and incubated under appropriate conditions to allow expression of genes whose products confer the selective advantage (normally, 37°C for 20 min). Samples (10 and 100 μl) were spread on selective agar and incubated overnight.

The polymerase chain reaction

DNA was amplified using the polymerase chain reaction (PCR) essentially as described by Saiki *et al.* (1988) using either Taq or $Vent_R$ DNA polymerase and the buffers supplied with the enzymes. The reaction mixture contained oligonucleotide primers (200 nM), dNTPs (50 μM), buffer (20 mM $(NH_4)SO_4$, 75 mM Tris.HCl pH 8.8, 0.1% v/v Tween 20, 2.5 mM $MgCl_2$), template and 50 U ml^{-1} Taq DNA polymerase, or primers (400 nM), dNTPs (200 μM), buffer (10 mM KCl, 10 mM $(NH_4)SO_4$, 20 mM Tris.HCl pH 8.8, 0.1% v/v Triton X100, 2 mM $MgCl_2$), template and 20 U ml^{-1} $Vent_R$ DNA polymerase. The template was usually phage DNA (approximately 1 ng μl^{-1}) although phage lysates, colonies, plasmids or products of a primary PCR, were also used as a source of the DNA template. In general, amplification with $Vent_R$ DNA polymerase required more template than Taq. Amplification was performed in an OmniGene thermal cycler (Hybaid Ltd.) and a general program profile consisted of three stages. The DNA was denatured for 5 min to provide single-stranded template, followed by multiple cycles of denaturation (1 min), annealing (40 sec) and extension (40 sec), with the last extension stage increased to 10 min.

DNA denaturation was at 96°C for $Vent_R$, but 94°C for the less thermo-tolerant enzyme, Taq. The choice of annealing temperature is dependent on the DNA sequence, and was estimated to be 5°C lower than the melting temperature of the sequence. The melting temperature T_m was estimated using the equation:

$$T_m \sim 4[G+C] + 2[A+T]$$

The choice of annealing temperature was the factor that appeared to have the greatest influence on the amount of product obtained. Extension from the primers was done at 72°C, and the final incubation of the reaction at 72°C for 10 min assured that most of the products were full-length. The denaturation-annealing-extension stage was generally repeated 30 times, although this was reduced to 20-25 times when the fidelity of replication was considered important, e.g. site-directed mutagenesis.

End-labelling of oligonucleotides

Oligonucleotides were labelled for use in gel retardations by using T4 polynucleotide kinase (PNK) to transfer a labelled-phosphate from [γ-^{32}P]dATP to the 5' end of the DNA (Midgley and Murray, 1985). The oligonucleotide was mixed in buffer (20 µl reaction volume; 500 mM Tris.HCl pH 7.5, 100 mM $MgCl_2$, 50 mM DTT) with [γ-^{32}P]dATP (1 µl, ~3 000 Ci mmol^{-1}) and T4 PNK (50 fold dilution in 25 mM Tris.HCl pH 7.5, 10 mM 2-mercaptoethanol, 50% v/v glycerol), and incubated at 37°C for 45 min. The enzyme was inactivated by heating to 68°C for 10 min, and the unincorporated label removed using a NuTrap probe-purification push column (Stratagene Ltd.). The oligonucleotide was eluted in TE buffer (70 µl) and annealed to an unlabelled, complimentary oligonucleotide to generate double stranded DNA for gel retardation.

Products of the PCR reaction were 'end-labelled' using 'cold' dATP if they were cloned as blunt-ended fragments. The reaction conditions were similar, but 30 mM ATP was substituted for [γ-^{32}P]dATP and the products were not column-purified.

Nucleotide sequence analysis

Nucleotide sequence was determined by chain-termination sequencing (Sanger *et al.*, 1977) using Sequenase 2.0, and the method described is based on the recommended procedure (Amersham International plc.). Double-stranded DNA template (3 to 5 µg in 10 µl) was mixed with an oligonucleotide primer (1 to 50 pmoles) and denatured by heating to 100°C for 4 min, followed by rapid freezing in dry ice. A reaction mix was made by combing DTT (2 µl, 0.1 M), Sequenase labelling mix (0.7 µl; 1.5 µM dGTP, 1.5 µM dCTP, 1.5 µM dTTP), Sequenase reaction buffer (2 µl; 200 mM Tris.HCl pH 7.5, 100 mM $MgCl_2$, 250 mM NaCl), [^{35}S]dATPαS (0.5 µl; 400 Ci mmol^{-1}) and Sequenase 2.0 T7 DNA polymerase diluted eight-fold in enzyme dilution buffer (2.5 µl; 10 mM Tris.HCl pH 7.5, 5 mM DTT, 0.5 mg ml^{-1} bovine serum albumin). The reaction mixture was added to the thawed template-primer and incubated at 20°C for 4 min. Samples (4 µl) were added to each of 4 ddNTP termination mixes (each termination mix contained 80 µM dGTP, dATP, dCTP and dTTP in 50 mM NaCl with 8 µM of either ddGTP, ddATP, ddCTP, or ddTTP) and incubated at 37°C before the reaction was halted by the addition of stop solution (4 µl; 95% v/v formamide, 20 mM EDTA, 0.05% w/v bromophenol blue, 0.05% w/v xylene cyanol FF). The DNA solutions were denatured by heating to 100°C for 3 min, and the products separated by polyacrylamide gel electrophoresis (PAGE) in the presence of urea.

When reactions used template DNA amplified by the PCR, dimethyl sulphoxide (DMSO) was added to each stage of the reaction (10% v/v) to help prevent problems related to secondary structure formation (Winship, 1989).

DNA-denaturing PAGE

The products of the nucleotide sequencing reactions were separated on a 6% polyacrylamide gel using SequiGen apparatus (BioRad). A stock solution was made containing acrylamide/bis-acrylamide (21 ml; 40% w/v, 19:1), urea (59.5 g), TBE buffer (3.5 ml, 20x), dH_2O (to 140 ml) and stored in the dark at 4°C. The base of the gel plates was sealed with a small volume (10 ml) of the urea-acrylamide mix polymerised with ammonium persulphate (AMPS; 125 µl, 10% w/v) and TEMED (50 µl), and the gel formed by pouring a solution of urea-acrylamide mix (30 ml), AMPS (125 µl, 10% w/v), and TEMED (50 µl) between the two plates. The acrylamide was allowed to polymerise and electrophoresis done in 1x TBE buffer at a constant power of 40 W. The gel was pre-warmed to 55°C and the samples heated to 100°C (3 min), prior to loading. After electrophoresis, the gel was soaked in a solution containing methanol (10% v/v) and acetic acid (10% v/v) for 20 min, to fix the DNA and remove urea from the gel. The gel was dried at 80°C under vacuum for 90 min, and labelled-DNA visualised by autoradiography.

Autoradiography

Autoradiography was used to detect DNA labelled with ^{35}S in sequencing reactions, ^{32}P for gel retardations, and proteins from cells grown in the presence of [^{35}S]cysteine. Gels dried on filter paper were placed in direct contact with BioMax (detection of ^{35}S) or Cronex (detection of ^{32}P) X-ray film, and stored at 20°C (^{35}S) or -70°C (^{32}P) until developed using an automatic film processor (X-Ograph Ltd.).

Fluorography

Fluorography was used to detect [methyl-^{3}H]DNA, methylated by the *Eco*KI MTase in the presence of labelled-AdoMet. DNA was separated from AdoMet by non-denaturing PAGE, and the DNA fixed in the gel by soaking in a solution of methanol (10% v/v) and acetic acid (10% v/v) for 30 min with agitation. After fixing, the gel was covered in Amplify (Amersham International plc.), soaked for a further 30 min as before, and dried on a piece of blotting paper under vacuum at 80°C for 2 hr. The dried gel was placed in direct contact with Hyperfilm MP and stored at -70°C until the film was developed using an automatic film processor.

Manipulation of proteins

Liquid chromatography was carried out using a modular system purchased from either Pharmacia Biotech Ltd. or ISCO Inc., and high performance (high pressure) liquid chromatography (HPLC) used a modular system from Gilson Medical Electronics Inc. Spectroscopy was done on a Perkin Elmer Lambda 5 spectrophotometer, and proteins were further analysed by N-terminal sequence analysis (A. Cronshaw, Dept. Biochemistry) and matrix-assisted laser desorption ionisation, time of flight (MALDI-TOF) mass spectroscopy (K. Shaw, Dept. Chemistry).

Columns and media

The Resource S (1 ml, cation exchanger), Resource Q (1 ml, anion exchanger), Hi-Load 16/60 Superdex 200 (gel filtration), PD-10 (gel filtration), and Hi-Trap NHS activated (1 ml, affinity chromatography) pre-packed columns were purchased from Pharmacia Biotech Ltd.; The Econo-Pac Q (5 ml, anion exchanger) column was from BioRad Laboratories Ltd., and the Vydac C_{18} (110 x 4.6 mm, HPLC) column was from Anachem Ltd.

Diethylaminoethyl- (DEAE-) Sepharose (anion exchanger), Carboxymethyl- (CM-) Sepharose (cation exchanger) and Heparin-agarose were purchased from Sigma-Aldrich Company Ltd.; Sephadex G-50 was from Pharmacia Biotech Ltd, and Hydroxyapatite-BioGel was from BioRad Laboratories Ltd. When required, the resin was packed in a modular-system C-column from Pharmacia Biotech Ltd.

Determination of protein concentration in solution

The concentration of proteins in solution was estimated, from measurements of uv-light absorption, using the equation (Harris, 1987):

$$\text{protein (mg ml}^{-1}) \sim 1.55\ Ab_{280} - 0.76\ Ab_{260}$$

The concentration of Ral and the *Eco*KI MTase was calculated from the absorbance at 280 nm (Ab_{280}), using the molar extinction coefficient (ε) estimated from the sum of values of the absorbance (Mihalyi, 1970) of tyrosine, tryptophan and phenylalanine residues in the proteins (ε, at 280 nm of a 1 mg ml^{-1} solution, *Eco*KI MTase = 0.842, Ral = 1.62).

The absorption spectrum between 240 and 340 nm was routinely used to examined purity and aggregation of proteins in solution. The purity was estimated by taking the ratio of absorbance at 280 nm and 250 nm; aromatic residues, e.g. tryptophan, absorb at 280 nm whereas nucleotides, in particular, absorb at 250 nm. Particles in the

solution, e.g. protein aggregates, scatter light that is seen as a deviation in the base line at higher wavelengths.

Coupling of Ral to keyhole limpet hemocyanin

Peptides used as immunogens are usually coupled to a carrier protein to enhance the immune response to the peptide (Gnann *et al.*, 1989). Synthetic Ral may be sufficiently large not to require a carrier protein to stimulate an immune response, but it was coupled to the carrier keyhole limpet hemocyanin (KLH) as this was unlikely to decrease its immunogenic properties. KLH was activated with maleimidobenzoic acid-N-hydroxysuccinimide ester (MBS). An KLH solution was made in phosphate buffer (pH 7.2), dialysed overnight, and adjusted to a concentration of 20 mg ml^{-1}. A sample (250 µl) was removed, 55 µl phosphate buffer (10 mM, pH 7.2) added and mixed with 85 µl MBS (6 mg ml^{-1} in dimethylformamide, DMF). MBS was added dropwise to ensure that local concentrations of DMF did not cause precipitation of the KLH, and the mixture was stirred for 30 min at 20°C. A column packed with Sephadex G-25 (PD-10) was washed with phosphate buffer, pH 6.0 (50 mM), and the protein solution applied to separate the activated KLH from free MBS. The eluted protein (~1.5 ml) was mixed with an equal volume of synthetic Ral, dissolved in phosphate-buffered saline (PBS pH 7.5; ~2 mg unpurified synthetic Ral), and was left, with stirring, at 20°C (3 hr). The coupled KLH-Ral mixture was divided in two, and used to inoculate rabbits (R707 and R708; G. Ledbetter, Dept. Medical Microbiology; first inoculation mixed with Freund's complete adjuvant).

Coupling of Ral to NHS-activated Sepharose

Synthetic Ral (Brown *et al.*, 1993) was covalently coupled to agarose using a HiTrap N-hydroxysuccinimide- (NHS-) activated Sepharose column (1 ml). The column was washed with 6 volumes of HCl (1 mM, 4°C) and synthetic Ral (0.6 mg; AB72B and AB80B) applied in 1 column volume of a solution containing NaCl (0.5 M) and $NaHCO_3$ (0.2 M, pH 8.3) at 1 ml min^{-1}. The column was sealed and left at 4°C (4 hr) to allow coupling of the ligand. The NHS-Sepharose was deactivated, and excess ligand eluted from the column, by washing with 2 buffers alternately, buffer A (0.5 M ethanolamine, 0.5 M NaCl, pH 8.3) and buffer B (0.1 M acetic acid, 0.5 M NaCl, pH 4.0). Six column-volumes of buffer were passed through the column sequentially at 1 ml min^{-1} (A, B, A, left 30 min, B, A, B), neutralised with 2 column-volumes of Tris.HCl pH 8.0 (100 mM) and stored in 20% ethanol.

Gel retardation

Gel retardations were done essentially as described by Powell *et al.* (1993). The cofactor AdoMet (100 μM) was incubated separately with ^{32}P-labelled DNA and protein (0.2 nM DNA; 100 nM purified *Eco*KI MTase and 2 mM Ral) for 10 min at 4°C in buffer containing 20 mM Tris.HCl pH 8.0, 100 mM NaCl and 5% glycerol. Equal volumes of protein and DNA mixtures were combined and incubated for 10 min at 4°C, followed by the addition of purified Ral antibodies (10 μl). After 20 min at 4°C, complexes were separated by non-denaturing PAGE (5% w/v polyacrylamide) and labelled-DNA visualised by autoradiography.

Non-denaturing PAGE

Non-denaturing PAGE was used for gel retardations (5% w/v polyacrylamide) and for the separation of labelled-DNA and [methyl-^{3}H]AdoMet in MTase assays (12% w/v polyacrylamide). AMPS (100 μl, 10% w/v) and TEMED (40 μl) were added to a mixture of acrylamide (10.1 ml for a 12% gel, 4.2 ml for a 5% gel; 30% w/v acrylamide/bis-acrylamide [37.5:1]), TBE (2.5 ml, 20x) and dH_2O (made-up to 25 ml) and poured between two glass plates sealed with the spacer (1 mm; ATTO Corp.) held together with bulldog clips. After polymerisation, samples were applied to the gel in glycerol (10%) and electrophoresis done in 1x TBE buffer at a constant current of 30 mA. The gel-front was marked by applying a sample (5 μl) of Ficol loading buffer to an empty well.

Labelling of proteins

Proteins encoded by genes under the control of the T7 promoter were labelled with [^{35}S]cysteine essentially as described by Tabor (1990). Rifampicin specifically inhibits *E. coli* RNA polymerase, and genes under the control of T7 RNA polymerase are still expressed in the presence of the antibiotic. These gene products may be specifically labelled after addition of the antibiotic. BL21(DE3), or NM794(DE3), was transformed with the plasmid of interest and grown at 37°C with good aeration to an OD_{650} ~0.4 in L-broth supplemented with ampicillin (100 μg ml^{-1}). Cells from 1 ml of culture were harvested by centrifugation (11 000 x *g*, 5 min) and resuspended in M9-minimal media. The cells were harvested, as before, and resuspended in 1 ml M9-minimal media supplemented with 18 aa mix (5% v/v) and ampicillin. The culture was incubated, with aeration at 37°C for 60 min, and expression of the T7 RNA polymerase gene induced by the addition of 0.6 mM IPTG. After 20 min, cultures were warmed to 42°C and, 10 min later, rifampicin added (200 μg ml^{-1}). Cells were

incubated with rifampicin for 10 min at 42°C, as the antibiotic efficiently inhibits expression of host-encoded proteins at this temperature, and the culture was shifted to 37°C for 20 min, a temperature suitable for efficient labelling. A sample (500 µl) was removed, [^{35}S]cysteine added (10 µCi), and cells incubated for a further 5 min at 37°C. Cells were harvested by centrifugation (11 000 x *g*, 5 min), resuspended in 'cracking' buffer (100 µl) and heated to 100°C (5 min). Polypeptides were separated by Tris-tricine SDS-PAGE and labelled proteins visualised by autoradiography.

Tris-tricine SDS-PAGE

SDS-PAGE with a buffer containing tricine was used (Schägger and von Jagow, 1987), instead of the usual Tris-glycine buffer system (Laemmli, 1970), to achieve good separation of proteins in the 5 to 20 kDa range. This discontinuous method uses three gels, a stacking, spacer and separating gel. SDS is used to ensure a similar charge to mass ratio for each protein, hence, proteins are separated on the basis of molecular mass. To get good resolution, it is important that proteins in the sample reach the separating region of the gel at the same time, and this is achieved by the stacking of proteins in the stacking gel. The spacer gel also helps to stack low molecular weight proteins, although if high resolution was not thought to be a critical factor then this gel was omitted. tricine, instead of the usual glycine, also aids the stacking of proteins under 30 kDa.

The stacking gel contained 4% acrylamide (1 ml; all gels used a 49.5% w/v acrylamide/bis-acrylamide [32:1] stock solution), gel buffer (3.1 ml; 3 M Tris.HCl pH 8.45, 0.3% w/v SDS), and dH_2O (to 12.5 ml). The spacer gel contained 10% acrylamide (6.1 ml), gel buffer (10 ml) and dH_2O (to 30 ml). The separating gel contained 16.5% acrylamide (10 ml), gel buffer (10 ml), glycerol (4 ml) and dH_2O (to 30 ml).

The acrylamide solutions were polymerised by the addition of AMPS (100 µl, 10% w/v) and TEMED (10 µl) and poured between two sealed glass plates (ATTO Corp., or Hoffer Scientific Instruments for minigels). The stacking and spacer gels were made approximately twice the thickness of the wells, although the spacer gel was usually omitted when using minigels. A flat interface between different gels was achieved by overlaying each gel with butan-2-ol until polymerised, and removing the solvent before applying the next layer. Samples were mixed in 'cracking' buffer, and heated to 100°C (5 min) before applying to the gel. Electrophoresis was done with different anode and cathode buffers (0.2 M Tris.HCl pH 8.9 and 0.1 M Tris 0.1 M tricine pH 8.25, respectively), initially at a constant voltage of 30 V, until the sample

had completely entered the stacking gel, and then increased to 85 V overnight (150 V for a minigel, 4°C, 4 hr). Low molecular weight markers (Sigma-Aldrich Ltd.) were used as standards.

Staining of proteins with coomassie

Proteins were stained in an aqueous solution of ethanol (250 ml l^{-1}), formaldehyde (10 ml l^{-1}, 35%) and brilliant blue R (1.2 g l^{-1}), and the stain removed from the gel by soaking in a solution of methanol (10% v/v) and acetic acid (10% v/v). The destain was changed periodically and timings were arbitrary.

Western blotting

Proteins were transferred from a polyacrylamide gel to PVDF membrane by wet elecrophoretic transfer (Towbin *et al.*, 1979) using a TransBlot Cell (BioRad Laboratories Ltd.). Six pieces of blotting paper, PVDF membrane wetted with methanol, support pads from the cell, and the gel itself were soaked in transfer buffer (3 l; 25 mM Tris, 190 mM glycine) to remove any air bubbles, and assembled in a 'sandwich' while submerged in buffer. The gel and membrane were placed on 3 pieces of filter paper on top of a support pad, and the remaining 3 pieces of filter paper were placed on top of this. The other support pad was then placed on the top and the 'sandwich' placed in the buffer tank and covered with transfer buffer. The proteins were transferred by electrophoresis, taking care that the membrane was on the correct (positive) side of the assembled cell, at a constant voltage of 30 V overnight or, for a minigel, 40 V, 1.5 hr.

After electrophoresis, the apparatus was dismantled and the membrane used for immunodetection. When required, the transferred protein was visualised by coomassie staining using solutions suitable for the PVDF membrane (stain: 50% v/v methanol, 10% v/v acetic acid, 0.1% w/v brilliant blue R; destain: 50% v/v methanol, 10% v/v acetic acid). The membrane was blocked, to prevent non-specific adsorption of the antibodies during immunodetection, with a solution made from non-fat dried milk (10% w/v in Tris-buffered saline; TBS) for 1 hr at 37°C. The block was removed, and primary antibody (usually sera from R707 at 1:1000 in block) incubated with agitation at 20°C. After an hour, the primary antibody mixture was removed, the membrane washed 4 times in TBS and the secondary antibody added. The secondary antibody was anti-rabbit IgG labelled with horseradish peroxidase, and was used at 1:5000 in block. After incubation at 20°C (1 hr) with agitation, the antibody mixture was removed and the membrane washed as before. The secondary antibody was detected

using ECL reagents (Amersham International plc.); the conjugated enzyme causes oxidation of luminol, emitting light that may be detected with X-ray film. Equal volumes of the two ECL solutions were mixed and incubated with the membrane for 1 min. The membrane was blotted briefly, wrapped in Saranwrap (Dow Chemical Corp.), placed in direct contact with Hyperfilm and developed between 10 sec and 20 min later.

Analysis of DNA and protein sequences

DNA and protein sequences were analysed using the University of Wisconsin Genetic Computer Group package 8.0 (UWGCG 8.0, 1994).

Isoelectric

The Isoelectric program calculates the overall charge of a protein at different pH values and generates a plot of the net charge as a function of pH. The net charge of a protein is calculated on the basis of protonated (positively-charged) residues (K, L, H), the protonated amino terminus, deprotonated (negatively-charged) residues (Y, C, E, D), and the deprotonated carboxyl terminus. The isoelectric point is the pH at which the protein has no overall charge.

Compare and DotPlot

The Compare program compares two sequences and produces a file that may be plotted using DotPlot (Maizel and Lenk, 1981). Two sequences are compared and a point added to the file when a given number of matches (stringency) are found within a given range of the sequence (window). This comparison is repeated such that all regions are compared between the two sequences, and DotPlot creates a graph with points of similarity plotted against their position within each sequence.

Gap

The Gap program compares two sequences and creates an alignment with the largest number of matches. A penalty is specified for both a gap introduced in the alignment (gap weight), and the length of the gap (length weight). The penalty is such that, for any gap, the specified number of extra matches must be made (Needleman and Winsch, 1970).

Chapter three: Ral activity *in vivo*

Introduction

Restriction of unmodified λ phages by *E. coli* K-12 may be overcome at high multiplicities of infection (moi). Heip *et al.* (1974) observed two phenomena, a multiplicity effect and a rescue effect. The multiplicity effect was observed when cells were infected with unmodified 'helper' phages (λ.0) and the 'test' phage (λ*vir*.0) was added after adsorption of the helper phage. It was concluded that the multiplicity effect was due to an inactivation of the restriction complex and this conclusion is now understood at a molecular level. The *Eco*KI ENase does not turn-over after DNA cutting (Meselson and Yuan, 1968) and the multiplicity effect results from a saturation of the hosts restriction potential. The alleviation of restriction reaches a maximum at a moi of 12, therefore, on average, 12 unmodified phages are able to inactivate the host-encoded *Eco*KI ENase. There are five *Eco*KI target sites in the DNA of the helper phage, and binding of *Eco*KI ENase to each of these sites presumably removes that molecule from the pool of the free ENase in the host. It has been calculated that if 12 phages, each with five sites, saturate the restriction ability of *E. coli* K-12, there are approximately (12 x 5 =) 60 molecules of *Eco*KI ENase in each cell (Kelleher and Raleigh, 1994).

The rescue effect observed by Heip *et al.* (1974) has a different basis, as alleviation of restriction by *Eco*KI was only observed when the helper phage was modified (λ.K). Furthermore, this effect was dependent on expression of phage genes as restriction by *Eco*KI was unaffected by the addition of chloramphenicol, an antibiotic that inhibits protein synthesis. The gene product responsible for the rescue effect was termed Ral, for restriction alleviation (Zabeau *et al.*, 1980). Zabeau *et al.* (1980) used different helper phages to map the *ral* gene to between *c*III and *N* in the phage λ genome. Mutagenesis and mapping of the mutants located *ral* between 70.6 and 70.9% on the physical map of λ (Debrouwere *et al.*, 1980a) and when this region was sequenced an open reading frame encoding a polypeptide of 66 aa was assigned to *ral* (Ineichen *et al.*, 1981). The nucleotide sequence of one of the *ral* mutants, *ral*18, isolated by Debrouwere *et al.* (1980a) has been determined, and it contains two missense mutations (Loenen and Murray, 1986). A transversion from G to T changes the predicted translational initiation codon from ATG to ATT. There is also a change from G to A at base 73 that would cause an aa substitution, Glu to Lys, at this position (note: there is a typographical error in Loenen and Murray (1986) that reads a change at base 72. The laboratory notes of W. Loenen, and my own nucleotide sequence

analysis, confirm that base 73 is changed). The *ral* gene and the mutant allele, *ral*18, have been cloned in an expression vector. Minicells transformed with the plasmids and an *in vitro* transcription-translation system have been used to detect the *ral* gene product. A polypeptide of suitable size may be detected with the cloned wild-type *ral*, but no corresponding protein is detected with the mutant, *ral*18 (Loenen and Murray, 1986).

The effect of Ral on restriction by a number of different R-M systems has been examined, including the type I systems *Eco*KI, *Eco*BI (IA), *Eco*AI (IB) and *Eco*R124I (IC), the type II systems *Eco*RI and *Eco*RII, and the type III system *Eco*PI (Zabeau *et al.*, 1980; Loenen and Murray, 1986). Ral only alleviates restriction by the type IA R-M systems. These screens are rather limited, but there is no evidence to suggest that Ral affects any systems other than the type IA. The type IA R-M systems are also the only systems to have a strong preference for hemimethylated DNA *in vivo*. After a single round of infection of either an $r_B^- m_B^+$ or $r_K^- m_K^+$ host by unmodified phages, the progeny of λ*ral*$^+$ phages are efficiently modified, whereas the progeny of λ*ral*$^-$ phages are only poorly modified (Zabeau *et al.*, 1980). Similar results have been obtained with the cloned *ral* (Loenen and Murray, 1986).

DNA methylation has been examined *in vitro*, under conditions where DNA cutting is blocked, for *Eco*KI, *Eco*BI (type IA; Vovis *et al.*, 1974), *Eco*AI (type IB; Suri and Bickle, 1985), *Eco*R124I and *Eco*R124II (Price *et al.*, 1987a). The type IA ENases are the only prokaryotic enzymes that have a strong preference for hemimethylated DNA *in vitro*, although the type IC ENases also show some differences in the rate of methylation of an unmodified substrate versus a hemimethylated substrate. The MTases of *Eco*KI (Suri *et al.*, 1984; Dryden *et al.*, 1993) and *Eco*R124I (Taylor *et al.*, 1993) show similar results. The conclusion of this work is that the type IB (e.g. *Eco*AI) R-M systems methylate an unmodified target as efficiently as a hemimethylated substrate, the type IA (e.g. *Eco*KI) systems have a strong preference for hemimethylated DNA and the type IC systems (e.g. *Eco*R124II) are somewhat intermediate between the type IA and IB systems. It seems relevant that the type IA are the only systems to have a strong preference for hemimethylated DNA and are also the only systems affected by Ral.

Unmodified *ral*$^+$ phages (λ*ral*$^+$.0) are restricted as efficiently as *ral*$^-$ phages (λ*ral*$^-$.0) presumably as Ral is not produced sufficiently quickly to protect its own DNA (Zabeau *et al.*, 1980; Toothman, 1981). It has been suggested that Ral activity is advantageous to the progeny of phages that escape restriction (Zabeau *et al.*, 1980; Loenen and Murray, 1986). The progeny of unmodified *ral*$^-$ phages (λ*ral*$^-$.0) that escape restriction

by a type IA R-M system will only be poorly modified, whereas the progeny of *ral*$^+$ phages (λ*ral*$^+$.0) will be efficiently methylated and protected from restriction by bacteria with an R-M system of the same specificity. This hypothesis has been investigated with strains proficient for restriction by *Eco*KI using unmodified *ral*$^+$ and *ral*$^-$ phages, and monitoring modification following a single round of infection.

Single round of infection with unmodified phages

The degree of modification of progeny phages was examined after a single round of infection of a restriction-proficient host by unmodified *ral*$^+$ and *ral*$^-$ phages (λ.0). Unmodified phages are restricted, therefore, only a small number of progeny phages were expected after a single round of infection. Unless unadsorbed phages were discounted from the subsequent measurements of the degree of phage modification, these residual, unmodified, phages could potentially mask any effect of Ral. Traditionally, antiserum to λ phage particles has been used to inactivate unadsorbed phages, but the only antisera available were ineffective, possibly due to a low titre of λ-specific antibodies. Therefore, a genetic method was used based on the observation that Ral does not affect *Eco*RI (Kelleher *et al.*, 1991; A. Titheradge, unpublished). The host cells were transformed with a plasmid, pJH16, that encodes the *Eco*RI MTase (Heitman *et al.*, 1989) and the test strains used to titre the progeny phages were transformed with a plasmid that encodes the *Eco*RI R-M system (Yoshimori *et al.*, 1972). Only phages that have been passaged through the host strains will be modified against *Eco*RI, and unadsorbed, hence unmodified, phages will be efficiently restricted by *Eco*RI encoded by the test strains. Two test strains were used, $r_K^+ r_{RI}^+$ and $r_K^- r_{RI}^+$, therefore, the degree of *Eco*KI-modification of progeny phages was examined for those phages that had been through an infective cycle on the $r_K^+m_K^+ r_{RI}^-m_{RI}^+$ host.

Preliminary experiments were done using three different strains of *E. coli* K-12, C600, NM710 and Ymel. These strains were transformed with pJH16, to make them $r_K^+m_K^+$ $r_{RI}^-m_{RI}^+$, and infected (100 µl, moi of 1) with unmodified λ*ral*$^+$ (λNM324.0) or λ*ral*$^-$ (λNM325.0) phages. After 10 min the infected cells were harvested by centrifugation (11 000 x *g*, 5 min), to remove some of the unadsorbed phages, and resuspended in pre-warmed L broth (10 ml) supplemented with chloramphenicol to select for hosts that maintain the plasmid, pJH16. The culture was incubated to allow a single round of infection (70 min, 37°C with aeration) and a sample removed (1.5 ml). Cells were lysed by the addition of chloroform, and the titre of progeny phages determined on two test strains. The test strains were transformed with a plasmid encoding *Eco*RI, to make them either $r_K^+ r_{RI}^+$ (C600 RI) or $r_K^- r_{RI}^+$ (5K RI). The degree of K-specific

modification of the progeny phages was estimated by comparing the titre of phages on C600 RI versus that on 5K RI (the efficiency of plating, eop). The eop of progeny phages varies depending on both the phage and strain used (table 3.1). After an sri of NM710 or Ymel, transformed with pJH16, the progeny of λ*ral*$^+$.0 phages are modified against restriction by *Eco*KI more than the those of λ*ral*$^-$.0 phages. However, little difference is detected in the modification of the progeny of λ*ral*$^+$.0 and λ*ral*$^-$.0 phages after an sri of C600. The results with NM710 and Ymel support the hypothesis that the progeny of an unmodified *ral*$^+$ phage are more likely to become modified against restriction by a type IA R-M system than a *ral*$^-$ phage. The results with C600, however, do not support this hypothesis as the progeny phages are well modified irrespective of the *ral* genotype of the phage.

Many strains of *E. coli* K-12, including C600, contain a defective prophage termed Rac (Kaiser and Murray, 1979), but Ymel and NM710 have a deletion removing this element (G. King and N.E. Murray, unpublished). The Rac prophage is known to encode a Ral-like antirestriction function, termed Lar (Toothman, 1981). In the prophage, *lar* would not be expressed, however, Rac will excise in a small proportion of the hosts allowing expression of *lar* and this may be responsible for the masking of the Ral effect in C600. The cells that contain the excised Rac prophage, and hence express *lar*, will be phenotypically $r_K^-m_K^+$ and efficiently methylate unmodified DNA. NM710 is a Rac$^-$ derivative of C600 and the results with these strains, and phages of different *ral* genotype, indicate an effect of both Ral and Lar. The difference in the modification of progeny phages after an sri of NM710 and Ymel is difficult to explain, the Ral effect appears to be more pronounced after an sri of Ymel than NM710, and there are probably factors influencing the modification of phages other than Ral or Lar.

To further examine the activities of Ral and Lar, experiments were done using strains differing only in the region of the Rac prophage, and phages differing only with respect to *ral*. The *E. coli* K-12 strains used were NM777 and NM778, made by P1 transduction using CAG12081 (Singer *et al.*, 1989) as the donor of *fnr*501Δ(*fnr-lar*) to transduce W3110 and selecting for the presence of a linked transposon *zcj-3061*::Tn*10*. NM777 and NM778 both contain the transposon, but only NM777 is Lar$^-$ (King and Murray, 1995; N.E. Murray, unpublished). These two isogenic strains were transformed with pJH16 and modification was monitored after a single round of infection as described above, but using two test phages which differ only with respect to *ral* (λ*c*I60 and λ*ral*18*c*I60; table 3.2). Phages modified to protect against restriction by both *Eco*KI and *Eco*RI were used to ensure that the plating efficiencies of the test strains were comparable (table 3.3). The results with these isogenic strains and phages

(table 3.2) are similar to those obtained before (table 3.1) and demonstrate a role for both Ral and Lar. In a Lar$^-$ host (NM777), the progeny of *ral*$^+$ phages are better modified than *ral*$^-$ phages. The effect of Ral is masked in a Lar$^+$ host (NM778). The effect of Ral after a single round of infection of NM777 (table 3.2) is only about a four fold increase in the modification of *ral*$^+$ progeny phages, whereas the Ral effect after an sri of NM710 and Ymel is about 18 fold and 140 fold respectively. Therefore, there are probably other factors influencing the modification of progeny phages other than Ral and Lar. The *E. coli* chromosome does contain defective prophages and elements other than the Rac prophage (Campbell, 1977) and it is conceivable that one of these encodes a Ral-like function. Alternatively there may be a host-encoded function that influences modification. Regardless of the strain-specific differences, this work supports the hypothesis that the progeny of phages that escape restriction by a type IA R-M system are more likely to become modified if the phages are *ral*$^+$ than *ral*$^-$. This enhanced modification would be advantageous, as the progeny of *ral*$^+$ phages would be protected from restriction in a subsequent round of infection of a host with a R-M system of the same specificity.

Host	Phage	Plaques per plate (no dilution)		eop
$r_K^+m_K^+$ $r_{RI}^-m_{RI}^+$		$r_K^+r_{RI}^+$ (C600 RI)	$r_K^-r_{RI}^+$ (5K RI)	
C600 pJH16	λ*ral*$^+$.0	133, 117, 145, 173	203, 211, 173, 189	7.3×10^{-1}
	λ*ral*$^-$.0	131, 149, 178, 160	248, 289, 210, 243	6.2×10^{-1}
NM710 pJH16	λ*ral*$^+$.0	49, 46, 46, 46	581, 625, 586, 510	8.1×10^{-2}
	λ*ral*$^-$.0	4, 1, 5, 2	616, 627, 704	4.6×10^{-3}
Ymel pJH16	λ*ral*$^+$.0	42, 32, 48, 30	192, 163, 190, 171	2.1×10^{-1}
	λ*ral*$^-$.0	0, 0, 0, 0	184, 153, 186, 149	$<1.5 \times 10^{-3}$

Figure 3.1. The modification of progeny phages after a single round of infection (sri) of an *Eco*KI restriction-proficient host. See text for details. The *ral*$^+$ phage was λ*bio*1*imm*^{434}*c*I (λNM324) and the *ral*$^-$ phage was λ*bio*232*imm*^{434}*c*I (λNM325). The *E. coli* K-12 strains Ymel and NM710 are *lar*$^-$, whereas C600 is *lar*$^+$. The plasmid pJH16 encodes the *Eco*RI MTase and the test strains, C600 and 5K, were transformed with a plasmid that encodes the *Eco*RI R-M system.

Host	Phage	Plaques counted (no dilution)		eop
$r_K^+m_K^+$ $r_{RI}^-m_{RI}^+$		$r_K^+r_{RI}^+$ (C600 RI)	$r_K^-r_{RI}^+$ (5K RI)	
NM777 pJH16	λ*ral*$^+$.0	685, 724, 690, 717	743, 705, 696, 766	9.7×10^{-1}
	λ*ral*$^-$.0	146, 157, 165, 136	663, 672, 642, 657	2.3×10^{-1}
NM778 pJH16	λ*ral*$^+$.0	730, 751, 737, 736	722, 754, 751, 738	1.0
	λ*ral*$^-$.0	594, 632, 616, 617	666, 667, 648, 654	9.3×10^{-1}

Table 3.2. The modification of progeny phages after a single round of infection (sri) of an *Eco*KI restriction proficient host. See text for details. The *ral*$^+$ phage was λ*c*I60 (λNM1094) and the *ral*$^-$ phage was λ*ral*18*c*I60 (λNM1090). The *E. coli* K-12 strains NM777 and NM778 are *lal*$^-$ and *lar*$^+$ respectively, but otherwise isogenic.

Phage	C600 RI ($r_K^+r_{RI}^+$)	5K RI ($r_K^-r_{RI}^+$)
λ*ral*$^+$.KRI	2.9×10^{10}	2.2×10^{10}
λ*ral*$^-$.KRI	2.0×10^{10}	1.5×10^{10}

Table 3.3. Titre of *ral*$^+$ (λNM1094) and *ral*$^-$ (λNM1090) phages, modified against *Eco*KI and *Eco*RI, on the test strains C600 RI ($r_K^+r_{RI}^+$) and 5K RI ($r_K^-r_{RI}^+$); a control to ensure that the plating efficiencies of the test strains were comparable.

Chapter four: The Ral protein

Overexpression of the *ral* gene of bacteriophage λ

The *ral* gene of phage λ has been cloned in two plasmid expression vectors by Loenen and Murray (1986). pUC*ral* is a derivative of pUC13 (Messing, 1983), that contains an 400 bp *Hae*III-*Bam*HI insert including *ral* (figure 4.1a), and p*ral*5 is a derivative of pCQV2 (Queen, 1983) that contains an 4.2 kb *Bam*HI fragment from λ*bio*214 (figure 4.1b). The *ral* gene product was visualised using an *in vitro* transcription-translation system and in minicells transformed with p*ral*5, however, the product is present only as a very small proportion of the total cellular protein (Loenen and Murray, 1986).

To increase the level of *ral* expression, alternative expression vectors were tried. The ~400 bp *Eco*RI-*Hin*dIII fragment from pUC*ral* was cloned in the plasmid vector pT7-6 (Tabor, 1990), to make pGK1 (figure 4.2), and the nucleotide sequence of *ral* in pGK1 was confirmed by double-strand sequencing using Sequenase. Transcription of *ral* in pGK1 is driven from the T7 promoter, and T7 RNA polymerase is encoded by a λDE3 prophage, but translational initiation is dependent on the ribosome binding site, and other DNA sequences, cloned with *ral*. Cell extracts were prepared from BL21(DE3) transformed with pGK1, proteins separated by SDS-PAGE and visualised with coomassie stain, however, no overexpression of *ral* could be detected.

Vectors that combine both a strong promoter and sequences required for translational initiation upstream of the cloned gene, have been successful in overexpression studies (Studier and Moffatt, 1986; Schauder *et al.*, 1987; Tabor, 1990). These vectors normally contain an *Nde*I target site to place the ATG start codon of the cloned gene in the correct position for maximal expression; the *Nde*I recognition sequence is 5' CATATG. An *Nde*I target site was introduced at the start of *ral* by using the PCR, with a primer containing an *Nde*I site overlapping the start codon (figure 4.3a). The template used in these reactions was pUC*ral* and the other primer annealed to the -40 region in the vector (figure 4.3b). The 270 bp fragment amplified by the PCR was partially sequenced, by direct sequencing using Sequenase, to confirm that the fragment contained *ral*, converted to a blunt-ended fragment with Klenow DNA polymerase and 5'-end phosphates added using polynucleotide kinase (see chapter 2). This 270 bp fragment including *ral* was ligated to pUC18 cut with *Sma*I, to generate pGK3 (figure 4.3c). The *ral* gene was then transferred to each of two vectors that provided a strong, well-controlled, promoter and an efficient translational initiation region. The 220 bp *Nde*I-*Eco*RI *ral* fragment was cloned in pJLA503 (Schauder *et al.*,

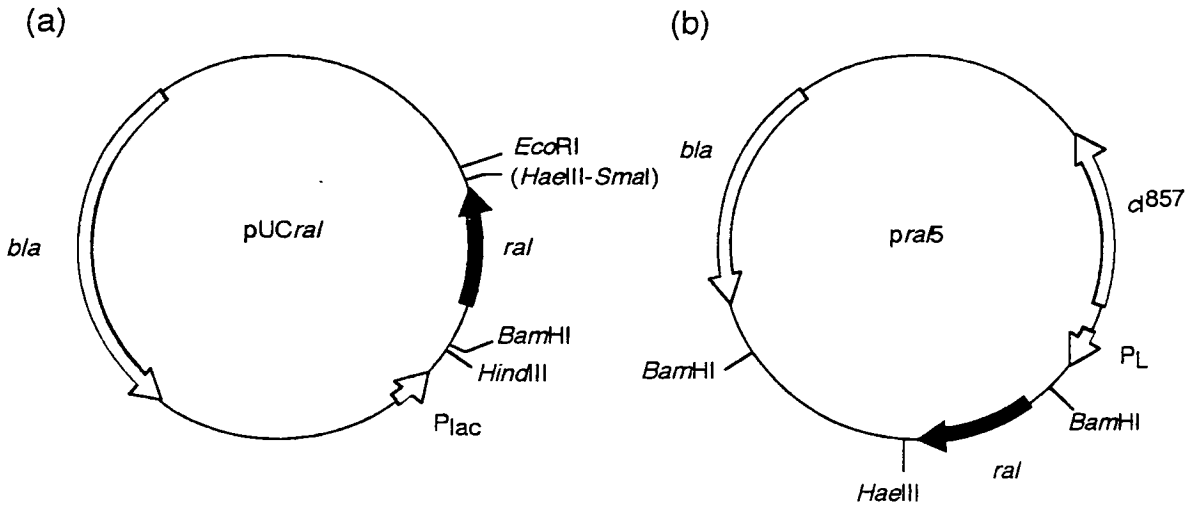


Figure 4.1. Plasmids constructed by Loenen and Murray (1986). (a) pUC*ral* (3.1 kb). (b) p*ral*5 (9.5 kb). See text for details. *bla*, gene encoding ampicillin resistance; P_{lac}, *lac* promoter; P_L, λ promoter; *c*I857, temperature sensitive allele encoding λ repressor; target sites in brackets, lost in cloning.

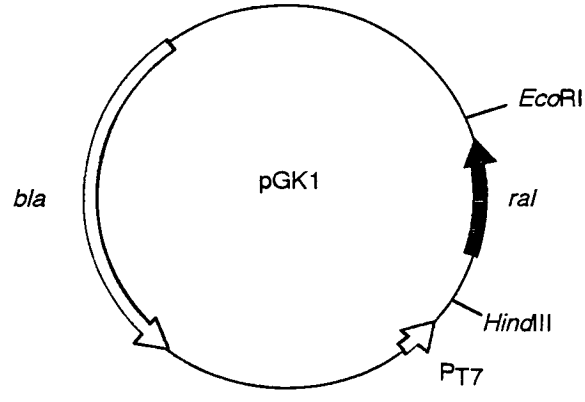
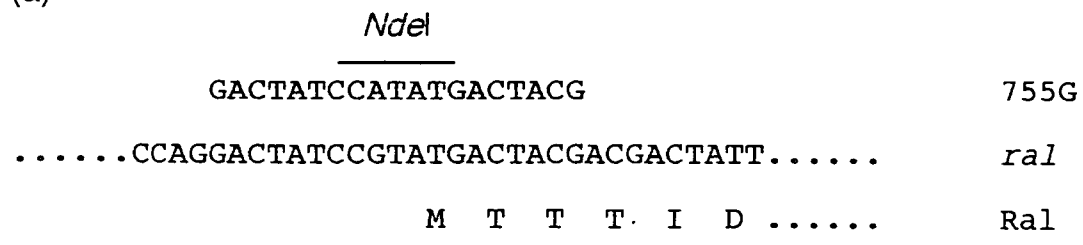


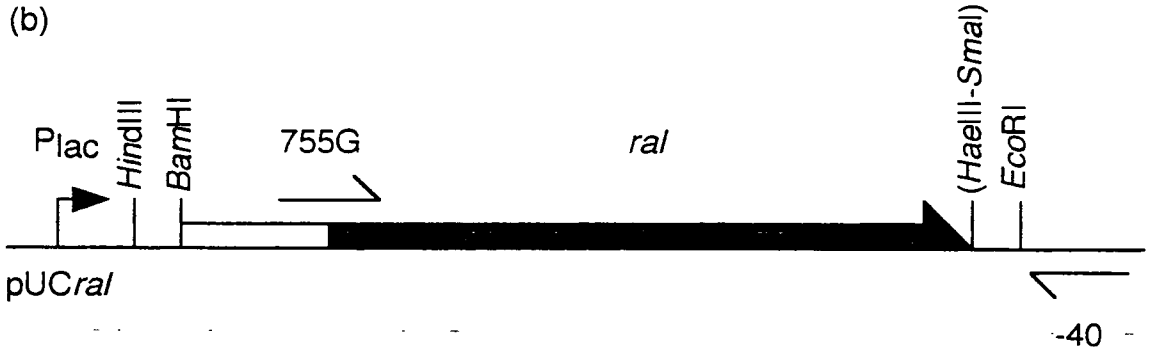
Figure 4.2. Plasmid pGK1. See text for details. *bla*, gene encoding ampicillin resistance; P_{T7}, T7 gene *10* promoter.

Figure 4.3. Generation of plasmid pGK3. (a) Nucleotide sequence at the start of the *ral* gene, the deduced aa sequence of its product, and the oligonucleotide, 755G, used to introduce an *Nde*I target site overlapping the translational initiation codon of *ral*. (b) The region of pUC*ral* amplified by the PCR using oligonucleotide primers 755G and a primer (642C) that anneals to the plasmid -40 region. (c) pGK3 (3 kb) containing the *ral* fragment amplified by the PCR. See text for further details. *bla*, gene encoding ampicillin resistance; P_{lac}, *lac* promoter; target sites in brackets, lost in cloning; -40, region to which the standard '-40' nucleotide sequencing primer anneals.

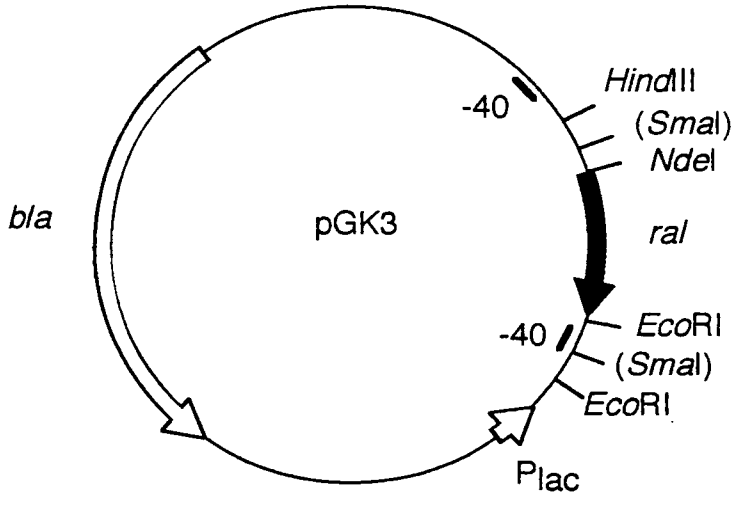
(a)


NdeI
GACTATCCATATGACTACG
755G
.....CCAGGACTATCCGTATGACTACGACGACTATT.....
ral
M T T T I D .....
Ral

(b)


Plac
HindIII
BamHI
755G
ral
(HaeIII-SmaI)
EcoRI
pUCral
-40

(c)


HindIII
(SmaI)
NdeI
-40
bla
pGK3
ral
EcoRI
-40
(SmaI)
EcoRI
Plac

1987) to generate pGK4 (figure 4.4a), and pT7-7 (Tabor, 1990) to generate pGK5 (figure 4.4b). pJLA503 contains the λ promoters P_L and P_R, upstream of the translational initiation region of *atpE*, and the λ gene cI^{857}. Transcription is repressed at 30°C by the product of cI^{857}, and induced by a temperature shift to 42°C (Schauder *et al.*, 1987). pT7-7 contains the promoter and upstream sequence of T7 gene *10* (Tabor, 1990). Transcription relies on T7 RNA polymerase, that is frequently provided by a λDE3 prophage (Studier and Moffatt, 1986). λDE3 was generated by cloning the genes *lacI*, and T7 gene *1* downstream of an IPTG-inducible *lac*UV5 promoter, in the *Bam*HI site of the λ vector D69 (Mizusawa and Ward, 1982; Studier and Moffatt, 1986). Insertion in the *Bam*HI site disrupts the λ *int* gene, that is required for both integration and excision, but lysogens may be made using a helper phage to provide the *int* gene product *in trans*. Therefore, stable lysogens may be made with λDE3 and transcription of T7 gene *1*, encoding T7 RNA polymerase, is induced by adding IPTG to the media. Overexpression of *ral* was examined as before, a product was detected in cell extracts of both BL21(DE3) pGK5, and NM522 pGK4, although the cells transformed with pGK5 appeared to give the higher level of expression. The complete nucleotide sequence of the insert in pGK5 was determined, by double-strand sequencing using Sequenase, and was identical to the published sequence (Ineichen *et al.*, 1981). To test whether the product of the cloned *ral* gene was active, a restriction-deficient modification-proficient host, NM526(DE3), was transformed with pGK5 and infected with an unmodified *ral*$^+$ phages. After a single round of infection, the degree of phage modification was estimated by measuring the efficiency of plating of progeny phages on a restricting host versus a non-restricting host. The progeny phages are efficiently modified after a single round of infection of the host transformed with pGK1, showing that the cloned *ral* gene encoded an active polypeptide (table 4.1).

Overexpression of *ral* and the solubility of its product were examined by growing BL21(DE3) pGK5 at 28, 37 or 42°C, and taking samples at various times after induction. The highest proportion of Ral in the soluble fraction appeared to be after three hr induction, with cells grown at 28°C (figure 4.5).

Vectors that produce a fusion polypeptide have been successful in increasing the level of expression of cloned genes, and in aiding the purification of the fusion proteins (for example, Smith and Johnson, 1988). The pET16b vector has a T7 promoter upstream of a translational initiation region and coding sequence for poly-histidine and a factor Xa proteolytic cleavage site. Immediately downstream of the coding sequence for the proteolytic cleavage site, is an *Nde*I site that can be used to clone the gene of interest.

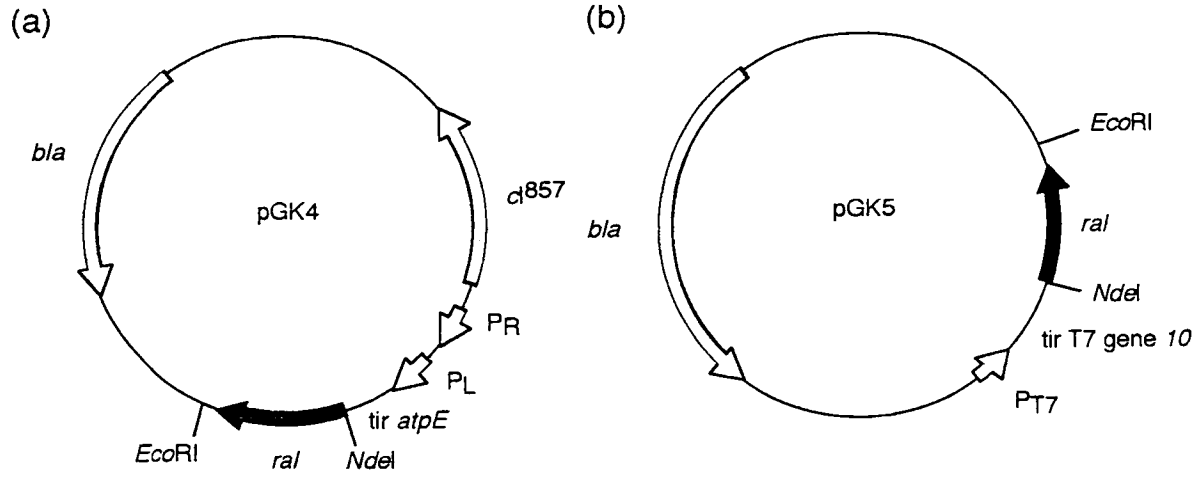


Figure 4.4. Plasmids that overexpress *ral*. (a) pGK4 (5.1 kb). (b) pGK5 (2.7 kb). *bla*, gene encoding ampicillin resistance; P_L and P_R, λ promoters; *c*I857, temperature sensitive allele encoding λ repressor; tir, translation initiation region.

Host (sri λ325.0)	eop
526(DE3) pT7-7	1.3×10^{-3}
526(DE3) pGK5	8.8×10^{-1}

Table 4.1. Modification enhancement by pGK5. A restriction-deficient, modification-proficient strain of *E. coli* K-12 lysogenic for λDE3 , NM526(DE3), was transformed with pT7-7 or pGK5 and infected with unmodified λ*ral*$^+$ phages (λNM325.0). After a single round of infection, the degree of phage modification was estimated by measuring the efficiency of plating (eop) of progeny phages on a restricting strain (C600) relative to that on a non-restricting strain (5K). See chapter 2 for details.

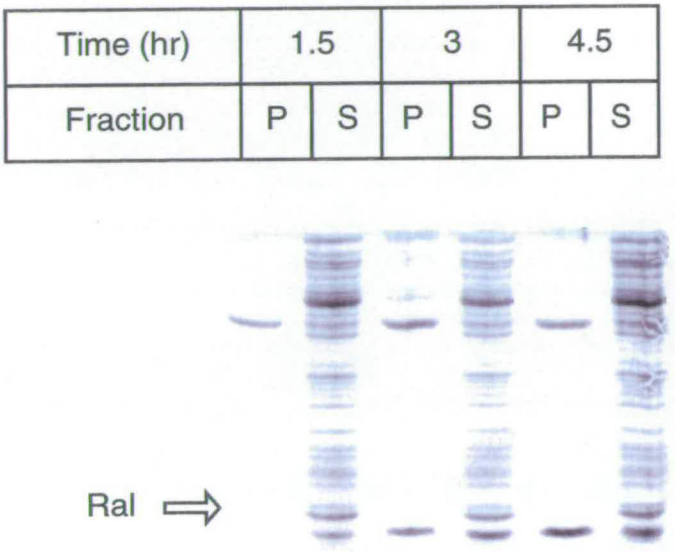


Figure 4.5. Overexpression of *ral*, and visualisation of its product. Cell extracts were prepared from BL21(DE3)pGK5, grown at 28°C for 1.5, 3, or 4.5 hours, and clarified by centrifugation (20 000 x *g*, for 20 min at 4°C). The insoluble (pellet) and soluble (supernatant) fractions were separated by Tris-tricine SDS-PAGE and visualised by coomassie staining of the gel, see chapter 2. P, pellet; S, supernatant; migration of Ral is indicated by the arrow.

The 220 bp *Nde*I-*Bam*HI fragment from pGK5 was cloned in pET16b, to generate pGK25 (figure 4.6a), however no overexpression of *ral* was detected and Ymel(DE3) transformed with pGK25 remains restriction-proficient (results not shown). The DNA of pGK25 gives the expected fragments when cut with a number of type II ENases, but the nucleotide sequence of the insert was not determined and there may be some subtle rearrangement or mutation. The fusion polypeptide contains an extra 21 aa residues and this may disrupt the activity of Ral.

One other plasmid was made that would allow the overexpression of both *ral* and the *hsdMS* genes of *Eco*KI in the same cell. The plasmid used to overexpress the *Eco*KI MTase, pPT7, is a derivative of pJF118 (Fürst *et al.*, 1986; Dryden *et al.*, 1993; P. Thorpe, unpublished) and contains the origin of replication from plasmid ColE1. pACYC184 contains the origin from plasmid p15A, and is compatible with ColE1-based plasmids (Chang and Cohen, 1978). The 560 bp *Bgl*II-*Bam*HI *ral* fragment from pGK5 was ligated to pACYC184 cut with *Bam*HI, to generate pGK24 (figure 4.6b). pGK24, therefore, contains the same arrangement of promoter, translational initiation region and *ral* gene as pGK5, but in a vector that will be maintained with pPT7. Modification of λ*ral*.0 enhanced by NM526(DE3) transformed with pGK24 and pPT7 (table 4.2, see comments in the legend), showing that an active Ral polypeptide is produced.

Production and purification of Ral Abs

The complete Ral protein has been made by solid-phase peptide synthesis, as part of a study on the synthesis and purification of polypeptides (Brown *et al.*, 1993). One of the problems of chemically synthesising a protein like Ral, is that a range of truncated peptides accumulate during the synthesis which makes it difficult to purify the full length product. Brown and colleagues (1993) have developed a method of attaching a large organic compound (Tbfmoc) to the N-terminus of the peptide, on the last cycle of synthesis, and this changes the elution time of the full length polypeptide by reverse-phase high performance (high pressure) liquid chromatography (HPLC). Thus, the complete peptide is easily purified and the Tbfmoc group cleaved to generate the full length protein. During the development of this system, a number of partially purified preparations of Ral were made that have been used to produce Ral anti-sera and to affinity purify specific antibodies (Abs).

Peptides are normally coupled to a carrier protein to enhance the immune response to the peptide, and synthetic Ral was coupled to keyhole limpet hemocyanin (KLH) using maleimidobenzoic acid-N-hydroxysuccinimide ester, that cross-links proteins through

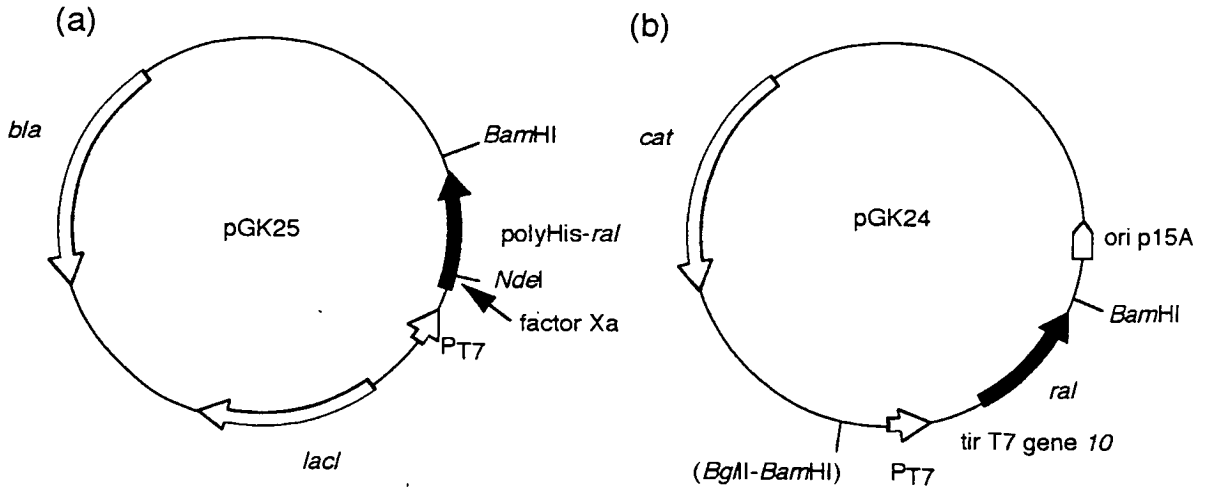


Figure 4.6. Other plasmids that overexpress *ral*. (a) pGK25 (5.9 kb). (b) pGK24 (4.5 kb). See text for details. *bla*, gene encoding ampicillin resistance; *cat*, gene encoding chloramphenicol resistance; P_{T7}, T7 gene *10* promoter; factor Xa, proteolytic cleavage site; tir, translation initiation region; ori p15A, origin of replication from plasmid p15A.

Host (sri λ325.0)	eop
526(DE3) pPT7 pACYC184	8.3×10^{-4}
526(DE3) pPT7 pGK24	3.5×10^{-1}
526(DE3) pJF118 pGK24	1.1

Table 4.2. Modification enhancement by pGK24. See text, section '*In vitro* methylation with cell extracts' for details. The degree of phage modification after a single round of infection (sri) of unmodified, *ral*$^+$, phage (λNM325) was estimated by measuring the efficiency of plating (eop) of progeny phages on a restricting strain (C600) relative to that on a non-restricting strain (NM679). Phages were titred by spotting 10 µl of diluted phage on a lawn of cells in BBL top layer, and the measurement of eop is not as accurate as the method described in chapter 2. However, this assay, based on sampling small volumes, is sufficiently accurate to indicate the enhancement of modification.

cysteine residues (Gnann *et al.*, 1989). Two rabbits (R707 and R708) were immunised with the Ral-KLH antigen, at days 0, 30, and 60. Test bleeds were taken from the pre-immunised rabbits and 10 days after each immunisation. The final test bleeds were used to detect Ral in Western blot analyses of the synthetic peptide and extracts from cells expressing *ral*. The third test bleeds were found to contain antibodies that recognised Ral, so the rabbits were sacrificed and sera collected (G. Leadbetter, Dept. Medical Microbiology). The Ral anti-sera and pre-immune sera contain a lot of other proteins, including other Abs that cross-react to *E. coli* proteins (figure 4.7). The Ral-specific Abs were purified by affinity chromatography (Harlow and Lane, 1988); a one ml Hi-Trap N-hydroxysuccinimide-activated Sepharose column was used to cross-link synthetic Ral to a solid support (see chapter 2). 0.1 ml of 1 M Tris.HCl pH 8.0 was added to 0.9 ml of Ral anti-sera (R707), and this was applied to the Ral-affinity column, pre-equilibrated with 100 mM Tris.HCl pH 8.0. The column was washed with 20 volumes 100 mM Tris.HCl pH 8.0, then 10 volumes 10 mM Tris.HCl pH 8.0. The Ral-specific Abs were eluted with 100 mM glycine pH 3.0, and 0.5 ml fractions were collected in tubes containing 50 µl 1 M Tris.HCl pH 8.0. One fraction was found to contain most of the eluted Ab, about 0.4 mg of protein estimated from the uv adsorption spectrum, and was stored in 50% v/v glycerol at -20°C.

The pre-immune sera, Ral anti-sera, the Abs that did not bind to the Ral-affinity column, and the affinity-purified Abs from R707, were compared by western blot analysis of cell extracts; extracts were prepared from induced cultures of BL21(DE3) transformed with pT7-7 or pGK5. Abs from the pre-immune sera cross-react to many *E. coli* proteins, but not to Ral, whereas the Ral anti-sera also recognised Ral (figure 4.7). The Abs that did not bind to the Ral-affinity column gave similar results to the pre-immune sera, but the Abs that bound to the column specifically recognised Ral and little else (figure 4.7). The other band seen with the Ral-specific Abs may be another protein that cross-reacts, or Ral that has not been fully reduced-denatured and is migrating anomalously by SDS-PAGE.

Purification of the *ral* gene product

Preliminary experiments were performed to identify which conditions and resins may be suitable to purify Ral. Chromatography columns were made from one ml syringes, blocked with glass wool and filled with a variety of different media. Depending on the media, different buffers and pHs were used:

- Diethylaminoethyl- (DEAE-) Sepharose with Tris-MES buffer, pH 8.0 and 6.5

Antibodies	Pre-immune sera		Ral anti-sera		Non-specific		Affinity purified	
pGK5	+	-	+	-	+	-	+	-
pT7-7	-	+	-	+	-	+	-	+

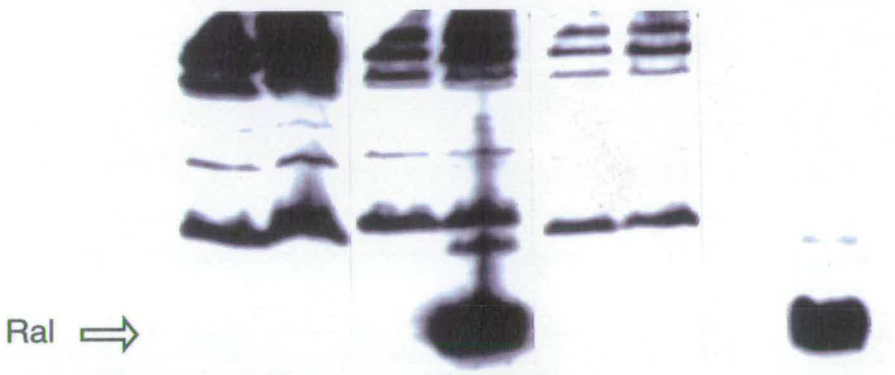


Figure 4.7. Cross-reactivity of antibodies to cell extracts. Cell extracts were prepared from BL21(DE3) transformed with pGK5 or pT7-7, separated by Tris-tricine SDS-PAGE, and analysed by Western Blot. See text for details. Arrow indicates migration of Ral.

- Carboxymethyl- (CM-) Sepharose with Tris-MES buffer, pH 8.0 and 6.5
- Heparin-agarose with Tris-MES buffer, pH 8.0
- Hydroxyapatite-BioGel with Piperazine buffer, pH 9.5

BL21(DE3) pGK5 or NM522 pGK4 were grown to an OD_{650} of 0.6, induced with 1 mM IPTG and harvested after three hr. For each column, 0.5 g of cells were resuspended in 1.5 ml of the relevant buffer, sonicated and the cell debris removed by centrifugation. The soluble cell extract was applied to the column, pre-equilibrated with buffer, and bound protein removed with buffer containing 1 M NaCl. Samples of the flow-through and bound fractions were separated by Tris-tricine SDS-PAGE, and analysed by Western blot with Ral antisera. Ral did not bind to heparin, hydroxyapatite, or DEAE, but some Ral did bind to CM at pH 6.5.

DEAE- and CM-Sepharose are anion and cation exchange resins respectively, and proteins bind by electrostatic interactions. The charges on the exchange media are balanced by counterions, e.g. Cl^- or Na^+, that are displaced by the protein, and usually the protein binds with an overall charge the same as the counterion it displaces. Anion-exchangers, like DEAE, are positively charged so a protein should bind if it has an overall negative charge, conversely cation-exchangers, like CM, are negatively charged. The isoelectric point of a protein is the pH at which the protein has no overall charge. Ion-exchange chromatography is generally performed one to 1.5 pH units above (anion-exchange) or below (cation-exchange) the isoelectric point of the protein, to ensure an overall negative or positive charge on the protein. The ion-exchanger remains ionised over a broad pH range, and bound proteins are eluted by increasing the ionic strength of the solution, thereby weakening the strength of the electrostatic interaction (Scopes, 1988). The isoelectric point of Ral was calculated to be 8.5 using the Isoelectric program of the UWGCG 8.0 package (figure 4.8).

The preliminary experiments with DEAE-Sepharose were probably performed at too low pH, and further attempts were made during the course of this work to bind Ral to an anion-exchanger. However, attempts at pH 9.5 (piperazine buffer, Econo-Pac Q column), pH 10.3 (diamino propane buffer), or pH 11.9 (phosphate buffer, Resource Q column) were unsuccessful, as approximately half of Ral in the sample only ever bound and did not give a clear elution profile. This may be due to the instability of Ral at these pHs or non-specific interactions with the media. Cation-exchange chromatography, however, has been extremely useful in the purification of Ral. Ral should be positively charged below pH 8.5, but it does not completely bind CM at pH 6.5, possibly because of local charge differences. However, unlike most *E. coli*

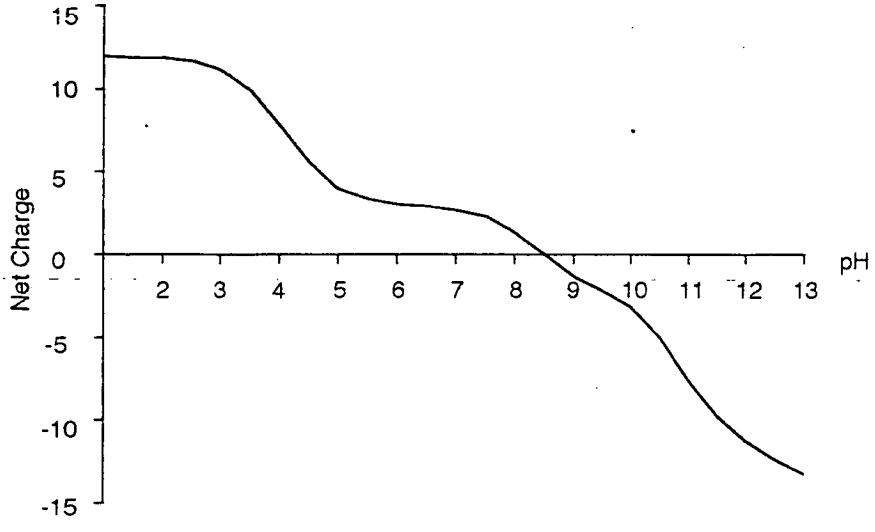


Figure 4.8. The predicted net charge of Ral plotted as a function of pH. The overall charge was calculated on the degree of protonation of positively-charged residues (K,R and H) and negatively-charged residues (Y, C, E and D) at a particular concentration of hydrogen ions (pH), using the Isoelectric program of the UWGCG 8.0 package.

proteins, Ral does bind to CM-Sepharose at pH 5.0, and this was used as the first step in the purification of Ral.

Ral bound to CM-Sepharose may be eluted using a salt gradient, which separates Ral from some other proteins, but the protein is then diluted. A simple method of concentrating protein is by ammonium sulphate precipitation, but for this to work the protein must already be at a concentration of approximately one mg ml^{-1}. Therefore, protein bound to the CM-Sepharose column was eluted in a step gradient of one M NaCl so that the protein concentration was suitable for Ral to be precipitated with $NH_4(SO_4)_2$. The salting out of proteins using high concentrations of salt is dependent on the surface charge distribution and polar interactions with the solvent. Water molecules are normally ordered around hydrophobic patches in proteins, and as the salt concentration increases, water molecules are attracted away from the protein to hydrate salt ions. Exposed hydrophobic regions will then cause aggregation of the proteins, and this precipitate may be harvested by centrifugation and dissolved in a small volume of low salt buffer.

Ral does not bind to heparin, hydroxyapatite, or a cation-exchanger so gel filtration was chosen as the next step in the purification of Ral. Gel filtration chromatography separates proteins according to size; smaller proteins are retarded in the molecular pore structure of the packed beads and elute later than large molecules which are retarded less and flow through the column faster. The separation achieved by gel filtration depends on loading a small sample volume, normally no more than 3% of the column volume (Scopes, 1988).

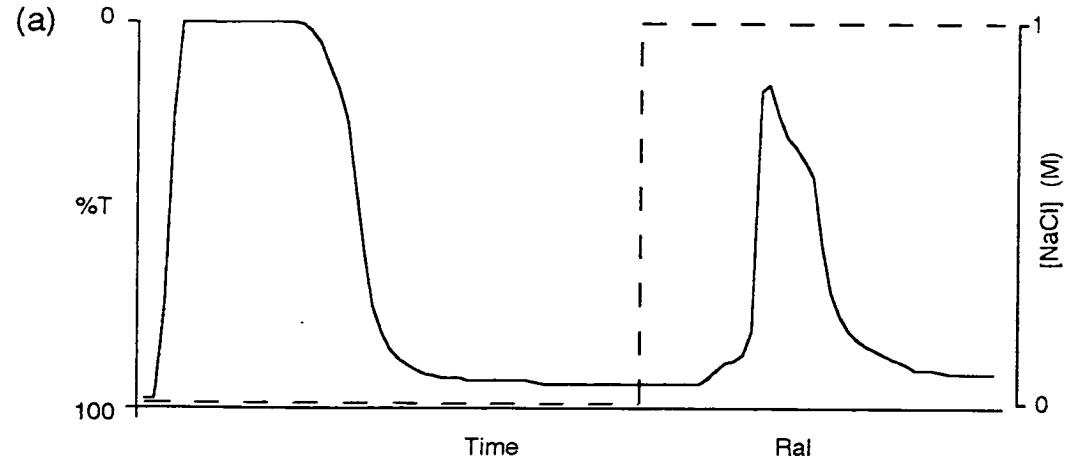
Two types of media have been used for gel filtration, that have different molecular mass separation ranges; Superdex 200, 10-300 kDa and Sephadex G-50, 1.5-30 kDa. On the basis of the working ranges, a column packed with Sephadex G-50 would be the better choice to purify Ral, however a Hi-Load (factory-packed) Superdex 200 column was available that provided better separation than the Sephadex column. Fractions containing Ral were collected from the Superdex 200 column and further purified using hydrophobic adsorbtion, reverse phase HPLC. HPLC uses columns of homogeneous small-beaded particles that can operate at high pressures. Hydrophobic adsorption chromatography, as its name suggests, relies on hydrophobic interactions between proteins and immobilised aliphatic chains, and reverse phase refers to the elution of bound macromolecules with organic solvents.

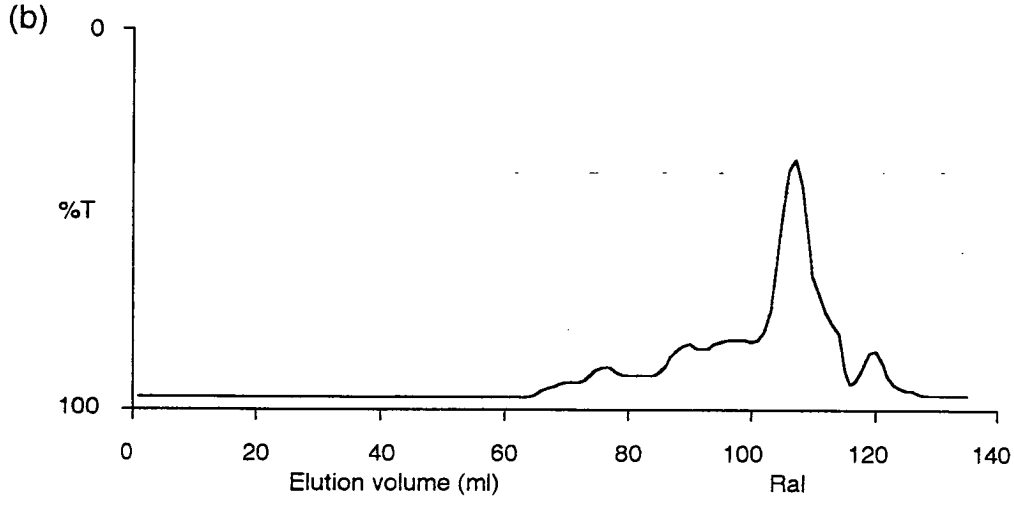
A generalised procedure for the overexpression of *ral* and the purification of its product is outlined below.

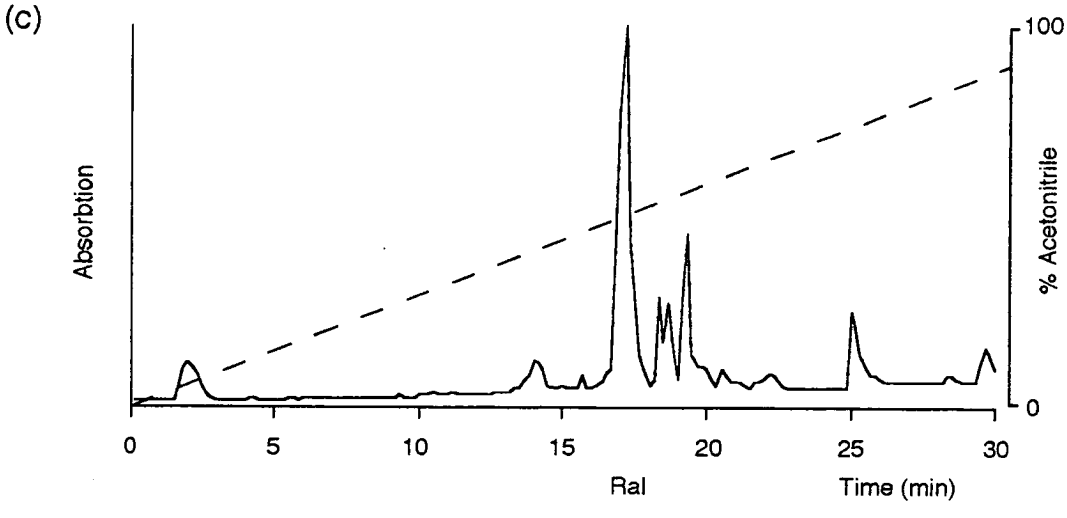
1. BL21(DE3) was transformed with pGK5 and a single colony grown for eight hr, at 37°C with good aeration, in an 20 ml bottle containing five ml of L broth supplemented with 100 μg ml^{-1} ampicillin.
2. Three ml of this culture were used to inoculate 100 ml L broth, supplemented with ampicillin, in a one l flask, and grown overnight at 37°C with good aeration.
3. 10, two l flasks containing 250 ml L broth, pre-warmed to 37°C and supplemented with ampicillin, were inoculated with eight ml of the overnight and grown at 28°C to an OD_{650} of 0.6.
4. Expression of *ral* was induced by adding IPTG to a final concentration of 0.6 mM.
5. Cells were harvested by centrifugation three hr after induction, and resuspended in 30 ml buffer (50 mM acetic acid pH 5.0) supplemented with 7mM β-mercaptoethanol, 20 μM benzamidine and 10 μM phenylmethylsulfonyl fluoride (benzamidine and PMSF are protease inhibitors).
6. The cell suspension was sonicated at 4°C and clarified by centrifugation at 37 000 x *g* for one hr.
7. The supernatant, i.e. soluble cell extract, was applied to a column packed with 40 ml CM-Sepharose, that had been pre-equilibrated with buffer.
8. The column was washed with buffer until no protein was detected in the flow through, and the bound protein eluted with a step gradient of one M NaCl, in buffer, at 60 ml hr^{-1} (figure 4.9).
9. Protein eluted from the CM-Sepharose column was precipitated by adding $NH_4(S0_4)_2$ to 80% saturation, and the precipitate was collected by centrifugation at 13 000 x *g* for 20 min.
10. The pellet was dissolved in 1.5 ml buffer, passed through a filter with a 0.2 μm pore size, and applied to a Hi-Load 16/60 Superdex 200 column pre-equilibrated with buffer.
11. Ral was eluted in buffer (figure 4.9b) and applied to a Vydac C_{18} (5 μm) column to further purify Ral by HPLC. The column was equilibrated in 0.1% v/v aqueous trifluoroacetic acid (TFA), 10% v/v acetonitrile and the protein was eluted with a gradient of 10 to 90% v/v acetonitrile in 0.1% v/v TFA over 30 min at a flow rate of 1 ml min^{-1} (figure 4.9c).
12. Fractions containing Ral were frozen in liquid nitrogen and dried under vacuum.

The most abundant protein in the sample applied to the C_{18} column elutes at the same place as synthetic Ral, and this fraction was further analysed by 'MALDI-TOF' mass spectroscopy (K. Shaw, Dept. Chemistry) and N-terminal peptide sequencing (A. Cronshaw, Dept. Biochemistry). The mass of the protein is 7464 kDa which is the

Figure 4.9. Elution profiles of Ral. (a) CM-Sepharose at pH 5.0, with a step gradient of 1 M NaCl. % Transmission (%T) at 280 nm is plotted as a function of salt concentration. (b) HiTrap Superdex 200 at pH 5.0. %T is plotted as a function of the elution volume. (c) Vydac C_{18} (HPLC) 10-90% v/v acetonitrile + 0.1% v/v TFA over 30 min. Absorbtion at 214 nm is measured as a function of time. The elution of Ral is indicated under the profiles. See text for further details.


(a)
0
%T
100
1
[NaCl] (M)
0
Time
RaI


(b)
0
%T
100
0
20
40
60
80
100
120
140
Elution volume (ml)
RaI


(c)
Absorbtion
100
% Acetonitrile
0
0
5
10
15
20
25
30
RaI
Time (min)

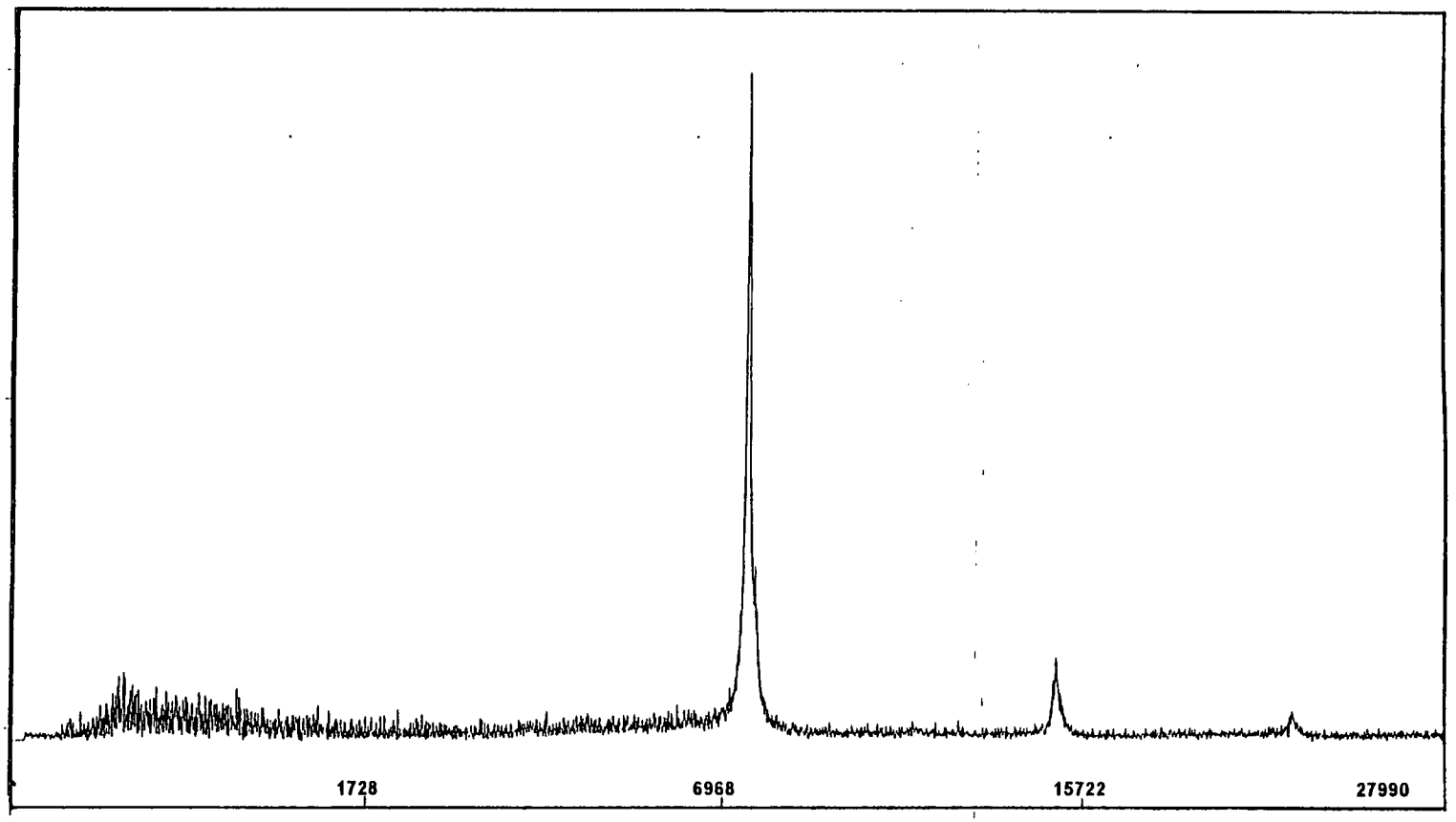
mass expected for Ral if the N-terminal methionine is missing, and a small peak is seen at 7605 kDa that corresponds to the mass of full-length Ral (figure 4.10). The loss of this methionine is a common postranslational modification in prokaryotes, and peptide sequencing confirmed that the N-terminal sequence started at the second residue of Ral (sequence obtained TTTID).

Analysis of Ral activity *in vitro*

The *Eco*KI MTase has been purified, and the protein efficiently modifies hemimethylated oligonucleotides containing the *Eco*KI target sequence, but only poorly modifies unmodified substrates *in vitro* (Dryden *et al.*, 1993). However, the binding affinities of the MTase for specific oligonucleotides are similar, irrespective of the methylation state of the DNA. The preference of the MTase for hemimethylated DNA, therefore, must be due to differences in catalysis and not binding affinity (Powell *et al.*, 1993). DNA footprinting of MTase:oligonucleotide complexes, by methylation interference, correlates differences in the protein:DNA interface with DNA of different methylation states. Without AdoMet, the methyl donor and cofactor, interference is detected for all substrates in the two halves of the recognition sequence and also at adjacent bases in the spacer region. The interference patterns do not change for hemimethylated or fully modified substrates when a complex is formed in the presence of AdoMet, but the footprints change with unmodified substrates; interference is restricted to bases in the recognition sequence (Powell and Murray, 1995). Methylation of the target sequence, therefore, changes the conformation of the protein:DNA interface and this conformational difference, in the presence of AdoMet, may be important in determining whether the DNA will be methylated or restricted (Powell and Murray, 1995).

Ral could both alleviate restriction and enhance modification of *Eco*KI if it altered the ENase so that an unmodified target was recognised as hemimethylated. Restriction would be alleviated, as hemimethylated targets are not a substrate for restriction, and modification would be enhanced, as *Eco*KI efficiently modifies hemimethylated DNA. If this is true, then the MTase should methylate an unmodified target as efficiently as a hemimethylated site in the presence of Ral, and the methylation interference footprint with Ral, MTase, AdoMet, and an unmodified substrate may look like the footprint of MTase with a modified substrate. Alternatively, Ral could enhance methylation to such an extent that restriction doesn't have time to occur or, if Ral blocked restriction, a small stimulation of methylation may explain the Ral phenotype. Modification is

Figure 4.10. 'MALDI-TOF' (matrix-assisted, laser desorbtion ionisation, time of flight) mass spectrum of Ral, with ionisation plotted as a function of mass. Main peak, 7464.5; shoulder, 7605.5; two other peaks at approximately 15 000 and 22 000 may be oligomers that have not been completely ionised.


1728
6968
15722
27990

enhanced by Ral even in the absence of HsdR, the subunit required for restriction, and consequently Ral activity was studied *in vitro* with purified Ral and the *Eco*KI MTase.

Purified *Eco*KI MTase was kindly donated by D. Dryden and L. Cooper (Dryden *et al.*, 1993). The elution of proteins with organic solvents, in reverse phase HPLC, can denature or inactivate proteins and, therefore, Ral activity was examined *in vitro* using Ral eluted from the Superdex 200 column. Typically, the total amount of protein eluted from the Superdex column, that contained Ral, was 4.3 mg (from a three l culture) at a concentration of 55 μM. The ratio of the absorbance at 280 nm to 250 nm was typically 2.3 (figure 4.11) and, from the HPLC elution profile, this sample contains greater than 90% Ral (figure 4.9c).

In vitro methylation

An *in vitro* methylation assay has been developed (Dryden *et al.*, 1993; Willcock *et al.*, 1994) based on the transfer of a ^{3}H-methyl group from radiolabelled AdoMet to a synthetic oligonucleotide substrate. Two complementary oligonucleotides were annealed to generate both an unmodified and hemimethylated *Eco*KI target site (408L and 409L and 860L:409L, respectively). The sequence of the oligonucleotides, and the position of the modified bases used to generate methylated targets, are included in figure 4.12 (Powell *et al.*, 1993; oligonucleotides were annealed by mixing equimolar quantities together, heating to 100°C for five minutes in a water bath and leaving to cool gradually to 25°C). The *Eco*KI MTase was applied to a PD-10 gel filtration column and eluted with 100 mM NaCl, 100 mM Tris.HCl pH 7.9 ('reaction buffer'), to remove glycerol and salt in the MTase storage buffer. The Ral storage buffer could not be changed in a similar way, as Ral is too small to be separated from the glycerol and salt using a PD-10 column, so Ral was diluted to the required concentration in reaction buffer. The components for the assay were typically 0.1 μM MTase, 0.5 μM Ral (if present), 15 mM β-mercaptoethanol, 200 μCi ml^{-1} [^{3}H]AdoMet (Amersham, specific activity approximately 75 Ci mmol^{-1}), and 2.25 μM annealed oligonucleotide (408L:409L or 860L:409L) in reaction buffer. AdoMet and DNA were mixed together in the reaction buffer and then added to the other components, as the order of addition can drastically affect the results of the assay (M. Winter and D. Dryden, unpublished). The assay mixture was incubated at 37°C, five μl samples removed at different time points and heated at 65°C for 10 minutes to stop the reaction and disrupt the protein:DNA complex. The oligonucleotides were separated from [^{3}H]AdoMet by PAGE (three μl of 20% v/v glycerol in reaction buffer was added to each tube, and loaded on a non-denaturing 12% w/v polyacrylamide gel), and labelled-DNA was

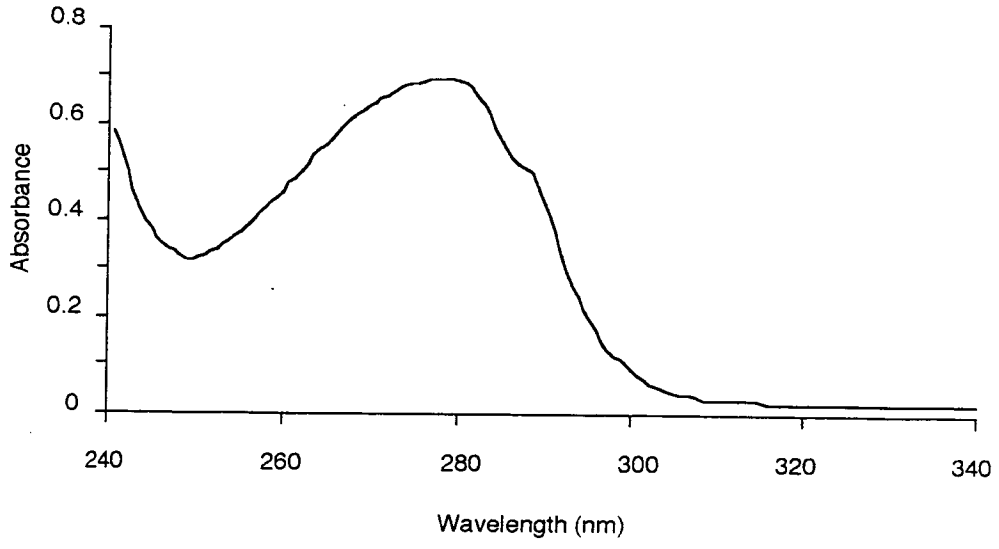


Figure 4.11. UV absorbtion spectrum of Ral between 240 nm and 340 nm.

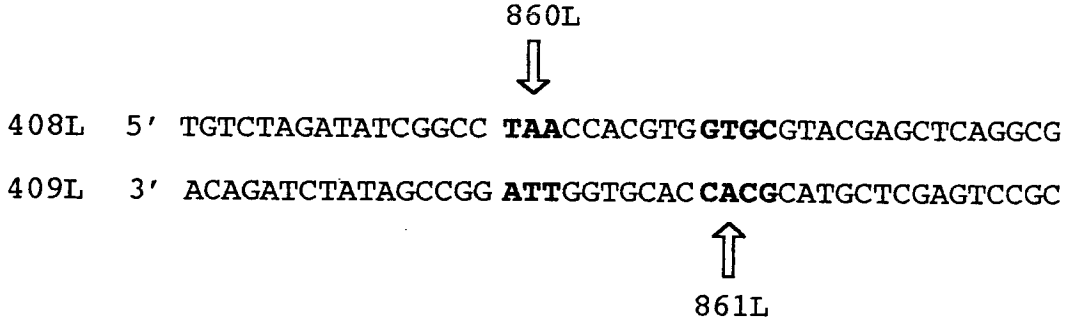


Figure 4.12. Oligonucleotides used to generate an *Eco*KI target site. Arrows, bases methylated by *Eco*KI and substituted with N^6-methyl adenine in 860L and 861L; bold, half-sites of the *Eco*KI recognition sequence.

visualised by fluorography (see chapter 2). No effect of Ral on methylation by *Eco*KI was detected (figure 4.13), although the assay is probably not sensitive enough to detect small differences in methylation activity.

Gel retardation

Gel retardation is used to show the interaction of proteins and DNA, and Powell *et al.* (1993) have used this method to calculate binding affinities of the *Eco*KI MTase for oligonucleotides of different methylation states. MTase bound to DNA causes a shift in the migration of the DNA by PAGE, that may be detected by autoradiography if the DNA is labelled with ^{32}P. If Ral interacts with the MTase a further shift may be expected, however, Ral is a small protein and the change in mobility of the complex probably would not be detected. Using Abs to Ral, it was hoped that a 'supershift' could be produced with a DNA:MTase:Ral:Ab complex. Gel retardations were performed essentially as described by Powell *et al.* (1993), with 50 nM MTase, 1 mM Ral and 0.1 nM unmodified ^{32}P-labelled DNA (annealed oligonucleotides, 408L:409L). The protein:DNA complexes were formed for 10 minutes at 4°C before 10 µl of the purified Ral Abs were added, and incubated for a further 15 min. Complexes were separated by PAGE and visualised by autoradiography (see chapter 2). A shift in the migration of the DNA was seen with MTase, but no extra shift was detected with Ral and the Abs (figure 4.14).

In vitro methylation with cell extracts

The purified Ral may be inactive, or some factor required for activity may be missing from the reaction mixtures, so methylation activity was examined in cell extracts of strains overexpressing the genes encoding the *Eco*KI MTase and Ral. BL21 is *hsdR*$^+$ and a ENase complex may be made if the cloned *hsdMS* genes are expressed, and this host is also *lar*$^+$ which may influence the methylation results (see chapter 3). Therefore, an *hsdR*Δ4*lar*$^-$ strain, NM526(DE3), was transformed with pairs of compatible plasmids, pPT7 and pACYC184, pPT7 and pGK24, or pJF118 and pGK24; pPT7 is the derivative of pJF118 that expresses *hsdMS* (Dryden *et al.*, 1993), and pGK24 is the derivative of pACYC184 that expresses *ral* (this chapter). Five ml L broth, supplemented with 100 µg ml^{-1} ampicillin and 40 µg ml^{-1} chloramphenicol, were inoculated with a single colony and grown at 37°C to an OD_{650} of 0.6. Transcription of the cloned genes was induced by adding IPTG to a final concentration of 0.6 mM, and the cells were harvested after two hr.

Figure 4.13. *In vitro* DNA methylation with purified proteins. *Eco*KI MTase activity was followed with unmodified (408L:409L) or hemimethylated (860L:409L) DNA substrate, over a range of times (1 min to overnight, o/n) with (+) or without (-) Ral. The labelled-DNA was separated from the [^{3}H]AdoMet by non-denaturing PAGE, and visualised by fluorography. See text for further details.

DNA	Unmodified										Hemimethylated									
Ral	+					-					+					-				
Time (min)	1	10	30	60	o/n	1	10	30	60	o/n	1	10	30	60	o/n	1	10	30	60	o/n

Figure 4.14. 'Supershift' gel retardation. See text for details. The migration of free DNA and DNA:MTase complex is indicated with arrows.

A sample of each culture was used to examine the Ral phenotype of the clones, 0.5 ml of culture was harvested by centrifugation and the cells resuspended in 100 µl 10 mM $MgSO_4$. Modification enhancement activity was examined after a single round of infection (see chapter 2) and the transformants gave the expected phenotypes (table 4.2, see comments in the legend). As the host is *hsdMS*$^+$, a strain containing the cloned *ral* gene always enhanced modification, but a host overexpressing *hsdMS*, in the absence of *ral*, does not have this phenotype.

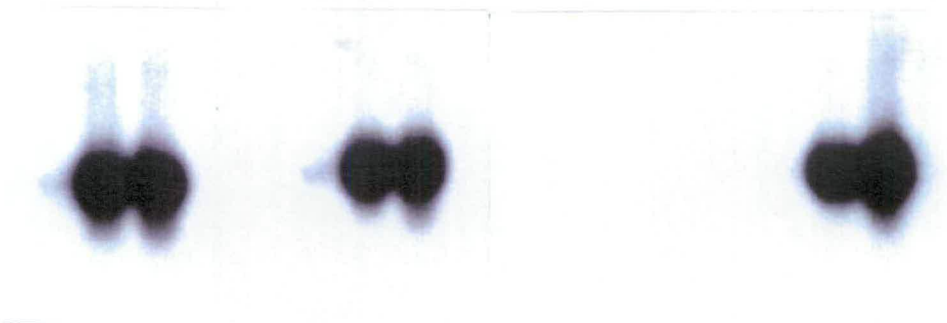
The remaining cultures were used to make cell extracts; pellets were resuspended in 200 µl 100 mM NaCl, 100 mM Tris.HCl pH 7.9, 7 mM β-mercaptoethanol, sonicated at 4°C and clarified by centrifugation at 20 000 x *g* for 30 min at 4°C. Methylation assays were performed as described previously, but using 40% (v/v) cell extract instead of purified protein. An additional DNA substrate was used that had a completely modified *Eco*KI target site (oligonucleotide pair, 860L:861L; figure 4.12) to confirm that the methylation of the oligonucleotides observed was due to *Eco*KI MTase activity. The hemimethylated targets were the only substrates to be efficiently methylated. Comparisons cannot be made between different cell extracts, as the concentration of MTase will vary, but Ral did not appear to have any effect on MTase activity (figure 4.15).

Discussion

The *ral* gene of bacteriophage λ has been overexpressed and its product purified, using a combination of ion exchange and gel filtration liquid chromatography, and reverse phase HPLC. The purified protein has the molecular mass and N-terminal sequence expected from the nucleotide sequence of *ral* (Ineichen *et al.*, 1981) without the N-terminal methionine. The first aa residue, incorporated at the initiation of translation, is always N-formyl-methionine, although the first codon is not always an ATG codon for methionine, and this residue is commonly removed post-translationally in *E. coli* (Hershey, 1987). Most of the Ral produced by cells overexpressing the gene at 37°C, is insoluble. Attempts to increase the solubility of Ral, by growing cells at lower temperatures, have been partially successful; more Ral appears in the soluble cell extract from cultures grown at 28°C than those grown at higher temperatures. Incubation temperatures lower than 28°C were not tried as no equipment was readily available that could aerate cultures in flasks below this temperature. Expression of *ral* at lower temperatures, however, may further increase the solubility of its product. The production of fusion polypeptides can increase the solubility, and aid the purification, of expressed gene products and further investigations with pGK25 or similar plasmids

Figure 4.15. *In vitro* DNA methylation with cell extracts. See text for details. MTase + Ral, protein from cells overexpressing both the genes encoding the MTase and Ral; MTase + Ral (mix), protein from cells overexpressing the genes encoding the MTase and cells expressing *ral* were mixed together after the extracts were made. =, unmodified target (408L:409L); ±, hemimethylated target (860L:409L); ‡, fully modified target (860L:409L).

Proteins	MTase					MTase + Ral					Ral					MTase + Ral (mix)				
DNA	- -		+ -		+ +	- -		+ -		+ +	- -		+ -		+ +	- -		+ -		+ +
Time (min)	1	5	1	5	5	1	5	1	5	5	1	5	1	5	5	1	5	1	5	5



may increase the yield of Ral. The present purification procedure provides ample protein for biochemical studies, but a higher yield would be useful for structural studies. In particular, Ral is sufficiently small to be amenable to analysis by nuclear magnetic resonance (NMR), although approximately 20 mg of protein would be required to attempt a full structural determination. These quantities of synthetic Ral are available, but attempts to solve the structure by 2D NMR have been unsuccessful (P. Barlow and A. Brown, unpublished). The Ral purified from cell extracts may be useful for comparisons with the synthetic Ral, to determine whether they have the same biochemical and structural properties. Larger quantities of Ral could be purified if the insoluble protein was denatured before purification, e.g. with urea. However, the denatured protein would have to be renatured, and it would be necessary to determine if the renatured protein had refolded correctly, although similar studies could be performed to those envisaged for the synthetic Ral.

The main reason for purifying Ral was to examine its molecular mechanism of action and its interaction with the *Eco*KI MTase. Large quantities of the *Eco*KI MTase have been purified (Dryden *et al.*, 1993) and a variety of biochemical techniques have been used to examine its activity (Powell *et al.*, 1993; Powell and Murray, 1995; Chen *et al.*, 1995). The *Eco*KI MTase has a strong preference for hemimethylated DNA *in vivo* and *in vitro* (Zabeau *et al.*, 1980; Suri *et al.*, 1984a; Dryden *et al.*, 1993) and may be regarded as a maintenance MTase, as it maintains the methylation patterns of the host after semi-conservative DNA replication. Mutants have been isolated and mapped to *hsdM* that no longer have this preference for hemimethylated DNA and efficiently methylate an unmodified substrate, consistent with *de novo* methylation activity (Kelleher *et al.*, 1991). An hypothesis to explain the Ral phenotype, is that Ral interacts directly with the MTase to stimulate MTase activity (Loenen and Murray, 1986), and the maintenance MTase of *Eco*KI could be changed to a *de novo* MTase.

Preliminary experiments with gel retardations using the *Eco*KI MTase, Ral, labelled oligonucleotides and affinity purified Ral Abs were unable to show an interaction between Ral and the MTase associated with its target. This technique relies on Ab binding to cause a 'supershift', and the extra shift will not occur if the epitopes, recognised by the Ab, are masked. Other methods have been used to try and show an interaction between Ral and the MTase, including PAGE of glutaraldehyde-crosslinked complexes and gel retardation followed by Western blot analysis, however, the results were inconclusive and more work would be required to refine the techniques before any conclusions could be drawn. Another method that may be useful is liquid chromatography, either gel filtration or affinity chromatography with immobilised

protein that would be similar to the method used for the purification of the Ral Abs. Gel retardation and chromatography techniques rely on a stable association during PAGE or on the column, whereas crosslinking of the MTase and Ral merely requires a close association in solution. Therefore, glutaraldehyde-crosslinking and detection of Ral by Western blot analysis, may be the most suitable technique to further analyse possible interactions between Ral and the MTase. Purified HsdM subunit is also available (M. Winter and D. Dryden, unpublished) that could be used to further define possible interactions.

Methylation assays, with purified MTase and synthetic oligonucleotide substrates, did not show any increase in *de novo* methylation when purified Ral was present. The *in vivo* observations clearly demonstrate that the modification of unmodified phage DNA is enhanced in the presence of Ral, and the lack of increased methylation *in vitro* could be due to a number of factors:

1. A component is missing.
2. Subunit rearrangements are required that do not take place *in vitro.*
3. The purified Ral is inactive, or the purified MTase has lost the activity required to interact with Ral.
4. Oligonucleotides are not a suitable substrate.
5. Impurities in the Ral preparation interfere with Ral.
6. The enhanced methylation is not sufficient to be detected in this assay.

Methylation assays with extracts from cells expressing the genes encoding the *Eco*KI MTase and Ral were designed to overcome problems associated with the first three points. Immediately before the extracts were made, bacteria were able to enhance the methylation of unmodified *ral*⁻ phages, however, enhanced methylation of unmodified DNA was not detected *in vitro*. Hemimethylated oligonucleotides are substrates for the *Eco*KI MTase, but in case the substrate did make a difference, attempts were made to examine the methylation of plasmids with cell extracts. Hemimethylated substrates could be made by annealing two linearised plasmids isolated from strains with different modification phenotypes, however the plasmids were not labelled sufficiently to be visualised by fluorography (results not shown). The concentration of *Eco*KI targets used in the MTase assay with oligonucleotides was 2.25 μM, and to achieve this concentration of sites with the plasmid (pRH3; Sain and Murray, 1980) would require approximately 4.5 μg μl^{-1} DNA in the reaction (the molecular mass of pRH3 is approximately 4 x 10^6, and it contains 2 *Eco*KI sites). The concentration of plasmid DNA actually used was 0.2 μg μl^{-1} and this could be increased, although technical

considerations would soon limit the amount of DNA in the reactions. If these experiments were continued it may be preferable to use a different method of separating the labelled DNA from AdoMet, e.g. gel filtration, and then detect the tritiated DNA in scintillant using a scintillation counter.

The purified Ral, and cell extracts, used for these experiments do contain other proteins and it is possible that one of these interferes with Ral activity *in vitro*, but not *in vivo*. The Ral protein could be further purified by liquid chromatography using DEAE, Heparin and Hydroxyapatite; Ral does not bind to these resins, but other proteins probably will, and Ral would be collected in the flow-through. HPLC using a cation-exchange column and elution with an aqueous salt gradient may also be useful, although attempts to separate Ral using a Resource S column have been unsuccessful as Ral elutes over a large salt concentration. The measure of methylation enhancement *in vivo* is much more sensitive than these *in vitro* experiments, and more extensive studies, possibly using techniques of greater sensitivity should indicate that Ral does have a stimulatory effect on methylation *in vitro*.

If Ral only slightly stimulates methylation, then it must also block DNA cutting. Small quantities of the purified *Eco*KI ENase are available (L. Cooper and D. Dryden, unpublished), but preliminary experiments indicate that purified Ral added to the ENase does not stop the cutting of unmodified plasmids *in vitro*. However, for similar reasons outlined for the methylation assays with purified proteins, an examination of DNA cutting with extracts from cells overexpressing the genes encoding the *Eco*KI ENase and Ral may be more informative. Other host-encoded exonucleases and endonucleases would be a problem in these experiments, although the circular plasmid DNA would be resistant to degradation, and non-specific cleavage could be detected with appropriate controls e.g. plasmid DNA prepared from a host expressing the genes encoding *Eco*KI.

Chapter five: The Lar activity of bacteriophage λ*reverse*

Introduction

The Ral function of bacteriophage λ is able to modulate the activity of the type IA R-M systems, both alleviating restriction and enhancing modification (Zabeau *et al.*, 1980). A number of lambdoid phages are known to encode a Ral-like function, including P22, ϕ21, and the recombinant phage λ*reverse* (λ*rev*; Toothman, 1981; Semerjian *et al.*, 1989). The nucleotide sequences of the *ral* genes of phage λ and ϕ21 are identical (Franklin, 1984). The nucleotide sequence of ϕ21 in the region including *ral* is from an λimm^{21} hybrid phage, and the sequence identified as the *ral* gene of ϕ21, may be sequence from λ. The *ral* genes of phages λ and P22 also share a high degree of sequence identity, approximately 75% at the nucleotide and predicted aa level (Semerjian *et al.*, 1989). As the predicted aa sequences of the three identified *ral* genes are so similar, little information may be gleaned about individual residues that may be important for structure and/or function. The gene encoding the Ral-like activity of λ*rev* has been termed *lar* (Toothman, 1981). This chapter describes the identification and characterisation of *lar* and its product.

To understand the nature of λ*rev*, and *lar*, it is necessary to explain a little about the development of phage λ and how this recombinant phage was isolated. Phage λ DNA enters the cell as a linear molecule with complementary, single-stranded ends that are circularised by the activity of host-encoded DNA ligase to form a double-stranded *cos* site. DNA replication initiates on the supercoiled circular DNA molecule and proceeds bi-directionally to create monomeric daughter DNA circles (θ replication). Packaging of λ DNA requires two *cos* sites, therefore, monomeric circles are not packaged, although homologous recombination mediated by the host-encoded RecA function or phage-encoded Red function, can form multimeric molecules that may be used to package DNA. The DNA normally forms concatemers by rolling-circle (σ) replication and is packaged after cutting at two suitably-spaced *cos* sites. For the phage to replicate by σ replication, the RecBCD nuclease must be inactivated by the phage-encoded Gam protein, presumably because RecBCD would degrade phage DNA from the unprotected single-stranded *cos* site (figure 5.1; Furth and Wickner, 1983; Murray, 1991).

The Fec (feckless) and Spi (sensitive to P2 interference) phenotypes are important in the isolation of λ*rev*, and some understanding of these phenotypes at a molecular level

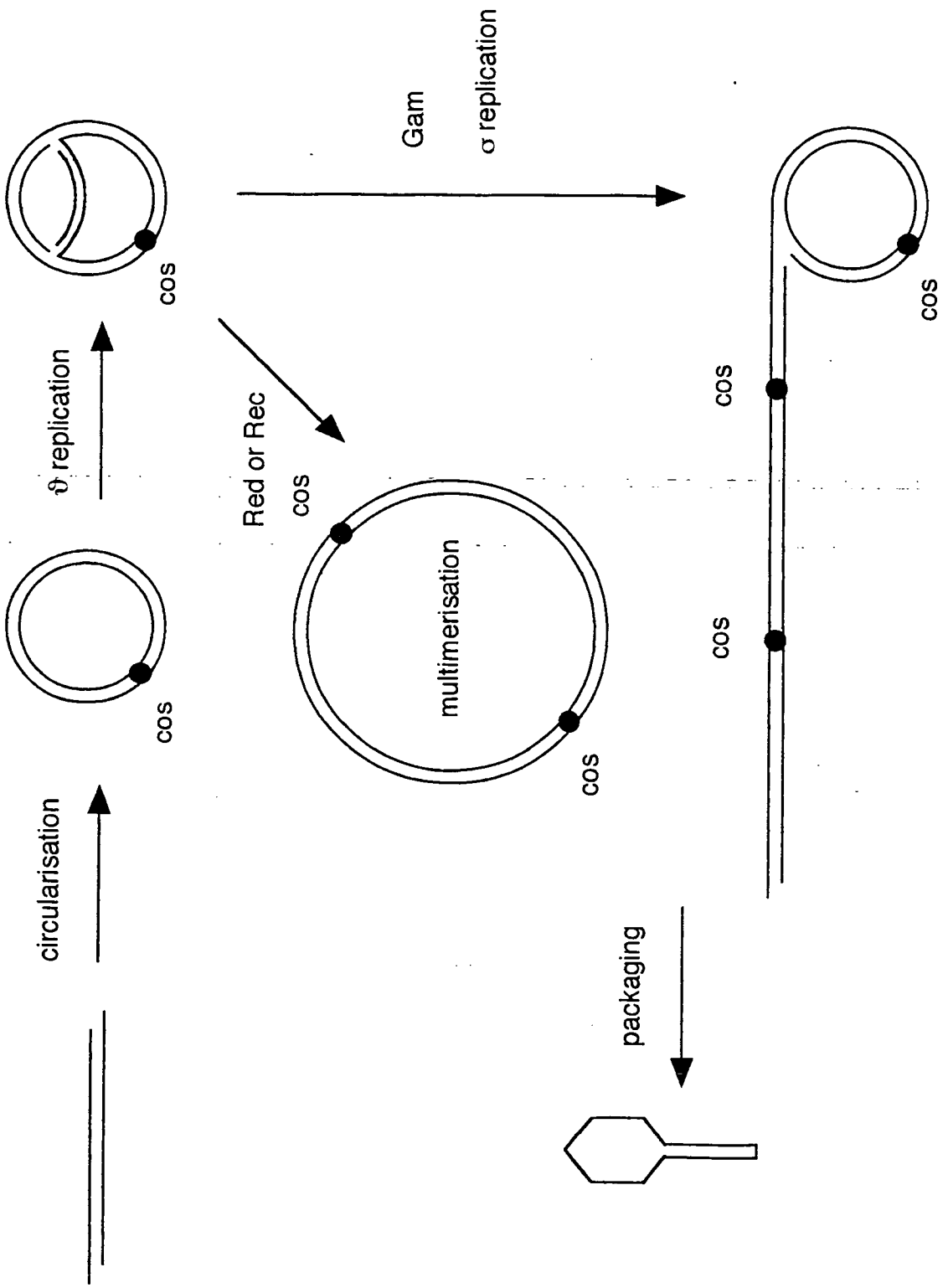


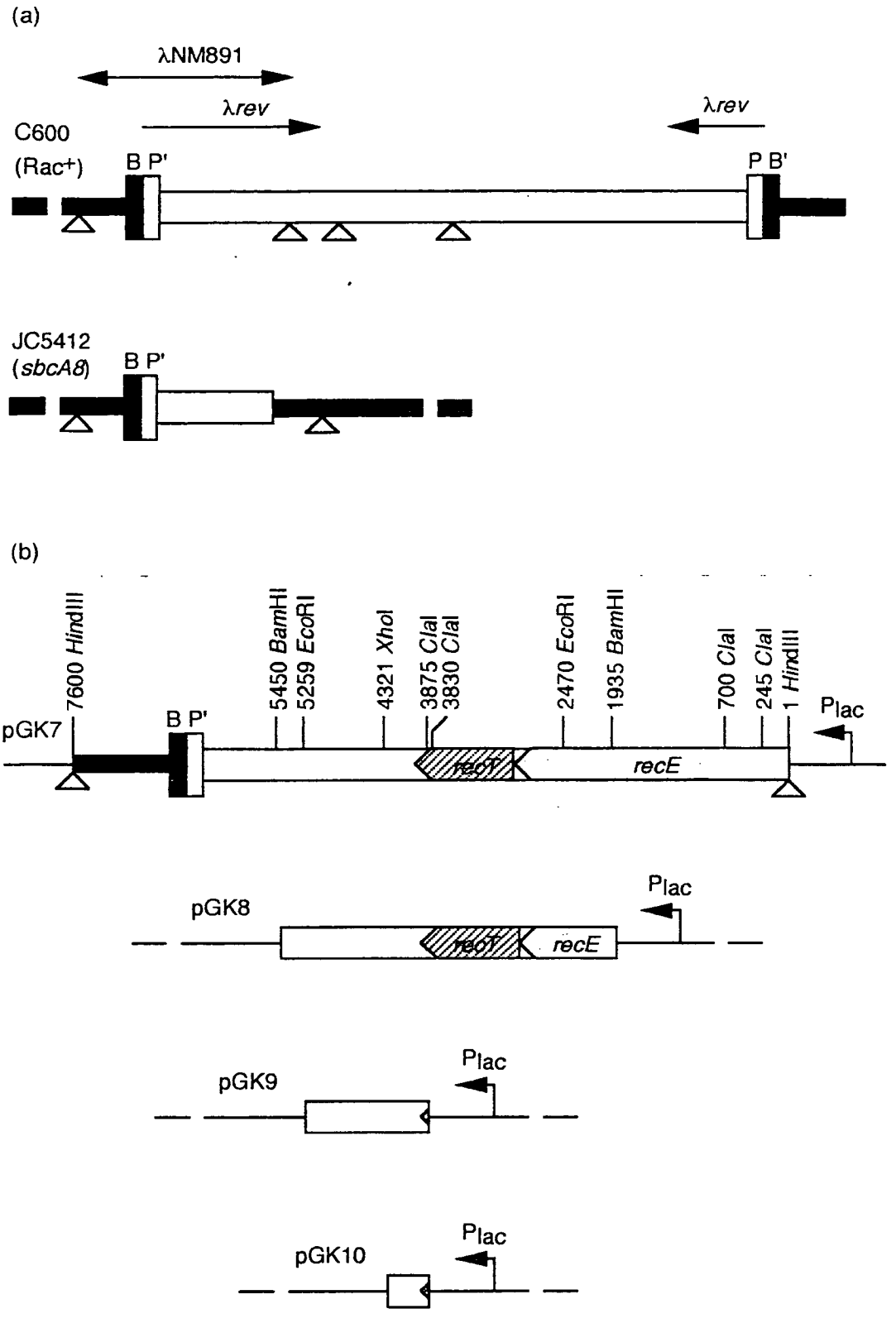
Figure 5.1. The lytic pathway of bacteriophage λ. See text for details

may be appreciated with the information outlined above. λ*red*$^-$*gam*$^-$ phages are unable to grow on a *recA* host, the Fec$^-$ phenotype (Manly *et al.*, 1969), as concatemeric DNA molecules cannot be formed. Wild-type λ is unable to grow on a P2 lysogen, the Spi$^+$ phenotype (Lederberg, 1957), and this is only partially understood. λ*red*$^-$*gam*$^-$ phages are Spi$^-$ and phage P2 cannot lysogenise a *recB* mutant (Sironi, 1969), suggesting that the inactivation of RecBCD nuclease by Gam is responsible for the Spi$^+$ phenotype. λ*red*$^+$*gam*$^-$ phages are also Spi$^+$ and the Fec$^+$/Spi$^+$ phenotypes, therefore, would appear to be mutually exclusive. However, Zissler and colleagues (1971) isolated *reverse* mutants of Fec$^-$ phages that were Fec$^+$, but Spi$^-$; these phages, termed λ*reverse*, could both grow on a *recA* host and a P2 lysogen. Heteroduplex mapping and genetic analysis has identified a region in the genome of λ*rev* that is substituted with DNA from *E. coli* that serves the same functions encoded by the early genes of phage λ, including integration, recombination, immunity and replication (Low, 1973; Gottesman *et al.*, 1974; Diaz *et al.*, 1979). Southern blot analysis of λ*rev* and different strains of *E. coli* (Kaiser and Murray, 1979), confirmed the hypothesis that the region of host DNA substituted in λ*rev* was derived from the *E. coli* Rac prophage (Low, 1973; Gottesman *et al.*, 1974). *E. coli sbcA* mutants (Templin *et al.*, 1972) express the *recE* and *recT* genes of the Rac prophage (Kushner *et al.*, 1974; Clark *et al.*, 1993), and the exonuclease activity encoded by λ*rev* and Rac is indistinguishable (Gillen *et al.*, 1977). The Red activity of bacteriophage λ is encoded by two genes, *redα* (*exo*) and *redβ* (*bet*). RecE exhibits a double-strand DNA exonuclease activity similar to Redα (Kushner *et al.*, 1974), and RecT, like Redβ, catalyses the renaturation of single-stranded DNA (Kolodner *et al.*, 1994). *E coli sbcA* mutants are unable to restrict the growth of unmodified phages, although they are genotypically *hsdRMS*$^+$ (Simmon and Lederberg, 1972), and it seems probable that λ*rev* and *sbcA* hosts express the same gene, termed *lar*, whose product is responsible for the alleviation of restriction (Toothman, 1981).

Identification of the *lar* gene

To identify the region of the Rac prophage that encodes *lar*, a 7.6 kb *Hin*dIII fragment from λNM891 (Kaiser and Murray, 1979), that includes DNA common to the *E. coli sbcA8* mutant and λ*rev*, was cloned in pUC18 (Yanisch-Perron *et al.*, 1985) to generate pGK7 (figure 5.2a). An *Eco*KI restriction-proficient strain (NM710) was transformed with pGK7 and the restriction alleviation activity of the plasmid was examined by comparing the titre of λv.0 on this strain versus a non-restricting strain (NM792). In this screen, a similar titre indicates restriction alleviation. pGK7 alleviated restriction, and the *lar* gene was further localised by cloning segments of the

Figure 5.2. Localisation of the region encoding *lar*. (a) Origin of λ*rev*, λNM891 and the Rac prophage integrated in the chromosome of the *E. coli* strain C600 and an *sbcA* mutant (*sbcA*8; JC5412). The DNA common to λ*rev* and the *sbcA*8 mutant is contained in λNM891. (b) The 7.6 kb *Hin*dIII fragment from λNM891 was cloned in pUC18 (Yanisch-Perron *et al.*, 1985) to generate pGK7. Segments of this 7.6 kb *Hin*dIII fragment were cloned and screened for restriction alleviation activity. The *Bam*HI fragment was cloned in pUC18 to generate pGK8. The *Cla*I-*Eco*RI fragment was excised from pGK8, converted to a blunt-ended fragment using Klenow polymerase, and ligated to pUC18 cut with *Sma*I, to generate pGK9. pGK10 was similarly made with the 500 bp *Xho*I-*Eco*RI fragment from pGK9. Thick line, *E. coli* chromosome; thin line, pUC18 DNA; open box, Rac DNA; PB', hybrid Rac attachment site; open triangles, *Hin*dIII target sites; P_{lac}, *lac*UV5 promoter in pUC18; DNA sequence numbering as Clark *et al.* (1993).


(a)
λNM891
λrev
λrev
C600
(Rac+)
B P'
P B'
JC5412
(sbcA8)
B P'
(b)
7600 HindIII
5450 BamHI
5259 EcoRI
4321 XhoI
3875 ClaI
3830 ClaI
2470 EcoRI
1935 BamHI
700 ClaI
245 ClaI
1 HindIII
Plac
pGK7
B P'
recT
recE
pGK8
Plac
recT
recE
pGK9
Plac
pGK10
Plac

7.6 kb *Hin*dIII fragment in pUC18 and screening for restriction alleviation activity (figure 5.2b). The *Bam*HI fragment was cloned in pUC18 to generate pGK8. The *Cla*I-*Eco*RI fragment was excised from pGK8, converted to a blunt-ended fragment using Klenow polymerase, and ligated to pUC18 cut with *Sma*I, to generate pGK9. pGK10 was similarly made with the 500 bp *Xho*I-*Eco*RI fragment from pGK9, and this 500 bp insert was the smallest DNA fragment that conferred restriction alleviation activity.

The region of the Rac prophage in pGK10 has been sequenced (Clark *et al.*, 1993; GenBank accession number L23927) and contains four open reading frames (*orf*s), *orfG* and *orfH'* (truncated) in one orientation (figure 5.3a), and *froR* and *froS* in the other (figure 5.3b). All of the clones that had Lar activity contained a plasmid insert oriented such that *orfG* should be transcribed from the *lac* promoter. In order to examine the orientation dependence of the restriction alleviation phenotype, the 500bp *Eco*RI-*Hin*dIII fragment from pGK10 was cloned in pUC19, to generate pGK11, then removed and recloned in pUC18, to generate pGK12. pGK11 and pGK12, therefore, differ only in the orientation of the 500 bp insert relative to the *lac* promoter. Restriction alleviation activity was unaffected by the orientation of the insert (table 5.1) suggesting that there is sufficient *lar* expression to alleviate restriction even when transcription of the gene is opposed by that from the *lac* promoter. In these experiments, transcription was neither enhanced nor opposed by activation of the *lac* promoter.

A 290 bp fragment including *orfG* (figure 5.3a/b) was amplified by the polymerase chain reaction (PCR) and cloned in pT7-6 to generate pGK15. This plasmid was found to both alleviate restriction and enhance modification by *Eco*KI (tables 5.2 and 5.3). *orfG* is the only complete *orf* transcribed from the T7 promoter in pGK15, although *froR* and *froS* are read in the opposite orientation. To test whether *orfG* encoded Lar, an amber codon was generated within *orfG* that would have no effect on the predicted aa sequence of *froS*. Site-directed mutagenesis was performed using the PCR with four oligonucleotide primers, two flanking primers and two complementary internal primers that overlap the codon to be changed and contains the nucleotide sequence required for the aa substitution (Higuchi, 1990). Amplification by the PCR was performed separately with each pair of correctly oriented flanking and internal primers, with a wild-type DNA template, to generate two products. The DNA template and excess primers were removed by separation through an agarose gel and the full-length fragment, containing the required substitution, was amplified by a second round of the PCR using the products of the first PCR as the template and the two flanking

```
(a)          HindIII
          GGAAGCTTCAGAGGAATAATTCAGC                          G2370
          ClaI
3875   5'ATCGATAATTCAGAGGAATAATTCAGCCTGGCGGTGTAATG
          I  D  N  S  E  E  *                                RecT
                                                  M          OrfG

3916   CACCGCCAACTTGAAATATTTTTTATGAGAAAAATTATGAGAT
        H  R  Q  L  E  I  F  F  M  R  K  I  M  R             OrfG

3959   ATGACAATGTTAAACCATGTCCATTTTGTGGTTGTCCATCAGT
       Y  D  N  V  K  P  C  P  F  C  G  C  P  S  V          OrfG

                                        TACCGAGCGTAGTGTAAC    G2371
4002   AACGGTGAAAGCCATTTCAGGATATTACCGAGCGAAGTGTAAC
         T  V  K  A  I  S  G  Y  Y  R  A  K  C  N            OrfG

4045   GGATGCGAATCCCGAACCGGTTATGGTGGAAGTGAAAAAGAAG
        G  C  E  S  R  T  G  Y  G  G  S  E  K  E             OrfG

4088   CACTCGAAAGATGGAATAAACGAACCACTGGAAATAATAATGG
       A  L  E  R  W  N  K  R  T  T  G  N  N  N  G          OrfG

                                        GGTGATAACTTTTCCTTCT   G2372
4131   AGGTGTTCATGTATAAAATTACCGCCACTATTGAAAAGGAAGG
         G  V  H  V  *                                      OrfG
                  M  Y  K  I  T  A  T  I  E  K  E  G         OrfH
       EcoRI
       TAAGGG                                               G2372
4174   TGGCACTCCTACTAACTGGACAAGATATTCAAAATCTAAACTA
         G  T  P  T  N  W  T  R  Y  S  K  S  K  L           OrfH

4217   ACGAAATCAGAATGCGAAAAAATGCTCTCAGGTAAAAAAGAAG
        T  K  S  E  C  E  K  M  L  S  G  K  K  E            OrfH

4260   CAGGCGTTTCCAGAGAGCAGAAAGTAAAACTGATAAATTTTAA
       A  G  V  S  R  E  Q  K  V  K  L  I  N  F  N          OrfH

                           XhoI
4303   TTGCGAGAAACTTCAGTCCTCGAG 3'
         C  E  K  L  Q  S  S                                OrfH
```

Figure 5.3. Nucleotide sequence of the *Cla*I-*Xho*I fragment encoding *lar*, and the predicted aa sequence of possible gene products. (a) Sequence in the orientation co-directional with *recET*, containing *orfG* and *orfH*. (b) The complementary sequence, in the 5'-3' orientation, containing *froS* and *froR*. 'Flanking' primers used for the PCR and one of the internal primers used to generate the K42Am substitution (codon in bold) are shown. Target sites for type II ENases are underlined.

(b)

*Xho*I

4326 5′CTCGAGGACTGAAGTTTCTCGCAATTAAAATTTATCAGTTT

4285 TACTTTCTGCTCTCTGGAAACGCCTGCTTCTTTTTTACCTGAG

4242 AGCATTTTTTCGCATTCTGATTTCGTTAGTTTAGATTTTGAAT

*Eco*RI

GGGAATTCTTCCTTTTCAATAGT G2372

4199 ATCTTGTCCAGTTAGTAGGAGTGCCACCTTCCTTTTCAATAGT

GG G2372

4156 GGCGGTAATTTTATACATGAACACCTCCATTATTATTTCCAGT

M N T S I I I S S FroS

4113 GGTTCGTTTATTCCATCTTTCGAGTGCTTCTTTTTCACTTCCA

G S F I P S F E C F F F T S FroS

CAATGT**GAT**GCGAGCCA G2371

4070 CCATAACCGGTTCGGGATTCGCATCCGTTACA**CTT**CGCTCGGT

T I T G S G F A S V T **L** R S V FroS

T G2371

4027 AATATCCTGAAATGGCTTTCACCGTTACTGATGGACAACCACA

I S * FroS

M A F T V T D G Q P Q FroR

3984 AAATGGACATGGTTTAACATTGTCATATCTCATAATTTTTCTC

N G H G L T L S Y L I I F L FroR

CGA G2370

3941 ATAAAAAATATTTCAAGTTGGCGGTGCATTACACCGCCAGGCT

I K N I S S W R C I T P P G FroR

*Hin*dIII

CTTAATAAGGAGACTTCGAAGG G2370

*Cla*I

3898 GAATTATTCCTCTGAATTATCGAT 3′

* FroR

Host	Titre λ.0
NM794	4.5×10^{10}
NM777 pUC18	2.4×10^{7}
NM777 pGK11	2.1×10^{10}
NM777 pGK12	3.8×10^{10}

Table 5.1. Orientation dependence of Lar activity. Restriction alleviation activity was determined for each host by measuring the titre of unmodified λ*rat*$^{-}$ phages (λNM325.0) normalised against the titre of modified λ (λNM325.K). This corrects for variation resulting from factors other than restriction, see chapter 2.

Host	eop of progeny phages
NM526(DE3) pT7-6	4.6×10^{-3}
NM526(DE3) pGK15	1.0

Table 5.2. Modification enhancement activity of *lar*. A restriction deficient, modification proficient strain (NM526) was transformed with the plasmid of interest and infected with unmodified λ phages (λNM325.0). After a single round of infection, the degree of phage modification was estimated by measuring the efficiency of plating (eop) of progeny phages on a restricting strain (Ymel) relative to that on a non-restricting derivative (NM526), see chapter 2.

Host	Lar	Titre λ.0
NM777(DE3) pT7-6		3.3×10^{7}
NM777(DE3) pGK15	Wild-type	4.0×10^{10}
NM777(DE3) pGK16	K42Am	2.5×10^{7}
NM777(DE3) pGK17	M1I	3.1×10^{10}
NM777(DE3) pGK18	M10I	2.8×10^{10}
NM777(DE3) pGK19	M14I	1.5×10^{10}
NM777(DE3) pGK20	M1Am	3.6×10^{10}
NM777(DE3) pGK21	M10Am	2.8×10^{10}
NM777(DE3) pGK22	M14Am	1.1×10^{7}
NM777(DE3) pGK23	M10Am M14I	2.6×10^{7}
Ymel(DE3) pT7-6		1.7×10^{7}
Ymel(DE3) pGK15	Wild-type	2.7×10^{10}
Ymel(DE3) pGK16	K42Am	2.2×10^{10}
Ymel(DE3) pGK22	M14Am	2.1×10^{7}
Ymel(DE3) pGK23	M10Am M14I	8.9×10^{6}

Table 5.3. Restriction alleviation activity of *lar* mutants. Restriction alleviation activity was determined for each host by measuring the titre of unmodified λ*ral*$^{-}$ phages (λNM325.0) normalised against the titre of modified λ (λNM325.K), see chapter 2.

primers (figure 5.4). The flanking and internal primers used for the site-directed mutagenesis of *orfG*, described below, are shown in figure 5.3 and 5.5, and an example of the separation of the products generated by the first and second round of the PCR is shown in figure 5.6. The DNA fragments generated by the PCR were cloned in pT7-6 using the *Hin*dIII and *Eco*RI sites in the flanking primers, and the nucleotide sequence determined using Sequenase (figure 5.7, see comments in the legend). To test whether *orfG* encoded *lar*, codon 42 of *orfG* was changed from AAG to TAG (K42Am; figure 5.3a) and this is a synonymous change in *froS* (the codon for L35 changes from CTT to CTA; figure 5.3b). The resulting plasmid (pGK16; K42Am) was examined for restriction alleviation activity in both a *sup°* (NM777) and *supF* (Ymel) host (table 5.3). Restriction alleviation was quantified by measuring the titre of unmodified λ on a restriction-proficient host transformed with the plasmid of interest, and the ability to enhance modification was examined by a single round of infection of *ral*⁻ phages on restriction-deficient, modification-proficient transformants. Ymel (r_K^+, *supF*), NM526 ($r_K^- m_K^+$ derivative of Ymel) or NM777 (r_K^+, *sup°*), lysogenic for λDE3, was the host in these experiments, and the phage was λNM325 that contains a *bio* substitution that removes a region of phage DNA including both the *int* and *ral* genes. It is important that the phage is *ral*⁻, as the progeny of a *ral*⁺ phages would be efficiently modified in a single round of infection, and the phage is *inf* to ensure that it does not act as a helper phage to allow λDE3 to excise from the host chromosome. Restriction was not alleviated in a *sup°* host transformed with pGK16, but was alleviated in a *supF* host. The open reading frame *orfG*, therefore, encodes a Ral-like antirestriction activity that is almost certainly the Lar activity identified by Toothman (1981).

Examination of the DNA sequence at the start of *lar* identified a number of possible translational start codons (figure 5.3a). The DNA codon ATG encodes the most common translational initiation codon (Gold, 1988; McCarthy and Gualerzi, 1990), and there are three of these close to the start of *orfG* at positions 1, 10, and 14. To identify which ATG was the translational initiation codon, each ATG codon was mutated individually to one for isoleucine; the first ATG codon was changed to ATA (pGK17) and codons 10 and 14 to ATC (pGK18 and pGK19 respectively; figure 5.5). It was anticipated that a mutation to the ATG codon required for translational initiation would abolish Lar activity. However, all three mutants retained restriction alleviation activity (table 5.3) indicating that a single ATG codon was unlikely to be wholly responsible for translational initiation.

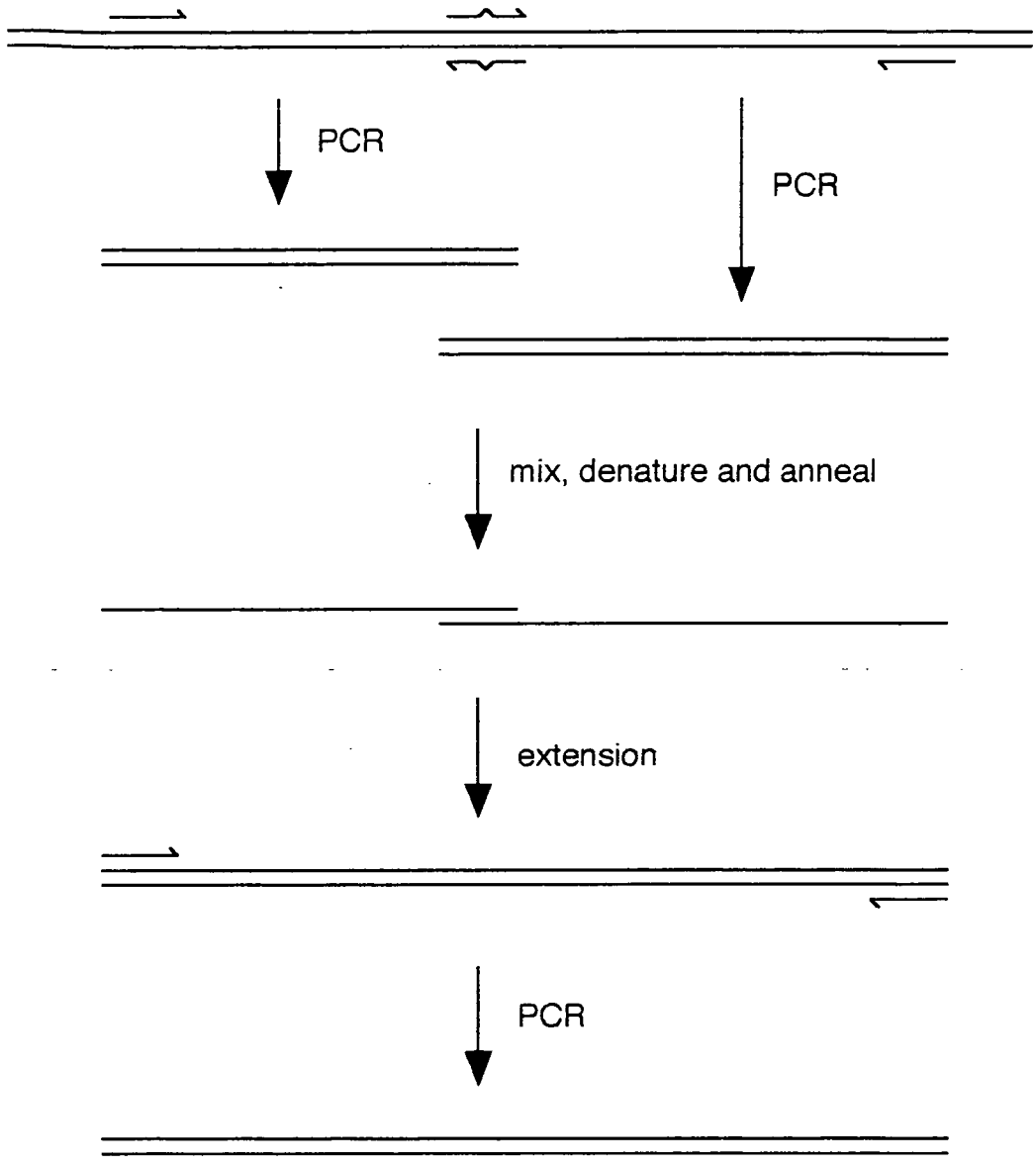


Figure 5.4. Site-directed mutagenesis using the PCR. See text for details. Primers are drawn as short lines, arrows indicate 5'-3' orientation and a 'bulge' indicates base mismatches; parallel lines, the DNA template.

Figure 5.5. The primers used to generate mutations at the start of the *lar* gene. The nucleotide and predicted aa sequence at the start of *lar* are shown. Bold type indicates the codons changed, and primer designations in brackets are the complementary oligonucleotides used in the PCR.

CTGGCGGTGTA**ATG**CACCGCCAACTTGAAATATTTTTT**ATG**AGAAAAATT**ATG**AGATATGACAATGTA *lar*

M H R Q L E I F F **M** R K I **M** R Y D N V aa

Isoleucine substitutions

G5859 GTGTA**ATA**CACCGCCAAC (G5860)

G7230 TTT**ATC**AGAAAAATT**ATG**AG (G7231)

G3789 GAAAAATT**ATC**AGATATGAC (G3788)

Amber codons

G8267 GTGTA**TAG**CACCGCCAAC (G8266)

H1207 TTT**TAG**AGAAAAATT**ATG**AG (H1208)

H1832 GAAAAATT**TAG**AGATATGAC (H1831)

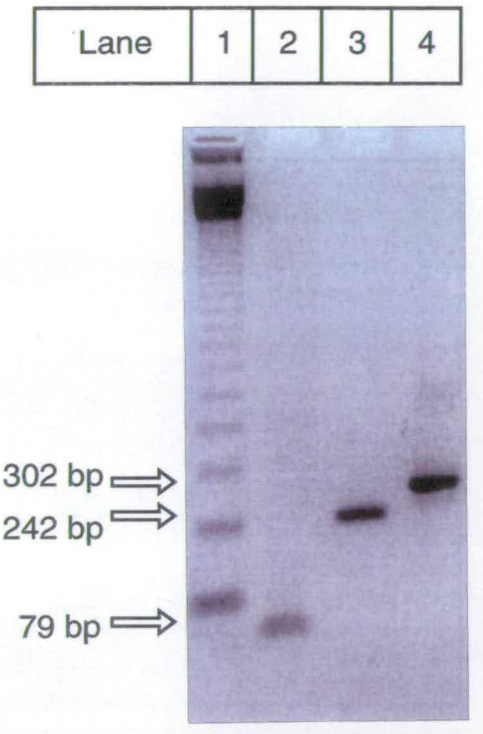
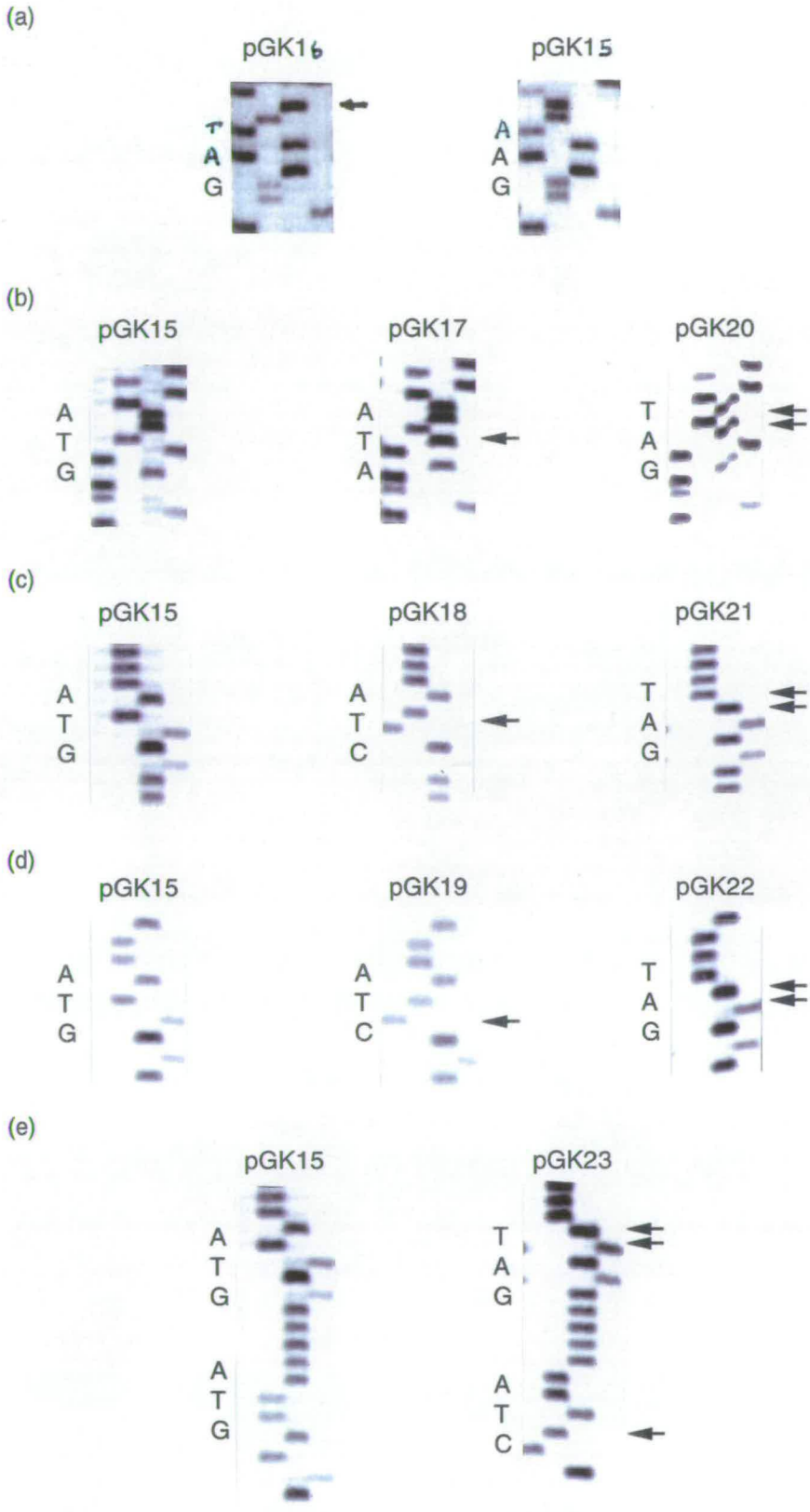


Figure 5.6. Site-directed mutagenesis using the PCR - an example of the three products. Lane 1, 123 bp ladder; lane 2, fragment amplified using primers G7231 and G2370; lane 3, fragment amplified using primers G2372 and G7230; lane 4, fragment amplified using primers G2372 and G2370 with DNA purified from lanes 2 and 3 as the template. Arrows indicate the size of the fragments amplified by the PCR.

Figure 5.7. The nucleotide sequence of the codons changed by site-directed mutagenesis. (a) pGK15 (wild-type) and pGK16 (K42Am). (b) pGK15 (wild-type), pGK17 (M1I) and pGK20 (M1Am). (c) pGK15 (wild-type), pGK18 (M10I) and pGK21 (M10Am). (d) pGK15 (wild-type), pGK19 (M14I) and pGK22 (M14Am) (e) pGK15 (wild-type) and pGK23 (M10Am M14I). In all parts, sequence is read GATC, the bases changed are indicated with an arrow and the codon is written at the side. Note, the sequence is of the reverse strand to *lar*. The nucleotide sequence of each fragment amplified by the PCR was determined using Sequenase. pGK17 and pGK20 both contain an additional mutation (G-A), outside any ORF, immediately 5' to the primers used in the PCR. Mutants could not be isolated that did not contain an additional mutation at this position. All other plasmids contain the expected sequence.


(a)
pGK16
pGK15
T
A
G
A
A
G
(b)
pGK15
pGK17
pGK20
A
T
G
A
T
A
T
A
G
(c)
pGK15
pGK18
pGK21
A
T
G
A
T
C
T
A
G
(d)
pGK15
pGK19
pGK22
A
T
G
A
T
C
T
A
G
(e)
pGK15
pGK23
A
T
G
A
T
G
T
A
G
A
T
C

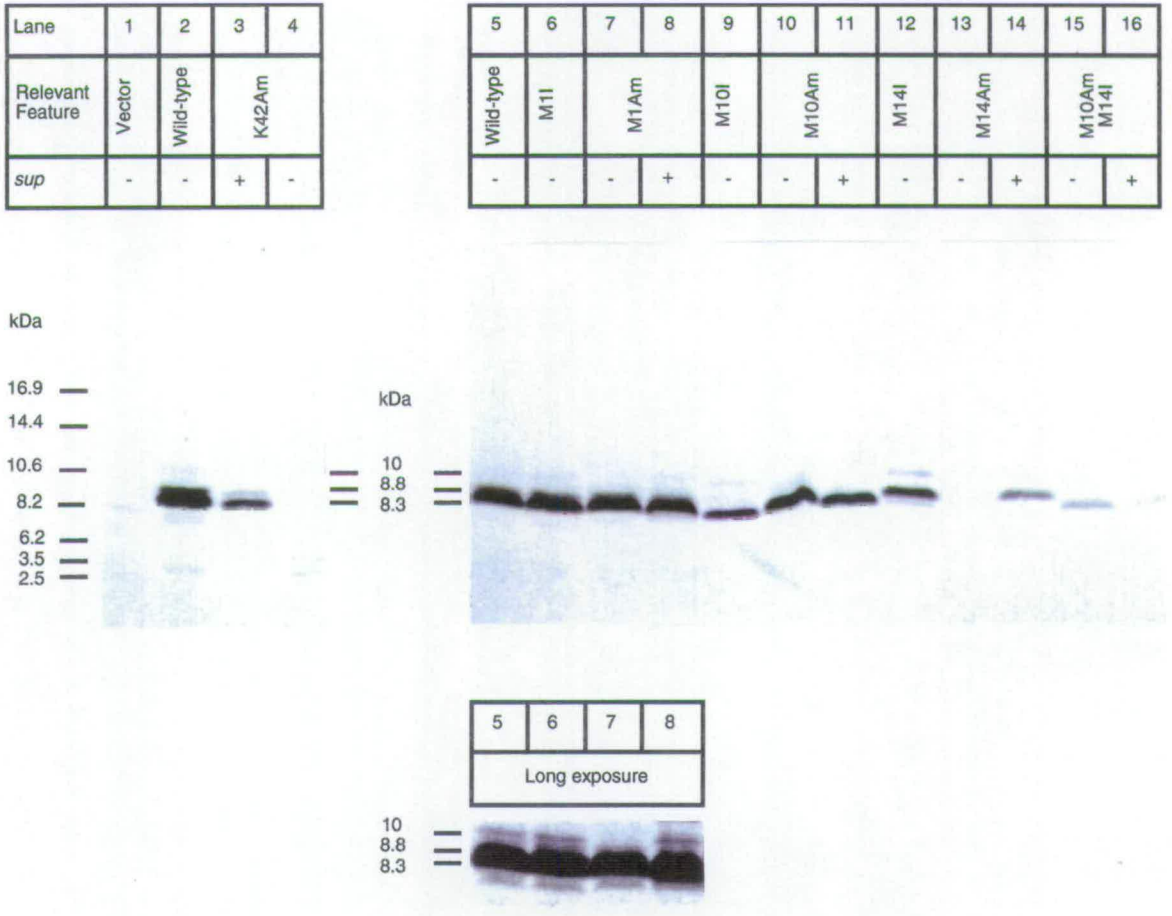
To identify the minimum coding sequence required for restriction alleviation activity, each of the three ATG codons was mutated to an amber codon (figure 5.5). An amber codon is recognised as a translational 'stop' codon in a *sup°* host whereas, in a *supF* host, a tyrosine residue is substituted at this position. An amber mutation in either the first (M1; pGK20) or second (M10; pGK21) ATG codons did not affect activity, but a change to the third ATG codon (M14; pGK22) abolished activity, both in a *sup°* and *supF* host. This showed that translation of the DNA sequence upstream of the second ATG codon was not required for restriction alleviation. To ascertain whether the third ATG codon could be mutated to one for isoleucine (ATC) when no translation occurred upstream of the second methionine codon, a double mutant was made (M10Am M14I; pGK23). The double mutant was generated by site directed mutagenesis using the PCR, with the same primers used to generate pGK19 (M14I; figure 5.5) but with pGK21 (M10Am) as the template. This double mutant is unable to alleviate restriction in either a *sup°* or *supF* host (table 5.3). To further the interpretation of these results, the polypeptides produced by each of the *lar* clones were examined.

Identification of the *lar* gene product

The products of the *lar* clones were visualised by the selective labelling of proteins encoded by genes under the control of the T7 promoter (Tabor, 1990). *E. coli* RNA polymerase, but not T7 RNA polymerase, is inactivated by the antibiotic Rifampicin, and the products of genes transcribed by T7 RNA polymerase may be selectively labelled if cultures expressing the T7 gene *1* are first treated with Rifampicin and then ^{35}S-labelled aa added to the media (see chapter 2). Different sized polypeptides were detected depending on both the plasmid and the suppressor genotype of the host (figure 5.8). The host strains alone, BL21(DE3)*sup°* or BL21(DE3)*supF*, produced no labelled protein (data not shown), nor did pT7-6 in either the *sup°*(NM777, lane 1) or *supF* host (Ymel, data not shown). pGK15 (wild type *lar*; lanes 2 and 5) produced 2 main products, with estimated molecular mass 8.8 kDa, and 8.3 kDa. The smaller of these is the more abundant.

The proteins produced by the *lar* mutants were similarly examined. In a *sup°* host, pGK16 (K42Am) produced no polypeptides of an appropriate size for Lar (lane 4), although a number of smaller products were seen, consistent with truncated forms of Lar. In contrast, in a *supF* host (lane 3) both the 8.8 kDa and 8.3 kDa proteins were produced by pGK16. This correlates well with the observed phenotype; pGK16 does not alleviate restriction in a *sup°* host, but does in a *supF* host.

Figure 5.8. Visualisation of the *lar* gene products. (a) Lane 1, BL21(DE3) pT7-6 (vector); lane 2, BL21(DE3) pGK15 (wild-type); lane 3, NM793(DE3) pGK16 (K42Am); lane 4, BL21(DE3) pGK16. (b) Lane 5, BL21(DE3) pGK15 (wild-type); lane 6, BL21(DE3) pGK17 (M1I); lane 7, BL21(DE3) pGK20 (M1Am); lane 8, NM793(DE3) pGK20; lane 9, BL21(DE3) pGK18 (M10I); lane 10, BL21(DE3) pGK21 (M10Am); lane 11, NM793(DE3) pGK21; lane 12, BL21(DE3) pGK19 (M14I); lane 13, BL21(DE3) pGK22 (M14Am); lane 14, NM793(DE3) pGK22; lane 15, BL21(DE3) pGK23 (M10Am M14I); lane 16, NM793(DE3) pGK24. (c) A longer exposure of the autoradiograph showing the 10 kDa Lar polypeptide in lanes 5-8.


Lane
1
2
3
4
Relevant Feature
Vector
Wild-type
K42Am
sup
-
-
+
-
5
6
7
8
9
10
11
12
13
14
15
16
Wild-type
M1I
M1Am
M10I
M10Am
M14I
M14Am
M10Am M14I
-
-
-
+
-
-
+
-
-
+
-
+
kDa
16.9
14.4
10.6
8.2
6.2
3.5
2.5
kDa
10
8.8
8.3
5
6
7
8
Long exposure
10
8.8
8.3

Both pGK18 (M10I; lane 9) and pGK21 (M10Am), in a *sup°* (lane 10) or *supF* host (lane 11), produced only the 8.3 kDa polypeptide, from which it is concluded that the second ATG (M10) is the initiation codon for the 8.8 kDa protein.

The third ATG codon (M14) is the major translational start of *lar*, encoding the 8.3 kDa protein, as this polypeptide was not detected in either a *sup°* or *supF* host when the third ATG codon was mutated to an amber codon (M14Am; pGK22; lanes 13 and 14 respectively). Nevertheless, pGK19 (M14I; lane 12) and pGK23 (M10Am M14I; lanes 15 and 16) both produced small quantities of the 8.3 kDa protein. The ATC codon in pGK19 and pGK23, therefore, appears to be used as a poor translational start in these mutants. It seemed unlikely that the codons between the second and third ATG codons could encode the translational start of the 8.3 kDa protein as this polypeptide was not visualised in a *supF* host transformed with pGK22 (M14Am), and yet the 8.8 kDa polypeptide (starting at M10) was detected (lane 14).

A 10 kDa polypeptide, specific to the *lar* clones, can be visualised after longer exposures of the autoradiograph (figure 5.8c). The initiation codon for this low abundance polypeptide must precede the first ATG codon since, in a *supF* host, pGK20 (M1Am) produced the 10 kDa, 8.8 kDa and 8.3 kDa proteins (lane 8) but, in a *sup°* host, this clone only produced the two smaller polypeptides (lane 7). pGK17 (M1I) did not affect the polypeptides produced (lane 6).

In vitro methylation with cell extracts

Lar activity was examined *in vitro* using cell extracts of NM519 transformed with pPT7 to methylate oligonucleotides. NM519 is a derivative of *E. coli* K-12 that constitutively expresses *lar* (*sbcA23*), and experiments were conducted as described for cells expressing *ral* and the genes cloned in pPT7 that encode the *Eco*KI MTase (chapter 4). The *in vivo* results show that NM519 cells exhibit enhanced methylation, but no evidence of this is found *in vitro* (table 5.4; comments in legend); a hemimethylated oligonucleotide is efficiently modified whereas an unmodified oligonucleotide is only poorly methylated (figure 5.9). It may be that Lar, like Ral, only has a slight stimulatory affect on MTase activity that would not be detected in this assay (see chapter 4 for discussion).

Comparisons between the predicted aa sequences of Ral and Lar

The minimum aa sequence required for Lar activity starts at the methionine encoded by the third ATG codon (M14). The nucleotide and predicted aa sequences of the *lar*

Host	eop of progeny phages
NM519 pPT7	4.5×10^{-1}
NM519 pJF118	9.8×10^{-1}

Table 5.4. Modification enhancement by *E. coli sbcA*23. The degree of modification after a single round of infection (sri) of unmodified phages (λ325.0) was estimated by measuring the efficiency of plating (eop) of progeny phages on a restricting strain (C600) relative to that on a non-restricting strain (NM679). Phages were titred by spotting 10 µl of diluted phage on a lawn of cells in BBL top layer, and the measurement of eop is not as accurate as the method described in chapter 2. However, this assay, based on sampling small volumes, is sufficiently accurate to indicate the enhancement of modification.

Protein	Lar					Lar + MTase				
DNA	- -		+ -		+ +	- -		+ -		+ +
Time	1	5	1	5	5	1	5	1	5	5

Figure 5.9. *In vitro* DNA methylation with cell extracts. See text for details. =, unmodified target (408L:409L); ±, hemimethylated target (860L:409L); ‡, fully modified target (860L:409L).

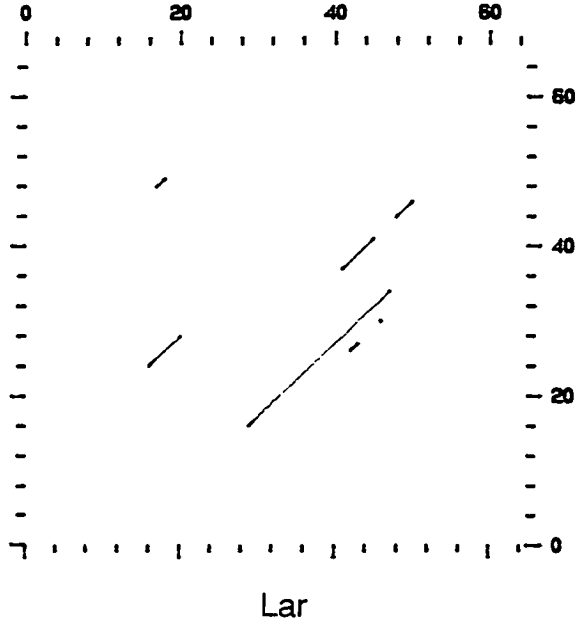
gene, starting at the third ATG codon, were compared to those of *ral* using the UWGCG Compare and DotPlot programs (figure 5.10a/b). The two proteins are much more dissimilar than λ Ral and P22 Ral (Semerjian *et al.*, 1989), although a number of regions do align. The lack of similarity at the aa level cannot be explained by a frame-shift mutation in one of the genes as no diagonal offset lines are seen on the DotPlot of the nucleotide sequences (figure 5.10b). The predicted aa sequences of Ral and Lar were also compared using the UWGCG Gap program (figure 5.10c) and this alignment may indicate an important structural and/or functional role for the residues CNGC and W--RTT in the two polypeptides. The region of *lar* encoding W--RTT is aligned in both nucleotide and aa sequence.

Discussion

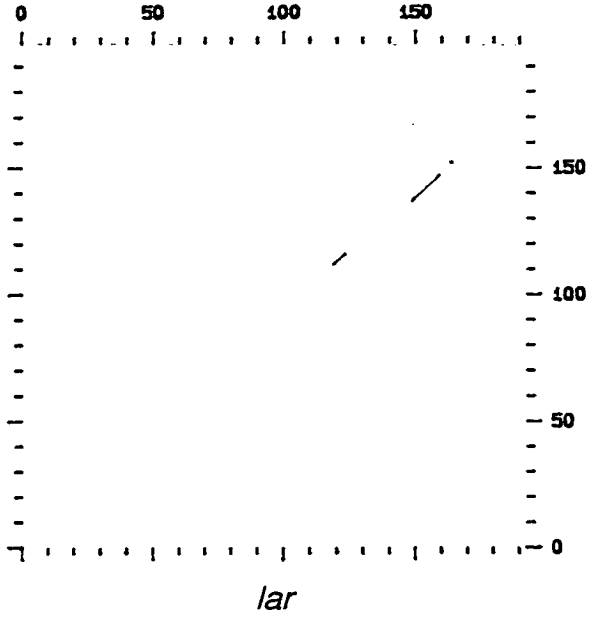
The *lar* gene, responsible for restriction alleviation and modification enhancement activity of the *E. coli* Rac prophage, has been identified and although it has been sequenced previously, its function was unknown and it was termed *orfG* (Clark *et al.*, 1993). The *lar* gene, cloned in an expression plasmid, has three translation initiation codons producing polypeptides of estimated molecular mass 10 kDa, 8.8 kDa and 8.3 kDa. pGK21 (M10Am) alleviates restriction in a *sup°* host and yet only produces the 8.3 kDa protein; demonstrating that this is the only form of Lar that is required for activity. The translational initiation codon for this protein appears to be codon 14 (ATG). However, when codon 14 is changed to an ATC (pGK19 and pGK23) the 8.3 kDa polypeptide is still produced, although as is apparent from the labelling experiments with pGK19 (M14I; figure 5.8b, lane 12), the amount of the 8.3 kDa polypeptide is much reduced relative to the 8.8 kDa protein. It has been shown that ATC can be used to initiate protein synthesis (Romero and Garcia, 1991) and results would indicate that, in pGK19 and pGK23, this codon can be used as a poor translational initiation codon. Restriction was not alleviated by pGK23 (M10Am M14I). It is likely that insufficient Lar was produced to alleviate restriction in the absence of providing T7 RNA polymerase to induce expression of the *lar* gene. The restriction alleviation and modification enhancement assays provided clear phenotypes with uninduced, wild-type *lar* (pGK15; tables 5.2 and 5.3). It seemed preferable to use these conditions to examine Lar activity rather than those in which the cell physiology is perturbed by the gross overproduction of Lar dependent upon T7 RNA polymerase. The initiation codon for the 10 kDa form of Lar appears to be before the first ATG codon (M1), and that for the 8.8 kDa polypeptide is the second ATG codon (M10). The 8.3 kDa protein is the most abundant, although it is not possible to predict

Figure 5.10. Comparisons between the *ral* and *lar* genes and products. (a) The predicted aa sequences of Ral and Lar (starting at M14) were compared using the UWGCG Compare program (window size, 30; stringency, 9) and plotted using DotPlot. (b) The nucleotide sequences of *ral* and *lar* (starting at codon 14) were compared using the UWGCG Compare program (window size, 21; stringency, 14) and plotted using DotPlot. (c) The predicted aa sequences of Ral and Lar (starting at M14) were compared using the UWGCG Gap program (gap weight, 3.0; length weight, 0.1; lines indicate identity; dots indicate similarity; % similarity, 52%; % identity, 25%). See text for further details.

(a)



(b)



(c)

```
Lar   MRYDNVKPCPFCGCPSVTVKAISGYYRAKCNGCESRTGYGG 54
                 ...:. .. :| :: :||||. ..:: .
Ral   ............MTTTIDKNQWCGQFK.RCNGCKLQSECMV 28

      SEKEAL.........ERWNKRTTG.........NNNGGVHV 77
      ...| :          ::|. |||:          ||.::
      KPEEMFPVMEDGKYVDKWAIRTTAMIARELGKQNNKAA... 66
```

what forms would be produced from the Rac prophage and whether the different forms have any biological significance.

The Rac prophage includes gènes that can serve the same functions as the early genes of phage λ, including integration, recombination, immunity and replication (Low, 1973; Gottesman *et al.*, 1974; Diaz *et al.*, 1979). The *recE* and *recT* genes from Rac can substitute for the general recombination genes of λ, *exo* and *bet* respectively (Kushner *et al.*, 1974; Clark *et al.*, 1993). In λ, *ral* is upstream of the recombination genes, whereas in the Rac prophage, *lar* is immediately downstream of *recT*. The *ral*, *exo* and *bet* genes are all transcribed from P_L in λ and it is logical to expect that they are also cotranscribed in Rac. Other genes in this transcriptional unit include those required by λ for excision (*xis*) and integration (*int*); analogous genes in the Rac prophage may be encoded in the DNA sequence downstream of *lar* (figure 5.11). Despite the rearrangement of gene order in Rac compared to λ, the DNA sequences of *ral* and *lar* show a 47% identity suggesting that the two genes have a common origin.

Ral and Lar are both able to alleviate restriction and enhance modification of the type IA R-M systems *in vivo* and an examination of their predicted aa sequences may, therefore, help identify key residues involved in structure and/or function. Cysteine residues are commonly used to stabilise the structure of proteins, either forming disulphide bridges or associating with divalent metal ions (Branden and Tooze, 1991). Lar contains a sequence of cysteine residues, Cys - X_2 - Cys - X_{17} - Cys - X_2 - Cys, that resembles a zinc-binding motif. Ral does not contain the same motif but it does contain a symmetrical sequence, Cys - X_5 - Cys - X_2 - Cys - X_5 - Cys, which may serve a similar role. When the aa sequences of Ral and Lar are compared a number of regions align. The identities CNGC and W--RTT may be relevant. The cysteines in the sequence CNGC may be of structural relevance, whereas the W--RTT residues may indicate a region of functional importance. The best conservation at the nucleotide level is seen in the region encoding W--RTT. Ral is amenable to structural analysis by nuclear magnetic resonance, and polypeptide produced by solid-phase synthesis (Brown *et al.*, 1993) or *in vivo* from the cloned gene (chapter 4) could be used for this purpose. Given the structure of Ral, the motifs identified should help to further distinguish important residues and define targets for site-directed mutagenesis.

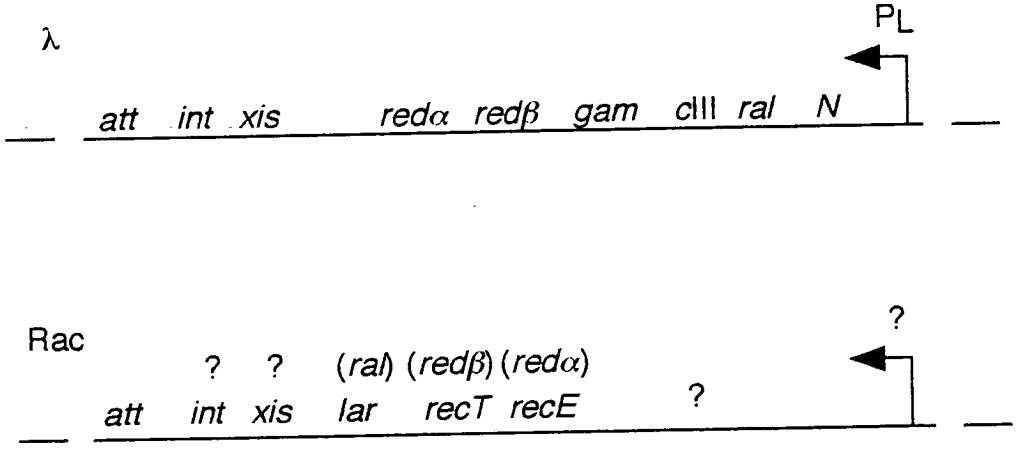


Figure 5.11. The early region of phage λ and the similar region from the Rac prophage. See text for details. In brackets, the products of the genes from phage λ that are functionally similar to those encoded by the Rac prophage; ?, unknown or predicted.

Chapter six: Conclusions and general discussion

Conclusions and general discussion

The Ral function of bacteriophage λ is able to alleviate restriction and enhance methylation by type IA R-M systems (Zabeau *et al.*, 1980; Loenen and Murray, 1986). Ral is not produced to protect its own DNA, but unmodified phages that do escape restriction are more likely to become methylated if they are *ral*$^+$ (see chapter 3). Therefore, the progeny of λ*ral*$^+$ phages that escape restriction are efficiently methylated, whereas the progeny of λ*ral*$^-$ phages are only poorly modified. This would confer an advantage as the progeny of a *ral*$^+$ phage would not be restricted by a host with an R-M system of the same specificity. A number of hypotheses have been proposed to explain the activity of Ral at a molecular level.

Debrouwere *et al.* (1980b) concluded, from experiments performed with defective prophages, that Ral had an anti-termination function and antagonised the transcription termination factor Rho. They suggested that Rho may form a complex with *Eco*KI and that this complex was dissociated by Ral bound to either Rho or *Eco*KI. The *Eco*KI MTase has been purified to homogeneity (Dryden *et al.*, 1993) and only the *hsdM* and *hsdS* gene products are required to form a MTase that modifies hemimethylated DNA efficiently but only methylates unmodified DNA poorly. Rho, therefore, is not required for the 'maintenance' methylation activity of *Eco*KI. Rho is involved in transcriptional termination, and Ral could antagonise this activity (Debrouwere *et al.*, 1980b). However, Loenen and Murray (1986) found that the cloned *ral* gene did not make a strain phenotypically Rho$^-$, and it has been suggested that another gene, *Ea*10, expressed in the defective *ral*$^+$ prophage used by Debrouwere *et al.* (1980b), is responsible for the observed anti-termination activity (Court and Oppenheim, 1983). *Ea*10 encodes a single-strand binding protein of unknown function, although the single-strand binding protein of *E. coli* plays a central role in DNA metabolism and is involved in DNA replication, repair and recombination (Meyer and Laine, 1990).

Zinkevich *et al.* (1987) suggest that Ral affects transcription or translation so that production of one of the subunits of the MTase is increased. It is difficult to see why an increase in only one of the subunits required to form the MTase should lead to an increase in methylation. Overexpression of the genes encoding the *Eco*KI MTase does not enhance methylation (Loenen and Murray, 1986; also see table 4.2) and the only form of *Eco*KI MTase shown to be active is $HsdM_2S_1$ (Dryden *et al.*, 1993; L. Cooper and D. Dryden, unpublished). Zinkevich *et al.* (1987) specifically labelled proteins

from *ral*$^+$ and *ral*$^-$ phages that encoded the *Eco*KI MTase, and concluded that *ral* greatly stimulated synthesis of the HsdM protein, as this subunit could only be detected from the *ral*$^+$ phage. However, the authors do not indicate how HsdM and HsdS were identified, and other proteins are known to migrate with HsdS on SDS-polyacrylamide gels which do not appear to be resolved in this work. In particular, the region cloned includes *mcrB* whose product co-migrates with HsdS (Sain and Murray, 1980). Critically the *ral*$^-$ and *ral*$^+$ phages used by Zinkevich *et al.* (1987) rely on different promoters for the transcription of *hsdMS*. Transcription of *hsdMS* in the *ral*$^-$ phage is driven by the promoter present in *E. coli* K-12, P_{mod}, whereas transcription of *hsdMS* in the *ral*$^+$ phage may also be driven by the λ promoter, P_L. The protein identified as HsdS, therefore, may be a number of unresolved proteins, including McrB, and the higher level of expression of *hsdM* in the *ral*$^+$ phage may be due to transcription from a strong λ promoter.

An attractive explanation for the molecular mechanism of Ral activity, is that Ral interacts directly with the R-M complex (Zabeau *et al.*, 1980; Toothman, 1981; Loenen and Murray, 1986). If Ral changed the ENase so that unmodified DNA was recognised as hemimethylated, then this would explain both phenotypes. Hemimethylated DNA is not a substrate for the ENase so the DNA would not be cut, but it is a good substrate for the MTase so the DNA would be efficiently methylated. Ral cannot only block restriction as Ral is required for efficient methylation of unmodified DNA, even in the absence of HsdR, the subunit required for restriction (Loenen and Murray, 1986). The *ral* gene has been overexpressed and its product purified (chapter 4). The purified *Eco*KI MTase (Dryden *et al.*, 1993) and Ral have been used to examine methylation enhancement *in vitro* and similar experiments have been done with extracts from cells expressing *hsdMS* and *ral*. Preliminary results, following the transfer of a [^{3}H]methyl group from AdoMet to a DNA substrate, do not show any effect of Ral; Ral does not appear to greatly enhance MTase activity, and unmodified DNA is only methylated poorly. This is inconsistent with Ral changing *Eco*KI so that unmodified DNA is recognised as hemimethylated. Ral may only slightly stimulate methylation and this probably would not be detected in the assay used. It is not at all clear how much Ral would need to stimulate methylation to produce the observed phenotypes, but presumably if methylation was not greatly enhanced then DNA cutting must be actively blocked by Ral. This could be achieved in a number of ways; Ral could block the formation of the ENase complex, $HsdR_2M_2S_1$, or prevent a process required specifically for restriction, e.g. DNA translocation. To examine restriction alleviation *in vitro*, experiments could be

performed with purified *Eco*KI ENase, a small quantity of which is available (L. Cooper and D. Dryden, unpublished), although the use of cell extracts would overcome some of the problems associated with using purified proteins, not least the possible inactivation of the proteins during purification.

The work described has done little to explain the molecular mechanism of Ral activity, but the identification of the *lar* gene that encodes the Ral-like activity of the *E. coli* Rac prophage (chapter 5), has identified a number of conserved residues in the proteins that may be important for structure and/or function. The *ral* genes of phage λ and P22 show 75% identity (Semerjian *et al.*, 1989), whereas *lar* shows only 47% identity with the *ral* gene of λ. The predicted aa sequences of Ral and Lar share only 25% identity, 52% similarity, and only two short regions align. In particular, the identity W--RTT may indicate residues important for function. Site-directed mutagenesis could be performed on these residues, although this may be a little presumptuous in the absence of any supporting structural or biochemical evidence. The identification of other genes that encode Ral-like activities would be useful to see if the regions of identity between *ral* and *lar* were present, and confirm their possible importance. A collection of lambdoid phages are available that could be screened for Ral activity (Highton *et al.*, 1990), and the most interesting genes would be those that shared little overall identity with either *ral* or *lar*. Southern analysis of DNA prepared from the collection of lambdoid phages, with probes from the *ral* and *lar* genes, would identify phages with similar DNA sequence. Those phages that encoded a Ral-like function, but did not contain DNA that hybridised to probes from *ral* or *lar*, would probably be the most interesting. The genes encoding the function that alleviated restriction could be cloned by a similar method to the one employed in the isolation of *lar* (see chapter 5). Fragments from the phage DNA cut with a type II ENase, e.g. *Sau*3A, would be cloned in a plasmid expression vector and used to transform a strain proficient for *Eco*KI restriction. Clones would be screened for the ability to restrict unmodified phages (λ*vir*.0), and the plasmids purified from the hosts that are phenotypically r_K^-. Further analyses, including the modification phenotype, should identify genes encoding Ral-like functions and the determination of the nucleotide sequence would provide more information on important residues. Another approach to identify important residues is random mutagenesis. A modified PCR has been used, that reduces the fidelity of replication by Taq DNA polymerase, to saturate regions of DNA with changes (Leung *et al.*, 1989; Zhou *et al.*, 1991; Spee *et al.*, 1993). A method could be optimised to ensure one or two aa substitutions for every *ral* clone. Phenotypic and sequence analyses of each clone, would give information both on

important residues and those that could be changed without effecting Ral activity. If the alleviation of restriction and modification enhancement phenotypes of Ral involve distinct molecular interactions, then some very interesting mutants may also be isolated that effect only restriction or modification.

References

Abadjieva, A., J. Patel, M. Webb, V. Zinkevich and K. Firman. 1993. A deletion mutant of the type IC restriction endonuclease *Eco*R124I expressing a novel DNA specificity. Nucleic Acids Res. **21**: 4435-4443.

Amitsur, M., R. Levitz and G. Kaufmann. 1987. Bacteriophage T4 anticodon nuclease, polynucleotide kinase and RNA ligase reprocess the host lysine transfer RNA. EMBO J. **6**: 2499-2503.

Amitsur, M., I. Morad, D. Chapman-Shimshoni and G. Kaufmann. 1992. Hsd restriction-modification proteins partake in latent anticodon nuclease. EMBO J. **11**: 3129-3134.

Appleyard, R.K. 1954. Segregation of new lysogenic types during growth of a doubly lysogenic strain derived from *Escherichia coli* K-12. Genetics. **39**: 440-452.

Arber, W. 1965a. Host specificity of DNA produced by *Escherichia coli*. V. The role of methionine in the production of host specificity. J. Mol. Biol. **11**: 247-256.

Arber, W. 1965b. Host controlled modification of bacteriophage. Ann. Rev. Microbiol. **19**: 365-377.

Arber, W. 1994. Bacteriophage transduction. *In* **Encyclopaedia of Virology**. R.G. Webster and A. Granoff (eds), Academic Press Ltd. pp. 107-112.

Arber, W. and D. Dussoix. 1962. Host specificity of DNA produced by *Escherichia coli*. I. Host controlled modification of bacteriophage λ. J. Mol. Biol. **5**: 18-36.

Arber, W. and M.L. Morse. 1965. Host specificity of DNA produced by *Escherichia coli*. II. Effects on bacterial conjugation. Genetics. **51**: 137-148.

Arber, W. and S. Linn. 1969. DNA modification and restriction. Ann. Rev. Biochem. **38**: 467-500.

Arber, W. and D. Wauters-Willems. 1970. Host specificity of DNA produced by *Escherichia coli*. XII. The two restriction and modification systems of strain 15T⁻. Molec. Gen. Genetics. **108**: 203-217.

Argos, P. 1985. Evidence for a repeating domain in type I restriction enzymes. EMBO. J. **4**: 1351-1355.

Athanasiadis, A., M. Vlassi, D. Kotsifaki, P.A. Tucker, K.S. Wilson and M. Kokkinidis. 1994. Crystal structure of *Pvu*II endonuclease reveals extensive structural homologies to *Eco*RV. Struct. Biol. **1**: 469-475.

Bandyopadhyay, P.K., F.W. Studier, D.L. Hamilton and R. Yuan. 1985. Inhibition of the type I restriction-modification enzymes *Eco*B and *Eco*K by the gene *0.3* protein of bacteriophage T7. J. Mol. Biol. **182**: 567-578.

Barcus, V.A. 1993. The type I restriction systems of *Escherichia coli.* PhD Thesis. University of Edinburgh, UK.

Barcus, V.A. and N.E. Murray. 1995. *In* **Population Genetics of Bacteria.** S. Baumberg *et al.* (eds), SGM, Cambridge University Press. pp. 31-58.

Barcus, V.A., A.J.B. Titheradge and N.E. Murray. 1995. Alleles and pseudoalleles at the *hsd* locus in natural populations of *E. coli.* Genetics. (in press).

Belogurov, A.A. and E.P. Delver. 1995. A motif conserved among the type I restriction-modification enzymes and antirestriction proteins: a possible basis for mechanism of action of plasmid-encoded antirestriction functions. Nucleic Acids Res. **23**: 785-787.

Bertani, G. and J.J. Weigle. 1953. Host controlled variation in bacterial viruses. J. Bacteriol. **65**: 113-121.

Bickle, T.A. 1993. The ATP-dependent restriction enzymes. *In* **Nucleases.** E.M. Linn *et al.* (eds), Cold Spring Harbor Laboratory Press, NY. pp. 89-109

Bickle, T.A., C. Brack and R. Yuan. 1978. ATP-induced conformational changes in the restriction endonucleases from *Escherichia coli* K-12. Proc. Royal Acad. Sci. **75**: 3099-3103.

Bickle, T.A. and D.H. Krüger. 1993. Biology of DNA restriction. Biol. Rev. **57**: 434-450.

Blumenthal, R.M., S.A. Gregory and J.S. Cooperider. 1985. Cloning of a restriction-modification system from *Proteus vulgaris* and its use in analysing a methylase-sensitive phenotype in *E. coli.* J. Bacteriol. **2**: 501-509.

Borck, K., J.D. Beggs, W.J. Brammar, A.S. Hopkins and N.E. Murray. 1976. The construction *in vitro* of transducing derivatives of phage lambda. Molec. Gen. Genet. **146**: 199-207.

Boye, E. and A. Lϕbner-Olesen. 1990. The role of *dam* methyltransferase in control of DNA replication of *E. coli.* Cell. **62**: 981-989.

Boyer, H.W. 1964. Genetic control of restriction and modification in *Escherichia coli.* J. Bacteriol. **88**: 1652-1660.

Boyer, H.W. 1971. DNA restriction and modification mechanisms in bacteria. Ann. Rev. Microbiol. **25**: 153-176.

Boyer, H.W. and D. Roulland-Dussoix. 1969. A complimentation analysis of the restriction and modification of DNA in *Escherichia coli.* J. Mol. Biol. **41**: 459-472.

Brammar, W.J., N.E. Murray and S. Winton. 1974. Restriction of λ*trp* bacteriophages by *Escherichia coli* K. J. Mol. Biol. **90**: 633-647.

Brandon, C. and J. Tooze. 1991. *In* **Introduction to Protein Structure**. New York: Garland Publishing.

Brown, A.R., S.L. Irving and R. Ramage. 1993. Affinity purification of synthetic peptides and proteins on porous graphitised carbon. Tet. Lett. **34**: 7129-7132.

Brunel, F. and J. Davison. 1979. Restriction insensitivity in bacteriophage T5. III. Characterisation of *Eco*RI-sensitive mutants by restriction analysis. J. Mol. Biol. **128**: 527-543.

Bullas, L.R. and C. Coulson. 1975. DNA restriction and modification systems in *Salmonella.* III. SP, a *Salmonella potsdam* system allelic to the SB system in *Salmonella typhimurium.* Mol. Gen. Genet. **139**: 177-188.

Bullas, L.R., C. Coulson and A. Van Pel. 1976. DNA restriction and modification systems in *Salmonella.* J. Gen. Microbiol. **95**: 166-172.

Bullas, L.R., C. Colson and B. Neufeld. 1980. Deoxyribonucleic acid restriction and modification systems in *Salmonella*: Chromosomally located systems of different serotypes. J. Bacteriol.. **141**: 275-292.

Burckhardt, J., J. Weisemann, D.L. Hamilton and R. Yuan. 1981a. Complexes formed between the restriction endonuclease *Eco*K and heteroduplex DNA. J. Mol. Biol. **153**: 425-440.

Burckhardt, J., J. Weisemann and R. Yuan. 1981b. Characterisation of the DNA methylase activity of the restriction enzyme from *Escherichia coli* K. J. Biol. Chem. **256**: 4024-4032.

Burland, V., G. Plunkett III, D.L. Daniels and F.R. Blattner. 1993. DNA-sequence analysis of 136 kilobases of the *E. coli* genome - organisational symmetry around the origin of replication. Genomics. **16**: 551-561.

Campbell, A. 1977. Defective bacteriophages and incomplete prophages. *In* **Comprehensive Virology**. H. Fraenkel-Conrat and R. Wagner (eds), Plenum Press, NY. pp. 259-328.

Chandrasegaran, S. and H.O. Smith. 1988. Amino acid sequence homologies among 25 restriction endonucleases and methylases. *In* **Structure and expression**. R.H. Sarma and M.H. Sarma (eds), Adenine Press Inc., NY. pp. 149-156.

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

Chen, A., L.M. Powell, D.T.F. Dryden, N.E. Murray and T. Brown. 1995. Tyrosine 27 of the specificity polypeptide of *Eco*KI can be UV crosslinked to a bromodeoxyurideine-substituted DNA target sequence. Nucleic Acids Res. **23**: 1177-1183.

Cheng, X., S. Kumar, J. Pósfai, J.W. Pflugrath and R.J. Roberts. 1993. Crystal structure of the *Hha*I methylase complexed with S-adenosyl-L-methionine. Cell. **74**: 299-307.

Cheng, X., K. Balendiron, I. Schildkraut and J.E. Anderson. 1994. Structure of *Pvu*II endonuclease with cognate DNA. EMBO J. **13**: 3927-3935.

Clark, A.J., V. Sharma, S. Brenowitz, C.C. Chu, S. Sandler, L. Satin, A. Templin, I. Berger and A. Cohen. 1993. Genetic and molecular analyses of the C-terminal region of the *recE* gene from the Rac prophage of *E. coli* K-12 reveal the *recT* gene. J. Bacteriol. **175**: 7673-7682.

Clewell, D.B. and D.R. Helinski. 1969. Supercoiled circular DNA-protein complex in *E. coli*: purification and induced conversion to an open circular DNA form. Proc. Natl. Acad. Sci. USA. **62**: 1159-1166.

Cooper, L.P. and D.T.F. Dryden. 1994. The domains of a type I DNA methyltransferase, interactions and role in recognition of DNA methylation. J. Mol. Biol. **236**: 1011-1021.

Court, D. and A.B. Oppenheim. 1983. Accessory genes. *In* **Lambda II**. R.W. Hendrix *et al.* (eds), Cold Spring Harbor Laboratory Press. NY. pp. 251-277.

Cowan, G.M., A.A.F. Gann and N.E. Murray. 1989. Conservation of complex DNA recognition domains between families of restriction enzymes. Cell. **56**: 103-109.

Craigie, J. and C.H. Yen. 1938. The demonstration of types of *B. typhosus* by means of preparations of type II Vi phage. Canadian J. Public Health. **29**: 448-463.

Dartois, V., O. De Backer and C. Colson. 1993. Sequence of the *Salmonella typhimurium Sty*LTI restriction-modification genes. Homologies with *Eco*PI and *Eco*P15I type III R-M systems and presence of helicase domains. Gene **127**: 105-110.

Daniel, A.S., F.V. Fuller-Pace, D.M. Legge and N.E. Murray. 1988. Distribution and diversity of *hsd* genes in *Escherichia coli* and other enteric bacteria. J. Bacteriol. **170**: 1775-1782.

Davison, J. and F. Brunel. 1979a. Restriction insensitivity in bacteriophage T5. I. Genetic characterisation of mutants sensitive to *Eco*RI restriction. J. Virology. **29**: 11-16.

Davison, J. and F. Brunel. 1979b. Restriction insensitivity in bacteriophage T5. II. Lack of *Eco*RI modification in $T5^+$ and T5*ris* mutants. J. Virology. **29**: 17-20.

Day, R.S. 1977. UV-induced alleviation of K-specific restriction of bacteriophage λ. J. Virology. **21**: 1249-1251.

Debrouwere, L., M. Zabeau, M. Van Montagu and J. Schell. 1980a. The *ral* gene of phage λ. II. Isolation and characterisation of *ral* deficient mutants. Molec. Gen. Genet. **179**: 75-80.

Debrouwere, L., M. Van Montagu and J. Schell. 1980b. The *ral* gene of phage λ. III. Interference with *E. coli* ATP dependent functions. Molec. Gen. Genet. **179**: 81-88.

Dharmalingam, K. and E.B. Goldberg. 1976. Mechanism, localisation and control of restriction cleavage of phage T4 and λ chromosomes *in vivo*. Nature. **260**: 406-410.

Dharmalingam, K., H.R. Revel and E.B. Goldberg. 1982. Physical mapping and cloning of bacteriophage T4 anti-restriction endonuclease gene. J. Bacteriol. **149**: 694-699.

Diaz, R., P. Barnsley and R.H. Pritchard. 1979. Location and characterisation of a new replication origin in the *E. coli* K-12 chromosome. Molec. Gen. Genet. **175**: 151-157.

Dixon, D.A. and S.C. Kowalczykowski. 1993. The recombination hotspot χ is a regulatory sequence that acts by attenuating the nuclease activity of the *E. coli* RecBCD enzyme. Cell. **73**: 87-96.

Dixon, D.A., J.J. Churchill and S.C. Kowalczykowski. 1994. Reversible inactivation of the *E. coli* RecBCD enzyme by the recombination hotspot χ *in vitro*: Evidence for functional inactivation or loss of the RecD subunit. Proc. Natl. Acad. Sci. USA. **91**: 2980-2984.

Dryden, D.T.F., L.P. Cooper and N.E. Murray. 1993. Purification and characterisation of the methyltransferase from the type I restriction and modification system of *Escherichia coli* K-12. J. Biol. Chem. **268**: 13228-13236.

Dryden, D.T.F., S. Sturrock and M. Winter. 1995. Structural modelling of a type I DNA methyltransferase. Nature Struct. Biol. (in press).

Dunn, J.J. and F.W. Studier. 1983. The complete nucleotide sequence of bacteriophage T7 DNA, and the locations of T7 genetic elements. J. Mol. Biol. **166**: 477-535.

Dussoix, D. and W. Arber. 1962. Host specificity of DNA produced by *Escherichia coli*. II. Control over acceptance of DNA from infecting phage λ. J. Mol. Biol. **5**: 37-49.

Eddy, S.R. and L. Gold. 1992. Artificial mobile DNA element constructed from the *Eco*RI endonuclease gene. Proc. Natl. Acad. Sci. USA. **89**: 1544-1547.

Efimova, E.P., E.P. Delver and A.A. Belogurov. 1988a. 2-Aminopurine and 5-bromouracil induce alleviation of type I restriction in *Escherichia coli*: mismatches function as inducing signals? Molec. Gen. Genet. **214**: 317-320.

Efimova, E.P., E.P. Delver and A.A. Belogurov. 1988b. Alleviation of type I restriction in adenine methylase (*dam*) mutants of *Escherichia coli*. Molec. Gen. Genet. **214**: 313-316.

Endlich, B. and S. Linn. 1981. Type I restriction enzymes. *In* **The enzymes Vol. XIV**. P.D. Boyer (ed), Academic press, NY. pp. 137-156.

Eskin, B. and S. Linn. 1972a. The deoxyribonucleic acid modification and restriction enzymes of *Escherichia coli* B. II. Purification, subunit structure, and catalytic properties of the restriction endonuclease. J. Biol. Chem. **247**: 6183-6191.

Eskin, B. and S. Linn. 1972b. The deoxyribonucleic acid modification and restriction enzymes of *Escherichia coli* B. III Studies of the restriction adenosine triphosphatase. J. Biol. Chem. **247**: 6192-6196.

Eskridge, R.W., H. Weinfeld and K. Paigen. 1967. Susceptibility of different coliphage genomes to host-controlled variation. J. Bacteriol. **93**: 835-844.

Franklin, N. 1984. Conservation of genome form but not sequence in the transcription antitermination determinants of bacteriophages λ, ϕ21 and P22. J. Mol. Biol. **181**: 75-84.

Fuller-Pace, F.V., L.R. Bullas, H. Delius and N.E. Murray. 1984. Genetic recombination can generate altered restriction specificity. Proc. Natl. Acad. Sci. USA. **81**: 6095-6099.

Fürste, J.P., W. Pansegrau, R. Frank, H. Blocker, P. Scholz, M. Bagdasarian and E. Lanka. 1986. Molecular cloning of the plasmid RP4 primase region in a multi-host-range *tacP* expression vector. Gene. **48**: 119-131.

Furth, M.E. and S.H. Wickner. 1983. DNA replication. *In* **Lambda II**. R.W. Hendrix *et al.* (eds), Cold Spring Harbor Laboratory Press. NY. pp. 145-173.

Gann, A.A.F., A.J.B. Campbell, J.F. Collins, A.F.W. Coulson and N.E. Murray. 1987. Reassortment of DNA recognition domains and the evolution of new species. Mol. Microbiol. **1**: 13-22.

Gillen, J.R., A.E. Karu, H. Nagaishi and A.J. Clark. 1977. Characterisation of the deoxyribonuclease determined by lambda reverse as exonuclease VIII of *E. coli*. J. Mol. Biol. **113**: 27-41.

Glover, S.W. 1970. Functional analysis of host specificity mutants in *Escherichia coli*. Genet. Res. Camb. **15**: 237-250.

Glover, S.W. and C. Colson. 1969. Genetics of the host-controlled restriction and modification in *Escherichia coli*. Genet. Res. Camb. **13**: 227-240.

Gnann, J.W., L.L. Smith and M.B.A. Oldstone. 1989. Custom-designing synthetic peptide immunoassays for distinguishing HIV type 1 and type 2 infections. Methods Enzymol. **178**: 693-714.

Gold, L. 1988. Posttranscriptional regulatory mechanisms in *E. coli*. Ann. Rev. Biochem. **57**: 199-233.

Gorbalenya, A.E. and E.V. Koonin. 1991. Endonuclease R subunits of type I and type III restriction-modification enzymes contain a helicase-like domain. FEBS Lett. **291**: 277-281.

Gottesman, M.M., M.E. Gottesman, S. Gottesman and M. Gellert. 1974. Characterisation of bacteriophage λ*reverse* as an *E. coli* phage carrying a unique set of host-derived recombination functions. J. Mol. Biol. **88**: 471-487.

Gough, J.A. and N.E. Murray. 1983. Sequence diversity among related genes for recognition of specific targets in DNA molecules. J. Mol. Biol. **166**: 1-19.

Gubler, T. and T.A. Bickle. 1991. Increased protein flexibility leads to promiscuous protein-DNA interactions in type IC restriction-modification systems. EMBO J. **10**: 951-957.

Gubler, T., M., D. Braguglia, J. Meyer, A. Piekarowitcz and T.A. Bickle. 1992. Recombination of constant and variable modules alters DNA sequence recognition by type IC restriction-modification enzymes. EMBO J. **11**: 233-240.

Haberman, A. 1974. The bacteriophage P1 restriction endonuclease. J. Mol. Biol. **89**; 545-563.

Hadi, S.M. and R. Yuan. 1974. Complimentation *in vitro* by mutant restriction enzymes from *Escherichia coli* K. J. Biol. Chem. **249**: 4580-4586.

Hadi, S.M., T.A. Bickle and R. Yuan. 1975. The role of S-Adenosylmethionine in the cleavage of deoxyribonucleic acid by the restriction endonuclease from *Escherichia coli* K. J. Biol. Chem. **250**: 4159-4164.

Harlow, E. and D. Lane. 1988. Immunoaffinity purification of antibodies on an antigen column. *In* **Antibodies: a laboratory manual**. Cold Spring Harbor Laboratory Press, NY. pp. 313-315.

Harris, D.A. 1987. Spectrophotometric assays. *In* **Spectrophotometry and Spectrofluorimetry**. D.A. Harris and C.L. Bushford (eds). IRL Press, Oxford. pp. 49-90.

Hastings, J.R.B. and K.S. Kirby. 1966. The nucleic acids of *Drosophila melanogaster*. Biochem. J. **100**: 532-539.

Hattman, S. 1970. DNA methylation of T-even bacteriophages and of their nonglucosylated mutants: its role in P1-directed restriction. Virology. **42**: 359-367.

Hattman, S. 1982. DNA methyltransferase-dependent transcription of the phage Mu *mom* gene. Proc. Natl. Acad. Sci. USA. **79**: 5518-5521.

He, M., A. Wilde and M.A. Kaderbhai. 1990. A simple single-step procedure for small-scale preparation of *E. coli* plasmids. Nucleic Acids Res. **18**: 1660.

Heery, D.M. and L.K. Dunican. 1989. Improved efficiency M13 cloning using electroporation. Nucleic Acids Res. **17:** 8006.

Hegna, I.K., E.S. Karlstrϕm, R. Lopez, T. Kristensen and A. Kolstϕ. 1992. A type-III DNA restriction and modification system in *B. cereus*. Gene. **114**: 149-150.

Heip, J., B. Rolfe and J. Schell. 1974. Abolition of host cell restriction by high multiplicity of phage infection. Virology. **59**: 356-370.

Heitman, J. and P. Model. 1987. Site-specific methylases induce the SOS DNA repair response in *E. coli*. J. Bacteriol. **169**: 3243-3250.

Heitman, J., J. Treisman, N.G. Davis and M. Russel. 1989. Cassettes of the f1 intergenic region. Nucleic Acids Res. **17**: 4413.

Hershey, J.W.B. 1987. Protein Synthesis. *In* ***Escherichia coli*** **and** ***Salmonella typhimurium*****: Cellular and Molecular Biology**. F.C. Neidhardt *et al.*, (eds). ASM, Washington DC. pp. 613-647.

Highton, P.J., Y. Chang and R.J. Myers. 1990. Evidence for the exchange of segments between genomes during the evolution of lambdoid bacteriophages. Mol. Microbiol. **4**: 1329-1340.

Higuchi, R. 1990. Recombinant PCR. *In* **PCR Protocols.** M.A. Innis *et al.* (eds), Academic Press ltd. pp. 177-196.

Hill, C., L.A. Miller and T.R. Klaenhammer. 1991. *In vivo* genetic exchange of a functional domain from a type II methylase between lactococcal plasmid pTR2030 and a virulent bacteriophage. J. Bacteriol. **173**: 4363-4370.

Hiom, K. and S.G. Sedgwick. 1991. Simple methods for making *Eco*K and McrA restrictionless mutants. Nucleic Acids Res. **19**:2502.

Holmes, D.S. and M. Quigley. 1981. A rapid boiling method for the preparation of bacterial plasmids. Anal. Biochem. **114**: 193-197.

Hubacek, J. and S.W. Glover. 1970. Complimentation analysis of temperature sensitive host specific mutations in *Escherichia coli*. J. Mol. Biol. **50**: 111-127.

Humbelin, M., B. Suri, D.N. Rao, D.P. Hornby, H. Eberle, T. Pripfl, S. Kenel and T.A. Bickle. 1988. Type III DNA *res*triction and *mod*ification systems *Eco*P1 and *Eco*P15. J. Mol. Biol. **200**: 23-29.

Iida, S., J. Meyer, B. Bächi, M. Stalhammer-Carlemalm, S. Schrickel, T.A. Bickle and W. Arber. 1983. DNA restriction-modification genes of the phage P1 and plasmid p15B. J. Mol. Biol. **165**: 1-18.

Iida, S., M.B. Streiff, T.A. Bickle and W. Arber. 1987. Two DNA antirestriction systems of bacteriophage P1, *darA*, and *darB*: characterisation of *darA*$^-$ phages. Virology. **157**: 156-166.

Ineichen, K., J.C.W. Shephard and T.A. Bickle. 1981. The DNA sequence of the phage lambda genome between P_L and the gene *bet*. Nucleic Acids Res. **9**: 4639-4653.

Janulaitis, A., R. Kazlauskiene, L. Lazareviciute, R. Gilvonauskaite, D. Steponaviciene, M. Jagelavicius, M. Petrusyte, J. Bitinaite, Z. Vezeviciute, E. Kiuduliene and V. Butkus. 1988. Taxonomic specificity of restriction-modification enzymes. Gene. **74**: 229-232.

Janulaitis, A., M. Petrusyte, Z. Maneliene, S. Klimasauskas and V. Butkus. 1992. Purification and properties of the *Eco*57I restriction endonuclease and methylase - prototypes of a new class (type-IV). Nucleic Acids Res. **20**: 6043-6049.

Kahmann, R. 1984. The *mom* gene of bacteriophage Mu. Curr. Topics Microbiol. Immunol. **108**: 29-47.

Kahmann, R., A. Seiler, F.G. Wulczyn and E. Pfaff. 1985. The *mom* gene of bacteriophage Mu: a unique regulatory scheme to control a lethal function. Gene. **39**: 61-70.

Kaiser, K. and N.E. Murray. 1979. Physical characterisation of the "Rac prophage" in *E. coli* K-12. Molec. Gen. Genet. **175**: 159-174.

Kan, N.C., J.A. Lautenberger, M.E. Edgell and C.A. Hutchison III. 1979. The nucleotide sequence recognised by the *Escherichia coli* K12 restriction and modification enzymes. J. Mol. Biol. **130**: 191-209.

Kannan, P.R., G.M. Cowan, A.S. Daniel, A.A.F. Gann and N.E. Murray. 1989. Conservation of organisation in the specificity polypeptides of two families of type I restriction enzymes. J. Mol. Biol. **209**: 335-344.

Karlin, S., C. Burge and A.M. Campbell. 1992. Statistical analysis of counts and distributions of restriction sites in DNA sequences. Nucleic Acids Res. **20**: 1363-1370.

Kauc, L. and A. Piekarowicz. 1978. Purification and properties of a new restriction endonuclease from *Haemophilus influenzae*. Eur. J. Biochem. **92**: 417-426.

Kelleher, J.E. and E.A. Raleigh. 1991. A novel activity in *E. coli* K-12 that directs restriction of DNA modified at CG nucleotides. J. Bacteriol. **173**: 5220-5223.

Kelleher, J.E., A.S. Daniel and N.E. Murray. 1991. Mutations that confer *de novo* activity upon a maintenance methylase. J. Mol. Biol. **221**: 431-440.

Kelleher, J.E. and E.A. Raleigh. 1994. Response to UV damage be four *E. coli* K-12 restriction systems. J. Bacteriol. **176**: 5888-5896.

Kelly, T.J. and H.O. Smith. 1970. A restriction enzyme from *H. influenzae*. II. Base sequence of the recognition site. J. Mol. Biol. **51**: 393-409.

Kessler, C. and V. Manta. 1990. Specificity of restriction endonucleases and DNA modification methyltransferases - a review. Gene. **92**: 1-248.

Kim, Y., J.C. Grable, R. Love, P.J. Greene and J.M. Rosenberg. 1990. Refinement of *Eco*RI endonuclease crystal structure: a revised protein chain tracing. Science. **249**. 1307-1309.

King, G. and N.E. Murray. 1994. Restriction enzymes in cells, not Eppendorfs. Trends Microbiol. **2**: 465-469.

King, G. and N.E. Murray. 1995. Restriction alleviation and modification enhancement by the *E. coli* Rac prophage. Mol. Microbiol. **16**: 769-777.

Klimasauskas, S., A. Timinskas, S. Menkevicius, D. Butkiene, V. Butkus and A. Janulaitis. 1989. Sequence motifs characteristic of DNA [cytosine-N4] methyltransferases: similarity to adenine and cytosine-C5 DNA-methylases. Nucleic Acids Res. **17**: 9823-9832.

Klimasauskas, S., S. Kumar, R.J. Roberts and X. Cheng. 1994. *Hha*I methyltransferase flips its target base out of the DNA helix. Cell. **76**: 357-369.

Kneale, G.G. 1994. A symmetrical model for the domain structure of type I DNA methyltransferases. J. Mol. Biol. **243**: 1-5.

Kolodner, R., S.D. Hall and C. Luisideluca. 1994. Homologous pairing proteins encoded by the *E. coli recE* and *recT* genes. Mol. Microbiol. **11**: 23-30.

Kong, H., S.E. Roemer, P.A. Waite-Rees, J.S. Benner, G.G. Wilson and D.O. Nwankwo. 1994 Characterisation of *Bcg*I, a new kind of restriction-modification system. J. Biol. Chem. **269**: 683-690.

Korona, R. and B.R. Levin. 1993. Phage-mediated selection and the evolution and maintenance of restriction-modification. Evolution. **47**: 565-575.

Korona, R., B. Korona and B.R. Levin. 1993. Sensitivity of naturally occurring coliphage to type-I and type-II restriction and modification. J. Gen. Microbiol. **139**: 1283-1290.

Krüger, D.H., C. Schroeder, S. Hansen and H.A. Rosenthal. 1977. Active protection by bacteriophage T3 and T7 against *E. coli* B and K-specific restriction of their DNA. Molec. Gen. Genet. **153**: 99-106.

Krüger, D.H., M. Reuter, S. Hansen and C. Schroeder. 1982. Influence of phage T3 and T7 gene functions on a type III (*Eco*P1) DNA restriction-modification system *in vivo*. Molec. Gen. Genet. **185**: 457-461.

Krüger, D.H. and T.A. Bickle. 1983. Bacteriophage survival: multiple mechanisms for avoiding the deoxyribonucleic acid restriction systems of their hosts. Microbiol. Rev. **47**: 345-360.

Krüger, D.H., C. Schroeder, M. Santibanez-Koref and M. Reuter. 1989. Avoidance of DNA methylation. Cell. Biophysics. **15**: 87-95.

Krüger, D.H., T.A. Bickle, M. Reuter, C.D. Pein and C. Schroeder. 1990. DNA methylation and restriction processes in *E. coli*: insights by use of bacterial viruses T3 and T7. *In* **Nucleic acid methylation**. G.A. Clawson *et al.* (eds), Alan R. Liss Inc. NY. pp. 113-124.

Kühnlein, U., S. Linn and W. Arber. 1969. Host specificity of DNA produced by *Escherichia coli*. XI. In vitro modification of phage fd replicative form. Proc. Natl. Acad. Sci. USA. **63**: 556-567.

Kühnlein, U. and W. Arber. 1972. Host specificity of DNA produced by *Escherichia coli*. XII. The role of nucleotide methylation in *in vitro* B-specific modification. J. Mol. Biol. **63**: 9-19.

Kushner, S.R., H. Nagaishi and A.J. Clark. 1974. Isolation of exonuclease VIII: the enzyme associated with the *sbcA* indirect suppressor. Proc. Natl. Acad. Sci. USA. **71**: 3593-3597.

Kuzminov, A. E. Schabtach and F.W. Stahl. 1994. χ sites in combination with RecA protein increase the survival of linear DNA in *E. coli* by inactivating exoV activity of RecBCD nuclease. EMBO J. **13**: 2764-2776.

Labahn, J., J. Granzin, G. Schluckebier, D.P. Robinson, W.E. Jack, I. Schildkraut and W. Saenger. 1994. Three dimensional structure of the adenine specific DNA methyltransferase M.*Taq*I in complex with the cofactor S-adenosylmethionine. Proc. Natl. Acad. Sci. USA. **91**: 10957-10961.

Laemmli, U.K. 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature. **227**: 680-685.

Lange, C., A. Jugel, J. Walter, M. Noyer-Weidner and T.A. Trautner. 1991. 'Pseudo' domains in phage-encoded DNA methyltransferases. Nature. **352**: 645-648.

Lautenberger, J.A. and S. Linn. 1972. The deoxyribonucleic acid modification and restriction enzymes of *Escherichia coli* B. J. Biol. Chem. **247**: 6176-6182.

Lautenberger, J.A., N.C. Kan, D. Lackey, S. Linn, M.H. Edgell and C.A. Hutchinson III. 1978. Recognition site of *Escherichia coli* B restriction enzyme on ϕXsB1 and simian virus 40 DNAs: An interrupted sequence. Proc. Natl. Acad. Sci. USA. **75**: 2271-2275.

Lederberg, S. 1957. Suppression of the multiplication of heterologous bacteriophage in lysogenic bacteria. Virology. **3**: 496-513.

Lennox, E.S. 1955. Transduction of linked genetic characters of the host by bacteriophage P1. Virology. **1**: 190-206.

Leung, D.W., E. Chen and D.V. Goeddel. 1989. A method for random mutagenesis of a defined DNA segment using a modified polymerase chain reaction. Technique. **1**: 11-15.

Levin, B.R. 1986. Restriction-modification immunity and the maintenance of genetic diversity in bacterial populations. *In* **Evolutionary Processes and Theory**. S. Karlin and E. Nero (eds), Academic Press Inc. NY. pp. 669-688.

Levin, B.R. 1988. Frequency-dependent selection in bacterial populations. Philos. Trans. Royal Soc. London. **319**: 459-472.

Levin, B.R. and R.E. Lenski. 1985. Bacteria and phage: a model system for the study of the ecology and co-evolution of hosts and parasites. *In* **Ecology and genetics of host-parasite interactions**. D. Rollison and R.M. Anderson (eds), Academic Press, NY. pp. 227-242.

Levitz, R., D. Chapman, M. Amitsur, R. Green, L. Snyder and G. Kaufmann. 1990. the optional *E. coli prr* locus encodes a latent form of phage T4-induced anticodon nuclease. EMBO J. **9**: 1383-1389.

Linn, S. and W. Arber. 1968. Host specificity of DNA produced by *Escherichia coli.* In vitro restriction of phage fd replicative form. Proc. Natl. Acad. Sci. USA. **59**: 1300-1306.

Loenen, W.A.M. and N.E. Murray. 1986. Modification enhancement by the restriction alleviation protein (Ral) of bacteriophage λ. J. Mol. Biol. **190**: 11-22.

Loenen, W.A.M., A.S. Daniel, H.D. Braymer and N.E. Murray. 1987. Organisation and sequence of the *hsd* genes of *Escherichia coli* K-12. J. Mol. Biol. **198**: 159-170.

Low, K.B. 1973. Restoration of the *rac* locus of recombinant forming ability in *recB*⁻ and *recC*⁻ merozygotes of *E. coli* K-12. Mol. Gen. Genetic **122**: 119-130.

Lunnen, K.D., J.M. Barsomian, R.R. Camp, C.O. Card, S. Chen, R. Croft, M.C. Looney, M.M. Meda, L.S. Moran, D.O. Nwankwo, B.E. Slatko, E.M. Van Cott and G.G. Wilson. 1988. Cloning type II restriction and modification genes. Gene. **74**: 25-32.

Luria, S.E. and M.L. Human. 1952. A non-hereditary, host-induced variation of bacterial viruses. J. Bacteriol. **64**: 557-569.

Maizel, J.V. and R.P. Lenk. 1981. Enhanced graphic matrix analysis of nucleic acid and protein sequences. Proc. Natl. Acad. Sci. USA. **78**: 7665-7669.

Makino, O., F. Kawamura, H. Saito and Y. Ikeda. 1979. Inactivation of restriction endonucleases *Bam*Nx after infection with phage ϕNR2. Nature. **277**: 64-66.

Makino, O., H. Saito and T. Ando. 1980. *Bacillus subtilis* phage ϕ1 overcomes host-controlled restriction by producing *Bam*Nx inhibitor protein. Molec. Gen. Genet. **179**: 463-468.

Manly, K.F., E.R. Signer and C.M. Radding. 1969. Nonessential functions of bacteriophage λ. Virology. **37**: 177-188.

Marinus, M.G. 1987. DNA methylation in *E. coli.* Ann. Rev. Genet. **21**: 113-131.

McCarthy, J.E.G. and C. Gualerzi. 1990. Translational control of prokaryotic gene expression Trends Genet. **6:** 78-85.

McKane, M. and R. Milkman. 1995. Transduction, restriction and recombination patterns in *E. coli*. Genetics. **139**: 35-43.

Meisel, A., T.A. Bickle, D.H. Krüger and C. Schroeder. 1992. Type III restriction enzymes need two inversely orientated recognition sites for DNA cleavage. Nature. **355**: 467-469.

Meisel, A., P. Mackeldanz, T.A. Bickle, D.H. Krüger and C. Schroeder. 1995. Type-III restriction endonucleases translocate DNA in a reaction driven by recognition site-specific ATP hydrolysis. EMBO J. **14**: 2958-2966.

Meister, J., M. MacWilliams, P. Hubner, H. Jutte, E. Skrzypek, A. Piekarowicz and T.A Bickle. 1993. Macroevolution by transposition: Drastic modification of DNA recognition by a type I restriction enzyme following Tn5 transposition. EMBO J. **12**: 4585-4591.

Meselson, M. and R. Yuan. 1968. DNA Restriction enzyme from *E.coli*. Nature. **217**: 1110-1114.

Meselson, M., R. Yuan and J. Heywood. 1972. Restriction and modification of DNA. Ann. Rev. Biochem. **41**: 447-466.

Messing, J. 1983. New M13 vectors for cloning. Methods Enzymol. **101**: 20-78.

Meyer, R.R. and P.S. Laine. 1990. The single-strand DNA-binding protein of *E. coli*. Microbiol. Rev. **54**: 342-380.

Midgley, C.A. and N.E. Murray. 1985. T4 polynucleotide kinase; cloning of the gene (*pse*T) and amplification of its product. EMBO J. : 2695-2703.

Mihalyi, E. 1970. Numeric values of the absorbances of the aromatic amino acids in acid, neutral and alkaline solutions. *In* **Handbook of Biochemistry**. H.A. Sober (ed). CRC Press, Ohio. pp. B75-B77.

Milkman, R. and M. McKane-Bridges. 1993. Molecular evolution of the *E. coli* chromosome. IV. Sequence comparisons. Genetics. **133**: 455-468.

Miller, J.H. 1992. P1 transduction. *In* **A short course in bacterial genetics**. Cold Spring Harbor Laboratory Press, NY. pp. 263-278

Mizusawa, S. and D.F. Ward. 1982. A bacteriophage λ vector for cloning with *Bam*HI and *Sau*3A. Gene. **20**: 317-322.

Moffat, B.A. and F.W. Studier. 1988. Entry of bacteriophage T7 DNA into the cell and escape from host restriction. J. Bacteriol. **170**: 2095-2105.

Morad, I., D. Chapman-Shimshoni, M. Amitsur and G. Kaufmann. 1993. Functional expression and properties of the $tRNA^{Lys}$-specific anticodon nuclease encoded by *E. coli prrC*. J. Biol. Chem. **268**: 26842-26849.

Murray, N.E. 1991. Special uses of λ phage for molecular cloning. Methods Enzymol. **204**: 280-301.

Murray, N.E., P.L. Batten and K. Murray. 1973. Restriction of bacteriophage λ by *Escherichia coli* K. J. Mol. Biol. **81**: 395-407.

Murray, N.E., S.A. Bruce and K. Murray. 1979. Molecular cloning of the DNA ligase gene from bacteriophage T4. II. Amplification and preparation of the gene product. J. Mol. Biol. **132**: 493-505.

Murray, N.E., J.A. Gough, B. Suri and T.A. Bickle. 1982. Structural homologies among type I restriction-modification systems. EMBO J. **1**: 535-539.

Murray, N.E., A.S. Daniel, G.M. Cowan and P.M. Sharp. 1993. Conservation of motifs within the unusually variable polypeptide sequences of the type I restriction and modification enzymes. Mol. Microbiol. **9**: 133-143.

Nagaraja, V., M. Stieger, C. Nager, S.M. Hadi and T.A. Bickle. 1985a. The nucleotide sequence recognised by *Eco*D type I restriction and modification enzyme. Nucleic Acids Res. **13**: 389-399.

Nagaraja, V., J.C.W. Shepherd, T. Pripfl and T.A. Bickle. 1985b. Two type I restriction enzymes from *Salmonella* species. Purification and DNA sequences. J. Mol. Biol. **182**: 579-587.

Naito, T., K. Kusano and I. Kobayashi. 1995. Selfish behaviour of restriction-modification systems. Science. **267**: 897-899.

Nathan, P.D. and J.E. Brooks. 1988. Characterisation of clones of the *Bam*HI methyltransferase gene. Gene. **74**: 35-36.

Needleman, S.B. and C.D. Wunsch. 1970. A general method applicable to the search for similarities in the amino acid sequence of two proteins. J. Mol. Biol. **48**: 443-453.

Newman, M., T. Strzelecka, L.F. Dorner, I. Schildkraut and A.K. Aggarwal. 1994. Structure of restriction endonuclease *Bam*HI and its relationship to *Eco*RI. Nature. **368**: 660-664.

Nikolskaya, I.I., M.I. Tediashvili, N.G. Lopatina, T.G. Chanishvili and S.S. Debov. 1979. Specificity and functions of guanine methylase of *Shigella sonnei* DDVI phage. Biochim. Biophys. Acta. **561**: 232-239.

Noyer-Weidner, M., B. Pawlek, S. Jentsch, U. Günthert and T.A. Trautner. 1981. Restriction and modification in *Bacillus subtilis*: gene coding for *Bsr*R-specific modification methyltransferase in the temperate bacteriophage ϕ3T. J. Virology. **38**: 1077-1080

Noyer-Weidner, M., R. Diaz and L. Reiners. 1986. Cytosine specific DNA modification interferes with plasmid establishment in *E. coli* K-12: involvement of *rglB*. Molec. Gen. Genet. **205**: 469-475.

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

Old, R., K. Murray and G. Roizes. 1975. Recognition sequence of restriction endonuclease III from *Haemophilus influenzae*. J. Mol. Biol. **92**: 331-339.

Pao, C.C. and J.F. Speyer. 1973. Order of injection of T7 bacteriophage DNA . J. Virology. **11**: 1024-1026.

Pause, A. and N. Sonenberg. 1992. Mutational analysis of a DEAD-box RNA helicase - the mammalian translation initiation-factor eIF-4A. EMBO J. **11**: 2643-2654.

Piekarowicz, A. and J. Kalinowska. 1974. Host specificity of DNA in *Haemophilus influenzae*: similarity between host-specificity types of *Haemophilus influenzae* Re and Rf. J. Gen. Microbiol. **81**: 405-411.

Piekarowicz, A. and J.D. Goguen. 1986. The DNA sequence recognised by *Eco*DXXI restriction endonuclease. Eur. J. Biochem. **154**: 295-298.

Pittard, J. 1964. Effect of phage-controlled restriction on genetic linkage in bacterial crosses. J. Bacteriol. **87**: 1256-1257.

Pósfai, J., A.S. Bhagwat, G. Pósfai and R.J. Roberts. 1989. Predictive motifs derived from cytosine methyltransferases. Nucleic Acids Res. **17**: 2421-2435.

Powell, L.M., D.T.F. Dryden, D.F. Willcock, R.H. Pain. and N.E. Murray. 1993. DNA recognition by the *Eco*K methyltransferase. The influence of DNA methylation and the cofactor S-adenosyl methionine. J. Mol. Biol. **234**: 60-71.

Powell, L.M. and N.E. Murray. 1995. S-adenosyl methionine alters the DNA contacts of the *Eco*KI methyltransferase. Nucleic Acids Res. **23**: 967-974.

Prakash-Cheng, A., B. Valinluck and J. Ryu. 1991. Genomic *hsd*-Mu(*lac*) operon fusion mutants of *E.coli* K-12. Gene. **99**: 9-14.

Prakash-Cheng, A. and J. Ryu. 1993. Delayed expression of in vivo restriction activity following conjugal transfer of *E.coli* hsd_K (restriction-modification) genes. J. Bacteriol. **175**: 4905-4906.

Prakash-Cheng, A., S.S. Chung and J. Ryu. 1993. The expression and regulation of hsd_K genes after conjugative transfer. Molec. Gen. Genet. **241**: 491-496.

Price, C. and T.A. Bickle. 1986. A possible role for DNA restriction in bacterial evolution. Microbiol. Sci. **3**: 296-299.

Price, C., T. Pripfl T.A. Bickle. 1987a. *Eco*R124 and *Eco*R124/3. The first members of a new family of type I restriction and modification systems. Eur. J. Biochem. **167**: 111-115.

Price, C., J.C.W. Shepherd and T.A. Bickle. 1987b. DNA sequence recognised by a new family of type I restriction enzymes: A unique relationship between two different DNA specificities. EMBO J. **6**: 1493-1497.

Price, C., J. Linger, T.A. Bickle, K. Firman and S.W. Glover. 1989. Basis for changes in DNA recognition by *Eco*R124 and *Eco*R124/3 type I DNA restriction and modification enzymes. J. Mol. Biol. **205**: 115-125.

Queen, C. 1983. A vector that uses phage signals for efficient synthesis of proteins in *E. coli.* J. Mol. Appl. Genet. **2**: 1-10.

Raleigh, E.A. 1992. Organisation and function of the *mcrBC* genes of *Escherichia coli* K-12. Mol. Microbiol. **6**: 1079-1086.

Raleigh, E.A. and G.G. Wilson. 1986. *E.coli* K-12 restricts DNA containing 5-methylcytosine. Proc. Natl. Acad. Sci. USA. **83**: 9070-9074.

Raleigh, E.A., R. Trimarchi and H. Revel. 1989. Genetic and physical mapping of the *mcrA* (*rglA*) and *mcrB* (*rglB*) loci of *E. coli* K-12. Genetics. **122**: 279-296.

Ravetch, J.V., K. Horiuchi and N.D. Zinder. 1978. Nucleotide sequence of the restriction site for the restriction-modification enzyme of *Escherichia coli* B. Proc. Natl. Acad. Sci. USA. **75**: 2266-2270.

Revel, H.R. 1967. Restriction of non-glucosylated T-even bacteriophage: properties of permissive mutants of *E. coli* B and K-12. Virology. **31**: 688-701.

Revel, H.R. 1983. DNA modification: glucosylation. *In* **Bacteriophage T4**. C.K. Mathews *et al.* (eds), ASM, Washington DC. pp. 156-165.

Revel, H.R. and S.E. Luria. 1970. DNA-glucosylation in T-even phage: Genetic determination and role in phage-host interaction. Annu. Rev. Genet. **4**: 177-192.

Roberts, J.W. and C.W. Roberts. 1975. Proteolytic cleavage of bacteriophage lambda repressor in induction. Proc. Natl. Acad. Sci. USA. **72**: 147-151.

Roberts, R.G. 1976. Restriction endonucleases. CRC Crit. Rev. Biochem. **4**: 123-164.

Roberts, R.J. and S.E. Halford. 1993. Type II restriction endonucleases. *In* **Nucleases**. S.M. Linn *et al.* (eds), Cold Spring Harbor Laboratory Press, NY. pp. 35-88.

Roberts, R.J. and D. Macelis. 1993. Restriction enzymes and their isoschizomers. Nucleic Acids Res. **21**: 3125-3137.

Romero, A. and P. Garcia. 1991. Initiation of translation at AUC, AUA and AUU codons in *Escherichia coli*. FEMS Microbiol. Lett. **84:** 325-330.

Rosamond, J., B. Endlich and S. Linn. 1979. Electron microscopic studies of the mechanism of action of the restriction endonuclease of *Escherichia coli* B. J. Mol. Biol. **129**. 619-635.

Roulland-Dussoix, D. and H.W. Boyer. 1969. The *Escherichia coli* B restriction endonuclease. Biochim. Biophys. Acta. **195**: 219-229.

Saha, S and D.N. Rao. 1995. ATP hydrolysis is required for DNA cleavage by *Eco*PI restriction enzyme. J. Mol. Biol. **247**: 559-567.

Saiki, S.J., D.H. Gelfand, S. Stoffel, S.J. Scharf, R. Higuchi, G.T. Horn, K.B. Mullis and H.A. Erlich. 1988. Primer-directed enzymatic amplification of DNA with a thermostable DNA-polymerase. Science. **239**: 487-491.

Sain, B. and N.E. Murray. 1980. The *hsd* genes of *E.coli* K12. Molec. Gen. Genet. **180**: 35-46.

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

Sanger, F., A.R. Coulson, G.F. Hong, D.F. Hill and G.B. Peterson. 1982. Nucleotide sequence of bacteriophage λ DNA. J. Mol. Biol. **162**: 729-773.

Schauder, B., H. Blocker, R. Frank and J.E.G. McCarthy. 1987. Inducible expression vectors incorporating the *Escherichia coli atpE* translational initiation region. Gene. **52**: 279-283.

Schroeder, C., H. Jurkschat, A. Meisel, J. Reich and D. Krüger. 1986. Unusual occurrence of *Eco*P1 and *Eco*P15 recognition sites and counterselection of type-II methylation and restriction sequences in bacteriophage T7 DNA. Gene. **45**: 77-86.

Scopes, R.K. 1988. *In* **Protein Purification: Principles and Practice** (second edition). Springer-Verlag.

Semerjian, A.V., D.C. Mallory and A.R. Poteete. 1989. Genetic structure of the bacteriophage P22 P_L Operon. J. Mol. Biol. **207**: 1-13.

Shägger, H. and G. von Jagow. 1987. Tricine-sodium dodecyl sulphate-polyacrylamide gel electrophoresis for the separation of proteins in the range from 1 to 100 kDa. Anal. Biochem. **166:** 368-379.

Sharp, P.M. 1986. Molecular evolution of bacteriophage: Evidence of selection against the recognition sites of host restriction enzymes. Mol. Biol. Evol. **3**: 75-83.

Sharp, P.M., J.E. Kelleher, A.S. Daniel, G.M. Cowan and N.E. Murray. 1992. Roles of selection and recombination in the evolution of type I restriction-modification systems in enterobacteria. Proc. Natl. Acad. Sci. USA. **89**: 9836-9840.

Sharrocks, A.D. and D.P. Hornby. 1991. Transcriptional analysis of the restriction and modification genes of bacteriophage P1. Mol. Microbiol. **5**: 685-694.

Siksnys, V., N. Zareckaja, R. Vaisvila, A. Timinskas, P. Stakenas, V. Butkas and A. Janulaitis. 1994. CAATTG-specific restriction-modification *mun*I genes from *Mycoplasma*: sequence similarities between R.*Mun*I and R.*Eco*RI. Gene. **142**: 1-8.

Simmon, V.F. and S. Lederberg. 1972. Degradation of bacteriophage lambda deoxyribonucleic acid after restriction by *Escherichia coli* K-12. J. Bacteriol. **112**:161-169.

Sironi, G. 1969. Mutants of *E. coli* unable to be lysogenised by the temperate bacteriophage P2. Virology. **37**: 163-176.

Smith, D.B. and K.S. Johnson. 1988. Single-step purification of polypeptides expressed in *E. coli* as fusions with glutathione S-transferase. Gene. **67**: 31-40.

Smith, G.R. 1988. Homologous recombination in prokaryotes. Microbiol. Rev. **52**: 1-28.

Smith, G.R., S.M. Kunes, D.W. Schultz, A. Taylor and K.L. Triman. 1981. Structure of Chi hotspots of generalised recombination. Cell. **24**: 429-436.

Smith, H.O. and K.W. Wilcox. 1970. A restriction enzyme from *Hemophilus influenzae*. J. Mol. Biol. **51**: 379-391.

Smith, H.O. and D. Nathans. 1973. A suggested nomenclature for bacterial host modification and restriction systems and their enzymes. J. Mol. Biol. **81**: 419-423.

Snyder, L. 1995. Phage-exclusion enzymes: a bonanza of biochemical and cell biology reagents? Mol. Microbiol. **15**: 415-420.

Spee, J.H., W.M. de Vos and O.P. Kuipers. 1993. Efficient random mutagenesis method with adjustable mutation frequency by the use of PCR and dITP. Nucleic Acids Res. **21**: 777-778.

Spoerel, N., P. Herrlich and T.A. Bickle. 1979. A novel bacteriophage defence mechanism: the anti-restriction protein. Nature. **278**: 30-34.

Stahl, M.M., I. Kobyashi, F.W. Stahl and S.K. Huntington. 1983. Activation of Chi, a recombinator, by the action of an endonuclease at a distant site. Proc. Natl. Acad. Sci. USA. **80**: 2310-2313.

Streiff, M.B., S. Iida and T.A. Bickle. 1987. Expression and proteolytic processing of the *darA* antirestriction gene product of bacteriophage P1. Virology. **157**: 167-171.

Studier, F.W. 1972. Bacteriophage T7. Science. **176**: 367-376.

Studier, F.W. 1975. Gene *0.3* of bacteriophage T7 acts to overcome the DNA restriction system of the host. J. Mol. Biol. **94**: 283-295.

Studier, F.W. and N.R. Movva. 1976. SAMase gene of bacteriophage T3 is responsible for overcoming host restriction. J. Virology. **19**: 136-145.

Studier, F.W. and B.A. Moffatt. 1986. Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. J. Mol. Biol. **189**: 113-130.

Studier, F.W. and P.K. Bandyopadhyay. 1988. Model of how type I restriction enzymes select cleavage sites in DNA. Proc. Natl. Acad. Sci. USA. **85**: 4677-4681.

Suri, B., V. Nagaraja and T.A. Bickle. 1984a. Bacterial DNA modification. Curr. Top. Microbiol. Immunol. **108**: 1-9.

Suri, B., J.C.W. Shepherd and T.A. Bickle. 1984b. The *Eco*A restriction and modification system of *Eschericia coli* 15T⁻: enzyme structure and DNA recognition sequence. EMBO J. **3**:575-579.

Suri, B. and T.A. Bickle. 1985. *Eco*A: The first member of a new family of type I restriction and modification systems. J. Mol. Biol. **186**; 77-85.

Swinton, D., S. Hattman, P.F. Crain, C.-S. Cheng, D.L. Smith and J.A. McCloskey. 1983. Purification and characterisation of the unusual deoxynucleotide, α-*N*-(9-β-D-2'-deoxy-ribofuranosylpurin-6-yl) glycinamide, specified by the phage Mu modification function. Proc. Natl. Acad. Sci. USA. **80**: 7400-7404.

Szybalski, W., S.C. Kim, N. Hasan and A.J. Podhajska. 1991. Class-IIs restriction enzymes-a review. Gene. **100**: 13-26.

Tabor, S. 1990. Expression using the T7 RNA polymerase/promoter system. *In* **Current protocols in molecular biology.** F.A. Ausubel *et al.* (eds), Greene publishing and Wiley-Interscience, NY. pp. 16.2.1

Takano, T., T. Watanabe and T. Fukasawa. 1966. Specific inactivation of infectious λ DNA by sonicates of restrictive bacteria with R factors. Biochem. Biophys. Res. Comm. **25**: 192-198.

Tao Tao, T., J.C. Bourne and R.M. Blumenthal. 1991. A family of regulatory genes associated with type-II restriction-modification systems. J. Bacteriol. **173**: 1367-1375.

Taylor, I., D. Watts and G. Kneale. 1993. Substrate recognition and selectivity in the type IC DNA modification methylase M.*Eco*R124I. Nucleic Acids Res. **21**: 4929-4935.

Templin, A., S.R. Kushner and A.J. Clark. 1972. Genetic analysis of mutations indirectly suppressing *recB* and *recC* mutations. Genetics. **72:** 205-215.

Toothman, P. 1981. Restriction alleviation by bacteriophage lambda and lambda reverse. J. Virology. **38**: 621-631.

Towbin, H., T. Staehelin and J. Gordon. 1979. Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. Proc. Natl. Acad. Sci. USA. **76**: 4350-4354.

Tyndall, C., J. Meister and T.A. Bickle. 1994. The *E. coli prr* region encodes a functional type IC DNA restriction system closely integrated with an anticodon nuclease gene. J. Mol. Biol. **237**: 266-274.

Van Pel, A. and C. Colson. 1974. DNA restriction and modification systems in *Salmonella*. Mol. Gen. Genetics. **135**: 51-60.

Vovis, G.F., K. Horiuchi and N.D. Zinder. 1974. Kinetics of methylation of DNA by a restriction endonuclease from *Escherichia coli* B. Proc. Natl. Acad. Sci. USA. **71**: 3810-3813.

Vovis, G.F. and N.D. Zinder. 1975. Methylation of f1 DNA by a restriction endonuclease from *Escherichia coli* B. J. Mol. Biol. **95**: 557-568.

Walder, R.Y., J.A. Walder and J.E. Donelson. 1984. The organisation and complete nucleotide sequence of the *Pst*I restriction-modification system. J. Biol. Chem. **259**: 8015-8026.

Walker, J.E., M. Saraste, M.J. Runswick and N.J. Gay. 1982. Distantly related sequences in the α- and β- subunits of ATP synthase, myosin, kinases and other ATP-requiring enzymes and a common nucleotide binding fold. EMBO J. **1**: 945-951.

Warner, H.R. and D.P. Snustad. 1983. T4 DNA nucleases. *In* **Bacteriophage T4**. C.K. Mathews *et al.* (eds), ASM, Washington DC. pp. 103-109.

Wilkins, B. 1995. Gene transfer by bacterial conjugation: diversity of systems and functional specialisations. *In* **Population Genetics of Bacteria**. S. Baumberg *et al.* (eds), Cambridge University Press. pp. 59-88.

Willcock, D.F., D.T.F. Dryden and N.E. Murray. 1994. A mutational analysis of the two motifs common to adenine methyltransferases. EMBO J. **13**: 3902-3908.

Wilson, G.G. and N.E. Murray. 1991. Restriction and modification systems. Ann. Rev. Genet. **25**:585-627.

Winkler, F.K., D.W. Banner, C. Oefner, D. Tsernoglou, R.S. Brown, S.P. Heathman, R.K. Bryan, P.D. Martin, K. Petratos and K.S. Wilson. 1993. The crystal structure of *Eco*RV endonuclease with cognate and non-cognate DNA fragments. EMBO J. **12**: 1781-1795.

Winship, P.R. 1989. An improved method for directly sequencing PCR amplified material using dimethyl sulphoxide. Nucleic Acids Res. **17**: 1266.

Wood, W.B. 1966. Host specificity of DNA produced by *Escherichia coli*: Bacterial mutations affecting the restriction and modification of DNA. J. Mol. Biol. **16**: 118-133.

Wulczyn, F.G., M. Bolker and R. Kahmann. 1989. Translation of the bacteriophage Mu *mom* gene is positively regulated by the phage *com* gene-product. Cell. **57**: 1201-1210.

Wyatt, G.R. and S.S. Cohen. 1952. A new pyrimidine base from bacteriophage nucleic acids. Nature. **170**: 1072-1073.

Yanisch-Perron, C., J. Vieria and J. Messing. 1985. Improved M13 phage cloning vectors and host strains - nucleotide sequences of the M13mp18 and pUC18 vectors. Gene. **33:** 103-109.

Yoshimori, R., D. Roulland-Dussoix and H.W. Boyer. 1972. R factor-controlled restriction and modification of deoxyribonucleic acid: restriction mutants. J. Bacteriol. **112**: 1275-1279.

Yuan, R. 1981. Structure and mechanism of multifunctional restriction endonucleases. Ann. Rev. Biochem. **50**: 285-315.

Yuan, R., T.A. Bickle, W. Ebbers and C. Brack. 1975. Multiple steps in recognition by restriction endonucleases from *E.coli* K. Nature. **256**: 556-560.

Yuan, R., D.L. Hamilton and J. Burckhardt. 1980. DNA translocation by the restriction enzyme from *E.coli* K. Cell. **20**: 237-244.

Zabeau, M., S. Friedman, M. Van Montagu and J. Schell. 1980. The *ral* gene of phage λ. I. Identification of a non-essential gene that modulates restriction and modification in *E. coli.* Molec. Gen. Genet. **179**: 63-73.

Zhou, Y., X. Zhang and R.H. Ebright. 1991. Random mutagenesis of gene-sized DNA molecules by use of PCR with Taq DNA polymerase. Nucleic Acids Res. **19**: 6051.

Zinkevich, V.E., A.M. Alekseev and V. I. Tanyashin. 1986. Influence of bacteriophage λ*ral* gene on the level of synthesis of the restriction endonuclease *Eco*K β-subunit. Molec. Biol. (Moscow). **20**: 1345-1351.

Zissler, J., E.R. Singer and F. Schaefer. 1971. The role of recombination in growth of bacteriophage lambda. *In* **The bacteriophage lambda.** A.D. Hershey (ed), Cold Spring Harbor Laboratories, New York. pp. 455-475.

Appendix

Publications

King, G. and N.E. Murray. 1994. Restriction enzymes in cells, not Eppendorfs. Trends Microbiol. **2**: 465-469.

King, G. and N.E. Murray. Modification enhancement and restriction alleviation by bacteriophage λ. Abstract from the 3rd NEB workshop on biological DNA modification. Gene. **157**: 225

King, G. and N.E. Murray. 1995. Restriction alleviation and modification enhancement by the *E. coli* Rac prophage. Mol. Microbiol. **16**: 769-777.

Restriction enzymes in cells, not eppendorfs

Gareth King and Noreen E. Murray

Restriction enzymes are essential reagents to molecular biologists, but their relevance to bacterial populations is less obvious. Most bacteria encode restriction and modification systems and these are commonly considered to be a barrier to phage infection. Current evidence also supports a more general role for them in genetic recombination.

G. King and N.E. Murray are in the Institute of Cell and Molecular Biology, University of Edinburgh, Edinburgh, UK EH9 3JR.

Bacterial restriction systems are often likened to the vertebrate 'immune' system because both distinguish 'self' from 'non-self' and both attack molecules that they designate as foreign. Restriction systems, unlike immune systems, confine their attack to DNA, and their mode of distinction between 'self' and 'non-self' relies on sequence-specific modification. The absence of specific modification can identify DNA as foreign and lead to its restriction. The diversity of restriction and modification (R–M) systems provides a battery of tools for modern molecular biology and many varied substrates for analysing key aspects of molecular recognition, particularly protein–DNA interactions. Recently solved structures of DNA–protein co-crystals for both restriction and modification enzymes have considerably aided our understanding of the DNA sequence specificity and catalytic mechanisms of both enzymes[1–4]. The DNA sequence specificity is not imposed by any previously identified motif that occurs in other DNA-binding proteins, but involves amino acids that are widely separated in the primary protein sequence. In this brief review, we confine our discussion to biological aspects of restriction and modification, and focus on the enteric bacteria and their phages, although we believe that the information broadly applies to bacterial systems in general.

R–M systems

Classical R–M systems have two components, a DNA endonuclease (restriction enzyme) and a DNA methyltransferase (modification enzyme), both of which recognize the same DNA target sequence. Only unmethylated target sites are substrates for the restriction enzyme, while hemimethylated sites are substrates for the cognate methylase, and so the fully methylated state of 'self' DNA is regenerated after replication.

The DNA of an infecting phage will only carry the modification required to escape cleavage if it has been propagated in a host with an R–M system of the same specificity. A single unmodified target in a phage chromosome can elicit cutting and is sufficient to inhibit, or 'restrict', phage propagation, although additional targets increase the level of restriction. This can be quantified by measuring the efficiency of plating (e.o.p.) of a phage on a restricting strain relative to that on a nonrestricting strain. The degree of restriction varies depending on the phage and the restriction system challenged, but the e.o.p. is typically reduced to between 10^{-1} and 10^{-5} by a single system. However, bacteria may have more than one system, in which case phage restriction may approach the product of the restriction abilities of the individual systems. Phage restriction is not generally absolute, and phage genomes that do escape cleavage normally become modified, so that progeny phages are protected against restriction by bacteria with an R–M system of the same specificity.

Variations in the e.o.p. of phages consistent with the activity of R–M systems were described as early as 1938[5], but an explanation based on the host-dependent modification of phages was first suggested in the early 1950s[6,7]. A decade later, it was shown that DNA is the substrate for modification, and that restriction is associated with the breakdown of unmodified DNA. Restriction enzymes were then purified on the basis of their endonucleolytic activity[8]. Classical R–M systems are now divided into three types on the basis of enzyme complexity, cofactor requirements and position of DNA cleavage (see Fig. 1; for recent reviews, see Refs 1,2,9–11), although new systems are being discovered that do not fit readily into this classification[12,13]. Some of the systems described in the 1950s are not classical R–M systems[6], but consist only of a restriction enzyme that is activated by specifically modified DNA. Interest in these modification-dependent restriction enzymes was stimulated when they were 'rediscovered' as the enzymes that hampered the cloning of genes encoding specific DNA methylases[14].

Type II R–M systems are the simplest. Restriction and modification activities are performed by separate enzymes that recognize short palindromic sequences, and the restriction enzymes cut within this sequence[1]. In contrast, a type I R–M system is a complex of three polypeptides (HsdR, HsdM and HsdS) with both activities. HsdS confers sequence specificity, and the restriction enzyme cuts the DNA at a variable distance from its recognition sequence. The type III enzymes are intermediate in complexity. They consist of two polypeptides, both of which are required to cleave DNA, but only the DNA-recognition subunit is required for modification[15].

Type I

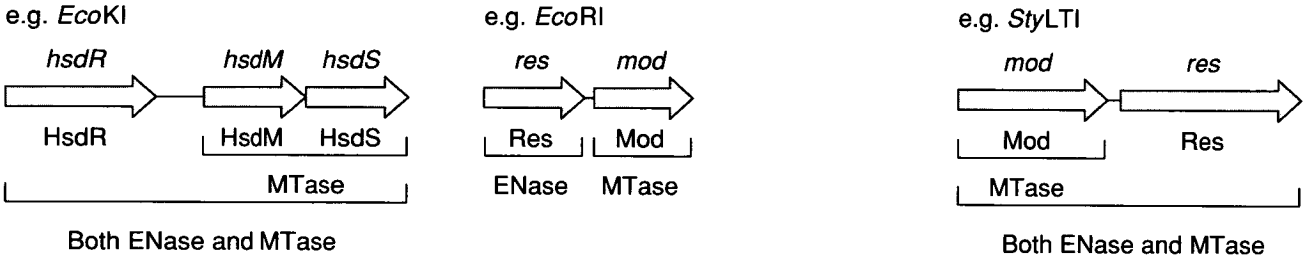
- 20 members
- Hetero-oligomeric enzymes
- Require ATP hydrolysis for restriction
- Cut DNA at various distances from the recognition sequence
- DEAD-box proteins

Type II

- Over 2400 members
- ENase and MTase separate enzymes
- Cut DNA within recognition sequence

Type III

- 5 members
- Hetero-oligomeric ENase
- ATP required for restriction
- Cut DNA close to recognition sequence
- DEAD-box proteins



Modification-dependent

- 4 members
- Cut modified DNA
- No cognate MTase

e.g. McrBC

Type IIs

- Over 35 members
- Similar to type II, but cut DNA outside recognition sequence

e.g. *Fok*I

Not classified

e.g. *Eco*571

- One polypeptide is both an ENase and MTase

e.g. *Bcg*I

- Heterodimeric ENase and MTase

Fig. 1. The distinguishing characteristics and organization of the genetic determinants and subunits of the different types of restriction and modification (R–M) systems. R–M systems are classified on the basis of their complexity, cofactor requirements and position of DNA cleavage with respect to their DNA target sequence[9]. The type I, II and III systems are the classical R–M systems. The restriction enzymes of the type I and III systems contain motifs characteristic of 'DEAD'-box proteins[49,50], and may therefore be helicases. The type IIs systems[51] are a subgroup of the type II systems that cleave DNA outside their recognition sequence. Some systems do not fit readily into the present classification[12,13] and *Eco*571 has been tentatively termed a type IV system[12]. The modification-dependent systems are not classical R–M systems, as they have no cognate methylase and only cut DNA that contains certain modifications. ENase, restriction endonuclease; MTase, methyltransferase.

Occurrence of R–M systems

R–M systems may be encoded by bacterial chromosomes or by their plasmids, and occasionally by prophages, but, surprisingly, they are also encoded by the viruses of the eukaryote *Chlorella*[16]. Of the R–M systems identified, type II systems are by far the most numerous, with over 2400 members[17]. This predominance probably reflects their importance to molecular biologists and the efficient screens that have been developed to identify them, rather than representing their prevalence in nature. There is no effective general screen for all types of R–M systems, but a few surveys give us some information on their distribution and frequency in nature. Of over 1000 strains of enteric bacteria, mainly *Escherichia coli*, that were screened for DNA-sequence-specific endonucleases, 25% were positive and may therefore have a type II R–M system[18]. A collection of 72 *E. coli* strains that represent the genotypic diversity of this species is now available[19]. A screen using probes for genes encoding type I R–M systems identified them in 17 of the 36 strains tested[20]. Most type I systems have been identified in the commonly studied enteric bacteria, although representatives are now being found in other bacteria species[21]. Estimates of the frequency of occurrence of R–M systems are limited by the screens available, but it is likely that the majority of bacteria do have a system *Escherichia coli* C is the only natural isolate that has been studied extensively and found not to have a R–M system[14], but, like other bacteria, it can readily acquire a plasmid-encoded system.

The type I R–M systems that have been identified in natural isolates of Enterobacteriaceae currently have 20 different specificities, 17 of which are believed to be encoded by alleles of the chromosomal *hsd* locus which was originally identified in *E. coli* K12. Some alleles of *hsdR* and *hsdM* encode polypeptides with very similar amino acid sequences (80–99% identity) even when from different species, but regions of *hsdS* always differ considerably if they confer different specificities. Other alleles encode remarkably dissimilar polypeptides (20–35% identity), irrespective of the host species, and retain little evidence of common ancestry[20]. This unusual allelic diversity is shared by other genes that are concerned with differentiating 'self' from 'non-self', for example, the major histo

compatibility complex class II locus of mammals and the self-incompatibility genes of plants[22]. The selective pressure imposed on a bacterial population for variation in the sequence specificity of R–M systems may be the challenge by phages[23].

Phage antirestriction systems

The significance of R–M systems to bacteriophages is supported by the numerous ways in which phages can escape restriction. These include the absence or masking of restriction sites and the inactivation or modulation of R–M enzymes[11]. Sequence analyses of bacteriophage genomes have shown that some phages (for example, the *Bacillus* phage Φ29) have significantly fewer targets for type II restriction enzymes of potential hosts than expected[9,11,24], and a recent analysis of naturally occurring coliphages suggests that these target sequences are selected against[25].

The coliphages T3 and T7 encode an early gene, *ocr*, the product of which binds type I R–M enzymes and abolishes both restriction and modification activities. The *ocr* gene is expressed before targets in the phage genome are accessible to host restriction enzymes[26], so that these phages are protected from restriction and modification by type I systems. Bacteriophage P1 also protects its DNA from type I restriction, but the antirestriction function, Dar, does not interfere with modification. The Dar proteins are co-injected with encapsidated DNA, so that any DNA packaged in a P1 head is protected. This allows efficient generalized transduction to occur between strains with different type I R–M systems. How Dar functions is unclear, but it has been suggested that Dar bound to DNA might hinder the DNA translocation that type I enzymes require for cleavage[27].

Bacteriophage λ encodes a very specialized antirestriction function, Ral, that modulates the activity of only some type I R–M systems, by alleviating restriction and enhancing modification. The systems influenced by Ral are those that have a modification enzyme with a strong preference for hemimethylated DNA. Ral may act by changing the activity of the complex from a maintenance methylase to a *de novo* methylase[28]. Unmodified λ DNA is restricted because *ral* is not normally expressed before the genome is cut by the host R–M system, but phages that escape restriction and express *ral* are protected by the enhanced modification of their DNA (G. King and N.E. Murray, unpublished).

Some phages are made resistant to many types of R–M system by the presence of glucosylated hydroxymethylcytosine (hmC) in their DNA, for example, the *E. coli* T-even phages and the *Shigella* phage DDVI. This modification also identifies phage DNA and allows selective degradation of host DNA by phage-encoded nucleases[29]. Nonglucosylated T-even phages are resistant to classical R–M systems because their DNA still contains the modified base hmC. These phages can be restricted by the methylation-dependent systems McrA and McrBC, although T-even phages encode a protein, Arn, that protects superinfecting phages from McrBC restriction. It has been suggested that phages have evolved DNA containing hmC to counteract classical R–M systems, and that host-encoded modification-dependent endonucleases are a response to this phage adaptation. In this evolutionary story, the Arn protein and glucosylation of hmC would be the latest mechanisms that render T-even phages totally resistant to most R–M systems[11].

Evolution of new specificities

The type I R–M systems are suited to changing their specificity because the same subunit, HsdS, defines the specificity of both restriction and modification activities. Sequence comparisons have shown that the type I *hsdS* genes consist of two variable regions separated by a central conserved region. These systems recognize a DNA target sequence of two half sites, each 3–5 nucleotides long, separated by a nonspecific spacer of 6–8 nucleotides. Several spontaneous mutants have demonstrated how new specificities may be generated in nature[11,20], and have identified key aspects of DNA recognition by type I systems (see Fig. 2 for details). The two variable regions of *hsdS* encode two target-recognition domains (TRDs), each of which recognizes a particular half site in the target sequence. New combinations of variable domains and minor changes in the length of coding sequence that separates the variable regions can both generate new specificities. Functional type I R–M systems have been identified in *hsdS* deletion mutants encoding only one TRD (Refs 30,31). It seems probable that *hsdS* originally encoded only one TRD, with gene duplication and divergence leading to the present arrangement[9]. It is now apparent that TRDs may be assembled in different ways to generate new specificities, but it is not at all clear how new TRDs may evolve by the accumulation of point mutations.

Transfer of R–M systems

R–M systems are transferred between hosts and even between species, as implied by DNA sequence comparisons of the chromosomal *hsd* alleles[22], and the existence of plasmid-encoded R–M systems, which may be maintained in a broad host range. Self-transmissible plasmids may also encode antirestriction functions, in particular *ard*, the product of which is functionally similar to T7 Ocr and has regions of sequence identity with this protein[32,33]. If an R–M system is transferred to a new host, the DNA of the recipient must be protected from cutting, which may be achieved by ensuring that a functional methylase is produced before the restriction enzyme. Several systems have been shown to have separate promoters for the genes encoding the methylase and endonuclease, which provide the potential for differential gene expression[9]. Some of the genes encoding type II enzymes are transcriptionally regulated by repressor-like C proteins of about 100 amino acids with a characteristic helix–turn–helix DNA-binding motif[34,35]. C proteins have a much greater degree of similarity to each other, and to some other gene repressors, than do the R–M systems with which they are associated. Perhaps the genes encoding C proteins have been recruited independently into R–M systems[34].

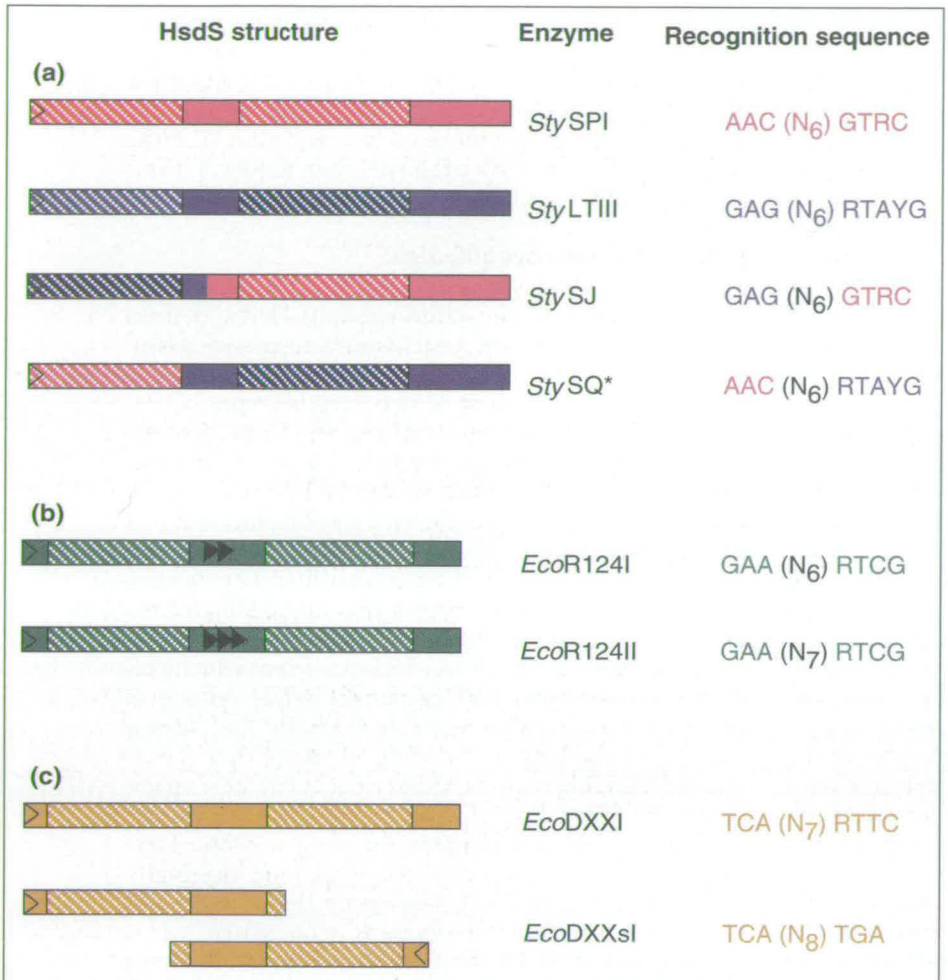


Fig. 2. Evolution of new type I restriction and modification (R–M) specificities[11,20]. **(a)** Rearrangement of variable regions within *hsdS* results in R–M systems with hybrid specificities. *Sty*SJ arose by recombination between the central conserved regions of the *hsdS* genes of *Sty*SPI and *Sty*LTIII, two naturally occurring R–M systems, and recognizes a hybrid sequence of the two parental systems. This demonstrates that there are two target-recognition domains (TRDs) located in separate regions of the HsdS subunit. The *Sty*SQ* *hsdS* gene contains only the amino-terminal variable region of *Sty*SPI and the remaining coding region of *Sty*LTIII. The amino-terminal variable region, therefore, encodes the TRD responsible for recognition of the trinucleotide component of the recognition sequence and, by analogy, the carboxy-terminal variable region encodes the TRD that recognizes the tetranucleotide component. **(b)** Illustrates changing DNA specificity by altering the length of the *hsdS* central conserved region. *Eco*R124I and *Eco*R124II differ only in the number of repeats of a short sequence within the central conserved region of *hsdS*, with *Eco*R124II having an extra repeat. The *Eco*R124II recognition sequence contains the same two half sites as *Eco*R124I, but has an extra base in the non-specific spacer. **(c)** An active R–M system may have only one kind of TRD. A transposon-insertion mutant of the *hsdS* gene of *Eco*DXXI encodes a system with a new specificity, *Eco*DXXsI, that has an HsdS polypeptide with only one TRD. *Eco*DXXsI recognizes a bipartite, symmetrical sequence, with each half site corresponding to the sequence recognized by the single TRD. This is consistent with two truncated HsdS polypeptides replacing one HsdS polypeptide with two TRDs in the R–M enzyme complex. In all parts, colours indicate the origin of regions of the HsdS polypeptide; open arrow, HsdS amino terminus; hatched box, variable region; open box, conserved region; solid arrow, repeat; R, purine; Y, pyrimidine; N, any base.

The *hsd* genes of type I systems also have two promoters, one for the genes encoding the methylase and another for the gene encoding the additional subunit that is required for restriction, but no differences have been detected in transcription from the two promoters[36–38]. However, when the *hsd* genes are transferred by conjugation, there is a lag before restriction activity is detected, providing time for modification to occur[39]. A mutant has been isolated that cannot act as a recipient for functional *hsd* genes by conjugation, presumably because the host DNA is degraded by an active restriction complex immediately on transfer[40]. Restriction activity is not controlled at the transcriptional level, and characterization of this mutant should help to identify the mechanism by which R–M activity is regulated.

Possible roles for R–M systems

The wide distribution and extraordinary diversity of R–M systems suggest that they have an important role in the bacterial community. Usually, this is considered to be protection against phage infection. However, laboratory studies on the population dynamics of bacteria and phages indicate that R–M systems provide only a transitory advantage to bacteria[41]. Essentially, a 'novel' R–M system can assist bacteria in colonizing a new habitat in which phages are present, but this advantage is soon lost as phages become modified and bacteria acquire mutations conferring resistance to the infecting phages. One possible advantage of a barrier provided by an R–M system, rather than a mutation conferring phage resistance, is that one R–M system can protect simultaneously against a variety of phages. Maintenance of one or more R–M systems may be preferable to maintaining resistance to multiple phages because phage-adsorption sites commonly have additional functions, and their disruption may compromise the fitness of the bacterium in nature. This is not supported by the laboratory experiments in which mutations conferring resistance to multiple phages were prevalent[41]. It seems relevant that the restriction barrier is generally incomplete, irrespective of the mechanism of DNA transfer, and that the fate of phage and bacterial DNA cut by restriction enzymes is likely to be different. Fragmentation of a phage chromosome leads to restriction of phage growth, while some of the DNA fragments generated from a bacterial chromosome may be rescued by homologous recombination with the host genome. A protective role for R–M systems against phage infection does not exclude an additional role in genetic recombination.

Linear DNA fragments are extremely vulnerable to degradation by bacterial nucleases. It is not unexpected,

therefore, that experiments have shown that the overall level of recombination is severely depressed in the presence of an active R–M system[27,42]. Similarly, linkage is reduced[20]. However, the suggestion that the DNA ends generated by restriction enzymes are recombinogenic[43,44] is gaining experimental support. Restriction *in vivo* can generate DNA ends that initiate recombination[45], and recent data give some clues as to how the apparent competition between degradation and recombination may be influenced[46]. One analysis of transductants suggests that R–M systems may influence the pattern of DNA fragments incorporated into the recipient chromosome, creating a mosaic of short sequences in the genome[47]. This is consistent with the finding that the chromosomes of naturally occurring *E. coli* differ in short regions of DNA sequence[48] and the idea that R–M systems are involved in influencing the flux of genetic material between bacterial populations. Current technology permits a refined analysis of DNA transfer; the restriction analysis of DNA amplified by the polymerase chain reaction can be used to trace the origin of even short segments of chromosomal DNA. Applying these techniques to examine recombinant bacteria, where the parental strains are of defined recombination and R–M genotypes, now allows the influence of restriction on genetic transfer to be analysed in detail, at least in the laboratory.

Acknowledgements

We thank Vicki Barcus, Annette Titheradge, Brian Wilkins, Melissa McKane and Roger Milkman for making manuscripts available before publication, and Lynn Powell, David Leach and Francis Fuller-Pace for helpful discussions. Space limitations mean that we have cited reviews for much of the earlier work; we apologize to many workers for the lack of individual citations. Work in this laboratory is supported by the Medical Research Council.

References

1 Roberts, J.R. and Halford, S.E. (1993) in *Nucleases* (Linn, S.M., Lloyd, R.S. and Roberts, R.J., eds), pp. 35–88, Cold Spring Harbor Laboratory Press
2 Halford, S.E. *et al.* (1993) in *Nucleic Acids and Molecular Biology* (Vol. 7) (Eckstein, F. and Lilley, D.M.J., eds), pp. 47–69, Springer-Verlag
3 Cheng, X. *et al.* (1994) *EMBO J.* 13, 3927–3935
4 Klimasauskas, S. *et al.* (1994) *Cell* 76, 357–369
5 Craigie, J. and Yen, C.H. (1938) *Can. J. Public Health* 29, 448–463
6 Luria, S.E. and Human, M.L. (1952) *J. Bacteriol.* 64, 557–569
7 Bertani, G. and Weigle, J.J. (1953) *J. Bacteriol.* 65, 113–121
8 Arber, W. (1971) in *The Bacteriophage Lambda* (Hershley, A.D., ed.), pp. 83–96, Cold Spring Harbor Laboratory Press
9 Wilson, G.G. and Murray, N.E. (1991) *Annu. Rev. Genet.* 25, 585–627
10 Heitman, J. (1993) in *Genetic Engineering* (Vol. 15) (Setlow, J.K., ed.), pp. 57–108, Plenum Press
11 Bickle, T.A. and Krüger, D.H. (1993) *Biol. Rev.* 57, 434–450
12 Janulaitis, A. *et al.* (1992) *Nucleic Acids Res.* 20, 6043–6049
13 Kong, H. *et al.* (1994) *J. Biol. Chem.* 269, 683–690
14 Raleigh, E.A. (1992) *Mol. Microbiol.* 6, 1079–1086
15 Bickle, T.A. (1993) in *Nucleases* (Linn, S.M., Lloyd, R.S. and Roberts, R.J., eds), pp. 89–109, Cold Spring Harbor Laboratory Press
16 Burbank, D.E. *et al.* (1990) *Virology* 176, 311–315
17 Roberts, J.R. and Macelis, D. (1993) *Nucleic Acids Res.* 21, 3125–3137
18 Janulaitis, A. *et al.* (1988) *Gene* 74, 229–232
19 Ochman, H. and Selander, R.K. (1984) *J. Bacteriol.* 157, 690–693
20 Barcus, V. and Murray, N.E. in *Population Genetics of Bacteria* (*Soc. Gen. Microbiol. Symp. 52*) (Baumberg, S. *et al.*, eds), Cambridge University Press (in press)
21 Dybvig, K. and Yu, H. (1994) *Mol. Microbiol.* 12, 547–560
22 Sharp, P.M. *et al.* (1992) *Proc. Natl Acad. Sci. USA* 89, 9836–9840
23 Levin, B.R. (1986) in *Evolutionary Processess and Theory* (Karlin, S. and Nero, E., eds), pp. 669–688, Academic Press
24 Karlin, S., Burge, C. and Campbell, A.M. (1992) *Nucleic Acids Res.* 20, 1363–1370
25 Korona, R., Korona, B. and Levin, B.R. (1993) *J. Gen. Microbiol.* 139, 1283–1290
26 Moffat, B.A. and Studier, F.W. (1988) *J. Bacteriol.* 170, 2095–2105
27 Iida, S. *et al.* (1987) *Virology* 157, 156–166
28 Loenen, W.A.M. and Murray, N.E. (1986) *J. Mol. Biol.* 190, 11–22
29 Warner, H.R. and Snustad, D.P. (1983) in *Bacteriophage T4* (Mathews, C.K. *et al.*, eds), pp. 103–109, American Society for Microbiology
30 Meister, J. *et al.* (1993) *EMBO J.* 12, 4585–4591
31 Abadjieva, A. *et al.* (1993) *Nucleic Acids Res.* 21, 4435–4443
32 Wilkins, B.M. in *Population Genetics of Bacteria* (*Soc. Gen. Microbiol. Symp. 52*) (Baumberg, S. *et al.*, eds), Cambridge University Press (in press)
33 Belogurov, A.A., Delver, E.P. and Rodzevich, O.V. (1993) *J. Bacteriol.* 175, 4843–4850
34 Tao, T., Bourne, J.C. and Blumenthal, R.M. (1991) *J. Bacteriol.* 173, 1367–1375
35 Siksnys, V. *et al.* (1994) *Gene* 142, 1–8
36 Loenen, W.A.M. *et al.* (1987) *J. Mol. Biol.* 198, 159–170
37 Prakash-Cheng, A., Valinluck, B. and Ryu, J. (1991) *Gene* 99, 9–14
38 Sharrocks, A.D. and Hornby, D.P. (1991) *Mol. Microbiol.* 5, 685–694
39 Prakash-Cheng, A. and Ryu, J. (1993) *J. Bacteriol.* 175, 4905–4906
40 Prakash-Cheng, A., Chung, S.S. and Ryu, J. (1993) *Mol. Gen. Genet.* 241, 491–496
41 Korona, R. and Levin, B.R. (1993) *Evolution* 47, 556–575
42 Arber, W. and Morse, M.L. (1965) *Genetics* 51, 137–148
43 Endlich, B. and Linn, S. (1985) *J. Biol. Chem.* 260, 5720–5728
44 Price, C. and Bickle, T.A. (1986) *Microbiol. Sci.* 3, 296–299
45 Stahl, M.M. *et al.* (1983) *Proc. Natl Acad. Sci. USA* 80, 2310–2313
46 Kuzminov, A., Schabtach, E. and Stahl, F.W. (1994) *EMBO J.* 13, 2764–2776
47 McKane, M. and Milkman, R. *Genetics* (in press)
48 Milkman, R. and Bridges, M.M. (1993) *Genetics* 133, 455–468
49 Gorbalenya, A.E. and Koonin, E.V. (1991) *FEBS Lett.* 291, 277–281
50 Murray, N.E. *et al.* (1993) *Mol. Microbiol.* 9, 133–143
51 Szybalski, W. *et al.* (1991) *Gene* 100, 13–26

Chance and selection in the evolution of barley mildew

James K.M. Brown

Populations of the barley powdery mildew fungus are genetically very diverse. However, when a new resistance gene is introduced into barley to control mildew, the population of the pathogen may respond by rapid growth of a few virulent clones. These phases of rapid clonal evolution cause radical changes in the frequencies of mildew genotypes.

J.K.M. Brown is in the Cereals Research Dept, John Innes Centre, Colney Lane, Norwich, UK NR4 7UH.

In 1986, a new winter barley variety, Pipkin, was included in the list of cereal varieties recommended to be grown in England and Wales[1]. Although Pipkin's excellent malting quality was its main attraction to farmers, it also had outstanding resistance to powdery mildew. Mildew is caused by a fungus, *Erysiphe graminis* f. sp. *hordei*, and is the most serious disease of barley in northern Europe. Until 1988, Pipkin had a mildew-resistance rating of 9, the maximum possible. However, as a result of trials carried out in 1988, its rating fell to 4 in 1989 and then to 2 in 1990, indicating that Pipkin had become very susceptible to powdery mildew and should not be grown widely for this reason (Fig. 1).

The resistance to mildew of Pipkin, and of several other barley varieties introduced around the same time, is conferred by the gene *Mla13*. In plants, resistance to specific strains of pathogens is under genetic control. Therefore, the *Mla13* resistance gene became ineffective when the combination of the virulence alleles *Va13* and *V(Ru3)* in the pathogen (Box 1), which enable mildew to overcome the resistance, reached a significant frequency. (For brevity, virulence on *Mla13* is denoted *Va13* hereafter.)

The speed with which mildew adapted to *Mla13* is neither unexpected nor unusual, as many other mildew-resistance genes that had been used previously in barley breeding also remained effective for only a few years (Fig. 1) because of the evolution of matching virulences[2]. Current research in several countries in Europe is aimed at understanding the structure of populations of *E. graminis* f. sp. *hordei* and its implications for the evolution of the pathogen.

Emergence of virulent clones

The heavy mildew infections that first appeared on *Mla13* varieties early in the summer of 1988 provided an opportunity to study a mildew population in the process of adapting to a new resistance gene[3]. Mildew isolates were sampled throughout the UK and Ireland from farms and field trials and by trapping airborne spores. Samples were taken from several *Mla13* barley varieties, mainly Pipkin and two spring varieties, Digger and Sherpa (Fig. 2).

The range of characteristics that can be used to classify mildew isolates is limited because the fungus cannot be cultured on artificial media. Nevertheless, useful variation in virulences on different barley varieties (Box 1), responses to fungicides and genetic fingerprints can be exploited. The DNA probe that was used for fingerprinting *Va13* isolates hybridizes to repetitive sequences dispersed throughout the *E. graminis* f. sp. *hordei* genome, and can therefore be used to determine whether a group of isolates belongs to a single clone[4]. Two clones predominated among *Va13* isolates. The more common of the two was found over a range of 650 km, while the other was isolated in three counties, over a range of 400 km. A third clone was comparatively rare, but was found in three places over 250 km. Several isolates with unique genotypes were also identified (Fig. 2)[3]. It was concluded that the breakdown of the *Mla13* resistance had largely been caused by the two common clones of *E. graminis* f. sp. *hordei*.

Single clones have caused the breakdown of two other mildew resistances and the partial loss of effectiveness of a fungicide. The variety Triumph, which was introduced in 1980, carried an effective resistance gene, *Ml(Ab)*, as well as an ineffective one, *Mla7* (Fig. 1). A survey in 1985 indicated that *Ml(Ab)* had been overcome by just one clone of the mildew pathogen[5,6]. In the 1980s in north-east Scotland, morpholine-type fungicides were used to control mildew on a susceptible variety, Golden Promise. Mildew isolates with particularly low sensitivity to these fungicides were sampled from this region in 1988 and 1990, and probably constituted a single clone because they shared an unusual set of virulences[7,8]. The most spectacular increase of a clone occurred in 1986. Until that year, several important barley varieties had either *Mla7*, which was often combined with *Mlk*, or *MlLa*. No popular variety had carried both *Mla7* and *MlLa*. However, in 1986, 10% of spring barley seed sales were of the variety Klaxon, which had *Mla7*, *Mlk* and *MlLa*, while 1.3% were of the variety Doublet (*Mla7* and *MlLa*). Until 1986, these combinations of resistances were effective, because very few mildew isolates had the combined virulences *Va7* and *VLa* (Box 1). However, in 1986, the frequency of one clone with this virulence combination rose from less than 1% to 36% between June and October in random samples of isolates from the airborne spore population[9].

Molecular Microbiology (1995) **16**(4), 769–777

Restriction alleviation and modification enhancement by the Rac prophage of *Escherichia coli* K-12

Gareth King and Noreen E. Murray*
Institute of Cell and Molecular Biology, University of Edinburgh, Darwin Building, Kings Buildings, Edinburgh EH9 3JR, UK.

Summary

Bacteriophage λ encodes an antirestriction function, Ral, which is able to modulate the activity of the *Escherichia coli* K-12 restriction and modification system, *Eco*KI. Here we report the characterization of an analogous function, Lar, expressed by *E. coli sbcA* mutants and the hybrid phage λ*reverse*. *E. coli sbcA* mutants and λ*reverse* both express genes of the Rac prophage, and we have located the *lar* gene immediately downstream of *recT* in this element. The *lar* gene has been cloned in an expression plasmid, and a combination of site-directed mutagenesis and labelling of plasmid-encoded proteins has enabled us to identify a number of translational products of *lar*, the smallest of which is sufficient for restriction alleviation. Lar, like Ral, is able both to alleviate restriction and to enhance modification by *Eco*KI. Lar, therefore, is functionally similar to Ral and the nucleotide sequences of their genes share 47% identity, indicating a common origin. A comparison of the predicted amino acid sequences of Lar and Ral shows only a 25% identity, but a few short regions do align and may indicate residues important for structure and/or function.

Introduction

Bacteriophages use numerous mechanisms to escape host-encoded restriction and modification (R-M) systems (see Bickle and Krüger, 1993, and references therein). The activity of the *Escherichia coli* K-12 R-M enzyme, *Eco*KI, is modulated by a number of anti-restriction systems, including the Ocr activity of phages T3 and T7 (Studier, 1975), and the Ral activity of phage λ (Zabeau *et al.*, 1980). Ocr has been shown to competitively bind *Eco*KI, and block both restriction and modification of the phage (Bandyopadhyay *et al.*, 1985). However, Ral is different; it alters the activity of the restriction complex so that unmodified DNA that escapes restriction is efficiently methylated (G. King and N. Murray, unpublished). *Eco*KI is a member of the Type IA R-M enzymes, which are the only prokaryotic DNA methyltransferases (MTases) shown to have a strong preference for hemimethylated DNA *in vivo*, and are also the only systems in which restriction is known to be alleviated by Ral (Loenen and Murray, 1986). One *ral* homologue, that of phage P22, has been sequenced; the predicted amino acid (aa) sequences of the two gene products are almost identical (Semerjian *et al.*, 1989). *E. coli sbcA* hosts (Templin *et al.*, 1972) and a derivative of phage λ, designated λ*reverse* (λ*rev*; Zissler *et al.*, 1971), express a Ral-like activity (Simmon and Lederberg, 1972; Toothman, 1981). In *sbcA* mutants, the *recE* and *recT* genes, the general recombination functions of a defective lambdoid prophage, Rac, have been activated (Kushner *et al.*, 1974; Clark *et al.*, 1993; Kolodner *et al.*, 1994). λ*rev* also expresses the *recE* and *recT* genes (Gottesman *et al.*, 1974) acquired by a recombination event between λ and the excised Rac prophage (Kaiser and Murray, 1979). The gene encoding the Ral-like activity of the Rac prophage has been termed *lar* (Toothman, 1981).

The Type IA R-M systems are encoded by three genes: *hsdR*, *hsdM*, and *hsdS*. The HsdM and HsdS subunits together form a MTase that methylates two adenine residues in complementary strands of a bipartite DNA recognition sequence. In the presence of the HsdR subunit the complex can also act as an endonuclease, binding to the same target sequence but cutting the DNA some distance from this site. Whether the DNA is cut or modified depends on the methylation state of the target sequence. When the target site is unmodified, the DNA is cut. When the target site is hemimethylated, that is when only one adenine residue in the target is methylated, the complex acts as a maintenance MTase modifying the DNA so that both strands become methylated. This maintenance MTase activity ensures that the pattern of DNA methylation is preserved after replication. The enzyme recognizes a completely modified target sequence, but dissociates from the DNA (Suri *et al.*, 1984; Dryden *et al.*, 1993; Powell *et al.*, 1993).

Experiments dependent on expression of *ral* from a multicopy plasmid suggest that Ral alleviates restriction by changing the activity of *Eco*KI from a maintenance MTase to a *de novo* MTase, which efficiently methylates

Received 22 December, 1994; revised 20 February, 1995; accepted 24 February, 1995. *For correspondence. Tel (0131) 6505374; Fax (0131) 6683870.

unmodified DNA (Loenen and Murray, 1986). The efficient methylation of unmodified DNA is also observed with Ral and the *Eco*KI MTase alone, and mutations have been isolated in *hsdM* which mimic the activity of Ral (Kelleher *et al.*, 1991). We are interested in the molecular mechanisms of these activities, and in defining important domains or residues within Ral and *Eco*KI.

Here we report the identification of *lar*, the gene

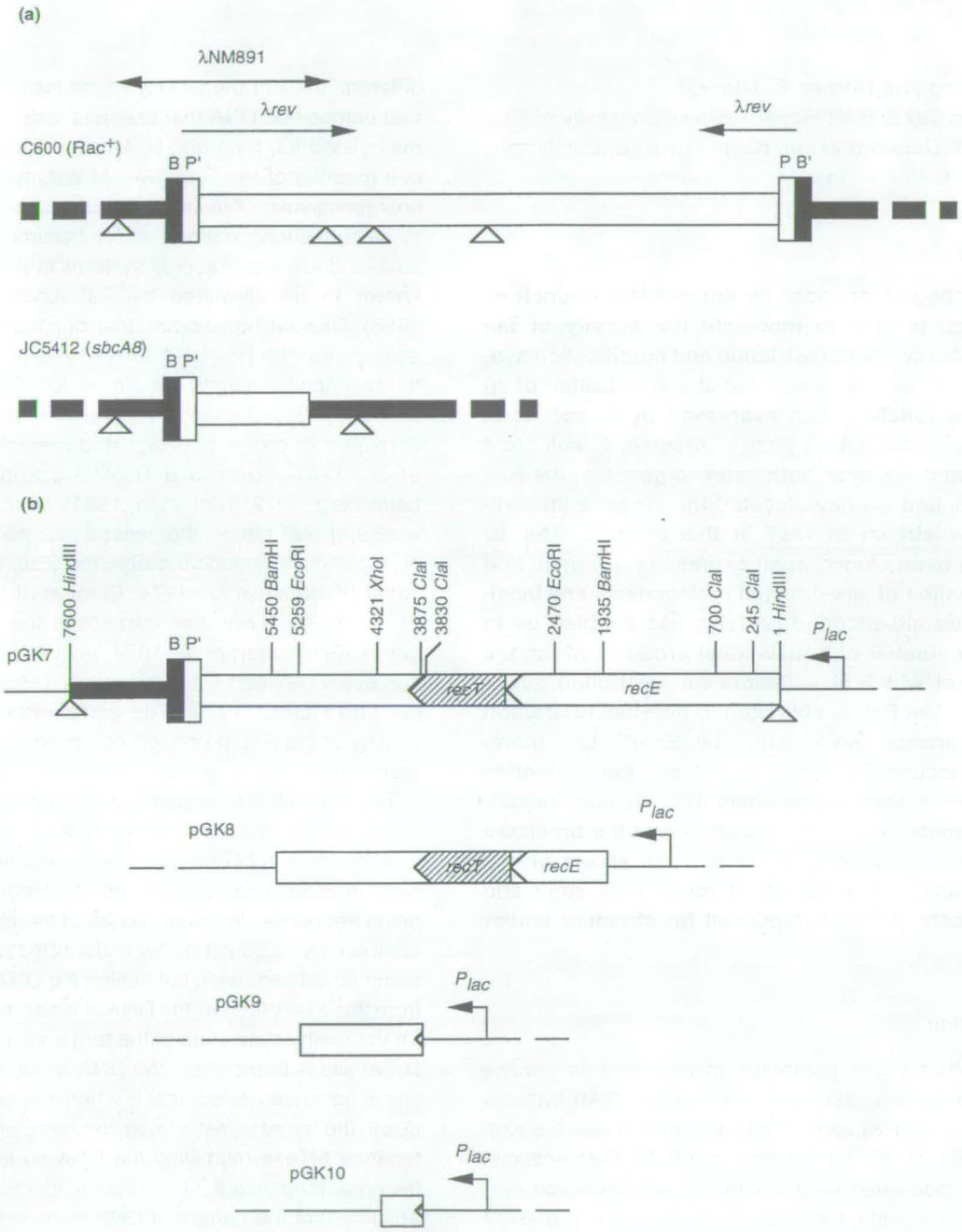


Fig. 1. Localization of the region encoding *lar*.
a. Origin of λ*rev*, λNM891 and the Rac prophage integrated in the chromosome of the *E. coli* strain C600 and an *sbcA* mutant (*sbcA8*; JC5412). The DNA common to λ*rev* and the *sbcA8* mutant is contained in λNM891.
b. The 7.6 kb *Hin*dIII fragment from λNM891 was cloned in pUC18 (Yanisch-Perron *et al.*, 1985) to generate pGK7. Segments of this 7.6 kb *Hin*dIII fragment were cloned and screened for restriction-alleviation activity. The *Bam*HI fragment was cloned in pUC18 to generate pGK8. The *Cla*I–*Eco*RI fragment was excised from pGK8, converted to a blunt-ended fragment using Klenow polymerase, and ligated to pUC18 cut with *Sma*I, to generate pGK9. pGK10 was similarly made with the 500 bp *Xho*I–*Eco*RI fragment from pGK9. Thick line, *E. coli* chromosome; thin line, pUC18 DNA; open box, Rac DNA; PB', hybrid Rac attachment site; open triangles, *Hin*dIII target sites; p_{lac}, *lacUV5* promoter in pUC18; DNA sequence numbering as Clark *et al.* (1993).

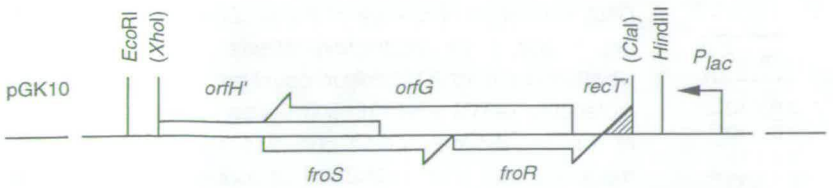


Fig. 2. Schematic showing the open reading frames within the 500 bp *Cla*I–*Xho*I DNA fragment cloned in pGK10. Predicted length of gene products; OrfG, 77 aa; OrfH', 66 aa (truncated); FroR, 39 aa; FroS, 40 aa. Brackets indicate restriction sites in the Rac prophage lost during cloning.

responsible for Ral-like activity in *E. coli sbcA* hosts and λ*rev* (Toothman, 1981). We have cloned *lar* and used an expression plasmid to identify its gene product. A comparison of the predicted aa sequences of Ral and Lar has highlighted a number of common motifs.

Results

Identification of the region encoding lar

To identify the region of the Rac prophage that encodes *lar*, a 7.6 kb *Hin*dIII fragment from λNM891 (Kaiser and Murray, 1979) that includes DNA common to the *E. coli sbcA8* mutant and λ*rev* was cloned in pUC18 to generate pGK7 (Fig. 1). An *Eco*KI restriction-proficient strain (NM710) was transformed with pGK7 and the restriction-alleviation activity of the plasmid was examined by comparing the titre of λv.0 on this strain versus a non-restricting strain (NM792). In this screen, a similar titre indicates restriction alleviation. pGK7 alleviated restriction, and the *lar* gene was further localized by cloning segments of the 7.6 kb *Hin*dIII fragment in pUC18 (Fig. 1; see legend for experimental details). The smallest DNA fragment that conferred restriction-alleviation activity was a 500 bp insert in pGK10.

The region of the Rac prophage in pGK10 has been sequenced (Clark *et al.*, 1993; GenBank Accession Number L23927) and contains four open reading frames (ORFs), *orfG* and *orfH'* (truncated) in one orientation, and *froR* and *froS* in the other (Fig. 2). All of the clones that had Lar activity contained a plasmid insert oriented such that the *lac* promoter produces a transcript of *orfG*. In order to examine the orientation dependence of the restriction-alleviation phenotype, the 500 bp *Eco*RI-*Hin*dIII fragment from pGK10 was cloned in pUC19, to generate pGK11, then removed and recloned in pUC18, to generate pGK12. pGK11 and pGK12, therefore, differ only in the orientation of the 500 bp insert relative to the *lac* promoter. Restriction-alleviation activity was unaffected by the orientation of the insert (Table 1), suggesting that there is sufficient *lar* expression to alleviate restriction even when the gene is in the opposite orientation to the *lac* promoter. In these experiments, transcription was neither enhanced nor opposed by activation of the *lac* promoter.

Table 1. Orientation dependence of *lar* activity.

Host	Titre λ.0
NM794	4.5×10^{10}
NM777 pUC18	2.4×10^{7}
NM777 pGK11	2.1×10^{10}
NM777 pGK12	3.8×10^{10}

Restriction-alleviation activity was determined for each host by measuring the titre of unmodified λ phages (λNM325.0) normalized against the titre of modified λ. This corrects for variation resulting from factors other than restriction (see the *Experimental procedures*).

Table 2. Modification-enhancement activity of *lar*.

Host	Eop of progeny phages
NM526(DE3) pT7-6	4.6×10^{-3}
NM526(DE3) pGK15	1.0

A restriction-deficient, modification-proficient strain (NM526) was transformed with the plasmid of interest and infected with unmodified λ phages (λNM325.0). After a single round of infection, the degree of phage modification was estimated by measuring the efficiency of plating (eop) of progeny phages on a restricting strain (Ymel) relative to that on a non-restricting derivative (NM526) (see the *Experimental procedures*).

Identification of the gene encoding lar

A 290 bp fragment including *orfG* (nucleotides 3883–4172, Clark *et al.*, 1993) was amplified by the polymerase chain reaction (PCR) and cloned in pT7-6 to generate pGK15. This plasmid was found both to alleviate restriction and to enhance modification by *Eco*KI (Tables 2 and 3). *orfG* is the only complete ORF transcribed from the T7 promoter in pGK15, although *froR* and *froS* are read in the opposite orientation. To test whether *orfG* encoded *lar*, an amber codon was generated within *orfG* which would have no effect on the predicted aa sequence of *froS*. Codon 42 of *orfG* was changed from AAG to TAG. The resulting plasmid (pGK16; K42Am) was examined for restriction-alleviation activity in both a *sup*° (NM777) and *supF* (Ymel) host (Table 3). Restriction was not alleviated in a *sup*° host, but was alleviated in a *supF* host. The open reading frame *orfG*, therefore, encodes a Ral-like anti-restriction activity that is almost certainly the Lar activity identified by Toothman (1981).

Localization of the Lar start codon

Examination of the DNA sequence at the start of *lar*

Table 3. Restriction-alleviation activity of *lar* mutants.

Host	Lar	Titre λ.0
NM777(DE3) pT7-6		3.3×10^7
NM777(DE3) pGK15	Wild type	4.0×10^{10}
NM777(DE3) pGK16	K42Am	2.5×10^7
NM777(DE3) pGK17	M1I	3.1×10^{10}
NM777(DE3) pGK18	M10I	2.8×10^{10}
NM777(DE3) pGK19	M14I	1.5×10^{10}
NM777(DE3) pGK20	M1Am	3.6×10^{10}
NM777(DE3) pGK21	M10Am	2.8×10^{10}
NM777(DE3) pGK22	M14Am	1.1×10^7
NM777(DE3) pGK23	M10Am M14I	2.6×10^7
YmeI(DE3) pT7-6		1.7×10^7
YmeI(DE3) pGK15	Wild type	2.7×10^{10}
YmeI(DE3) pGK16	K42Am	2.2×10^{10}
YmeI(DE3) pGK22	M14Am	2.1×10^7
YmeI(DE3) pGK23	M10Am M14I	8.9×10^6

Restriction-alleviation activity was determined for each host by measuring the titre of unmodified λ phages (λNM325.0) normalized against the titre of modified λ (see the *Experimental procedures*).

identified a number of possible translational start codons. The DNA codon ATG encodes the most common translational initiation codon (see McCarthy and Gualerzi, 1990, and references therein), and there are three of these close to the start of *orfG* at positions 1, 10, and 14 (see Fig. 3). To identify which ATG was the translational initiation codon, we mutated each ATG codon individually to one for isoleucine; the first ATG codon was changed to ATA (pGK17) and codons 10 and 14 were changed to ATC (pGK18 and pGK19, respectively). We anticipated that a mutation to the ATG codon required for translational initiation would abolish Lar activity. However, all three mutants retained restriction-alleviation activity (Table 3), indicating that a single ATG codon was unlikely to be wholly responsible for translational initiation.

To identify the minimum coding sequence required for restriction-alleviation activity, each of the three ATG codons was mutated to an amber codon. An amber codon is recognized as a translational 'stop' codon in a *sup*° host, whereas, in a *supF* host, a tyrosine residue is substituted at this position. An amber mutation in either the first (M1; pGK20) or second (M10; pGK21) ATG codons did not affect activity, but a change to the third ATG codon (M14; pGK22) abolished activity, both in a *sup*° and *supF* host. This showed that translation of the DNA sequence upstream of the second ATG codon was not required for restriction alleviation. To ascertain whether the third ATG codon could be mutated to one for isoleucine (ATC) when no translation occurred upstream of the second methionine codon, a double mutant was made (M10Am M14I; pGK23). This double mutant is unable to alleviate restriction in either a *sup*° or a *supF* host. To further the interpretation of these results, the polypeptides produced by each of the *lar* clones were examined.

```
AAGCTTCAGAGGAATAATTCAGCCTGGCGGTGTAATGCACCGC   3915
                                  M  H  R     Lar

CAACTTGAAATATTTTTTATGAGAAAAATTATGAGATATGAC    3960
 Q  L  E  I  F  F  M  R  K  I  M  R  Y  D     Lar
```

Fig. 3. The DNA and predicted aa sequence of the 5′ region of *lar*. The first 6 bp in this sequence comprise the *Hin*dIII recognition site, introduced by the PCR and used to clone *lar*. DNA sequence numbering is as in Clark *et al.* (1993).

Identification of the lar *gene product*

The products of the *lar* clones were visualized by the selective labelling of proteins encoded by genes under the control of the T7 promoter (Tabor, 1990). Differently sized polypeptides were detected depending on both the plasmid and the suppressor genotype of the host (Fig. 4, a and b). The host strains alone, BL21(DE3)*sup*° or BL21(DE3)*supF*, produced no labelled protein (data not shown), nor did pT7-6 in either the *sup*°(lane 1) or *supF* host (data not shown). pGK15 (wild-type *lar*; lanes 2 and 5) produced two main products, with estimated molecular masses of 8.8 kDa, and 8.3 kDa; the smaller of these is the more abundant.

The proteins produced by the *lar* mutants were similarly examined. In a *sup*° host, pGK16 (K42Am) produced no polypeptides of an appropriate size for Lar (lane 4), although a number of smaller products were seen, consistent with truncated forms of Lar. In contrast, in a *supF* host (lane 3) both the 8.8 kDa and 8.3 kDa proteins were produced by pGK16. This correlates well with the observed phenotype; pGK16 does not alleviate restriction in a *sup*° host, but does in a *supF* host.

Both pGK18 (M10I; lane 9) and pGK21 (M10Am), in a *sup*° (lane 10) or *supF* host (lane 11), produced only the 8.3 kDa polypeptide, from which we concluded that the second ATG (M10) is the initiation codon for the 8.8 kDa protein.

The third ATG codon (M14) is the major translational start of *lar*, encoding the 8.3 kDa protein, as this polypeptide was not detected in either a *sup*° or a *supF* host when the third ATG codon was mutated to an amber codon (M14Am; pGK22; lanes 13 and 14, respectively). Nevertheless, pGK19 (M14I; lane 12) and pGK23 (M10Am M14I; lanes 15 and 16) both produced small quantities of the the 8.3 kDa protein. The ATC codon in pGK19 and pGK23, therefore, appears to be used as a poor translational start in these mutants. It seemed unlikely that the codons between the second and third ATG codons could encode the translational start of the 8.3 kDa protein as this polypeptide was not visualized in a *supF* host transformed with pGK22 (M14Am), and yet the 8.8 kDa polypeptide (starting at M10) was detected (lane 14).

A 10 kDa polypeptide, specific to the *lar* clones, can be

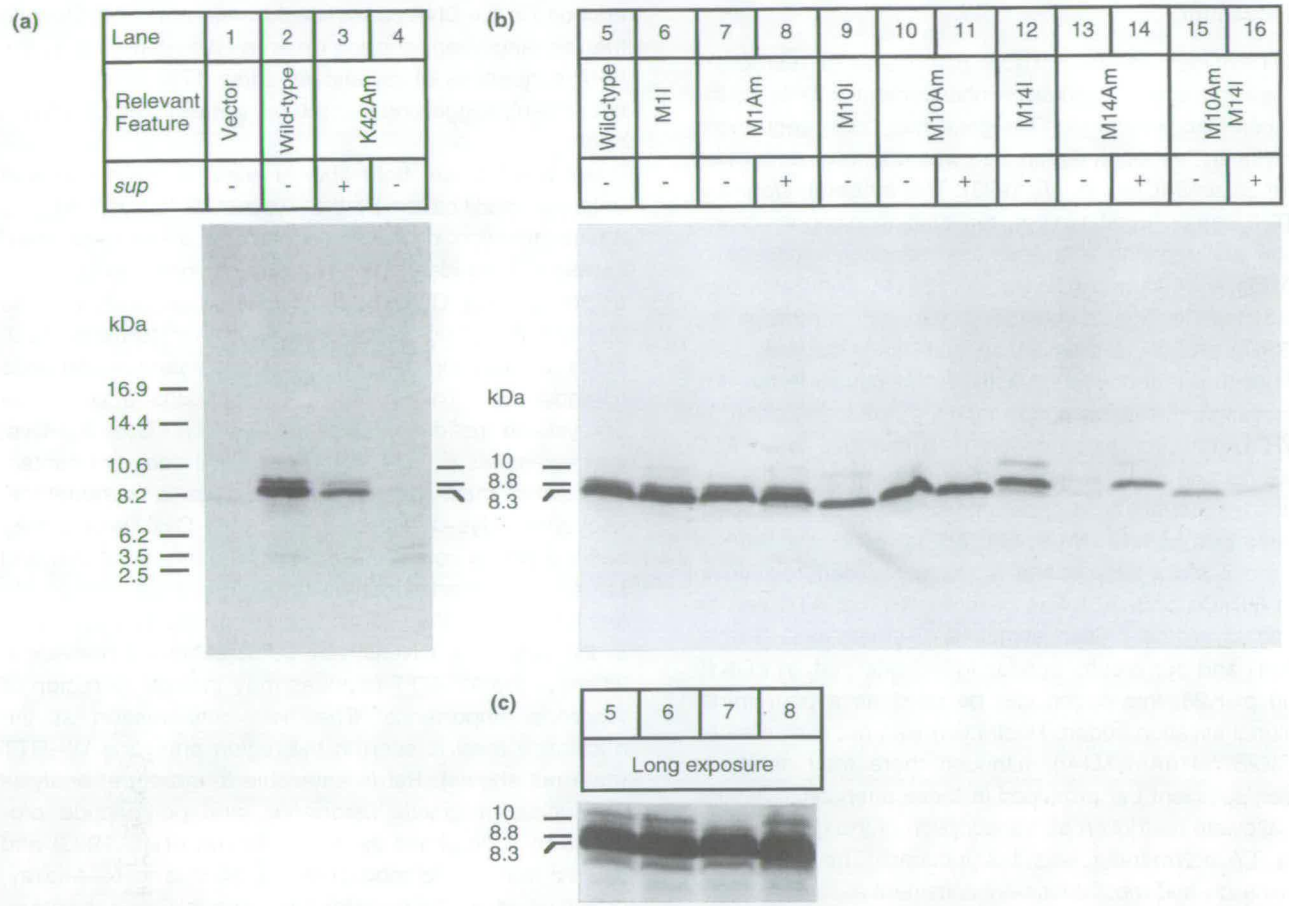


Fig. 4. Visualization of the *lar* gene products.
a. Lane 1, BL21(DE3) pT7-6 (vector); lane 2, BL21(DE3) pGK15 (wild-type); lane 3, NM793(DE3) pGK16 (K42Am); lane 4, BL21(DE3) pGK16.
b. Lane 5, BL21(DE3) pGK15 (wild-type); lane 6, BL21(DE3) pGK17 (M1I); lane 7, BL21(DE3) pGK20 (M1Am); lane 8, NM793(DE3) pGK20; lane 9, BL21(DE3) pGK18 (M10I); lane 10, BL21(DE3) pGK21 (M10Am); lane 11, NM793(DE3) pGK21; lane 12, BL21(DE3) pGK19 (M14I); lane 13, BL21(DE3) pGK22 (M14Am); lane 14, NM793(DE3) pGK22; lane 15, BL21(DE3) pGK23 (M10Am M14I); lane 16, NM793(DE3) pGK24.
c. A longer exposure of the autoradiograph showing the 10 kDa Lar polypeptide in lanes 5–8.

visualized after longer exposures of the autoradiograph (Fig. 4c). The initiation codon for this low-abundance polypeptide must precede the first ATG codon since, in a *supF* host, pGK20 (M1Am) produced the 10 kDa, 8.8 kDa and 8.3 kDa proteins (lane 8) but, in a *sup*° host, this clone only produced the two smaller polypeptides (lane 7). pGK17 (M1I) did not affect the polypeptides produced (lane 6).

```
Lar MRYDNVKPCPFCGCPSVTVKAISGYYRAKCNGCESRTGYGG 54
                ...:. .. :| :: :||||. ..:: .
Ral ............MTTTIDKNQWCGQFK.RCNGCKLQSECMV 37

    SEKEAL.........ERWNKRTTG.........NNNGGVHV 77
    ...| :           ::|. |||:           ||.::
    KPEEMFPVMEDGKYVDKWAIRTTAMIARELGKQNNKAA... 66
```

Fig. 5. Comparison of the predicted amino acid sequence of Ral and Lar (starting at M14) (see text). Gap weight 3.0; length weight 0.1; lines indicate identity; dots indicate similarity; % similarity, 52%; % identity, 25%.

Comparisons between the predicted amino acid sequences of Ral and Lar

The minimum aa sequence required for Lar activity starts at the methionine encoded by the third ATG codon (M14). The predicted aa sequence of this polypeptide was compared with that predicted for Ral using the GAP Program of the University of Wisconsin Genetics Computer Group (GCG 7.3, 1991; see Fig. 5). The two proteins are much more dissimilar than λ Ral and P22 Ral (Semerjian *et al.*, 1989). A number of regions do align, and these may indicate important structural and/or functional roles for residues within the two proteins.

Discussion

We have identified *lar*, the gene responsible for restriction-alleviation and modification-enhancement activity of the *E. coli* Rac prophage. This gene has been sequenced previously, although its function was unknown and it was termed *orfG* (Clark *et al.*, 1993). The *lar* gene, cloned in pT7-6, can be translated from three initiation codons to produce polypeptides with estimated molecular masses of 10 kDa, 8.8 kDa and 8.3 kDa. pGK21 (M10Am) alleviates restriction in a *sup*° host and yet only produces the 8.3 kDa protein, demonstrating that this is the only form of Lar that is required for activity. We conclude that the translational initiation codon for this protein is codon 14 (ATG). However, when codon 14 is changed to an ATC (pGK19 and pGK23) the 8.3 kDa polypeptide is still produced, although, as is apparent from the labelling experiments with pGK19 (M14I; Fig. 4B, lane 12), the amount of the 8.3 kDa polypeptide is much reduced relative to the 8.8 kDa protein. It has been shown that ATC can be used to initiate protein synthesis (Romero and Garcia, 1991) and our results appear to indicate that, in pGK19 and pGK23, this codon can be used as a poor translational initiation codon. Restriction was not alleviated by pGK23 (M10Am M14I), although there may not have been sufficient Lar produced in these phenotypic assays to alleviate restriction as transcription of the gene encoding T7 polymerase was not induced. The restriction-alleviation and modification-enhancement assays provided clear phenotypes with uninduced, wild-type *lar* (pGK15). It seemed preferable to use these conditions to examine Lar activity rather than those in which the cell physiology is perturbed by the gross overproduction of Lar, dependent upon T7 RNA polymerase. The initiation codon for the 10 kDa form of Lar appears to be before the first ATG codon (M1), and that for the 8.8 kDa polypeptide is the second ATG codon (M10). The 8.3 kDa protein is the most abundant, although it is not possible to predict what forms would be produced from the Rac prophage and whether the different forms have any biological significance.

The Rac prophage includes genes that can serve the same functions as the early genes of phage λ, including integration, recombination, immunity and replication (Low, 1973; Gottesman *et al.*, 1974; Diaz *et al.*, 1979). The *recE* and *recT* genes from Rac can substitute for the general recombination genes of λ, *exo* and *bet*, respectively (Kushner *et al.*, 1974; Clark *et al.*, 1993). In λ, *ral* is upstream of the recombination genes, whereas in the Rac prophage, *lar* is immediately downstream of *recT*. The *ral*, *exo* and *bet* genes are all transcribed from p_L in λ and it is logical to expect that they are also cotranscribed in Rac. Other genes in this transcriptional unit include those required by λ for excision (*xis*) and integration (*int*); analogous functions in the Rac prophage may be encoded in the DNA sequence downstream of *lar*. Despite the rearrangement of gene order in Rac relative to λ, the DNA sequences of *ral* and *lar* show 47% identity (data not shown), suggesting that the two genes have a common origin.

Ral and Lar are both able to alleviate restriction and enhance modification by the Type IA R-M systems and an examination of their predicted aa sequences may, therefore, help identify key residues involved in structure and/or function. Cysteine residues are commonly used to stabilize the structure of proteins, either forming disulphide bridges or associating with divalent metal ions (Branden and Tooze, 1991). Lar contains a sequence of cysteine residues, $Cys-X_2-Cys-X_{17}-Cys-X_2-Cys$, that resembles a zinc-binding motif. Ral does not contain the same motif but it does contain a symmetrical sequence, $Cys-X_5-Cys-X_2-Cys-X_5-Cys$, which may serve a similar role. When the aa sequences of Ral and Lar are compared, a number of regions align. The identities CNGC and W--RTT may be relevant. The cysteines in the sequence CNGC may be of structural relevance, whereas the W--RTT residues may indicate a region of functional importance. The best conservation at the nucleotide level is seen in the region encoding W--RTT (data not shown). Ral is amenable to structural analysis by nuclear magnetic resonance, and polypeptide produced by solid-phase synthesis (Brown *et al.*, 1993) and *in vivo* from the cloned gene (G. King and N. Murray, unpublished) will be used for this purpose. Given the structure of Ral, the motifs identified here should help us to further distinguish important residues and define targets for site-directed mutagenesis.

Experimental procedures

Strains, phages, and microbiological methods

Bacterial strains and phages are described in Table 4. Media and general methods have been described previously (Murray *et al.*, 1977). NM710 was used as the host for the preparation of K-modified lysates (λ.K), and NM679 was used as the host for unmodified lysates (λ.0). Bacteria were made lysogenic for λDE3 by a co-infection with λNM75, to provide Int *in trans*, and lysogens with the immunity of phage 21 were selected by plating on L-agar seeded with λ*imm*^{21}*cI* (λNM507) and a derivative with the host range of φ80 (λNM508).

NM793(DE3) was made by a series of P1 transductions. An efficient donor of *supF*, tagged with a closely linked transposon, was generated by the P1-mediated transfer of *trpC*::Tn*10* from a *sup*° strain (W3110*trpC*::Tn*10*) to the *supF* strain, Ymel. Tetracycline-resistant transductants of Ymel were selected on L-agar supplemented with 10 μg ml^{-1} tetracycline, and screened for the retention of *supF* by their ability to suppress λ*Sam7* (λNM123) but not λ*Pam3* (λNM435). Phage P1 grown on this host (NM787) was used to transduce BL21(DE3). Tetracycline-resistant transductants

Table 4. Strains and phages.

Strain/Phage[a]	Relevant features	Source
E. coli K-12 derivative		
NW2	Δ(*mrr–hsd–mcrB*)	Woodcock *et al.* (1989)
CAG12081	*zcj3061*::Tn*10fnr501*	Singer *et al.* (1989)
W3110	*sup*°	C. Yanofsky
W3110*trpC*::Tn*10*	*trpC*::Tn*10*	C. Yanofsky
NM679	W3110Δ(*mrr–hsd–mcr*)	See text
NM777[b]	W3110*lar*	See text
NM794	NM777*hsdR*Δ*4*	See text
LE451	C600*recAsrI*::Tn*10*Δ(*rac*)	Diaz *et al.* (1979)
NM710	C600*srI*::Tn*10*Δ(*rac*)	See text
NM792	NM710*hsdR*Δ*4*	See text
YmeI[b]	*supF*	Lab. collection
NM526[b]	YmeI *hsdR*Δ*4*	See text
NM787	YmeI *trpC*::Tn*10*	See text
E. coli B derivative		
BL21[b]	*sup*°*hsdS*	Studier and Moffat (1986)
NM793[b]	BL21*supF*	See text
Phage		
P1*kc*		Lennox (1955)
DE3	λ*int*::*lacI lac*UV5-T7gene*1imm*^{21}	Studier and Moffat (1986)
NM22	λ*Nam7Nam53cI26*	Lab. collection
NM75	λ*b2red3imm*$^{\lambda}$	Lab. collection
NM123	λ*cI857Sam7*	Lab. collection
NM243	λ*vir*	Lab. collection
NM325	λ*bio232*Δ(*int–ral*)*imm*^{434}*cI*	Lab. collection
NM435	λ*Pam3cI26*	Lab. collection
NM507	λ*b2imm*^{21}*cI*	Lab. collection
NM508	λ*h*^{80}*imm*^{21}*cI*	Lab. collection
NM891	λ*recE*	Kaiser and Murray (1979)

a. NM numbers refer to stock numbers in this laboratory.
b. Lysogenic for λDE3, where appropriate.

were selected for as before and screened for the presence of *supF* by their ability to suppress λ*Nam7Nam53* (λNM22), a phage chosen to avoid complications of complementation by the resident *imm*^{21} prophage. Phage P1 was grown on the *trpC*$^{+}$ host and used to regenerate a prototrophic strain (NM793). These were selected on M9-minimal agar and screened for the presence of the supressor mutation as described above.

The *lar* and *hsd* deletion derivatives of W3110 were also made by P1 transduction. The former (NM777) using CAG12081 (Singer *et al.*, 1989) as the donor of *fnr501*(Δ*fnr–lar*), and the latter (NM679) by first introducing the temperature-sensitive lethal mutation *dnaC325* by cotransduction with *zjj*::Tn*10*. *dnaC*$^{+}$ recombinants were selected following transduction with a P1 lysate made on the *hsd* deletion strain, NW2 (Woodcock *et al.*, 1989). Transductants selected at 42°C were screened for Hsd^{-}, McrB^{-} and tetracycline sensitivity.

The *hsdR*Δ4 derivatives of YmeI (NM526), NM710 (NM792) and NM777 (NM794) were made by transferring *hsdR*Δ4 from a λ*hsd* phage to the bacterial chromosome (Gough and Murray, 1983). NM710 is a *recA*$^{+}$ derivative of the *rac* deletion strain LE451 (Diaz *et al.*, 1979), made by transferring the *recA*$^{+}$ allele from a λ*recA*$^{+}$ phage to the host chromosome.

Restriction-alleviation tests

Bacteria transformed with plasmids were grown at 37°C until mid-log phase, in L-broth supplemented with 10 mM $MgSO_4$ and 100 μg ml^{-1} ampicillin. Samples (0.2 ml) were infected with appropriate dilutions of λNM325.0 or λNM325.K for 15 min, and plated in top layer on BBL-agar to assay infective centres. Plaques were counted after overnight incubation and the titre of λNM325.0 was normalized against the titre of λNM325.K.

Modification-enhancement tests

Bacteria transformed with plasmids were grown as outlined above. Samples (0.2 ml) were infected with λNM325.0 at a multiplicity of infection of 0.1, and incubated for 15 min to allow adsorption of phages. The infected cells were harvested and resuspended in 10 ml of L-broth supplemented with 100 μg ml^{-1} ampicillin. Cultures were incubated with aeration at 37°C for 70 min to allow a single round of infection, and the proportion of modified progeny phage was estimated from the titre on NM777 relative to that on a restriction-deficient derivative (NM794). λNM325.K was used as a control to ensure that the plating efficiencies of the indicator bacteria were comparable.

Plasmids, DNA manipulation and sequencing

Preparation, cloning and analysis of DNA was carried out essentially as described by Midgley and Murray (1985). Competent bacteria were prepared by electroporation (Heery and Dunican, 1989) and small-scale DNA isolation performed by the method of He *et al.* (1990). Details of the construction of plasmids used to localize restriction-alleviation activity are given in the legend to Fig. 1. The PCR, and site-directed mutagenesis using the PCR, were performed as described by Higuchi (1990), with a concentration of 50 μM dNTPs, and 0.2 μM primers. DNA fragments generated by the PCR were cloned in pT7-6 (Tabor, 1990), which does not have a ribosome-binding site upstream of the polylinker and therefore relies on sequences cloned downstream of the T7 promoter to initiate translation. Template DNA used in the PCR was λNM891 unless otherwise stated. The oligonucleotide primers used to generate the insert in pGK15, and flanking primers used for site-directed mutagenesis, are 5′-GGAAGCTTCAGAGGAATAATTCAG and 5′-GGGAATTCTTCCTTTTCAATAGTGG. Complementary pairs of primers were used to generate mutations and only one primer of each pair is described. The primers used to generate the mutations are; pGK16 5′-TACCGAGCGTAGTGTAAC, pGK17 5′-GTGTAATACACCGCCAAC, pGK18 5′-TTTATCAGAAAAATTATGAG, pGK19 5′-GAAAAATTATCAGATATGAC, pGK20 5′-GTGTATAGCACCGCCAAC, pGK21 5′-TTTTAGAGAAAAATTATGAG, pGK22 5′-GAAAAATTTAGAGATATGAC. The −40 universal primer, 5′-GTTTTCCCAGTCACGAC, was used as a flanking primer with pGK11 DNA template in the PCR to generate the fragment cloned in pGK17 and pGK20; the

usual flanking primer annealed too close to the mutagenic primer to generate a manageable product. pGK21 was used as the template in the PCR to generate the insert in pGK23 with the mutagenic primers used for pGK19, in order to generate a double mutant. All plasmid inserts generated by the PCR were analysed by double-strand DNA sequencing, using Sequenase 2.0 (Amersham Life Science). pGK17 and pGK20 both have an additional mutation (G to A) outside any ORF, immediately 5′ to the PCR primer. We were unable to isolate mutants using these primers without a mutation at this position. All other plasmids contain the expected sequence (Clark *et al.*, 1993).

Analysis of polypeptides

Gene products were visualized by the selective labelling of polypeptides encoded by genes under the control of the T7 promoter, essentially as described by Tabor (1990). BL21(DE3) and NM793(DE3) were used as the hosts for these experiments and expression of the T7 polymerase gene was induced by the addition of 0.6 mM IPTG (Studier and Moffat, 1986). Samples were denatured and reduced in 'cracking' buffer (50 mM Tris-HCl pH 6.8, 2% SDS, 0.1% bromophenol blue, 10% glycerol, 10% 2-mercaptoethanol, 100 mM DTT), the polypeptides separated by Tricine–SDS–PAGE (Shägger and von Jagow, 1987), and visualized by autoradiography.

Acknowledgements

We thank S. Tabor and C. Yanofsky for providing strains and plasmids, A. J. Clark for sharing DNA sequences with us prior to publication, A. Coulson for helpful discussions on sequence alignments and comparisons, and L. Powell, M. O'Neill and P. Thorpe for critical reading of the manuscript. This work was supported by grants from the Medical Research Council.

References

Bandyopadhyay, P.K., Studier, F.W., Hamilton, D.L., and Yuan, R. (1985) Inhibition of the type I restriction-modification enzymes *Eco*B and *Eco*K by the gene *0.3* protein of bacteriophage T7. *J Mol Biol* **182:** 567–578.

Bickle, T.A., and Krüger, D.H. (1993) *Biol Rev* **57:** 434–450.

Brandon, C., and Tooze, J. (1991) Structural motifs of eukaryotic transcription factors. In *Introduction to Protein Structure*. New York: Garland Publishing, pp. 113–127.

Brown, A.R., Irving, S.L., and Ramage, R. (1993) Affinity purification of synthetic peptides and proteins on porous graphitized carbon. *Tet Lett* **34:** 7129–7132.

Clark, A.J., Sharma, V., Brenowitz, S., Chu, C.C., Sandler, S., Satin, L., Templin, A., Berger, I., and Cohen, A. (1993) Genetic and molecular analyses of the C-terminal portion of the *recE* gene from the Rac prophage of *E. coli* K12 reveal the *recT* gene. *J Bacteriol* **175:** 7673–7682.

Diaz, R., Barnsley, P., and Pritchard, R.H. (1979) Location and characterisation of a new replication origin in the *E. coli* K12 chromosome. *Mol Gen Genet* **175:** 151–157.

Dryden, D.T.F., Cooper, L.P., and Murray, N.E. (1993) Purification and characterisation of the methyltransferase from the type I restriction and modification system of *Escherichia coli* K12. *J Biol Chem* **268:** 13228–13236.

Gottesman, M.M., Gottesman, M.E., Gottesman, S., and Gellert, M. (1974) Characterisation of bacteriophage λ*reverse* as an *E. coli* phage carrying a unique set of host-derived recombination functions. *J Mol Biol* **88:** 471–487.

Gough, J.A., and Murray, N.E. (1983) Sequence diversity among related genes for recognition of specific targets in DNA molecules. *J Mol Biol* **166:** 1–19.

He, M., Wilde, A., and Kaderbhai, M.A. (1990) A simple single-step procedure for small-scale preparation of *E. coli* plasmids. *Nucl Acids Res* **18:** 1660

Heery, D.M., and Dunican, L.K. (1989) Improved efficiency M13 cloning using electroporation. *Nucl Acids Res* **17:** 8006.

Higuchi, R. (1990) Recombinant PCR. In *PCR Protocols*. Innis, M.A., Gelfand, D.H., Sninsky, J.J., and White, T.J. (eds). San Diego: Academic Press, pp. 177–196.

Kaiser, K., and Murray, N.E. (1979) Physical characterisation of the 'Rac prophage' in *E. coli* K-12. *Mol Gen Genet* **175:** 159–174.

Kelleher, J.E., Daniel, A.S., and Murray, N.E. (1991) Mutations that confer *de novo* activity upon a maintenance methylase. *J Mol Biol* **221:** 431–440.

Kolodner, R., Hall, S.D., and Luisideluca, C. (1994) Homologous pairing proteins encoded by the *Escherichia coli* *recE* and *recT* genes. *Mol Microbiol* **11:** 23–30.

Kushner, S.R., Nagaishi, H., and Clark, A.J. (1974) Isolation of exonuclease VIII: the enzyme associated with the *sbcA* indirect suppressor. *Proc Natl Acad Sci USA* **71:** 3593–3597.

Lennox, E.S. (1955) Transduction of linked genetic characters of the host by bacteriophage P1. *Virology* **1:** 190–206.

Loenen, W.A.M., and Murray, N.E. (1986) Modification enhancement by the restriction alleviation protein (Ral) of bacteriophage λ. *J Mol Biol* **190:** 11–22.

Low, K.B. (1973) Restoration of the *rac* locus of recombinant forming ability in *recB*$^-$ and *recC*$^-$ merozygotes of *E. coli* K12. *Mol Gen Genet* **122:** 119–130.

McCarthy, J.E.G., and Gualerzi, C. (1990) *Trends Genet* **6:** 78–85.

Midgley, C.A., and Murray, N.E. (1985) T4 polynucleotide kinase — cloning of the gene (*pseT*) and amplification of its product. *EMBO J* **4:** 2695–2703.

Murray, N.E., Brammar, W.J., and Murray, K. (1977) Lambdoid phages that simplify the recovery of *in vitro* recombinants. *Mol Gen Genet* **150:** 53–61.

Powell, L.M., Dryden, D.T.F., Willcock, D.F., Pain, R.H., and Murray, N.E. (1993) DNA recognition by the *Eco*K methyltransferase. The influence of DNA methylation and the cofactor S-adenosyl-L-methionine. *J Mol Biol* **234:** 60–71.

Romero, A., and Garcia, P. (1991) Initiation of translation at AUC, AUA and AUU codons in *Escherichia coli*. *FEMS Microbiol Lett* **84:** 325–330.

Semerjian, A.V., Mallory, D.C., and Poteete, A.R. (1989) Genetic structure of the bacteriophage P22 p_L operon. *J Mol Biol* **207:** 1–13.

Shägger, H., and von Jagow, G. (1987) Tricine-sodium

dodecyl sulphate-polyacrylamidegel electrophoresis for the separation of proteins in the range from 1 to 100 kDa. *Anal Biochem* **166:** 368–379.

Simmon, V.F., and Lederberg, S. (1972) Degradation of bacteriophage lambda deoxyribonucleic acid after restriction by *Escherichia coli* K-12. *J Bacteriol* **112:** 161–169.

Singer, M., Baker, T.A., Schnitzler, G., Deischel, S.M., Goel, M., Dove, W., Jaacks, K.J., Grossman, A.D., Erickson, J.W., and Gross, C.A. (1989) A collection of strains containing genetically linked alternating antibiotic resistance elements for genetic mapping of *Escherichia coli*. *Microbiol Rev* **53:** 1–24.

Studier, F.W. (1975) Gene *0.3* of bacteriophage T7 acts to overcome the DNA restriction system of the host. *J Mol Biol* **94:** 283–295.

Studier, F.W., and Moffat, B.A. (1986) Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. *J Mol Biol* **189:** 113–130.

Suri, B., Nagaraja, V., and Bickle, T.A. (1984) Bacterial DNA modification. *Curr Top Microbiol Immunol* **108:** 1–9.

Tabor, S. (1990) Expression using the T7 RNA polymerase/promoter system. In *Current Protocols in Molecular Biology*. Ausubel, F.A., Brent, R., Kingston, R.E., Moore, D.D., Seidman, J.G., Smith, J.A., and Struhl, K. (eds). New York: Greene Publishing and Wiley-Interscience, pp. 16.2.1–16.2.11.

Templin, A., Kushner, S.R., and Clark, A.J. (1972) Genetic analysis of mutations indirectly suppressing *recB* and *recC* mutations. *Genetics* **72:** 205–215.

Toothman, P. (1981) Restriction alleviation by bacteriophage lambda and lambda reverse. *J Virology* **38:** 621–631.

Woodcock, D.M., Crowther, P.J., Doherty, J., Jefferson, S., DeCruz, E., Noyer-Weidner, M., Smith, S.S., Michael, M.Z., and Graham, M.W. (1989) Quantitative evaluation of *Escherichia coli* host strains for tolerance to cytosine methylation in plasmid and phage recombinants. *Nucl Acids Res* **17:** 3469–3478.

Yanisch-Perron, C., Vieria, J., and Messing, J. (1985) Improved M13 phage cloning vectors and host strains — nucleotide sequences of the M13mp18 and pUC18 vectors. *Gene* **33:** 103–109.

Zabeau, M., Friedman, S., Van Montagu, M., and Schell, J. (1980) The *ral* gene of phage λ. I. Identification of a non-essential gene that modulates restriction and modification in *E. coli*. *Mol Gen Genet* **179:** 63–73.

Zissler, J., Singer, E.R., and Schaefer, F. (1971) The role of recombination in growth of bacteriophage lambda. In *The Bacteriophage Lambda*. Hershey, A.D. (ed.). Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press, pp. 455–475.