LEABHARLANN CHOLÁISTE NA TRÍONÓIDE, BAILE ÁTHA CLIATH | TRINITY COLLEGE LIBRARY DUBLIN
Ollscoil Átha Cliath | The University of Dublin

Terms and Conditions of Use of Digitised Theses from Trinity College Library Dublin

Copyright statement

All material supplied by Trinity College Library is protected by copyright (under the Copyright and Related Rights Act, 2000 as amended) and other relevant Intellectual Property Rights. By accessing and using a Digitised Thesis from Trinity College Library you acknowledge that all Intellectual Property Rights in any Works supplied are the sole and exclusive property of the copyright and/or other IPR holder. Specific copyright holders may not be explicitly identified. Use of materials from other sources within a thesis should not be construed as a claim over them.

A non-exclusive, non-transferable licence is hereby granted to those using or reproducing, in whole or in part, the material for valid purposes, providing the copyright owners are acknowledged using the normal conventions. Where specific permission to use material is required, this is identified and such permission must be sought from the copyright holder or agency cited.

Liability statement

By using a Digitised Thesis, I accept that Trinity College Dublin bears no legal responsibility for the accuracy, legality or comprehensiveness of materials contained within the thesis, and that Trinity College Dublin accepts no liability for indirect, consequential, or incidental, damages or losses arising from use of the thesis for whatever reason. Information located in a thesis may be subject to specific use constraints, details of which may not be explicitly described. It is the responsibility of potential and actual users to be aware of such constraints and to abide by them. By making use of material from a digitised thesis, you accept these copyright and disclaimer provisions. Where it is brought to the attention of Trinity College Library that there may be a breach of copyright or other restraint, it is the policy to withdraw or take down access to a thesis while the issue is being resolved.

Access Agreement

By using a Digitised Thesis from Trinity College Library you are bound by the following Terms & Conditions. Please read them carefully.

I have read and I understand the following statement: All material supplied via a Digitised Thesis from Trinity College Library is protected by copyright and other intellectual property rights, and duplication or sale of all or part of any of a thesis is not permitted, except that material may be duplicated by you for your research use or for educational purposes in electronic or print form providing the copyright owners are acknowledged using the normal conventions. You must obtain permission for any other use. Electronic or print copies may not be offered, whether for sale or otherwise to anyone. This copy has been supplied on the understanding that it is copyright material and that no quotation from the thesis may be published without proper acknowledgement.

"Regulation of Antigen 43, a phase-variable autoaggregation factor of *Escherichia coli*"

A thesis submitted to Trinity College Dublin
for the degree of Ph.D.

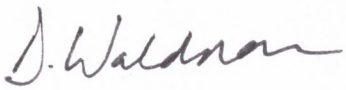
Denise Ethel Waldron

2002

TRINITY COLLEGE
18 MAR 2003
LIBRARY DUBLIN
THESIS
7214

DECLARATION

I hereby declare that this thesis represents entirely my own work, and has not previously been submitted for a degree at this or any other university. Trinity College library may lend or copy this thesis upon request.



Denise Waldron

ACKNOWLEDGEMENTS

First and foremost I want to express my gratitude to Peter and Charlie for giving me so much of their time and attention over the last few years. I feel very lucky to have had two supervisors to look out for me and help keep the project on track. I especially want to thank them for always showing a genuine concern for my health and well-being. I greatly appreciate the support given by everyone in prep room, who helped to keep things running smoothly from day-to-day. Thanks to Mary, Deirdre, Andy, Brian, Paul, Clare, Gerry, Louise, Fionnuala and Rob for making the bunker lab a fun place to work. Mary has always been extremely helpful and interested in my work and has definitely helped to keep me (reasonably) sane throughout my Ph.D. Thanks to all in the CJD lab (Matt, Orla, Ronan, Sorcha, Padraig, Arlene, Stephen, Christophe and Susan) for integrating me into their group and being such a great laugh. I especially want to thank Fionnuala and Orla, both of whom I will miss very much, for being such great friends and always encouraging me to be more positive. Thanks also to my flatmates Ellie, Alison, Susan and Sarah for making Russell Avenue a nice place to come home to (despite our disastrous renovations). I couldn't have wished for a better bunch of people to live with. Finally, I want to thank my sister Karen and my parents for all the financial support, which helped to keep a roof over my head and keep the bank manager off my back.

SUMMARY

The Antigen 43 protein (Ag43) of *Escherichia coli,* encoded by the *agn43* gene, has previously been shown to be expressed in a phase-variable manner. A plasmid-borne fusion of the *agn43* regulatory region to the reporter gene *lacZ,* was constructed and used to study the regulation of Ag43 at the level of transcription. β-galactosidase expression, under the control of the *agn43* promoter was phase-variable in a background containing functional OxyR and Locked ON in an *oxyR* mutant, confirming that Ag43 phase variation is regulated, by OxyR, at the transcriptional level. In an *oxyR,* Locked ON background, the *agn43* promoter was induced in the stationary phase of growth, and was regulated in response to the environmental cues of temperature and oxygen availability. Primer extension analysis identified the *agn43* transcriptional start site and located the *agn43* promoter upstream of three 5'-GATC-3' sequences, which lie within the proposed OxyR binding site. Expression of *agn43* was elevated in an *rpos* mutant indicating that *agn43* is negatively regulated by RpoS. Phase variation of *agn43* was altered by an *lrp* mutation, in a manner dependent on the expression of Type-1 fimbriae, which probably involves a change in the redox state of the OxyR repressor. It has been suggested previously that the phase-variable expression of Ag43 results from a competition between Dam methylase and the OxyR repressor for these sites. The 5'-GATC-3' sequences were inactivated for methylation by site-directed mutagenesis and all possible combinations of inactive and active sites were assessed for effects on phase-variable expression of the *agn43* gene. Inactivation of any 5'-GATC-3' site individually had no effect; at least two sites had to be inactivated to disrupt the normal pattern of expression. Studies of OxyR interaction with *agn43* DNA showed that methylation of any two 5'-GATC-3' sites was necessary and

sufficient to block binding of the repressor. It was also found that the adenines of the second and third 5'-GATC-3' sites are required for OxyR binding, demonstrating that the sites for Dam methylation and for repressor binding are associated intimately. This is consistent with a competition model in which Dam and OxyR share a preference for specific DNA sequences in the regulatory region of the *agn43* gene

PUBLICATION

Denise E. Waldron, Peter Owen and Charles J. Dorman (2002). Competitive interation of the OxyR DNA-binding protein and the Dam methylase at the antigen 43 gene regulatory region in *Escherichia coli.* Mol Microbiol **44:** 509-520.

CONTENTS

		page
Chapter 1	**General Introduction**	1
1	Introduction	2
1.1	The *E. coli* outer membrane	5
1.2	Autotransporters	8
1.3	Ag43	10
1.3.1	Biochemical properties of Ag43	10
1.3.2	Phase-variable properties of Ag43	12
1.3.3	Biological functions of Ag43	14
1.3.3.1	Ag43 as an adhesin	14
1.3.3.2	Ag43 in autoaggregation and biofilm formation	15
1.3.3.3	Ag43 and bacteriophage infection	18
1.4	Mechanisms of phase and antigenic variation	19
1.4.1	Vmp outer membrane lipoproteins of Bor. hermsii	20
1.4.2	Shiga-like toxin II	20
1.4.3	Type IV pilin of *N. gonorrhoeae*	21
1.4.4	Opa proteins of *N. meningitidis* and *N. gonorrhoeae*	21
1.4.5	Opc of *N. meningitidis*	23
1.4.6	Phase variation by DNA inversion	23
1.5	Regulation of Ag43	24
1.5.1	Dam	24
1.5.1.1	Methylation of 5'-GATC-3' sites by Dam	25
1.5.1.2	Distribution of 5′-GATC-3′ sequences in the bacterial chromosome	26
1.5.1.3	Biological functions of Dam	27
1.5.2	OxyR	28
1.5.2.1	Regulation by OxyR	29
1.5.2.2	DNA binding properties of OxyR	30
1.5.3	OxyR and Dam in *agn43* regulation	32
1.5.4	Dam-dependent regulation of the phage Mu *mom* gene	33
1.5.5	Regulatory cross talk between Type-1 fimbriae and Ag43	35
1.6	Dam-dependent regulation of other operons	37
1.6.1	Phase variation of the *E. coli pap* operon	37
1.6.2	Dam-dependent phase variation of the *sfa, daa, foo* and *pef* operons	40
1.6.3	Phase variation of Type-1 fimbriae	41
1.7	Aims and Objectives	44

Chapter 2	**Materials and Methods**	47
2.1	Chemicals and growth media	47
2.1.1	Chemicals, enzymes and radionucleotides	47
2.1.2	Growth conditions and media	47
2.2	Bacterial strains, plasmids and bacteriophage	48
2.2.1	Bacterial strains and plasmids	48
2.2.2	Bacteriophage	48
2.3	Oligonucleotides	49
2.4	Genetic techniques	49
2.4.1	Use of the generalized transducing phage P1cml	49
2.4.2	Preparation of a high-titre P1cml lysate	49
2.4.3	Transduction with P1*cml*	50
2.4.4	Transformation of *E. coli* with plasmid DNA	51
2.4.4.1	Preparation of competent cells by the calcium-chloride method	51
2.4.4.2	Transformation of calcium chloride competent cells	51
2.4.4.3	Preparation of electrocompetent cells	52
2.4.4.4	Electroporation of electrocompetent cells	52
2.5	Assays based on spectrophotometry	53
2.5.1	Monitoring the growth of bacterial cultures	53
2.5.2	Assay of β-galactosidase activity	53
2.5.3	Determination of protein concentration by Bradford assay	54
2.6	Isolation of plasmid DNA	54
2.7	Purification of chromosomal DNA	55
2.8	DNA amplification by polymerase chain reaction	55
2.9	*In vitro* manipulation of DNA	55
2.9.1	Restriction endonuclease cleavage of DNA	55
2.9.2	Phosphatase treatment of restriction endonuclease cleaved DNA	56
2.9.3	Purification of linear DNA from agarose gels	56
2.9.4	Ligation of DNA	56
2.9.5	Dam methylation of PCR products	57
2.9.6	Ethanol precipitation of DNA	57
2.9.7	Phenol/chloroform extraction	57
2.9.8	5'-end labelling of DNA using $[\gamma^{32}P]$AT	58
2.10	DNA sequencing	58
2.11	Denaturing polyacrylamide gel electrophoresis	59
2.11.1	Pouring the sequencing gel	59
2.11.2	Loading and running the gel	60
2.11.3	Treatment of the gel prior to autoradiography	60

2.11.4	Autoradiography	60
2.12	Manipulation of RNA *in vitro*	60
2.12.1	RNA extraction	60
2.12.2	Primer extension analysis	61
2.12.2.1	Annealing of the primer and template	61
2.12.2.2	cDNA synthesis	62
2.12.2.3	Purification of the cDNA products	62
2.13	Site directed mutagenesis	62
2.13.1	Quickchange	62
2.13.2	Unique-site elimination	63
2.14	Determination of switching frequencies	64
2.15	Southern blotting	65
2.15.1	Generation of DIG-labelled probes by PCR	65
2.15.2	Transfer of DNA	65
2.15.3	Hybridisation of the probe to immobilised targets	66
2.15.4	Chemiluminescent detection of bound probe	66
2.16	Preparation of crude protein extracts	67
2.17	Electrophoretic mobility shift assays (EMSA)	67
2.18	Colony immunoblotting	68
2.19	*In vivo* footprinting by methylation protection assay	69
Chapter 3	**Cloning and analysis of the *agn43* promoter region**	70
3.1	Introduction	71
3.2	Results	72
3.2.1	Generation of an *agn43-lacZ* transcriptional fusion	72
3.2.2	Transformation of strains CSH50 and CSH50*oxyR* with plasmid pAg43	74
3.2.3	Coordinated expression of chromosomally-encoded Antigen 43 with the *agn43-lacZ* fusion	75
3.2.4	The effect of environmental conditions on *agn43* promoter function	76
3.2.4.1	The effect of growth phase on *agn43* promoter activity	76
3.2.4.2	The effect of temperature on *agn43* promoter activity	77
3.2.4.3	The effect of osmolarity and pH on *agn43* promoter activity	77
3.2.4.4	The effect of anaerobiosis on *agn43* promoter	77

	activity	
3.2.5	Mapping of the transcriptional start site of *agn43*	78
3.2.6	Promoter analysis	78
3.2.7	Methylation state of the *agn43* promoter region in Phase ON and Phase OFF cells	80
3.3	Discussion	82
Chapter 4	**The identification of factors affecting *agn43* expression**	86
4.1	Introduction	87
4.1.1	pBSL181	88
4.2	Results	90
4.2.1	The effect of an *rpoS* mutation on Ag43 expression	90
4.2.2	Phase variation of *agn43* is altered in an lrp mutant	91
4.2.3	Generation, by transposon mutagenesis, of mutants with altered *agn43-lacZ* expression	93
4.2.3.1	Mutagenesis of strain CSH50	93
4.2.3.2	Integrity of the reporter plasmid in mutants	95
4.2.3.3	PCR analysis of *oxyR* and *rpoS* genes in mutants	96
4.2.3.4	Marker rescue	96
4.2.3.5	Cloning and sequencing of mutated DNA from M4	97
4.3	Discussion	98
Chapter 5	**Mutational analysis of the 5′-GATC-3′ sites and the OxyR binding site within the *agn43* regulatory region**	103
5.1	Introduction	104
5.2	Results	107
5.2.1	Site-directed mutagenesis of 5'-GATC-3' sites	107
5.2.1.1	The Quickchange procedure	107
5.2.1.2	The Unique-site elimination method	108
5.2.2	Strategy for the generation of mutant constructs	109
5.2.2.1	Construction of plasmids pDEW10-pDEW16	110
5.2.2.2	Construction of plasmids pDEW20-pDEW25	111
5.2.3	Analysis of constructs pDEW10-pDEW16 in CSH50	112
5.2.4	The effect of OxyR overexpression on pDEW10-pDEW16	114

5.2.5	Disruption of 5′-GATC-3′ sites with alternative mutations	116
5.2.6	Phenotypic analysis of mutant constructs pDEW20-25 in strains CSH50 and CSH50OxyR^{++}	117
5.2.7	The effect of a *dam* mutation on mutant phenotypes	118
5.2.8	The effect of an *oxyR* mutation on mutant phenotypes	120
5.2.9	Binding of OxyR to *agn43* DNA	120
5.2.9.1	Binding of OxyR to the wildtype *agn43* regulatory region	122
5.2.9.2	Binding of OxyR to unmethylated DNA from mutant constructs	123
5.2.9.3	Binding of OxyR to methylated DNA from mutant constructs	124
5.2.10	Footprinting analysis of OxyR binding	126
5.3	Discussion	127
Chapter 6	**General discussion**	140
	Bibliography	153

FIGURES AND TABLES

Figure	Previous page
1.1	8
1.2	18
1.3	26
1.4	32
1.5	32
1.6	33
1.7	35
1.8	36
1.9	39
1.10	39
1.11	43
3.1	73
3.2	74
3.3	75
3.4	76
3.5	77
3.6	78
3.7	79
3.8	79
3.9	79
3.10	81
3.11	85
4.1	89
4.2	90
4.3	90
4.4	92
4.5	96
4.6	97
4.7	102
5.1	108
5.2	108
5.3	108
5.4	111
5.5	116
5.6	121
5.7	122
5.8	125

5.9 125
5.10 135
5.11 138

Tables **Previous page**

1.1 23
1.2 28
2.1 69
2.2 69
2.3 69
3.1 77
3.2 77
3.3 77
4.1 92
4.2 95
5.1 109
5.2 109
5.3 113
5.4 116
5.5 118
5.6 119
5.7 119

ABREVIATIONS

Bacterial species

E. coli	*Escherichia coli*
Bor. hermsii	*Borrelia hermsii*
N. meningitidis	*Neisseria meningitidis*
N. gonorrheae	*Neisseria gonorrheae*
P. fluorescens	*Pseudomonas fluorescens*
S. typhimurium	*Salmonella typhimurium*
V. cholerae	*Vibrio cholerae*

Others

α^{43A}	α^{43} subunit encoded by *agn43A* gene
α^{43B}	α^{43} subunit encoded by *agn43B* gene
β^{43A}	β^{43} subunit encoded by *agn43A* gene
β^{43B}	β^{43} subunit encoded by *agn43B* gene
Adomet	s-adenosylmethionine
Ag43	Antigen 43
Ap	ampicillin
BSA	bovine serum albumin
Cah	calcium-binding Antigen 43 homologue
CAP	catabolite activator protein
Cm	chloramphenicol
Dam	deoxyadenosine methyl transferase

DEPC	diethylpyrocarbonate
DIG	digoxigenin
DMS	dimethylsulphate
DTT	dithiothreitol
EHEC	enterohaemmoragic *E. coli*
EMSA	electrophoretic mobility shift assay
EPEC	enteropathogenic *E. coli*
H-NS	histone nucleoid-like structuring protein
IHF	integration host factor
IPTG	isopropyl-β-D-galactopyranosidase
Km	kanamycin
LRP	leucine-responsive regulatory protein
MU	Miller units
ONPG	o-nitrophenyl-β-D-galactopyranosidase
Opa	opacity protein
PCR	polymerase chain reaction
SD	standard deviation
SDS PAGE	sodium-dodecyl sulphate polyacrylamide gel electrophoresis
SLT	shiga-like toxin
T4 PNK	T4 polynucleotide kinase
Tc	tetracycline
TEMED	N, N, N', N'-tetramethylethylene diamine
Vmp	variable membrane protein
X-Gal	5-bromo-4-chloro-3-indoyl-β-D-galactosidase

Chapter 1

General Introduction

1 Introduction

Escherichia coli (*E. coli*) is a species of the bacterial family *Enterobacteriaceae*, which consists of both commensal and pathogenic strains. The organism is a facultatively anaerobic Gram-negative rod that lives in the intestinal tracts of mammals, including humans. *E. coli* is spread between hosts via the faecal-oral route. The majority of *E. coli* strains are non-pathogenic. However, there are a large number of variants that cause infectious disease. Pathogenic *E. coli* are responsible for intestinal infections, urinary tract infections and neonatal meningitis. The ability of pathogenic stains to cause infection is often dependent on the presence of virulence determinants on the cell surface (Miller *et al.*, 1994).

The ability of bacterial populations to adapt rapidly to environmental changes is an important aspect of microbial survival (DiRita and Mekalanos, 1989). The expression of particular proteins may only be required at certain times. As such, the ability to regulate the expression of proteins in response to growth conditions is important from an economical perspective and leads to increased bacterial competitiveness (Goetz *et al.*, 1989). The presence or absence of these environmental signals affects different regulatory proteins that in turn regulate the level of expression of a particular gene.

The phase-variable manner in which some virulence factors are expressed further enhances the contribution that they make to survival of the bacterial population, particularly when encountering a change in environment (van der Woude *et al.*, 1996). Phase variation is the reversible switching between expressing (Phase ON), and non-expressing (Phase OFF) states and is a property of some fimbriae, pili, flagella and a number of outer membrane proteins (Owen *et al.*, 1996). Phase variation can be

reversible (Phase ON ↔ Phase OFF) or irreversible (Phase ON→ Phase OFF). Another property that leads to phenotypic alterations is antigenic variation, which involves the expression of one or more of a number of different phenotypes (Saunders, 1990; Deitsch *et al.*, 1997).

Bacterial populations are better able to deal with sudden environmental changes if they are phenotypically heterogeneous. The likelihood of survival of the bacterial population is increased if a subpopulation has some selective advantage on entering a new environment. Thus, it seems that phase variation occurs to increase the survival of the population as a whole and not to improve the prospects of an individual cell. This seems contrary to basic principles of evolution which dictate that the survival of each cell within the population, and hence of the population as a whole, depends on the exhibition of a single optimum phenotype. However, phenotypic homogeneity would be detrimental to the survival of a population, which would be less able to deal with sudden changes in environmental conditions. It follows that a major function of phase and antigenic variation could be to create populations capable of surviving within more than one environment (Saunders, 1989). It is a fact that entire populations of infecting bacteria can be derived from a single clone (Moxon and Murphy, 1978). Considering this, efficient mechanisms for the generation of phenotypic diversity seem vital for pathogens, which can encounter sudden changes that may threaten their survival.

Bacterial surface proteins are likely to be highly immunogenic and are easy targets for protective antibodies (Saunders, 1989). It has been shown that *Neisseria gonorrhoeae* are more resistant to phagocytosis in the absence of Opa proteins or the presence of particular Opa variants (Virji and Heckels, 1986). Thus, the most obvious role for

phase variation of antigenic proteins is to facilitate the evasion of host immune response. Phase variation may also enhance bacterial survival by other means. For example flagellar synthesis is switched off following biofilm formation in order to avoid its potential destabilizing effect on the biofilm structure (Prignent-Combart *et al.*, 1999). Also switching the expression of an adhesin OFF may allow the bacteria to disperse and colonize additional target sites.

A number of studies have been published which indicate that phase variation occurs *in vivo*. Swanson *et al.* (1988) recovered Opa^+ gonococci from humans originally inoculated with Opa^- bacteria. Similarly Fim^+ cells were recovered from mouse urinary tracts originally inoculated with Fim^- *E. coli* (Struve and Krogfelt, 1999; see Sections 1.4.4 and 1.6.3 for descriptions of Opa and Fim phase variation, respectively).

Another possible role for phase variation is to allow the bacteria to colonise different tissues within the host. For example Opa variants of *Neisseria* show different tissue specificities. Furthermore, analysis of clinical isolate has shown that a correlation exists between individual variants and their sites of infection (Virji *et al.*, 1991; Nassif *et al.*, 1993).

Phase variation may also have a role in facilitating bacterial dissemination between hosts. For example switching off expression of an adhesin might aid the dissemination within the host and transmission to another host. The causative agent of relapsing fever, *Borrelia hermsii*, exemplifies the way in which antigenic variation aids persistence, reinfection and spread within and between hosts (see Section 1.4.1). After exposure to one serotype a host immune response is triggered and this overcomes the

infection. In order to circumvent this immune response the bacteria adopt a new serotype leading to a relapse of the illness. Thus, antigenic variation facilitates increased longevity within the host thereby increasing the chances of dissemination. The subject of this dissertation is regulation of the phase-variable expression of the most abundant phase-variable protein, Antigen 43 (Ag43), in the outer membrane of *E. coli*.

1.1 The *E. coli* outer membrane

The outermost layer of the *E. coli* cell envelope is formed by the outer membrane, which is a lipid bilayer about 7 nm in thickness. The membrane is composed of phospholipids, lipoproteins, integral outer membrane proteins and lipopolysaccharide (LPS). The phospholipids, comprising phosphatidylethanolamine, phosphatidylglycerol and cardiolipin, are located mainly in the inner leaflet of the outer membrane. In the outer layer phospholipids are largely replaced by LPS. Thus, unlike the cytoplasmic membrane, the lipid bilayer of the outer membrane is highly assymetrical in structure. Also embedded in the inner monolayer are lipoproteins, some of which (e.g. the Braun lipoprotein) covalently anchor the outer membrane to the peptidoglycan below (outer membrane structure is reviewed by (Nikaido, 1988).

LPS is a high-molecular-weight, strongly negatively-charged molecule that consists of three regions (a) an inner Lipid A region, which inserts into the membrane, (b) a middle core of 8-12 variable sugars and 3-8 phosphate residues, and (c) the O antigen that extends outwards from the cell surface and consists of 3-5 sugar units that are repeated a variable number of times. Mutants that lack the O antigen are referred to as 'rough' strains because of their characteristic colony morphology. The LPS core contains

several unique sugars (heptoses and ketodeoxyoctonic acid), which are not found elsewhere in nature (Heinrichs *et al.*, 1998). LPS functions as an endotoxin when released from the outer membrane and triggers innate immune defences such as inflammation, fever and phagocytosis. Another likely function for LPS is to add strength and rigidity to the outer membrane (Rietschel *et al.*, 1998).

The most important function of the outer membrane is to serve as a protective barrier by acting as a molecular sieve. Thus, the outer membrane slows or prevents the entry of harmful substances, such as bile salts and antibiotics. The outer membrane is more permeable than the plasma membrane and permits the passage of small molecules like glucose and other monosaccharides. This is due to the presence of porins, β-barrel structures, which span the outer membrane to form channels through which small molecules of < 600-700 Da may pass. Larger molecules must be transported across the outer membrane by specific carriers. There are three classes of porins in the outer membrane comprising (a) general porins, which have minimal substrate specificity, (b) specific porins that are substrate-selective and have specific binding site within the channel, and (c) gated specific porins which have channels that are normally closed but are thought to open upon binding of a specific substrate (Nikaido, 1992; Koebnik *et al.*, 2000; Delcour, 2002).

Approximately 50% of the outer membrane's mass consists of proteins, which comprise lipoproteins and integral outer membrane proteins. All integral outer membrane proteins probably consist of β-pleated sheets in the form of a closed barrel, called a β barrel (Koebnick *et al.*, 2000). Integral outer membrane proteins include several that are expressed constitutively at high levels, such as the general porins and OmpA. Other

integral outer membrane proteins are expressed strongly only when needed. Included in this category are the TonB-dependent receptors (e.g. FhuA and FepA), components of protein export systems I to III, the autotransporters, various enzymes (e.g. phospholipase A) and proteins involved in the biosynthesis of fimbriae and flagella (Koebnik *et al.*, 2000). Outer membrane proteins are essential for the adaption of bacteria to the hostile host environment and are therefore important virulence factors (Lin *et al.*, 2002). Outer membrane proteins have diverse functions and include those involved in iron uptake (e.g. FhuA and FepA; Braun *et al.*, 1998), antimicrobial peptide resistance (e.g. OmpT; Stumpe *et al.*, 1998), serum resistance (e.g. OmpX and OmpA; Lin *et al.*, 2002), muti-drug resistance (e.g. TolC; Koronakis *et al.*, 2000) and bile resistance (e.g. OmpF and OmpC; Thanassi *et al.* 1997). The functions performed by many of outer membrane proteins are, however, not yet known.

Anchored to the *E. coli* outer membrane are regular surface structures such as flagella and fimbriae (Fernandez and Berenguer, 2000). Flagella are helical rigid structures that protrude form the cell and rotate to provide motility. A typical *E. coli* cell has 6 - 8 flagella anchored at random positions on the cell surface. Fimbriae, also known as pili, constitute another major class of surface appendage. Fimbriae are shorter and stiffer than flagellae and form hair-like structures on the bacterial surface. They are usually involved in the adherence of bacterial cells to surfaces and are major determinants of bacterial virulence (Sauer *et al.*, 2000). They are highly antigenic and the expression of certain types, such as Pap pili and the Type-1 fimbriae, is subject to phase variation (Blomfield, 2001; see Sections 1.6.1 and 1.6.3). Another membrane-anchored structure that is subject to phase variation is the surface protein Antigen 43 (Ag43). Ag43 is a bipartite protein complex and is anchored to the cell surface via its β-subunit, which is

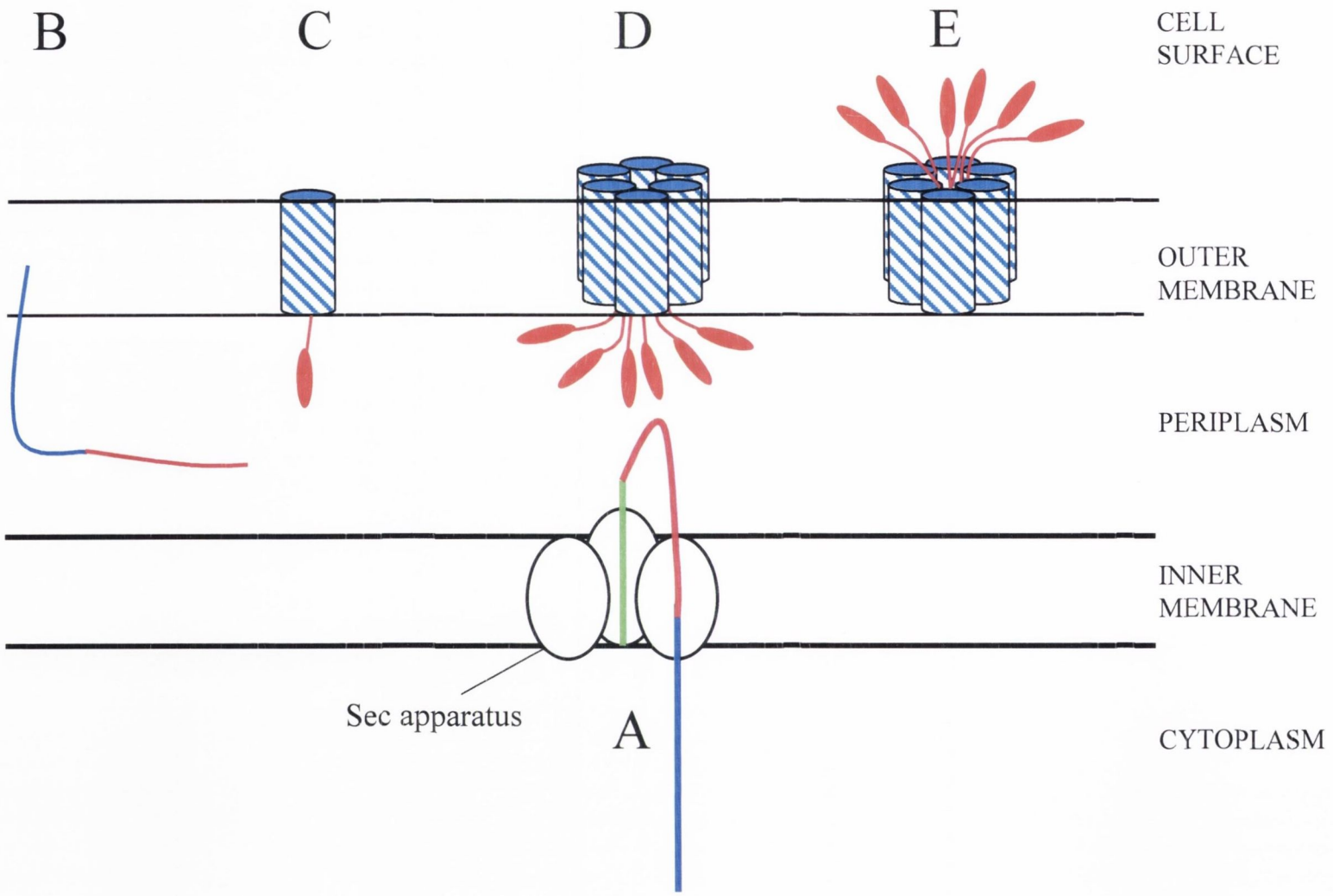
an integral outer membrane protein. Ag43 belongs to the family of autotransporter proteins, which are characterised by the fact that all of the information for transport to the outer membrane and secretion through the cell envelope is contained within the protein itself (Henderson *et al.*, 1998). The secretion of autotransporters is discussed in the next section.

1.2 Autotransporters

Nearly all members of the autotransporter family are comprised of three domains, which have all of the features required to mediate transport of the protein across the inner and outer membranes. These are (a) an N-terminal leader peptide, (b) the surface localized 'mature' protein, termed the passenger domain or α domain, and (c) a C-terminal β domain that mediates secretion through the outer membrane (Jose *et al.*, 1995).

It is thought that inner membrane translocation is mediated by the Sec system (Figure 1.1), because of the presence, within most autotransporters, of putative Sec-dependent signal sequences. The features of these signal sequences are (a) an N domain of positively charged amino acids, (b) a hydrophobic region of neutral amino acids, and (c) a C domain signal peptidase recognition site (Izard and Kendall, 1994). As yet, however, there is no conclusive evidence to support the assumption that the Sec system is involved in autotransporter export. Many of the autotransporters have unusually long signal sequences that are at least 47 amino acids in length (Henderson *et al.*, 1998). The function of this extended signal peptide is not yet known.

Figure 1.1 The proposed method of autotransporter expression. The polypeptide precursor is composed of three domains viz. a signal region (green), a passenger domain (red) and a C-terminal β-barrel-forming domain (blue). (A) Translocation through the inner membrane is thought to involve an interaction between the N-terminal signal and the Sec system. (B) Having crossed the inner membrane the C-terminal domain inserts into the outer membrane. (C) The C-terminal domain forms a β-barrel pore within the outer membrane. (D) Several such β barrels are assembled in this manner to form a ring-shaped oligomeric complex. (E) the passenger domains are translocated through a central hydrophilic pore formed by the oligomeric complex.


B
C
D
E
CELL SURFACE
OUTER MEMBRANE
PERIPLASM
INNER MEMBRANE
CYTOPLASM
Sec apparatus
A

After transport across the inner membrane (via the Sec apparatus) and cleavage of the signal peptide, the autotransporter is thought to exist as a periplasmic intermediate. It is, however, unclear how the protein could avoid extensive degradation by proteases in the periplasm. A periplasmic chaperone may be involved in protecting the intermediate from degradation. Alternatively inner and outer membrane translocation may be temporally coupled such that time spent in the periplasm is minimal. At present there is no experimental evidence to support either of these possibilities (Henderson *et al.*, 1998).

It is thought that the β domain of the autotransporter spontaneously inserts into the outer membrane in a β-barrel configuration. Until recently it was thought that the β barrel acts as a channel for translocation of the passenger domain (Suzuki *et al.*, 1995; Jose *et al.*, 1995, 1996; Loveless and Saier, 1997). Recently Veiga *et al.* (2002) provided evidence that autotransporter proteins assemble in a ring-shaped oligomeric structure within the outer membrane and that export of the passenger domains occurs through a common central channel shared by different subunits of this complex (Figure 1.1).

Maurer *et al.* (1999) have suggested the existence, within autotransporters, of an additional domain, termed the linker, between the passenger and β-barrel domains that is essential for the translocation of the N-terminal. Deletions within this linker region prevented surface expression of the cholera toxin B (CTB) subunit fused to the C-terminal domain of AIDA-1 (Maurer *et al.*, 1999).

Once it reaches the cell surface the passenger domain undergoes one of three events viz. (a), it is cleaved from the β-barrel/linker domain and is released into the extracellular milieu as is the case for gonococcal IgA1 protease (Pohlner *et al.*, 1987), (b), it remains non-covalently attached to the C-terminal domain following cleavage, as is the case for Ag43 (Henderson and Owen, 1999), or (c), cleavage does not occur resulting in a product with fused passenger and β-barrel domains, as is thought to occur for Hsr of *Helicobacter mustelae* (O'Toole *et al.*, 1994).

1.3 Ag43

1.3.1 Biochemical properties of Ag43

Ag43 was discovered during a crossed-immunoelectrophoresis analysis of membrane vesicles isolated from *E. coli* ML308-225, a derivative of a human intestinal isolate which has been used extensively in the study of membrane structure (Owen and Kaback, 1978). Ag43 is a prominent outer membrane protein and is found in most strains of *E. coli*. It is thought to be the most abundant phase variable outer membrane protein in *E. coli* and exists in numbers exceeding 50,000 per cell. Ag43 is a hetero-oligomeric complex composed of two distinct protein subunits, termed α^{43} and β^{43}, which are present in 1:1 stoichiometry. The primary product of the *agn43* gene, which encodes Ag43, is processed via an N-terminal signal peptidase into a 987-amino acid protein. This is thought to be followed by an autocatalytic processing event that results in an α subunit of 499 amino acids and a β subunit of 488 amino acids. The α^{43} subunit is surface expressed whereas β^{43} is an integral membrane protein with pronounced heat-modifiability (Owen *et al.*, 1987).

The β^{43} subunit of Ag43 is thought to form a β-barrel pore, which is a complex protein structure composed of multiple amphipathic antiparallel β-sheets surrounding a water-filled channel (Schirmer *et al.*, 1995; Henderson and Owen, 1999). The β-barrel structure probably facilitates the translocation of α^{43} to the cell surface. Following translocation, α^{43} remains bound to the cell surface via a non-covalent interaction with β^{43}, and can be selectively detached from the outer membrane by brief heating to 60°C (Owen, 1986; Owen *et al.*, 1987, 1996; Caffrey and Owen, 1989; Henderson and Owen, 1999). α^{43} shows extensive sequence identity with AIDA-1, which is an adhesin of diffuse adhering enteropathogenic *E. coli* (Benz and Schmidt, 1992; Henderson and Owen, 1999). Recently Torres *et al.* (2002) characterized a protein of enterohaemorrhagic *E. coli* (EHEC) serotype O157:H7 that shares homology with Ag43, which they called Cah (calcium-binding Antigen 43 homologue). The α^{43} and β^{43} subunits of Ag43 contain no detectable carbohydrate, acyl groups or inter-/intra-molecular disulfide bonds. Nearest neighbour analysis indicates that Ag43 is situated near to FepA, the ferric-enterochelin receptor (Owen *et al.*, 1987, 1996).

Initial studies by Henderson and Owen (1999) resulted in the detection of just one *agn43* gene and product within the strain ML308-225. The M_r values, predicted from sequence initially obtained from this strain, were 49 789 and 51 642 for α^{43} and β^{43}, respectively. More recent studies by Roche *et al.* (2001) revealed that this strain contains two *agn43* genes, which were subsequently termed *agn43A* and *agn43B*. Independent sequencing of both *agn43* genes from ML308-225 revealed that the original sequence data obtained by Henderson and Owen (1999) contained inaccuracies, which were most likely due to cross amplification of the duplicate gene copies (Roche *et al.*, 2001). The sequence of the surface expressed α^{43} entities, encoded by *agn43A*

and *agn43B*, differ by three amino acids. The α subunit encoded by *agn43B* contains alanines residues at positions 226 and 230, where the *agn43A* product contains threonine residues. The predicted M_rs for the α^{43A}, α^{43B}, β^{43A} and β^{43B} subunits were 49 904, 49 795, 51 554 and 51 633, respectively. It is possible that the amino acid differences between the gene products result in changes in O-glycosylation patterns. The differences do alter the molecular size, as evidenced by sodium-dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE), as well as the apparent epitope expression of α^{43}. It is therefore conceivable that the presence of duplicate copies of *agn43* may allow the bacteria to express antigenically distinct forms of the protein (Roche *et al.*, 2001). Thus, in certain strains, Ag43 expression may be subject to both phase variation and antigenic variation.

Analysis of a panel of enteropathogenic *E. coli* (EPEC) strains by colony immunoblotting, Western immunoblotting and immunofluorescence microscopy (Meehan, 1994) revealed the presence of phase variable anti-α^{43} cross-reacting proteins in the M_r range 54,000 to 94,000, which were localized to the outer membrane. Furthermore analysis of EPEC strains revealed the presence of multiple copies of the *agn43* gene (Owen *et al.*, 1996; Roche *et al.*, 2001). In light of these discoveries antigenic variation of Ag43 seems to be a likely feature of *E. coli* infection.

1.3.2 Phase-variable properties of Ag43

Initial attempts at purification of the *agn43* product were continually complicated by fluctuations in the level of Ag43 in the outer membrane fractions being studied (Caffrey and Owen, 1989). Further investigation of this phenomenon revealed the presence of $Ag43^+$ and $Ag43^-$ variants within single colony isolates of strain ML308-225.

Screening of populations derived from single colonies revealed that the bacteria were switching reversibly between expressing (Phase ON) and non-expressing (Phase OFF) states. This phase variation may be observed by colony immunoblotting, immunofluorescence and immunoelectron microscopy or by direct observation of colony morphology. In liquid media, phase switching occurs at frequencies of ~2.2 X 10^{-3} per cell per generation (ON to OFF) and ~1 X 10^{-3} per cell per generation (OFF to ON; Owen *et al.*, 1996).

The *agn43* gene is present as a single copy, between minute 44.6 and 44.8, on the chromosome of *E. coli* K-12 (Blattner *et al.*, 1997). As mentioned earlier, the strain ML308-225, from which Ag43 was initially characterized, possesses duplicate copies of the gene, termed *agn43*A and *agn43*B (Roche *et al.*, 2001). Both *agn43* genes in strain ML308-225 are subject to phase variation, although they undergo phase switching at different rates. The *agn43A* gene switches ON to OFF at a frequency approximately two-fold faster than *agn43B*, and undergoes OFF to ON switching approximately eight times slower than *agn43B*. These differences may be due to differences in sequence between the regulatory regions of the two genes. However, it is not known whether any such sequence variation exists. Although the promoter of the *agn43B* gene has been sequenced attempts at sequencing the promoter region of *agn43A* have been unsuccessful (Roche *et al.*, 2001)

Analysis of EPEC strains has also revealed the presence of a 94- kDa protein, which cross-reacts strongly with anti-α^{43} antibodies. Cleavage of the 94-kDa protein *in situ* with trypsin generates two membrane-associated products. One is a 55-kDa protein, which, like α^{43}, can be selectively released from the outer membrane by heating while

the other product, like β^{43}, is heat-modifiable. Despite reacting strongly with anti-α^{43} antibodies the 94-kDa protein reacts poorly with anti-β^{43} antibodies and its subunits are clearly distinct from those of Ag43 in terms of M_r (Owen *et al.*, 1996). Interestingly, like Ag43 the 94-kDa protein is regulated by Dam. Furthermore, the proteins undergo switching at similar rates (Henderson, 1996). Thus, it is tempting to speculate that the 94-kDa protein is actually an Ag43 precursor. However, there is at present no evidence to support a precursor-product relationship between these two proteins.

1.3.3 Biological functions of Ag43

Initial observations were suggestive of a role for Ag43 in adhesion of bacteria to host cells (Henderson, 1996; Owen *et al.*, 1996). It has since emerged that the probable function of Ag43 is in the formation of cell aggregates and biofilms through Ag43-Ag43 intercellular interactions (Henderson *et al.*, 1999; Danese *et al.*, 2000; Kjærgaard *et al.*, 2000). More recently, a possible role for Ag43 in the modulation of bacteriophage susceptibility within the gut has been suggested (Gabig *et al.*, 2002). The evidence that has led to these suggestions is discussed below.

1.3.3.1 Ag43 as an adhesin

Ag43 has many features that are suggestive of a role as an adhesin, viz., the presence of an RGD motif in both subunits, phase-variable surface expression and shared sequence identity with known adhesions such as AIDA-I (Caffrey and Owen, 1989; Henderson and Owen, 1999). These observations catalysed studies to investigate the effect of Ag43 expression on adhesion of bacteria to a variety of cell lines (Henderson, 1996; Owen *et al.*, 1996). These studies revealed that populations Phase ON for Ag43 expression were 3-8 fold (depending on cell line tested) more adherent than Phase OFF

populations. Furthermore, adherence of Phase ON populations was inhibited, in a dose dependent manner, by purified α^{43}. These results indicated a possible role for Ag43 in adhesion. However, no data to support this hypothesis has ever been published by others.

1.3.3.2 Ag43 in autoaggregation and biofilm formation

In 1980 Diderichsen reported the existence, in *E. coli* K-12, of a property that was subject to high frequency phase variation. Many strains were observed to contain two variants, which could be distinguished by their ability or inability to autoaggregate in liquid media. The gene responsible for these phenotypes mapped to between 43 and 44 minutes on the *E. coli* chromosome and was named *flu*. Thus, the genotypes of the aggregating and non-aggregating variants were termed *flu*$^-$ and *flu*$^+$, respectively. Diderichsen also noted that colony morphology was dependent on the *flu* state. *flu*- variants of *E. coli* K-12 formed colonies that were flat and irregular, with a frizzy, rough surface. In contrast, colonies derived from *flu*$^+$ variants were small and raised with a glossy surface.

More recently, Henderson *et al.*, (1997a) identified Ag43 as the product responsible for the *flu*-related properties described by Diderichsen (1980). Cells that were Phase ON and Phase OFF for Ag43 expression gave rise to Flu$^-$ and Flu$^+$ colonies, respectively. In addition, Phase ON cells tended to autoaggregate and settle out of liquid media whereas Phase OFF cells did not. The gene encoding Ag43 mapped to the same position as *flu* establishing beyond doubt that Ag43 is the *flu* product (Henderson *et al.*, 1997). This observation was later confirmed by Hasman *et al.* (1999) through the construction of a *flu:tetR* mutant.

Diderichsen (1980) also identified part of the regulatory mechanism controlling *flu* expression. An *E. coli* mutant strain, BD1302, which had a deletion in the 89-min region, was locked in the frizzy (Flu^-) phenotype. Subsequently Warne *et al.* (1990) identified the mutated gene as encoding a Lys-R-type regulator, which they termed Mor. Later this product was found to be identical to OxyR (Henderson and Owen, 1997; 1999).

Ag43 acts as self-recognising adhesin and the autoaggregation it promotes results from intercellular Ag43-Ag43 interactions (Hasman *et al.*, 2000). The expression of Type-1 fimbriae, which is also subject to phase variation, has been shown to abolish Ag43-mediated autoaggregation. It is thought that the protrusion of these fimbriae, which are long rigid organelles, prevents the cells from making the close contact that is required for Ag43-Ag43 interactions to occur (Hasman *et al.*, 1999). Ag43 can be expressed in a functional form in *Pseudomonas fluorescens* leading to autoaggregation and changes in colony morphology. Furthermore, interspecies aggregation occurs when *E. coli* are mixed with Ag43-expressing pseudomonads (Kjærgaard *et al.*, 2000a,b).

The phase variably expressed Type-1 fimbriae are also capable of mediating cell aggregation (Pratt and Kolter, 1998). Unlike Ag43-mediated aggregates, which settle out of solution, Type 1 fimbriae-mediated aggregates form a pellicle at the liquid air interface (Hasman *et al.*, 2000). The expression of Ag43 is coordinated with that of Type-1 fimbriae, as further discussed in Section 1.5.5. Thus, by coordinating the phase variable expression of both Type-1 fimbriae and Ag43, bacteria can ensure that

members of the population occupy various niches within a static liquid environment such that the overall survival of the population is enhanced (Hasman *et al.*, 2000).

In their natural environments the majority of microbes associate with a surface by forming a complex structure, which is known as a biofilm. These structures may be defined as matrix-enclosed bacterial populations adherent to each other and/or to non-bacterial surfaces or interfaces. They consist of exopolysaccharide-encased microbial pillars separated by channels, which allow nutrients to circulate and waste to be disposed of. In nature biofilms usually consist of multiple species each of which has a distinct niche within the community (Donlan and Costerton, 2002).

There are several steps involved in the formation of a biofilm. A single bacterium associates transiently with other microbes attached to a surface and/or with the surface itself. The initial association facilitates the occurrence of more stable interactions, which are required for the formation of the three-dimensional biofilm structure. Individual cells may subsequently detach from the structure and resume a planktonic state and go on to colonize distal sites (Kjærgaard *et al.*, 2000a; Donlan, 2002).

Studies of cells within single-species biofilms have revealed that their profile of gene transcription differs to that of free-living, planktonic cells (Prigent-Combaret *et al.*, 1999). Ag43 expression has been shown to be necessary for the formation of biofilms by *E. coli*, in minimal medium. This is probably due, simply, to an Ag43-mediated increase in intercellular binding, as Ag43 does not seem to promote cell attachment to abiotic surfaces (Danese *et al.*, 2000; Kjærgaard *et al.*, 2000a). Furthermore, expression of Ag43 in *P. fluorescens*, which normally contains no Ag43 homologue,

enhances the ability of this species to form biofilms and allows it to form multispecies biofilms with *E. coli* (Kjærgaard *et al.*, 2000b).

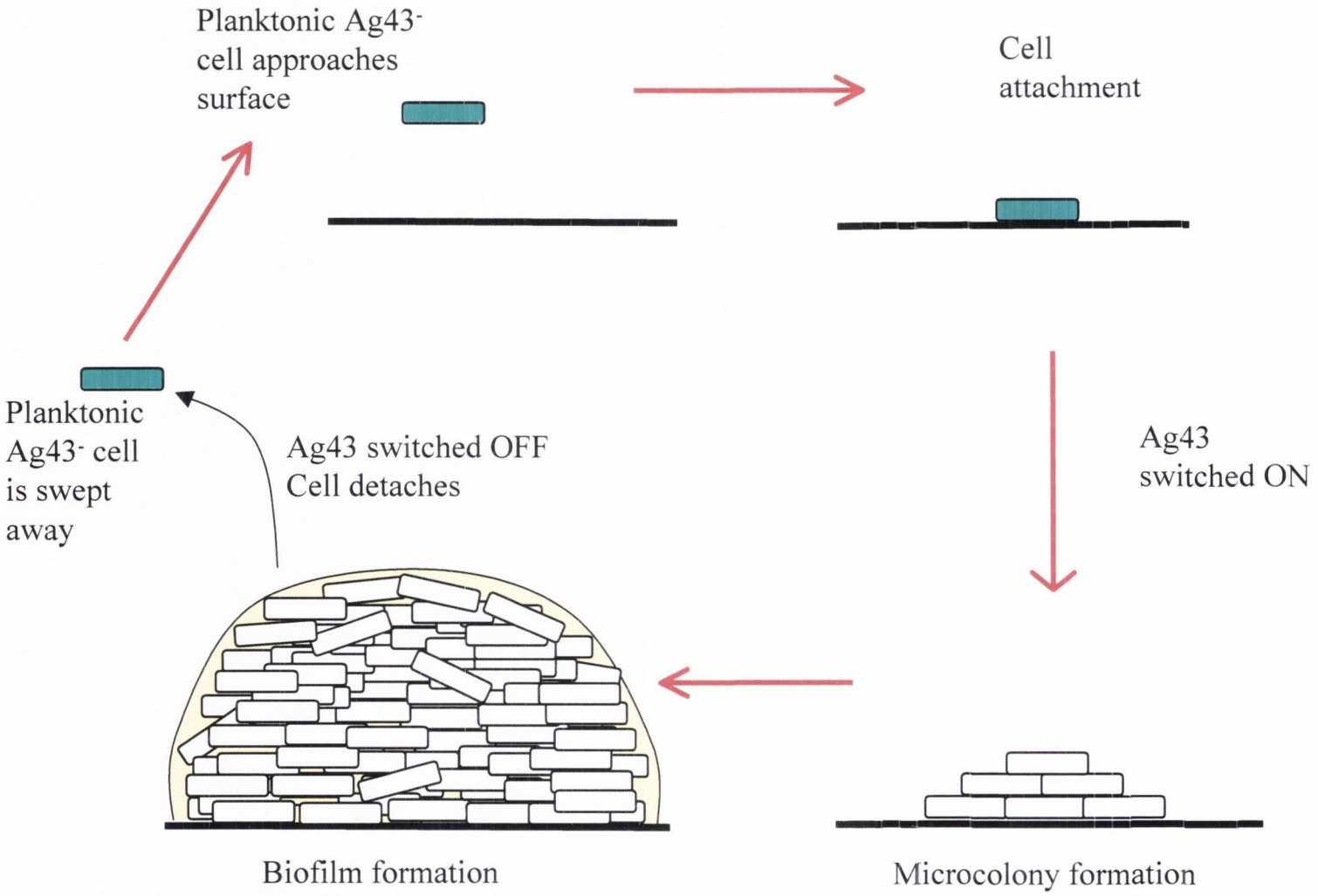
Members of bacterial biofilms have a number of advantages compared with free-swimming bacteria. They are not readily swept away by water currents, they have increased resistance to antimicrobial reagents, and the members of host located biofilms are often highly resistant to host defence mechanisms. Biofilm formation poses serious problems by facilitating the colonization of catheters and other medical devices (Costerton *et al.*, 1999).

One potential problem associated with life in a biofilm is that bacteria encased in the exopolysaccharide matrix may have difficulty in freeing themselves if conditions become less favourable. Phase switching of autoaggregative factors, such as Ag43, may thus facilitate the process of detachment from the biofilm and recolonization of distal sites (Danese *et al.*, 2000; Kjærgaard *et al.*, 2000a). The role of Ag43 phase variation in biofilm formation is illustrated in Figure 1.2.

1.3.3.3 Ag43 and bacteriophage infection

It has recently been reported that infection of *E. coli* by bacteriophages is inhibited in the presence of certain bile salts and carbohydrates when cells are in the Phase OFF state of Ag43 expression (Gabig *et al.*, 2002). In mammals *E. coli* is found in the lower bowel and in the small intestine. The small intestine contains high concentrations of bile salts, mainly cholate and deoxycholate, as well as carbohydrates, which are derived from the diet (Gabig *et al.*, 2002). Many phages require the presence of divalent cations to neutralise the negative charge of the bacterial cell wall so that they can adsorb

Figure 1.2 The possible effect of Ag43 phase variation on association with a biofilm. A free-swimming planktonic Ag43$^-$ cell (green) approaches a surface and subsequently attaches. Subsequent switching to the Ag43 Phase ON state (white) might facilitate microcolony formation mediated by Ag43-Ag43 intercellular interactions. The bacteria become encased in an exopolysaccharide matrix (yellow) leading to formation of the three-dimensional biofilm structure. Unfavourable environmental changes may result in Ag43 expression being switched OFF allowing Ag43$^-$ bacteria (green) to dissociate from the biofilm and resume a planktonic stage facilitating the colonization of other sites.


Planktonic Ag43- cell approaches surface
Cell attachment
Ag43 switched ON
Microcolony formation
Biofilm formation
Ag43 switched OFF Cell detaches
Planktonic Ag43- cell is swept away

efficiently (Hancock, 1997). It is thought that bile salts and carbohydrates inhibit phage adsorption by sequestering cations, since inhibition can be suppressed by addition of excess Mg^{++} or Ca^{++} ions to growth media. This inhibition can also be circumvented by the presence of Ag43 on the cell surface (Gabig *et al.*, 2002). It is thought that Ag43 serves as an initial site of phage-host interaction under limiting conditions of divalent cations since membrane fractions from $Ag43^{+}$, but not $Ag43^{-}$, cells were found to sequester λ phage (Gabig *et al.*, 2002).

As mentioned earlier in this chapter, phase variation is usually thought of as a mechanism that facilitates the evasion of host defences. The phase variation of Ag43 expression might also benefit the bacterium by affecting the susceptibility to phage infection within the gut (Gabig *et al.*, 2002). Switching to the Phase OFF state might prevent phage adsorption in the gut, thus allowing the survival of a fraction of any population endangered by phage infection (Gabig *et al.*, 2002). It has been suggested that inhibition of phage adsorption by bile salts and carbohydrates has possible implications for bacteriophage therapy in the treatment of uropathogenic *E. coli*. The inhibition of phage infection by bile salts in the gut may allow *E. coli* to be selectively eliminated from infected urinary tracts without compromising the intestinal gut flora (Gabig *et al.*, 2002).

1.4 Mechanisms of phase and antigenic variation

There are a number of distinct mechanisms controlling phase and antigenic variation in bacteria (Table 1.1). Some well-characterised examples are discussed in the following sections.

1.4.1 Vmp outer membrane lipoproteins of *Bor. hermsii*

Bor. hermsii is the causative agent of relapsing fever, which involves recurring bouts of illness, each corresponding to a new bacterial serotype (Coffey and Eveland, 1967). Variable membrane proteins (Vmps) are lipoproteins encoded by a family of 26 structurally-variable *vmp* genes, only one of which is expressed at any given time (Burman *et al.*, 1990; Restrepo *et al.*, 1992). The *vmp* gene being expressed dictates the bacterial serotype and is located near the right telomere of a 28-kb linear plasmid (Kitten and Barbour, 1990; Barbour *et al.*, 1991). A new serotype appears when this allele switches places with one located at a silent locus on a different plasmid. This allelic replacement results from a recombination event involving two identical sequences found in both the silent and expressed loci (Kitten and Barbour, 1990). A second mechanism of Vmp switching involves a deletion event involving the expressed Vmp gene and an adjacent inactive sequence, which results in a new gene fusion (Restrepo *et al.*, 1994). Both of these mechanisms are responsible for serotype swiches *in vivo* during *Bor. hermsii* infection (Restrepo and Barbour, 1994). Antigenic variation of Vmp proteins would be expected to increase bacterial longevity in the host, thereby increasing the likelihood of propogation and of spread to new hosts.

1.4.2 Shiga-like toxin II

Shiga-like toxins (SLT) are produced by various Enterohaemmorhagic *E. coli* (EHEC) strains, which cause haemorrhagic colitis and haemolytic uraemic syndrome. SLT toxins are toxic to both HeLa and Vero cells and are categorized into two antigenically distinct groups, SLT-I and SLT-II (reviewed by Karmali, 1989). SLT-I toxins appear to be functionally homogeneous. In contrast the SLT-II family contains two functionally distinct proteins, viz., SLT-II and SLT-IIv, which is toxic to Vero cells only (Samuel *et*

al., 1990). Antigenic variation of SLT-II proteins may allow the expression of proteins with different receptor specificity or different levels of cytotoxicity (Samuel *et al.*, 1990). Furthermore, the presence of duplicate copies of a gene may leave one copy free to generate random mutations, which could lead to the evolution of proteins with novel functions (Force *et al.*, 1999).

1.4.3 Type IV pilin of *N. gonorrhoeae*

The structural subunit of the gonococcal type IV pilus is expressed from the *pilE* locus. Untranscribed alleles of *pilE* (called *pilS*, for 'silent') are situated elsewhere on the chromosome. ON-to-OFF phenotypic switching may result from deletions in the 5' region of *pilE* (Segal *et al.*, 1985; Hill *et al.*, 1990). Expression may be restored via recombination events in which homologous sequences from any of the *pilS* loci are transferred to the *pilE* locus by nonreciprocal gene conversion (Haglbom *et al.*, 1985; Swanson *et al.*, 1987). Unequal recombination between *pilE* and *pilS* may occur resulting in the expression of 'L-pilin' variants due to the presence of multiple copies of *pilS* in the expression locus. Populations expressing L-pilin are non-piliated since these subunits are unable to assemble into functional pilin structures. Such bacteria may switch back to the piliated state following deletion of the extra pilin gene copies (Seifert, 1996). The Phase OFF state may also result from incorrect sequence-dependent processing of propillin, which results in the expression of S-pilin, a soluble truncated form of pilin that is secreted from the bacterium (Haas *et al.*, 1987).

1.4.4 Opa proteins of *N. meningitidis* and *N. gonorrhoeae*

The opacity proteins (Opa) of *N. gonorrhoeae* and *N. meningitidis* are subject to both phase variation and antigenic variation (Meyer *et al.*, 1990). Phase variable control of

Opa expression occurs at the level of translation via alterations in the 5' end of the reading frame. The DNA sequence encoding the signal peptide region of each *opa* gene contains a pentameric repeat, 5'-CTTCT-3', which is termed the coding repeat (CR). Slipped-strand mispairing leads to variation of the number of these repeats, which may shift the translational reading frame of the *opa* gene. A series of 6, 9 or 12 CRs results in the Phase ON state since the initiation codon is in-frame with the coding region, leading to expression of functional protein. Any other configuration of CRs leads to an out-of-frame initiation codon resulting in the expression of truncated proteins, which are subject to rapid degradation (Stern *et al.*, 1986).

Antigenic variation of Opa proteins is due to the presence of multiple structurally-variant *opa* genes throughout the chromosome. *N. meningitidis* and *N. gonorrhoeae* contain 4 and 11 copies of *opa*, respectively. The genes have large regions of near-perfect nucleotide sequence identity interspersed with three regions that vary in size and sequence between genes. One region, which encodes the amino terminus of the mature protein, exhibits a moderate degree of variability and is designated SV (semivariable). Two other regions, which are more highly variable, are designated HV1 and HV2 (hypervariable). Computer analysis predicts that Opa proteins form β-barrels within the outer membrane. The variable regions (SV, HV1, HV2) are predicted to lie within surface-exposed loops, leading to the suggestion that the diversification process is promoted by the immune system (Malorny *et al.*, 1998). This diversification is facilitated by three related genetic mechanisms: gene duplication, gene replacement, and partial non-reciprocal recombination (Hobbs *et al.*, 1994).

1.4.5 Opc of *N. meningitidis*

Like the Opa proteins, Opc of *N. meningitidis* is surface exposed, is thought to form a β barrel and is involved in the adhesion/ invasion process. In addition, slipped-strand mispairing is involved in the phase variation of both the Opc and Opa proteins. However, in contrast to Opa phase variation, which is controlled at the level of translation (see Section 1.4.4), Opc phase variation is controlled at the transcriptional level. In addition to reversible ON-to-OFF switching Opc expression exhibits volume control such that three switch states exist, viz., Opc^-, Opc^+ and Opc^{++}. Phase variation of Opc expression results from a strand slippage event within a stretch of cytosines that lies between the –10 and –35 regions of the promoter. A poly(C) tract of either 12 or 13 bases results in the Phase ON Opc^{++} state. Transcription is approximately 10-fold lower when the number of cytosines is 11 or 14 resulting in the Opc^+ phenotype. When the number of cytosines is either less than 10 or greater than 15 transcription is abolished resulting in the Opc^- phenotype (Sarkari *et al.*, 1994).

1.4.6 Phase variation by DNA inversion

Another mechanism that leads to phase variation involves site-specific DNA inversions that direct transcription either towards or away from a gene. One such system is found in *Salmonella typhimurium*, where flagellar phase variation is regulated by the Hin invertase (Heichman and Johnson, 1990). Such a system also regulates the phase variation of the *fim* operon in *E. coli*, which encodes Type-1 fimbriae. Since there is regulatory crosstalk between the *agn43* and *fim* systems, detailed discussion of the regulation of *fim* phase variation will be deferred to Section 1.6.3.

Table 1.1 Examples of bacterial phase variation (Adapted from Henderson *et al.*, 1999).

Moiety affected	locus	Property affected	Species	Mechanism
Capsular polysaccharide	*cap*	Capsular polysaccharide production	*H. influenzae*	Recombinational deletion
Type IV pili	*pilE*	Fimbrial expression	*N. meningitidis*/ *N. gonorrhoeae*	Recombinational deletion
P-fimbriae	*pap*	Fimbrial expression	*E. coli*	Dam and Lrp transcriptional control
S-fimbriae	*sfa*	Fimbrial expression	*E. coli*	Dam and Lrp transcriptional control
CS31A	*clp*	Fimbrial expression	*E. coli*	Dam and Lrp transcriptional control
Type-1 fimbriae	*fimA*	Fimbrial expression	*E. coli*	DNA inversion
Flagella	*hin*	Flagellar expression	*Salmonella* spp.	DNA inversion
Opa	*opa*	Adhesion /invasion	*N. meningitidis*/ *N. gonorrhoeae*	Strand slippage, translational
Type IV pilus	*pilC*	Fimbrial expression	*N. meningitidis*/ *N. gonorrhoeae*	Strand slippage, translational
Opc	*opc*	Adhesion/ invasion	*N. meningitidis*	Strand slippage, transcription
LKP-fimbriae	*hif*	Fimbrial expression	*H. influenzae*	Strand slippage, transcription

1.5 Regulation of Ag43

The phase-variable expression of Ag43 can be detected by immunofluorescence microscopy, Western immunoblotting, colony immunoblotting and by direct observation of colony morphology (See Figure 1 of Henderson *et al*, 1997; Meehan, 1995). This phase-variation has been shown to be independent of a wide range of well-documented regulatory proteins. A key observation was the lack of Ag43 expression in strains carrying mutations in the *dam* gene. Such strains are said to be Locked OFF (Henderson *et al.*, 1997). Equally significant was the observation that strains containing mutations in the *oxyR* gene were Locked ON for Ag43 expression, as were *dam oxyR* double mutants (Henderson and Owen, 1999).

The *dam* product deoxyadenosine methyl transferase (Dam) methylates the N^6 position of adenine at most of the about 19 000 5'-GATC-3' sites in *E. coli*. The OxyR protein belongs to the Lys-R like family of transcriptional regulators. It is best known for its role in activating the expression of antioxidant defence genes in response to oxidative stress. In view of their critical roles in regulating *agn43* gene expression Dam and OxyR are discussed in detail in the following sections.

1.5.1 Dam

Dam is a 32-kDa monomeric protein consisting of 278 amino acid residues (Herman and Modrich, 1982). In rapidly growing *E. coli* K-12 there are 130 molecules of Dam per cell (Boye *et al.*, 1992). The Dam protein of *E. coli* is related at the amino acid level to the corresponding Dam enzymes of phages P1, T1, T2 and T4 and to the *Dpn*II, *Eco*RV, and *Porphyromonas* methyltransferases (Lauster, 1989; Banas *et al.*, 1991). Dam homologues are widespread amongst enteric bacteria including *Salmonella* spp.,

Serratia marcescens, *Yersinia* spp., and *Vibrio cholerae*. Although *E. coli dam* mutants are viable, *dam* is an essential gene in *V. cholerae* and *Yersinia pseudotuberculosis* (Bale *et al.*, 1979; Julio *et al.*, 2001).

1.5.1.1 Methylation of 5'-GATC-3' sites by Dam

The *dam* product methylates the N^6 position of adenine at most of the 5'-GATC-3' sites in *E. coli* (Hattman *et al.*, 1978; Figure1.3). S-adenosylmethionine (AdoMet) supplies the methyl group for the Dam methylation reaction and is bound to a high affinity site on the Dam protein. When the initial Dam•AdoMet complex binds to a target 5'-GATC-3' site a fit-induced rearrangement of the protein around the DNA occurs. This results in the formation of an active complex capable of methylating the DNA to which it is bound (Bergerat and Guschlbauer, 1990). Each binding event results in the transfer of only one methyl group (Herman and Modrich, 1982; Urig *et al.*, 2002). The methylation of 5'-GATC-3' sites in exponentially growing cells occurs approximately 1 minute after passage of the replication fork (Barrass and Marinus, 1989; Campbell and Kleckner, 1990). This time lag before the methylation of newly synthesised DNA during replication results in the transient presence of hemi-methylated DNA thereby providing a time window during which certain cellular processes may be regulated in response to the cell cycle (Messer and Noyer-Weidner, 1988). There is little difference between the rate at which Dam modifies unmethylated and hemi-methylated DNA (Bergerat *et al.*, 1989).

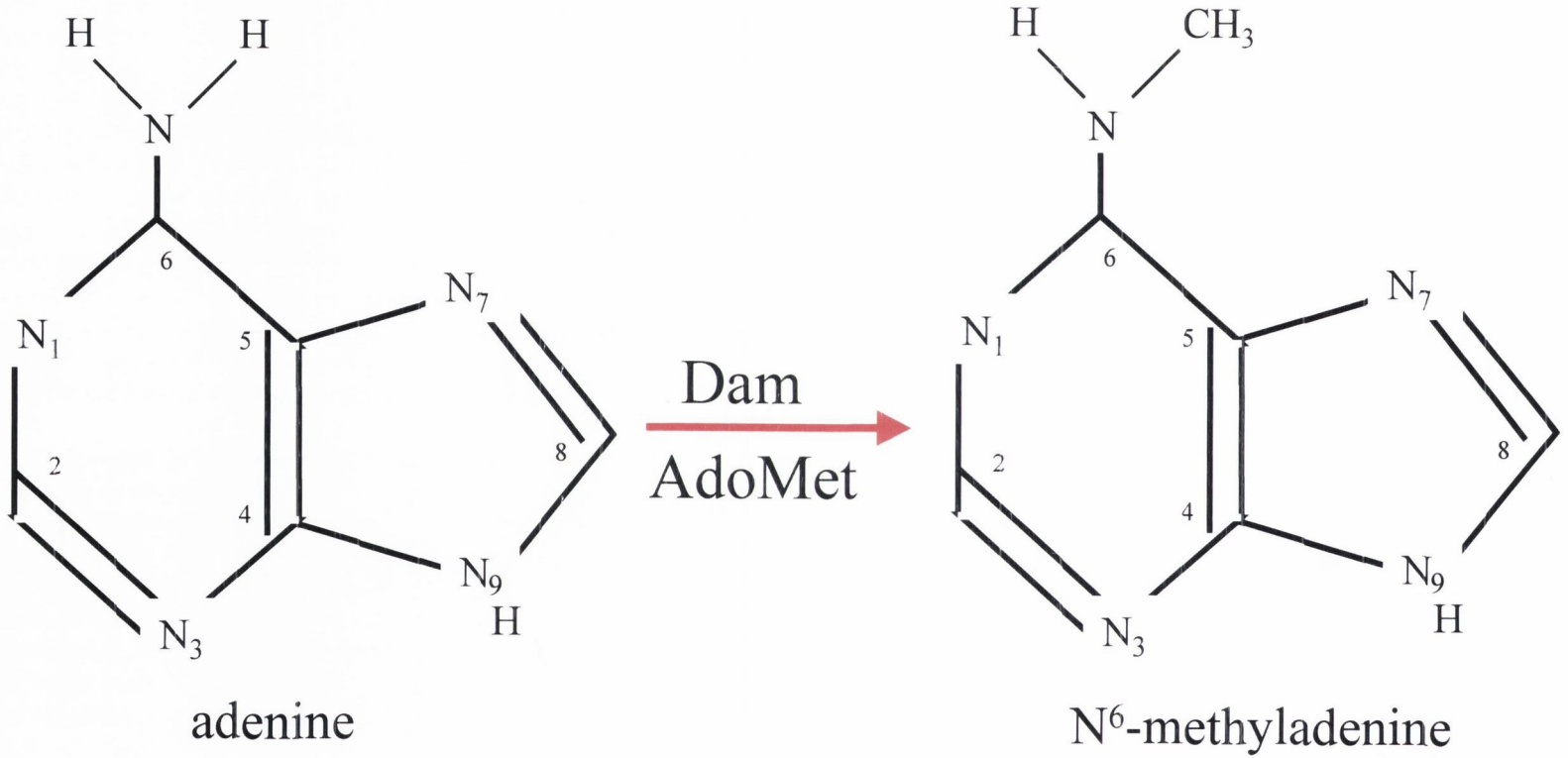
Recent studies have shown that Dam methylates DNA by a highly processive mechanism (Urig *et al.*, 2002). The enzyme is thought to slide along the DNA in a

process of linear diffusion carrying out consecutive reactions on the same DNA substrate. *In vitro* Dam scans 3000 bp of λ DNA on average before dissociating. Here the enzyme performs a random linear walk encountering many target sites more than once and methylating approximately 55 of a total 116 5'-GATC-3' sites. *In vivo,* however, it is quite possible that processive Dam methylation is directed and not random. It has been suggested that copies of Dam are located at a fixed point in the cell and that the DNA is pushed through them. This would allow Dam to scan the DNA in a directed fashion and one molecule could modify 3000 5'-GATC-3' sites processively (Urig *et al.*, 2002). The processive nature of DNA methylation should considerably accelerate the rate of methylation of newly synthesised DNA.

1.5.1.2 Distribution of 5′-GATC-3′ sequences in the bacterial chromosome

The *E. coli* K-12 genome comprises 4,639,221 bp and contains 19,123 5'-GATC-3' sites, which are subject to Dam-mediated methylation (Blattner *et al.*, 1997). The 5'-GATC-3' sequence is found more frequently in coding regions than in non-coding regions, consistent with the requirement for more frequent mismatch repair surveillance in translated sequences (Baras and Marinus, 1988). Genes encoding rRNA and tRNA have a particularly low 5'-GATC-3' content possibly to avoid secondary structures arising from palindromic sequences. There is never a distance greater than 2 kbp between successive 5'-GATC-3' sites allowing for Dam-directed mismatch repair to occur throughout the chromosome. Dam-mediated mismatch repair it is not efficient at distances greater than 2 kb (Baras and Marinus, 1988; Modrich, 1991).

Figure 1.3 Chemical structures of adenine and N^6-methyladenine. Dam catalyses the transfer of a methyl group from AdoMet to the N^6 position of adenine to give N^6 - methyladenine.



adenine

N^6-methyladenine

1.5.1.3 Biological functions of Dam

Strains deficient in Dam expression exhibit a wide variety of phenotypic traits and other properties (Palmer and Marinus, 1994; see Table 1.2). This vast array of difference scompared to wildtype reflects multiple functions for Dam and Dam-mediated methylation in the cell. Several major functions for Dam have been well documented and these are discussed below.

Dam has a central role in the methyl-directed mismatch repair system (Messon, 1988; Modrich, 1989). The transient presence of hemimethylated DNA behind the replication fork allows the mismatch repair system to discriminate between the template and nascent strands and correct any errors in the newly synthesized strand. In *dam* mutant strains repair occurs on either strand due to the inability of the mismatch repair system to discriminate between strands in the absence of methylation. As a result of this, *dam* strains have an increased frequency of spontaneous mutation (Marinus and Morris, 1975). Strains over expressing Dam also have a mutator phenotype due to premature methylation of newly synthesised DNA reducing the efficiency of mismatch correction (Marinus *et al.*, 1984).

Dam methylation helps to regulate the initiation of chromosome replication at *oriC*. The *E. coli oriC* region contains 10 times more 5'-GATC-3' sites than expected in a random sequence. Initiation of replication is inhibited when these sites are in the hemimethylated state (Russel and Zinder, 1987). Immediately after initiation both *oriC* and the *dnaA* gene remain hemimethylated for an unusually long time due to sequestration of 5'-GATC-3' sites by the SeqA protein (Taghbalout *et al.*, 2000). This keeps the origin in the hemimethylated inert state as well as reducing expression of

DnaA, which is required for replication initiation. Strains bearing mutations in the *dam* gene have uncoordinated DNA replication and overexpression of Dam interferes with the initiation process by reducing the time that *oriC* is hemimethylated (Boye *et al.*, 1988; Campbell and Kleckner, 1990).

Dam methylation regulates the expression of *agn43* and several other genes by altering the binding of regulatory proteins. Methylation is thought to block the binding of proteins directly by steric hindrance. It may also alter protein binding by another mechanism. Dam methylation has been shown to cause the DNA helix to unwind by 0.5 degrees per methyl group transferred (Cheng *et al.*, 1985). This is a relatively small conformational change. However, methylation of several closely grouped sites at *oriC* has been shown to result in a significant degree of DNA bending (Kimura *et al.*, 1989). It is conceivable that such changes in DNA secondary structure could perturb the alignment of sequences required for binding of transcriptional regulators.

The Dam-regulated genes identified so far include *trpR*, Tn*10* transposase, *dnaA*, *pap*, *sfa*, *daa*, *clp*, *agn43* and the *mom* gene of phage Mu (Hattman, 1982; Peterson *et al.*, 1985; Roberts *et al.*, 1985; Braun and Wright, 1986; van der Woude and Low, 1994; Henderson and Owen, 1999). The Dam-dependent regulation of the *mom* gene of phage Mu, and of the phase variable *pap*, *sfa*, *daa*, *foo* and *clp* operons, will be discussed later in this chapter.

1.5.2 OxyR

The *E. coli* OxyR protein is a 34-kDa, redox-sensitive global transcriptional regulator, which is best known for its role in the activation of antioxidant defense genes, under

Table 1.2 Altered physiological properties in a *dam* mutant (from Palmer and Marinus, 1994).

Property
A mutator phenotype
A hyper-recombination phenotype
Alleviation of *Eco*K restriction
Increased sensitivity to UV light
Control of phage P1 packaging into virions
Increased drug-induced mutagenesis
Asynchromous initiation of chromosome DNA replication
Failure of methylated plasmids to transform *dam* mutants with high efficiency
Increased precise excision and transposition of Tn*10* and other transposons
Increased spontaneous induction of lysogenic phages
Derepression of certain genes in the SOS regulon
Altered expression of certain chromosomal and nonchromosomal genes
Reduced Dam activity leading to a deduction in N^6 methyladenine in 5'-GATC-3' sequences
Suppression of certain Dam phenotypes by second site mutations in *mutS*, *mutH* and *mutL*

oxidizing conditions (Christman *et al.*, 1989). The genes induced by OxyR under conditions of oxidative stress include *katG* (encoding HP1 catalase), *ahpCF* (encoding an alkyl hydroperoxide reductase), *gorA* (encoding glutathione reductase) and an untranslated regulatory RNA, *oxyS* (Toledano *et al.*, 1994). Under normal conditions OxyR also acts as a transcriptional repressor, negatively regulating the expression of the *agn43*, *mom* and *oxyR* genes (Christman *et al.*, 1985; Bolker and Kahmann, 1989; Henderson and Owen, 1999).

1.5.2.1 Regulation by OxyR

OxyR is a member of the LysR-like family of transcriptional regulators. LysR-type proteins contain a helix-turn-helix motif, which is thought to be important for DNA binding (Schell, 1993). These proteins characteristically negatively regulate the expression of their own genes. They also protect unusually long DNA regions from DNase-1 digestion suggesting that they are multimeric. The oxidized and reduced forms of OxyR are predominantly tetrameric. Results of mutational analysis indicate that the C-terminal region of OxyR is involved in tetramerization (Kullik *et al.*, 1995). OxyR binds adjacent to the –35 box of the genes that it activates and is thought to contact the α subunit of RNA polymerase and cooperatively increase its binding (Tao *et al.*, 1995)

OxyR is regulated in response to growth phase and is maximally expressed in the exponential phase of growth (Gonzalez-Flecha and Demple, 1997). Although OxyR is an important regulator of antioxidant defence genes, transcription of its own gene is not affected by oxidative stress. The oxidation-dependent activation of certain genes by OxyR is due to posttranslational modification of the protein, not to changes in protein

levels. OxyR becomes activated by the formation of an intramolecular disuphide bond and deactivated by enzymatic reduction of the disulphide bond between two redox-active cysteines, C199 and C208 (Zheng *et al.*, 1998). The thiol-disulphide redox state of the cell cytosol is proposed to regulate the activation of OxyR (Åslund *et al.*, 1999).

In addition to the oxidised and reduced forms of OxyR recent studies have proved the existence of other stable species in which the redox-sensitive cysteine is alternatively modified (Kim *et al.*, 2002). S-nitrosylation gives rise to a modified form (S-NO) of OxyR. Oxidation can give rise to a mixed disulphide with glutathione (S-SG) or a sulfenic acid derivatives (S-OH), as well as the disuphide form of OxyR (S-S). These forms are transcriptionally active, but differ in structure, cooperative properties and DNA binding affinity. They also differ in their ability to activate *in vitro* transcription of the *katG* gene raising the possibility that these forms may be involved in differential gene regulation *in vivo* (Kim *et al.*, 2002).

1.5.2.2 DNA binding properties of OxyR

OxyR binds as a tetrameric complex specifically, and with high affinity, to the promoter regions of the genes it regulates. OxyR-binding sites are unusual in that they are particularly long usually spanning a 45-bp region. The observation that just 3 out of 45 bp were conserved among six charaterised binding sites led to the suggestion that OxyR may have a "degenerate recognition code" (Tartaglia *et al.*, 1992). An OxyR-binding motif has however since been proposed based on the binding of OxyR to oligonucleotides of random sequence (Toledano *et al.*, 1994). The binding motif for oxidized OxyR consists of four ATAGnt sequences spaced at 10-bp intervals, which the protein contacts in four adjacent major grooves on one face of the DNA helix (See

Figure 1.4). The entire sequence of the proposed motif is ATAGntnnnanCTATnnnnnnnATAGntnnnanCTAT. Although this motif was based on binding to synthetic DNA of random sequence it is highly matched at the *katG* and *oxyRS* natural binding sites (Toledano *et al.*, 1994)

Interestingly OxyR gives rise to distinct DNaseI footprints under oxidizing and reducing conditions with the reduced form of the protein having an extended binding site (Storz *et al.*, 1990). In fact the two forms of the protein require different DNA contacts for binding (Toledano *et al.*, 1994) and it is this unique property that allows OxyR to regulate different promoters under oxidizing and reducing conditions. For example, OxyR only binds and activates *katG* under conditions of oxidative stress when expression of this gene is required. It is thought that oxidized OxyR binds to four adjacent major grooves while the reduced form of the protein binds to two pairs of adjacent major grooves separated by one helical turn (Toledano *et al.*, 1994).

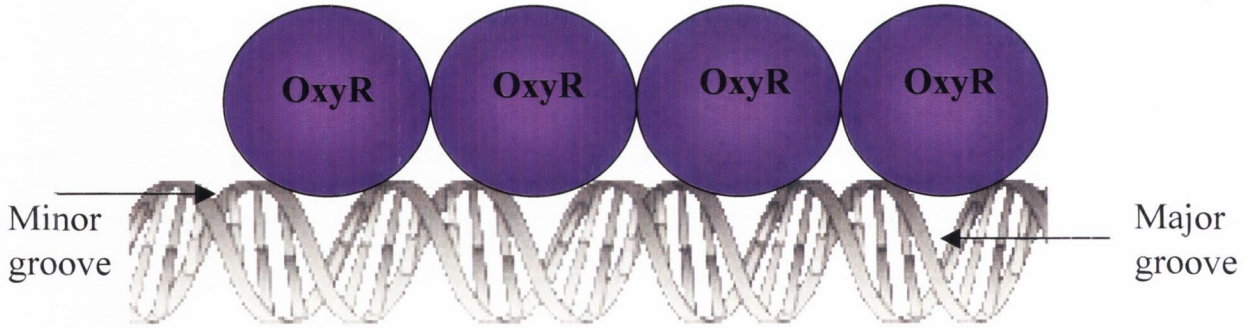
The protein is thought to be capable of shifting its contacts whilst remaining bound as its oxidation state changes. The repositioning of DNA contacts as OxyR is reduced is thought to occur asymmetrically with two of the OxyR subunits remaining bound whilst the other two are shifted by one helical turn. For this to happen, contact point 4 must be bound by different subunits under oxidizing and reducing conditions. It is thought that OxyR makes tight contact by binding four contacts of intermediate affinity. This could allow it to be zippered on and off the DNA as it shifts its DNA contacts in response to a change in oxidation state (Toledano *et al.*, 1994).

The asymmetrical repositioning of protein-DNA contacts allows OxyR to exert different regulatory effects at the divergent OxyR and OxyS promoters. Under oxidizing conditions OxyR can simultaneously act as a repressor of *oxyR* and an activator of *oxyS* at the same binding site. As the protein becomes reduced one of the subunits is repositioned over the *oxyS* –35 box where it prevents inappropriate activation of the *oxyS* gene (Figure 1.5; Toledano *et al.*, 1994). Hence, OxyR continues to repress its own transcription regardless of its oxidation state whilst regulating *oxyS* in response to oxidative stress.

1.5.3 OxyR and Dam in *agn43* regulation

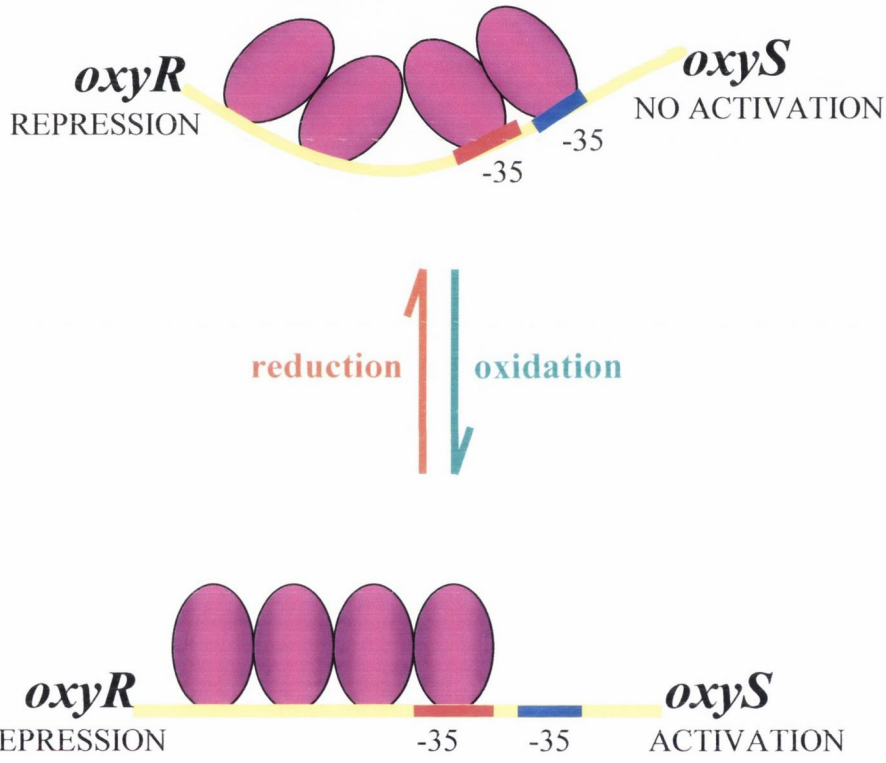
As well as observing that Ag43 phase variation was dependent on both the *dam* and *oxyR* gene products, Henderson and Owen (1999) showed that the regulatory region of the *agn43* gene, which encodes Ag43, contained a sequence that had high homology with the OxyR binding site of other genes regulated by OxyR, viz., the *mom* gene of phage Mu and the *oxyR* gene itself. Interestingly this region contains three 5'-GATC-3' sites, which have the potential to be methylated by the *dam* product (Henderson and Owen, 1999). As such, the recent findings that OxyR binds to the *agn43* regulatory region *in vitro* and that this binding is abrogated by methylation of the *agn43* 5'-GATC-3' sites (Haagmans and van der Woude, 2000), are not surprising. These 5'-GATC-3' sites have been observed to be fully methylated in an *oxyR* Locked ON background suggesting that they are protected from methylation by bound OxyR in Phase OFF cells. A correlation has been observed between the proportion of Phase ON cells in a population and the proportion of cells in the same population having methylated, or hemimethylated, 5'-GATC-3' sites. Conversely, the proportion of Phase OFF cells in a

Figure 1.4 The binding of an oxidised OxyR tetramer to 4 adjacent major grooves of the DNA helix. Subunits of the OxyR tetramer are represented by purple circles. The position of the major and minor grooves of the DNA are indicated. Below is the sequence of the OxyR binding motif proposed by Toledano *et al.* (1994). Upper case letters represent nucleotides with which OxyR is thought to make direct contact. Lower case letters (a and t) represent adenine and thymine residues, which are conserved but are thought not to contact OxyR directly. Positions that can be occupied by any nucleotide are represented by the letter n.



ATAGntnnnnanCTATnnnnnnnnATAGntnnnanCTAT

Figure 1.5 Model for regulation by oxidized and reduced OxyR at the divergent *oxyR* and *oxyS* promoters. The intergenic DNA region between *oxyR* and *oxyS* is represented by a yellow line. Subunits of the OxyR tetramer are represented by purple ovals. The -35 boxes of *oxyR*, and *oxyS*, are in red and blue, respectively. Upon oxidation two subunits of the OxyR tetrameric complex shift their DNA contacts by one helical turn exposing the *oxyS* promoter for activation (Figure adapted from Toledano *et al.*, 1994).


oxyR
REPRESSION
oxyS
NO ACTIVATION
-35
-35
reduction
oxidation
oxyR
REPRESSION
-35
-35
oxyS
ACTIVATION

population has been observed to correlate with the proportion of cells in which the *agn43* regulatory region is unmethylated (Haagmans and van der Woude, 2000).

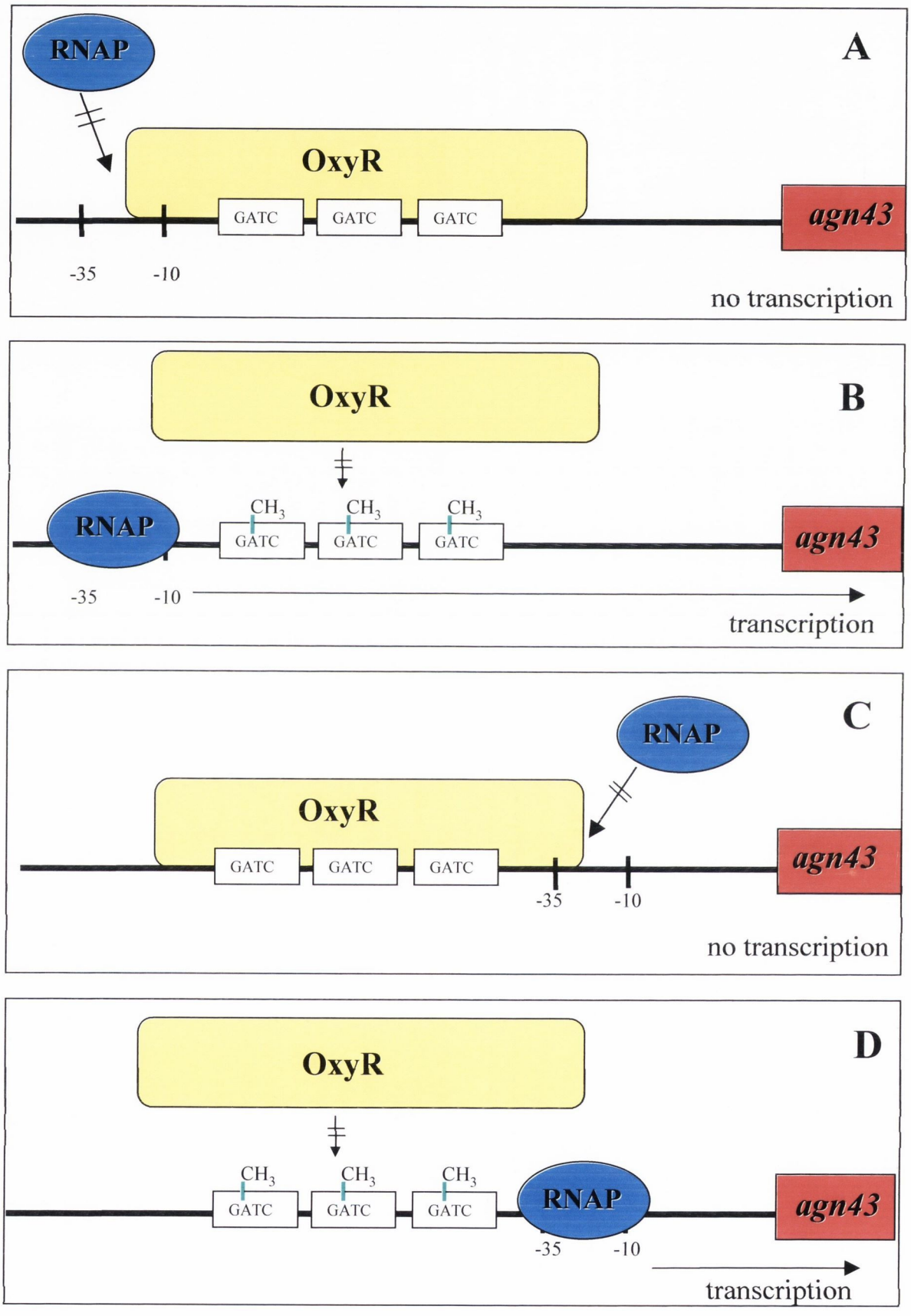
Together this evidence points strongly towards a model for regulation of phase variation in which binding of OxyR to unmethylated 5'-GATC-3' sites within the promoter region of *agn43* represses transcription, possibly by causing RNA polymerase to be excluded. Dam methylation of these sites would prevent binding of OxyR thereby allowing transcription of *agn43* to proceed. Competition between OxyR and Dam for unmethylated 5'-GATC-3' sites could result in switching between the Phase ON and Phase OFF states (Figure 1.6).

Such a system is known to regulate the *mom* gene of phage Mu, with which the promoter region of the gene encoding Ag43 shows extensive nucleotide sequence homology (Henderson and Owen, 1999). Regulation of *mom* and the possibility of similar regulation mechanisms for *mom* and *agn43* are discussed in the next section.

1.5.4 Dam-dependent regulation of the phage Mu *mom* gene

The *mom* gene of phage Mu encodes a DNA modification system, which makes both Mu DNA and cellular DNA resistant to the action of many type I and type II restriction endonucleases (Toussaint, 1976; Swinton *et al.*, 1983; Kahmann, 1984). Tight regulation ensures that Mom is expressed only in a short interval of phage development when the toxic effect of Mom-dependent modification can no longer interfere with phage development (Toussaint, 1977).

Figure 1.6 Model for the Phase OFF and Phase ON states of *agn43* transcription . In (A) and (B) the *agn43* promoter identified in this study is shown. In (C) and (D) the promoter proposed by Henderson and Owen (1999) is shown. (A) and (C), OxyR binds freely to the *agn43* regulatory region and blocks the binding of RNA polymerase (RNAP), resulting in the Phase OFF state. (B) and (D), Dam-methylation of 5'-GATC-3' sites within the *agn43* regulatory region blocks binding of OxyR leaving the *agn43* promoter accessible to RNA polymerase. Transcription proceeds resulting in the Phase ON state.


A
RNAP
OxyR
GATC
GATC
GATC
agn43
-35
-10
no transcription
B
OxyR
RNAP
CH3
CH3
CH3
GATC
GATC
GATC
agn43
-35
-10
transcription
C
RNAP
OxyR
GATC
GATC
GATC
agn43
-35
-10
no transcription
D
OxyR
CH3
CH3
CH3
GATC
GATC
GATC
RNAP
agn43
-35
-10
transcription

Transcription of the *mom* gene requires the action of a phage-encoded transactivator, MuC (Plasterk *et al.*, 1983; Hattman *et al.*, 1985). Dam methylation of three 5'-GATC-3' sites is also necessary for activation of *mom* transcription, thus the *mom* gene is only transcribed in strains that express Dam (Toussaint, 1977; Hatttman, 1982; Plasterk *et al.*, 1983).

The *mom* operon contains two divergent promoters *mom*P1 and *mom*P2, which overlap. The *mom*P1 promoter is a weak promoter and has a 19-bp suboptimal spacer between the –35 and –10 RNA polymerase binding sites. This promoter is also weakened by an intrinsic DNA distortion, the result of a tract of 6 successive T nucleotides in the spacer region (Basak *et al.*, 2001). RNA polymerase is unable to bind to *mom*P1 without the action of MuC, which binds as a dimer in two adjacent major grooves on one face of the double helix. MuC activates transcription by unwinding the DNA and reorienting the weak *mom*P1 promoter into a favourable conformation for successful binding by RNA polymerase (Basak and Nagaraja, 2001). In the absence of MuC, RNA polymerase binds instead to the *mom*P2 promoter, which directs transcription in a direction away from the *mom* gene. The purpose of *mom*P2 transcription is not known although the antisense RNA produced may prevent elongation of mRNA from the upstream gene into *mom* (Basak *et al.*, 2001). It has also been suggested that *mom*P2 acts as a sink for capturing RNA polymerase and holding it until required at the adjacent *mom*P1 promoter (Balke *et al.*, 1992).

Transcription of *mom* is strongly repressed by the OxyR protein in *dam*$^-$ but not in *dam*$^+$ cells (Bolker and Kahmann, 1989). OxyR represses transcription by binding to hemimethylated *mom* DNA, upstream of the *mom*P1, and interfering with the MuC-

mediated recruitment of RNA polymerase (Sun and Hattman, 1996, 1997). Dam methylation of the three 5-'GATC-3' sites blocks the binding of OxyR allowing RNA polymerase to bind the *mom*P1 promoter and drive transcription of the *mom* gene.

Mom expression is also regulated at a post-transcriptional level, by the phage-encoded Com protein. The *mom* mRNA translational start signals are contained within a stem loop structure, which inhibits translation. The binding of Com adjacent to this stem loop results in a change in RNA secondary structure that exposes the translation start signals (Hattman, 1999).

Interestingly, the *agn43* regulatory region shares extensive sequence homology with that of *mom* (Henderson and Owen, 1999). Comparative analysis shows a region displaying 65.5% identity at the nucleotide level. The *agn43* 5'-GATC-3' sites have a spatial distribution almost identical to those in the region of the *mom* promoter as shown in Figure 1.7 (Henderson and Owen, 1999). The similarities that exist between the regulatory regions of *agn43* and *mom*, and the fact that both genes are regulated by Dam and OxyR, suggested that they share similar mechanisms of regulation. The extensive nucleotide similarity between *agn43* and *mom* at sequences corresponding to the MuC binding domain (Bőlker *et al.*, 1989) also fuelled speculation for the involvement of an additional protein (apart from Dam and OxyR) in Ag43 regulation (Henderson and Owen, 1999).

1.5.5 Regulatory cross talk between Type-1 fimbriae and Ag43

agn43 expression has recently been shown to be regulated indirectly by the expression of Type-1 fimbriae (Schembri and Klemm, 2001). Type-1 fimbriae are thin adhesive

Figure 1.7 (A) Alignment of the regulatory regions from *agn43* and the *mom* gene of phage Mu. A dash (-) has been introduced in the *mom* regulatory sequence to improve the alignment. The 5'-GATC-3' sites in both the *agn43* and *mom* regulatory regions are highlighted in blue. The MuC binding domain of *mom* is underlined. Base identity is indicated by asterisks. The sequence shown corresponds the region in part (B) that is underlined. (B) Sequence of the *agn43* regulatory region. The translation start site is in bold and 5'-GATC-3' sites are highlighted in blue. The RNA polymerase contact sites identified in this study are in red (see Chapter 3). RNA polymerase contact sites proposed by Henderson and Owen (1999) are in green.

A

agn43 ATAAACGATCAATATTGATTTATTGGATCGTTTATATCGATCGATTAGCTAATAATAACCTTTGTCAGTAAC

mom AACGACGATCGA-ATCAATTAAATCGATCGGTAATACAGATCGATTATGCCCCAATAACCACACTCAACCCA

MuC

B

TTGTCCGTGCAATAGCTCAATAATAGAATAAAACGATCAATATCTATTTTATCGATCGTTTATATCGATCGATAAGCTAATAA

TAACCTTTGTCAGTAACATGCACAGATACGTACAGAAAGACATTCAGGGAACAACAGAACCACAATTCAGAAACTCCCACAGC

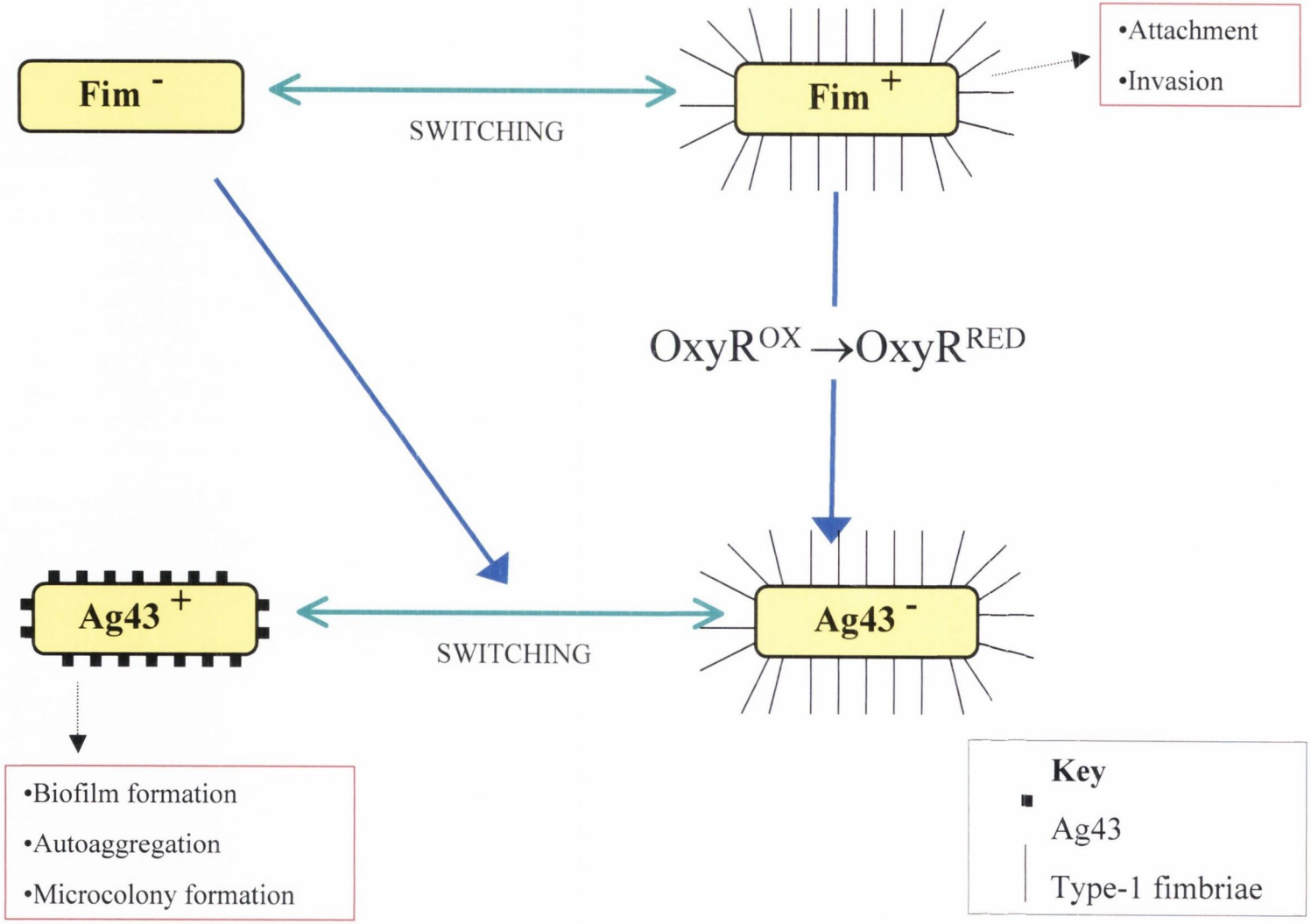
CGGACCTCCGGCACTGTAACCCTTTACCTGCCGGTATCCACGTTTGTGGGTACCGGCTTTTTTATTCACCCTCAATCTAAGGA

AAAGCTG**ATG**AAACG

organelles found on the surface of many strains of *E. coli* and other enterobacteriaceae (Klemm and Krogfelt, 1994). Their expression is subject to phase variation due to the inversion of a DNA region containing the promoter that drives transcription of the *fim* structural genes (Olsen and Klemm, 1994). Hence, bacteria alternate between fimbriated (expressing) and bald (non-expressing) states. A fimbriated bacterium has several hundred fimbriae radiating up to 1μm from its surface. An adhesin at the tip of the fimbriae recognises the mannose-containing glycoproteins that are present on the cell surface in many mammalian tissue types (Choudhury *et al.*, 1999). This allows the bacteria to attach to, and colonize, the surface of the urinary tract in humans (Connell *et al.*, 1996). It has been reported that the presence of Type-1 fimbriae on the cell surface abolishes Ag43-mediated autoaggregation, hence they are phenotypically dominant to Ag43 (Hasman *et al.*, 1999a). Their protrusion from the surface likely impedes the close contact that is required for Ag43 self-recognition and interaction (Schembri and Klemm, 2001).

Microarray-based studies revealed that *agn43* transcription is 20-fold greater in a *fim* deletion strain than in a strain Locked ON for the expression of Type-1 fimbriae (Schembri and Klemm, 2001). This is thought to be due to the effect that the mass formation of disulphide bonds, which occurs during fimbrial biosynthesis, has on the thiol-disulphide status of the cell. The alteration in cellular thiol-disulphide status, required to counteract the net oxidation effect of disulphide bond formation, would promote the reduction of OxyR (Schembri and Klemm, 2001). An increase in abundance of the thiol form of OxyR would result in down-regulation of *agn43* due to increased binding by this active form of the repressor (Figure 1.8).

Figure 1.8 Model of the OxyR-mediated one way signalling system that represses expression of Ag43 during the Phase ON state of *fim* expression. The fimbriated phenotype promotes the reduced form of OxyR, which represses transcription of the gene encoding Ag43. $OxyR^{OX}$, oxidized form of OxyR; $OxyR^{RED}$, reduced form of OxyR.


Fim -
SWITCHING
Fim +
•Attachment
•Invasion
OxyR^OX →OxyR^RED
Ag43 +
SWITCHING
Ag43 -
•Biofilm formation
•Autoaggregation
•Microcolony formation
Key
Ag43
Type-1 fimbriae

Since Type-1 fimbriae are phenotypically dominant, the production of thousands of copies of Ag43 would be a waste of resources and of no benefit to a fimbriated cell (Schembri and Klemm, 2001). Clearly, the existence of this one way signalling system between *fim* and *agn43* makes sense from an economical point of view. The costs in resources required for expression of these factors must be balanced with any potential benefits that they might provide. The coordinated and mutually exclusive nature of Type-1 fimbriae and Ag43 expression may aid in the process of infection as follows; the ON phase of fimbrial expression allows bacterial attachment while the OFF phase allows expression of Ag43, which enhances microcolony and biofilm formation. The OFF phase of Ag43 expression might allow bacteria to disperse leading to the colonization of more sites within the host (Schembri and Klemm, 2001). This system of regulatory cross talk is not unique in *E. coli*. It has recently been shown that expression of Type-1 fimbriae can be modulated by PapB and SfaB, which are regulators of the phase-variable S pilus and Pap systems, respectively. The regulation of Pap and S pilus, which involve Dam methylation, and of Type-1 fimbriae, is discussed further in the next section of this chapter (see Section 1.6).

1.6 Dam-dependent regulation of other operons

In addition to *agn43*, Dam regulates the expression of a large group of *E. coli* pili that are important in urinary tract infections (e.g. Pap, S, F1845 and F165[1] pili) and diarrheal infections (e.g. Afa, Cs31a and K88 pili). The phase-variable regulation of Pap, S, F1845 and F165(1) pili, are discussed in the following sections.

1.6.1 Phase variation of the *E. coli pap* operon

A well-characterised regulatory system involving Dam methylation is that of the *pap* operon, which encodes the pyelonephritis-associated pilus of *E. coli*. Pap pili play an important role in the pathogenesis of urinary tract infections by facilitating the adherence of uropathogenic *E. coli* to host epithelial cells (Hultgren *et al.*, 1993). Expression of Pap pili is subject to phase variation, which occurs via alterations in DNA-methylation patterns (Low *et al.*, 1987).

The *pap* operon contains nine genes whose products are involved in the assembly and expression of Pap pili, as well as two genes encoding transcriptional regulatory factors, PapB and PapI. The *papB* and *papI* genes are divergently transcribed from the *pBA* and *pI* promoters (Figure 1.9). The *papBA* transcript is polycistronic containing the *papB* gene and the *papA* gene, which encodes the major structural subunit of Pap fimbriae. Most *papBA* transcripts are terminated by a *rho*-dependent transcriptional terminator downstream of *papA*. However, some transcription continues into downstream genes, which code for assembly, minor fimbrin and adhesin proteins (Baga *et al.*, 1985).

Pap phase variation is controlled by the cooperative binding of the leucine responsive regulatory protein (Lrp) to binding sites in the regulatory region of the *pap* operon. Within this regulatory region are two 5′-GATC-3′ sites, termed 5′-GATC-3′$_{dist}$ and 5′-GATC-3′$_{prox}$, which are targets for Dam-mediated methylation (Figure 1.9). The *pap* 5′-GATC-3′ sites are contained within two different sets of binding sites for Lrp which are designated sites 4 and 5 and sites 1, 2 and 3 (Nou *et al.*, 1993). The Dam-mediated methylation of these 5'-GATC-3' sites modulates binding of Lrp to the *pap* promoter region (Nou *et al.*, 1993). Lrp has a dual role in *pap* regulation and acts as either an activator or a repressor of transcription, depending on which of the five sites it binds in

the *pap* promoter region. The methylation pattern of the 5'-GATC-3' sequences affects whether Lrp (a) binds to sites overlapping the *papBA* promoter (sites 1, 2 and 3) and blocks transcription by steric hindrance of RNA polymerase, or (b) binds to distant sites (sites 4 and 5) and allows transcription to proceed. Methylation of 5′-GATC-3′$_{dist}$ inhibits transcription by blocking the binding of Lrp to sites 4 and 5 (Hale *et al.*, 1998). In contrast methylation of 5′-GATC-3′$_{prox}$ activates transcription by inhibiting the binding of Lrp to sites 1, 2 and 3 thus leaving the *papBA* promoter accessible to RNA polymerase (Braaten *et al.*, 1994; van der Woude *et al.*, 1995). Conversely, the binding of Lrp to regions containing the *pap* 5′-GATC-3′ sites blocks Dam methylation by steric hindrance (Braaten *et al.*, 1991; Braaten *et al.*, 1994). PapI promotes the translocation of Lrp from sites 1, 2 and 3, to sites 4 and 5 by binding specifically to Lrp-*pap* complexes and altering the affinity of Lrp for the *pap* DNA (Kaltenbach *et al.*, 1995). The PapI-dependent translocation of Lrp results in differential protection of the two 5′-GATC-3′ sites from Dam, resulting in different methylation patterns for the Phase ON and Phase OFF states of Pap expression (Figure 1.10).

Catabolite activator protein (CAP) is essential for the activation of *papBA* transcription and binds at a site 215-bp upstream of the pBA promoter. CAP-dependent activation of *papBA* requires the binding of Lrp to sites 4 and 5 and may involve an interaction with the C-terminal domain of the α subunit of RNA polymerase (Weyand *et al.*, 2001). The *pap*-encoded PapB protein binds adjacent to the CAP-binding site and helps to activate *papI* transcription (Baga *et al.*, 1985). PapB also binds near the –10 of the pBA promoter where it has an autoregulatory repressive effect on transcription (Forsman *et al.*, 1989). PapB is also involved in the regulation of *papI*, as are Lrp, CAP and PapI itself (Forsman *et al.*, 1989, 1992). In addition, H-NS positively regulates the *pap*

Figure 1.9 Arrangement of the *pap* operon regulatory region. The relative locations of the *pBA* and *pI* promoters, the 5'-GATC-3'$_{dist}$ and 5'-GATC-3'$_{prox}$ sites, and the binding sites for PapB, CAP and Lrp, are shown.

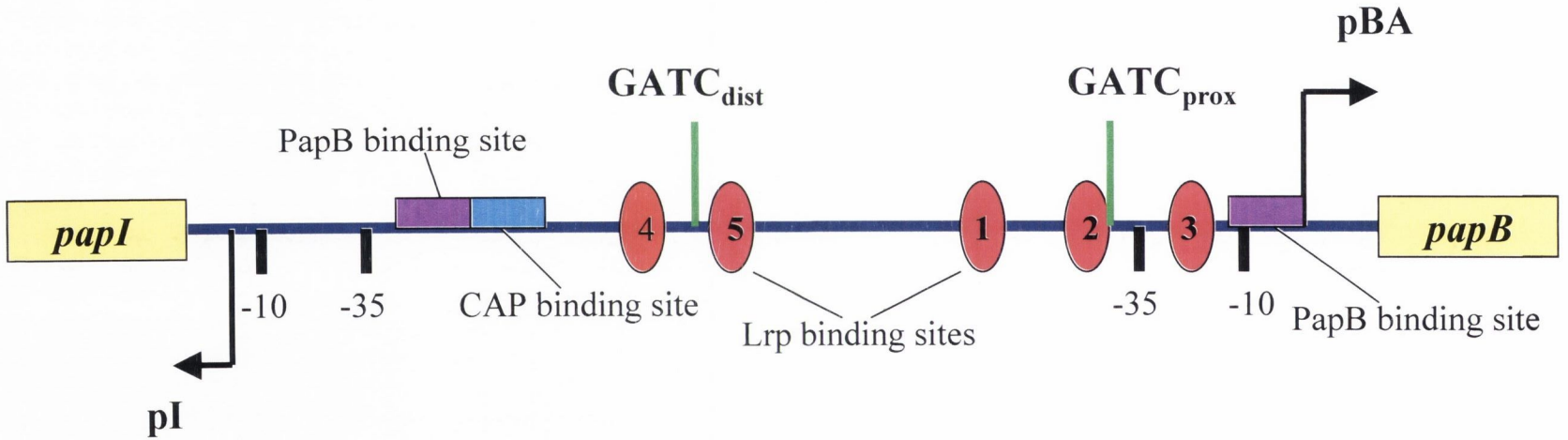
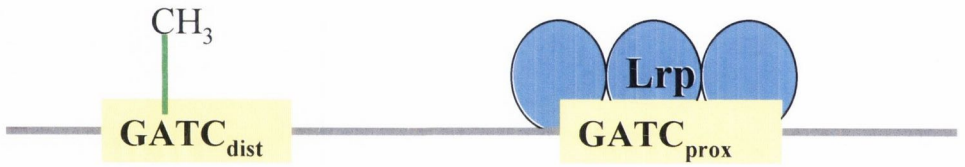
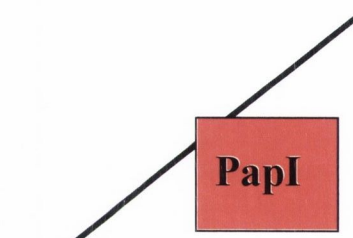

pBA
GATC dist
GATC prox
PapB binding site
papI
4
5
1
2
3
papB
-10
-35
CAP binding site
Lrp binding sites
-35
-10
PapB binding site
pI

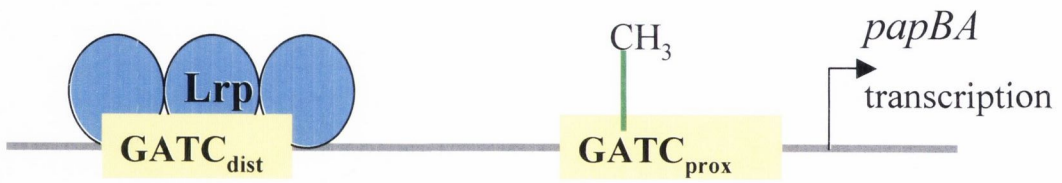
Figure 1.10 Regulation of the *pap* phase variation by Dam methylation. The Phase OFF methylation pattern results from binding of Lrp to sites (1, 2 and 3; See Figure 1.9) proximal to the *papBA* promoter, which blocks transcription and inhibits methylation of $GATC_{prox}$. PapI binds to Lrp increasing its affinity for distal binding sites (4 and 5; See Fig 1.9) thus facilitating methylation of 5'-GATC-3'$_{prox}$ forming the Phase ON methylation pattern.

PHASE OFF





PHASE ON



operon in response to the environmental cues of temperature, osmolarity and glucose as a carbon source, by an unknown mechanism (White-Ziegler *et al.*, 2000).

1.6.2 Dam-dependent phase variation of the *sfa*, *daa*, *foo and pef* operons

The *E. coli sfa* operon codes for S pili, which are associated with strains localized to the urinary tract in humans (Schmoll *et al.*, 1990). The *daa* operon codes for F1845 pili, which are associated with strains causing diarrhoea in humans (Bilge *et al.*, 1989). Like Pap, both F1845 pili and S pili are subject to phase-variable expression (van der Woude and Low, 1994). Both the *sfa* and *daa* operons code for proteins that have a high degree of homology with the *pap*-encoded regulatory proteins PapI and PapB. The *sfa*-encoded SfaC protein and *daa*-encoded DaaF protein share high sequence identity with PapI (Bilge *et al.*, 1993). The SfaB and DaaA proteins both have regions that are highly homologous with PapB. The regulatory regions of both these operons contain two 5′-GATC-3′ sequences, which are spaced such that they correspond to the $GATC_{prox}$ and $GATC_{dist}$ of the *pap* operon (van der Woude *et al.*, 1992). In both these operons the $GATC_{dist}$ site is non-methylated in Phase ON cells and the $GATC_{prox}$ site is non-methylated in Phase OFF cells. These methylation patterns are dependent on the binding of Lrp and in turn they modulated Lrp binding (van der Woude and Low, 1994). In addition the *daa* and pap operons are both regulated, in response to the same environmental signals, by the histone nucleoid-like structuring protein, H-NS (White-Ziegler *et al.*, 2000). Based on these similarities it is likely that the *sfa*, *daa* and *pap* operons share a common methylation-dependent regulatory mechanism for phase variation (van der Woude and Low, 1994).

Another phase-variable operon whose regulation shares features with that of *pap* is the *foo* operon of *E. coli* O115:F165 (Harel *et al.*, 2000). *foo* codes for the phase-variable F165(1) fimbriae, which are associated with septicaemia in pigs. This operon too codes for PapI and PapB homologues and contains two 5'-GATC-3' sites, which are separated by 102 bp. Repression of F165(1) pilus expression by leucine is indicative of a direct role for Lrp, which may be modulated by Dam methylation, in the regulation of *foo* (Harel *et al.*, 2000).

Expression of the phase-variable *E. coli* CS31A surface antigen, encoded by the *clp* operon, is also controlled by the Dam methylation of 5'-GATC-3' sites (Martin, 1996). Phase variation requires a PapI homologue as well as Lrp, which protects the *clp* 5'-GATC-3' sites from methylation by Dam. In addition to controlling phase variation of *clp*, Lrp regulates the actual level of transcription, along with a PapB homologue, ClpB (Martin, 1996).

Methylation-dependent gene expression is also found outside of *E. coli*. Expression of plasmid-encoded fimbriae (Pef) of *S. typhimurium* is subject to phase variable regulation that is dependent on Dam and Lrp (Nicholson and Low, 2000). Regulation of *pef* also involves homologues of the PapI and PapB regulatory proteins. Both Lrp and Dam are required for transcription while H-NS and the stationary phase sigma factor, RpoS, are negative regulators of *pef* (Nicholson and Low, 2000).

1.6.3 Phase variation of Type-1 fimbriae

The *fim* gene cluster encodes the structural components of Type-1 fimbriae, the fimbrial biosynthesis machinery and proteins involved in *fim* regulation (Klemm *et al.*, 1985).

Phase-variation of Type-1 fimbriae occurs via the inversion of a 314-bp DNA fragment containing the *fimA* promoter, which drives expression of the *fim* structural genes (McClain *et al.*, 1993; Olsen and Klemm, 1994). This invertible element is flanked by two 9-bp inverted repeats. When the *fimA* promoter is pointed towards the *fimA* gene transcription can proceed and cells will be fimbriate (Phase ON). When the promoter points in the opposite direction there is no *fimA* transcription and cells will be afimbriate (Phase OFF; McClain *et al.*, 1993).

Inversion of this 'switch' is under the control of two site-specific recombinases, FimB and FimE. FimE and FimB bind to half sites, which overlap with the two *fim* inverted repeats (Gally *et al.*, 1996). While both FimB and FimE mediate inversion of the *fim* switch, FimE promotes only ON→OFF switching whereas FimB promotes both ON→ OFF and OFF→ON switching (McClain *et al.*, 1993). The actions of both recombinases are modulated in response to growth temperature. FimB-promoted inversion occurs optimally between 37°C and 40°C whereas FimE-promoted inversion occurs more at temperatures below 37°C (Gally *et al.*, 1993). Thus, switching towards the ON state is favoured at 37°C, the internal temperature of the mammalian host.

For inversion to occur the two inverted repeats that flank the invertible element must be brought into contact with each other so that synapsis, strand exchange and rejoining may occur (Hallet and Sherratt, 1997). The *fim* system requires the action of additional proteins to help overcome topological constraints that would otherwise prevent the necessary configuration of DNA. This DNA manipulation role is fulfilled by Lrp, Integration Host Factor (IHF) and H-NS.

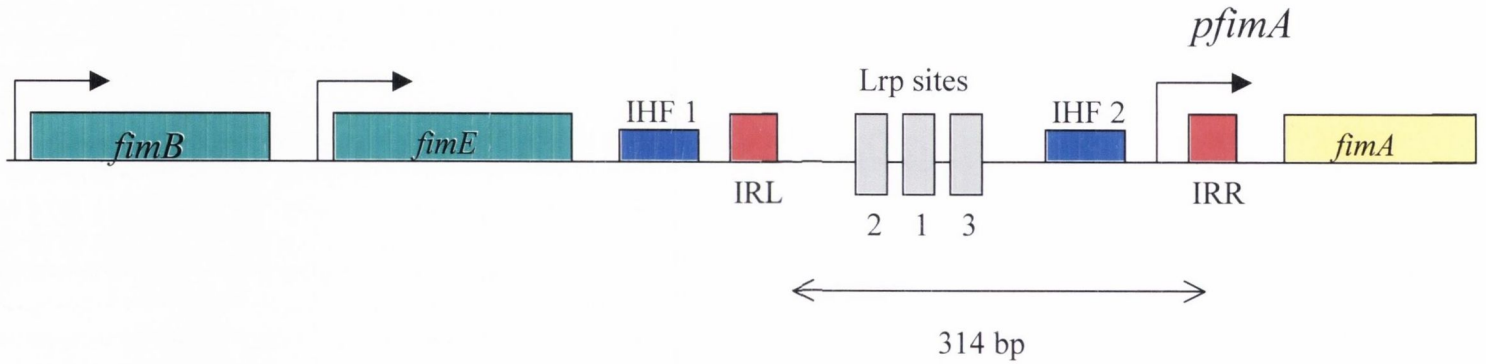
The action of IHF, H-NS and Lrp is required, in addition to FimB and FimE, for normal frequencies of *fim* switch inversion (Dorman and Higgins, 1987; Eisenstein *et al.*, 1987; Blomfield *et al.*, 1993). Lrp binds to three sites within the *fim* switch (Figure 1.11). Binding at just sites 1 and 2 results in a DNA conformation that is more likely to undergo inversion than that produced by binding to all three sites. Although leucine has an effect on Lrp bound at sites 1 and 2, it has a more pronounced affect on Lrp bound at site 3, promoting greater dissociation from this site and thereby stimulating inversion of the *fim* switch (Roesch and Blomfield, 1998).

The DNA bending protein, IHF, has an important role in the site-specific recombination of the *fim* invertible element. IHF is absolutely required for both FimE- and FimB-mediated switching and positively regulates the *fimA* promoter (Dorman and Higgins, 1987; Blomfield *et al.*, 1997). IHF binds to two distinct sites: one (IHF1 immediately downstream of *fimE* and the other (IHF2) within the invertible element (Figure 1.11). IHF is thought to be involved in wrapping and bending the switch DNA into a recombination-proficient conformation (Gally *et al.*, 1993; Blomfield *et al.*, 1997).

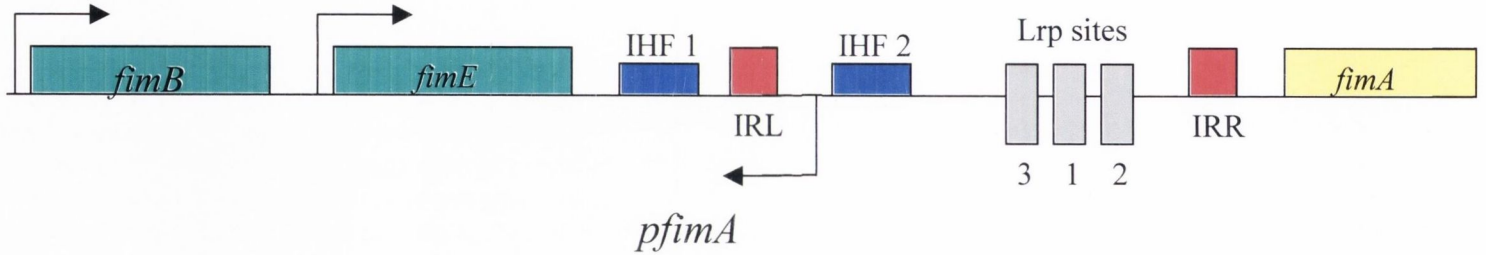
Strains bearing mutations in the gene encoding H-NS have 100-fold increased rates of *fim* switch inversion. This may occur via direct interaction with the *fim* switch or through regulation of the gene encoding FimB (Donato *et al.*, 1997). H-NS also silences transcription of the *fimA* promoter and has a greater repressive effect at 30°C than at 37°C.

Figure 1.11 Diagram showing the arrangement of the *fim* switch in orientations corresponding to the Phase ON and Phase OFF states of expression. The positions of the two IHF binding sites and three Lrp binding sites are shown. The right and left inverted repeats are labelled IRR and IRL, respectively and the *fimA* promoter is labelled *pfimA*.

Phase ON



Phase OFF



1.7 Aims and Objectives

When this work began no studies on the subject of *agn43* transcriptional regulation had yet been published. It had only been reported that expression of the Ag43 protein was affected by mutations in *oxyR* and *dam* and that the regulatory region of *agn43* contained putative binding sequences for these proteins (Henderson *et al.*, 1997). These observations were highly suggestive of the involvement of OxyR and Dam in the regulation of *agn43* transcription. Accordingly, the aims of the current study were to clarify the role of these proteins in regulation of Ag43 phase variation at the level of transcription. A possible role for Dam-mediated DNA methylation of 5'-GATC-3' sites in the regulation of *agn43* expression was particularly intriguing. Thus the role of the Dam target sequences was dissected by extensive mutagenesis and the effect of different DNA methylation patterns on binding and repression by OxyR was investigated (Chapter 5). In addition, to determine whether the mechanism of *agn43* regulation was comparable to that of other systems the *agn43* RNA polymerase contact sites were analysed by mutagenesis (Chapter 3).

As mentioned earlier, bacteria regulate the expression of certain virulence genes in response to different environmental cues. Thus, it was decided to determine whether *agn43* was subject to such environmental regulation (Chapter 3). Finally, in a bid to provide further insight into the mechanism for transcriptional regulation of *agn43*, the possible involvement of regulators other than Dam and OxyR in phase variation was also investigated (Chapter 4).

During the course of this project related studies, on the subject of *agn43* regulation, have been reported by others (Haagmans and van der Woude, 2000; Wallecha *et al.*,

2002; Correnti *et al.*, 2002). These will be discussed in the context of the current study in relevant chapters of this dissertation.

Chapter 2

Materials and Methods

2.1 Chemicals and growth media

2.1.1 Chemicals, enzymes and radionucleotides

All chemicals and reagents used in this study were purchased from BDH Chemicals Ltd., Gibco BRL, Fisons, Pharmacia, USB or Sigma Chemical Company Ltd. DNA restriction and modifying enzymes were purchased from New England Biolabs, Promega, Stratagene, or Boehringer Mannheim. All radionucleotides were supplied by Amersham International or NEN Life Sciences.

2.1.2 Growth conditions and media

Unless otherwise stated, bacteria were grown on L agar at 37°C or in L broth at 37°C with shaking. Materials for the preparation of growth media were supplied by Difco or Oxoid. Media were sterilised by autoclaving at 120°C for 20 min prior to use. Any solutions which would be damaged by autoclaving e.g. antibiotic solutions, were sterilised by filtration through 0.2μm Acrodisc® Syringe Filters (Gelman Sciences). Solid media were prepared by the addition of 15 g of agar per litre of medium prior to autoclaving. All quantities listed below are for the preparation of 1 litre of medium in distilled, deionised water. Media were supplemented with appropriate antibiotics as required.

L broth and L agar contained Oxoid tryptone (1%), Yeast extract (0.5%) and NaCl (O.5%). SOC medium, used following electroporation to increase the efficiency of transformation, contained 2% (w/v) Bactotryptone, 0.5% (w/v) Bacto yeast extract, 10 mM NaCl, 2.5 mM KCl, 10 mM $MgCl_2$, 10 mM $MgSO_4$ and 20 mM glucose.

Ampicillin (Ap) was prepared as a 100 mg/ml stock solution in ddH_20 and used in media at a final concentration of 100 μg/ml. Chloramphenicol (Cm) was prepared as a 50 mg/ml stock solution in 100% ethanol and used at a final concentration of 20 μg/ml. Kanamycin (Km) was prepared as a 25mg/ml stock solution in ddH_2O and used in media at a final concentration of 50 μg/ml. Tetracycline (Tc) was prepared as a 10 mg/ml stock solution in 100% ethanol and used at a final concentration of 10 μg/ml. All antibiotic solutions were stored at -20°C.

5-bromo-4-chloro-3-indoyl-β-D-galactosidase (X-Gal), a chromogenic substrate for β-galactosidase was prepared as a 20 mg/ ml stock solution in dimethyl formamide and stored in the dark at -20°C. X-Gal was used in media at a final concentration of 50 μg/ml. isopropyl-β-D-thiogalactopyranoside (IPTG) was prepared as a 100 mM stock solution in ddH_2O and stored at –20°C. IPTG was used in liquid media at a final concentration of 40 μM.

2.2 Bacterial strains, plasmids and bacteriophage

2.2.1 Bacterial strains and plasmids

All bacterial strains and plasmids used in this study are listed in Table 2.1 together with relevant details and source. All bacterial strains were derivatives of *Escherichia coli* K-12. Detail of the construction of plasmids resulting from this work will be described in the relevant chapters.

2.2.2 Bacteriophage

The bacteriophage used in this study was bacteriophage P1*cml* (see Section 2.4). Transducing lysates were routinely stored at 4°C in 5-ml volumes supplemented with 100µl chloroform.

2.3 Oligonucleotides

The sequence of oligonucleotides used in this study are listed in Table 2.2. Oligonucleotides were synthesised by MWG-Biotech Ltd.

2.4 Genetic techniques.

2.4.1 Use of the generalized transducing phage P1*cml*

P1*cml*, a derivative of the bacteriophage P1, was routinely used to transduce mutations from one bacterial strain to another. P1 packages between 110 and 115-kbp of bacterial DNA during lytic growth. Transducing particles represent between 0.3% and 6% of total phage particles in any P1 lysate. P1*cml* contains the temperature sensitive c1.100 repressor mutation allowing lysogenisation of bacteria at temperatures below 33°C but not above 37°C. This phage also contains a Tn*9* transposon, which encodes Cm resistance and allows the identification of P1*cml* lysogens as Cm resistant colonies.

2.4.2 Preparation of a high-titre P1*cml* lysate

P1*cml* lysogens of the appropriate donor strain were made by resuspending 3 colonies in 100 µl of L broth and plating onto L-agar plates containing Cm. 10 µl of a P1*cml* lysate, which had been prepared from the wildtype *E. coli* strain MC4100, were spotted onto the plate, which was then incubated at 30°C overnight. A high-titre lysate was then made following the procedure outlined by Silvahy *et al.* (1984). 50 µl of overnight culture were used to inoculate 10 ml of L broth. The culture was grown to an OD_{600} of

0.1 to 0.2 at 30°C with aeration, and then heat shocked by transfer to a static waterbath at 42°C h for 20 min. During heat shock the culture was periodically shaken for 3-4 seconds. The culture was then incubated at 37°C with aeration until the cells had visibly lysed. Typically lysis took 45-60 min. Following the addition of 100µl chloroform the culture was mixed and cellular debris was pelleted by centrifugation at 5000rpm for 10 minutes at 4°C. The supernatant constituting the high-titre lysate was removed and maintained sterile by the addition of 100 µl of chloroform.

2.4.3 Transduction with P1*cml*

A single colony of the recipient strain was used to inoculate 5 ml of L broth and and cells were grown to stationary phase at 37°C with aeration. Cells were pelleted by centrifugation at 5 000 g for 10 min, then resuspended in 2.5 mls of a solution containing 10mM $MgSO_4$ and 5mM$CaCl_2$. Recipient cells and phage were added to sterile test tubes in the ratios shown in Table 2.3.

The tubes were incubated in a static waterbath at 37°C for 30 min to allow the adsorption of phage to cells. 100 µl of 1 M sodium citrate was then added followed by 1 ml of L broth that had been prewarmed to 37°C. Tubes were then incubated for 1 h to allow phenotypic expression of the transduced marker. Cells were then pelleted by centrifugation at 5 000 g for 45 sec, resuspended in 100µl L broth and spread onto selection plates. These were then incubated overnight at 37°C. Plates corresponding to tubes 1 and 5 control for reversion and bacterial contamination of the lysate, respectively. Transductants from plates corresponding to tubes 2-4 inclusive were

single-colony purified then tested for the presence of a stable lysogen by streaking onto L plates containing Cm and incubating these at 30°C for 48 h.

2.4.4 Transformation of *E. coli* with plasmid DNA

Two separate methods for the transformation of *E. coli* with plasmid DNA were employed viz. (a) calcium-chloride induction of competence followed by heat-shock uptake of plasmid DNA, and (b) electroporation.

2.4.4.1 Preparation of competent cells by the calcium-chloride method

100 µl of an overnight culture of the strain to be transformed were used to innoculate 100 ml of L broth and the bacteria were grown to an OD_{600} of 0.4. The cells were pelleted by centrifugation at 2000 g for 5 min at 4 °C and the bacterial pellet was resuspended in 10 ml of ice-cold $CaCl_2$ solution (60 mM $CaCl_2$, 15% (v/v) glycerol, 10 mM PIPES, pH7). After incubation on ice for 30 min, the cells were centrifuged as described above and resuspended in 5 ml of ice-cold $CaCl_2$ solution. After a final centrifugation step the cells were resuspended in 2 ml of $CaCl_2$ solution, aliquoted into 100-µl volumes and stored at -70°C.

2.4.4.2 Transformation of calcium chloride competent cells

10 ng - 1 µg of the DNA to be transformed was added to a sterile test tube and incubated on ice for 5 min. 100 µl of competent cells were added and the mixture was left on ice for a further 10 min. The mixture was heat shocked at 42°C for 2 min (or at 37°C for 5 minutes) and then placed on ice for 2 min. 1 ml of L broth was added and the culture was incubated at 37°C for 1 h. 100-µl, 10-µl ad 1-µl quantities of the transformation mix were plated onto L agar containing the appropriate antibiotic for

selection. As a control for contamination, cells to which no DNA had been added were treated in the same way. Following overnight incubation at 37°C, single colonies were streaked to purity on fresh selection plates.

2.4.4.3 Preparation of electrocompetent cells

20 µl of an overnight culture of the recipient strain were used to inoculate 10 ml of L broth and the culture was grown at 37°C to an OD_{600} of 0.6. The bacterial culture was then incubated on ice for 15 minutes before the cells were pelleted by centrifugation at 3 500 g for 10 min. Cells were then resuspended in 10 ml of ice-cold sterile ddH_2O. The centrifugation step was repeated and the cells were resuspended in 5 ml of ice-cold ddH_2O. After another similar centrifugation step, cells were washed in 2 ml of ice-cold 10% (v/v) glycerol, pelleted and finally resuspended in 200 µl of ice-cold 10% (v/v) glycerol. The electrocompetent cells were aliquoted as 100-µl volumes and stored at -70°C.

2.4.4.4 Electroporation of electrocompetent cells

DNA to be transformed (100-500 ng in a volume not exceeding 4 µl) was added to an electroporation cuvette (EquiBio Ltd.) with a gap-width of 2 mm. The cuvette was placed on ice for 5 min before the addition of 70 µl of electrocompetent cells. After a further 2min incubation on ice, the cuvette was placed in the Gene Pulser chamber (BioRad). A pulse was delivered with the Gene Pulser, which was set to deliver 2.5 kV from the 25 mF capacitor. 1 ml of prewarmed SOC medium was added to the cuvette and the cells were transferred to a sterile test tube and incubated at 37°C for 1 h. Cells were pelleted by centrifugation (10 000 g for 2 min at 4°C) resuspended in 100µl of broth and plated onto L agar plated containing the appropriate antibiotic. Cells to

which no DNA had been added were treated in the same way and thus served as a control for contamination of the electrocompetent cells. After incubation of plates at 37°C overnight, transformants were streaked on fresh L agar plates containing the appropriate antibiotic.

2.5 Assays based on spectrophotometry

2.5.1 Monitoring the growth of bacterial cultures

The growth of bacterial cultures was monitored by measurement of the optical density of the culture at a wavelength of 600 nm (OD_{600}). 1 ml of the culture was transferred into a disposable cuvette (Greiner) and the OD_{600} value measured in a Genesys UV10 spectrophotometer. Cultures were diluted appropriately if the OD_{600} was greater than 1.0.

2.5.2 Assay of β-galactosidase activity

β-galactosidase activity of sodium dodecyl sulphate (SDS)-$CHCl_3$-permeabilised cells was assayed by an adaption of the method of Miller (1972). 75 µl of the bacterial culture to be assayed were transferred into duplicate Eppendorf tubes. Cells were permeabilised by adding 725 µl of Z-buffer (60 mM Na_2HPO_4, 40 mM NaH_2PO_4, 50 mM β-mercaptoethanol, 10 mM KCl, 1 mM $MgSO_4$, pH 7.0), 25 µl of $CHCl_3$ and 25 µl of 0.1% (w/v) SDS, and vortexing duplicate tubes for 10 seconds. Following a 10-min incubation at 28°C, 150 µl of the chromogenic β-galactosidase substrate o-nitrophenyl-β-D-galactopyranoside (ONPG; 4 mg/ml in ddH_2O) were added to each tube and incubation at 28°C continued until a yellow colour was obtained. The reaction was stopped by adding 375 µl of 1 M Na_2CO_3 and the tubes were centrifuged at 10 000g for 5 minutes to pellet cellular debris. 1 ml of each supernatant was placed in a plastic

cuvette and the OD_{420} was determined. The amount of β-galactosidase activity, expressed in Miller units (MU), was calculated as follows:

$$MU = \frac{OD_{420} \times 1000}{T \times V \times OD_{600}}$$

where T = Reaction time in minutes and V = Volume of cells added in ml (usually 0.75 ml).

Duplicate assays were performed on duplicate cultures and mean values, with standard deviations (SD) in parenthesis, are given throughout this study.

2.5.3 Determination of protein concentration by Bradford assay

Protein concentration was determined using the Bio-Rad protein assay according to the manufacturer's instruction. The kit is based on the method of Bradford (Bradford, 1976) and exploits the shift in the absorbance maximum of an acidic solution of Coomassie Brilliant Blue G-250 from 465 nm to 595 nm upon protein binding. Typically the protein concentration of lysates were computed from measurements made at 595nm using a standard curve, employing bovine serum albumin (BSA).

2.6 Isolation of plasmid DNA

The Wizard Plus SV Miniprep Kit (Promega) was used to extract plasmid DNA from 1-10-ml overnight cultures according to the manufacturer's instructions. For large-scale isolation of plasmid DNA the Qiagen Midi Prep Plasmid Purification Kit was used to prepare DNA from 100-ml overnight cultures.

2.7 Purification of chromosomal DNA

Chromosomal DNA was isolated from 5-ml overnight cultures using the Genomic DNA Purification Kit (Edge Biosystems), in accordance with the manufacturer's protocol.

2.8 DNA amplification by polymerase chain reaction

In vitro amplification of DNA products was carried out by the polymerase chain reaction (PCR) as described by Mullis and Faloona (1987). The recombinant form of the enzyme *Taq* DNA polymerase from the thermophilic eubacterium, *Thermus aquaticus* BM (Promega) was routinely used for PCR amplification. DNA amplification for cloning purposes was conducted using *Pfu* polymerase (Promega), which harbours a 3' to 5' proof-reading exonuclease activity resulting in a reduced error rate compared to *Taq* DNA polymerase.

2.9 *In vitro* manipulation of DNA

2.9.1 Restriction endonuclease cleavage of DNA

Typically 0.5-2.0 µg of DNA were cut with 10 units of restriction enzyme in a 50-µl volume containing the appropriate reaction buffer, supplied with the enzyme. For double digests involving the simultaneous cleavage of DNA by two different endonucleases the manufacturer's guidelines were used to select a suitable buffer in which both enzymes had adequate activity. Digests with incompatible enzymes were performed sequentially with an intervening ethanol precipitation step. Reactions were conducted at 37°C for 1-2 hours unless otherwise recommended by the supplier.

2.9.2 Phosphatase treatment of restriction endonuclease cleaved DNA

Removal of the 5' phosphate group of vector DNA prior to ligation prevents the self-ligation of vector DNA by T4 ligase, thereby resulting in an increased yield of ligated vector-insert DNA products. 2U of calf alkaline phosphatase (Boehringer Mannheim), along with the supplied buffer, were added to the reaction mix following restriction endonuclease cleavage of vector DNA. Following incubation at 37°C for 30 min, the reaction mixture was heated to 65°C for 15 min to inactivate the phophatase.

2.9.3 Purification of linear DNA from agarose gels

Linear DNA fragments for cloning or $\gamma^{32}P$ labelling were subject to electrophoresis in a 1.5% (w/v) agarose gel containing 1 x TAE. The gel was then placed in a solution of ethidium bromide (1µg/ml) for 10 min and then washed in ddH_2O for 5 min. DNA was visualized by brief exposure to UV light and then excised using a surgical blade. DNA was purified from the agarose slice using the Concert Rapid Gel Extraction System (GibcoBRL) following the protocol supplied by the manufacturer. The DNA was eluted in 20-50 µl ddH_2O as required.

2.9.4 Ligation of DNA

Bacteriophage T4 DNA ligase was routinely used to clone insert DNA into plasmid vectors throughout this work. In the presence of Mg^{2+} ions, T4 ligase promotes the ATP-dependent formation of phosphodiester bonds between the 5' phosphate of one DNA molecule and the 3' hydroxyl group of another. The Rapid DNA Ligation Kit (Boehringer Mannheim) was used for DNA ligation reactions in accordance with the manufacturer's instructions. Typically 10 µl of a 20-µl reaction were transformed into the appropriate *E. coli* strain.

2.9.5 Dam methylation of PCR products

1 μg of the unmethylated PCR product was incubated for 1 h at 37°C with 1 U Dam methylase and 80μM S-adenosyl methionine in the appropriate buffer (New England Biolabs). To estimate the extent of methylation of 5'-GATC-3' sites methylated and control unmethylated PCR products were then digested with *Mbo*I, which is specific for unmethylated 5′-GATC-3′ sites. The *Mbo*I-restricted PCR products were subjected to electrophoresis on a 2% (w/v) agarose gel and the degree of methylation, as indicated by protection from *Mbo*I digestion, was assessed.

2.9.6 Ethanol precipitation of DNA

0.1 volumes of 3 M sodium acetate (pH 5.2) and 2 volumes of 100% ethanol were added to the DNA suspension. The contents of the tube were mixed by inversion and incubated at -70°C for 1 hour before centrifugation at 15 000 g at 4°C for 30 min. The pellet was washed in 300 μl of 70% (v/v) ethanol, air-dried and resuspended in a suitable volume of ddH_2O.

2.9.7 Phenol/chloroform extraction

Proteinaceous debris was removed from DNA solutions by mixing with phenol.chloroform. An equal volume of phenol.chloroform was added to the DNA solution, which was then mixed by vortexing for 20 seconds. The layers were separated by centrifugation at 10 000 g for 5 minutes. The aqueous layer was removed and further extractions performed with phenol.chloroform as required. An equal volume of chloroform was then added, the tube mixed and the layers separated by centrifugation as before. Nucleic acids were then precipitated as described in section 2.9.6.

2.9.8 5'-end labelling of DNA using [γ^{32}P]ATP

Single-stranded DNA oligonucletide primers for use in primer extension, and double stranded PCR products for use in electrophoretic mobility shift assays, were labelled at their 5′ ends with [γ^{32}P]-dATP using bacteriophage T4 polynucleotide kinase (T4 PNK). 20 pmol of DNA was mixed with 50 µCi [γ^{32}P]-dATP, 2 µl X 10 kinase buffer (70 mM Tris-HCl, 7 mM $MgCl_2$, 5 mM dithiothreitol (DTT), pH7.6) and 15 U T4 PNK (10 U/µl) in a 1.5-ml screw-cap Eppendorf tube and incubated at 37°C for 30 min. Samples were then precipitated with 3 µl sodium acetate and 90 µl 100% ethanol at -70°C for one hour, and the labelled DNA washed with 70% (v/v) ethanol and finally resuspended in 20µl ddH_2O.

2.10 DNA sequencing

The dideoxy chain termination method described by Sanger *et al.* (1977) was used to determine the nucleotide sequence of DNA. This method involves the *in vitro* synthesis of a DNA strand complementary to a template by T7 DNA polymerase primed from a specific oligonucleotide. The synthesis reaction in terminated by the incorporation of a 2′ 3'-dideoxynucleotide 5'-triphosphate (ddNTP), which lacks the 3′-OH group necessary for chain elongation. dNTPs are mixed with one of each of the four chain terminating ddNTPs in four separate reactions. The dATP in each reaction is labelled with [^{35}S]dATP. Polymerisation results in termination in a fraction of the population of chains at each site where a dNTP can be incorporated. The products of each of the four reactions are radioactively labelled as they are synthesised. Reaction mixes are then subjected to electrophoris alongside each other on a high-resolution polyacrylamide gel allowing sequence discrimination following autoradiography. In this study 5 µg of plasmid DNA to be sequenced was dissolved in 32 µl ddH_2O and denatured by

incubation with 8 µl of 2 M NaOH for 10 min at room temperature. After ethanol precipitation (see Section 2.9.6) the DNA was pelleted by centrifugation for 15 minutes at 3000 g in a Sorval T-20 centrifuge at 4°C. The supernatant was removed, the DNA pellet was washed with 70% (w/v) ethanol and centrifugation was repeated. After being dried under vacuum the DNA was redissolved in 10 µl of ddH_20. Sequencing was conducted using the Pharmacia T7 polymerase sequencing kit according to the manufacturer's instructions. The sequencing reactions were analysed on a 6% polyacrylamide gel as described below.

2.11 Denaturing polyacrylamide gel electrophoresis

Sequencing and primer extension products were separated by electrophoresis using 6 % (w/v) polyacrylamide gels containing 8.3 M urea, in 1 x TBE buffer. A model S2 sequencing gel electrophoresis apparatus (Gibco BRL), and computer-controlled model 3000 XI power supply (Bio-Rad), set at 70 W, were used.

2.11.1 Pouring the sequencing gel

The gel was prepared by mixing 18 ml SequaGel (National Diagnostics) concentrate, 7.5 ml SequaGel Buffer and 49.5 ml SequaGel diluent. 300 µl of 10% (w/v) ammonium persulphate were added followed by 50 µl of N,N,N',N'-tetramethylethylene diamine (TEMED). The solution was mixed well and poured between two glass plates, separated by spacers, which had been cleaned with 100 % ethanol and taped together. The straight edge of a comb was inserted into the top of the molten gel and polymerisation was allowed to proceed for at least 2 h.

2.11.2 Loading and running the gel

The polymerised gel was assembled in the electrophoresis apparatus and 1 x TBE was added to the top and bottom reservoirs. The comb was removed and any unpolymerised gel mix was cleared from the top of the gel by flushing with electrophoresis buffer. The combs were inserted with the teeth penetrating the gel by approximately 2 mm. The gel was allowed to heat at a constant power setting of 70 W for 1 h. The sequencing reactions/primer extension products were heated to 90°C for 4 min before being were loaded into wells as 4 µl aliquots. The samples were subjected to electrophoresis at 70 W for 90 – 120 min.

2.11.3 Treatment of the gel prior to autoradiography

Once electrophoresis was complete, the gel apparatus was dismantled and the glass plates were separated. The gel was transferred to a piece of 3-mm filter paper (Whatmann), covered in Saran™ Wrap and dried under vacuum at 80°C for 1 h using a Savant SG Speedgel® gel drying system.

2.11.4 Autoradiography

Autoradiography was used to detect the radio emissions from [γ-^{32}P]ATP and α-^{35}S-dATP used in reiterative PCR, primer extension analysis and DNA sequencing. The X-ray film used in all cases was double-coated Hyperfilm-MP purchased from Amersham.

2.12 Manipulation of RNA *in vitro*

2.12.1 RNA extraction

Total RNA was extracted from stationary phase cultures using the Qiagen RNeasy Kit according to the manufacturer's instructions. Samples were stored at -70°C. To

inactivate any contaminating RNase, diethylpyrocarbonate (DEPC) to a final concentration of 0.1% (v/v) was added to all solutions used for subsequent RNA work except those containing Tris. Further measures to minimise RNase contamination included the wearing of gloves, using apparatus dedicated to RNA work and using microfuge tubes and pipette tips straight from the bag in which they were supplied (i.e. they had not been previously handled).

2.12.2 Primer extension analysis

Extension of a radiolabelled oligonucleotide with reverse transcriptase using mRNA as a template allowed mapping of the 5′ ends of mRNA transcripts. Reverse transcriptase synthesises cDNA in the 5′ to 3′ direction using an oligonucleotide to prime the synthesis. Extension stops when the 5′ end of the mRNA template is encountered resulting in a single-stranded DNA molecule of defined origin whose length is determined by the transcription start site of the mRNA transcript. The start site is identified by the electrophoresis of the labelled cDNA alongside a DNA sequencing ladder generated using the same oligonucleotide used to prime cDNA synthesis.

2.12.2.1 Annealing of the primer and template

Oligonucleotide primers complementary to sequences, estimated to be 150 bp from the predicted 5' ends of transcripts to be analysed, were labelled with [γ-^{32}P]ATP as described in section 2.9.8. 50 μg total cellular RNA and 2 pmol of [γ-^{32}P]ATP-labelled primer were mixed and incubated at 90°C for 3 minutes to denature RNA. The template and primer were then cooled on ice to allow annealing.

2.12.2.2 cDNA synthesis

To the solution containing annealed template-primer the following were added: 14 µl of ddH_20, 1 µl each of dATP, dTTP, dCTP and dGTP (each stock solution 100 mM), 10 µl of 100 mM DTT, 20 µl of 5 x First Strand Buffer (Gibco BRL). 1 U RNase inhibitor (Boehringer Mannheim) and 1 U Superscript™ II RT reverse transcriptase (Gibco BRL) were then added and primer extension was allowed to proceed by incubating at 42°C for 90 min. RNA was removed by incubating at 42°C for 10 minutes in the presence of 5 µl RNase (500 µg/ml; Boehringer Mannheim).

2.12.2.3 Purification of the cDNA products

The cDNA extension products were purified by phenol/chloroform and chloroform extractions as described in section 2.9.7. Following ethanol precipitation (Section 2.9.6) the DNA pellets were resuspended in 6 µl of ddH_2O and 4 µl of formamide dye mix (95% (v/v) formamide, 20mM EDTA, 0.05% (w/v) bromophenol blue, 0.05% (w/v) xylene cyanol FF). Extension products were analysed by denaturing polyacrylamide gel electrophoresis (Section 2.11). A sequencing ladder created using the same primer was run alongside.

2.13 Site directed mutagenesis

Two different methods were used for site directed mutagenesis of the *agn43* regulatory region viz. (a) The unique site elimination method and (b) The Quickchange method.

2.13.1 Quickchange

Site directed mutagenesis was conducted using the Stratagene® Quickchange™ Site-Directed Mutagenesis Kit in accordance with the manufacturer's instructions. Two

primers containing the desired mutation, each complementary to opposite strands of the vector, were extended during temperature cycling by using *pfu* DNA polymerase. This involved 20 cycles of denaturation (94°C for 45 sec), annealing (60°C-70°C for 45 sec) and extension (72° C for 2 min per kb of vector DNA). The annealing temperature was adjusted as required depending on the melting temperature of mutagenic primers. Incorporation of the mutagenic primers generated a mutant plasmid containing staggered nicks. This product of the reaction was treated with *Dpn*I, which being specific for methylated and hemimethylated DNA, digested parental DNA leaving the unmethylated mutation-containing product intact. The nicked product was then transformed into XL1-Bue competent cells. The presence of the desired mutation was confirmed by sequencing (see Section 2.10).

2.13.2 Unique-site elimination

This method was conducted using the USE mutagenesis kit (Amersham), which is based on the unique-site elimination method of Deng and Nickloff (1992). With this method the selection of plasmids containing the desired mutation is based on the elimination of a unique restriction site elsewhere on the plasmid. This method requires two primers that anneal to different regions on the same strand of the template plasmid. The mutagenic primer contained the desired mutation whilst the other (USE primer) contained a mutation that eliminated a unique *Sal*I restriction site within the multiple cloning site of pAg43. Following denaturation at 100°C for 5 min, primers were annealed to the template DNA by chilling on ice for 5 min. Primers were then extended at 37°C for 1 h using T4 DNA polymerase. The synthesis of a new DNA strand incorporating both the mutagenic and USE primers gave rise to a plasmid containing one mutagenized strand and one non-mutagenized strand. In order to linearize any parental DNA, the product

was digested with the enzyme *Sal*I, whose restriction site had been eliminated by the USE primer. Digested DNA was transformed into the *mutS* strain of *E. coli*, which is deficient in mismatch repair. Replication of the circular plasmid in this strain gives rise to one wild type molecule and one molecule containing the desired mutations on both strands. Plasmid DNA was prepared from this strain and digested with *Sal*I, which linearized the wildtype plasmid leaving only mutagenized plasmid. DNA was transformed into XL-1, then recovered and sequenced manually (see Section 2.10) to confirm the presence of mutations.

2.14 Determination of switching frequencies

L agar containing X-Gal (Section 2.1.2) was used for the detection of *agn43-lacZ* expression in strains containing the plasmid pAg43. On this medium, colonies composed predominantly of cells Phase ON for *agn43-lacZ* expression appear blue, due to the production of β-galactosidase under the control of the *agn43* promoter, conversely colonies composed predominantly of Phase OFF colonies appear white. Five white and five blue colonies of each strain to be examined were serially diluted, plated and then incubated for 16 h at 37°C. The total viable count (N) and the number of colonies having switched from the original phenotype of the innoculum (M), were estimated. Based on the assumption that predominantly Phase ON colonies and predominantly Phase OFF colonies are derived from Phase ON and Phase OFF cells, respectively, the following equation was used to calculate the frequency of Phase switching.

Switching frequency (per cell per generation)= $1 - \sqrt[g]{1 - M/N}$

Where g=number of generation of growth.

2.15 Southern blotting

Specific sequences in electrophoretically separated fragments of DNA were detected according to the principle devised by Southern (1975). In brief, DNA is digested with the restriction endonucleases prior to electrophoresis on an agarose gel. DNA is subsequently denatured then transferred from the gel to a nylon membrane to which it is then fixed by UV crosslinking. A digoxigenin (DIG)-labelled DNA probe, complimentary to the sequence of interest is then allowed to hybridise to corresponding fragment(s) within the membrane. A chemiluminescent immunoassay allows the detection of the DIG moiety of the hybridised probe. Autoradiography is used to locate the positions and to estimate relative abundances of the DNA fragments complimentary to the probe.

2.15.1 Generation of DIG-labelled probes by PCR

DIG was incorporated into a double-stranded DNA probe by PCR. This was carried out using *Taq* DNA polymerase as described in Section 2.8 except that 5 µl PCR Dig Probe Synthesis Mix (2 mM dATP, 2 mM dCTP, 2 mM dGTP, 1.3 mM dTTP and 0.7 mM alkali-labile DIG-dUTP, pH 7.0) (Boehringer Mannheim) were used instead of the standard dNTPs.

2.15.2 Transfer of DNA

After digestion with appropriate restriction endonucleases DNA was separated by electrophoresis on a 1.5 % (w/v) agarose gel (Section 2.9.3). DNA was denatured *in situ* then transferred to a positively charged nylon membrane by capillary transfer according

to the method outlined by Sambrook *et al.* (1989). DNA was then fixed to the membrane by UV-crosslinking. This involved exposing the membrane to 50 mJoules of UV radiation administered using a Bio-Rad Gene Linker™ UV chamber.

2.15.3 Hybridisation of the probe to immobilised targets

The membrane was prehybridised at 68°C for 1 h in 20 ml of hybridisaton solution [5 x SSC (750 mM NaCl, 75 mM Na-citrate, pH 7.0), 1.0% (w/v) blocking reagent (Boehringer Mannheim), 0.1% (w/v) N-lauroylsarcosine, 0.02% (w/v) SDS]. After the initial prehybridisation step, DIG-labelled double-stranded DNA probes, diluted 1/1000 in hybridisation solution, were hybridised to the immobilised DNA targets overnight at 68°C. Two 5-min washes were performed at room temperature with 50 ml of 2 x SSC containing 0.1% (w/v) SDS followed by two washes with 50 ml 0.5 x SSC containing 0.1% (w/v) SDS for 15 min at 68°C.

2.15.4 Chemiluminescent detection of bound probe

Chemiluminescent detection of the DIG-labelled probe hybridised to its immobilised target was achieved using the DIG Luminescent Detection Kit (Boehringer Mannheim) according to the manufacturer's instructions. The nylon membrane was initially incubated with blocking reagent to prevent non-specific binding of antibody to the membrane. The membrane was then incubated with an anti-DIG-alkaline phosphatase-antibody conjugate (anti-DIG-AP-conjugate) used at a 1:20 000 dilution. Following several washin steps, chemiluminescent substrate (CDP-Star™) was added and the signal was detected by exposure of the membrane to X-ray film (Amersham).

2.16 Preparation of crude protein extracts

10-ml cultures of the desired strains were grown to an OD_{600} of 1.5. After centrifugation (5 000 g, 4°C, 10 min) cells were resuspended in 250 µl Sonication Buffer [10% (w/v) sucrose, 50 mM Tris-HCl (pH 7.5), 100 mM NaCl, 1 mM EDTA, 5 mM DTT]. 7.5 µl of 10 mg/ml lysozyme solution were then added and the cells were mixed gently by with the aid of a pipette. Samples were then sonicated using an MSE Soniprep Sonicator (Sanyo) set to deliver 3 pulses at amplitude setting 5 for 10 sec, allowing 30 sec cooling between each pulse. Cell debris was removed by centrifugation at 25 000 g for 20 min at 4°C. The supernatant fraction, constituting the crude protein extract, was used on the same day for electrophoretic mobility shift assays after quantitation of total protein concentration by the Bradford assay (see Section 2.5.3).

2.17 Electrophoretic mobility shift assays (EMSA)

Inserts from constructs were amplified using the primers Ag43F and Ag43R (see Table 2.2) and aliquots of PCR products were Dam methylated *in vitro* (see Section 2.9.5). Unmethylated and methylated probes were end labelled with [γ^{32}P]-dATP as described in Section 2.9.10. Radiolabelled DNA was purified using a NICK™ column (Pharmacia) and eluted in 200 µl TE buffer [10 mM Tris-HCl, 1 mM EDTA, pH 8.0]. Crude protein extracts (Section 2.16) containing 4 ng total protein were incubated at room temperature for 10 min with 2 µl of radiolabelled probe in BS buffer [25 mM Tris-HCl (pH 7.7), 50 mM KCl, 5 mM $MgCl_2$, 5% (w/v) glycerol, 50 µg/ml BSA, 1mM DTT, 0.5 mg/ml herring sperm DNA]. The final volume of the reaction mix was 20 µl. The samples were subjected to electrophoresis on a native 6 % (v/v) polyacrylamide gel to allow the resolution of free DNA probe from that bound by protein. Following separation of the glass plates, the gel was transferred to a piece of Whatman 3-mm filter

paper, then covered in Saran™ wrap and dried under vacuum at 80°C for 2 h using a Savant SG Speedgel® geldrying system. Visualization by autoradiography involved exposing the dried gel for 16 hours to Hyperfilm -MP (Amersham).

2.18 Colony immunoblotting

2-ml overnight cultures of the desired strain were subjected to 10-fold serial dilutions in L broth. 100-µl aliquots of the 10^{-5} and 10^{-6} dilutions were spread onto L agar plates, which were subsequently incubated at 37°C overnight. Circular pieces of nitrocellulose were placed on the surface of the agar plates, which were then incubated at 4°C for 5 min. The nitrocellulose was removed and adherent colonies washed off by shaking in a petri dish using three changes of TS-buffer [10 mM Tris, 0.9% (w/v) NaCl, pH 7.4]. To prevent non-specific antibody binding, the nitrocellulose was blocked by incubation in 10 ml of 5% (w/v) Marvel® at room temperature for 2 h with shaking. The blot was then incubated in primary antibody solution [1:2000 dilution of anti-α^{43} in 10 ml of 5% (w/v) Marvel] at room temperature, for 90 min, with shaking. This incubation was followed by three 10-min washes in TS-buffer with shaking. The blot was then incubated in secondary antibody solution [1:1000 dilution of horseradish peroxidase conjugated goat anti-rabbit IgG in 5% (w/v) Marvel®] for 1 h with shaking. This was followed by three 10-min washes in TS buffer as above. Finally the blot was developed by incubation in 20 ml zymogram solution [18 ml 50 mM Tris-HCl (pH 7.4), 2 ml 0.6% (w/v) 4-chloro-1-napthol, 10 µl 30% H_2O_2]. When satisfactory development had been achieved, the blot was rinsed briefly in H_2O and allowed to dry.

2.19 *In vivo* footprinting by methylation protection assay

Overnight cultures of CSH50 (pAg43) and CSH50*ΔoxyR* (pAg43) were diluted 1:100 into 3 ml of L broth and incubated at 37°C until an OD_{600} of 1.0 was reached. The cells were then harvested by centrifugation at 5 000 g and resuspended in 100 µl of L broth containing 30mM dimethylsulphate (DMS). Methylation was allowed to proceed for 2 minutes at room temperature before the addition of 3 ml of ice cold DMS stop solution (10mM Tris-HCl (pH8.0), 1mM EDTA, and 40mM NaCl). The DMS-treated cells were harvested by centrifugation and washed once with 1ml of ice-cold DMS stop solution. Finally genomic DNA was isolated using the genomic DNA purification kit (Gibco BRL), according to the manufacturer's instructions. DMS-modified sites were identified on each strand by using the polymerase chain reaction to extend 5′ end-labelled oligonucleotides. Linear PCR was preceded by a 5-min incubation at 95°C to break the unstable glycosidic bonds of methylated deoxynucleotides. All PCR reactions were performed in a volume of 100 µl, containing 150 ng of DNA, 0.5 pmol of end-labelled primer, 10µl of X10 buffer [500 mM KCl, 100 mM Tris-HCl (pH 9), 1% Triton and 15 mM $MgCl_2$], 200 µM each dNTP and 2.5 units of *Taq* DNA polymerase (Promega). After 30 cycles of denaturation (95°C), annealing (50°C) and extension (72°C), the resultant extension products were phenol-chloroform extracted, ethanol precipitated (see Sections 2.9.7 and 2.9.6) and resuspended in 6 µl of ddH_2O and 4µl of formamide dye mix [95% (v/v) formamide, 0.05% (w/v) bromophenol blue, 0.05% (w/v) xylene cyanol FF, 20 mM EDTA]. Extension products were analyzed by denaturing 6% (w/v) polyacrylamide gel electrophoresis (see Section 2.11) and methylated residues identified by comparison with a DNA sequence ladder generated using the same primer.

Table 2.1 *E. coli* strains and plasmids used in this study

Strain or plasmid	Relevant characteristics[a, b]	Reference/source
E. coli		
AAEC 189	*ΔfimB-H ΔrecA*	Blomfield *et al.* (1991)
BD1302	metB *Δ(ppc-argECBH oxyR) relA spoT* (λ^+)	Diderischen (1980)
BD1302*dam*	BD1302 *Δdam::kan*	Henderson and Owen (1999)
CSH50	*ara Δ(pro-lac) rpsL thi*	Koch *et al.* (1988)
CSH50*ΔoxyR*	CSH50 *ΔoxyR::kan*	This study
CSH50 OxyR^{++}	CSH50 containing pAQ5	This study
CSH50*dam*	CSH50 *Δdam::kan*	This study
CSH50*lrp*	CSH50 *Δlrp::tet*	Charles Dorman
GS05	*ΔoxyR::kan*	Kullik *et al.* (1995)
NM522*mutS*	*thi,SupE,hsdMS-mcrB)5,(lac-proAAB) [mutS::*Tn*10]*	Amersham
XL1-Blue	*recA1 endA1 gyrA96 thi-1 hsdR17 supE44 relA1 lac* [F' *proAB lacI^{q}Z ΔM15 Tn10*];Tc^r; used for plasmid propagation	Stratagene
ZK1000	ZK126 *Δrpos::kan*	Bohannon *et al.* (1991)
ZK126	W3110 *ΔlacU169 tna-2 Rpos*$^+$	Connell *et al.* (1987)
DW1	CSH50 *Δdam::kan* containing pDEW10	This study
DW2	CSH50 *Δdam::kan* containing pDEW11	This study

Strain or plasmid	Relevant characteristics[a, b]	Reference/source
DW3	CSH50 *Δdam::kan* containing pDEW12	This study
DW4	CSH50*Δdam::kan* containing pDEW13	This study
DW5	CSH50*Δdam::kan* containing pDEW14	This study
DW6	CSH50*Δdam::kan* containing pDEW15	This study
DW7	CSH50*Δdam::kan* containing pDEW16	This study
DW8	CSH50*Δdam::kan* containing pDEW20	This study
DW9	CSH50*Δdam::kan* containing pDEW21	This study
DW10	CSH50*Δdam::kan* containing pDEW22	This study
DW11	CSH50*Δdam::kan* containing pDEW23	This study
DW12	CSH50*Δdam::kan* containing pDEW24	This study
DW13	CSH50*Δdam::kan* containing pDEW25	This Study
Plasmids		
pQF50	Promoterless *lacZ* vector, Ap^r	Farinha and Kropinski (1990)
pUC18	Cloning vector; Ap^r	Stratagene
pAQ5	*oxyR* wt in pACYC184; Cm^r, Ap^r	Kullik *et al.* (1995)
pAQ25	*oxyR* wt in pKK177-3 with altered Shine-Dalgarno sequence to allow overproduction of OxyR; Ap^r	Storz *et al.* (1990)
pBSL181	Suicide vector for transposon mutagenesis, Cm^r, Ap^r	Alexeyev and Shokolenko (1995)
pAg43	*agn43* promoter region and part of coding region cloned upstream of *lacZ* in pQF50; Ap^r	This study

Strain or plasmid	Relevant characteristics[a, b]	Reference/source
pDEW1	pAg43 with -35 mutation in *agn43* promoter; Ap^r	This study
pDEW2	*agn43* promoter region and part of coding region cloned in pUC18; Ap^r	This study
pDEW3	pDEW2 GTTC(1); Ap^r	This study
pDEW4	pDEW2 GTTC(2); Ap^r	This study
pDEW5	pDEW2 GTTC(3); Ap^r	This study
pDEW6	pDEW3 GTTC(2); Ap^r	This study
pDEW7	pDEW3 GTTC(3); Ap^r	This study
pDEW8	pDEW2 GTTC(2), GTTC(3); Ap^r	This study
pDEW9	pDEW3 GTTC(2), GTTC(3); Ap^r	This study
pDEW10	pAg43 GTTC(1); Ap^r	This study
pDEW11	pAg43 GTTC(2); Ap^r	This study
pDEW12	pAg43 GTTC(3); Ap^r	This study
pDEW13	pAg43 GTTC(1), GTTC(2); Ap^r	This study
pDEW14	pAg43 GTTC(1), GTTC(3); Ap^r	This study
pDEW15	pAg43 GTTC(2), GTTC(3); Ap^r	This study
pDEW16	pAg43 GTTC(1), GTTC(2), GTTC(3); Ap^r	This study
pDEW17	pDEW2 CATC(2); Ap^r	This study
pDEW18	pDEW2 GATG(3); Ap^r	This study
pDEW19	pDEW3 CATC(2); Ap^r	This study
pDEW20	pAg43 CATC(2); Ap^r	This study
pDEW21	pAg43 GATG(3); Ap^r	This study
pDEW22	pAg43 GTTC(1), CATC(2); Ap^r	This study

Strain or plasmid	Relevant characteristics[a, b]	Reference/source
pDEW23	pAg43 G<u>T</u>TC(1), GATG(3); Ap^r	This study
pDEW24	pAg43 <u>C</u>ATC(2), GATG(3); Ap^r	This study
pDEW25	pAg43 G<u>T</u>TC(1), <u>C</u>ATC(2), GATG(3); Ap^r	This study

[a] Ap^r, ampicillin resistant; Cm^r, chloramphenicol resistant; Tc^r, tetracycline resistant; wt, wild type

[b] For mutated plasmid constructs (pDEW3-25), numbers in brackets indicate which 5'-GATC-3' sites have been mutated. Altered bases are underlined.

Table 2.2 Oligonucleotides used in this study

Name[a]	Sequence[b]
Ag43F	5'-CGCGGGATCCTGAACCTGTCGTGACTTGATG-3'
Ag43R	5'-GGCCAAGCCTTGTCATGTGATTCCATACCAG-3'
Ag43p	5'-GTGGTTCTGTTGTTTCCCTG-3'
GTTC1t	5'-CAATAATAGAATAAACGTTCAATAATTTCCTATTTTATCG-3'
GTTC1b	5'-CGATTAAAATAAAGATATTGAACGTTTTATTCTATTATTG-3'
GTTC2t	5'-CAATATCCCCTATTTTATCGTTCGTTTATATCGATC-3'
GTTC2b	5'-GATCGATATAAACGAACGATAAAATAGGGGATATTG-3'
GTTC3t	5'-TATCGATCGTTTATATCGTTCGATAAGGCTAATAATAACC-3'
GTTC3b	5'-GGTTATTATTAGCCTTATCGAACGATATAAACGATCGATA-3'
GTTG2+3t	5'-CTATTTTATCGTTCGTTTATATCGTTCGATAAGC-3'
GTTG2+3b	5'-GCTTATCGAACGATATAAACGAACGATAAAATAG-3'
GATG3t	5'-CGTTTATATCGATGGATAAGCTAATAATAACC-3'
GATG3b	5'-GGTTATTATTAGCTTATCCATCGATATAAACG-3'
CATC2t	5'-CAATATCTATTTTATCCATCGTTTATATCGATC-3'
CATC2b	5'-GATCGATATAAACGATGGATAAAATAGATATTG-3'
P-35t	5'-CGTTACTGTCTCTGACGTCCGTGCAATAGC-3'
P-35b	5'-GCTATTGCACGGACGTCAGAGACAGTAACG-3'
USE	5'-CCATGGACGCGTGACGTGGACTCTAGAGGATCC-3'
DamF	5'-CTGCGCGTAATCTGCTGCTTGC-3'
DamR	5'-CAGCTGCCAGGGCTTTTGCGG-3'
OxyRF	5'-ATGAATATTCGTGATCTTGAGTACC-3'
OxyRR	5'-TTAAACCGCCTGTTTTAAAACTTTATCG-3'

Name[a]	Sequence[b]
CatF	5'-TCACTGGATATACCACCGTTGATATATCC-3'
CatR	5'-TACGCCCCGCCCTGCCACTCATCGCAGTAC-3'
RpoSF	5'- ATGAGTCAGA ATACGCTGAAAGTTCATGA-3'
RposR	5'- TACTCGCGGAACAGCGCTTCGATATTCAGC-3'
CatExtn	5'-GTACTGCGATGAGTGGCAGGGCGGGGCG-3'

[a] Abbreviations; F and R, forward and reverse primers, respectively; t and b, corresponding nomenclature for mutagenic primers

[b] Altered bases of mutagenic oligonucleotides are underlined

Table 2.3 Volumes of recipient cells, P1*cml* lysate and L broth that were used for P1*cml* transduction

Tube no.	Volume recipient cells	Volume P1*cml* lysate	Volume L broth
1	100 µl	0	100 µl
2	100 µl	10 µl	90 µl
3	100 µl	50 µl	50 µl
4	100 µl	100 µl	0
5	0	100 µl	100 µl

Chapter 3

Cloning and analysis of the *agn43* promoter region

3.1 Introduction

Within the regulatory region of the *agn43* is a sequence that is highly similar to the OxyR-binding sites of other genes regulated by OxyR, the *mom* gene of phage Mu and the *oxyR* gene itself. Interestingly this region contains three 5′-GATC-3′ sites, which have the potential to be methylated by the *dam* gene product (Henderson *et al.*, 1997). As such, the findings that OxyR binds to the *agn43* regulatory region *in vitro* and that this binding is abrogated by methylation of the *agn43* 5′-GATC-3′ sites (Hagmaans and Van der Woude, 2000) are not surprising. These 5′-GATC-3′ sites have been observed to be fully methylated in an *oxyR* Locked ON background suggesting that they are protected from methylation by bound OxyR in Phase OFF cells. A correlation has been observed between the proportion of Phase ON cells and the proportion of cells having methylated (or hemimethylated) 5′-GATC-3′ sites, in a population of cells. Conversely, the proportion of Phase OFF cells in a population has been observed to correlate with the proportion of cells in which the *agn43* regulatory region is unmethylated (Haagmans and Van der Woude, 2000). Together this evidence points strongly towards a model for regulation of phase variation in which binding of OxyR to unmethylated 5'-GATC-3' sites within the promoter region of *agn43,* represses transcription, possibly by causing RNA polymerase to be excluded. Dam methylation of these sites would prevent binding of OxyR thereby allowing transcription of *agn43* to proceed. Competition between OxyR and Dam for unmethylated 5'-GATC-3' sites could result in switching between the Phase ON and Phase OFF states. Such a system is known to regulate the *mom* gene of phage Mu with which the promoter region of the gene encoding *agn43* shows extensive nucleotide sequence homology (Bolker and Kahmann, 1986; Seiler *et al.*, 1989; Henderson and Owen, 1999). Successful transcription of the *mom* gene requires the dam methylation of three 5′-GATC-3′ sites within the OxyR-binding region (Kahmann, 1983;

Plasterk *et al* 1983). These 5′-GATC-3′ sites have a spatial distribution almost identical to that of the corresponding sites within the *agn43* promoter region (Henderson and Owen, 1999).

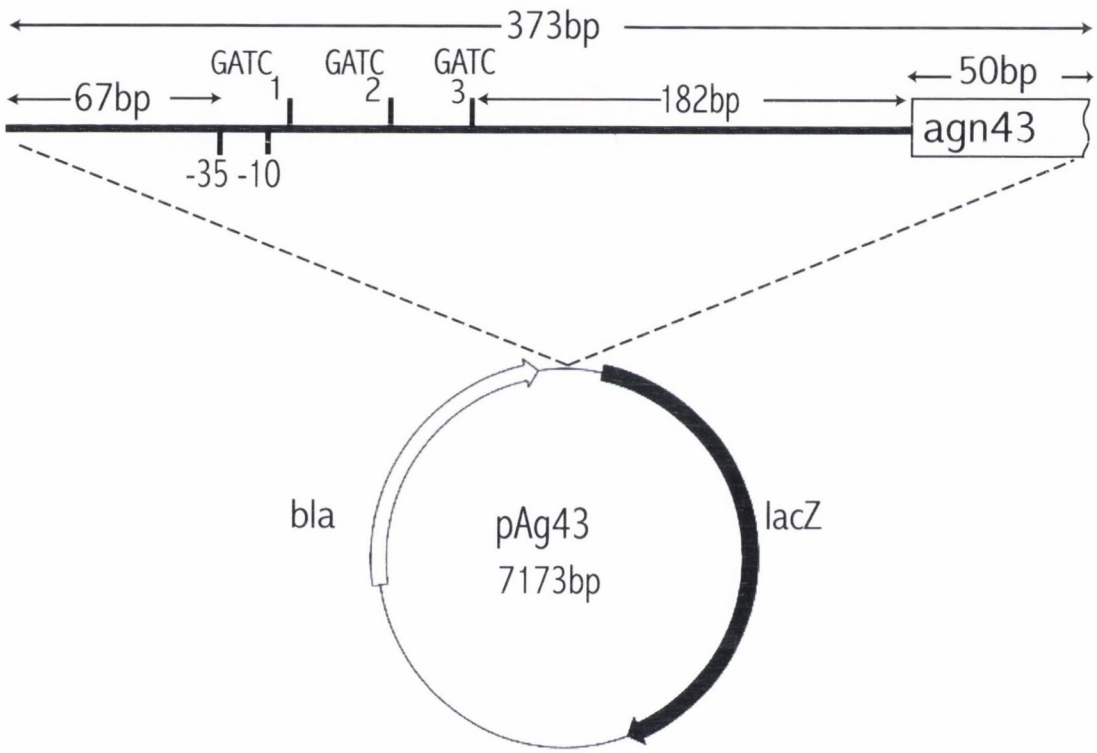
This chapter describes the construction of a transcriptional fusion of the *agn43* regulatory region to a promoterless *lacZ* gene, thus allowing the study of *agn43* regulation at the level of transcription. To determine whether regulation of this plasmid-based system reflects that of the native *agn43* gene, the correlation between expression of the *agn43-lacZ* transcriptional fusion and the expression of chromosomally encoded Ag43 protein is examined. Based on the extensive sequence similarities between the regulatory regions of *agn43* and of the *mom* gene of phage Mu, Henderson and Owen suggested that the *agn43* promoter is situated downstream of the three 5′-GATC-3′ sites (Henderson and Owen, 1999). To determine whether this is indeed the case, the *agn43* transcription start site is mapped in this chapter. Mutagenic analysis reveals that *agn43* regulation differs fundamentally from that of *mom* since, contrary to expectations, RNA polymerase binds upstream of the three Dam targets. Virulence genes are often regulated in response to environmental stimuli. Accordingly, the effect of different environmental parameters on *agn43* trasncription is assessed. It is demonstrated that the *agn43* promoter is highly responsive to certain environmental conditions. Finally, in order to further understand the critical role played by Dam methylase in *agn43* regulation, the methylation state of 5′-GATC-3′ sites, within the cloned *agn43* regulatory region, is assessed.

3.2 Results

3.2.1 Generation of an *agn43-lacZ* transcriptional fusion

To facilitate routine transcriptional studies of *agn43,* its regulatory region was fused to the promoterless *lacZ* gene in plasmid pQF50 (Table 2.1), placing transcription of *lacZ* under the control of the *agn43* promoter. The 6.8-kbp low-copy-number reporter plasmid, pQF50, utilises the promoterless *lacZ* gene from pCB267 (Schneider and Back, 1986). In this plasmid the first 15 bp of the *lacZ* gene have been replaced with the Shine-Dalgarno sequence and the first 12 bases of the *E. coli* lipoprotein (*lpp*) gene. The hybrid β-galactosidase is functionally indistinguishable from the native protein (Nakamura and Inouye, 1982). A multiple cloning site derived from pUC18 is situated upstream of *lacZ.* The plasmid contains a functional *bla* gene for selection on ampicillin. A 373-bp DNA fragment encompassing 323 bp of the regulatory region and 50 bp of the coding region of *agn43* was amplified by PCR from K-12 strain BD1302 (Table 2.1). The 5′ ends of oligonucleotides that primed from the 5′(Ag43F) and 3′(Ag43R) ends of the desired DNA fragment contained *Bam*HI and *Hin*dIII sites, respectively (Table 2.2). The amplified fragment was cloned into the *Bam*HI-*Hin*dIII site of pQF50 creating an *agn43-lacZ* transcriptional fusion and thereby making β-galactosidase expression contingent on *agn43* promoter activity. The resulting plasmid, pAg43, was sequenced to ensure that no mutations had been introduced during PCR amplification of the DNA insert. None were observed. Figure 3.1 shows a map of the plasmid pAg43.

Figure 3.1 Map of the plasmid pAg43. *lacZ*, promoterless *lacZ* gene; *bla*, β-lactamase gene. The three 5′-GATC-3′ sites of the *agn43* insert and the –10 and –35 RNA polymerase recognition sites identified in this study are shown. Distances between these features are shown in base pairs.

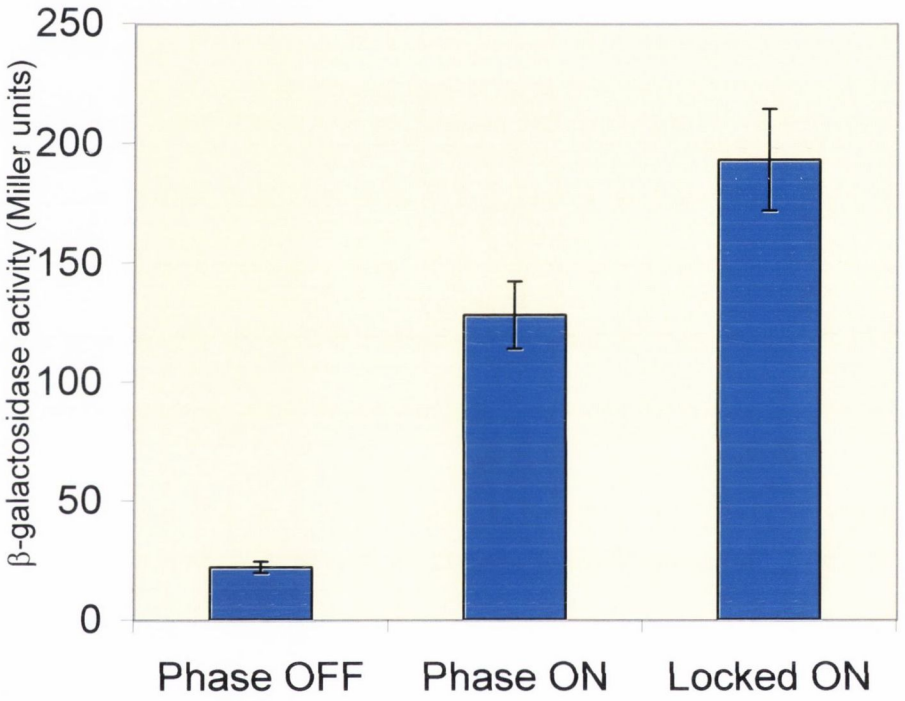

373bp
GATC 1
GATC 2
GATC 3
67bp
182bp
50bp
agn43
-35 -10
bla
pAg43
7173bp
lacZ

3.2.2 Transformation of strains CSH50 and CSH50*oxyR* with plasmid pAg43

The reporter plasmid, pAg43, was introduced into the *E. coli* K-12 strain CSH50, which is phase-variable for chromosomally encoded Ag43 expression. CSH50 is a *lacZ* knockout and is therefore unable to express β-galactosidase. As such the β-galactosidase activity of pAg43 transformants could be attributed only to expression of the plasmid borne *agn43-lacZ* transcriptional fusion. The CSH50 (pAg43) transformants gave rise to blue, white and blue-and-white sectored colonies when grown on agar plates containing the chromogenic substrate 5-bromo-4-chloro-3-indolyl-β-D-galactosidase (X-Gal), characteristic of phase variable expression of the *agn43-lacZ* fusion. The rate of switching between Lac^+ (ON) and Lac^- (OFF) phenotypes was 10^{-2} to 10^{-3} per cell per generation in both directions. This was in good agreement with the switching frequencies observed previously for Ag43 expression from the chromosome in other *E. coli* strains, as measured by immunoflourescence microscopy (Owen *et al.*, 1996: Roche *et al.*, 2001). β-galactosidase assays were conducted to determine the relative levels of transcription in Phase ON and Phase OFF cultures of CSH50 (pAg43) (Figure 3.2). The levels of *agn43-lacZ* activity in Phase ON and Phase OFF stationary phase, test tube cultures were 123 (SD 13) MU and 21 (SD 3) MU, respectively. β-galactosidase activity in the Phase OFF culture results from a minority of cells in the population having switched to the Phase ON state.

Having established that the *agn43-lacZ* fusion was expressed in a phase variable manner in strain CSH50, the effect of an OxyR mutation on expression of the fusion was tested. This required that the plasmid be placed in a suitable genetic background i.e. a strain Locked ON for Ag43 expression but unable to express β-galactosidase. In order to create such a strain, the Δ*oxyR::kan* mutation from the Locked ON strain GSO5 (lac^+)

Figure 3.2 Levels of β-galactosidase activity in Phase ON and Phase OFF cultures of CSH50 (pAg43) and in the Locked ON strain CSH50*oxyR* (pAg43). Assays were conducted on 5 separate stationary phase cultures in duplicate. 3-ml cultures were grown in test tubes at 37°C with shaking. β-galactosidase activity is expressed in Miller units (MU).


250
200
150
100
50
0
β-galactosidase activity (Miller units)
Phase OFF
Phase ON
Locked ON

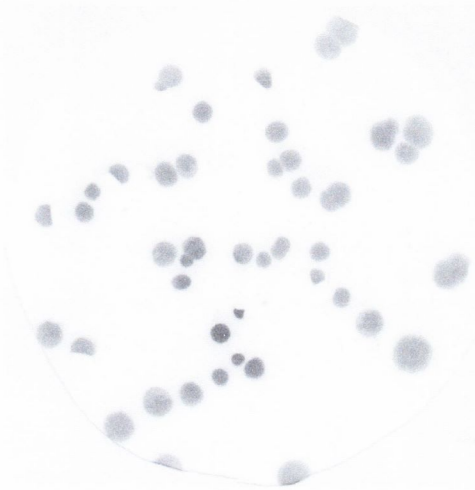
was introduced by P1*cml* transduction into the lac⁻ strain CSH50. The plasmid pAg43 was introduced into the resulting strain, CSH50*ΔoxyR* by transformation. The transformants gave rise exclusively to blue colonies on L agar containing X-Gal, indicating that expression of the *agn43-lacZ* fusion was Locked ON just as the chromosomal copy of the *agn43* gene was Locked ON. The level of *agn43* transcription in CSH50*ΔoxyR* (189 (SD 20) MU) was significantly higher than that in a Phase ON culture of CSH50 (123 (SD 11) MU) as shown in Figure 3.2. This is because all cells in the population of CSH50*ΔoxyR* were Locked ON for *agn43,* whereas a significant portion of the CSH50 Phase ON population had undergone switching to the Phase OFF state. These results confirmed that OxyR exerts its regulatory effect at the level of transcription (Henderson and Owen, 1999; Haagmans and Van der Woude, 2000). Thus it seems that, in all important respects, the artificial system in plasmid pAg43 is a faithful representation of the phase variable control of the native *agn43* gene.

3.2.3 Coordinated expression of chromosomally-encoded Antigen 43 with the *agn43-lacZ* fusion

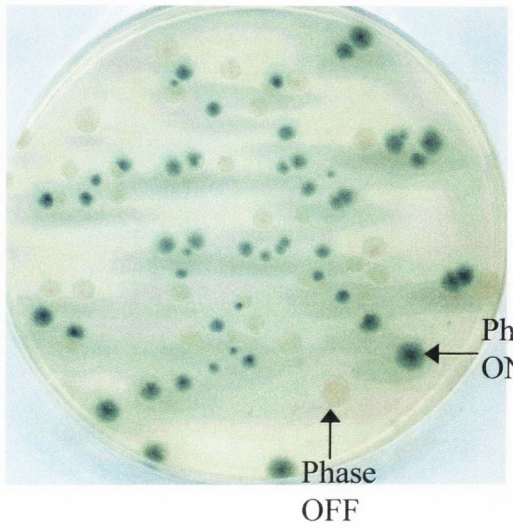
It was decided to determine whether there existed any correlation between transcription of the plasmid-borne *agn43-lacZ* fusion, and expression of chromosomally encoded Antigen 43 protein. Colonies of CSH50 (pAg43) were grown on L agar containing X-Gal, which allowed *agn43-lacZ* transcription to be monitored. Western colony immunoblotting, using antibodies specific for the alpha subunit of Ag43, was then conducted to detect surface expression of the chromosomally encoded Ag43 protein. Colony immunoblotting of CSH50 (pAg43) revealed an approximately 95% correlation (Based on screening of ~5000 colonies) between surface expression of chromosomally-encoded Ag43 and transcription of the multicopy *agn43-lacZ* fusion (Figure 3.3). This

Figure 3.3 The correlation between surface expression of Ag43 and transcription of the *agn43-lacZ* fusion. The phase-variable strains CSH50 (pAg43) and CSH50*ΔoxyR* (pAg43) [A and B, respectively] were analysed by Western colony immunoblotting using antibodies specific for the α^{43} subunit of Ag43 (A1 and B1, respectively) and by growth on L agar supplemented with X-Gal (A2 and B2, respectively). Note that Lac^- colonies in A2 (arrowed) are negative for Ag43 expression (A1) and that Lac^+ colonies of CSH50*ΔoxyR* (pAg43) and CSH50 (pAg43) express Ag43.

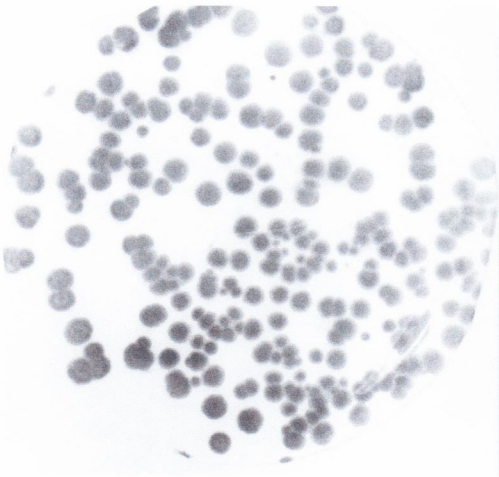
A1



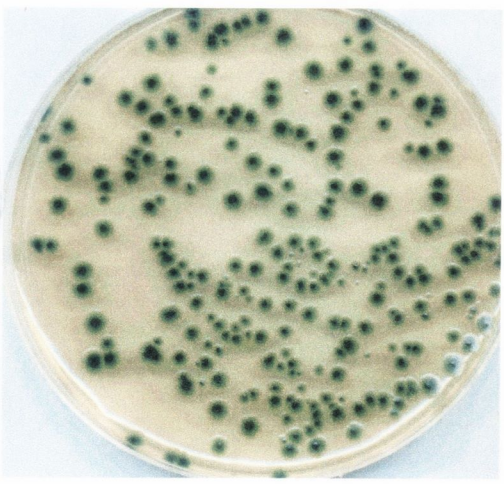
A2



B1



B2



correlation was also evident upon examination of colony morphology, i.e. blue colonies displayed the Flu^- phenotype (flat, rough) while white colonies were Flu^+ (raised, glossy). Furthermore, overnight broth cultures derived from blue colony inocula sedimented much more rapidly than did cultures derived from white colonies, consistent with autoaggregation caused by Ag43 expression in blue colonies (Data not shown).

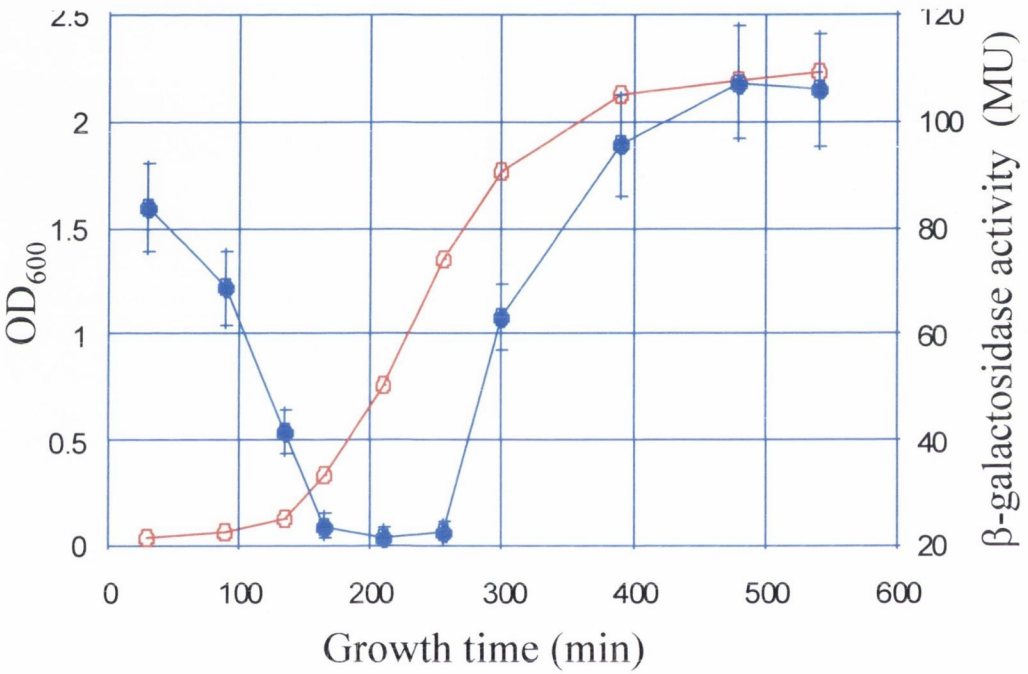
3.2.4 The effect of environmental conditions on *agn43* promoter function

A series of experiments was conducted to examine the effects of environmental parameters on the regulation of *agn43* expression. In these experiments, β-galactosidase assays were used to monitor activity of the *agn43* promoter fusion in the Locked ON background CSH50*ΔoxyR* under a range of growth conditions. A Locked ON strain was used so that *agn43* promoter activity could be monitored in the absence of Phase switching.

3.2.4.1 The effect of growth phase on *agn43* promoter activity

β-galactosidase activity was first assayed at different stages of bacterial growth in L broth. An approximately 4 to 5-fold increase in *agn43-lacZ* expression was observed in stationary phase compared to log phase of growth as shown in Figure 3.4. β-galactosidase expression was also found to be high in the early stages of bacterial growth due to high levels of *agn43-lacZ* transcription in the stationary phase (1%) inoculum. These results indicate that the *agn43* promoter is subject to growth-phase dependent regulatory mechanisms.

Figure 3.4 The effect of growth phase on *agn43-lacZ* activity in CSH50*ΔoxyR* containing pAg43. 100-ml cultures, inoculated with 1 ml of overnight culture, were grown at 37°C with aeration in 2 L baffled flasks. β-galactosidase assays were performed in triplicate on samples taken at different stages of growth. Results plotted (in blue) are mean activities expressed in Miller units. Bacterial growth is plotted in red.


2.5
2
1.5
1
0.5
0
OD600
120
100
80
60
40
20
β-galactosidase activity (MU)
0
100
200
300
400
500
600
Growth time (min)

3.2.4.2 The effect of temperature on *agn43* promoter activity

To determine the effect of growth temperature on *agn43* promoter activity, cultures derived from the same inoculum were grown at 37°C or 28°C and β-galactosidase activity was monitored during the subsequent growth cycle. A significant increase in *agn43* promoter activity, which varied between 3 and 5 fold depending on the density of cells, was apparent when cells were grown at 37°C compared to 28°C (See Figure 3.5). This is in good agreement with the temperature effect previously observed by Western blot analysis of whole cell lysates with anti-α^{43} antibodies (Meehan, 1995).

3.2.4.3 The effect of osmolarity and pH on *agn43* promoter activity

Changes in osmolarity and pH have been found to influence the transcription of a number of virulence genes (Mahan *et al.*, 1996; White-Zeigler *et al.*, 2000). For this reason the effect of these environmental stimuli on *agn43* transcription was examined. To assess the effect of osmolarity on *agn43* promoter activity, cells were grown in NaCl-free L broth and in L broth supplemented with 0.1 M, 0.3M or 0.6M NaCl. Osmolarity was observed to have no significant effect on *agn43-lacZ* expression (see Table 3.1). To assess the effect of pH, Cells were grown to stationary phase in L broth adjusted to an initial pH of 5, 6 7 or 8. β-galactosidase assays indicated that *agn43* promoter function is largely unaffected by pH (Table 3.2).

3.2.4.4 The effect of anaerobiosis on *agn43* promoter activity

To assess the effect of anaerobiosis on *agn43* promoter activity, cultures derived from a common inoculum of CSH50*ΔoxyR* (pAg43) were grown under aerobic and semi-anaerobic conditions. To create aerobic conditions, 100-ml cultures were grown in 2 l baffled flasks at 37°C with shaking. To mimic a semi-anaerobic environment, 20-ml

Figure 3.5 The effect of growth temperature on *agn43-lacZ* activity in CSH50*ΔoxyR* containing pAg43. 100-ml cultures were grown at 37°C, or at 28°C, with aeration in 2 l baffled flasks. β-galactosidase assays were performed in triplicate on samples taken at different stages of growth. Results plotted are mean activities expressed in Miller units.

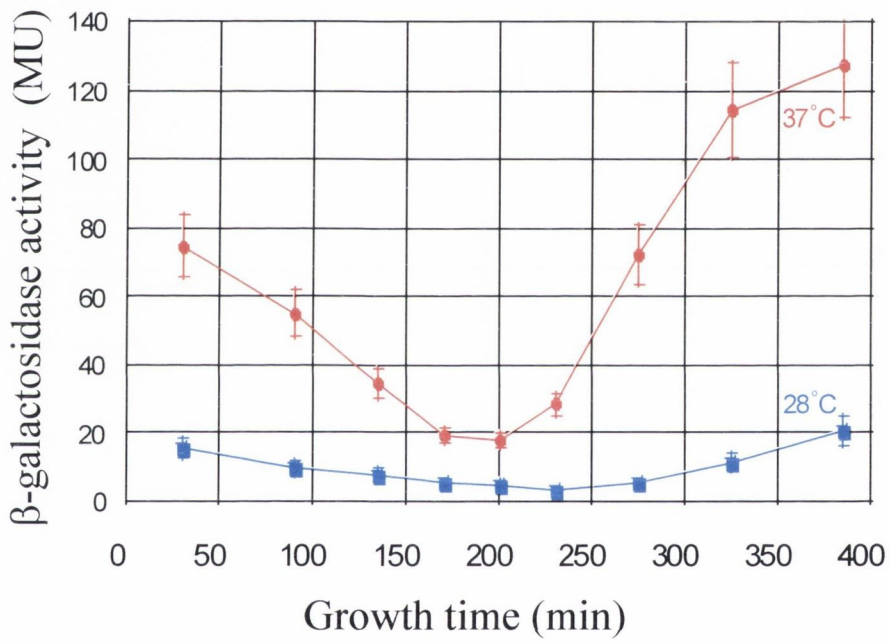

β-galactosidase activity (MU)
140
120
100
80
60
40
20
0
0
50
100
150
200
250
300
350
400
Growth time (min)
37°C
28°C

Table 3.1 The effect of osmolarity on *agn43-lacZ* expression in the Locked ON strain CSH50*ΔoxyR*. (pAg43).

Concentration NaCl in L broth (M)	β-galactosidase activity [a] (Miller Units)
0.0	117(10)
0.1	122(13)
0.3	131(16)
0.6	110(15)

[a]Assays were performed on duplicate stationary phase cultures, in duplicate. Results shown are averages with standard deviations in parenthesis.

Table 3.2 The effect of pH on *agn43-lacZ* expression in the Locked ON strain CSH50*ΔoxyR* (pAg43).

pH	β-galactosidase activity [a] (Miller Units)
5	130(15)
6	117(9)
7	105(14)
8	121(16)

[a]Assays were performed on duplicate stationary phase cultures, in duplicate. Results shown are averages with standard deviations in parenthesis.

Table 3.3 The effect of aerobic and semi-anaerobic growth conditions on *agn43-lacZ* expression in CSH50*ΔoxyR* containing pAg43.

Growth conditions	β-galactosidase activity[a] (Miller Units)
Aerobic[b]	127(25)
Semi-anaerobic[c]	292(35)

[a]Assays were conducted on 5 separate stationary phase cultures in duplicate.[b] 100-ml cultures were grown in 2-l baffled flasks at 37°C with shaking. [c]20-ml cultures were grown in sealed test tubes at 37°C without shaking. Results shown are averages with standard deviations in parenthesis.

cultures were grown in sealed test tubes at 37°C without shaking. β-galactosidase expression was compared when the cultures were in stationary phase rather than when at the same OD_{600} so as to ensure that any effects were due to anaerobiosis and not to differences in growth rate. Semi-anaerobic stress resulted in a 2-2.5-fold increase in *agn43-lacZ* expression compared to aerobic growth conditions (Table 3.3).

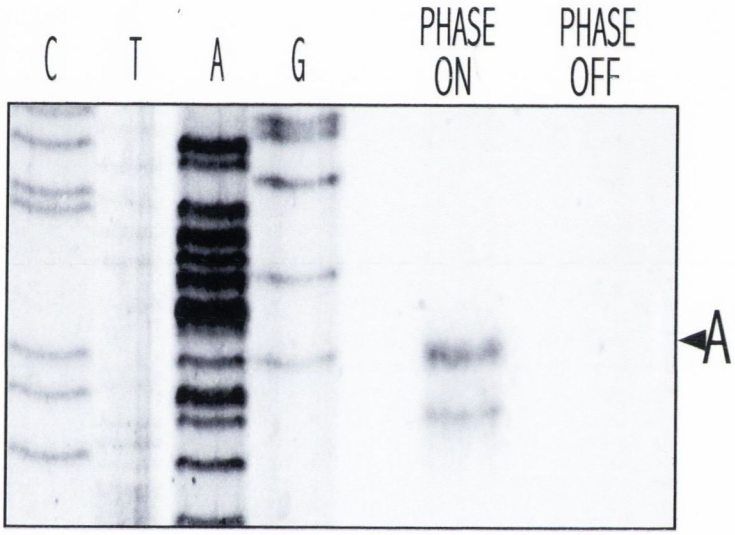
3.2.5 Mapping of the transcriptional start site of *agn43*

Primer extension analysis was conducted in order to locate the transcriptional start site of the *agn43 gene*. Phase ON and from Phase OFF cultures of CSH50 carrying the plasmid pAg43 were selected based on colony colour on X-Gal. RNA was then isolated from these cultures and primer extensions were performed (see Figure 3.6). Two primer extension products were detected with RNA from a culture Phase ON for *agn43-lacZ* expression. The absence of these products when using RNA from a Phase OFF culture identified them as cDNA copies of *agn43* mRNA. Band A in Figure 3.6, corresponds to a transcript initiated at the G of the first 5′-GATC-3′ site (i.e. 5'-GATC-3' site 1 in Figure 3.1) and is taken to be the true start site. The lower band is probably caused by the reverse transcriptase 'falling off' the RNA prematurely.

3.2.6 Promoter analysis

The consensus sequences for the –10 and –35 boxes of *E. coli* σ^{70}-dependent promoters are TATAAT and TTCTTG, respectively. When the region upstream of the transcription start site was examined for possible promoter sequences, candidates for the –10 and –35 hexamers were identified. The –10 candidate TAGAAT, was located 6 bp upstream of the start site (+1) and is separated from the –35 candidate, TTGTCC, by a spacer of 17 bp (Figure 3.7). The putative –35 box was disrupted by site directed mutagenesis in which a *Bsa*HI recognition site was introduced, altering the sequence

Figure 3.6 Primer extension analysis of RNA isolated from Phase OFF and Phase ON cultures of CSH50(pAg43). A sequencing ladder generated using the same primer is alongside. Band A corresponds to the likely transcriptional start site of *agn43*.


C
T
A
G
PHASE ON
PHASE OFF
A

from CTTGTCC to GACGTCC. Digestion with the enzyme *Bsa*HI was used to identify a plasmid, named pDEW1, carrying this mutation (Figure 3.8). Introduction of the desired mutation shortened the size of a ~1100-bp DNA fragment, created by *Bsa*HI digestion, due to the presence of the additional *Bsa*HI restriction site. Sequencing confirmed the presence of this mutation and additionally revealed that a 4th base pair substitution had unintentionally been introduced in the spacer region between the -35 and -10 sites (see Figure 3.7). *agn43-lacZ* transcription from the mutated plasmid was abolished as indicated by the exclusively white colour of transformed colonies on lactose indicator plates in both phase variable (CSH50), and Locked ON (CSH50Δ*oxyR*), backgrounds (data not shown). Primer extensions were then conducted to confirm the effect of this mutation. RNA was isolated from Phase ON and Phase OFF cultures of CSH50 (pAg43) and from cultures of CSH50Δ*oxyR* containing either the wildtype plasmid pAg43 or the mutated plasmid pDEW1 (Figure 3.9). As expected *agn43* transcripts were detected for RNA isolated from Phase ON CSH50 (pAg43) and from Locked ON CSH50Δ*oxyR* (pAg43) but not from Phase OFF CSH50 (pAg43). Importantly however, no *agn43* transcript was detectable when RNA was isolated from an OxyR$^-$ strain carrying the mutated plasmid, despite the fact that this strain is Locked ON for chromosomally encoded Ag43 expression. These results place the *agn43* promoter upstream of the three 5′-GATC-3′ sites contained within the proposed binding site for OxyR.

Figure 3.7 Sequence of the *agn43* promoter region. The –10 and –35 RNA polymerase contact sites are underlined and labelled. The transcription start site is indicated by an angled arrow. The four base substitutions of pDEW1 are indicated by red arrows. The –10 and –35 sequences previously suggested by Henderson and Owen (Henderson and Owen, 1999) are in green.


GAC
T
+1
TCTCTCTTGTCCGTGCAATAGCTCAATAATAGAATAAAACGATCAATATCTATTTTATCGATCG
-35
-10
1
2
TTTATATCGATCGATAAGCTAATAATAACCTTTGTCAGTAACATGCACGATACGTACAGAAA
3

Figure 3.8

A. *Bsa*HI restriction digests of the plasmids pAg43 and pDEW1, which contains a 3-bp substitution in the suspected –35 region. Sizes (in bp) are indicated on the DNA size standard alongside.

B. Sequencing of the promoter regions from pDEW1 and control pAg43. The CTT→TGG promoter mutation in pDEW1, which was introduced by site directed mutagenesis, is indicated by a red arrow. Also indicated by a red arrow is a C→T mutation, which was unintentionally introduced to the promoter spacer region.

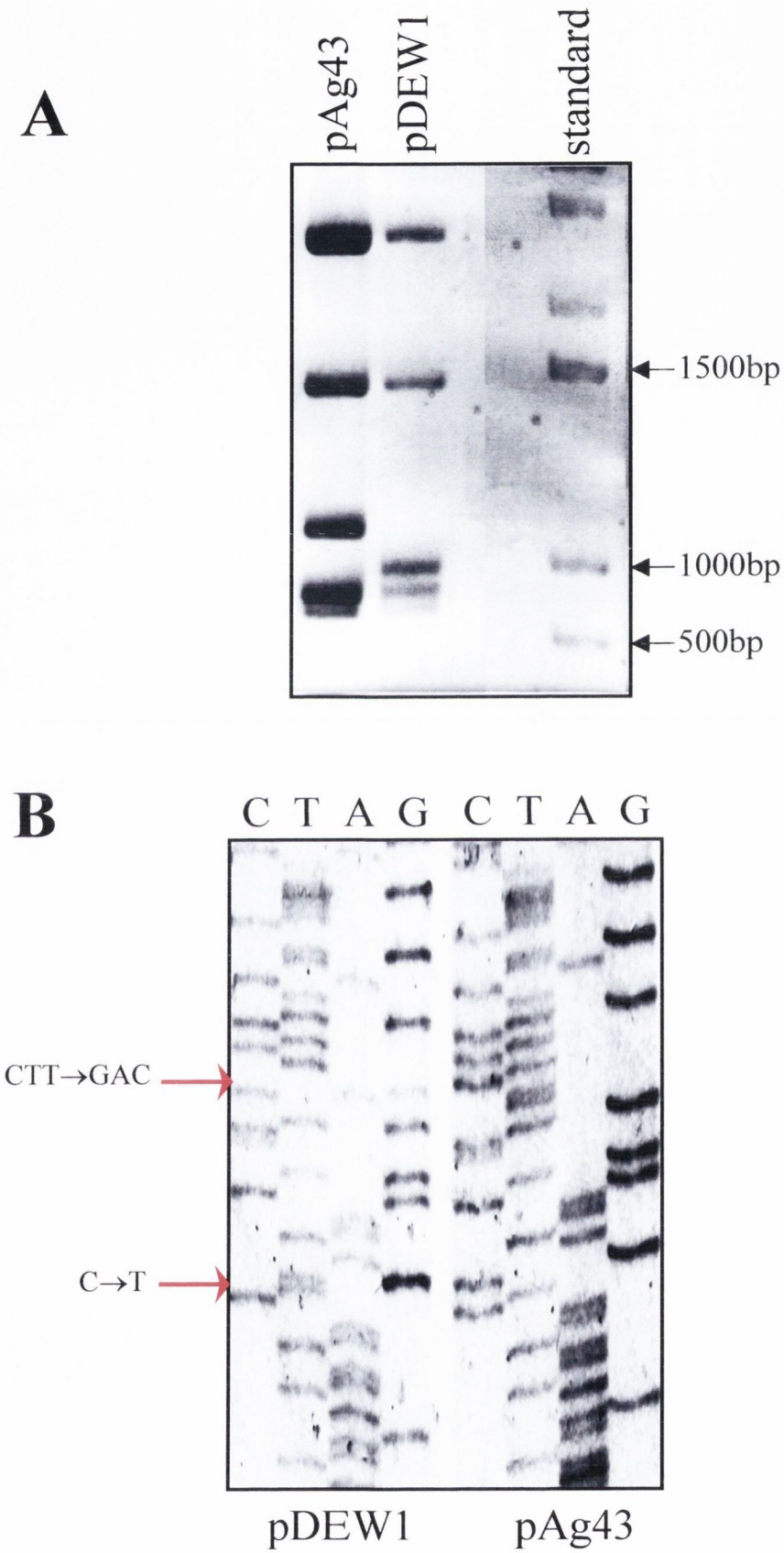

A
pAg43
pDEW1
standard
1500bp
1000bp
500bp
B
C T A G C T A G
CTT→GAC
C→T
pDEW1
pAg43

Figure 3.9 Primer extension analysis of RNA isolated from Phase OFF (lane 1) and Phase ON (lane 2) cells of CSH50 (pAg43) and from cells of CSH50*ΔoxyR* (pAg43) with (lane 3) and without (lane 4) a 4-bp substitution in the suspected *agn43* promoter region. A sequencing ladder generated using the same primer is alongside.

G A T C 1 2 3 4
CSH50 (pAg43) Phase OFF
CSH50 (pAg43) Phase ON
CSH50 ΔoxyR (pDEW1)
CSH50 ΔoxyR (pAg43)

3.2.7 Methylation state of the *agn43* promoter region in Phase ON and Phase OFF cells

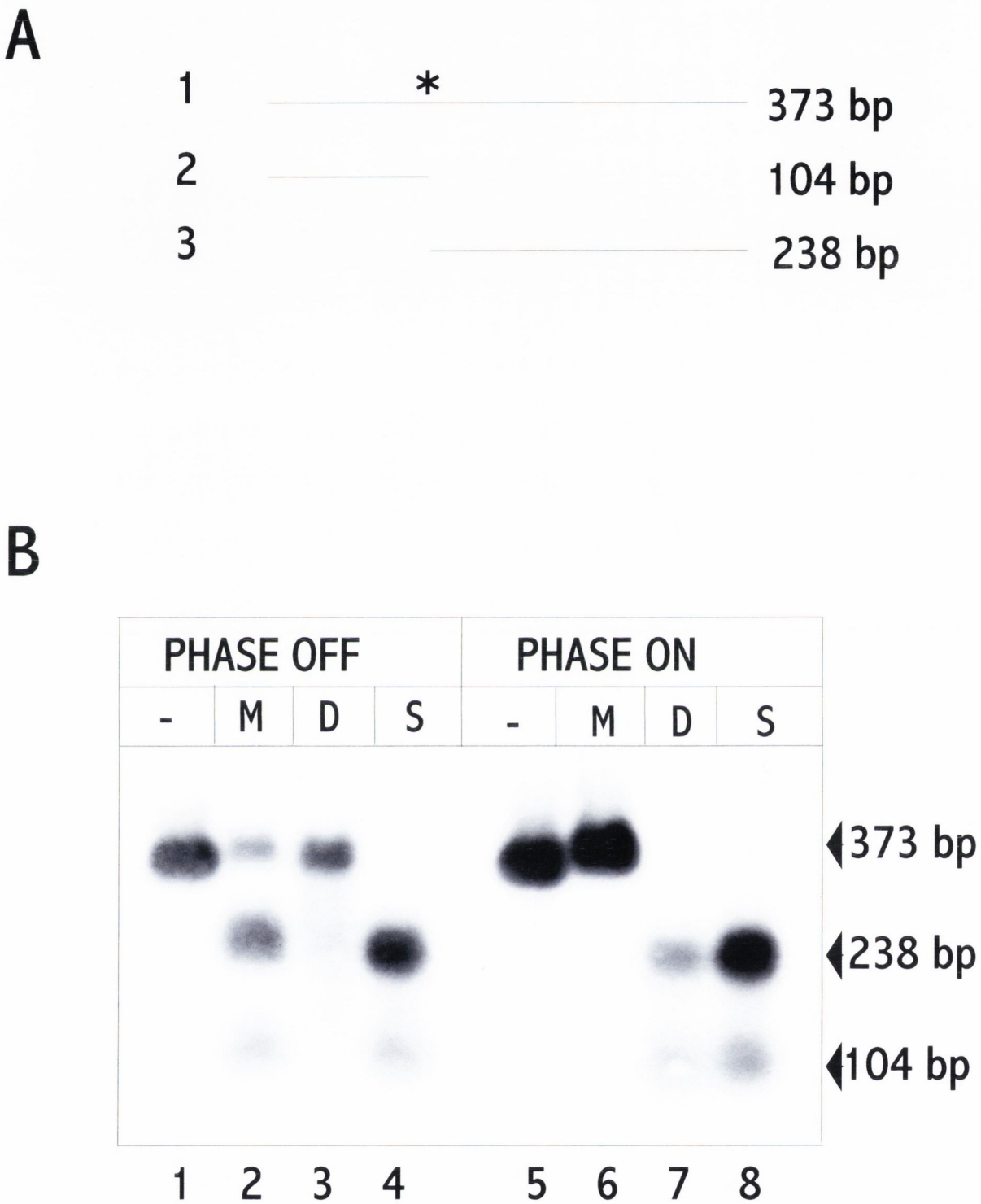
Since the *agn43* gene is subject to regulation by Dam, it was decided to determine whether the 5′-GATC-3′ sites of its promoter were differentially methylated in Phase ON and Phase OFF cultures. Plasmid DNA from Phase ON and from Phase OFF cultures of CSH50 (pAg43), was first digested with *Bam*HI and *Hin*dIII yielding 373-bp fragments containing the regulatory region of the *agn43-lacZ* fusion. The methylation state of the *agn43* 5'-GATC-3' sites was then determined by assessing their sensitivity to digestion with the enzymes *Mbo*I, *Dpn*I or *Sau*3A. The enzyme *Sau*3A cuts at all 5′-GATC-3′ sites regardless of their methylation state, whereas *Dpn*I and *Mbo*I cut only when the DNA is fully methylated and fully unmethylated, respectively. A DNA probe, which was generated using the primers Ag43R and Ag43F (Table 2.2) and spanned the entire cloned *agn43* regulatory region was used in Southern blot analysis of the restricted fragments (see Figure 3.10). Any cleavage with *Mbo*I, *Dpn*I or *Sau*3A should give rise to just two visible bands, within the size range tested, since the three target 5'-GATC-3' sites are in close proximity Figure 3.10A). DNA from the Phase OFF culture was resistant to *Dpn*I cleavage, indicating that none of the 5′-GATC-3′ sites were fully methylated. Cleavage with *Mbo*I gave rise to two hybridising fragments indicating that all three sites were hemimethylated or unmethylated with at least one of the 5′-GATC-3′ sites being completely unmethylated. DNA from a Phase ON culture was resistant to *Mbo*I cleavage indicating that none of the three 5′-GATC-3′ sites was fully unmethylated. Cleavage of Phase ON DNA by *Dpn*I indicated that all three 5′-GATC-3′ sites were methylated and at least one site was fully methylated. It is not possible to distinguish between cleavage at the individual 5′-GATC-3′ sites, due to

their close proximity. However, this experiment does show that the presence of nonmethylated, and fully methylated, 5′-GATC-3′ sites correlates with the Phase OFF and Phase ON transcriptional states of the *agn43-lacZ* fusion, respectively.

Figure 3.10 Analysis of the *in vivo* methylation state of the 5′-GATC-3′ sites in the *agn43* regulatory region.

A. A schematic representation of the restriction fragments detected after digests. Frament 1 is the *Bam*HI-*Hin*dIII fragment encompassing the regulatory region of *agn43-lacZ*. The location of the 5'-GATC-3' sequences is denoted by an asterisk. Fragments 1 , 2 and 3 are 373 bp, 104 bp and 238 bp, respectively.

B. Southern blot of plasmid DNA from Phase OFF (lanes 1-4) and Phase ON (lanes 5-8) cultures of CSH50(pAg43) probed with the *agn43* regulatory region. Plasmid DNA was digested with *Bam*HI and *Hin*dIII and an additional third enzyme as indicated: M, *Mbo*I: D, *Dpn*I: S, *Sau3*A.


A
1
*
373 bp
2
104 bp
3
238 bp
B
PHASE OFF
PHASE ON
-
M
D
S
-
M
D
S
373 bp
238 bp
104 bp
1
2
3
4
5
6
7
8

3.3 Discussion

The Antigen 43 protein of *Escherichia coli* is expressed in a phase-variable manner, and this phase-variation has been shown to be dependent on the OxyR transcriptional regulator and on Dam. A simple model for regulation of Ag43 phase variation is one in which the OxyR protein binds to unmethylated 5′-GATC-3′ sequences in the region of the *agn43* promoter, and represses transcription by causing the exclusion of RNA polymerase. Methylation of these sites would block binding of OxyR, thereby allowing RNA polymerase to bind, and *agn43* transcription to proceed. Phase switching would result from competition between OxyR and Dam for newly synthesized, unmethylated DNA.

In order to test the efficacy of this model, a plasmid borne transcriptional fusion of the *agn43* regulatory region to *lacZ* was constructed. Phase variable expression of β-galactosidase, under the control of the *agn43* promoter, resulted in the growth of Lac^+ and Lac^- colonies. In 95% of colonies, the Lac^+ phenotype was observed to correlate with surface expression of chromosomally encoded Ag43, indicative of co-expression of the multicopy *agn43-lacZ* fusion and the chromosomal copy of *agn43*. Co-regulation of multiple copies of *agn43* in the cell suggests that *agn43* phase switching is linked to the physiological state of the cell, and is not simply a random event dependent on the outcome of competition between Dam and OxyR for DNA. (Simple competition would result in colonies with uniform β-galactosidase expression, the level of which would be dependent on the relative affinities of Dam and OxyR for the DNA.). The molecular basis of the synchrony between phase switching on plasmids and on the chromosome may be related to synchrony of DNA replication and the associated generation of hemimethylated

DNA. There have been reports that plasmids of the ColE1 family (of which pQF50 is a member) can influence the cell cycle through a *trans*-acting small RNA molecule (Sharpe *et al*, 1999).

This work shows that Ag43 is not only subject to phase variation but also to a second type of regulation, that of the *agn43* promoter in response to environmental conditions. It appears that the *agn43* promoter is subject to some growth phase dependent regulatory mechanism, since expression of β-galactosidase, under the control of the *agn43* promoter, was strongly induced in the stationary phase of growth. In addition *agn43* promoter activity was approximately 4 fold greater when cells were grown at 37°C, the internal temperature of most mammalian hosts, compared to 28°C. Temperature has been shown to be an important regulator of virulence gene expression in several genera of bacteria (Mekalanos, 1992). Thermoregulation of the phase variable, Dam-regulated, *pap* and *daa* operons is mediated by H-NS (White-Zeigler *et al.*, 2000). Although Ag43 expression has previously been found to be unaffected by an H-NS mutation, the effect of such a mutation on expression at low temperatures has not yet been investigated. Environmental changes can affect the expression of genes via alterations in global DNA superhelicity. However, it is unlikely that the *agn43* promoter is sensitive to changes in the level of supercoiling since its activity was unaffected by substantial changes in osmolarity, which should affect the degree of global supercoiling. Promoter activity was elevated 2-2.5 fold when cells were grown under semi anaerobic, compared to aerobic, conditions. The bacteria may respond to oxygen deprivation in the internal environment of the host intestine by increasing expression of Ag43. It is likely that the bacterium uses the environmental signals to

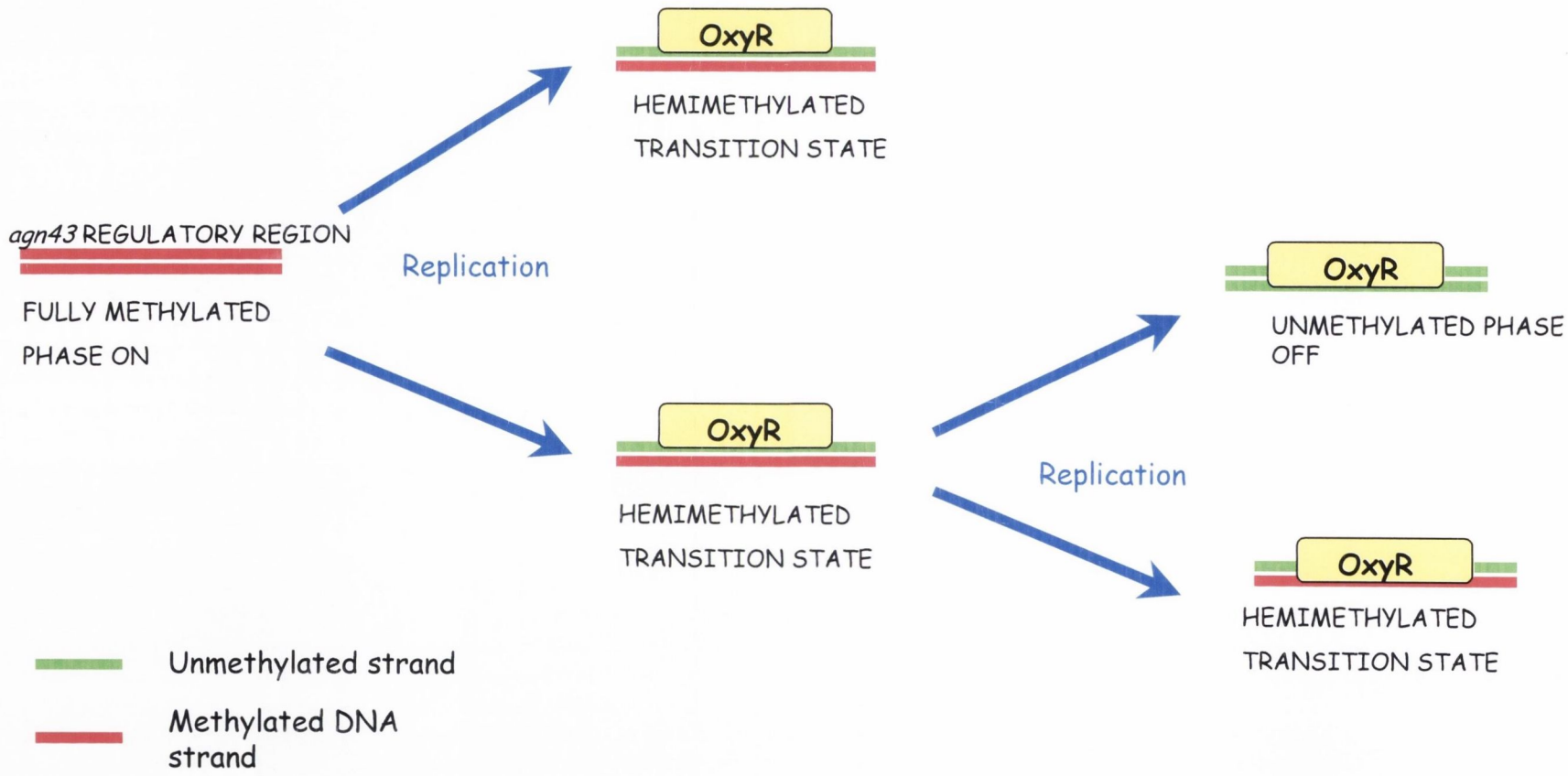
determine whether it is within a host so that, by regulating virulence gene expression accordingly, it can more efficiently utilize its resources.

The *agn43* transcriptional start site and promoter were located upstream of three 5'-GATC-3' sites, which lie within the proposed OxyR binding site. Thus the mechanism of *agn43* regulation differs greatly from that of the *mom* gene (see Section 1.5.4) despite the high level of promoter sequence similarity. The -10 site of the *agn43* promoter lies within the region that OxyR has recently been reported to protect from DNase1 digestion (Correnti *et al.*, 2002). It is therefore likely that bound OxyR exerts its regulatory effect by blocking binding of RNA polymerase to the *agn43* promoter. Methylation of 5'-GATC-3' sites would prevent OxyR from binding, leaving the *agn43* promoter accessible to RNA polymerase. Under conditions of oxidative stress OxyR becomes activated by the reversible formation of a disulphide bond (Zheng *et al.* 1998). DNase1 footprints of oxidized OxyR are reportedly shorter than those of reduced OxyR and as such OxyR may not overlap with the *agn43* promoter when in the oxidized state. The protein is thought to be capable of shifting its DNA contacts whilst remaining bound, as it is activated in response to oxidative stress, or deactivated (Toledano *et al.* 1994). Thus upon oxidation OxyR might be expected to vacate the *agn43* promoter resulting in a switch to the Phase On state of transcription. Whether OxyR represses *agn43* by blocking the binding of RNA polymerase to the promoter, or by blocking the movement of bound RNA polymerase along the DNA, remains to be determined.

Haagmans and van der Woude (2000) previously demonstrated a correlation between the Phase ON and Phase OFF states of a chromosomal *agn43-lacZ* fusion

with the presence of methylated, and nonmethylated, 5'-GATC-3' sequences, respectively (Haagmans and van der Woude, 2000). This correlation also exists for the plasmid borne *agn43-lacZ* fusion constructed in this study. The fully unmethylated state of the pAg43 DNA prepared from Phase OFF cultures could arise as follows; OxyR (or some other methylation-blocking factor) would bind to newly synthesised hemimethylated DNA preventing it from being fully methylated by Dam. Passage of the replication fork during a subsequent round of replication would result in the formation of fully unmethylated DNA to which the protein would remain bound, and of hemimethylated DNA to which another copy of OxyR could bind, etc. (Figure 3.11). Indeed, it has recently been shown that OxyR binding can protect all three 5′-GATC-3′ sites from methylation *in vitro*, and that hemimethylated *agn43* DNA is bound by OxyR, albeit with lower affinity than unmethylated DNA (Correnti *et al.*, 2002). We are unable to distinguish between individual 5′-GATC-3′ sites by means of Southern blot, due to their close proximity. However, the role of individual 5′-GATC-3′ sites and of different methylation patterns has been examined in this study and will be discussed in Chapter 5.

Figure 3.11 Model for the formation of the different DNA methylation patterns corresponding to the Phase ON and Phase OFF states of *agn43* transcription. A green line represents an unmethylated strand of the *agn43* regulatory region. A red line represents a methylated strand of the *agn43* regulatory region.


OxyR
HEMIMETHYLATED
TRANSITION STATE
agn43 REGULATORY REGION
Replication
FULLY METHYLATED
PHASE ON
OxyR
UNMETHYLATED PHASE
OFF
OxyR
HEMIMETHYLATED
TRANSITION STATE
Replication
OxyR
HEMIMETHYLATED
TRANSITION STATE
Unmethylated strand
Methylated DNA
strand

Chapter 4

The identification of factors affecting *agn43* expression

4.1 Introduction

It has recently been shown that the presence of both Dam and OxyR is sufficient for *agn43* to be expressed phase-variably *in vitro* (Wallecha *et al.*, 2002). This fact, however, does not discount the possibility that other factors are involved in regulating *agn43* transcription *in vivo*. Results described in Chapter 1 showed that there are two distinct levels of transcriptional regulation since, in the absence of phase variation, the *agn43* promoter was highly responsive to environmental cues. Thus, it seemed possible that additional unknown regulators might affect the actual level of *agn43* transcription in cells expressing Ag43. Henderson and Owen (1999) showed, by, colony immunoblotting, that Ag43 was expressed phase variably by a number of strains carrying mutations in genes encoding different global regulators. The strains tested contained deletions in *hns, crp, cya, gyrA, gyrB, fis, topA* and *rpoS*. At first sight this would appear to rule out involvement of these genes in *agn43* expression. However, although these strains could still express Ag43 there may have been subtle alterations in the level of transcription, and hence protein expression, which were not detectable by the assay methods used.

The extensive nucleotide sequence similarities between the regulatory regions of *agn43* and the *mom* gene of phage Mu were also suggestive of the involvement of an additional factor in *agn43* regulation (Henderson and Owen, 1999). Like *agn43*, the *mom* gene is regulated by OxyR and Dam methylation. However, regulation of *mom* also involves an accessory factor MuC, which activates transcription by helping to melt DNA in the region of the poor *mom* promoter. The sequence comprising the MuC binding site is highly conserved in the regulatory region of *agn43* (see Section 1.5.4; Figure 1.7). This similarity raises the possibility of the

involvement of an additional analogous protein, which acts at this conserved region, in the regulation of *agn43*. However, results reported in Chapter 1 showed that regulation of *mom* and *agn43* differ in that, the *agn43* promoter lies upstream rather than downstream of three 5'-GATC-3' sequences. This fundamental difference makes it seem unlikely that any such protein would have a role analogous to that of MuC i.e. the unwinding of DNA in the promoter region.

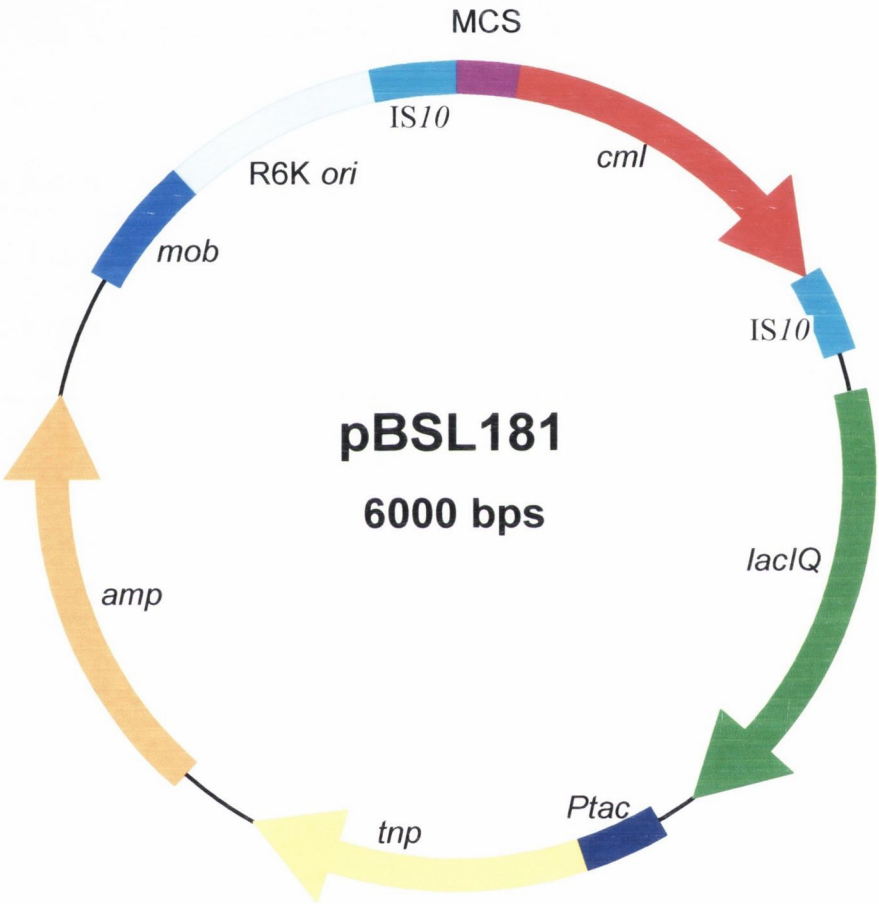
The stationary phase induction of *agn43-lacZ* expression described in Chapter 3 suggested that *agn43* might be subject to regulation by the alternative sigma factor RpoS, which is encoded by the *rpoS* gene. The possibility of RpoS involvement in *agn43* expression is explored in this chapter. Another protein with a possible role in *agn43* regulation is Lrp, which acts as a methylation-blocking factor in the Dam-dependent regulation of a number of other genes (see Section 1.6). The possibility that Lrp may be involved in *agn43* expression is therefore explored here as well. In addition, transposon mutagenesis is employed in an attempt to identify other unknown factors involved in the regulation of *agn43*. This experimental approach is outlined in greater detail below.

4.1.1 pBSL181

Tn*10*-based minitransposons are a useful tool for insertion mutagenesis of the bacterial chromosome. In this study the suicide vector, pBSL181 (Alexeyev and Shokolenko, 1995), which contains a mini-Tn*10* derivative, is employed in an attempt to identify unknown factors that affect *agn43* transcription (Figure 4.1). The suicidal properties of this vector are determined by the γ-*ori* of the R6K plasmid, whose activity is restricted to those *E. coli* strains, which can provide the π

protein, in *trans*. pBSL181 contains a *cml* gene (encoding chloramphenicol resistance) which is flanked by 110-bp inverted repeats consisting of the outermost 70-bp of IS*10* right embedded in 40 bp of the λ *cI* gene. The plasmid also contains the Tn*10* transposase-encoding gene, *tnp* with two mutations, *ats1* and *ats2*, under the control of the P_{tac} promoter. Induction of the *tnp* gene with IPTG results in the transposition of the *cml* gene and flanking IS*10* sequences into the chromosome of a transformed cell. The two mutations in *tnp* drastically reduce the 'hot spot' specificity of the wildtype Tn*10* transposase, such that transposition could be considered random (Kleckner *et al.*, 1991). pBSL181 contains a 300-bp region from the RP4 plasmid which allows conjugative transfer of the vector into a wide range of Gram-negative bacteria. The plasmid also contains the multiple cloning site of pBluescriptII to facilitate the chromosomal insertion of cloned DNA fragments.

Figure 4.1. Map of the suicide plasmid pBSL181, which was used for mutagenesis in this study. R6K, *ori* of the R6K plasmid; *cml*, chloramphenicol acetyl transferase gene; *tnp*, Tn*10* transposase containing *ats1* and *ats2* mutations; IS*10*, 110-bp inverted repeats consisting of 70 bp of IS*10* right embedded in 40bp of the λ *cI* gene; MCS, Multiple cloning site of pBluescriptII; *mob*, 300-bp *mob* region from RP4; Amp, ampicillin resistance gene; P_{tac}, hybrid *trp-lac* P_{tac} promoter.


MCS
IS10
cml
R6K ori
mob
IS10
pBSL181
6000 bps
lacIQ
amp
Ptac
tnp

4.2 Results

4.2.1 The effect of an *rpoS* mutation on Ag43 expression

Experiments described in Chapter 3 revealed that expression of the *agn43* gene is induced in the stationary phase of growth. It was therefore decided to investigate whether the stationary phase sigma factor, RpoS, might be responsible for this growth phase-dependent regulation. Accordingly, *agn43-lacZ* expression was compared in two *E. coli* strains that differ only in their ability to express RpoS, namely ZK1000 and ZK126 (see Table 2.1). Strain ZK1000 does not express RpoS, because of a kanamycin insertion in its *rpoS* gene. The reporter plasmid pAg43 (see Figure 3.1) could be used to monitor *agn43* promoter activity in ZK1000 and ZK126 as both strains are Lac$^-$.

Upon transformation of these strains with the plasmid pAg43 it became apparent, from colony colour on X-Gal, that *agn43-lacZ* expression was significantly elevated in the *rpoS* mutant, ZK1000 (pAg43). β-galactosidase assays were conducted in order to quantitate the positive effect of the *rpoS* mutation on *agn43-lacZ* transcription. β-galactosidase activity was elevated approximately four-fold in both Phase ON and Phase OFF colonies of the *rpoS* mutant ZK1000 (pAg43) compared to wildtype ZK126 (pAg43) (see Figure 4.2).

To determine whether this effect could be detected at the level of protein expression immunoblotting was performed using antisera specific for the surface expressed α^{43} subunit of Ag43 (Figure 4.3). Dot immunoblots using doubling dilutions of cultures of strains ZK126 and ZK1000 revealed an approximately 8-fold increase in

Figure 4.2 *agn43-lacZ* expression in Phase ON and Phase OFF cultures of ZK126 (pAg43) and its *rpoS* derivative ZK1000 (pAg43). 5-ml test-tube cultures were grown in L broth to stationary phase at 37°C, with shaking. Duplicate β-galactosidase assays were performed, in triplicate. Results are expressed as Miller units (MU).

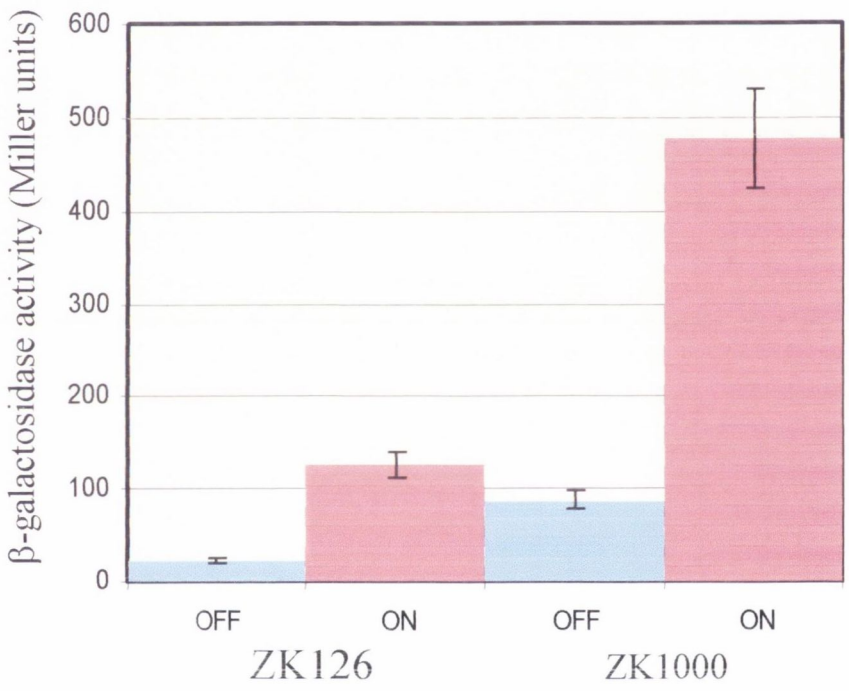
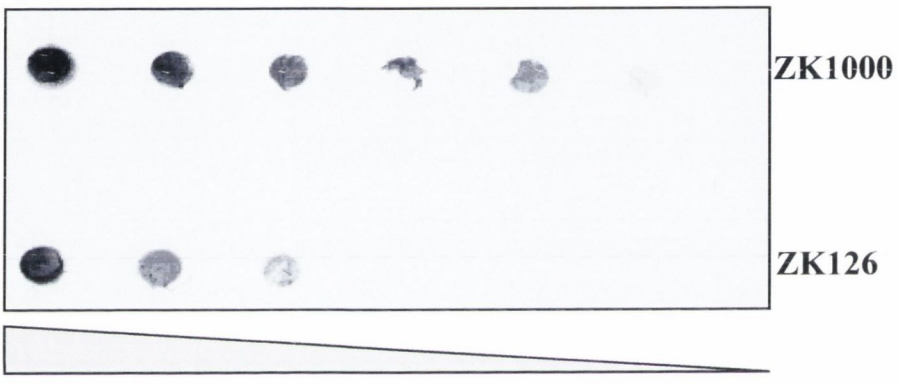

β-galactosidase activity (Miller units)
600
500
400
300
200
100
0
OFF
ON
OFF
ON
ZK126
ZK1000

Figure 4.3 Dot immunoblot analysis of Antigen 43 expression in the strains ZK126 and ZK1000. Anti-α^{43} antibodies were adsorbed to doubling dilutions of stationary phase cultures of the strains ZK126 ($RpoS^+$) and ZK1000 ($RpoS^-$).


ZK1000
ZK126
[cells]

expression of Ag43 in the *rpoS* mutant relative to wildtype, in reasonable agreement with the transcriptional effect seen using the reporter plasmid, pAg43.

4.2.2 Phase variation of *agn43* is altered in an *lrp* mutant

Lrp has been shown to act as a methylation-blocking factor in the Dam-dependent regulation of a number of other genes including *pap* (van der Woude and Low, 1994; see Section 1.6). The possibility of Lrp involvement in *agn43* regulation was therefore addressed by monitoring *agn43* expression in strains containing a mutation in the *lrp* gene. Previous studies involving immunoflouresence microscopy revealed the presence of Phase ON and Phase OFF cells in a strain containing a mutation in the *lrp* gene. This showed that Ag43 expression was still subject to phase variation and would tend to rule out a major role for this regulator in phase switching (Henderson and Owen, 1999). However, it is conceivable that this *lrp* mutation had a more subtle effect on *agn43* transcription, which went undetected by immunofluorescence microscopy. Accordingly, in the current study, the effect of an *lrp* mutation on *agn43* transcription was assessed.

agn43-lacZ expression was monitored in the isogenic Lac$^-$ strains CSH50 and CSH50*lrp* using the reporter plasmid, pAg43. Monitoring of *agn43-lacZ* switch frequencies in these strains revealed that OFF to ON switching in CSH50*lrp* was in the order of 2×10^{-5} per cell per generation (Table 4.1). This was extremely retarded compared to OFF to ON switching in the parent strain, CSH50, which was in the order of 1×10^{-3} per cell per generation. The *lrp* mutation had no significant effect on the rate of switching from Phase ON to Phase OFF. As shown in Table 4.1, the rates of ON to OFF switching were comparable in these strains.

To determine whether switching of Ag43 protein was also retarded in the *lrp* strain colony immunoblots were performed using anti-α^{43} antibodies. Phase OFF cultures of CSH50 and its *lrp* derivative were diluted and grown on L agar plates. Colony immunoblots revealed a significant difference, between strains, in the number of colonies having switched to the Phase ON state (Figure 4.4). OFF to ON switching of Ag43 protein was significantly retarded compared to wildtype in CSH50*lrp*, just as switching of the *agn43-lacZ* transcriptional fusion had been.

It has been reported that the expression of Type-1 fimbriae, and the net oxidation that results, has a negative effect on *agn43* transcription, which is thought to be mediated by reduction of OxyR (Schrembri and Klemm, 2001; see Section 1.5.5). Lrp is involved in phase variation of the phase-variable *fimA* operon and *fim* switching is retarded (in both directions) in strains containing mutations in the *lrp* gene. It was therefore possible that the observed alterations in *agn43* expression in the Lrp$^-$ mutant were an indirect effect resulting from the retardation of *fim* switching. To explore this possibility it was decided to examine the effect of an *lrp* mutation on *agn43-lacZ* expression in a strain unable to express Type 1 fimbriae.

The *lrp::*Tn*10* mutation from CSH50*lrp* was introduced into the Lac$^-$ *fim* deletion strain AAEC189, by P1*cml* transduction (Table 2.2). When transformed with the plasmid pAg43 the wildtype strain AAEC189 and its *lrp* derivative underwent phase variation of *agn43-lacZ* at comparable rates (Table 4.1). This indicated that the retardation of Ag43 switching seen earlier in CSH50*lrp* was dependent on the

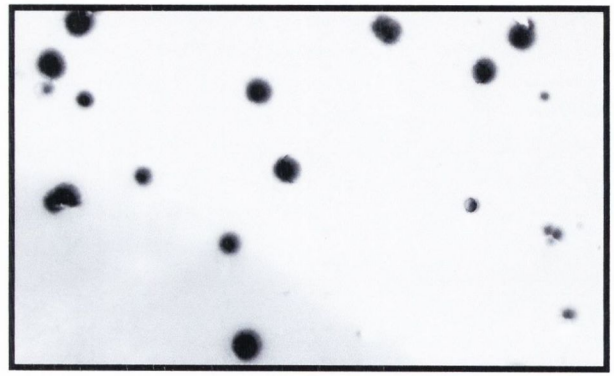
Table 4.1 Comparison of the *agn43-lacZ* switching rates of strains CSH50 and AAEC189 with those of their *lrp* derivatives.

Switch direction	CSH50	CSH50*lrp*	AAEC189	AAEC189*lrp*
ON → OFF	2.0×10^{-3}	3.8×10^{-3}	1.6×10^{-3}	1.9×10^{-3}
OFF → ON	1.4×10^{-3}	2.6×10^{-5}	2.0×10^{-3}	1.8×10^{-3}

Figure 4.4 Colony blot analysis of Ag43 expression in strains CSH50 and CSH50*lrp* using anti ~ α^{43} antibodies.. Cultures derived from Phase OFF colonies were diluted and plated on L agar. Plates containing similar numbers of colonies were used for immunoblotting.

CSH50*lrp*

CSH50



expression of Type 1 fimbriae. Thus it appears that Lrp has an indirect effect on *agn43* expression, which is mediated by *fim* probably through OxyR. Prolonged expression of fimbriae, due to retardation of ON to OFF switching in CSH50*lrp*, would be expected to result in the reduction of cellular OxyR. A resultant increase in abundance of the reduced (repressive) form of the protein could explain the retarded OFF to ON switching of Ag43, which was observed in CSH50*lrp*.

4.2.3 Generation, by transposon mutagenesis, of mutants with altered *agn43-lacZ* expression

Transposon mutagenesis was performed, using pBSL181 (Figure 4.1), in an attempt to identify unknown regulators of *agn43* expression. After transformation of the target strain (containing pAg43) with this vector, IPTG induction of the transposase gene results in transposition of the *cml* gene and flanking IS sequences into a random site (see Section 4.1.1). The suicide plasmid is then lost due to an inability to replicate in the absence of π protein. Mutagenized colonies were selected on L agar supplemented with chloramphenicol and X-Gal, the latter allowing the level of *agn43-lacZ* expression to be monitored. Insertion of the transposon into any non-essential chromosomal gene involved in *agn43* regulation should result in a detectable change in transcription of the plasmid-borne *agn43-lacZ* fusion.

4.2.3.1 Mutagenesis of strain CSH50

It was anticipated that phase variation of *agn43-lacZ* might cause some mutants with altered expression to remain undetected. In an attempt to avoid such screening problems, the Locked ON strain CSH50*ΔoxyR* (pAg43) was initially used as a target for mutagenesis. Approximately 10 000 mutagenized colonies were screened

for decreased *agn43-lacZ* expression on L agar containing X-Gal. Several colonies in which β-galactosidase expression was visibly Locked OFF, and one colony in which expression was down-regulated, were selected. Further analysis revealed that in all cases the Locked OFF phenotype resulted from disruption of the reporter plasmid by the transposon insertion. In the down-regulated mutant, the transposon had inserted into a chromosomal site. However, it became apparent that the resulting mutation had decreased β-galactosidase activity by altering the reporter plasmid's copy number and not by affecting *agn43-lacZ* transcription (data not shown). Since transposon mutagenesis of strain CSH50*ΔoxyR* did not generate any interesting mutants the wildtype strain CSH50 was subjected to mutagenesis.

It was possible that the mutation of any unknown gene affecting *agn43* transcription might only alter expression in an $OxyR^+$ phase-variable strain. The wildtype strain CSH50 (pAg43) was therefore mutagenised in the same manner as CSH50*ΔoxyR*. Unfortuntely, the phase-variable nature of *agn43-lacZ* expression in CSH50 had the potential to limit the effectiveness of colony screening. For example, increased *agn43-lacZ* transcription might not be detected if the corresponding mutation happened to arise in a Phase OFF colony. Furthermore, mutations decreasing the level of *agn43-lacZ* expression might be indistinguishable from Phase OFF colonies. Despite these limitations, CSH50 (containing pAg43) was subjected to transposon mutagenesis with pBSL181, and approximately 20,000 colonies were screened for visible alterations in *agn43-lacZ* expression. Colonies (seven in number) in which β-galactosidase expression was visibly increased were selected for further analysis. The switch phenotypes of these mutants (M1-M7) were assessed and the level of *agn43-lacZ* expression determined by β-galactosidase

assay. The data are presented in Table 4.2. In one of the mutants (M2) *agn43-lacZ* expression was subject to phase switching and in six mutants expression was Locked ON. Expression was most significantly elevated in the Locked ON mutants, M3 and M6, by approximately 9 and 7 fold respectively, compared to wildtype (Table 4.2).

4.2.3.2 Integrity of the reporter plasmid in mutants

To determine whether increased β-galactosidase expresssion was due to mutation of the *agn43-lacZ* fusion rather than a mutation in a regulatory gene, pAg43 DNA was prepared from each mutant and used to transform wildtype CSH50. Transformation of CSH50 with plasmid DNA from M1, M3 and M6 reproduced the mutant phenotype indicating that, in these cases, the reporter plasmid had been disrupted by a transposon insertion. The right IS element of the transposon from pBSL181 has a strong outward reading promoter (P^{out}). It is likely that the high level of expression in M1, M3 and M6 resulted from insertion of the transposon into *agn43-lacZ* such that the P^{out} promoter was driving transcription of *lacZ*. In view of this, these mutants were eliminated from further analyses.

The remaining four mutants (M2, M4, M5 and M7) were tested for chromosomal insertion of the transposon by PCR analysis using the primers CatF and CatR (Chapter 2, Table 2.2). Amplification using genomic DNA isolated from each of the above mutants resulted in a ~ 900-bp product corresponding to the *cml* gene (Figure 4.5A). No PCR product was obtained for genomic DNA that had been isolated from wildtype CSH50. These results indicated that the transposon had successfully inserted into the bacterial chromosome of M2, M4, M5 and M7.

Table 4.2 Switch phenotypes of CSH50 (pAg43) and of mutants generated by transposon mutagenesis.

Strain	Switch phenotype	β-galactosidase activity[a] (MU)
CSH50(pAg43)	Switching	130 (14)
M1	Locked ON	203 (21)
M2	Switching	396 (24)
M3	Locked ON	1098 (77)
M4	Locked ON	209 (19)
M5	Locked ON	189 (16)
M6	Locked ON	903 (53)
M7	Locked ON	198 (20)

[a] Assays were performed on duplicate cultures, in triplicate. Results shown are averages with standard deviation in parenthesis and are expressed in Miller Units (MU). In switching phenotypes, the β-galactosidase values refer to the Phase ON state.

4.2.3.3 PCR analysis of *oxyR* and *rpoS* genes in mutants

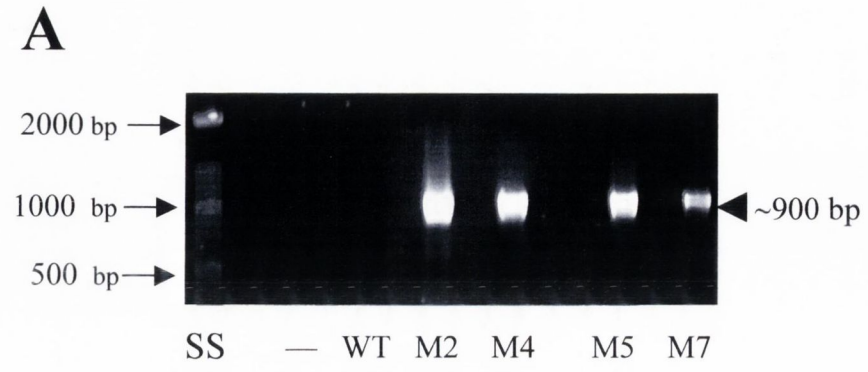
PCR analysis was performed in order to determine whether altered *agn43-lacZ* transcription resulted from mutations in genes encoding any known transcriptional repressors. Genomic DNA was used as a template for the amplification of the *oxyR* gene, using primers OxyR19F and OxyR19R, which anneal to each end of the *oxyR* open reading frame (Table 2.2). Agarose electrophoresis indicated that the *oxyR* genes of mutants M5 and M7 were approximately 1 kb longer than the wildtype ~ 900-bp gene (Figure 4.5B). The Locked ON phenotypes of M5 and M7 could therefore be attributed to the disruption of the *oxyR* gene by a transposon insertion.

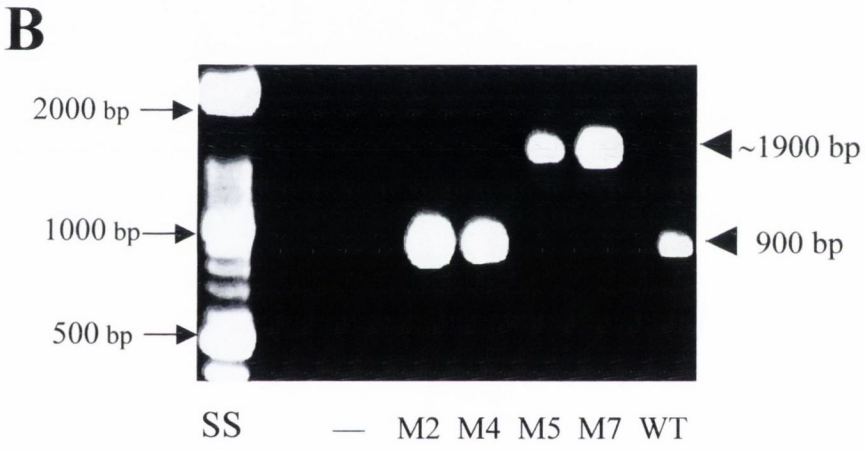
Previous results in this study had shown that Ag43 expression was elevated in an *rpoS* mutant (see Section 4.2.1). Amplification of the *rpoS* coding sequence from each mutant, using primers rposF and rposR (Chapter 2, Table 2.2), revealed that that of M2 had been disrupted by transposon insertion. As shown in Figure 4.5C, amplification of the *rpoS* gene from this mutant resulted in a PCR product approximately 1 kb greater in length than the wildtype 640-bp gene. The increase in *agn43-lacZ* transcription in this mutant could therefore be attributed to the presence of an *rpoS* mutation. As a result of these experiments M5, M7 and M2 were eliminated from further analyses.

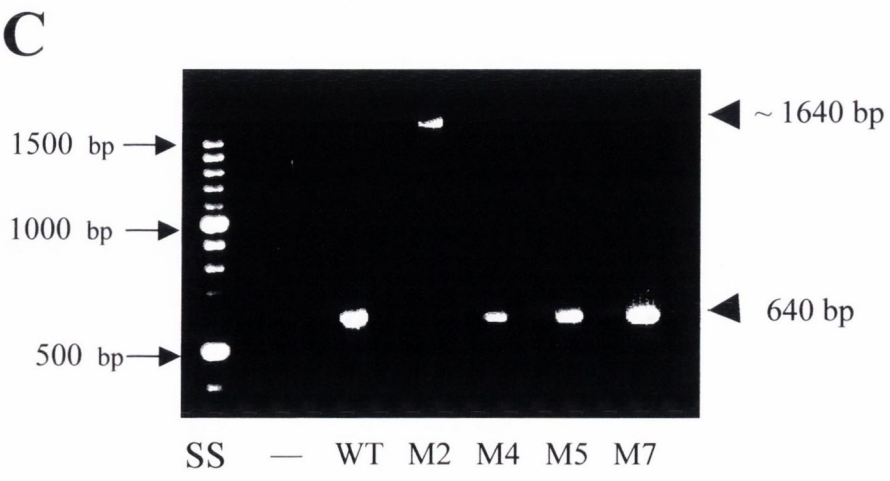
4.2.3.4 Marker rescue

To determine whether elevated *agn43-lacZ* transcription in the remaining mutant (M4) was the result of transposon insertion, and not spontaneous mutation, the chloramphenicol marker was moved, by P1 transduction, into the wildtype strain CSH50(pAg43). The mutant Locked ON phenotype was reproduced by all

Figure 4.5 PCR analysis of genomic DNA isolated from transposon mutants (M2, M4, M5 and M7). A dash (—) indicates negative control lanes, in which no template DNA was used. In lanes labelled 'WT', non-mutagenised CSH50 was used to provide template DNA for PCR. **A**. Amplification of the *cml* gene using primers CatF and CatR. **B**. Amplification of the *oxyR* gene using primers OxyR19F and OxyR19R. **C**. Amplification of the *rpoS* gene using primers rposR and rposF.


A
2000 bp
1000 bp
500 bp
~900 bp
SS — WT M2 M4 M5 M7


B
2000 bp
1000 bp
500 bp
~1900 bp
900 bp
SS — M2 M4 M5 M7 WT


C
1500 bp
1000 bp
500 bp
~ 1640 bp
640 bp
SS — WT M2 M4 M5 M7

transductants (data not shown), indicating that the mutation that altered *agn43-lacZ* expression resulted from insertion of the *cml* gene.

4.2.3.5 Cloning and sequencing of mutated DNA from M4

Genomic DNA from M4 was restricted with *Bam*HI and ligated with pUC18, which had been cut with the same enzyme. The transposon contains only one *Bam*HI restriction site within its multiple cloning site, which is upstream of the *cml* gene. Cloning with this enzyme should therefore result in a certain amount of downstream chromosomal DNA being cloned into pUC18 along with the *cml* marker. Clones containing the transposon and adjacent chromosomal DNA from M4 were selected for on L agar containing chloramphenicol, and termed pUCM4. The region downstream of the transposon in pUCM4 was sequenced (Lark Technologies Inc.) using the primer CatExtn, which reads outwards from the 3' end of the *cml* gene (Table 2.2).

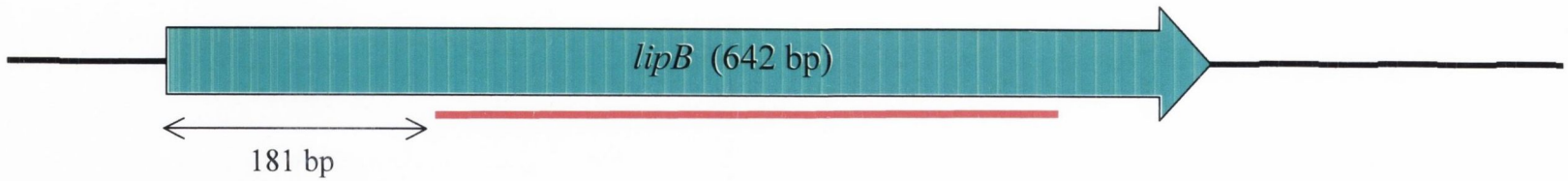
BLAST analysis of the sequence obtained from pUCM4 revealed that it was 100 % identical to part of the coding sequence of the *lipB* gene of *E. coli* (Figure 4.6). This indicated that the increased *agn43-lacZ* expression in M4 could be attributed to insertion of the transposon into the *lipB* coding sequnce. The *lipB* product has been identified as a negative regulator of Dam expression (Vaisvila *et al.*, 2000). As such, mutation of the *lipB* gene would be expected to result in an increase in the level of Dam expression. This provides an explanation for the Locked ON expression of *agn43-lacZ* in M4. Elevated levels of Dam should increase the rate of methylation resulting in a decrease in OxyR-mediated repression of *agn43*

Figure 4.6 (A) Sequence of the region downstream of the transposon in plasmid pUCM4, which was obtained by sequencing with primer CatExtn. (B) This sequence corresponds to a region (indicated by a red line) of the *lipB* open reading frame (green) starting 181 bp from *lipB* translational start site.

A

CCGGGTGATATTCCGGTGATCCAGAGCGATCGCGGTGGGCAGGTGACTTATCACGGGCCGGGGCAACAGGTGAT
GTATGTGTTGCTTAACCTGAAACGCCGTAAACTCGGTGTGCGTGAACTGGTGACCTTGCTTGNGCAAACAGTGGT
GAATACCCTGGCTGAACTGGGTATAGAAGCGCATCCTCGGGCTGACGCGCCAGGTGTCTATGTTGGGGAAAAGA
AAATTTGCTCACTGGGTTTACGTATTCGACGCGGTTGTTCATTCCACGGTCTGGCATTAAACGTCAATATGGATCT
TTCACCATTTTTACGTATTAATCCTTGTGGGNATGCCGGAATGGAAATGGCTAAAATATCACAATGGAAACCCGA
AGCGACGACTANTAATATTGCTCCACGTNTACTGGAAAATATTTTNGCGCTACTAAACAATCCGGNCTTCGNATN
TNTTNCCNNT

B



4.3 Discussion

Alterations in the levels of different sigma factors allow the *E. coli* cell to alter the specificity of RNA polymerase and transcribe different sets of genes. An increase in the level of the sigma factor, RpoS, during the stationary phase of *E. coli* growth, causes the induction of many genes which are require for stationary-phase adaptation (Hengge-Aronis, 1996). The dependence of different genes on RpoS varies from 2-fold to over 100-fold (Schellhorn *et al.*, 1998). It was therefore decided to determine whether RpoS might be responsible for the stationary phase induction of *agn43* transcription, which was described in Chapter 3. The superinduction of Ag43 expression in an *rpoS* mutant indicated that, while not responsible for the stationary phase induction of *agn43,* RpoS does have a negative regulatory effect on this gene. In *E. coli* several genes are negatively regulated by RpoS, including the *fim* operon, which encodes the phase variable Type 1 fimbriae (Dove *et al.*, 1997). In *S. typhimurium*, RpoS represses expression of Pef fimbriae, which are expressed phase variably by a mechanism involving Dam and Lrp (Nicholson and Low, 2000). It has recently been reported that an *rpoS* mutation enhances biofilm formation in *E. coli* (Corona-Izquierido and Membrillo-Hernandez, 2002). Ag43 is involved in biofilm formation (Danese *et al.*, 2000). However the RpoS-regulated biofilm formation mentioned above was promoted by an unidentified extracellular factor (Corona-Izquierido and Membrillo-Hernandez, 2002) and could not therefore be related to an increase in surface expression of Ag43.

The dual regulation of several genes by OxyR and RpoS has previously been reported (Mukhopadhyay and Schellhorn, 1997). OxyR itself appears to be

negatively regulated by RpoS since its expression is increased in *rpoS* mutants (Gonzalez-Flecha and Demple, 1997). Such an increase in OxyR expression should result in a decrease in *agn43* transcription. Thus, the observed increase in Ag43 expression in an *rpoS* mutant indicates that the regulation of *agn43* by RpoS is OxyR-independent. Farewell *et al.* (1998) proposed that competition between sigma factors for limiting amount of core RNA polymerase may account for RpoS-dependent repression of σ^{70}-dependent genes. A mutation in *rpoS* resulted in a 50% increase in expression of the genes studied (Farewell *et al.*, 1998). The 300% induction of Ag43 in an *rpoS* background is much greater by comparison and therefore may not be attributable to changes in sigma factor competition. Negative regulation of Ag43 by RpoS is more likely to be indirect, occurring via an RpoS-dependent *trans*-acting repressor that has not yet been identified.

As discussed in Chapter 1, it seems likely that the OFF phase of *agn43* expression is brought about via a transition state during which the hemimethylated 5'-GATC-3' sites are bound by a protein that protects them from full methylation. Although OxyR binds unmethylated *agn43* DNA and has recently been shown to bind hemimethylated *agn43* DNA *in vivo* (Correnti *et al.*, 2002) it was still possible that an additional protein was involved in protecting the *agn43* 5′-GATC-3′ sites from Dam methylation. Lrp acts as a methylation-blocking factor in the Dam-dependent regulation of the *pap*, *sfa* and *daa* operons, which, like *agn43,* are expressed in a phase variable manner (van der Woude and Low, 1994). For this reason Lrp was investigated as a possible regulator of *agn43* expression. Ag43 phase variation was altered in the absence of Lrp with the retardation of OFF to ON switching resulting in a bias towards the OFF phase. The absence of this OFF bias in the *fim*$^-$ Lrp$^-$

strain, however, indicated that it was an indirect effect probably due to the constitutive expression of Type 1 fimbriae and the resulting net cellular oxidation, which is thought to favour the reduced (*agn43*-repressing) form of OxyR.

Transposon mutagesis was conducted in an attempt to identify any other factors that might affect the expression of the *agn43* gene. The phase variable nature of *agn43* expression placed limitations on the efficiency of detection by colony screening. Mutagenesis of the Locked ON OxyR mutant was performed, but did not result in the detection of any interesting mutants. However, transposon mutagenesis of the wildtype strain CSH50 did result in the generation of a mutation that detectably altered the expression of *agn43-lacZ*. This mutation mapped to the *lipB* gene, which codes for the 24k-Da protein LipB. LipB is thought to be involved in the ligation of lipoyl groups to lipoate-dependent ezymes, such as the pyruvate dehydrogenases (Morris *et al.*, 1995). Previously, LipB has been shown to be a negative regulator of *dam* gene expression. Dam is subject to growth phase-dependent regulation and is expressed at five-fold higher levels in exponential than in stationary phase of growth (Rasmussen *et al.*, 1994). Thus, the level of Dam is adjusted to match the increased rate of DNA synthesis in exponential phase. In *lipB* mutants Dam expression remains constitutively high at all stages of growth (Vaisvila *et al.*, 2000).

Thus, the Locked ON phenotype of the transposon mutant M4 could be attributed to the loss of LipB-mediated repression of Dam expression. Increased levels of Dam would enhance transcription of *agn43* by increasing the rate of methylation and hence decreasing the likelihood of binding by OxyR. Consistent with this,

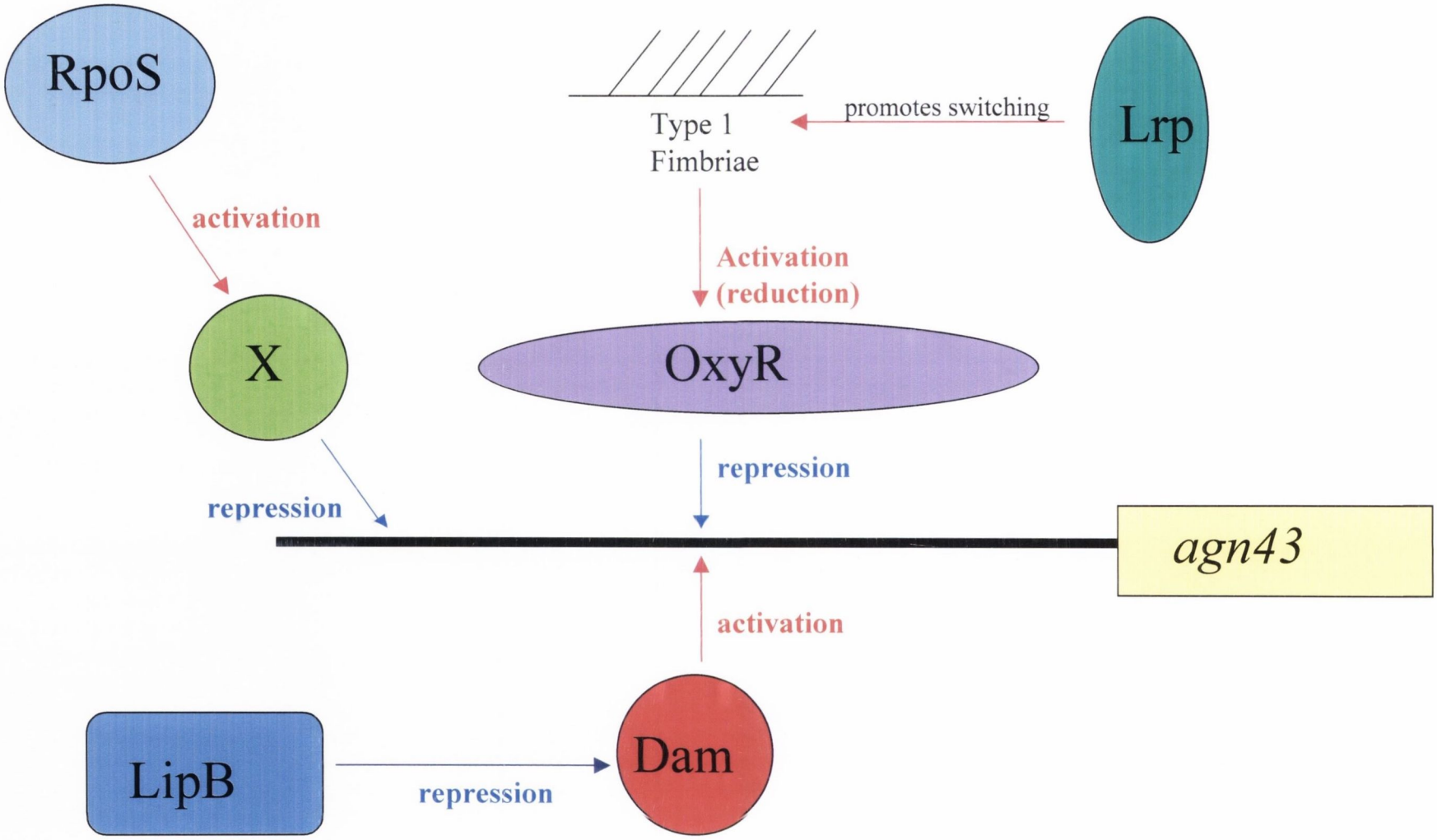
overexpression of Dam from a multicopy plasmid results in Locked ON transcription of *agn43* (Haagmans and van der Woude, 2000). Wallecha *et al.* (2002) have recently reported that Dam activates *agn43* transcription by an additional mechanism, which is independent of OxyR. This unknown mechanism may also contribute to the increased *agn43-lacZ* transcription in the *lipB* mutant.

It is comforting that *oxyR* and *rpoS* mutants were generated and detected in this study. However, only one *rpoS* mutant and two *oxyR* mutants were detected. Statistically speeking, this is not a sure indication that the entire genome has been covered. It is therefore possible that other factors involved in *agn43* regulation were never hit and disrupted by a transposon. It is possible, too, that genes encoding other regulators were mutated but their effects were masked due to phase variation of *agn43-lacZ*. Also, the effects of some mutations on *agn43-lacZ* expression might not have been large enough to result in an increase in β-galactosidase expression that was visible on X-Gal. Furthermore, it is only possible to identify non-essential genes in this study, as the insertion of a transposon into an essential gene would prove lethal to the bacterium.

Work presented in this chapter described the identification of three factors RpoS, Lrp and LipB, that affect the regulation of *agn43*, albeit indirectly. Figure 4.7 summarises the likely routes by which each of these proteins exerts its regulatory effect on the *agn43* gene. The stationary phase sigma factor, RpoS, most likely activates an, as yet unidentified, repressor of *agn43*, which is termed 'X' in Figure 4.7. Lrp controls the phase variation of Type-1 fimbriae, which affect transcription of *agn43* by altering the redox sate of OxyR. Finally the LipB protein is responsible

for the growth phase dependent repression of Dam, a protein that activates *agn43* transcription by blocking binding of the OxyR repressor.

Figure 4.7 Model for the mechanisms by which Lrp, RpoS and LipB regulate expression of the *agn43* gene. Lrp controls the switching of Type 1 fimbriae, which affects the cellular thiol-disulphide status and redox state of the OxyR repressor. RpoS activates an unidentified repressor of *agn43* (X). LipB represses expression of the Dam activator during the stationary phase of growth.


RpoS
activation
X
repression
Type 1
Fimbriae
promotes switching
Lrp
Activation
(reduction)
OxyR
repression
agn43
activation
LipB
repression
Dam

5.1 Introduction

The *agn43* regulatory domain contains a region that has high homology with OxyR-binding sites of other genes and contains three 5′-GATC-3′ sites, which are targets for methylation by Dam (Henderson and Owen, 1999). The Dam protein is present at approximately 130 copies per cell in rapidly growing *E. coli* K-12. It catalyses the AdoMet-dependent methylation of the amino group of adenine of the palindromic sequence 5′-GATC-3′. This tetranucleotide is the only sequence to which Dam can bind specifically. Dam methylates hemimethylated and unmethylated 5′-GATC-3′ targets with equal efficiency, but has no effect on fully methylated DNA (Modrich, 1987). Only one methyl group is transferred with each binding event regardless of whether the DNA is hemi-methylated or unmethylated (Herman and Modrich, 1982).

Dam methylation has multiple functions in the cell including regulation of the initiation of chromosome replication at *oriC*. A time lag before Dam methylates newly synthesised DNA results in the transient presence of hemimethylated DNA behind the replication fork. Initiation of replication is inhibited when 5'-GATC-3' sites in the region of *oriC* are in the hemimethylated state (Russell and Zinder, 1987). Thus, by catalysing the transition of DNA from the hemimethylated state to the fully methylated state, Dam controls the timing of chromosome replication. Dam also has a central role in the methyl-directed mismatch repair system (Modrich, 1989). The presence of hemimethylated DNA allows the mismatch repair system to discriminate between the template and nascent strands and correct any errors in the newly synthesized strand.

The transient presence of hemimethylated DNA may also allow the expression of Dam-regulated genes to be regulated in response to the cell cycle. Dam has been identified as a regulator of several chromosomal and non-chromosomal genes including *dnaA*, *pap* and the *mom* gene of phage Mu. Dam is also involved in regulation of *agn43* and transcription is Locked OFF in strains bearing mutations in the *dam* gene (Henderson and Owen, 1999). In Chapter 3 it was demonstrated that a correlation exists between the methylation state of three 5'-GATC-3' sites, upstream of *agn43*, and the state of *agn43* transcription. These sites were methylated and unmethylated in Phase ON and Phase OFF cells, respectively. The *agn43* 5'-GATC-3' sites lie within a region that has high homology with OxyR-binding sites upstream of the *mom* and *oxyR* genes. This suggests that Dam exerts its regulatory role by precluding OxyR binding, and hence OxyR-mediated repression of *agn43* transcription.

Dam-mediated methylation of 5′-GATC-3′ sites destabilizes the double helix due to the decreased stability of methyladenine-thymine interactions and as such a methylated duplex is more easily melted. Furthermore, methylation of several closely grouped 5'-GATC-3' sites at *oriC* has been shown to result in a significant degree of DNA bending (Kimura *et al.*, 1989). Methylation-induced DNA bending could possibly affect binding of transcriptional regulators and thus protein expression. Alternatively, Dam methylation may block protein binding by steric hindrance as the methyl group of methyladenine protrudes out of the major groove. This is the likely mechanism by which methylation of the *agn43* 5′-GATC-3′ sequences abrogates binding of the OxyR repressor, since the methylation of just

three sites seems unlikely to result in a sufficient degree of bending to block protein binding.

It has not previously been determined how many of the three *agn43* 5′-GATC-3′ sites must be available for methylation in order for binding of OxyR to be perturbed. It has been suggested that any asymmetry introduced by flanking sequences to the symmetrical 5′-GATC-3′ may give some orientation to the interaction of Dam with the DNA and may influence the order in which nearby sites are methylated (Bergerat *et al.*, 1989). Hence not all of the 5'-GATC-3' sites may have an equal role in the phase variation of *agn43*.

In Chapter 3 it was shown that a correlation exists between the methylation state of the *agn43* 5′-GATC-3′ sites and the phase state of *agn43* transcription. In this chapter the roles of the three 5′-GATC-3′ sites in Ag43 phase switching are dissected by performing site directed mutagenesis of the target sequences in all possible combinations. A series of reporter plasmids are constructed which contain combinations of mutations corresponding to all the possible methylation patterns, which might affect *agn43* expression. The switch phenotypes of these constructs are assessed in a number of isogenic strains differing in their ability to express OxyR and Dam. The ability of OxyR to bind to methylated DNA from mutant constructs is examined in order to determine the effects of different methylation patterns on OxyR-mediated repression. To discriminate between switch phenotypes due to methylation dependent, and methylation-independent disruption of OxyR binding, unmethylated mutated *agn43* DNA was also tested as a substrate for OxyR binding.

5.2 Results

5.2.1 Site-directed mutagenesis of 5'-GATC-3' sites

The three 5′-GATC-3′ sites within the *agn43* regulatory region were disrupted by site-directed mutagenesis to permit their role in phase-variable expression of the *agn43* gene to be assessed. Two different procedures were used to construct mutated plasmids. These were (a) the Quickchange method and (b) the unique site elimination (USE) method.

5.2.1.1 The Quickchange procedure

Mutagenesis of the *agn43* promoter was originally conducted in the plasmid pAg43. However, product yields in this vector were generally very poor. Accordingly, in order to increase efficiency, the *Bam*HI-*Hin*dIII fragment containing the 373-bp *agn43* regulatory region was subcloned from pAg43 into the corresponding site in the smaller vector, pUC18 to give pDEW2 (Figure 5.1). The site-directed mutagenesis procedure utilized the Quickchange kit (Stratagene) and is summarised in Figure 5.2. Two complementary primers, containing the desired mutation were annealed to different strand of the denatured template plasmid (Figure 5.2). The proofreading DNA polymerase *Pfu* was used to extend the mutagenic primers around the remainder of the template, resulting in circular nicked strands. The resulting product was treated with the restriction endonuclease *Dpn*1, which cuts only at methylated 5'-GATC-3'sites. This enzyme digested the methylated template DNA leaving intact the unmethylated newly synthesised DNA, which contained the desired mutation (Figure 5.2). Annealing of the newly synthesised complimentary strands resulted in a plasmid containing staggered nicks, which was then transformed into *E. coli* XL-1 blue. The X-LI blue cells repair the nicks in the

mutated plasmid. DNA was prepared from transformed colonies and sequenced to validate the presence of the desired mutation. Each mutant was digested with *Bam*HI and *Hin*dIII to release the 373-bp insert, which was then subcloned into the *Bam*HI-*Hin*dIII site of pQF50. This created a transcriptional fusion of each mutated regulatory region to *lacZ*, allowing the effect of each (combination of) mutation(s) on *agn43* transcription to be assessed by comparison of expression in mutant plasmids with that in pAg43.

5.2.1.2 The Unique-site elimination method

Three mutant plasmids were created using the USE mutagenesis kit (Pharmacia), which is based on the unique site elimination method of Deng and Nickloff (1992). The benefit of this method was that it did not require the subcloning of the *agn43* regulatory region between vectors. The steps involved in the procedure are summarised in Figure 5.3. With this method the selection of constructs carrying a desired mutation is based on the elimination of a unique restriction site elsewhere on the plasmid. The procedure requires two primers, which anneal to different regions on the same strand of the template plasmid. One primer contains the desired mutation whilst another contains a mutation that eliminates a unique restriction site (the *Sal*I site of pAg43 was used in this study). The target plasmid is heat denatured and the mutagenic primers are annealed to the same strand of the single-stranded plasmid DNA. Both primers become incorporated as a new strand of DNA is synthesised by T4 DNA polymerase. Hence, the non-selectable mutation of interest is coupled to a mutation that can be selected for by restriction digest. The product was then digested with the enzyme *Sal*I whose restriction site had been eliminated by the USE primer. Parental plasmid DNA is linearized by the

Figure 5.1 Map of the plasmid pDEW2 in which site-directed mutagenesis of 5′-GATC-3′ sites was performed. *rep* (pMB1), origin of replication of plamid pMB1; *bla*, β-lacatamase gene; *agn43*, 373 bp of the *agn43* regulatory region subcloned from pAg43.

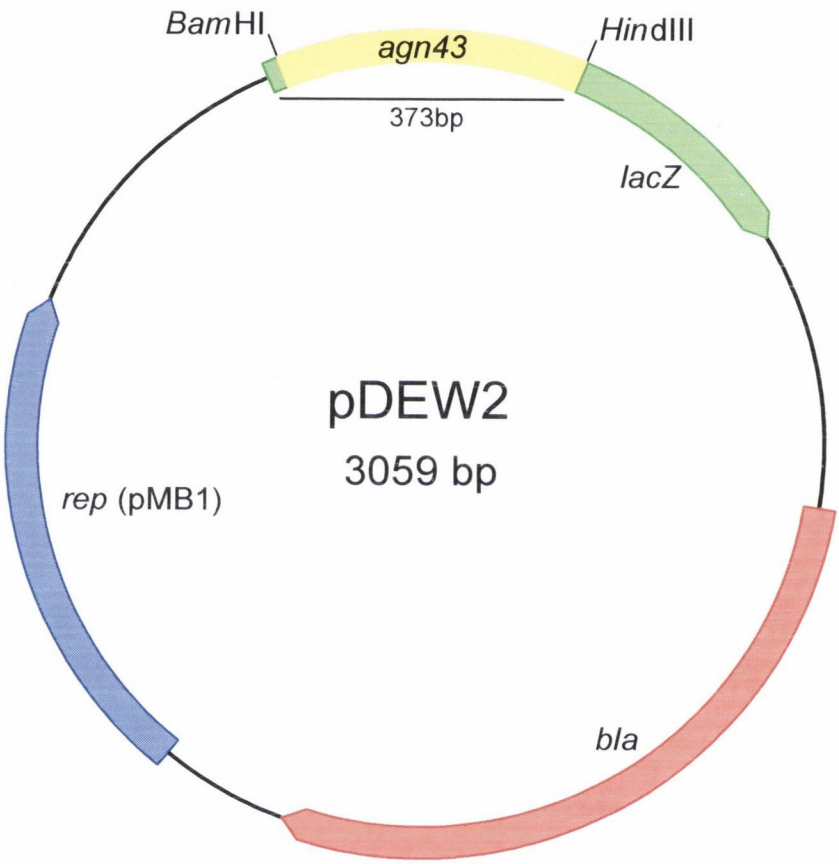

BamHI
agn43
HindIII
373bp
lacZ
pDEW2
3059 bp
rep (pMB1)
bla

Figure 5.2 Summary of the Quickchange site-directed mutagenesis method. **A**. Primers containing the desired mutation anneal to the same sequence on opposite strands of the denatured parental plasmid. **B**. Temperature cycling with *Pfu* polymerase extends the mutagenic primers resulting in nicked circular strands. **C**. The methylated non-mutated parental DNA is digested with *Dpn*I. **D**. The mutated plasmid is transformed into *E. coli* X-L1 blue, which repairs the staggered nicks. **E**. The insert containing the desired mutation is subcloned into pQF50 recreating a transcriptional fusion of the mutated *agn43* regulatory region to *lacZ*.

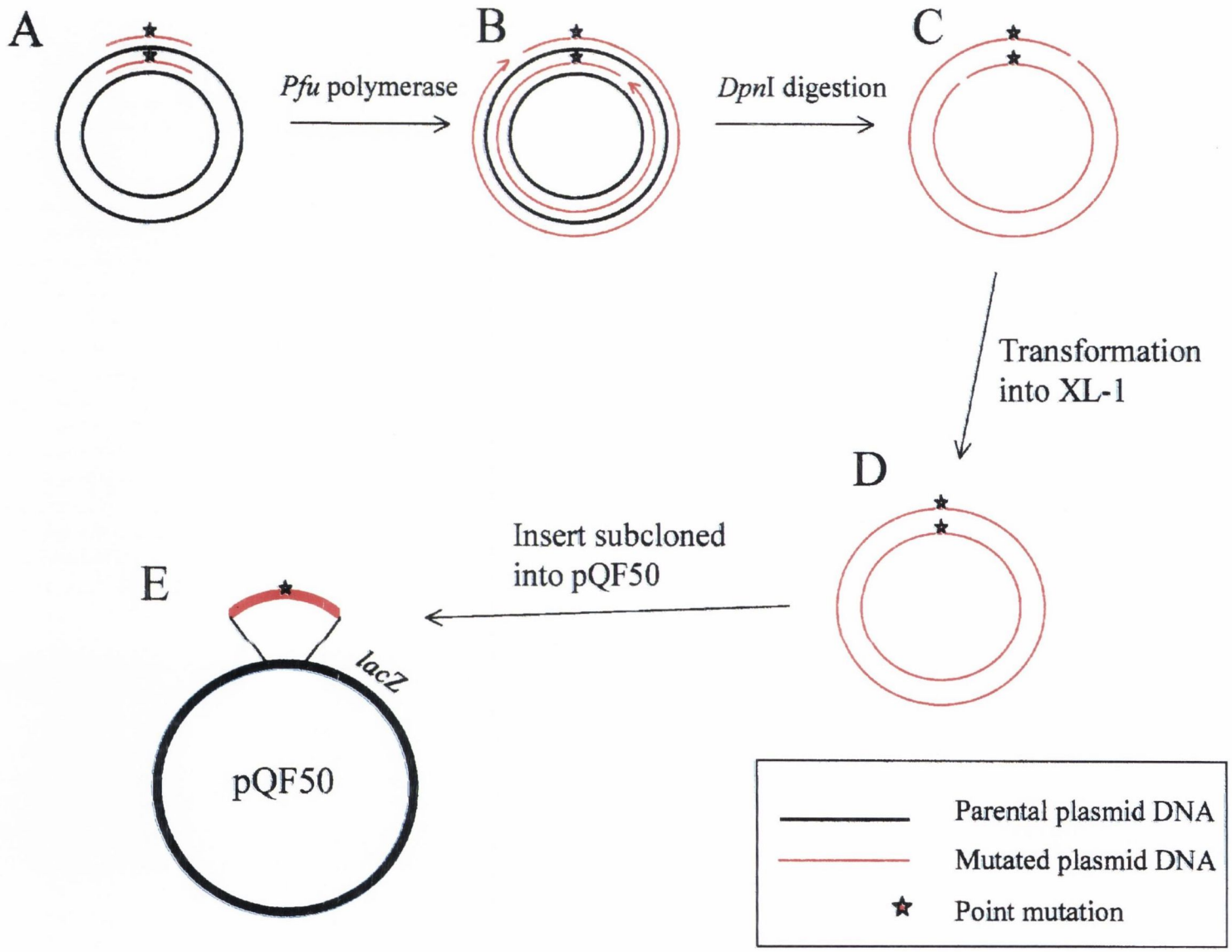
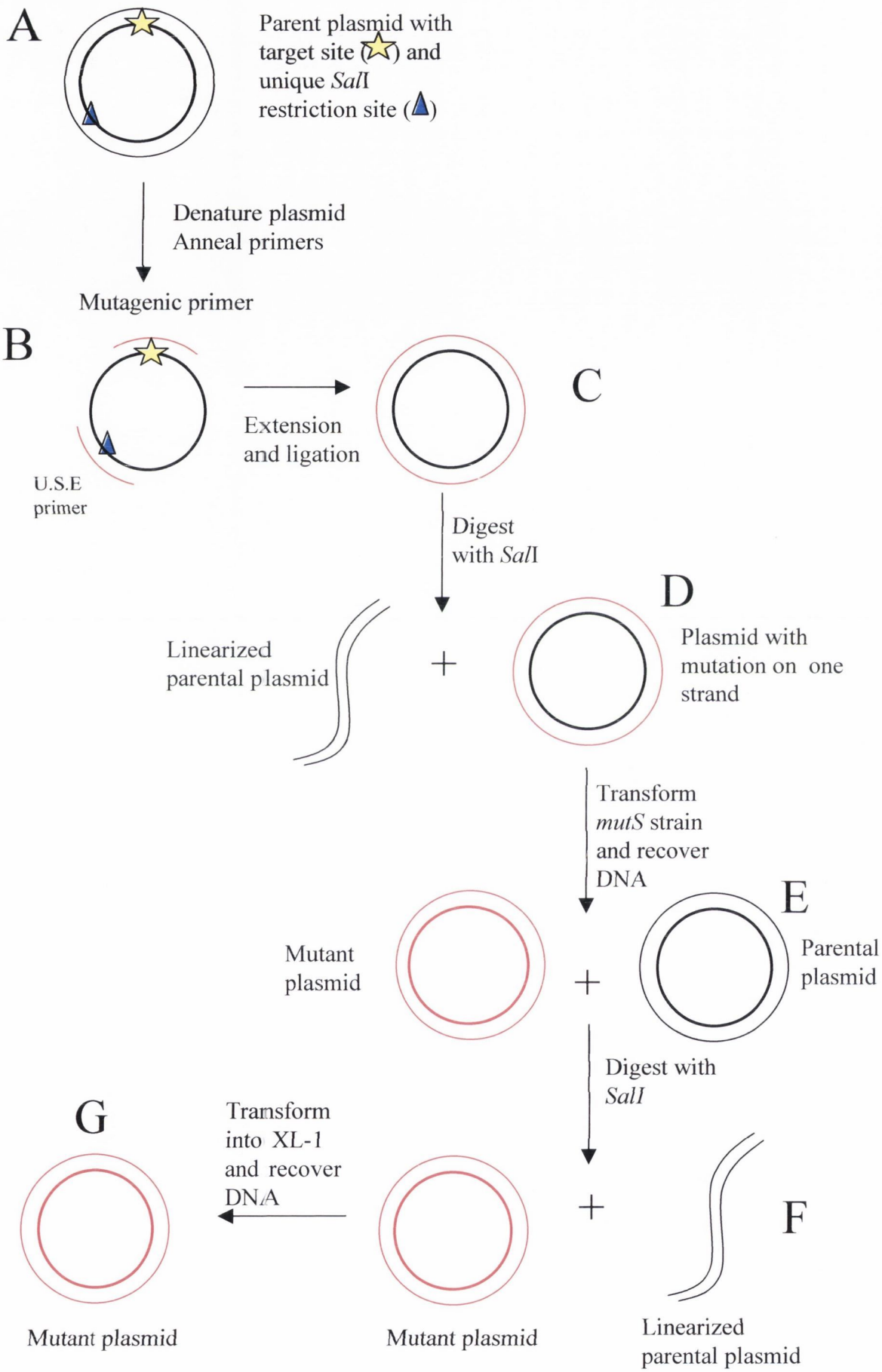

A
B
C
Pfu polymerase
DpnI digestion
Transformation into XL-1
D
Insert subcloned into pQF50
E
lacZ
pQF50
Parental plasmid DNA
Mutated plasmid DNA
Point mutation

Figure 5.3 Unique-site elimination method of site-directed mutagenesis. **A**. The template plasmid contains a unique *Sal*I restriction site and the target site to be mutated. **B**. The plasmid is denatured and the mutagenic primer and USE primer are annealed. **C**. Extension with T4 polymerase and ligation with T4 ligase results in a plasmid containing the desired mutations on one strand only. **D**. Digestion with *Sal*I linearizes the parental plasmid. **E**. The plasmid is transformed into a *mutS* strain. DNA recovered from transformants consists of both parental plasmid and mutant plasmid. **F**. Digestion with *Sal*I linearizes the parental plasmid. **G**. The DNA is transformed into *E. coli* X-L1 blue. DNA recovered from transformants consists of mutant plasmid only.


A
Parent plasmid with target site ( ) and unique *Sal*I restriction site ( )
Denature plasmid
Anneal primers
Mutagenic primer
B
U.S.E primer
Extension and ligation
C
Digest with *Sal*I
D
Linearized parental plasmid
+
Plasmid with mutation on one strand
Transform *mutS* strain and recover DNA
E
Mutant plasmid
+
Parental plasmid
Digest with *SalI*
G
Transform into XL-1 and recover DNA
+
F
Mutant plasmid
Mutant plasmid
Linearized parental plasmid

enzyme. However, plasmid DNA containing the USE mutation is resistant to digestion. Digestion with SalI leaves only plasmid molecules that contain the wild type sequence on one strand and the mutated sequence on the other. The digested DNA was then transformed into a *mutS* strain, which is deficient in misatch repair and therefore able to propogate the mutated plasmid. Transformation efficiency is 100- to 1000-fold greater for circular DNA than for linearized DNA so transformants are more likely to contain the mutant plasmid. Replication of each plasmid in this strain gives rise to one wild type daughter molecule and one mutated daughter molecule, which contains the desired mutations on both strands. Plasmid DNA was then recovered from the *mutS* strain and subject to digestion with *Sal*I, whose target site had been eliminated in the mutant plasmid. This removed the wildtype plasmid leaving only DNA that contained the desired mutation on both strands. DNA was transformed into *E. coli* X-L1 and plasmid DNA recovered from transformed colonies was sequenced manually to confirm the presence of the desired mutation(s).

5.2.2 Strategy for the generation of mutant constructs

The sequences of mutagenenic primers used for the construction of plasmids with mutated 5'-GATC-3' sites are listed in Table 5.1. The nature of the mutations in each construct and the names of the parental plasmid and primers used in each case are presented in Table 5.2.

5.2.2.1 Construction of plasmids pDEW10-pDEW16

To render the 5'-GATC-3' sites impervious to methylation by Dam methylase, a T residue was substituted for the A in each site singly or in combination with

Table 5.1 The sequence of mutagenic primers used in this study

Name	Sequence[a]
GTTC1t	5'-CAATAATAGAATAAACG**T**TCAATAATTTCCTATTTTATCG-3'
GTTC1b	5'-CGATTAAAATAAAGATATTGA**A**CGTTTTATTCTATTATTG-3'
GTTC2t	5'-CAATATCCCCTATTTTATCG**T**TCGTTTATATCGATC-3'
GTTC2b	5'-GATCGATATAAACGA**A**CGATAAAATAGGGGATATTG-3'
GTTC3t	5'-TATCGATCGTTTATATCG**T**TCGATAAGGCTAATAATAACC-3'
GTTC3b	5'-GGTTATTATTAGCCTTATCGA**A**CGATATAAACGATCGATA-3'
GTTC2+3t	5'-CTATTTTATCG**T**TCGTTTATATCG**T**TCGATAAGC-3'
GTTC2+3b	5'-GCTTATCGA**A**CGATATAAACGA**A**CGATAAAATAG-3'
GATG3t	5'-CGTTTATATCGAT**G**GATAAGCTAATAATAACC-3'
GATG3b	5'-GGTTATTATTAGCTTATCCAT**C**GATATAAACG-3'
CATC2t	5'-CAATATCTATTTTATC**C**ATCGTTTATATCGATC-3'
CATC2b	5'-GATCGATATAAACGAT**G**GATAAAATAGATATTG-3'
USE	5'-CCATGGACGCGTGACGTGGACTCTAGAGGATCC-3'

[a] Altered bases of mutagenic oligonucleotides are in bold and are underlined

Table 5.2 Strategy for the generation of constructs with mutated 5'-GATC-3' sites. For each mutated product the parental plasmid and primers used for construction, and the nature of any mutations introduced, are given. For plasmids pDEW10-pDEW22 the name of the intermediate product, from which the mutated *agn43* region was subcloned into pQF50, is also given.

Mutated product		Mutations in 5'-GATC-3' sites			Parental	Primers used
Intermediate	Final	Site 1	Site 2	Site 3	plasmid	
pDEW3	pDEW10	A→T			pDEW2	GTTC1t GTTC1b
pDEW4	pDEW11		A→T		pDEW2	GTTC2t GTTC2b
pDEW5	pDEW12			A→T	pDEW2	GTTC3t GTTC3b
pDEW6	pDEW13	A→T	A→T		pDEW3	GTTC2t GTTC2b
pDEW7	pDEW14	A→T		A→T	pDEW3	GTTC3t GTTC3b
pDEW8	pDEW15		A→T	A→T	pDEW2	GTTC2+3t GTTC2+3b
pDEW9	pDEW16	A→T	A→T	A→T	pDEW3	GTTC2+3t GTTC2+3b
pDEW17	pDEW20		G→C		pDEW2	CATC2t CATC2b
pDEW18	pDEW21			C→G	pDEW2	GATG3t GATC3b
pDEW19	pDEW22	A→T	G→C		pDEW3	CATC2t CATC2b
	pDEW23	A→T		C→G	pDEW10	GATG3t USE
	pDEW24		G→C	C→G	pDEW20	GATG3t USE
	PDEW25	A→T	G→C	C→G	pDEW22	GATG3t USE

substitutions in neighbouring sites. Target 5′-GATC-3′ sites in plasmid pDEW2 were disrupted by introduction of an A-to-T substitution in each site singly. Primer pairs GTTC1t and GTTC1b, GTTC2t and GTTC2b and GTTC3t and GTTC3b were used to disrupt the first, second and third 5′-GATC-3′ sites, respectively, generating plasmids pDEW3, pDEW4 and pDEW5, respectively (Tables 5.1 and 5.2).

In order to disrupt the 5′-GATC-3′ sites in combination the following strategy was used. Primer pairs GTTC2t and GTTC2b, and GTTC3t and GTTC3b were used individually to disrupt the second and third 5'-GATC-3' sites in plasmid pDEW3 resulting in pDEW6 and pDEW7, respectively. Primers GTTC2+3t and GTTC2+3b were used to disrupt the second and third 5′-GATC-3′ sites in plasmid pDEW2 resulting in pDEW8. Finally primers GTTC2+3t and GTTC2+3b were used to disrupt second and third 5′-GATC-3′ sites in pDEW3 to produce pDEW9 in which all three 5′-GATC-3′ sites contained within the cloned *agn43* regulatory region were disrupted. The presence of each desired mutation was confirmed by manual sequencing. Sequencing gels showing the mutations in several of these constructs are shown in Figure 5.4.

In order to allow the effect of these mutations on *agn43* transcription to be assessed, each mutated *agn43* regulatory region sequence had to be cloned into plasmid pQF50 upstream of a promoterless *lacZ* gene. Accordingly, the *Bam*HI-*Hin*dIII fragments from each of the mutant plasmids were individually subcloned into the corresponding site in the vector pQF50 creating transcriptional fusions of the mutated *agn43* regulatory regions to *lacZ*. Subcloning from plasmids pDEW3-9 into pQF50 resulted in the final products pDEW10-16, respectively (Table 5.2).

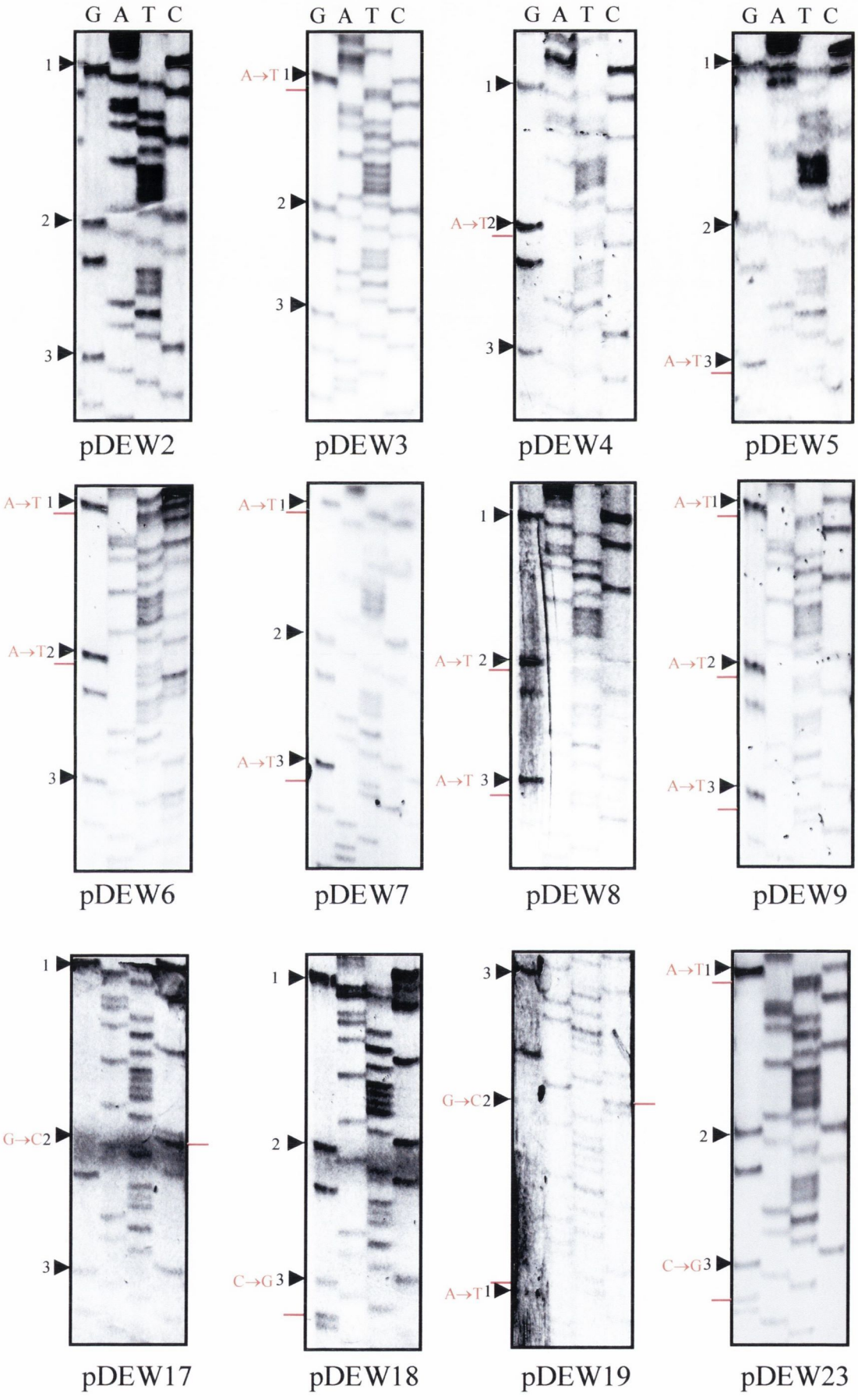
These plasmids only differed from the wild type reporter plasmid pAg43 at their 5'-GATC-3' sites as indicated in Table 5.2. As such, any differences in *agn43-lacZ* expression compared to pAg43 could be attributed only to disruption of the 5'-GATC-3' sites.

5.2.2.2 Construction of plasmids pDEW20-pDEW25

For reasons described later in this chapter it was decided to disrupt methylation of the second and third 5′-GATC-3′ sites by mutating residues other than adenine. Accordingly, the second 5′-GATC-3′ site was altered to 5′-**C**ATC-3′ and the third site was altered to 5′-GAT**G**-3′ in all possible combinations. The primers CATC2t and CATC2b were used to introduce G to C mutations into site 2 of pDEW2 generating the single mutant pDEW17 (Table 5.2). In a similar manner primers GATG3t and GATG3b were used to introduce a C-to-G mutation in site 3 of pDEW2 to construct plasmid pDEW18 (Table 5.2). Primers CATC2t and CATC2b were utilised again to mutate site 2 of pDEW3 generating a double mutant, pDEW19, in which 5'-GATC-3' sites 1 and 2 were disrupted. The presence of each desired mutation was confirmed by manual sequencing (Figure 5.4). Mutated inserts from each of the mutant plasmids were subcloned into the vector pQF50 creating a transcriptional fusion of the mutated *agn43* regulatory region to *lacZ*. Subcloning from plasmids pDEW17-19 into pQF50 resulted in constructs pDEW20-22, respectively.

The remaining plasmids, pDEW23-25, were constructed by the unique-site elimination method (Section 5.2). The USE primer is complementary to part of the pAg43 multiple cloning site and contains a mutation that eliminates the unique *Sal*I

Figure 5.4 Sequencing gels showing the location of different mutations used to disrupt the 5'-GATC-3' sites of plasmids constructed in this study. Plasmid DNA, corresponding to each mutant construct, was prepared form XL-1 and sequenced using the Pharmacia T7 sequencing kit, as described in Chapter 2 (see Section 2.6). With the exception of pDEW22, all constructs were sequenced using the primer Ag43R (Table 2.2), which had been used initially to amplify the cloned *agn43* regulatory region. This primer anneals to the forward strand approximately 270 bp downstream of the region of interest and generates sequence that corresponds to the coding strand when read from top to bottom. pDEW22 was sequenced using primer Ag43F which was also used initially to amplify the cloned *agn43* regulatory region but anneals to the reverse strand approximately 80 bp upstream of the region of interest. Thus the sequence data obtained for the coding strand of pDEW22 reads from bottom to top. Prior to electrophoresis, sequencing reactions were loaded onto the gel in such a way that bands in lanes from left to right correspond to G, A, T and C residues, within the coding strand, respectively. This particular order makes the intact 5'-GATC-3' sites clearly visible, thus facilitating localization of the region of interest within the extended sequencing ladder obtained. All 5'-GATC-3' sites (mutated and non-mutated) within the sequences shown are labelled at the G residue by arrowheads and numbers, which indicate the position of each site in a manner consistent with that used throughout this study. The nature of each mutation is indicated in red beside the respective mutated 5'-GATC-3' site and the position of the mutated base is indicated with a red dash. Due to a common artefact of sequencing, bands corresponding to a single position could be seen within 2 lanes, on some occasions. However bands corresponding to the known nucleotides were always present and these regions did not generally overlap with the target sites for mutagenesis. The 'C' reaction for pDEW6 initially failed resulting in bands at almost every position. This reaction was repeated, as were others when there was any doubt surrounding the presence or absence of mutations.


G A T C
1
2
3
pDEW2
G A T C
A→T 1
2
3
pDEW3
G A T C
1
A→T2
3
pDEW4
G A T C
1
2
A→T3
pDEW5
A→T 1
A→T2
3
pDEW6
A→T 1
2
A→T3
pDEW7
1
A→T 2
A→T 3
pDEW8
A→T1
A→T2
A→T3
pDEW9
1
G→C2
3
pDEW17
1
2
C→G 3
pDEW18
3
G→C2
A→T1
pDEW19
A→T1
2
C→G3
pDEW23

restriction site. The USE primer was used along with GATG3t to construct the double mutant pDEW23, which contained mutations in sites 1 and 3, from pDEW10. These primers were also used to combine a C→G mutation in site 3 with the G→C mutation in site 2 of pDEW22, generating the double mutant, pDEW24. Finally, the same primers were employed to create the triple mutant pDEW25, from pDEW22 (Table 5.2). pDEW25 was similar to the other triple mutant (pDEW16), in that all three 5'-GATC-3' sites were inactive as targets for Dam methylation. However, as described later, the different points at which the 5'-GATC-3' sites were disrupted, had drastically different effects on transcription of *agn43-lacZ*.

5.2.3 Analysis of constructs pDEW10-pDEW16 in CSH50

The plasmids pDEW10 to pDEW16, were transformed into the *lac*⁻ strain CSH50 (Table 2.1) to allow comparison of their phenotypes with that arising from the wild type *agn43* regulatory region in plasmid pAg43. CSH50 is phase variable for expression of chromosomally encoded Ag43 protein and for expression of the wild type *agn43-lacZ* transcriptional fusion. Therefore any differences in the switch phenotype of mutant plasmids could be attributed to the presence of A-to-T mutations in the 5'-GATC-3' sites. The switch phenotype of each mutated plasmid was assessed by growth of CSH50 transformants on L agar containing X-Gal. The actual level of *agn43-lacZ* transcription in Phase ON colonies was determined by means of β-galactosidase assay (see Section 2.5.2 of Materials and Methods). The phenotypes of mutants pDEW10-pDEW16 in CSH50 are presented in Table 5.3. and also in Table 5.7 for comparison with other results.

Despite the disruption of its first 5'-GATC-3' site, pDEW10 was able to undergo phase variation in CSH50 and *agn43-lacZ* transcription in this mutant was

comparable to that in wild type pAg43. The phenotype of pDEW10 showed that the first 5′-GATC-3′ site was not required absolutely for *agn43-lacZ* to be expressed phase variably. Interestingly, all other combinations of mutations (pDEW11 to pDEW16) resulted in constitutive Locked ON expression of β-galactosidase, and not the Locked OFF phenotype one might have expected from the elimination of the Dam target sequences (Table 5.3). A Locked OFF phenotype was expected since these mutations preclude the Dam methylation that is required to block binding of the OxyR repressor. The level of *agn43-lacZ* expression in Locked ON mutants varied depending on the (combination of) mutations introduced. The highest level of expression was displayed by pDEW15, which contains A-to-T mutations in both 5'-GATC-3' sites 2 and 3. This suggested that the effects of the A-to-T mutations in sites 2 and 3, on *agn43-lacZ* transcription, were cumulative in pDEW15. The lowest level of β-galactosidase expression was displayed by pDEW14, which contains A-to-T mutations in 5'-GATC-3' sites 1 and 3 (Table 5.3). The fact that expression in pDEW14 was lower than in pDEW12 (site 3 mutated), indicated that the up-regulatory effect of an A-to-T mutation in site 3 was counteracted somewhat by the additional mutation in site 1. Similarly pDEW13 (sites 1 and 2 mutated) expressed *agn43-lacZ* at a lower level than pDEW11 (site 2 mutated), suggesting that an A-to-T mutation in site 2 could also be counteracted by the same mutation in site 1 (Table 5.3).

The Locked ON phenotype of the triple mutant, pDEW16, suggests that this phenomenon was a methylation-independent effect, since Dam methylase activity is known to be required for Ag43 expression in an $OxyR^+$ background (Henderson and Owen, 1999). A possible explanation for expression of *agn43-lacZ* in the absence

Table 5.3 The effects of mutations in the 5'-GATC-3' sites of *agn43-lacZ* on switch phenotypes and β-galactosidase activity in strain CSH50

Construct	Methylation target[a]			Phenotype in CSH50	
	1	2	3	Switch phenotype	β–Gal[b]
pAg43	GATC	GATC	GATC	Switching	156(18)
pDEW10	G**T**TC	GATC	GATC	Switching	176(18)
pDEW11	GATC	G**T**TC	GATC	Locked ON	180(23)
pDEW12	GATC	GATC	G**T**TC	Locked ON	164(17)
pDEW13	G**T**TC	G**T**TC	GATC	Locked ON	124(15)
pDEW14	G**T**TC	GATC	G**T**TC	Locked ON	119(11)
pDEW15	GATC	G**T**TC	G**T**TC	Locked ON	205(21)
pDEW16	G**T**TC	G**T**TC	G**T**TC	Locked ON	193(16)

a. Substituted bases in each construct are in bold and underlined.

b. β-Gal, β-galactosidase activity. Assays were performed on duplicate cultures, in triplicate. Results shown are averages with standard deviation in parenthesis and are expressed in Miller units (MU). In switching phenotypes, the β-galactosidase values refer to the Phase ON state.

of Dam-mediated methylation, was that pDEW16 contained mutations that prevented OxyR from functioning as a repressor of *agn43*. Thus, it seemed likely that the Locked ON phenotypes of pDEW11-pDEW16 were due to methylation-independent disruption of OxyR-mediated repression, induced by the A-to-T mutations in 5'-GATC-3' sites 2 or 3.

5.2.4 The effect of OxyR overexpression on pDEW10-pDEW16

To explore the possibility that the A-to-T mutations in the Locked ON plasmids had decreased the affinity of the DNA for OxyR binding, this protein was over-supplied to the mutant plasmids through overexpression from the plasmid pAQ5. (pAQ5 contains the *oxyR* gene and promoter sequences, and confers chloramphenicol resistance; Table 2.1). This involved the co-transformation of CSH50 with each plasmid construct and pAQ5. CSH50 transformants were isolated on L agar supplemented with both ampicillin and chloramphenicol, which select for the presence of reporter plasmid and pAQ5, respectively. The switch phenotype of transformants was then assessed on L agar supplemented with X-Gal and the level of *agn43-lacZ* expression in each construct was determined by means of β-galactosidase assay. The results are presented in Tables 5.4 and 5.7.

Overexpression of OxyR resulted in some decrease in β-galactosidase expression and restored the phase variability of the plasmids bearing single mutations (pDEW11 and pDEW12). In contrast the double mutants, pDEW13 and pDEW14, became Locked ON in a state of relatively low expression, suggesting that, when only one 5′-GATC-3′ site was intact, OxyR was able to bind, albeit with poor affinity. Expression of *agn43-lacZ* in plasmids pDEW15 and pDEW16 also

decreased significantly when OxyR was over expressed, but transcription was still very much Locked ON.

These results supported the idea that the Locked ON phenotypes of pDEW11 to pDEW16 were due to a decrease in affinity, of the mutated *agn43* regulatory regions, for OxyR. In addition, the restoration of phase variation in the single mutants, pDEW11 and pDEW12, indicated that, in the absence of this effect, only 2 intact 5'-GATC-3' sites were required for phase variation of *agn43* to occur.

Although OxyR binding to DNA is specific and occurs with high affinity, only 3 bases are conserved in the OxyR-binding regions of 7 different promoters (Toledano *et al.*, 1994). However, based on the binding of OxyR to oligonucleotides of random sequence, a consensus OxyR-binding motif has previously been proposed comprising the sequence ATAGntnnnanCTATnnnnnnnATAGntnnnanCTAT. An OxyR tetramer is thought to contact each of these conserved elements on one face of the double helix (Toledano *et al.*, 1994). Like *agn43,* the *oxyS* promoter is subject to repression by reduced OxyR protein, and this promoter has been found to have an extended version of the consensus that includes a fifth OxyR contact point (Toledano *et al.*, 1994). As shown in Figure 5.5 both this extended motif, and the motif for oxidised OxyR, can be matched when aligned with the regulatory region of *agn43.* The alignment predicts that adenine residues of the 2nd and 3rd 5′-GATC-3′ sites are likely to make contact with OxyR, due to overlap with OxyR contact elements I and II. Consistent with this hypothesis, all plasmids in which either of these residues had been mutated were Locked ON for *agn43-lacZ* transcription in CSH50.

Moreover, plasmids pDEW15 and pDEW16, the constructs in which both of these residues had been altered, had the highest level of *agn43-lacZ* expression (Tables 5.3 and 5.4). That the A-to-T mutation in site 1 of pDEW10 had no effect on phase switching is consistent with this residue being uninvolved in contact with OxyR as predicted from the alignment in Figure 5.5.

5.2.5 Disruption of 5′-GATC-3′ sites with alternative mutations

Results presented in the previous sections (5.2.3 and 5.2.4) are consistent with the suggestion that A-to-T mutations at 5'-GATC-3' sites 2 and 3 not only preclude Dam methylation but also abrogate binding of OxyR to the *agn43* regulatory region.

In an attempt to disrupt methylation of the 2nd and 3rd GATC sites without affecting OxyR binding directly, nucleotides less likely than adenine to make contact with OxyR (on the basis of the alignment in Figure 5.5) were mutated. The G of site 2 and the C of site 3 were selected for mutation since neither of these residues overlap with any of the predicted OxyR contact points (Figure 5.5). Thus, mutating these residues should inactivate these sites for Dam methylation, without disrupting the OxyR-binding motif. Accordingly 5'-GATC-3' site 2 was altered to CATC and site 3 was altered to GATG as described in section 5.2.1.1 A second series of constructs were produced that contained all possible combinations of these mutations and the A→T mutation in site 1. It will be recalled that an A-to-T mutation in 5'-GATC-3' site 1 does which not appear to affect OxyR binding (Table 5.2 and Figure 5.5).

Table 5.4 The effect of OxyR overexpression on the switch phenotypes of plasmids containing mutations in the 5'-GATC-3' sites of *agn43-lacZ*.

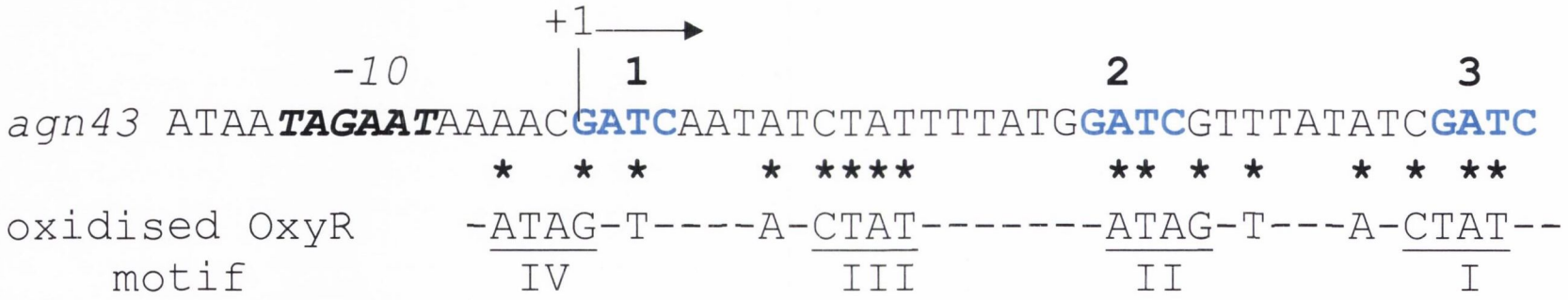
Construct	Methylation target[a]			Phenotype in CSH50 OxyR^{++}	
	1	2	3	Switch phenotype	β–Gal[b]
pAg43	GATC	GATC	GATC	Switching	145(16)
pDEW10	G**T**TC	GATC	GATC	Switching	154(19)
pDEW11	GATC	G**T**TC	GATC	Switching	152(16)
pDEW12	GATC	GATC	G**T**TC	Switching	89(9)
pDEW13	G**T**TC	G**T**TC	GATC	Locked ON	51(8)
pDEW14	G**T**TC	GATC	G**T**TC	Locked ON	61(7)
pDEW15	GATC	G**T**TC	G**T**TC	Locked ON	142(15)
pDEW16	G**T**TC	G**T**TC	G**T**TC	Locked ON	101(11)

a. Substituted bases in each construct are in bold and underlined.

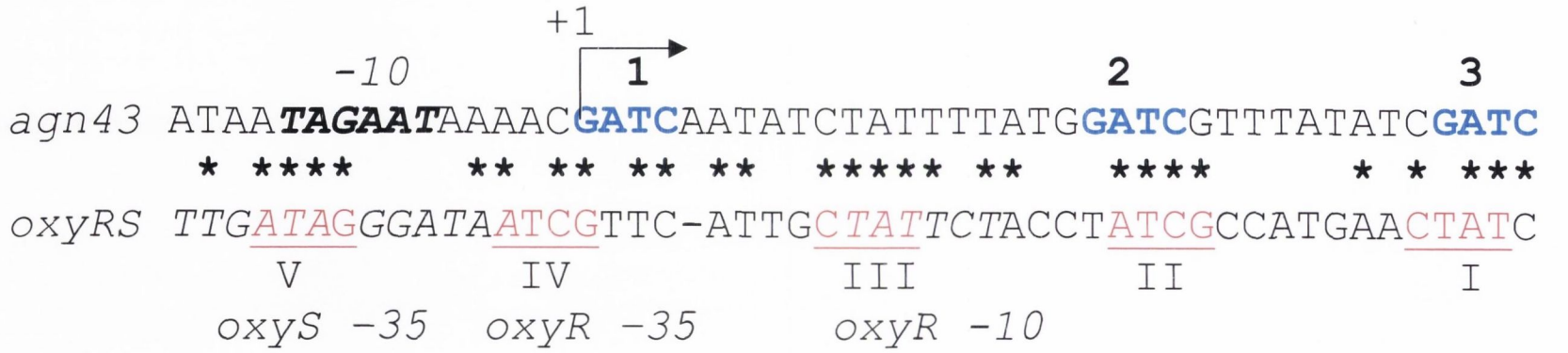
b. β-Gal, β-galactosidase activity. Assays were performed on duplicate cultures, in triplicate. Results shown are averages with standard deviation in parenthesis and are expressed in Miller units (MU). In switching phenotypes the β-galactosidase values refer to the Phase ON state.

Figure 5.5 Alignment of *agn43* regulatory region with proposed OxyR-binding motifs. 5′-GATC-3′ sites in *agn43* are in blue and are numbered. The transcriptional start site is indicated as +1 and is shown by an angled arrow. The asterisks indicate conserved residues in the aligned DNA sequences. (A) Alignment of *agn43* with proposed binding motif for Oxidised OxyR. (B) Alignment of *agn43* with *oxyRS* sequence. Underlined and labelled with Roman numerals are the nucleotides that make contact with OxyR. The –10 and -35 RNA polymerase recognition sites are in italics and are labelled. In (B), a dash (-) represents a space introduced between OxyR contact points III and IV to improve alignment with the *agn43* regulatory region. Here the OxyR contact points are highlighted in red.

A


+1
-10
1
2
3
agn43 ATAATAGAATAAAACGATCAATATCTATTTTATGGATCGTTTATATCGATC
oxidised OxyR motif
-ATAG-T----A-CTAT-------ATAG-T---A-CTAT--
IV
III
II
I

B


+1
-10
1
2
3
agn43 ATAATAGAATAAAACGATCAATATCTATTTTATGGATCGTTTATATCGATC
oxyRS TTGATAGGGATAATCGTTC-ATTGCTATTCTACCTATCGCCATGAACTATC
V
IV
III
II
I
oxyS -35
oxyR -35
oxyR -10

5.2.6 Phenotypic analysis of mutant constructs pDEW20-25 in strains CSH50 and CSH50OxyR^{++}

The effect of each combination of mutations on *agn43* regulation was assessed by analysis of plasmid constructs pDEW20-25 in the Lac$^-$ strain CSH50 (Table 5.5). Singly, the G-to-C and C-to-G mutations resulted in a phase-variable phenotype in strain CSH50 rather than the Locked ON phenotype produced by the A-to-T mutations in the same sites (Tables 5.3 and 5.7). However, β-galactosidase activity in Phase ON colonies of CSH50 containing these plasmids was (pDEW21 and pDEW22) was slightly lower than that in Phase ON colonies containing the wild type plasmid pAg43 (see Table 5.3). Overexpression of OxyR from plasmid pAQ5 did not significantly affect expression in these single mutants (Table 5.5). The restoration of phase variation resulting from the alternative mutations in pDEW20 and pDEW21 indicated that the adenine residues of the second and third 5′-GATC-3′ sites were probably important for OxyR recognition and that methylation of any two remaining sites was sufficient to allow phase variation of *agn43*. When combined with an A-to-T mutation in the first site, the G-to-C mutation in the second site (pDEW22) resulted in *agn43-lacZ* expression being locked at a low level (27 MU). This suggested that methylation of the third 5′-GATC-3′ site alone was not enough fully to block binding of OxyR and result in the Phase ON state. When OxyR was overexpressed this plasmid became Locked OFF with levels of β-galactosidase expression of <10 MU (Table 5.5).

Wild type CSH50 cells containing pDEW23 (first and third 5′-GATC-3′ sites mutated) or pDEW24 (second and third sites mutated) gave rise to colonies that appeared Lac$^-$ on L agar containing X-Gal, and β-galactosidase activity was less

than 10 MU (Table 5.5). Expression of *agn43-lacZ* was restored to its full level in an *oxyR* mutant (Table 5.6), indicating that the low level of expression in a wild type background was due to binding of singly methylated DNA by OxyR. In the final mutant, pDEW25, all three 5'-GATC-3' sites were disrupted in a manner thought unlikely to affect OxyR binding directly. (An A-to-T mutation was introduced into the first site and G-to-C and C-to-G substitutions were made in the second and third sites, respectively.) Expression of *agn43-lacZ* from this mutant plasmid in CSH50 was Locked OFF. This was reminiscent of the OFF phenotype of the *agn43* gene in a *dam* mutant. The phenotype associated with pDEW25 was consistent with OxyR binding freely to the *agn43* regulatory region in the absence of Dam-mediated methylation. As expected, overexpression of OxyR had no effect on expression of pDEW23-25 (Table 5.5), which presumably, were already fully bound, and repressed, by OxyR when this protein was present at wild type levels.

5.2.7 The effect of a *dam* mutation on mutant phenotypes

To determine whether their phenotypes were methylation dependent all constructs were examined in a *dam* strain. Initially the *dam::kan* mutation from BD1302*dam* (Table 2.1) was introduced into the *lac*$^-$ strain CSH50 by P1*cml* transduction. However, the resulting strain CSH50*dam* however could not be transformed, presumably due to the inhibition of plasmid replication by the hemimethylated DNA that must result from the first round of replication in the absence of Dam methylase (Russell and Zinder, 1987). This problem was overcome by transforming CSH50 with each of the plasmid constructs prior to introduction of the *dam* mutation. The resultant derivatives DW1-DW14 (Table 2.1) were then assessed for *agn43-lacZ* expression by means of β-galactosidase assay. The results are

Table 5.5 The switch phenotypes of pDEW20-pDEW25 in strains CSH50 and CSH50OxyR^{++}

Construct	Methylation target[a]			Phenotype in CSH50		Phenotype in CSH50OxyR^{++}	
	1	2	3	Switch phenotype	β–Gal[b]	Switch phenotype	β–Gal[b]
pDEW20	GATC	**C**ATC	GATC	Switching	125(17)	Switching	116(10)
pDEW21	GATC	GATC	GAT**G**	Switching	127(21)	Switching	121(9)
pDEW22	G**T**TC	**C**ATC	GATC	Locked ON	27(6)	Locked OFF	<10
pDEW23	G**T**TC	GATC	GAT**G**	Locked OFF	<10	Locked OFF	<10
pDEW24	GATC	**C**ATC	GAT**G**	Locked OFF	<10	Locked OFF	<10
pDEW25	G**T**TC	**C**ATC	GAT**G**	Locked OFF	0	Locked OFF	<10

a. Substituted bases in each construct are in bold and underlined.

b. β-Gal, β-galactosidase activity. Assays were performed on duplicate cultures, in triplicate. Results shown are averages with standard deviation in parenthesis and are expressed in miller units (MU). In switching phenotypes the β-galactosidase values refer to the Phase ON state.

presented in Tables 5.6 and 5.7, where each derivative is described by the name of the relevant plasmid in CSH50*dam*.

The plasmid pAg43, which is phase variable in a wildtype background, became Locked OFF for *agn43-lacZ* expression in the absence of Dam. This can be attributed to the OxyR repressor binding freely to the intact 5′-GATC-3′ sites in the absence of Dam-mediated methylation. Similarly, pDEW10 in which site 1 has been disrupted became Locked OFF in the absence of Dam. The mutation in site 1 of this plasmid was not thought to have affected the ability of OxyR to bind and block *agn43* transcription. Plasmids pDEW11-pDEW16 which contain A-to-T mutations in site 2 or site 3 were Locked ON in a wildtype background. Results had indicated that this was due to the Dam-independent disruption of the OxyR binding motif. Consistent with this idea, the Locked ON phenotype of each plasmid was unaffected by introduction of a *dam* mutation. This indicates that OxyR is unable to bind when adenines in the second or third 5′-GATC-3′ sites have been mutated, even in the absence of any methylation (e.g. of site 1). The phase variable plasmids pDEW20 and pDEW21, like the wildtype plasmid, were Locked OFF in a *dam* background presumably due to the unperturbed binding of OxyR to an intact and unmethylated binding motif. The construct pDEW22 was Locked ON at a very low level in CSH50. This transcription was abolished in the absence of methylation indicating that it had been due to Dam-dependent disruption of OxyR repression. Mutations in the plasmids pDEW23, pDEW24 and pDEW25 are thought to abolish methylation without disrupting the OxyR binding motif. As anticipated, these constructs remained Locked OFF for *agn43* transcription in the absence of Dam due presumably to unperturbed binding of the OxyR repressor.

Table 5.6 The effects of mutations in the 5'-GATC-3' sites of *agn43-lacZ* on switch phenotypes and β-galactosidase activity in strains CSH50*dam* and CSH50*ΔoxyR*.

Construct	Methylation target[a]			Phenotype in CSH50*dam*		Phenotype in CSH50*ΔoxyR*	
				Switch phenotype	β–Gal[b]	Switch phenotype	β–Gal[b]
pAg43	GATC	GATC	GATC	Locked OFF	<10	Locked ON	188(19)
pDEW10	G**T**TC	GATC	GATC	Locked OFF	< 10	Locked ON	203(20)
pDEW11	GATC	G**T**TC	GATC	Locked ON	163(17)	Locked ON	198(21)
pDEW12	GATC	GATC	G**T**TC	Locked ON	156(15)	Locked ON	196(22)
pDEW13	G**T**TC	G**T**TC	GATC	Locked ON	121(10)	Locked ON	205(21)
pDEW14	G**T**TC	GATC	G**T**TC	Locked ON	126(13)	Locked ON	206(23)
pDEW15	GATC	G**T**TC	G**T**TC	Locked ON	188(20)	Locked ON	201(22)
pDEW16	G**T**TC	G**T**TC	G**T**TC	Locked ON	189(22)	Locked ON	186(20)
pDEW20	GATC	**C**ATC	GATC	Locked OFF	<10	Locked ON	198(19)
pDEW21	GATC	GATC	GAT**G**	Locked OFF	<10	Locked ON	191(18)
pDEW22	G**T**TC	**C**ATC	GATC	Locked OFF	<10	Locked ON	204(21)
pDEW23	G**T**TC	GATC	GAT**G**	Locked OFF	<10	Locked ON	203(22)
pDEW24	GATC	**C**ATC	GAT**G**	Locked OFF	<10	Locked ON	207(23)
pDEW25	G**T**TC	**C**ATC	GAT**G**	Locked OFF	<10	Locked ON	180(21)

a. Substituted bases in each construct are in bold and underlined.

b. β-Gal, β-galactosidase activity. Assays were performed on duplicate cultures, in triplicate. Results shown are averages with standard deviation in parenthesis and are expressed in Miller units (MU). In switching phenotypes the β-galactosidase values refer to the Phase ON state.

Table 5.7 Switch phenotypes of mutant constructs in CSH50, CSH50OxyR^{++}, CSH50*dam* and CSH50*ΔoxyR*

Construct	Methylation target[a]			Phenotype in CSH50 OxyR^{++}		Phenotype in CSH50OxyR^{++}		Phenotype in CSH50*dam*		Phenotype in CSH50ΔoxyR	
	1	2	3	Switch	β–Gal[b]	Switch	β–Gal[b]	Switch	β–Gal[b]	Switch	β–Gal[b]
pAg43	GATC	GATC	GATC	Switching	156(18)	Switching	145(16)	Locked OFF	<10	Locked ON	188(19)
pDEW10	G**T**TC	GATC	GATC	Switching	176(18)	Switching	154(19)	Locked OFF	< 10	Locked ON	203(20)
pDEW11	GATC	G**T**TC	GATC	Locked ON	180(23)	Switching	152(16)	Locked ON	163(17)	Locked ON	198(21)
pDEW12	GATC	GATC	G**T**TC	Locked ON	164(17)	Switching	89(9)	Locked ON	156(15)	Locked ON	196(22)
pDEW13	G**T**TC	G**T**TC	GATC	Locked ON	124(15)	Locked ON	51(8)	Locked ON	121(10)	Locked ON	205(21)
pDEW14	G**T**TC	GATC	G**T**TC	Locked ON	119(11)	Locked ON	61(7)	Locked ON	126(13)	Locked ON	206(23)
pDEW15	GATC	G**T**TC	G**T**TC	Locked ON	205(21)	Locked ON	142(15)	Locked ON	188(20)	Locked ON	201(22)
pDEW16	G**T**TC	G**T**TC	G**T**TC	Locked ON	193(16)	Locked ON	101(11)	Locked ON	189(22)	Locked ON	186(20)
pDEW20	GATC	**C**ATC	GATC	Switching	125(17)	Switching	116(10)	Locked OFF	<10	Locked ON	198(19)
pDEW21	GATC	GATC	GAT**G**	Switching	127(21)	Switching	121(9)	Locked OFF	<10	Locked ON	191(18)
pDEW22	G**T**TC	**C**ATC	GATC	Locked ON	27(6)	Locked OFF	<10	Locked OFF	<10	Locked ON	204(21)
pDEW23	G**T**TC	GATC	GAT**G**	Locked OFF	<10	Locked OFF	<10	Locked OFF	<10	Locked ON	203(22)
pDEW24	GATC	**C**ATC	GAT**G**	Locked OFF	<10	Locked OFF	<10	Locked OFF	<10	Locked ON	207(23)
pDEW25	G**T**TC	**C**ATC	GAT**G**	Locked OFF	0	Locked OFF	<10	Locked OFF	<10	Locked ON	180(21)

a. Substituted bases in each construct are in bold and underlined.

b. β-Gal, β-galactosidase activity. Assays were performed on duplicate cultures, in triplicate. Results shown are averages with standard deviation in parenthesis and are expressed in Miller units (MU). In switching phenotypes the β-galactosidase values refer to the Phase ON state.

5.2.8 The effect of an *oxyR* mutation on mutant phenotypes

To determine whether the decreased transcription seen in a number of the mutants was due to binding and repression by OxyR, the phenotype of each mutant was assessed in the strain CSH50*ΔoxyR*. In this strain expression of all constructs was restored to a level comparable with that of wild type as shown in Tables 5.6 and 5.7. This indicated that the low levels of expression observed for plasmids pDEW22-pDEW24 was due to OxyR binding to DNA, which was probably methylated at the single remaining site. These results also indicate that the Locked OFF phenotype of the triple mutant pDEW25 in CSH50 (Table 5.5) was due to OxyR binding freely to an intact motif in the absence of any Dam methylation.

5.2.9 Binding of OxyR *to agn43* DNA

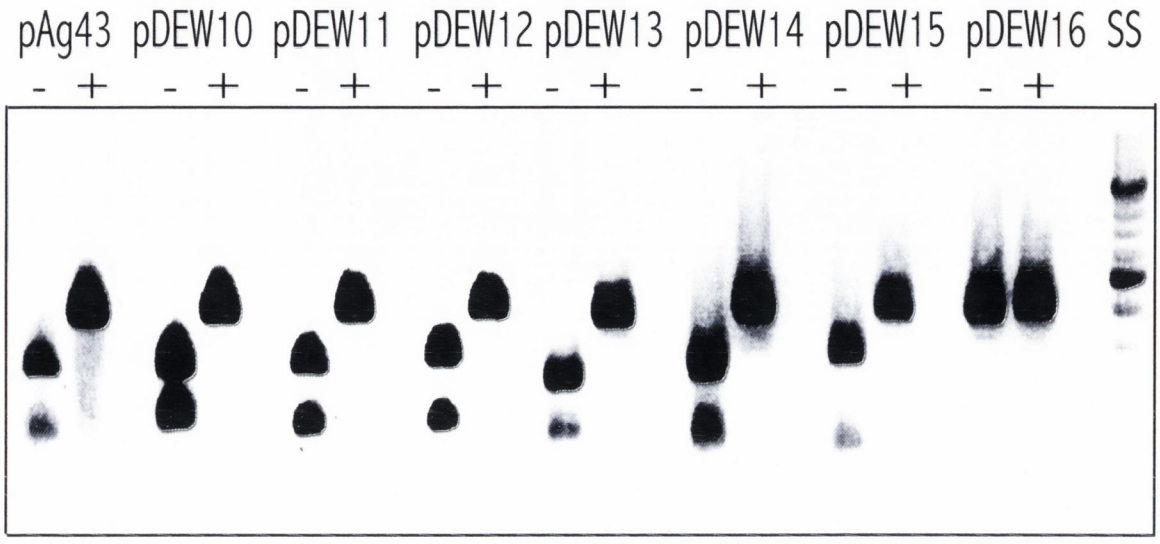
In the previous sections it was suggested that the Locked ON phenotypes of mutant plasmids was due to disruption of the OxyR binding motif, and was therefore Dam-independent. It was also thought possible that the Locked OFF phenotype, which was displayed by two constructs (pDEW23 and pDEW24), was caused by OxyR binding to DNA that was methylated at one 5'-GATC-3' site. To explore these possibilities, the ability of OxyR to bind the mutated *agn43* regulatory region from each mutant construct was assessed. This was achieved as described below.

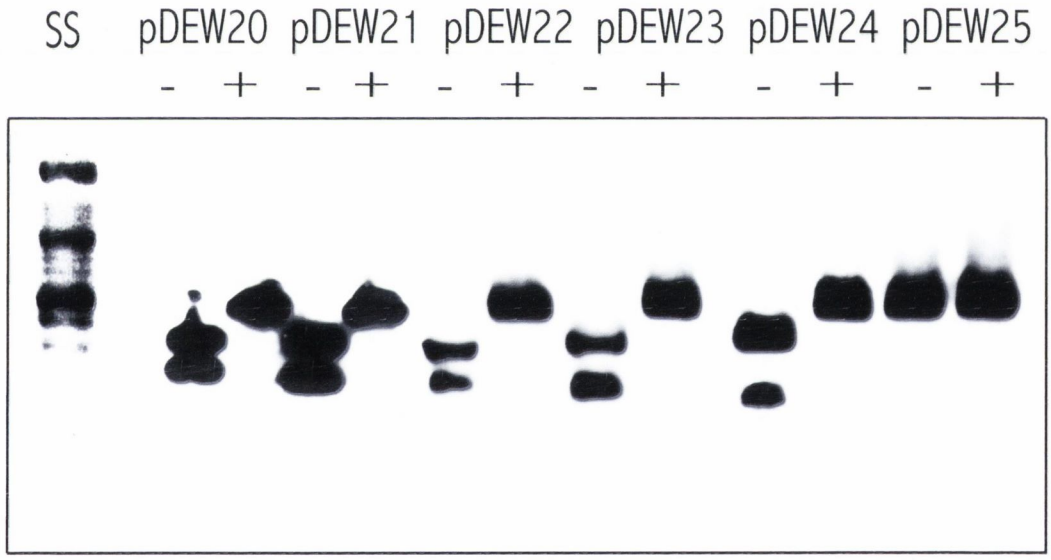
The binding of OxyR to *agn43* regulatory region was examined by means of EMSA. The inserts from all plasmid constructs, with mutations corresponding to different methylation patterns, were amplified by PCR. 373-bp probes, corresponding to the entire cloned *agn43* regulatory region were used and were amplified from each of the plasmid constructs using the primers Ag43R and Ag43F

(Table 2.2). Aliquots were then Dam methylated *in vitro* by incubation with Dam methylase and AdoMet. PCR products were then treated with the restriction enzyme *Mbo*I, which cuts only at unmethylated 5′-GATC-3′ sites. Resistance to digestion by *Mbo*I indicated that probes were methylated at all intact 5′-GATC-3′ sites. Methylated DNA was protected from *Mbo*I digestion while each unmethylated product was cleaved to give two visible DNA fragments (Figure 5.6). This is because the three 5′-GATC-3′ sites are close together in the middle of the 373-bp probe and only two fragments were resolvable by electrophoresis after cleavage with *Mbo*I. As anticipated, DNA amplified from constructs containing one or more intact 5′-GATC-3′ sites generated two bands before *Mbo*I digestion and a single band after *Mbo*I digestion. However, unlike DNA from other mutant constructs, that amplified from pDEW16 and pDEW25 was resistant to cleavage by *Mbo*I both before and after incubation with Dam methylase (see Figure 5.6). This is because 5'-GATC-3' is the target site for both *Mbo*I AND Dam methylase. Thus, elimination of these sites by mutation resulted in resistance to cleavage by *Mbo*I and to methylation by Dam.

Both methylated and unmethylated DNA probes amplified from each construct were end-labelled with γ^{32}P dATP. The probes were then incubated with extracts from BD1302*dam* or from IPTG-treated BD1302*dam* containing the OxyR overexpressing plasmid pAQ25. Dam$^-$ strains were used for the production of crude extracts to ensure that the methylation state of probes was unchanged during EMSA. Furthermore, BD1302*dam* does not express functional OxyR, due to a mutation in the *oxyR* gene. The plasmid pAQ25 contains the *oxyR* gene under the

Figure 5.6 Assessment of the methylation state of DNA probes. *agn43* DNA amplified from each construct was digested with *Mbo*I before (-) or after (+) incubation with Dam methylase. Fragments were resolved by agarose electrophoresis. Positions of size standards are indicated.


pAg43 pDEW10 pDEW11 pDEW12 pDEW13 pDEW14 pDEW15 pDEW16 SS
- + - + - + - + - + - + - + - + - +


SS pDEW20 pDEW21 pDEW22 pDEW23 pDEW24 pDEW25
- + - + - + - + - + - + - +

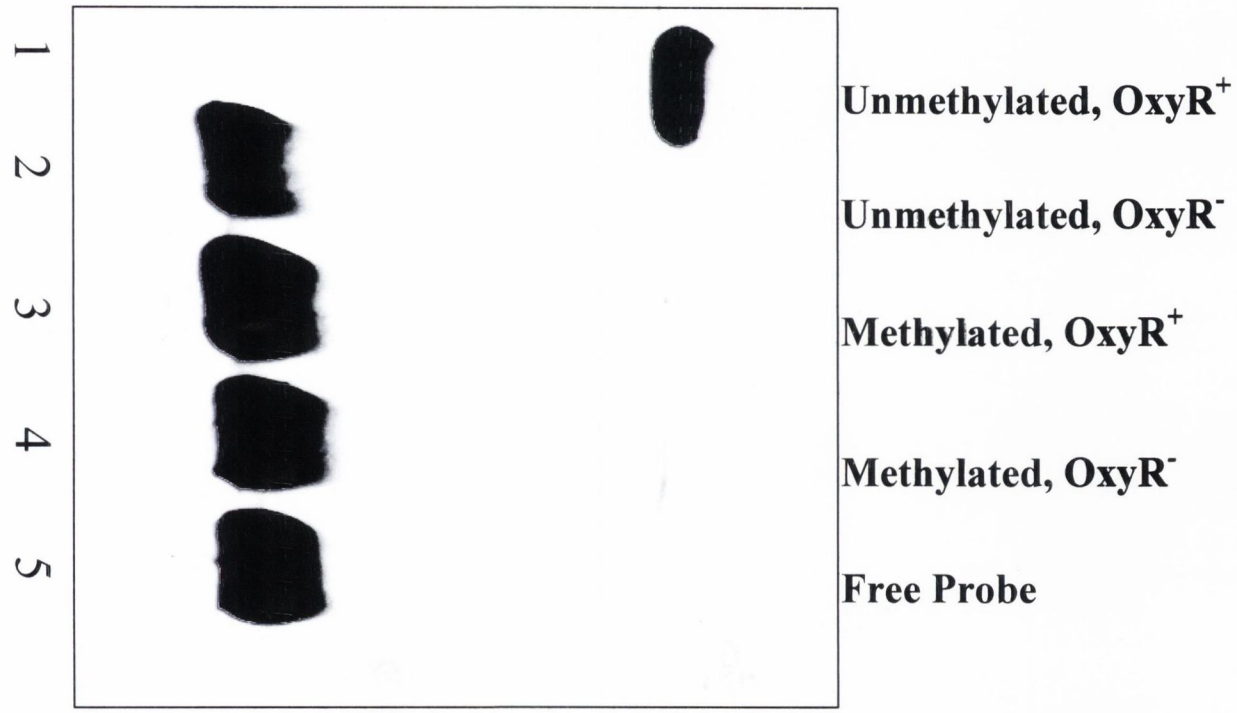
control of an IPTG-inducible P_{tac} promoter. Thus, the use of these strains ensured that any observed DNA retardation could be attributed only to binding by OxyR

5.2.9.1 Binding of OxyR to the wildtype *agn43* regulatory region

Before examining the binding of OxyR to DNA from mutant plasmids, it was necessary to ensure that OxyR could bind the wild type *agn43* regulatory region *in vitro*. Therefore probes spanning the entire cloned *agn43* regulatory region were amplified from the wild type reporter plasmid, pAg43. Both unmethylated and methylated probes were incubated in the presence or absence of OxyR. Their gel mobility was compared to that of free probe, which had not been incubated with any cell extract.

For wild type *agn43* DNA a shift was observed only with unmethylated DNA in the presence of OxyR (Figure 5.7, lane 1). No shift was observed when fully methylated DNA was incubated in the presence of OxyR (Figure 5.7, lane 3). This confirmed that Dam-mediated methylation of 5′-GATC-3′ sites in the *agn43* regulatory region prevents binding of OxyR *in vitro*. There was no retardation when either unmethylated or methylated DNA were incubated with extracts from OxyR$^-$ cultures (Figure 5.7, lanes 2 and 4). This indicated that no components of the cell extracts (other than OxyR) were capable of binding and shifting the probes, under the conditions of the experiment. These results provide direct evidence that OxyR binds to unmethylated *agn43* DNA and that methylation of three 5'-GATC-3' sites within the *agn43* regulatory region inhibits this binding.

Figure 5.7 The binding of OxyR to methylated and unmethylated *agn43* DNA was examined using EMSA. Lane 1, unmethylated DNA with $OxyR^+$ extract; Lane 2, unmethylated DNA with $OxyR^-$ extract; Lane 3, methylated DNA with $OxyR^+$ extract; Lane 4, methylated DNA, $OxyR^-$ extract; Lane 5, free probe, no extract.


1
2
3
4
5
Unmethylated, OxyR+
Unmethylated, OxyR-
Methylated, OxyR+
Methylated, OxyR-
Free Probe

5.2.9.2 Binding of OxyR to unmethylated DNA from mutant constructs

Unmethylated DNA was amplified from each mutant using primers Ag43R and Ag43F and examined for binding by OxyR *in vitro* (Figure 5.8). Unmethylated DNA from wild type pAg43, and from single mutants pDEW10, pDEW20 and pDEW21 was bound indicating that point mutations in these constructs do not have a direct effect on OxyR binding. This was expected since these constructs were all able to undergo OxyR-dependent phase variation (Table 5.7).

Unmethylated DNA from the double mutants, pDEW22, pDEW23 and pDEW24, and from the triple mutant pDEW25, which were Locked OFF for *agn43* transcription, was also bound by OxyR *in vitro*. That the G to C mutation in site 2 and C to G mutation in site 3 had no effect on OxyR binding indicates that these residues are not involved in OxyR binding. Thus, the Locked OFF phenotypes of pDEW22-25 and the Phase OFF state of pDEW20 and pDEW21 was most likely due to OxyR binding to an intact recognition motif. Only DNA from constructs containing mutations in the adenines of the 2nd or 3rd 5′-GATC-3′ sites (pDEW11-pDEW16) remained unbound in the presence of OxyR. This indicates that, as predicted from the alignment, these residues are part of the OxyR binding motif.

OxyR is subject to oxidation during preparation of cell extracts. Thus these experiments examine the ability of the mutated *agn43* regulatory regions to bind the oxidized form of the protein. However, the mutations that affected OxyR binding were in regions of *agn43* predicted to make contact with both the oxidized and reduced forms of the DNA-binding protein. As such, mutations in the adenines of 5′-GATC-3′ sites 2 and 3 would also be expected to eliminate binding of the

reduced (repressive) form of OxyR. Hence, these results confirmed that the Locked ON phenotype of constructs pDEW11–pDEW16, which contain mutations in adenines of 5′-GATC-3′ site 2 or 3, resulted from the methylation-independent disruption of OxyR binding.

5.2.9.3 Binding of OxyR to methylated DNA from mutant constructs

To determine which of the possible *agn43* methylation patterns can block binding of OxyR, EMSAs were performed using methylated DNA from each mutant construct (Figure 5.9). Methylated DNA amplified from plasmids pAg43, pDEW10-pDEW16, pDEW20 and pDEW21 remained unbound in the presence of OxyR. OxyR was unable to bind to DNA from pDEW11-pDEW16, even in its unmethylated state (Figure 5.8). Therefore the lack of binding to DNA from these mutants cannot be attributed to Dam methylation of the probes.

The single mutants pDEW10, pDEW20 and pDEW21 had been bound by OxyR when DNA was unmethylated, indicating that the OxyR binding motif, in each of these plasmids, is intact (Figure 5.8). However, methylated DNA from these single mutants was not retarded in the presence of OxyR (Figure 5.9) showing that methylation of any 2 remaining 5′-GATC-3′ sites is sufficient to block binding of OxyR (Figure 5.9). This explains the observation that pDEW10, pDEW20 and pDEW21 could phase-variably express *agn43-lacZ* in and $OxyR^+$ strain (Table 5.7).

DNA from the double mutants pDEW22-24 was shifted by OxyR in the methylated state. This showed that OxyR can actually bind to *agn43* DNA that is methylated at any of the three 5′-GATC-3′ sites singly. *agn43-lacZ* activity in pDEW22,

pDEW23 and in pDEW24 was relatively low in an OxyR$^+$ strain and full expression was restored in an *oxyR* background (Table 5.7). It is likely that binding of singly methylated DNA, in pDEW22, pDEW23 and pDEW24, by OxyR is responsible for the dramatic reduction in *agn43* transcription in a wildtype background. *agn43-lacZ* expression in pDEW22 was Locked ON at a very low level (compared to wildtype) whereas pDEW23-25 were completely Locked OFF. This suggests that OxyR binds less freely to pDEW22, which is only methylated at 5′-GATC-3′ site 3, than to the unmethylated Locked OFF triple mutant, pDEW25.

Both unmethylated and "methylated" DNA from the Locked OFF triple mutant pDEW25 were also bound by OxyR (Figures 5.8 and 5.9). Of course, this construct could not actually be methylated as all three Dam methylation targets had been eliminated. The binding of pDEW25 by OxyR confirms that the A residue of 5′-GATC-3′ site 1, the G of 5′-GATC-3′ site 2 and the C of 5′-GATC-3′ site 3 are not important residues for DNA recognition by OxyR.

From these results it would appear that methylation of two or more 5′-GATC-3′ sites is required to block binding of OxyR to the *agn43* regulatory region. This was to be expected since only constructs containing at least two intact 5′-GATC-3′ sites were able to phase-variably express *agn43-lacZ* at close to full level in the OxyR$^+$ Dam$^+$ strain, CSH50 (Table 5.7).

Figure 5.8 The binding of OxyR to unmethylated *agn43* DNA. Electrophoretic mobility shift assays were conducted as described in Chapter 2. Amplified DNA from each construct was incubated in the absence of any extract (lanes indicated by *), in the presence of extracts from CSH50*ΔoxyR* (lanes indicated by -) or in the presence of extracts from OxyR overproducing strain CSH50 containing pAQ25 (lanes indicated by +). The numbers 43 and 10-25 indicate DNA amplified from pAg43 and pDEW10-pDEW25, respectively.

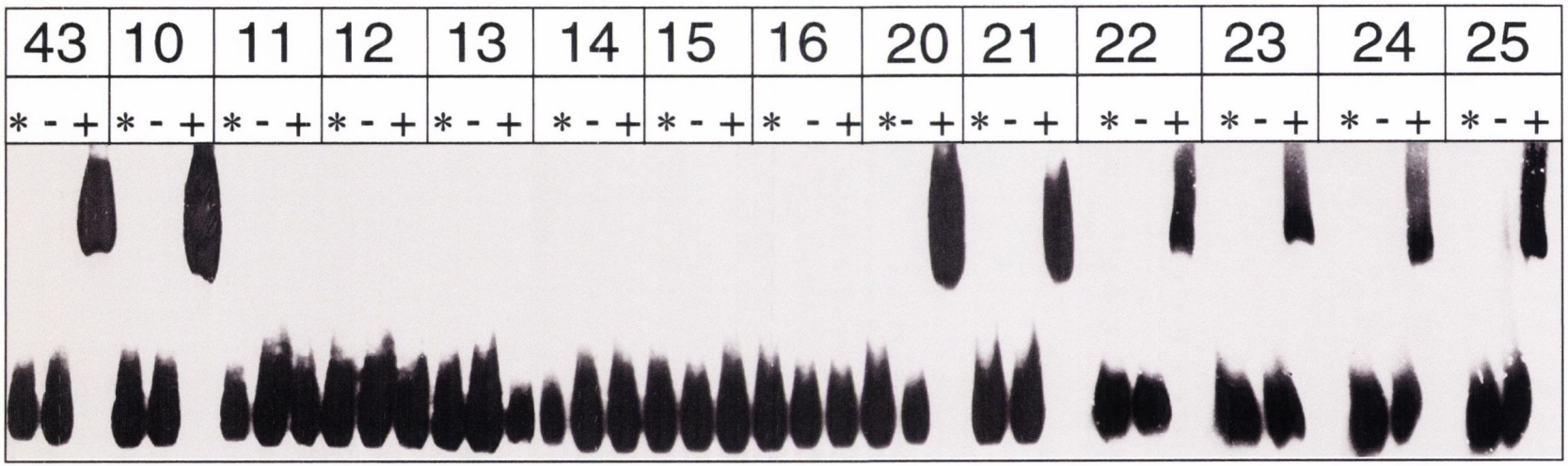
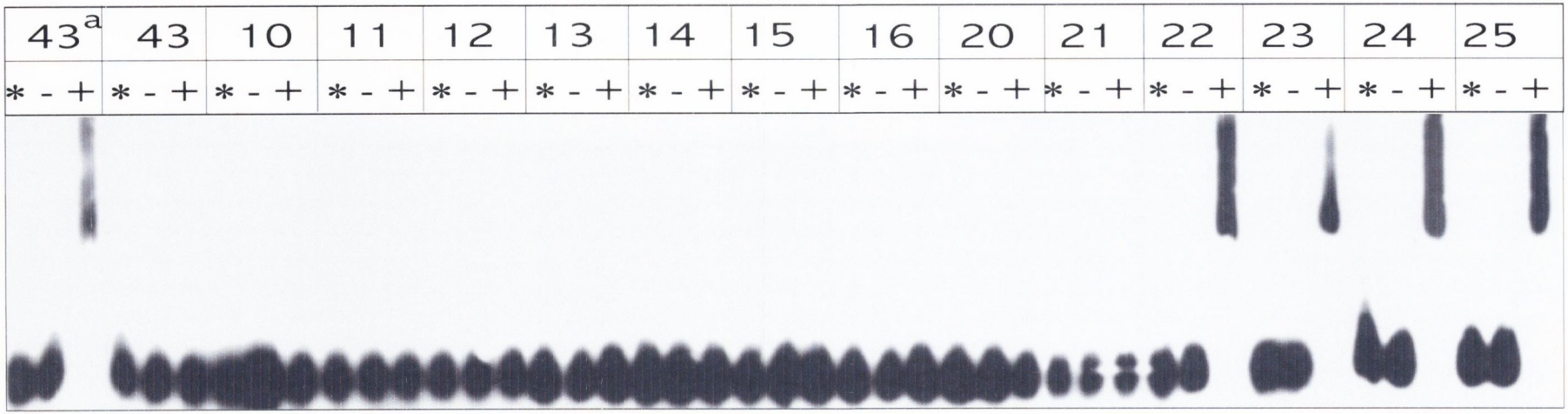

43
10
11
12
13
14
15
16
20
21
22
23
24
25
* - +

Figure 5.9 The binding of OxyR to methylated *agn43* DNA. Amplified DNA from each construct was incubated in the absence of any extract (lanes indicated by *), in the presence of extracts from CSH50*ΔoxyR* (lanes indicated by -) or in the presence of extracts from OxyR-overproducing strain CSH50 containing pAQ25 (lanes indicated by +). The numbers 43 and 10-25 indicate DNA amplified from pAg43 and pDEW10-pDEW25, respectively.

[a] Unmethylated DNA from pAg43 was included as a positive control.


43[a]
43
10
11
12
13
14
15
16
20
21
22
23
24
25
* - +

5.2.10 Footprinting analysis of OxyR binding

Footprinting analysis was conducted in an attempt to further characterise the binding of OxyR to the *agn43* regulatory region *in vivo*. Genomic DNA isolated from strains CSH50 and CSH50*ΔoxyR* was treated with DMS, which methylates DNA at A residues, and was used as a template for the extension of primers annealing to each strand of the *agn43* regulatory region. Incubation of DMS-methylated DNA at 95°C causes breakage of the unstable glycosidic bonds of methylated deoxynucleotides, resulting in the generation of DNA fragments which, when used as a template for primer extension, give rise to a ladder of DNA bands. Bound protein may protect certain residues from methylation by DMS resulting in the loss of the corresponding fragments from the ladder of primer extension products, resulting in a visible footprint.

Despite numerous attempts, over the course of several months, these studies proved to be inconclusive. Although DNA was prepared from a Phase OFF culture of the $OxyR^+$ strain, and was therefore probably bound by OxyR, no apparent footprint was obtained. A possible explanation for this is that the large OxyR tetramer does not bind tightly enough to the DNA helix to block access of the small DMS molecule to the *agn43* regulatory region. On one occasion results were suggestive of DNA protection at 5′-GATC-3′ site 2. However these results were not reproducible and for this reason footprinting studies were discontinued.

5.3 Discussion

It is thought that Dam methylates 5'-GATC-3' sites by a processive mechanism since closely grouped sites are methylated in succession (Urig *et al.*, 2002). Some sites do, however, appear to be methylated preferentially (Bergerat *et al.* 1989). In this respect it is interesting to note that, although the tetranucleotide 5′-GATC-3′ is the consensus target sequence for Dam methylase, there is evidence that bases flanking this sequence are involved in the recognition event. It has also been suggested that any asymmetry introduced by flanking sequences to the symmetrical 5′-GATC-3′ may give some orientation to the interaction of Dam with the DNA and may influence the order in which nearby sites are methylated (Bergerat *et al.*, 1989). Thus each of the three *agn43* 5′-GATC-3′ sites may not be equally important in phase variation. It has previously been reported that a correlation exists between the proportion of Phase ON cells and the proportion of cells having methylated, or hemimethylated, *agn43* 5′-GATC-3′ sites, in a population (Haagmans and van der Woude, 2000). However, the 5′-GATC-3′ sites were never analysed individually and it was unclear was not how many sites were necessary in order for phase variation of *agn43* to occur.

In this study the role of individual 5'-GATC-3' sites within the *agn43* regulatory region was investigated by mutation of critical adenine residues. Dam can bind non-specifically to DNA but does so with only 1% of the affinity with which it binds to the sequence 5'-GATC-3'. Dam, however, is not capable of methylating any residues other than the adenines of these sites (Theiking *et al.*, 1997). Thus, mutation of the adenine residues rendered the sites inactive for specific binding and methylation by Dam, due to disruption of the Dam target sequence.

In the current study, an A-to-T substitution in the first 5'-GATC-3' of the *agn43* regulatory region (pDEW10) had no detectable effect on phase variation of *agn43-lacZ*. This indicates that the first 5'-GATC-3' site is not required for phase variation of *agn43* and suggested that methylation of the 2nd and 3rd sites alone is sufficient to block binding of OxyR protein resulting in the Phase ON state. This idea was supported by results of EMSA experiments, which confirmed that OxyR was unable to bind *in vitro* to DNA that was methylated at sites 2 and 3.

Phase-variation of *agn43-lacZ* was abolished by A-to-T substitutions in the 2nd or the 3rd 5′-GATC-3′ sites individually or when these occurred in combination with similar substitutions at another site. However, these mutations involving the 2nd and 3rd 5′-GATC-3′ sites always resulted in a Locked ON state and not in the Locked OFF state one might expect to result from elimination of the Dam target sequences. It will be recalled that Dam is required for Ag43 expression in an $OxyR^+$ background such as CSH50 (Henderson and Owen, 1999). That the Locked ON phenotype was displayed by the triple mutant pDEW16, which contains no Dam targets, indicated that it results from a Dam-independent regulatory effect. The effects of the GATC to G**T**TC mutations on *agn43* transcription observed in pDEW11-16 are highly reminiscent of the Dam-independent effects observed after mutation of the corresponding nucleotides in the regulatory region of the *mom* gene of phage Mu (Seiler *et al.*, 1986). The regulatory regions of *mom* and *agn43* show extensive nucleotide sequence homology (Henderson and Owen, 1999) with successful transcription of the *mom* gene requiring the Dam methylation of three 5′-GATC-3′ sites within its OxyR-binding region (Kahmann, 1983; Plasterk *et al.*,

1983). These 5′-GATC-3′ sites have a spatial distribution almost identical to that of the corresponding sites within the *agn43* promoter region (Henderson and Owen,1999; see Section 1.5.4).

It seemed likely that theA-to-T mutations at sites 2 and 3 directly affect the ability of OxyR to bind to the *agn43* regulatory region and probably do so by interfering with sequences important for recognition by the protein. Although OxyR binding is specific and occurs with high affinity, only 3 bases are conserved in the OxyR-binding regions of 7 different promoters (Tartaglia *et al.*, 1992). However, an OxyR-binding motif has previously been proposed based on the binding of OxyR to oligonucleotides of random sequence. OxyR is thought to bind the DNA at four ATAG motifs positioned on one side of the double helix (Toledano *et al.*, 1994). As shown in Figure 5.5, this motif can be matched at 16 bases out of 20 when aligned with the regulatory region of *agn43*. Importantly, this alignment predicts that the adenine residues of the 2nd and 3rd 5'-GATC-3' sites make contact with OxyR. This provides a convincing explanation for the Locked ON phenotype of constructs in which either or both of these residues have been mutated. The importance of these residues in OxyR recognition was confirmed by EMSAs, which showed that unmethylated *agn43* DNA from Locked ON mutants pDEW11-pDEW16 was unable to bind OxyR. The significant drop in *agn43-lacZ* expression from these 'Locked ON' mutants observed when OxyR was made more available to the DNA (via pAQ5), is also consistent with the importance of these adenines in OxyR recognition.

It seems that an increase in the availability of OxyR revealed, at least partially, the hidden effects of inactivation of the 5'-GATC-3' sites as Dam targets. In an $OxyR^{++}$ background, previously Locked ON constructs pDEW11 and pDEW12, bearing single mutations, were phase variable for *agn43-lacZ* expression. This suggested that inactivation of any methylation target singly is not sufficient to abolish phase variation. It also indicated that methylation of any two sites alone is sufficient to block binding of the OxyR transcriptional repressor. Based on the alignment in Figure 5.5 it seems likely that OxyR makes actual contact with the 2^{nd} and 3^{rd} methylation sites themselves. It is understandable that the protrusion of a bulky $-CH_3$ group from the actual region of the DNA which the protein contacts, would seriously hinder its ability to bind.

In an $OxyR^{++}$ background the previously Locked ON plasmid pDEW14, in which only site 2 is intact, became Locked in a state of considerably lower expression. This suggested that methylation of the second site alone is not sufficient to fully block binding of OxyR since if it were it should result in the Phase ON state of *agn43* expression. It is likely that the lower level of β-galactosidase expression observed was caused by the mutation of adenine in site 3 resulting in disruption of the OxyR-binding motif. Overexpression of OxyR may not have been sufficient to fully alleviate this effect and a minority of the multiple copies of *agn43-lacZ* may have remained unbound and been transcribed.

Similarly, expression of *agn43* in pDEW13, which contains only the third 5'-GATC-3' site intact, was significantly lower in the presence of excess OxyR. Again, this suggested that methylation of one site alone is not sufficient to block

OxyR and allow the phase variation of *agn43*. There was, however, still a low level of β-galactosidase activity in the $OxyR^{++}$ background and as such these mutants could not be considered fully Locked OFF. It is conceivable that OxyR binds freely to the *agn43* regulatory region when one site is methylated, but does so with low affinity in pDEW13 and pDEW14 due to disruption of its binding motif.

Since adenines of the 2nd and 3rd 5′-GATC-3′ sites are probably involved in OxyR recognition it is not surprising that pDEW15 in which both of these residues have been abolished, had the highest level of *agn43-lacZ* expression. Expression in this mutant was decreased significantly by the overexpression of OxyR, as it had been in the other Locked ON mutants. However, the *agn43-lacZ* transcription in pDEW15 was so high in a wildtype background that, despite the significant decrease, expression was still very much Locked ON in the $OxyR^{++}$ background. This decrease in expression was not enough, perhaps, to reveal the 'true' effects that elimination of methylation at these sites would have. However, this does reflect the importance of the adenines of the 2nd and 3rd sites in allowing repression by OxyR. Similarly pDEW16, in which adenines of all three 5′-GATC-3′ sites were abolished, had a very high level of *agn43-lacZ* expression. Expression of *agn43* in the triple mutant was significantly lower in an $OxyR^{++}$ background than that of the double mutant, suggesting that the negative effect of motif disruption on OxyR binding, was counteracted somewhat by the additional mutation in 5′-GATC-3′ site 1.

The results of studies involving a multicopy plasmid encoding the *ahp* gene of *S. typhimurium*, which is bound by OxyR, have suggested that OxyR can be titrated by multiple copies of its binding site (Francis *et al.*, 1997). The combination of a

reduction in the affinity of OxyR for its mutated *agn43* binding site and its titration by the multiple copies of this site is probably responsible for the high level of expression in constructs pDEW11-pDEW16. However, the fact that wildtype pAg43 can switch to the Phase OFF state indicates that when the OxyR-binding site is intact, titration of OxyR by the low copy number plasmids used in this study does not limit its binding. However, in cells containing pDEW11-pDEW16, it is possible that titration of OxyR prevents it from binding to all copies of the mutated *agn43* site, even though its binding is probably not blocked by methylation. The substantial decrease in *agn43-lacZ* expression in cells containing these plasmids when OxyR was over-expressed, is highly supportive of this.

In an attempt to inactivate the *agn43* 5′-GATC-3′ sites for Dam-mediated methylation without affecting OxyR binding directly nucleotides less likely than adenine to make contact with OxyR were mutated. Both the G-to-C mutation in site 2 of pDEW20 and the C-to-G mutation in site 3 of pDEW21 resulted in a phase variable phenotype rather than the Locked ON phenotype produced by A to T mutations in the same sites. The restoration of phase variation resulting from these mutations is consistent with the Locked ON phenotype of constructs pDEW11-pDEW16 being a methylation-independent effect. The mutations in sites 2 and 3 of pDEW20 and pDEW21 should inactivate these sites for methylation as affectively as A-to-T mutations. As discussed earlier, an A-to-T mutation in 5′-GATC-3′ site 1 (pDEW10) had no effect on phase variation of *agn43-lacZ*. Furthermore, methylated DNA from this single mutant did not bind OxyR *in vitro*. The phase variable phenotype also displayed by single mutants pDEW20 and pDEW21 suggested that methylation of any two sites alone is sufficient to block binding of

OxyR. This was confirmed by results of EMSAs, which showed that OxyR was unable to bind to methylated DNA from any of these single mutants.

When the A-to-T mutation in 5'-GATC-3' site 2 of pDEW13 was replaced with a G-to-C mutation (pDEW22) *agn43-lacZ* became locked in a state of low expression in a strain containing only wild-type levels of OxyR. As mentioned earlier, the G-to-C mutation should inactivate this site for Dam-mediated methylation as effectively as an A-to-T mutation. The 5-fold difference in *agn43-lacZ* activity between pDEW13 and pDEW22 therefore seemed likely to be a methylation-independent effect again supporting the idea that an A-to-T mutation in site 2 affects the ability of OxyR to repress *agn43* transcription. The low level of β-galactosidase activity in cells containing pDEW22 appeared to indicate that methylation of the remaining third 5'-GATC-3' site alone, is not sufficient to block OxyR binding completely and result in the Phase ON state. There was however still a low level of *agn43-lacZ* expression, which must be accounted for. Low β-galactosidase expression in a colony could result from a minority of cells being Phase ON for *agn43* expression. Alternatively, it could result from all cells in the colony expressing *agn43-lacZ* at a uniform, low level. Low levels of β-galactosidase activity in an individual cell could result from a minority of the multiple copies of *agn43-lacZ* being fully expressed or from all copies of *agn43-lacZ* being expressed at a low level. If this effect results from all copies of the gene being expressed at a low level it would indicate that the repression of *agn43* by OxyR is not an 'all or nothing' effect. The β-galactosidase activity in colonies of cells containing pDEW22 may be due to OxyR binding to the methylated *agn43* regulatory region in such a way that it does not fully block RNA polymerase's binding to, or perhaps movement along, the

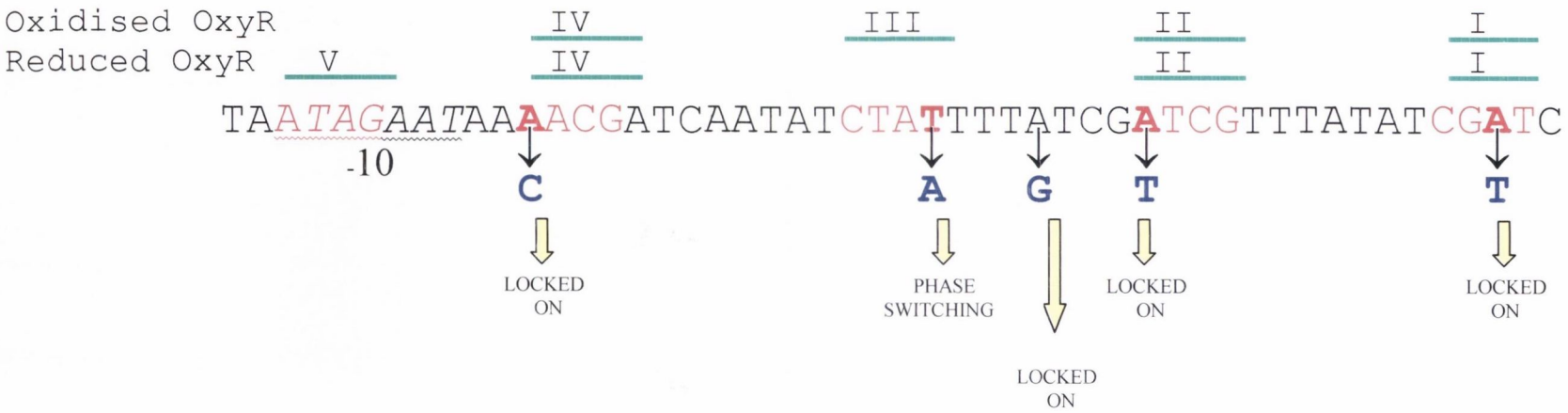
DNA. In EMSA experiments methylated DNA from pDEW22 was fully bound by OxyR. However, the phenotype of pDEW22 suggests that *in vivo* this binding is limited (compared to that in the Locked OFF double mutants), and does not fully block transcription.

OxyR exists almost exclusively in the reduced state under normal growth conditions; under conditions of oxidative stress, OxyR undergoes reversible formation of a disulphide bond, allowing it to activate transcription of genes required for oxidative stress defence (Zheng *et al.*, 1998; Åslund *et al.*, 1999). The *agn43* gene is repressed when its regulatory region binds reduced OxyR, a feature it shares with the regulatory regions of *mom* and *oxyRS*, the other members of the OxyR regulon that are repressed by reduced OxyR (Toledano *et al.*, 1994). Both oxidized and reduced OxyR can bind the *agn43* regulatory region *in vitro* and the reduced form of the protein has been shown to be sufficient for phase variation of *agn43 in vivo* (Haagmans and Van der Woude, 2000; Schrembri and Klemm, 2001). DN*ase*1 footprint analysis of OxyR binding to the *oxyRS* regulatory region revealed a total of five OxyR contacts on the DNA (Toledano *et al.*, 1994). These are labeled I-V in Figure 5.5. Reduced OxyR makes contact at positions I, II, IV, and V, resulting in transcription repression whereas oxidized OxyR makes contact at positions I, II, III, and IV, resulting in transcription activation (Toledano *et al.*, 1994). A comparison of the aligned *agn43* and *oxyRS* regulatory regions reveals that the sequence motifs corresponding to all five contact points are conserved in *agn43* (Figure 5.5). It is notable that contact point V overlaps with the Pribnow box of the *agn43* promoter. For this reason, reduced OxyR may obstruct RNA polymerase binding or impede its movement along DNA whereas oxidised OxyR

might be expected to vacate contact point V in *agn43* by analogy with its interactions at the *oxyRS* regulatory region.

Results of a recent continuation of this study have been consistent with this hypothesis. OxyR contact points III and IV were disrupted by introduction of a single base pair substitution (Kelly *et al.*, unpublished results). A construct containing a mutation in contact point III was able to undergo OxyR-dependent phase variation indicating that, as predicted, this region is not involved in binding of the reduced (repressive) form of the protein. However, a single base pair substitution in putative contact point IV resulted in a Locked ON phenotype reflecting the importance of this region in the binding of reduced OxyR. In addition an A residue situated between contact points II and III, which is highly conserved amongst OxyR binding sites (Zheng *et al.*, 2002), was mutated resulting in a Locked ON phenotype (Kelly *et al.*, unpublished results). It is likely that this mutation caused critical changes in local DNA architecture such that the surrounding points, with which OxyR makes direct contact, were wrongly positioned for successful binding by the repressor. The effects of all mutations that have been created in the OxyR binding Motif are summarised in Figure 5.10. Putative contact point V has never been disrupted since the effects of any such mutation on OxyR binding would probably be masked due to disruption of the *agn43* Pribnow box. The *agn43* OxyR DNase-I footprint has recently been published and is reported to extend from –15 to +43 supporting the idea that these regions are involved in OxyR binding (Wallecha *et al.*, 2002).

Figure 5.10 The location of mutated nucleotides thought to be involved in OxyR binding and their effect on *agn43* switch phenotype. Putative contact points for oxidised and reduced OxyR on agn43 DNA are indicated by Roman numerals I-IV. Nucleotides thought to make contact with OxyR are in Red. Bases that have been substituted in the *agn43* sequence are in blue and the effect of these individual base substitutions on switch phenotype is indicated.


Oxidised OxyR
Reduced OxyR
V
IV
IV
III
II
II
I
I
TAATAGAATAAAACGATCAATATCTATTTTATCGATCGTTTATATCGATC
-10
C
A
G
T
T
LOCKED ON
PHASE SWITCHING
LOCKED ON
LOCKED ON
LOCKED ON

In summary, there can be little doubt from the combined results presented here (in particular the Locked ON phenotypes displayed by pDEW11-16 in both wild type and *dam* backgrounds, the absence of OxyR binding to methylated/unmethylated DNA from these constructs, and the presence of a consensus OxyR binding site within the *agn43* regulatory region) that OxyR is capable of binding to the *agn43* promoter and that the adenine residues of 5′-GATC-3′ sites 2 and 3 are critically involved in OxyR recognition.

In addition, work presented in this chapter examined the role of individual 5′-GATC-3′ sites and different methylation patterns in *agn43* regulation, since even closely situated 5′-GATC-3′ sites can be differentially methylated and may influence each other. This mutational inactivation was performed in a manner unlikely to affect OxyR binding viz a viz by A-to-T, G-to-C and by C-to-G mutations in 5′-GATC-3′ sites 1-3, respectively. These mutations rendered the sites inactive for specific binding and methylation by Dam, due to disruption of the Dam target sequence.

The results of the phenotypic analysis (Table 5.7) and band shift assays performed on these mutated derivatives provide strong evidence that methylation of any two 5′-GATC-3′ sites is sufficient to block binding of the OxyR repressor. Recently Wallecha *et al.*, (2002) reported results of a similar study involving *agn43-lacZ* fusions that were inserted into the chromosome as a lysogen. They too mutated residues that were uninvolved in OxyR binding and found that a mutation in 5′-GATC-3′ site 1 had no effect on phase variation, in agreement with our results. However mutations in sites 2 or 3 resulted in a Locked OFF state unless Dam was

overexpressed, suggesting that, at wildtype levels, Dam was not methylating the remaining sites for some unknown reason. Transcriptional activation of these mutants in a Dam-overexpressing strain suggested that methylation of the two remaining sites was sufficient to block binding of OxyR and result in the Phase ON state. Consistent with this idea, and with our EMSA results, Wallecha *et al.* (2002) found that methylation of any two remaining sites does block OxyR binding *in vitro*. Our results indicate that this methylation, and hence transcriptional activation, also occurs *in vivo* in single mutants. In our study *agn43-lacZ* was expressed phase-variably by three single mutants pDEW10, pDEW20, and pDEW21, although the level of expression was slightly lower than wildtype when sites 2 or 3 were mutated. The reasons for the lack of methylation *in vivo* observed by Wallecha *et al.* may be due to differences in Dam expression levels between different strains. It is also possible their lysogen-borne transcriptional fusions were less available to cellular Dam than our plasmid-borne fusions. For example, architectural features of surrounding chromosomal DNA might limit Dam's accessibility to the lysogen-borne transcriptional fusion. Alternatively the phenotypes observed *in vivo* by Wallecha *et al.* (2002) may be somehow related to the fact that the mutations they used to disrupt the 2nd and 3rd 5'-GATC-3' sites were different form those used by us. Despite these phenotypic differences, the results of both studies show that methylation of any two 5'-GATC-3' sites in the *agn43* regulatory region is sufficient to block binding of, and repression by, OxyR.

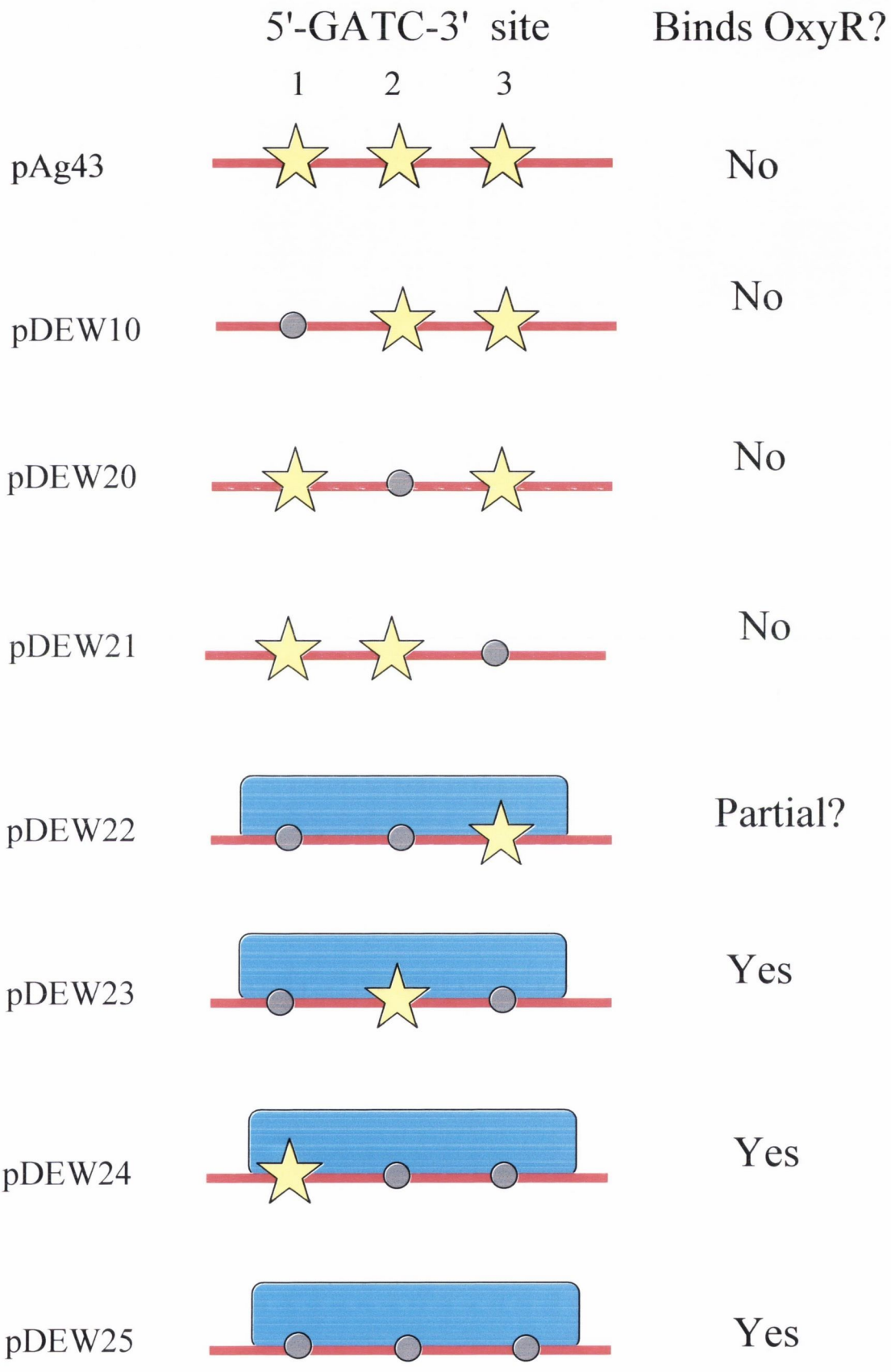
The Locked OFF phenotypes of the double mutants pDEW23 and pDEW24, and the constitutive low level transcription displayed by pDEW22, indicated that two 5′-GATC-3′ sites must remain available for phase variation to occur. These

phenotypes, along with bandshift analysis, indicate that OxyR can bind to singly methylated DNA blocking *agn43* transcription fully if sites 1 or 2 are methylated and partially if site 3 is methylated. Thus, it seems that the 5'-GATC-3' sites of *agn43* are redundant to some degree since no one site is absolutely required for phase variation to occur. However, methylation of any two sites is necessary and sufficient to block binding of OxyR *in vitro* indicating that this redundancy is only partial. The proposed effects of different methylation patterns on the binding of OxyR to the *agn43* regulatory region is summarised in Figure 5.11.

Evidence has recently been reported that Dam methylation affects *agn43* transcription by some additional mechanism since *in vitro agn43* transcription was further enhanced by Dam methylation in the absence of OxyR (Wallecha *et al.*, 2002). It has been suggested that these OxyR-independent effects are due to changes in DNA architecture brought about by Dam methylation (Wallecha *et al.*, 2002). It has also recently been reported that the *agn43* phase switch is biased OFF in strains containing a mutation in the *seqA* gene (Correnti *et al.*, 2002). Although it has been shown that the SeqA protein binds to hemimethylated *agn43* 5'-GATC-3' sites *in vitro* it is thought that the OFF bias in a *seqA* mutant is, at least in part, an indirect result of changes in the cell physiology due to an increased number of replication forks decreasing the availability of Dam (Correnti *et al.*, 2002).

As mentioned earlier, it seems likely that phase switching of Ag43 expression is not simply due to straight forward competition between Dam and OxyR for unmethylated DNA in the *agn43* regulatory region. There may be other hidden variables affecting the interactions of each of these proteins with the DNA. It has

Figure 5.11 The effect on OxyR binding of Dam methylation patterns corresponding to different constructs. Unmethylated 5′-GATC-3′ sites are represented by grey circles. Methylated 5′-GATC-3′ sites are represented by yellow stars. A blue oblong represents the OxyR tetramer.


5'-GATC-3' site
Binds OxyR?
1
2
3
pAg43
No
pDEW10
No
pDEW20
No
pDEW21
No
pDEW22
Partial?
pDEW23
Yes
pDEW24
Yes
pDEW25
Yes

been reported that 5′-GATC-3′ sites embedded in AT rich sequences are methylated less rapidly than are other sites (Bergerat *et al.*, 1989). The 5′-GATC-3′ sites involved in *agn43* regulation lie within a particularly AT rich region. This may be important for 'fine-tuning' the antagonistic interactions of Dam and OxyR with the DNA.

Chapter 6

General Discussion

The phase-variable nature of Ag43 expression was reported some time ago by Caffrey and Owen (1989) yet nothing was published in the primary literature regarding the likely mechanism of Ag43 phase variation until 1996. Then, Owen and colleagues reported their findings that Ag43 expression was Locked ON in both an *oxyR* mutant and in an *oxyR dam* double mutant and was Locked OFF in a *dam* mutant (Owen *et al.*, 1996; Henderson *et al.*, 1997a; Henderson and Owen, 1999). Significantly, within the regulatory region of *agn43* they noticed a 43-bp domain that had high homology with the OxyR-binding sites of other genes, and contained three 5'-GATC-3' sites, which are targets for Dam methylation. Taken together these observations were strongly indicative of a mechanism for phase variation in which Dam and OxyR regulate *agn43* at the level of transcription. A simple model for such a mechanism is one in which the OxyR protein binds to unmethylated 5'-GATC-3' sequences in the region of the *agn43* promoter, blocks methylation, and represses transcription by blocking the binding of RNA polymerase. Methylation of these sites would block binding of OxyR by steric hindrance, thereby allowing RNA polymerase to bind and transcribe the *agn43* gene.

The purpose of the current study was to provide insight into the regulatory mechanism controlling *agn43* transcription with respect, in particular, to the roles of OxyR and Dam in phase variation. To facilitate such an investigation it was necessary to construct a reporter system, which would allow the routine detection of transcription from the *agn43* promoter. Thus, Chapter 3 describes the construction of a plasmid-based *agn43-lacZ* transcriptional fusion, which proved to be a vital tool for transcriptional analyses conducted throughout this study. Phase-variable expression of *lacZ* expression under the control of the *agn43* promoter, confirmed

that Ag43 phase variation is regulated at the transcriptional level. Furthermore, the constitutive expression of this reporter fusion in an *oxyR* mutant confirmed that OxyR is involved in the repression of *agn43* transcription. DNA-binding studies described in Chapter 5 showed that OxyR is capable of binding the unmethylated *agn43* promoter *in vitro*. A lack of binding when the 5'-GATC-3' sites were methylated is supportive of our simple model in which Dam-methylation activates *agn43* transcription by blocking binding of OxyR.

In order to determine the mechanism by which OxyR represses *agn43* transcription it was necessary to locate the *agn43* promoter. Based on sequence analysis Henderson and Owen (1999) predicted that the *agn43* promoter was situated downstream of the three 5'-GATC-3' sites contained within the proposed OxyR-binding site. Primer extension studies and mutational analysis, documented in Chapter 3, revealed that this is not the case and that instead *agn43* transcription starts further upstream, at the G of the first 5'-GATC-3' site. Thus, the promoter lies upstream of the region to which OxyR binds as it (OxyR) represses transcription of *agn43*. It is therefore apparent that, despite sequence similarities, the regulatory mechanism of *agn43* differs fundamentally from that of the *mom* gene, whose promoter lies downstream of its three 5'-GATC-3' sites.

There are two possible mechanisms by which OxyR might repress transcription from an upstream promoter. OxyR might act as a transcriptional roadblock preventing the movement of bound RNA polymerase along the DNA helix. Alternatively, repression might result from occlusion of the promoter by OxyR such that RNA polymerase is unable to bind to its contact sites due to steric hindrance.

The alignment, made in Chapter 5, of the *agn43* regulatory region with that of *oxyR*S indicates that the reduced form of OxyR makes contact with sequences that overlap with the –10 box of the *agn43* promoter. The disruption of OxyR binding that resulted from mutation of residues whose selection was based on this alignment supports the idea that the contact sites for OxyR are as predicted. Furthermore, Correnti *et al.* (2002) have recently published a DNase I footprint of OxyR, which overlaps with the *agn43* Pribnow box identified in this study. Thus, it seems likely that reduced OxyR does occlude the *agn43* promoter and that this is the mechanism by which it exerts its negative regulatory effect. This hypothesis is consistent with the observation that the *agn43* promoter region is bound by both reduced and oxidized (which has a shorter footprint) forms of OxyR, but is repressed only by the reduced form.

Results in Chapter 3 demonstrated that the methylation state of the *agn43* regulatory region correlates with the state of *agn43-lacZ* transcription. Thus, DNA isolated from Phase ON cells was methylated or hemimethylated while that from Phase OFF cultures was fully unmethylated. The methylation of DNA from Phase ON cells is consistent with the activation of *agn43* transcription by Dam methylation, a central aspect of our model for phase variation. The fully unmethylated state of DNA from Phase OFF cells is consistent with the idea that bound OxyR protects 5'-GATC-3' sites from Dam methylation. Indeed it has recently been reported that OxyR binding can protect all three 5'-GATC-3' sites from methylation *in vitro*, and that hemimethylated *agn43* DNA is bound by OxyR (Correnti *et al.*, 2002). OxyR probably binds to newly synthesised DNA, maintaining its hemimethylated state, until passage of the replication fork results in the generation of fully unmethylated

DNA to which another copy of OxyR could bind. An extension of this study might involve examining the effect of hemimethylated DNA production on the kinetics of Ag43 phase variation. This could be achieved through use of a temperature sensitive *dnaC* mutation, which allows every cell in a culture to be blocked at the point of replication initiation. It is thus possible to synchronise the reinitiation of DNA synthesis in every member of the bacterial population. One could then observe the effect that reinitiation of DNA synthesis (and the concomitant production of hemimethylated DNA), may have on Ag43 phase variation.

Since Dam has a leading role in our model for the regulation of *agn43* phase variation, analysis of its 5'-GATC-3' target sites became a major part of this study. It has been suggested that all 5'-GATC-3' sites are not equally susceptible to methylation and there is evidence that sequences surrounding these targets are involved in Dam recognition. Furthermore, introduction of asymmetry around the symmetrical target sequence may give some orientation to the interaction of Dam with the DNA thereby influencing the order in which local 5'-GATC-3' sites are targeted (Bergerat *et al.*, 1989). These possibilities could have implications for the roles of the three 5'-GATC-3' in *agn43* regulation. If even closely situated 5'-GATC-3' sites can influence each other and are differentially methylated then it follows that the three 5'-GATC-3' sites within the *agn43* regulatory region may not be equally important in *agn43* regulation. Thus experimentation recorded in Chapter 5 examined the role of individual 5'-GATC-3' sites and different methylation patterns in regulation of *agn43* gene expression. This was performed by mutational inactivation of target sites in a manner unlikely to affect OxyR binding viz. by A-to-T, G-to-C and by C-to-G mutations in 5'-GATC-3' sites 1-3,

respectively. Thus, these mutations (a) rendered the sites inactive for specific binding and methylation by Dam, due to disruption of the Dam target sequence, yet (b) retained the potential of the promoter to bind the OxyR repressor. Accordingly, a series of reporter constructs were produced which contained different combinations of mutations corresponding to all possible methylation patterns, which might have an effect on *agn43* transcription.

In Chapter 5, the effect of these mutations on transcription from the *agn43* promoter was assessed in a number of strains that differed in their ability to express OxyR or Dam. The results of this study prompted analysis of the ability of OxyR to bind to methylated and unmethylated DNA from each mutant construct. Taken together, the results of phenotypic analysis, *agn43-lacZ* expression and *in vitro* OxyR-binding studies performed on these mutated derivatives provide strong evidence, (a) that any of the three 5'-GATC-3' site can be inactivated singly without abolishing phase variation; (b) that two sites must remain available for phase variation to occur; and (c) that OxyR can bind to singly methylated DNA blocking *agn43* transcription fully if sites 1 or 2 are methylated and partially if site 3 is methylated. The fact that no single site is absolutely required for phase variation to occur means that there is a certain level of redundancy associated with the 5'-GATC-3' sites of the *agn43* promoter region. However, this redundancy is only partial since methylation of any two sites is necessary and sufficient to block binding of OxyR *in vitro*. Redundancy in the *agn43* regulatory sequences may impart resilience to the system that allows subtle evolutionary changes to take place whilst maintaining phase variation.

Wallecha *et al.* (2002) have recently reported that Dam activates *agn43* by some additional mechanism since *in vitro* transcription of *agn43* was enhanced when DNA was Dam methylated. Furthermore, Haagmans and van der Woude (2002) previously observed that *agn43* transcription was 10-fold lower in a *dam oxyR* double mutant than in an *oxyR* mutant. It was not possible to measure transcription from our reporter plasmid, in an *oxyR dam* strain, as a Lac$^-$ double mutant was not available. Studies by Haagmans and van der Woude (2000) involved a transcriptional fusion of 680-bp of the *agn43* regulatory region to *lacZ* on a λ lysogen in the *E. coli* chromosome. β-galactosidase expression from this transcriptional fusion was ~ 10-fold greater than expression from our reporter plasmid pAg43. The *agn43* sequence in our transcriptional fusion starts just 50 bp upstream of the –35 RNA polymerase contact site. Thus, it is quite possible that additional regulators normally bind upstream of the *agn43* promoter to regions that lie beyond the 5' boundary of our *agn43-lacZ* fusion.

Mutational analysis of 5'-GATC-3' sites also provided valuable information regarding the specific nucleotides that are involved in the recognition by OxyR of the *agn43* regulatory region. Thus, the Locked ON phenotypes displayed by pDEW11-pDEW16 (which contained A-to-T mutations in 5'-GATC-3' sites 2 or 3) in both wildtype and *dam* backgrounds and the absence of OxyR binding to unmethylated/methylated DNA from these constructs indicates that the adenine residues of 5'-GATC-3' sites 2 and 3 are essential for OxyR recognition. These residues were contained within sequences that, based on our sequence alignment, were predicted to be involved OxyR recognition. Unfortunately, attempts to further examine the binding of OxyR by *in-vivo* studies involving DMS footprinting were

unsuccessful. However the *in-vitro* binding of OxyR to the *oxyRS* regulatory region has previously been characterized by means of hydroxyl radical footprinting (Toledano *et al.*, 1994). Similar experiments involving *agn43* DNA might help to determine whether the sequences with which OxyR makes contact are those predicted in this study. OxyR usually exists in the reduced form *in vivo* but is subject to oxidation *in vitro*. Despite this, it is possible to examine the binding of the reduced (repressing) form of the protein *in vitro*, by using a mutant strain expressing OxyR that is locked in the reduced conformation. Thus, one could examine the binding of both forms of the protein and determine whether their contact points within the *agn43* regulatory region are actually distinct. By confirming that only reduced-OxyR binds to regions overlapping the promoter, such experimentation would support our explanation for the inability of oxidized-OxyR to repress *agn43* (i.e. that it does not occlude the promoter).

Experiments presented in Chapter 3, involving analysis of transcription of *agn43-lacZ* in a Locked ON strain, revealed that the activity of the *agn43* promoter is affected by both the degree of oxygen availability and by growth temperature. Thus, *agn43* transcription is subject not only to the ON/OFF switch of phase variation but also to a type of 'volume control', whereby the actual level of transcription in Phase ON cells is environmentally regulated. It is likely that bacteria use these environmental signals of temperature and anaerobiosis to determine whether they are within the host. By regulating Ag43 expression accordingly they may manage their cellular resources more efficiently.

Promoter activity was elevated 2.5 fold when bacteria where grown under semi-anaerobic compared to aerobic conditions. In addition, *agn43* promoter activity was approximately 4-fold greater when cells were grown at 37°C, the internal host temperature, than 28°C. Evidence indicates that pathogenic *E. coli* strains exist within biofilms within the host intestine (Cassels and Wolf, 1995) and that pathogenic strains of *E. coli* can express Ag43 or Ag43 homologues (Owen *et al.*, 1996; Roche *et al.*, 2001; Torres et al., 2002). The *E. coli* bacterium may associate with these biofilm structures by increasing Ag43 expression in response to temperature signals and oxygen deprivation in the intestinal environment. Western blot analysis could confirm whether these environmentally-triggered transcriptional effects are reflected at the level of Ag43 protein expression. Additional flow-chamber-based studies might determine whether any such alterations in the level of Ag43 expression actually affect the ability of *E. coli* to engage in biofilm formation.

The stationary-phase induction of *agn43* promoter activity, described in Chapter 3, indicates that *agn43* is subject to some growth phase-dependent regulatory mechanism. The most influential regulator of stationary-phase induced genes is the alternative sigma factor, RpoS, which activates the expression of almost one hundred genes. Accordingly, the possibility of a role for RpoS in *agn43* transcription was investigated in Chapter 4. Contrary to expectations, experimentation revealed that Ag43 expression is up-regulated in an RpoS mutant and must therefore be subject to repression by RpoS in wildtype strains. It is most likely that RpoS regulates *agn43* indirectly by activating transcription of some unidentified *trans*-acting repressor. RpoS expression is elevated when bacteria are grown at temperatures below 30°C. An increase in RpoS-mediated repression may

therefore be responsible for the decrease in *agn43* promoter activity which was seen when cultures where grown at 28°C (Chapter 3).

Transposon mutagenesis was conducted in an attempt to identify this factor, and any others, which might be involved in the regulation of *agn43* transcription. This study did not result in the identification of any previously unknown regulators of *agn43* transcription. All mutants in which transcription was detectably elevated contained mutations in *rpoS* or *oxyR* with the exception of one that contained a mutation in the *lipB* gene, which encodes a regulator of Dam.

The results of this study do not discount the possibility that additional regulators are involved in regulation of *agn43* transcription. The identification of *agn43* regulators through mutation is limited in several respects. It is only possible to detect mutations in non-essential genes since mutation of essential genes is lethal. In addition, it was only possible to screen for mutants in which the level of transcription differed visibly and not for any in which phase variation rates may have been altered e.g. Lrp or SeqA mutants. Furthermore, screening for alterations in the level of transcription was complicated by the phase-variable nature of *agn43* transcription, and interesting mutations may have gone undetected.

A larger-scale mutagenesis programme may be more likely to result in the detection of additional mutants with altered *agn43* expression. Rather than screening many thousands of colonies, one could screen pools of mutagenized cells for altered expression of Ag43 using fluorescently-tagged antibodies. Also, transposon mutagenesis of a *dam oxyR* double mutant might lead to the identification of

additional factors that function independently of Dam and OxyR in the regulation of *agn43* expression.

If transcription of *agn43* is dependent on the outcome of competition between two regulatory proteins for the *agn43* promoter region, then Ag43 phase switching might be described as a semi-random event. Some degree of randomness within the *agn43* regulatory system should ensure that bacterial subpopulations with alternative phenotypes are generated, even when cells are subject to identical environmental conditions. Since Ag43-mediated autoaggregation is blocked by Type-1 fimbriae, Ag43 expression is unbeneficial in cells that are Phase ON for *fim* (Schembri and Klemm, 2001). By employing OxyR as a methylation-blocking factor *agn43* may exploit the net oxidation effect that results from the expression of Type-1 fimbriae so that switching becomes less random and is biased towards the OFF phase. Thus the *E. coli* bacterium appears to have evolved a cunning strategy for avoiding the wasteful expression of large quantities of Ag43 that are unable to function in aggregation.

The *E. coli* strain ML308-225 possesses duplicate copies of the *agn43* gene (*agn43A* and *agn43B*), which are 98% identical at the nucleotide and predicted protein levels (Roche *et al.*, 2001). Differences in the protein sequence of the surface expressed α^{43} subunits suggest that Ag43 may be subject to antigenic variation as well as phase variation. Analysis of EPEC clinical isolates indicated that they contain varying numbers of *agn43* alleles and express multiple Ag43 homologues (Roche *et al.*, 2001). Both *agn43* gene copies in ML308-225 are subject to phase variation, however they undergo phase switching at different rates.

It seems likely that sequence differences within the regulatory regions of the *agn43A* and *agn43B* gene copies are responsible for these differences in switching rate. At present only the regulatory region of the *agn43B* gene has been sequenced. Sequencing the promoter region from *agn43A* and comparing it with that of *agn43B* could provide valuable information regarding the sequences and likely mechanisms involved in regulating expression of the two Ag43 variant proteins.

The second *agn43* gene copy in ML308-225 must provide the bacteria with some type of fitness advantage as it has been stably maintained. Such an advantage might result from the ability to express functionally similar but antigenically different proteins, thereby evading the immune response without sacrificing valuable autoaggregative properties. Alternatively, having two *agn43* genes that are independently and differentially regulated may allow Ag43 to be expressed under a wider range of conditions. It is not known whether both *agn43A* and *agn43B* are expressed concurrently, however it appears that expression of the *agn43B* gene is favoured under normal laboratory growth conditions (Roche *et al.*, 2001). It is, however, possible that the *agn43A* gene is expressed primarily under a different set of environmental conditions.

This study has gone someway in furthering our understanding of the mechanisms that contribute to the regulation of *agn43* transcription. However, it seems quite likely that phase switching of Ag43 expression is not due only to straight forward competition between Dam and OxyR for unmethylated DNA in the *agn43* regulatory region. Additional regulatory factors could function, either independently or through interactions with OxyR and/or Dam, in *agn43* regulation.

The identification of any such factors is certainly vital if we are to fully understand the regulatory mechanisms controlling *agn43* transcription.

Bibliography

Alexeyev, M. F. and Shokolenko, I. N. (1995) Mini-Tn*10* transposon derivatives for insertion mutagenesis and gene delivery into the chromosome of Gram-negative bacteria. Gene. **160:** 59-62.

Åslund F, Zheng M, Beckwith J, and Storz, G. (1999) Regulation of the OxyR transcription factor by hydrogen peroxide and the cellular thiol-disulfide status. *Proc Natl Acad Sci U S A* **96:** 6161-6165.

Baga, M., Goransson, M., Normark, S., and Uhlin B. E. (1985) Transcriptional activation of a *pap* pilus virulence operon from uropathogenic *Escherichia coli.* *EMBO J* **4:** 3887-3893.

Bale, A., d'Alarcao, M., and Marinus, M. G., (1979) Characterization of DNA adenine methylation mutants of *Escherichia coli* K12. *Mutat Res* **59:** 157-165.

Banas, J.A., Ferretti, J.J., and Progulske-Fox, A. (1991) Identification and sequence analysis of a methylase gene in *Porphyromonas gingivalis. Nucleic Acids Res* **19:** 4189-4192.

Barbour, A.G., Burman, N., Carter, C.J., Kitten, T., Bergstrom, S. (1991) Variable antigen genes of the relapsing fever agent *Borrelia hermsii* are activated by promoter addition. *Mol Microbiol* **5:** 489-493.

Barrass, F., and Marinus, M.G. (1989) The Great GATC: DNA methylation in *E. coli. Trends Genet* **5:** 139-143.

Basaak, S., and Nagaraja, V. (1998) Transciptional Activator C Protein-mediated unwinding of DNA as a possible mechanism for *mom* gene activation. *J Mol Biol* **284:** 893-902.

Benz, A., and Schidt, M. A. (1992) Isolation and serologic characterization of AIDA-I, the adhesin mediating the diffuse adherence phenotype of the diarrhea-associated *Escherichia coli* strain 2787 (O126:H27). *Infect Immun* **60:** 13-18.

Bergerat, A., and Guschlbauer, W. (1990) The double role of methyl donor and allosteric effector of S-adenosyl-methionine for Dam methylase of *E. coli*. *Nucleic acids Res* **18:** 4369-4375.

Bergerat, A., Kriebardis, A., and Guschlbauer, W. (1989) Preferential site-specific hemimethylation of GATC sites in pBR322 DNA by Dam methyltransferase from *Escherichia coli*. *J Biol Chem* **264:** 4064-4070.

Bilge SS, Clausen CR, Lau W, Moseley SL. (1989) Molecular characterization of a fimbrial adhesin, F1845, mediating diffuse adherence of diarrhea-associated *Escherichia coli* to HEp-2 cells. *J Bacteriol* **171:** 4281-4289.

Bergerat, A., Guschlbauer, W., and Fazkerly, G. V. (1991) Allosteric and catalytic binding of S-adenosylmethionine to *Escherichia coli* DNA Adenine Methyltransferase monitored by ^{3}H NMR. *Proc. Natl Acad Sci USA* **88:** 6394-6397.

Bilge, S.S., Apostol, J.M., Fullner, K.J., and Moseley, S.L. (1993) Transcriptional organization of the F1845 fimbrial adhesin determinant of *Escherichia coli*. *Mol Microbiol* **7:** 993-1006.

Blattner, F.R., Plunkett, G., Bloch, C.A., Perna, N.T., Burland, V., Riley, M., Collado-Vides, J., Glasner, J.D., Rode, C.K., Mayhew, G.F., Gregor, J., Davis, N.W., Kirkpatrick, H.A., Goeden, M.A., Rose, D.J., Mau, B., and Shao, Y. (1997) The complete genome sequence of *Escherichia coli* K-12. Science **277:** 1453-74.

Blomfield, I.C. (2001) The regulation of *pap* and type 1 fimbriation in *Escherichia coli*. *Adv Microb Physio* **45:** 1-49.

Blomfield, I.C., Calie, P.J., Eberhardt, K.J., McClain, M.S., and Eisenstein, B.I. (1993) Lrp stimulates phase variation of type 1 fimbriation in *Escherichia coli* K-12. *J Bacteriol* **175:** 27-36.

Blomfield, I.C., Kulasekara, D.H., and Eisenstein, B.I. (1997) Integration host factor stimulates both FimB- and FimE-mediated site-specific DNA inversion that controls phase variation of type 1 fimbriae expression in *Escherichia coli*. *Mol Microbiol* **23:** 705-717.

Blomfield, I.C., McClain, M.S., and Eisenstein, B.I. (1991) Type 1 fimbriae mutants of *Escherichia coli* K12: characterization of recognized afimbriate strains and construction of new *fim* deletion mutants. *Mol Microbiol* **5:** 1439-1445.

Blyn, L. B., Braaten, B., and Low, D. (1994) Regulation of *pap* pilin phase variation by a mechanism involving differential Dam methylation states. *EMBO J* **9:** 4045-4054.

Bohannon, D.E., Connell, N., Keener, J., Tormo, A., Espinosa-Urgel, M., Zambrano, M.M., and Kolter, R. (1991) Stationary-phase-inducible "gearbox" promoters: differential effects of *katF* mutations and role of sigma 70. *J Bacteriol* **173:** 4482-4492.

Bolker, M., and Kahmann, R. (1989) The *Escherichia coli* regulatory protein OxyR discriminates between methylated and unmethylated states of the phage Mu *mom* promoter *EMBO J* **8**: 2403-2410.

Bolker, M., Wulczyn, F.G., and Kahmann, R. (1989) Role of bacteriophage MuC protein in activation of the *mom* gene promoter. *J Bacteriol* **171:** 2019-2027.

Boye, E., Marinus, M.G., and Lobner-Olesen, A. (1992) Quantitation of Dam methyltransferase in *Escherichia coli*. *J Bacteriol* **174:** 1682-1685.

Braaten, B.A., Blyn, L.B., Skinner, B.S., and Low, D.A. (1991) Evidence for a methylation-blocking factor (*mbf*) locus involved in *pap* pilus expression and phase variation in *Escherichia coli*. *J Bacteriol* **5:**1789-1800.

Braaten, B.A., Nou, X., Kaltenbach, L.S., and Low, D.A. (1994) Methylation patterns in *pap* regulatory DNA control pyelonephritis-associated pili phase variation in *E. coli*. *Cell* **76:** 577-588.

Braaten, B.A., Platko, J.V., van der Woude, M.W., Simons, B.H., de Graaf, F.K., Calvo, J.M., and Low, D.A. (1992) Leucine-responsive regulatory protein controls the expression of both the *pap* and *fan* pili operons in *Escherichia coli*. *Proc Natl Acad Sci U S A.* **89:** 4250-4254.

Braun, V., Hantkem K., and Koster, W. (1998) Bacterial iron transport: mechanisms, genetics, and regulation. *Met Ions Biol Syst* **35:** 67-145.

Braun, R.E., and Wright, A. (1986) DNA methylation differentially enhances the expression of one of the two *E. coli dnaA* promoters *in vivo* and *in vitro*. *Mol Gen Genet* **202:** 246-250.

Campbell, J.L., and Kleckner, N. (1990) *E. coli oriC* and the *dnaA* gene promoter are sequestered from dam methyltransferase following the passage of the chromosomal replication fork. *Cell* **62:** 967-979.

Cassels, F.J., and Wolf, M.K. (1995) Colonization factors of diarrheagenic *E. coli* and their intestinal receptors. *J Ind Microbiol* **15:** 214-226.

Cheng, S.C., Herman, and G., Modrich, P. (1985) Extent of equilibrium perturbation of the DNA helix upon enzymatic methylation of adenine residues. *J Biol Chem* **260:** 191-194.

Choudhury, D., Thompson, A., Stojanoff, V., Langermann, S., Pinkner, J., Hultgren, S.J., and Knight, S.D. (1999) X -ray structure of the FimC-FimH chaperone-adhesin complex from uropathogenic *Escherichia coli. Science* **285:** 1061-1066

Christman, M.F., Storz, G., and Ames, B.N. (1989) OxyR, a positive regulator of hydrogen peroxide-inducible genes in *Escherichia coli* and *Salmonella typhimurium*, is homologous to a family of bacterial regulatory proteins. *Proc Natl Acad Sci U S A* **86:** 3484-3488.

Coffey, E.M., and Eveland, W.C. (1967) Experimental relapsing fever initiated by *Borrelia hermsi* II. Sequential appearance of major serotypes in the rat. *J Infect Dis* **117:** 29-34.

Connell, I., Agace, W., Klemm, P., Schembri, M., Marild, S., and Svanborg, C. (1996) Type 1 fimbrial expression enhances *Escherichia coli* virulence for the urinary tract. *Proc Natl Acad Sci U S A* **93:** 9827-9832.

Connell N, Han Z, Moreno F, Kolter R. (1987) An *E. coli* promoter induced by the cessation of growth. Mol Microbiol **1:** 195-201.

Correnti, J., Munster, V., Chan, T., and van der Woude, M. (2002) Dam-dependent phase variation of Ag43 in *Escherichia coli* is altered in a *seqA* mutant. *Mol Microbiol* **44:** 521-532.

Corona-Izquierdo, F.P., and Membrillo-Hernandez, J. (2002) A mutation in *rpoS* enhances biofilm formation in Escherichia coli during exponential phase of growth. *FEMS Microbiol Lett* **211:** 105-110.

Costerton, J.W., Stewart, P.S., and Greenberg, E.P. (1999) Bacterial biofilms: a common cause of persistent infections. *Science* **284:** 1318-1322.

Danese, P.N., Pratt, L.A., Dove, S.L., and Kolter, R. (2000) The outer membrane protein, antigen 43, mediates cell-to-cell interactions within *Escherichia coli* biofilms. *Mol Microbiol* **37:** 424-432.

Delcour, A.H. (2002) Structure and function of pore-forming beta-barrels from bacteria. *J Mol Microbiol Biotechnol* **4:**1-10.

Deng, W.P., and Nickoloff, J.A. (1992) Site-directed mutagenesis of virtually any plasmid by eliminating a unique site. *Anal Biochem* **200:** 81-88.

Deitsch, K.W., Moxon, E.R., and Wellems, T.E. (1997) Shared themes of antigenic variation and virulence in bacterial, protozoal, and fungal infections. *Microbiol Mol Biol Rev* **61:** 281-93.

Diderischsen, B. (1980) *flu,* a metastable gene controlling surface properties of *Escherichia coli. J Bacteriol* **141:** 858-867.

DiRita, V.J., and Mekalanos, J.J. (1989) Genetic regulation of bacterial virulence. *Annu Rev Genet* **23:** 455-482.

Donlan, R. M. (2002) Biofilms: microbial life on surfaces. *Emerg Infect Dis* **8:** 881-890.

Donlan, R.M., and Costerton, J.W. (2002) Biofilms: survival mechanisms of clinically relevant microorganisms. *Clin Microbiol Rev* **15:** 167-193.

Dorman, C.J., and Higgins, C.F. (1987) Fimbrial phase variation in *Escherichia coli*: dependence on integration host factor and homologies with other site-specific recombinases. *J Bacteriol* **169:** 3840-3848.

Dove, S.L., Smith, S.G., and Dorman, C.J. (1997) Control of *Escherichia coli* type 1 fimbrial gene expression in stationary phase: a negative role for RpoS. *Mol Gen Genet* **254:** 13-20.

Eisenstein, B.I., Sweet, D.S., Vaughn, V., and Friedman, D.I. (1987) Integration host factor is required for the DNA inversion that controls phase variation in *Escherichia coli*. *Proc Natl Acad Sci U S A* **84:** 6506-6510.

Farewell, A., Kvint, K., and Nystrom, T. (1998) Negative regulation by RpoS: a case of sigma factor competition. **29:** 1039-1051.

Farinha, M.A., and Kropinski, A.M. (1990) Construction of broad-host-range plasmid vectors for easy visible selection and analysis of promoters. *J Bacteriol* **172:** 3496-3499.

Fernandez, L.A., and Berenguer, J. (2000) Secretion and assembly of regular surface structures in Gram-negative bacteria. *FEMS Microbiol Rev* **24:** 21-44.

Force, A., Lynch, M., Pickett, F.B., Amores, A., Yan, Y.L., and Postlethwait, J. (1999) .Preservation of duplicate genes by complementary, degenerative mutations. *Genetics* **151:** 1531-1545.

Forsman, K., Goransson, M., and Uhlin, B.E. (1989) Autoregulation and multiple DNA interactions by a transcriptional regulatory protein in *E. coli* pili biogenesis. *EMBO J* **8:** 1271-1278.

Forsman, K., Sonden, B., Goransson, M., and Uhlin, B.E. (1992) Antirepression function in *Escherichia coli* for the cAMP-cAMP receptor protein transcriptional activator. *Proc Natl Acad Sci U S A* **89:** 9880-9884.

Francis, K. P., Taylor, P. D. Inchley, C.H., and Gallagher, M.P. (1997) Identification of the *ahp* operon of *Salmonella typhimurium* as a macrophage-induced locus. *J Bacteriol* **179:** 4046-4048.

Gabig, M., Herman-Antosiewicz, A. Kwiatkowsa, M. Los, M., Thomas, M.S., and Węgrzyn, G. (2002) The cell surface protein Ag43 facilitates phage infection of *Escherichia coli* in the presence of bile salts and carbohydrates. *Microbiology* **148:** 1533-1542.

Gally, D.L., Bogan, J.A., Eisenstein, B.I., and Blomfield, I.C. (1993) Environmental regulation of the *fim* switch controlling type 1 fimbrial phase variation in *Escherichia coli* K-12: effects of temperature and media. *J Bacteriol* **175:** 6186-6193.

Gonzalez-Flecha, B. and Demple, B. (1997) Transcriptional Regulation of the *Escherichia coli oxyR* Gene as a Function of Cell Growth. *J Bacteriol* **179:** 6181-6186.

Haagmans, W., and van der Woude, M. (2000) Phase variation of Ag43 in *Escherichia coli:* Dam dependent methylation abrogates OxyR binding and OxyR-mediated repression of transcription. *Mol Microbiol* **35:** 877-887.

Haas, R., and Meyer, T.F. (1987) Molecular principles of antigenic variation in Neisseria gonorrhoeae. *Antonie Van Leeuwenhoek* **53:** 431-434.

Hallet, B., and Sherratt, D.J. (1997) Transposition and site-specific recombination: adapting DNA cut-and-paste mechanisms to a variety of genetic rearrangements. *FEMS Microbiol Rev* **21:** 157-178.

Hagblom, P., Segal, E., Billyard, E., and So, M. (1985) Intragenic recombination leads to pilus antigenic variation in *Neisseria gonorrhoeae*. *Nature* **31:** 5156-5158.

Hancock, R.E. (1997) The bacterial outer membrane as a drug barrier. *Trends Microbiol* **5:** 37-42.

Hasman, H., Chakraborty, T., and Klem, P. (1999) Antigen 43-mediated autoaggregation of *Escherichia coli* is blocked by Fimbriation. *J Bacteriol* **181:** 4834-4845.

Hasman, H., Schembri, M.A., and Klemm, P. (2000) Antigen 43 and type 1 fimbriae determine colony morphology of *Escherichia coli* K-12. *J Bacteriol* **182:** 1089-1095.

Hattman, S., Brooks, J.E., and Masweker, M. (1978) Sequence specificity of the P1 modification methylase controlled by the *Escherichia coli dam* gene. *J Mol Biol* **126:** 367-380.

Hattman, S. (1999) Unusual transcriptional and translational regulation of the bacteriophage Mu *mom* operon. *Pharmacol Ther* **84:** 367-388.

Hattman, S., and Sun, W. (1997) *Escherichia coli* OxyR modulation of bacteriophage Mu *mom* expression in Dam^+ cells can be attributed to its ability to bind hemimethylated *Pmom* promoter DNA. *Nucleic Acids Res* **25:** 4385-4388.

Harel, J., Daigle, F., Forget, C., Tessier, M.C., Crost, C., and Martin, C. (2000) Phase variation of F165(1) (Prs-like) fimbriae from *Escherichia coli* causing septicaemia in animals. *Can J Microbiol* **46:** 1101-1107.

Heichman, K.A., and Johnson, R.C. (1990) The Hin invertasome: protein-mediated joining of distant recombination sites at the enhancer. *Science* **249:** 511-517.

Heinrichs, D.E., Yethon , J.A., and Whitfield, C. (1998) Molecular basis for structural diversity in the core regions of the lipopolysaccharides of *Escherichia coli* and *Salmonella enterica*. *Mol Microbiol* 30: 221-232.

Henderson, I. R., and Owen, P. (1999) The major phase-variable protein of *Escherichia coli* structurally resembles the IgA-protease class of exported protein and is regulated by a novel regulatory mechanism involving Dam and OxyR. *Mol Microbiol* **181:** 2131-2141.

Henderson, I. R., Owen, P., and Nataro, J.P. (1999) Molecular switches – the ON and OFF of bacterial phase variation. *Mol Microbiol* **33:** 919-932.

Henderson, I. R., Meehan, M., and Owen, P. (1997a) A novel regulatory mechanism for a novel phase-variable outer membrane protein of *Escherichia coli*. *Adv Exptl Med Biol* **412**: 349-355.

Henderson, I.R., Meehan, M., and Owen, P. (1997b) Antigen 43, a phase-variable bipartite outer membrane protein, determines colony morphology and autoaggregation in *Escherichia coli* K-12. *FEMS Microbiol Lett* **149:**115-120.

Henderson, I. R., Navarro-Garcia, F., and Nataro, J.P, (1998) The great escape: structure and function of the autotransporter proteins. *Trends Microbiol* **6:** 370-378.

Hengge-Aronis, R. (1996) Back to log phase: sigma S as a global regulator in the osmotic control of gene expression in *Escherichia coli. Mol Microbiol* **21:** 887-893.

Herman, B., and Modrich, P. (1982) *Escherichia coli* Dam methylase: Physical and catalytic properties of the homogeneous enzyme. *J Biol Chem* **257:** 2605-2612.

Hill, S.A., Morrison, S.G., and Swanson, J. (1990) The role of direct oligonucleotide repeats in gonococcal pilin gene variation. *Mol Microbiol* **4:** 1341-1352.

Hobbs, M.M., Seiler, A., Achtman, M., and Cannon, J.G. (1994) Microevolution within a clonal population of pathogenic bacteria: recombination, gene duplication

and horizontal genetic exchange in the *opa* gene family of *Neisseria meningitidis*. *Mol Microbiol* **12:** 171-180.

Hultgren, S.J., Abraham, S., Caparon, M., Falk, P., St Geme, J.W., and Normark S. (1993) Pilus and nonpilus bacterial adhesins: assembly and function in cell recognition. *Cell* **73:** 887-901.

Izard, J. W., and Kendall, D.A. (1994) Signal peptides: exquisitely designed transport promoters. *Mol Microbiol* **13:** 765-773.

Jose, J., Jahnig, F., and Meyer, T.F. (1995) Common structural features of IgA1 protease-like outer membrane protein autotransporters. *Mol Microbiol* **18:** 378-380.

Julio, S.M., Heithoff, D.M., Provenzano, D., Klose, K.E., Sinsheimer, R.L., Low, D.A., and Mahan, M.J. (2001) DNA adenine methylase is essential for viability and plays a role in the pathogenesis of *Yersinia pseudotuberculosis* and *Vibrio cholerae*. *Infect Immun* **69:** 7610-7615.

Kahmann, R. (1983) Methylation regulates the expression of a DNA modification function encoded by phage Mu. *Cold spring harbor Symp Quant Biol* **47:** 639-646.

Kaltenbach, L.S., Braaten, B.A., and Low, D.A. (1995) Specific binding of PapI to Lrp-*pap* DNA complexes. *J Bacteriol* **177:** 6449-6455.

Karmali, M.A. (1989) Infection by verocytotoxin-producing *Escherichia coli*. *Clin Microbiol Rev* **2:** 15-38.

Kelly, D., Waldron, D.E., and Owen, P. (2002) Unpublished data.

Kim, S.O., Merchant, K., Nudelman, R., Beyer, W.F., Keng, T., DeAngelo, J., Hausladen, A., and Stamler, J.S. (2002) OxyR: a molecular code for redox-related signaling. *Cell* **109:** 383-396.

Kleckner, N., Bender, J., and Gottesman, S. (1991) Uses of transposons with emphasis on Tn*10*. *Methods Enzymol* **204:** 139-180.

Koch, C., Vandekerckhove, J., and Kahmann, R. (1988) *Escherichia coli* host factor for site-specific DNA inversion: cloning and characterization of the *fis* gene. *Proc Natl Acad Sci* U S A **85:** 4237-4241.

Kimura, T., Asai, T., Imai, M., and Takanami, M. (1989) Methylation strongly enhances DNA bending in the replication origin region of the *Escherichia coli* chromosome. *Mol Gen Genet* **219:** 69-74.

Kitten, T., and Barbour, A.G. (1990) Juxtaposition of expressed variable antigen genes with a conserved telomere in the bacterium *Borrelia hermsii*. *Proc Natl Acad Sci U S A.* **87:** 6077-6081.

Kjærgaard, K., Schembri, M.A., Ramos, C., Molin, S., and Klemm, P. (2000) Antigen 43 facilitates formation of multispecies biofilms. *Environ Microbiol* **2:** 695-702.

Klemm, P., Jorgensen, B.J., van Die, I., de Ree, H., and Bergmans, H. (1985) The *fim* genes responsible for synthesis of type 1 fimbriae in *Escherichia coli*, cloning and genetic organization. *Mol Gen Genet* **199:** 410-414.

Koch, C., Vandekerchove, J., and Kahmann, R. (1988) *Escherichia coli* host factor for site-specific DNA inversion: cloning and characterisation of the *fis* gene. *Proc. Natl. Acad. Sci. USA* **85:** 4237-4241.

Koebnik, R., Locher, K.P., and Van Gelder, P. (2000) Structure and function of bacterial outer membrane proteins: barrels in a nutshell. *Mol Microbiol* **37:** 239-253.

Koronakis, V., Sharff, A., Koronakis, E., Luisi, B., and Hughes, C. (2000) Crystal structure of the bacterial membrane protein TolC central to multidrug efflux and protein export. *Nature* **405:** 914-919.

Kullik, I., Stevens, J., Toledano, M.B., and Storz, G. (1995) Mutational analysis of the redox-sensitive transcriptional regulator OxyR: regions important for DNA binding and multimerization. *J Bacteriol* **177:** 1285-1289.

Kullik, I., Toledano, M.B., Tartaglia, L.A., and Storz, G. (1995) Mutational analysis of the redox-sensitive transcriptional regulator OxyR: regions important for oxidation and transcriptional activation. *J Bacteriol* **177:** 1275-1284.

Lauster, R. (1989) Close relationship between the HinfI and DpnA DNA-methyltransferase. *Nucleic Acids Res* **12:** 4402-4405.

Lin, J., Huang, S., and Zhang, Q. (2002) Outer membrane proteins: key players for bacterial adaptation in host niches. *Microbes Infect* **4:** 325-331.

Loveless, B. J., and Saier, M.H. (1997) A novel family of channel-forming, autotransporting, bacterial virulence factors. *Mol Membr Biol* **14:** 113-123.

Low, D., Robinson, E.N., McGee, Z.A., and Falkow, S. (1987) The frequency of expression of pyelonephritis-associated pili is under regulatory control. *Mol Microbiol* **1:** 335-346.

Malygin, E. G., Petrov, N.A., Gorgunov, Y.A, Kossykh, V.G., and Hattman, S. (1997) Interaction of the phage T4 Dam DNA-[N6-adenine] methyltransferase with oligonucleotides containing native or modified (defective) recognition sites. *Nucleic Acids Res* **25:** 4393-4399.

Malorny, B., Morelli, G., Kusecek, B., Kolberg, J., and Achtman, M. (1998) Sequence diversity, predicted two-dimensional protein structure, and epitope mapping of neisserial Opa proteins. *J Bacteriol* **180:** 1323-1330.

Marinus, M.G., Poteete, A., and Arraj, J.A. (1984) Correlation of DNA adenine methylase activity with spontaneous mutability in *Escherichia coli* K-12. *Gene* **28:** 123-125.

Marinus, M.G., and Morris, N.R. (1975) Pleiotropic effects of a DNA adenine methylation mutation (*dam-3*) in *Escherichia coli* K12. *Mutat Res* **28:** 15-26.

Martin, C. (1996) The *clp* (CS31A) operon is negatively controlled by Lrp, ClpB, and L-alanine at the transcriptional level. *Mol Microbiol* **21:** 281-292.

Maurer, J., Jose, J.J, and Meyer, T.F. (1999) Characterization of the essential transport function of the AIDA-I autotransporter and evidence supporting structural predictions. *J Bacteriol* **181:** 7014-7020.

McClain, M.S., Blomfield, I.C., Eberhardt, K.J., and Eisenstein, B.I. Inversion-independent phase variation of type 1 fimbriae in *Escherichia coli. J Bacteriol* **175:** 4335-4344.

Meehan, M. (1994) Studies of Antigen 43, a bipartite outer membrane protein of *Escherichia coli*. Ph.D. Thesis. Trinity College Dublin.

Mekalanos, J.J. (1992) Environmental signals controlling expression of virulence determinants in bacteria. *J Bacteriol* **174:** 1-7.

Messer, W., and Noyer-Weidner, M. (1988) Timing and targeting: the biological functions of Dam methylation in *E. coli*. *Cell* **54:** 735-737.

Meyer, T. F. (1990) Pathogenic *neisseriae*: a model of bacterial virulence and genetic flexibility. *Zentralbl Bakteriol* **274:** 135-154.

Miller, J. H. (1972) Experiments in Molecular Genetics, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Miller, J. H. (1992) A short course in bacterial genetics. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Miller V.L, and Stone, B.J. (1994) Molecular cloning of invasion genes from *Yersinia* and *Salmonella*. *Methods Enzymol* **236:** 546-51.

Modrich, P. (1987) DNA mismatch correction. *Annu Rev Biochem* **56:** 435-466.

Modrich, P. (1989) Methyl-directed DNA mismatch correction. *J Biol Chem* **264:** 6597-6600.

Morris, T.W., Reed, K.E., and Cronan, J.E. (1995) Lipoic acid metabolism in *Escherichia coli*: the *lplA* and *lipB* genes define redundant pathways for ligation of lipoyl groups to apoprotein. *J Bacteriol* **177:** 1-10.

Moxon, E.R., and Murphy, P.A. (1978) *Haemophilus influenzae* bacteremia and meningitis resulting from survival of a single organism. *Proc Natl Acad Sci U S A* **75:** 1534-1536.

Mukhopadhyay, S., and Schellhorn, H.E. (1997) Identification and characterization of hydrogen peroxide-sensitive mutants of *Escherichia coli*: genes that require OxyR for expression. *J Bacteriol* **179:** 330-338.

Mullis, K.B., and Faloona, F.A. (1987) Specific synthesis of DNA *in vitro* via a polymerase-catalyzed chain reaction. *Methods Enzymol* **155:** 335-350.

Nakamura, K., and Inouye, M. (1982) Construction of versatile expression cloning vehicles using the lipoprotein gene of *Escherichia coli*. *EMBO J* **1:** 771-775.

Nassif, X., Lowy, J., Stenberg, P., O'Gaora, P., Ganji, A, and So, M. (1993) Antigenic variation of pilin regulates adhesion of *Neisseria meningitidis* to humanepithelial cells. *Mol Microbiol* **8:** 719-725.

Nicholson, B., and Low, D. (2000) DNA methylation-dependent regulation of *pef* expression in *Salmonella typhimurium*. *Mol Microbiol* **35:** 728-742.

Nikaido, H. (1988) Structure and functions of the cell envelope of gram-negative bacteria. *Rev Infect Dis* **10:** 279-281.

Nikaido, H., and Saier M.H. (1992) Transport proteins in bacteria: common themes in their design. *Science* **258:** 936-942.

Nou, X., Skinner, B., Braaten, B., Blyn, L., Hirsch, D., and Low, D. (1993) Regulation of pyelonephritis-associated pili phase-variation in *Escherichia coli*: binding of the PapI and the Lrp regulatory proteins is controlled by DNA methylation. *Mol Microbiol* **7:** 545-553.

Olsen, P.B., and Klemm, P. (1994) Localization of promoters in the *fim* gene cluster and the effect of H-NS on the transcription of *fimB* and *fimE*. *FEMS Microbiol Lett* **116:** 95-100.

O'Toole, P.W., Austin, J.W., and Trust, T.J. (1994) Identification and molecular characterization of a major ring-forming surface protein from the gastric pathogen *Helicobacter mustelae*. *Mol Microbiol* **11:** 349-361.

Owen, P., and Kaback, H.R. (1978) Molecular structure of membrane vesicles from *Escherichia coli*. *Proc Natl Acad Sci U S A* **75:** 3148-3152.

Owen, P., Meehan, M., de Loughrey-Doherty, H., and Henderson, I.R. (1996) Phase-variable outer-membrane proteins in *Escherichia coli*. *FEMS Immunol. Med. Microbiol* **16:** 63-76.

Palmer, B.R., and Marinus, M.G. (1994) The *dam* and *dcm* strains of *Escherichia coli:* a review. *Gene* **143:** 1-12.

Peterson, K.R., Wertman, K.F., Mount, D.W., and Marinus, M.G. (1985) Viability of *Escherichia coli* K-12 DNA adenine methylase (*dam*) mutants requires increased expression of specific genes in the SOS regulon. *Mol Gen Genet* **201:** 14-19.

Plasterk, R., Vreiling, H.A., and Van de Putte, P. (1983) Transcription initiation of Mu *mom* depends on methylation of the promoter region and a phage encoded transactivator. *Nature* **301:** 344-347.

Pohlner, J., Halter, R., Beyreuther, K., and Meyer, T.F. (1987) Gene structure and extracellular secretion of *Neisseria gonorrhoeae* IgA protease. *Nature* **325:** 458-462.

Pratt, L.A., and Kolter, R. (1998) Genetic analysis of *Escherichia coli* biofilm formation: roles of flagella, motility, chemotaxis and type I pili. *Mol Microbiol* **30:** 285-293.

Prigent-Combaret, C., Prensier, G., Le Thi, T.T., Vidal, O., Lejeune, P., and Dorel, C. (2000) Developmental pathway for biofilm formation in curli-producing *Escherichia coli* strains: role of flagella, curli and colanic acid. *Environ Microbiol* **2:** 450-464.

Rasmussen, L.J., Marinus, M.G., and Lobner-Olesen, A. (1994) Novel growth rate control of *dam* gene expression in *Escherichia coli*. *Mol Microbiol* **12:** 631-638.

Restrepo, B.I., Kitten, T., Carter, C.J., Infante, D., and Barbour, A.G. (1992) Subtelomeric expression regions of *Borrelia hermsii* linear plasmids are highly polymorphic. *Mol Microbiol* **6:** 3299-3311.

Restrepo, B.I., and Barbour, A.G. (1994) Antigen diversity in the bacterium *Borrelia hermsii* through "somatic" mutations in rearranged *vmp* genes. *Cell* 78: 867-876.

Restrepo, B.I., Carter, C.J, and Barbour, A.G. (1994) Activation of a *vmp* pseudogene in *Borrelia hermsii*: an alternate mechanism of antigenic variation during relapsing fever. *Mol Microbiol* **13:** 287-299.

Rietschel, E.T., Schletter, J., Weidemann, B., El-Samalouti, V., Mattern, T., Zahringer, U., Seydel, U., Brade, H., Flad, H.D., Kusumoto, S., Gupta, D., Dziarski, R., and Ulmer, A.J. (1998) Lipopolysaccharide and peptidoglycan:

CD14-dependent bacterial inducers of inflammation. *Microb Drug Resist* **4:** 37-44.

Roberts, D., Hoopes, B.C., McClure, W.R., and Kleckner, N. (1985) IS*10* transposition is regulated by DNA adenine methylation. **43:** 117-130.

Roche A. J., McFadden, J. P., and Owen, P. (2001) Antigen 43, the major phase –variable protein of the *Escherichia coli* outer membrane, can exist as a family of proteins encoded by multiple alleles. *Microbiology* **147:** 161-169.

Roesch, P,L., and Blomfield, I.C. (1998) Leucine alters the interaction of the leucine-responsive regulatory protein (Lrp) with the *fim* switch to stimulate site-specific recombination in *Escherichia coli. Mol Microbiol* **27:** 751-761.

Russell, D.W., and Zinder, N.D. (1987) Hemimethylation prevents DNA replication in *E. coli. Cell* **50:** 1071-1079.

Sauer, F.G., Mulvey, M.A., Schilling, J.D., Martinez, J.J., and Hultgren, S.J. (2000) Bacterial pili: molecular mechanisms of pathogenesis. *Curr Opin Microbiol* **3:** 65-72.

Sambrook, J. E., Fritsch, F., and Maniatis, T. (1989) Molecular Cloning: A laboratory Manual, 2nd edition: Cold Spring Harbor Press, Cold Spring Harbor, New York.

Samuel, J.E., Perera, L.P., Ward, S., O'Brien, A.D., Ginsburg, V., and Krivan, H.C. (1990) Comparison of the glycolipid receptor specificities of Shiga-like toxin type II and Shiga-like toxin type II variants. *Infect Immun* **58:** 611-618.

Sarkari, J., Pandit, N., Moxon, E.R., and Achtman, M. (1994) Variable expression of the Opc outer membrane protein in *Neisseria meningitidis* is caused by size variation of a promoter containing poly-cytidine. *Mol Microbiol* **13:** 207-217.

Saunders, J.R. (1989) Antigenic variation. Modulating bacterial virulence. *Nature* **338:** 622-623.

Saunders, J.R. (1990) Genetic mechanisms for modulating virulence determinants on the bacterial surface. *Sci Prog* **74:** 279-290.

Schell, M.A. (1993) Molecular biology of the LysR family of transcriptional regulators. *Annu Rev Microbiol* **47:** 597-626.

Schellhorn, H.E., Audia, J.P., Wei, L.I., and Chang, L. (1998) Identification of conserved, RpoS-dependent stationary-phase genes of *Escherichia coli. J Bacteriol* **180:** 6283-6291.

Schembri, M. A., Hasman, H., and Klemm, P. (2000) Antigen 43 from *Escherichia coli* Induces Inter- and Intraspecies Cell Aggregation and Changes in Colony Morphology of *Pseudomonas fluorescens*. *J Bacteriol* **182:** 4789-4796.

Schembri, M.A., and Klemm, P. (2001) Coordinate gene regulation by fimbriae-induced signal transduction. *EMBO J* **20:** 3074-3081.

Schirmer, T., Keller, T.A., Wang, Y.F., and Rosenbusch, J.P. (1995) Structural basis for sugar translocation through maltoporin channels at 3.1 Å resolution. *Science* **267:** 512-514.

Schmoll, T., Ott, M., Oudega, B., and Hacker, J. (1990) Use of a wild-type gene fusion to determine the influence of environmental conditions on expression of the S fimbrial adhesin in an *Escherichia coli* pathogen. *J Bacteriol* **172:** 5103-5111.

Segal, E., Billyard, E., So, M., Storzbach, S., and Meyer, T.F. (1985) Role of chromosomal rearrangement in *N. gonorrhoeae* pilus phase variation. *Cell* **40:** 293-300.

Seifert, H. S. (1996) Questions about gonococcal pilus phase- and antigenic variation. *Mol Microbiol* **21:** 433-440.

Seiler, A., Blocker, H.. Frank, R., and Kahmann, R. (1986) The *mom* gene of bacteriophage Mu: the mechanism of methylation dependent expression. *EMBO J* **5:** 2719-2728.

Sharpe, M.E., Chatwin, H.M., Macpherson, C., Withers, H.L., and Summers, D.K. (1999) Analysis of the ColE1 stability determinant Rcd. *Microbiology* **145:** 2135-2144.

Southern, E.M. (1975) Detection of specific sequences among DNA fragments separated by gel electrophoresis. *Biotechnology* **24:** 122-139.

Stern, A,, Brown, M., Nickel, P., and Meyer, T.F. (1986) Opacity genes in *Neisseria gonorrhoeae*: control of phase and antigenic variation. *Cell* **47:** 61-71.

Storz, G., Tartaglia, L.A., and Ames, B.N. (1990) Transcriptional regulator of oxidative stress-inducible genes: direct activation by oxidation. *Science* **248:** 189-194.

Struve, C., and Krogfelt, K.A. (1999) *In vivo* detection of *Escherichia coli* type 1 fimbrial expression and phase variation during experimental urinary tract infection. *Microbiology* **145:** 2683-2690.

Stumpe, S., Schmid, R., Stephens, D.L., Georgiou, G., and Bakker, E.P. (1998) Identification of OmpT as the protease that hydrolyzes the antimicrobial peptide protamine before it enters growing cells of *Escherichia coli*. *J Bacteriol* **180:** 4002-4006.

Sun W., and Hattman, S. (1996) *Escherichia coli* OxyR protein represses the unmethylated bacteriophage Mu *mom* operon without blocking binding of the transcriptional activator C. *Nucleic Acids Res* **20:** 4042-4049.

Swanson, J., Barrera, O., Sola, J., and Boslego, J. (1988) Expression of outer membrane protein II by gonococci in experimental gonorrhea. *J Exp Med* **168:** 2121-2129.

Swanson, J., Robbins, K., Barrera, O., and Koomey, J.M. (1987) Gene conversion variations generate structurally distinct pilin polypeptides in *Neisseria gonorrhoeae*. *J Exp Med* **165:** 1016-1025.

Taghbalout, A., Landoulsi, A., Kern, R., Yamazoe, M., Hiraga, S., Holland, B., Kohiyama, M., and Malki, A. (2000) Competition between the replication initiator DnaA and the sequestration factor SeqA for binding to the hemimethylated chromosomal origin of *E. coli in vitro*. *Genes Cells* **5:** 873-884.

Tao, K., Zou, C., Fujita, N., and Ishihama, A. (1995) Mapping of the OxyR protein contact site in the C-terminal region of RNA polymerase alpha subunit. *J Bacteriol* **177:** 6740-6744.

Tartagalia, L., Gimeno, C.J. Storz G. and Ames, B. (1992) Multidegenerate DNA recognition by the OxyR transcriptional regulator. *J Biol Chem* **267:** 2038-2045.

Thanassi, D.G., Cheng, L.W., and Nikaido, H. (1997) Active efflux of bile salts by *Escherichia coli*. *J Bacteriol* **179:** 2512-2518.

Theiking, V., Dubois, S., Eritja, R., and Guschlbauer, W. (1997) Dam methylatransferase from *Escherichia coli*: kinetic studies using modified DNA oligomers: nonmethylated substrates. *J Biol Chem* **378:** 407-415.

Toledano, M., Kullick, I., Trinh, F., Baird, P., Schneider, T., and Storz, G. (1994) Redox-dependent shift of OxyR DNA Contacts along an extended DNA-Binding Site: A mechanism for differential promoter selection. Cell **78:** 897-909.

Torres, A. G., Perna, N.T., Burland, V., Ruknudin, A., Blattner, F.R., and Kaper, J.B. (2002) Characterization of Cah, a calcium-binding and heat-extractable autotransporter protein of enterohaemorrhagic *Escherichia coli*. *Mol Microbiol* **45:** 951-966.

Toussaint, A. (1977) DNA modification of bacteriophage Mu-1 requires both host and bacteriophage functions. *J Virol* **23:** 825-826.

Urig, S., Gowher, H., Hermann, A., Beck, C., Fatemi, M., Humeny, A., and Jeltsch, A. (2002) The *Escherichia coli* Dam DNA methyltransferase modifies DNA in a highly processive reaction. *J Mol Biol* **319:** 1085-1096.

Vaisvila, R., L. J. Rasmussen, A. Løbner-Oleson, von Freiesleben, U., and

Marinus, M.G. (2000) The LipB protein is a negative regulatory of *dam* gene expression in *Escherichia coli*. *Biochimica et Biophysica Acta* **1494:** 43-53.

van der Woude, M., Braaten, B., and Low, D. (1992) Evidence for global regulatory control of pilus expression in *Escherichia coli* by Lrp and DNA methylation: model building based on analysis of *pap*. *Mol Microbiol* 1992 **6:** 2429-2435.

van der Woude, M., Braaten, B., and Low, D. (1996) Epigenetic phase variation of the *pap* operon in *Escherichia coli*. *Trends Microbiol* **4:** 5-9.

van der Woude, M.W., Kaltenbach, L.S., and Low, D.A. (1995) Leucine-responsive regulatory protein plays dual roles as both an activator and a repressor of the *Escherichia coli pap* fimbrial operon. *Mol Microbiol* **17:** 303-312.

van der Woude, M., and Low, D. A. (1994) Leucine-responsive regulatory protein and deoxyadenosine methylase control the phase variation and expression of the *sfa* and *daa* pili operons in *Escherichia coli*. *Mol Microbiol* **11:** 605-618.

Veiga, E., Sugawar, E., Nikairo, H., de Lorenzo, V., and Fernandez, L.A. (2000). Export of autotransported proteins proceeds through an oligomeric ring shaped by C-terminal domains. *EMBO J* **21:** 2122-2131.

Virji, M,, Kayhty, H., Ferguson, D.J., Alexandrescu, C., Heckels, J.E., and Moxon, E.R. (1991) The role of pili in the interactions of pathogenic *Neisseria* with cultured human endothelial cells. *Mol Microbiol* **5:** 1831-1841.

Virji,M., and Heckels, J.E. (1986) The effect of protein II and pili on the interaction of Neisseria gonorrhoeae with human polymorphonuclear leucocytes. *J Gen Microbiol* **132:** 503-512.

Waldron, D.E., Owen, P., and Dorman, C.J. (2002) Competitive interaction of the OxyR DNA-binding protein and the Dam methylase at the antigen 43 gene regulatory region in *Escherichia coli. Mol Microbiol* **44:** 509-520.

Wallecha, A., Munster, V., Correnti, J., Chan, T., and van der Woude, M. (2002) Dam- and OxyR-dependent Phase Variation of *agn43*: Essential Elements and Evidence for a new Role of DNA methylation. *J Bacteriol* **184:** 3338-3347.

Warne, S.R., Varley, J.M., Boulnois, G.J., and Norton, M.G. (1990) Identification and characterization of a gene that controls colony morphology and auto-aggregation in *Escherichia coli* K12. *J Gen Microbiol* **136:** 455-462.

Weyand, N.J., Braaten, B.A., van der Woude, M., Tucker, J., and Low, DA. (2001) The essential role of the promoter-proximal subunit of CAP in *pap* phase variation: Lrp- and helical phase-dependent activation of *papBA* transcription by CAP from -215. *Mol Microbiol* **39:** 1504-1522.

Weyand, N. J., and Low, D. (2000) Lrp is sufficient for the establishment of the phase off *pap* DNA methylation pattern and repression of *pap* transcription *in vitro*. *J Biol Chem* **275:** 3192-3200.

White-Ziegler, C., Villapakkam, A., Kronaseki A., and Young, S. (2000) H-NS Controls *pap* and *daa* Fimbrial Transcription in *Escherichia coli* in Response to multiple environmental cues. *J Bacteriol* ***182:*** 6391-6400.

Zheng, F., Åslund, D., and Storz, G. (1998) Activation of the OxyR transcription factor by reversible disulphide bond formation. *Science* **279:** 1718-1721.