FLAVONOID DIRECTED REGULATION IN LACTOBACILLI

By

SANTOSH GURUNATHRAO PANDE

A DISSERTATION PRESENTED TO THE GRADUATE SCHOOL
OF THE UNIVERSITY OF FLORIDA IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

UNIVERSITY OF FLORIDA

2011

To my wife, parents and brothers for believing in me and providing unconditional love and support

ACKNOWLEDGMENTS

I would like to thank Dr. Graciela L. Lorca, chair of my PhD supervisory committee, for her support, encouragement and advice throughout my research program. I will always remain indebted to her for providing me an opportunity to work with her and for guiding me in writing the article and this dissertation.

I would also like to thank Dr. Claudio Gonzalez for his constructive criticism and encouragement to develop strong critical thinking skills. I would also like to acknowledge Dr. K. T. Shanmugam, Dr. Tony Romeo, and Dr Linda Bloom for serving on my PhD supervisory committee and for their valuable feedback on my research. I am thankful for both Fernando Pagliai and Chris L. Gardner for providing help whenever I needed their assistance. All the members of Drs. Lorca's and Gonzalez's laboratories were also instrumental in this work, especially Anastasia Potts, Clara Vu, and Kin Lai for their invaluable assistance during my research. I also would like to acknowledge Mariam Besharat, Nayantara Orekondy and Beverly Driver for their technical assistance. My special thanks to Dr. Boris Belitsky and Dr. Abraham Sonenshein for making the pLG103 plasmid available. I also thank University of Florida Alumni Foundation and the National Institutes of Health (NIH) for providing the financial support for my dissertation project. Finally, the words are not enough to thank my wife, parents, and brothers. Without their assistance, none of this would have been possible.

TABLE OF CONTENTS

page

ACKNOWLEDGMENTS 4

TABLE OF CONTENTS 5

LIST OF TABLES 8

LIST OF FIGURES 9

LIST OF ABBREVIATIONS 11

ABSTRACT 15

CHAPTER

1 INTRODUCTION 16

Flavonoids 16
Dietary Flavonoids and Health 16
Flavonoids Classification 17
Flavones 17
Flavonols 18
Flavanone 18
Flavan-3-ols 18
Anthocyanins 18
Isoflavones 18
Microbial Modification of Flavonoids 19
Modification of Flavonoids by Lactic Acid Bacteria (LAB) 20
Modulation of Gene Expression 21
Classification of Transcription Factors 25
Transcription Factors Responding to Flavonoids 27
Members of TetR Family Responsive to Flavonoids 27
Members of MarR Family Responsive to Flavonoids 28
Members of LuxR Family Responsive to Flavonoids 28
Members of LysR Family Responsive to Flavonoids 29
The LysR Family of Transcription Regulators (LTTRs) 30
Introduction 30
Domain Arrangement of LTTRs 30
LTTRs DNA Binding Sites 31
Mode of Regulation 32
Mode of Gene Regulation by LTTRs 32
Negative Gene Regulation by LTTRs 34
LrhA regulation of the *flhDC* operon 34
HexA regulation of virulence genes in *Erwinia carotovora* 34

Positive Gene Regulation by LTTRs 35
AtzR regulation of *atzDEF* 35
SpvR regulation of *spvABCD* 36
LTTR Autoregulation 36
Negative Autoregulation 37
AtzR 37
CysB 37
Positive Autoregulation 38
SpvR 39
LrhA 39
Project Design and Rationale 40

2 MATERIALS AND METHODS 53

Materials 53
Bacterial Strains and Plasmids 54
Competent Cells Preparation 54
DNA Manipulations 55
DNA Amplification 55
Cloning Techniques 56
Transformation 56
DNA Electrophoresis 57
Site Directed Mutagenesis 57
Construction of *lacZ* Fusions 58
RNA Extraction 59
cDNA Synthesis 59
Protein Techniques 60
Protein Overexpression in *E. coli* 60
Protein Purification 60
Protein Quantification 61
Protein Separation by SDS-PAGE 61
Size-Exclusion Chromatography 62
Small Molecule Screening by Differential Scanning Fluorometry 62
Electrophoretic Mobility Shift Assays (EMSAs) 63
DNase I Footprinting 64
β-galactosidase Assays 65
The Tools Used for Bioinformatics Studies 65

3 LVIS1989 (KAER), A LYSR TYPE TRANSCRIPTIONAL ACTIVATOR, UP REGULATES *LVIS1986*, *LVIS1987*, *LVIS1988* and *LVIS1989* IN RESPONSE TO KAEMPFEROL 74

Introduction 74
Results 76
The LysR Member, KaeR (LVIS1989) Interacts With Flavonoids *In Vitro* 76
Flavonoids are Involved in the Upregulation of *LVIS1986*, *LVIS1987*, *LVIS1988*, and *LVIS1989* mRNAs 77

Identification of KaeR Binding Region 79
KaeR Interacts With Two Binding Sites Separated by 280 bp 79
The Two Protected Regions are Required for the Formation of Higher Order Complexes *In Vitro* 80
KaeR Activity is Modulated by Kaempferol *In Vitro* 81
In Vivo Assessment of KaeR's Binding Sites on *LVIS1988* Expression 82
KaeR is Positively Autoregulated 84
Discussion and Conclusion 84

4 IDENTIFICATION OF KAER AMINO ACIDS INVOLVED IN LIGAND INTERACTION 99

Introduction 99
Results 101
KaeR Structure Modeling- 101
Identification of Residues that Modulate KaeR Activity 103
Effects of Mutations on the DNA Binding Properties of KaeR- 104
Effect of Mutations on the Oligomeric State of KaeR 104
Conclusions 105

5 SUMMARY AND CONCLUSIONS 123

Summary of Findings 123
Future Direction 125

APPENDIX

OPTIMIZATION OF KAER BINDING CONDITIONS 126

LIST OF REFERENCES 127

BIOGRAPHICAL SKETCH 140

LIST OF TABLES

Table		page
1-1	Flavonoids classification	42
1-2	Major transcription factors families in prokaryotes.	44
2-1	Strains used in this study.	66
2-2	Plasmids used in this study.	67
2-3	Primers used for electrophoretic mobility shift assays (EMSAs) and DNase I footprint assays.	68
2-4	Primers used to clone the genes for protein expression.	69
2-5	Primers used to clone the genes for *lacZ* reporter assays and sequencing reactions.	70
2-6	Primers used for quantitative real time PCR (qRT-PCR).	71
2-7	Primers used for site directed mutagenesis.	72
2-8	List of bioinformatics tools used.	73
3-1	Stabilization effect of ligand binding.	89
4-1	Summary of the KaeR homology modeling results.	107
4-2	Summary of KaeR mutation studies	108

LIST OF FIGURES

Figure		page
1-1	One component and two component response regulators.	46
1-2	Mechanism of transcriptional repression.	47
1-3	Mechanism of transcriptional activation.	48
1-4	Dual role of transcription factors.	49
1-5	LTTR domains arrangement	51
1-6	Classical mode of gene regulation by LTTRs.	52
3-1	Fluorescence based ligand screening assay.	90
3-2	Genomic environment of the *L. brevis* genes encoding LysR type transcriptional regulators.	91
3-3	Identification of the genes regulated by LVIS1989 (KaeR).	92
3-4	Identification of the KaeR binding region	93
3-5	Identification of KaeR binding sites.	94
3-6	Effect of kaempferol on KaeR binding to the different binding sites.	96
3-7	Gel filtration analysis of KaeR oligomerization in presence of kaempferol.	97
3-8	*In vivo* assessment of KaeR's binding sites.	98
4-1	CbnR and ArgP dimers.	109
4-2	CbnR tetramer (PDB # 1IZ1).	110
4-3	ArgP tetramer (PDB # 3ISP)	111
4-4	KaeR domains arrangement.	112
4-5	The effector binding domain of CynR (PDB # 3HFU) alignment to the structural model of KaeR.	113
4-6	Analysis of the azide ligand binding pocket in CynR.	114
4-7	Residues involved in ethanediol binding in CynR.	115
4-8	Summary of the putative KaeR residues involved in the ligand binding..	116

4-9 Electrophoretic mobility shift assays (EMSAs) to assess kaempferol effect on KaeR mutants. 117

4-10 Size exclusion chromotography to assess kaempferol effect on KaeR mutants. 120

4-11 CynR dimerization domain within the CynR tetramer. 122

A-1 Optimization of the KaeR binding conditions. 126

LIST OF ABBREVIATIONS

ABS	Activator binding site
Amp	Ampere
Ap^r	Ampicillin resistance
BGSC	Bacillus genetic stock center
B. subtilis	*Bacillus subtilis*
°C	Degree celcius
C4-HSL	C4-homoserine lactone
CAP	Catabolite activator protein
CBS	CysB binding site
cDNA	Complementary DNA
Chrom.	*Chromobacterium*
Cm^r	Chloramphenicol resistance
Comb.	*Combretum*
CPRG	Chlorophenol Red-β-D-galactopyranoside
CRP/Crp	Cyclic AMP receptor protein
CTD	Carboxy terminal domain
C-terminal	Carboxy terminal
DBD	DNA binding domain
DMSO	Dimethyl sulphoxide
dNTP	Deoxyribonucleotide triphosphate
DNA	Deoxyribonucleic acid
DTT	Dithiothreitol
EBD	Effector binding domain
E. coli	*Escherichia coli*

EDTA	Ethylenediamine tetra acetic acid
EGTA	Ethylene glycol tetra acetic acid
EMSA	Electrophoretic mobility shift assay
Erw.	*Erwinia*
Eub	*Eubacterium*
FPLC	Fast protein liquid chromatography
GIT	Gastro intestinal tract
GRAS	Generally regared as safe
HTH	Helix turn helix
IPTG	Isopropyl β-D-1-thiogalactopyranoside
ITC	Isothermal titration calorimetry
L.	*Lactobacillus*
L	Liter
LAB	Lactic acid bacteria
LB	Luria Bertani Medium
L. brevis	*Lactobacillus brevis*
LTTR	LysR type transcriptional regulator
M	Molar
µM	Micromolar
min	Minutes
ml	Milliliters
mM	Milimolar
MW	Molecular weight
N-terminal	Amino terminal
NA	Not Applicable

ND	Not determined
ng	Nanogram
nm	Nanometer
nM	nanoMolar
NO	Nitric oxide
OD600	Optical density at 600 nm
OFS	Organically farmed soybeans
ORF	Open reading frame
P. aeruginosa	*Pseudomonas aeruginosa*
PAGE	Polyacrylamide Gel Electrophoresis
PCR	Polymerase chain reaction
PDB	Protein data bank
Pg	Picogram
Poly (dI-dC)	Poly(deoxyinosinic-deoxycytidylic)
P. putida	*Pseudomonas putida*
qRT-PCR	Quantitaive real time polymerase chain reaction
R.	Rhizobium
RBS	Repressor binding site
RCSB	Resource for studying biological macromolecules
RD	Regulatory domain
RNA	Ribonucleic acid
RNAP	RNA polymerase
RPM	Revolutions per minute
SD	Shine-Dalgarno sequence
SDS	Sodium (Dodecyl) Sulfate

sec	Second
TCEP	Tris (2-carboxyethyl) phosphine
TF	Transcription factor
U	Units
UV	Ultra Violet
V	Volt
wHTH	Winged helix turn helix

Abstract of Dissertation Presented to the Graduate School
of the University of Florida in Partial Fulfillment of the
Requirements for the Degree of Doctor of Philosophy

FLAVONOID DIRECTED REGULATION IN LACTOBACILLI

By

Santosh Gurunathrao Pande

August 2011

Chair: Graciela L. Lorca
Major: Microbiology and Cell Science

The ability of transcription factors to respond to flavonoids as signal molecules was investigated in *Lactobacillus brevis*. Through *in vitro* screening of a small library of flavonoids, LVIS1989 (KaeR), a LysR type transcriptional regulator (LTTR), was identified as responsive to kaempferol. The modulation of KaeR activity by flavonoids was characterized *in vivo* and *in vitro*. DNase I footprint analysis identified the binding of KaeR at two distinctive sites, one in the intergenic region between *LVIS1988* and *LVIS1989* (-39 to +2) and another within *LVIS1988* (-314 to -353, from *kaeR* translation start point). Electrophoretic mobility shift assays (EMSAs) revealed that both binding sites are required for KaeR binding *in vitro*. Furthermore, KaeR–DNA interactions were stabilized by the addition of kaempferol (20 µM). *In vivo* qRT-PCR experiments performed in *L. brevis* confirmed that the divergently transcribed genes *LVIS1986, LVIS1987,* and *LVIS1988 and kaeR* are up-regulated in the presence of kaempferol, indicating the role of KaeR as a transcriptional activator. Transcriptional *lacZ* fusions using *Bacillus subtilis* as a surrogate host showed that expression of *kaeR* was induced by the presence of the flavonoid. These results indicate that KaeR belongs to a small and poorly understood family of LTTRs that are positively autoregulated by a ligand.

CHAPTER 1
INTRODUCTION

Flavonoids

Flavonoids are plant derived polyphenolic compounds which act as a first line of defence against environmental stress conditions. In plants, flavonoids are responsible for fruit and flower coloration. Flavonoids role in plant protection against environmental stress conditions has been extensively studied. In general, plant flavonoids are induced under excess light, cold, and drought conditions (Tatinni *et. al*., 2004; Li *et al.*,1993). Flavonoids quench the reactive oxygen species generated as a result of these stress conditions. Flavonoids also offer a competitive advantage to the plants. The allelopathic role of flavonoids has been shown in *Centaurea maculosa*. Bais *et al*. (2003) reported that *Centaurea maculosa* weed roots secrete (-)-catechin which inhibits the growth of the native plants. Flavonoids also possess antimicrobial properties against pathogenic fungi and bacteria which affect the plant health through foliage loss and reduced growth rate (Close and McArthur, 2002). The flavonoids are also responsible for the signaling of symbiotic relationship between leguminous plants and Rhizobia. Leguminous plants exude flavonoids which are sensed by Rhizobia activating the transcription of the genes involved in nodulation (Brencic and Winans, 2005).

Dietary Flavonoids and Health

Flavonoids possess antioxidant and radical scavenging properties (Rice-Evans *et al.*, 1997). Flavonoids are suggested to mediate anti-depressant or anti-parkinson's activity and possess anti-inflammatory properties (Jäger and Saaby, 2011).

The role of the green tea flavonoids (−)-epicatechin and its derivaties in cancer prevention has been suggested using mouse model studies (Mantena *et al.*, 2005;

Wang *et al.*, 1992; Baliga and Katiyar, 2006). In studies by Fraga *et.al*. (2011), it was suggested that flavonols lower blood pressure by maintaing optimal NO levels and decreasing the concentration of superoxide anions in blood vessels.

A 24 year study carried out in Finland using 9,959 human subjects ranging in age from 15 to 99, suggested that flavonoids could have the potential to reduce lung cancer incidence (Knekt *et al.*, 1997). Another study carried out over five years with 805 human subjects aged 64 to 84 showed that the consumption of flavonoids, in particular quercetin, kaempferol, myricetin, apigenin,and luteolin, was inversely proportional to mortality from coronary heart disease (Hertog *et al*.,1993). Flavonoids are also suggested to protect the tissues from lipid peroxidation which is thought to be involved in atherosclerosis, chronic inflammation and cancer (Le Marchand, 2002; Middleton *et al.*, 2000; Hollman and Katan, 1999). However, the studies carried out to prove anticarcinogenic and cardioprotective effects of flavonoids are not conclusive.

Flavonoids Classification

The basic structure of flavonoids consists of two benzene rings (A and B) connected through central pyrone ring (C). Flavonoids are classified into the following classes based upon the structural differences (Table 1-1)-

Flavones

The flavones structure is planar. It consists of a double bond between positions 2 and 3 of the C ring. The C-ring is connected to the B-ring at position 2. The B ring has a hydroxyl group at position 4'. Flavones are found in red peppers, apple skins and celeries (Griffiths and Smith, 1972b; Aura, 2005). The examples of flavones are luteolin and apigenin.

Flavonols

The flavonols have a planar ring structure. The B ring is connected to the pyrone ring at position 2. The C- ring fission occurs at positions between bonds 1 and 2, 3 and 4 as well as between 4 and 10. Examples of flavonols are quercetin, kaempferol and myricetin. These are found in onions, apples, red wine, broccoli and tea.

Flavanone

The Flavanone structure is not planar and lacks a double bond between positions 2 and 3 of the central pyrone ring. Flavanones are degraded through C-ring fission at positions between 1 and 2 as well as 4 and 10. Citrus fruits, grape fruits are abundant in flavanone. Examples of flavanone are hesperitin and naringenin.

Flavan-3-ols

The flavan-3-ols ring structure is not planar. The C ring of flavan-3-ols does not harbor a carbonyl group. Flavan-3-ols are found in fruits, tea, and wine. Some of the examples of flavan-3-ols are catechin, epicatechin and epigallocatechin.

Anthocyanins

The anthocyanin structure is planar. The central C ring lacks the carbonyl group found in flavones. There are double bonds between positions 1 and 2 as well as 3 and 4 of the C-ring. These are present in fruits and flowers and are responsible for fruit and flower coloration. Cherries, grapes, plums, raspberries, strawberries, peaches and apples are abundant in anthocyanins.

Isoflavones

The isoflavone structure is planar. It differs from the rest of the flavonoids since A-ring is connected to the central C-ring at position 3. Soybeans are rich in flavonoids. Examples of isoflavones are daidzein and genistein.

Microbial Modification of Flavonoids

Complex flavonoids cannot be absorbed directly in the gastrointestinal tracts of humans. Instead, they are further broken down by the gut microbiota into smaller compounds with higher biological activity (Selma *et al*., 2009). Flavonoids are usually conjugated to sugars and organic acids in their natural state. Hence, most studies on flavonoid degradation are carried out with flavonoid glycosides. These studies suggested that the first step of degradation involves deconjugation, i.e. cleavage of sugar moiety from the flavonoid producing aglycones (Aura, 2005). Flavonoiddegradation takes place through the central C-ring fission. The degradation products formed depend upon the position of the C ring cleavage and number of hydroxyl groups present (Selma *et al.*, 2009).

Some microorganisms isolated from human fecal sample are able to degrade flavonoid. For example, Winter *et al*. (1989) showed that *Clostridium* strains could cleave quercetin, kaempferol, naringenin to 3,4-dihydroxyphenyl acetic acid, 4-hydroxyphenyl acetic acid and phenylacetic acid respectively. *Eubacterium ramulus* could cleave genistein to 6-hydroxy-O-desmethylangolensin and 2-(4-hydroxyphenyl)-propionic acid *(*Schoefer *et al., 2002).* Similarly, Braune *et al*. (2001) suggested that this bacterium degrades luteolin to 3, 4-dihydroxyphenyl-propionic acid (from the B and C rings) and phloroglucinol (from A ring). Griffiths and Smith, (1972a) showed that the rat intestinal microbiota could cleave apigenin to p-hydroxyphenylpropionic acid, p-hydroxycinnamic acid and p-hydroxybenzoic acid whereas myricetin was degraded to 3-hydroxyphenyl acetic acid, 3,5-dihydroxyphenyl acetic acid and 3, 4, 5-trihydroxyphenyl-acetic acid (Griffiths and Smith, 1972b). Naringinin, a naringenin glycoside, is reported to be degraded to phloroglucinol and 3-phenylpropionic acid by colonic microflora

(Rechner *et al*., 2004). Catechin and epicatechin were cleaved to 3-hydroxyphenyl-propionic acid, 3', 4'-dihydroxyphenyl-Ɣ-valerolactone, 3-hydroxyphenyl- Ɣ-valerolactone, 3-hydroxyhippuric acid by human fecal microflora (Tzounis *et al.*, 2008). Anthocyanin is degraded to 3, 4-dihydroxybenzoic acid (Vitaglione *et al.*, 2007).

Taken together, these studies showed that the gut microbiota is involved in flavonoids degradation and suggest that degradation happens through the C-ring fission. However, the detailed mechanism is unknown.

Modification of Flavonoids by Lactic Acid Bacteria (LAB)

Lactic Acid Bacteria (LAB) is a group of bacteria that ferment hexoses primarily to lactic acid. These Gram-positive bacteria are non-pathogenic, microaerophilic or anaerobic, and do not sporulate. LAB group includes the genera *Lactobacillus* (Lb.), *Enterococcus* (E.), *Lactococcus* (Lc.), *Aerococcus* (A.), *Pediococcus* (P.), *Leuconostoc* (Ln.), *Vagococcus* (V.), *Carnobacterium* (C.), *Tetragenococcus* (T.), *Weissela* (W.) and *Oenococcus* (O.). LAB are usually found in diverse niches such as the gastrointestinal, vaginal and urogenital tracts, fermented fruits and vegetables, meat and dairy products. LAB are considered GRAS (generally regarded as safe) microorganisms, and many species are used as probiotics.

Most of the LAB are used as the starters for the dairy, meat and food fermentations (Leal-Sanchez *et.al*., 2003). Many genera are involved in the fermentation of olives, cabbage, cucumbers, eggplants, caper berries and grape must. *L. brevis* and *L. plantarum* are some of the *Lactobacillus* species isolated from these environments.

Interestingly, the flavonoid content in the berries help in inhibiting the growth of the Gram-negative and pathogenic bacteria without affecting the growth of lactobacilli (Puupponen-Pimiä *et al.*, 2001). There is scarce information on the LAB interaction with

the flavonoids. Most of the studies performed are limited to the deconjugation step. The formation of biologically active aglycone is due the high β-glucosidic activity of many LAB species (Di Cagno *et al*., 2010; Chun *et al.*, 2007; Pham and Shah, 2009; López de Felipe *et al.*, 2010). No studies are reported on the degradation of the aglycone form of the flavonoids.

Modulation of Gene Expression

Bacteria are exposed to the changing environmental conditions, such as nutrient and oxygen limitations, temperature changes and various other stress conditions. In order to quickly adapt to the new environment the genes involved in maintaining cell homeostasis, competition or survival should be expressed in temporal fashion.

In prokaryotes, protein levels can be controlled at the level of mRNA synthesis (transcriptional initiation, elongation or termination), synthesis of protein (translation) or post-translationally by protein modifications (such as phosphorylation). In bacteria, most genes are regulated at the level of transcription initiation by the use of alternative sigma factors, one component and two component systems. For example, σ^S upregulates around 500 genes involved in survival during starvation conditions and stationary phase in *Escherichia coli* (Weber *et al*., 2005).

Two-component systems are named after the two proteins involved in the pathway: a sensor kinase located inserted in the membrane and a response regulator in the cytoplasm. When an environmental signal (i.e., changes in pH, osmolarity) is recognized by the membrane bound sensor kinase, the protein is autophosphorylated. As a result, the sensor kinase transfers the phosphoryl group to the response regulatory protein, which then binds the DNA and in most cases activates transcription (Figure 1-1).

One component response regulatory systems are traditionally known as “transcription factors”. One component regulatory systems are the most abundant in prokaryotes (Ulrich *et al.*, 2005). In a one component regulatory system, the effector binding domain (EBD) and DNA binding domain (DBD) are present in the same protein. Once a signal molecule is present in the cellular environment, it interacts with the effector-binding domain of the protein. This interaction is thought to cause a structural rearrangement that leads to transcriptional activation or repression, depending on the particular gene and the environmental stimulus (Figure 1-1).

The number of transcription factors varies in different microorganisms. It was found that the percentage of transcription factors was correlated with the changing environmental conditions to which the microorganism is exposed to rather than the genome size (Ulrich *et al.*, 2005). Transcription factors are divided into more than 50 families based upon their similarity in the DBD (Minezaki *et al.*, 2006).

Transcription factors interact at the specific binding sites within the promoter region. Transcription factors can be found in different oligomeric forms such as, dimers, tetramers, or heterodimers. These proteins recognize on average 16-30 nt long palindromic or pseudopalindromic sequences on DNA (Schell, 1993; Rodionov, 2007). Binding of the transcription factor (TF) to the DNA may result in repression or activation of the gene expression.

Repressors may interact at the RNA polymerase (RNAP) binding site inhibiting the RNAP access to the promoter region by steric hindrance (Figure 1-2). For example, the binding of LacI overlaps the *lacZYA* promoter inhibiting the RNAP access to the promoter (Müller-Hill, 1996).

The repressor binding site also can be found further downstream of the promoter region that hinders the RNAP movement and blocks the transcription elongation (Figure 1-2B). For example, PurR in *E. coli* interacts at a site 242 bp downstream from transcription start site of *purB* (He and Zalkin, 1992).

Repression can also be achieved by looping mechanism which inhibits RNAP access to the promoter. In the *E. coli gal* operon, two operator sites are found. The two sites are located around -60 and +53 positions relative to the *galE* transcription start site (Müller-Hill, 1998). Repression is achieved by a looping mechanism due to the GalR repressor binding at the two operator sites.

Another mechanism of repression by steric hindrance can be exemplified by CytR. The transcription factor CytR modulates the interaction capability of the catabolite activator protein (CAP) (also known as cyclic AMP receptor protein, CRP) with the RNAP. CytR interacts at a site located between the two CAP binding sites which prevent further interactions with the α-subunit of the RNAP. Thus transcription initiation is prevented (Figure 1-2C) (Shin *et al.*, 2001). In absence of active repressor (in the presence of cytidine), CAP interaction with RNAP stabilizes RNAP binding to the promoter thus enhancing the transcription from the promoter (Figure 1-2F).

Small molecules modulate the affinity of the TF to the binding sequence. Corepressors increase the TF's affinity for the binding sequence, and inducers decrease the affinity (Figure 1-2). For example, allolactose acts as an inducer lowers LacI binding to the operator. Tryptophan acts as a corepressor by interacting with the inactive TrpR protein. The ligand interaction changes TrpR conformation which increases TrpR affinity to the P_{trp}.

Transcriptional activators usually interact upstream or adjacent to -35 elements depending upon the type of promoter. For example, CRP binds at position -41 in class II promoters (Zheng *et al.*, 2004) (Figure 1-3A). In class I promoters, CRP binds upstream of the -35 element at positions -62, -72 and -92 (Figure 1-3B). The mechanism of activation depends on the type of promoter as well. In class I promoters, activation is achieved by the interaction of the transcriptional regulator with the RNAP α- carboxy terminal domain (α- CTD). This allows the RNAP to strongly interact with the promoter (Figure 1-3B). In class II promoters gene activation is achieved by the interaction between the activator and the α-NTD as well as σ subunit of the RNAP (Figure 1-3A) (Browning and Busby, 2004).

Some transcription factors can behave as both repressors and activators under different conditions (Figure 1-4). For example, AtzR under standard growth conditions (explained below in detail) represses transcription from the promoter by inhibiting RNAP access. In the presence of cyanuric acid and under nitrogen limiting conditions, the AtzR interaction site is shifted further upstream to the -35 site. The shift in the binding site allows the RNAP access to the promoter and the genes *atzDEF* are transcribed (Porrúa *et al.*, 2010). Similarly, the OxyR transcription factor in reduced state binds at -35 site (extended form) (Figure 1-4A). Under oxidizing conditions, an allosteric change in the TF causes shift in the OxyR binding site (compact form) which allows RNAP interaction at the promoter (Toledano *et al.*, 1994) (Figure 1-4B).

Some repressors interact at several sites in the promoter and stimulate DNA looping. The DNA looping prevents RNAP access to promoter and thus inhibits transcription (Figure 1-4C). In the *araBAD* operon, AraC interacts at two sites and

bends the DNA. This looping inhibits RNAP access to the promoter (Figure 1-4C). In the presence of the inducer arabinose, the repressor changes conformation and allows activation of transcription from the promoter (Figure 1-4D). The *araBAD* operon is also under catabolite repression control (Schleif, 2003).

Another mechanism of transcription activation is found in the MerR family of transcription regulators, which act by inducing changes in the DNA conformation (Figure 1-4). Interaction of MerR with the promoter leads to a conformational change in the DNA. In the absence of Hg (II), the RNAP can weakly bind the promoter, and the open complex formation is inhibited (Figure 1-4E). In presence of Hg (II), a stable complex is formed between RNAP: DNA: MerR and Hg (II). Within the complex, the active MerR:Hg(II) reorients the promoter region to initiate transcription by allowing the formation of the open complex (Figure 1-4F) (Brown *et al.*, 2003; Heldwein and Brennan, 2001).

Some transcriptional activators do not require binding of the protein to the DNA. Under specific conditions, the TF interacts directly with the RNAP to initiate transcription without binding to the DNA. For example, the bacteriophage N4 single-stranded DNA binding protein (N4SSB) upregulates transcription from the N4 late promoters in *E. coli*. The gene expression is achieved by a direct interaction of the N4SSB protein with the carboxy terminus of the RNAP β subunit (Figure 1-3C) (Miller *et al*., 1997).

Classification of Transcription Factors

Bacterial transcription factors are classified based upon sequence identity and the presence of structural motifs involved in DNA recognition. In prokaryotes, the most common folds are the helix turn helix (HTH) and the winged helix turn helix (wHTH). The location of the DNA binding regions varies, and they can be found in the N-terminus

(i.e., Fur and LacI families), C-terminus (i.e., LuxR and OmpR families), or in the center (i.e., ArsR and MarR families) of the protein.

The HTH DNA-binding domain is the most common motif found in bacterial transcription factors. It consists of two helices, helix 2 and 3 connected through a short turn. Glycine is the most conserved residue in the turn that connects the two helices. helix 1 is necessary for stabilizing HTH structure. The helix 2 and helix 3 are positioned 120° from each other. The helix 3 interacts with the DNA at the major groove; hence it is called a recognition helix. The TetR, LacI and Fis families are some examples that contain this fold.

The winged helix turn helix motif is a variation of the HTH motif. It consists of an α/β structure comprising three helices (α1, α2, and α3), three beta strands (β1, β2, and β3), and two wings/turns (W1 and W2). The arrangement is α1-β1-α2-α3-β2-W1-β3-W2. The α2-α3 makes up the helix-turn-helix motif where the α3 is involved in the DNA interaction at the major groove (recognition helix). The structure of the wHTH domain structure resembles that of a butterfly with W1 and W2 representing the wings and α3 as the thorax. Hence, the motif is named winged helix turn helix. This fold can be found in the MarR, OmpR and IclR families.

Rodionov (2007) identified the number of transcription factors from 230 sequenced genomes of prokaryotic organisms. The predictions were based upon profile Hidden Markov Models (HMMs) (Kummerfeld and Teichmann, 2006) using the PFAM (Finn *et al.*,2006) and SUPERFAMILY (Wilson *et al.*, 2007) databases. The major transcription factor families from this study are summarized in Table 1-2. The study suggested that the LysR type transcription regulators are most abundant in prokaryotes followed by the

AraC and TetR families. However, the number of transcriptional regulators in different microorganisms do not necessarily follow a similar trend. For example, MarR and IclR are the largest families in *B. subtilis* (Moreno-Campuzano *et al.*, 2006) and *Bordetella* species (Molina-Henares *et al.*, 2006) respectively. However, in *L. brevis* the largest family is MarR (Lorca, unpublished).

Transcription Factors Responding to Flavonoids

There are very few transcription factor families known to respond to flavonoids. Most of the transcription factors that have been identified are believed to regulate genes required to survive the antimicrobial properties of flavonoids. These TFs have been discovered in several different families including TetR (LmrA, YxaF, and TtgR), MarR (YetL), LuxR (RhlR), and LysR (NodD). However, NodD is the only characterized protein that senses flavonoids in a specific mode. NodD mediates the upregulation of *Rhizobium* genes involved in nodulation.

The mode of regulation of the transcription factors responsive to flavonoids are described below.

Members of TetR Family Responsive to Flavonoids

TtgR from the non-pathogenic water and soil bacterium *Pseudomonas putida* negatively regulates *ttgABC* transcription. The *ttgR* gene is negatively autoregulated (Teran *et al.*, 2003). The TetR family members contain a N-terminal DNA binding domain and a C-terminal effector binding domain. The genes in the *ttgABC* operon encode for a multidrug efflux pump TtgABC which is involved in detoxification of structurally unrelated antimicrobial compounds like antibiotics, dyes, biocides thus conferring multidrug resistance to *P. putida* (Teran *et al*., 2003). Using gene reporter assays, it was observed that the expression of the *ttgABC* operon was also activated in

presence of the flavonoids. Direct binding experiments with TtgR using isothermal titration calorimetry (ITC) confirmed the results (Teran *et al.*, 2006). It was suggested that the response of TtgR from *P. putida* to the flavonoids is nonspecific and is functioning as a resistance mechanism to the antimicrobial properties of the flavonoids.

The *yxaF* gene found in *B. subtilis* is oriented in the same direction as that of its target genes *yxaGH* (Hirooka *et al.*, 2007). Microarray analysis and genome wide computational analysis were used to find the YxaF regulated genes. The gene *yxaG* encodes for a quercetin 2,3-dioxygenase (Bowater *et al.*, 2004), and *yxaH* encodes for an uncharacterized membrane protein. Using *yxaGH* promoter: *lacZ* fusion reporter assays it was suggested that quercetin acts as an inducer. Electrophoretic mobility shift assays (EMSA) using different flavonoids showed that fisetin, tamarixatin and galangin also inhibited the interaction of YxaF with the *yxaGH* promoter (Hirooka *et al*., 2007).

Members of MarR Family Responsive to Flavonoids

The MarR member *yetL* is divergently transcribed to *yetM,* which encodes a putative FAD-dependent monooxygenase (Hirooka *et al.*, 2009). DNA microarray analysis showed that transcription of *yetM* is enhanced in *yetL* mutants. EMSA studies suggested that YetL interacts with *yetL* and *yetM* promoters, but the addition of flavonoids prevents this interaction. These results were further supported by *yetL* and *yetM*: *lacZ* reporter assays. It was proposed that the regulation of *yetM* by YetL in response to flavonoids is a survival strategy used by *B. subtilis* in flavonoid rich environments (Hirooka *et al*., 2009).

Members of LuxR Family Responsive to Flavonoids

Vandeputte *et al*. (2010) demonstrated that bark extract from *Combretum albiflorum* could inhibit a signal transduction pathway that is mediated by quorum

sensing factors. HPLC analysis suggested that catechin was one of the active compounds in this bark extract. Using *Pseudomonas aeruginosa* and *E. coli* as biosensor strains, it was shown that catechin inhibited sensing of C4-HSL through the RhlR transcriptional regulator (Vandeputte *et al.*, 2010). In agreement with these observations, it has been demonstrated that the plant genes involved in flavonoid biosynthetic pathways were up regulated in response to homoserine lactones suggesting an antimicrobial role for flavonoids (Mathesius *et al.*, 2003).

Members of LysR Family Responsive to Flavonoids

NodD from *Rhizhobium leguminosarum* responds to flavonoids by inducing the expression of multiple genes involved in promoting nodulation in leguminous plants. There is a symbiotic association between *Rhizhobium* and plants. The plants get benefited due to the bacterial fixation of atmospheric nitrogen to ammonia (N_2 to NH_3), and the bacteria receive a carbon source from the plant. Multiple copies of the *nodD* gene are present in different strains of Rhizobia. One or two copies are present in *R. leguminosarum* bv. viciae and *Rhizobium sp* strain NGR234, respectively (Downie *et al.*, 1985). Three copies are present in *R. leguminosarum* bv. Phasoeli (Davis and Johnston, 1990). While most of the NodD proteins respond to flavonoids, in *R. meliloti* it was found that NodD3 does not require an effector molecule to activate gene expression (Swanson *et al.*, 1993).

Based on EMSA and DNase I footprint assays, it was suggested that the interaction of NodD3 with DNA binding sites (nod boxes) leads to DNA bending, and this activates transcription from the promoter (Fisher and Long, 1993). However, the effect of the flavonoids was not evident in this case, possibly because the overexpression of NodD3 was sufficient for the up regulation of *nodABC* genes

(Mulligan and Long, 1989). The ability of NodD to react to different flavonoids is related to the plant host range of the bacteria (Broughton and Perret, 1999).

It was also found that *nod* genes are also under negative regulation. In *R. meliloti*, NolR, a NodD homologue, negatively regulates the genes involved in nodulation. In addition, NodD interacts with a 47 bp sequence (containing the palindromic sequence ATC-N9-GAT) in *R. leguminasarum*. (Wang and Stacey, 1991; Fisher and Long, 1993). Variation from this motif in different strains also has been reported. Understanding of NodD interaction with flavonoids is mostly based upon genetic analysis and mutational studies. Thus, there are no direct binding experiments supporting that flavonoids bind to NodD (van Rhijn and Vanderleyden, 1995).

The LysR Family of Transcription Regulators (LTTRs)

Introduction

LTTRs are one component transcription factors with a DNA binding domain and effector binding domain in the same regulatory protein. LTTRs are the most abundant transcription factors in prokaryotes (Pareja *et al.*, 2006; Rodinov, 2007). They are involved in regulating genes required for diverse activities. Some examples of genes regulated by LTTRs include virulence, sporulation, antibiotic resistance, and DNA replication genes (Maddocks and Oyston, 1993).

Domain Arrangement of LTTRs

The domain arrangement is shown in Figure 1-5. The 90 amino acids in the N-terminal in LTTRs form the DNA-binding domain. Within this domain, residues 23 to 43 form the wHTH (Schell, 1993; Maddocks and Oyston, 2008). The wHTH motif is connected to linker helix through hinge 1. Hinge 1 allows flexibility to wHTH during interaction with DNA. Most LTTRs are 300 amino acids long with residues 90 to 300

being involved in the formation of the regulatory domain. The wHTH is connected to the regulatory domain by a linker helix through hinge 2. The regulatory domain is divided into two subdomains: regulatory domain I (RDI) and regulatory domain II (RDII) (Muraoka *et al.*, 2003; Tyrrell *et al.*, 1997). Usually the last 30-40 amino acids (C-terminal) are also part of RDI. Although the exact mechanism of this interaction is not known, it has been suggested that the C-terminal 40-60 amino acids are involved with DNA interaction (Schell *et al.*, 1990; Lochowska *et al.*, 2001). RDI and RDII are connected by two crossover regions (in Figure1-5 these are labeled 3a and 3b), which form a hinge or cleft region (hinge 3) (Muraoka *et al.*, 2003). It is suggested that residues from 95 to 210 are involved in ligand interaction in this cleft (Jorgensen and Dandanell, 1999). Hinge 3 is also suggested to play a role in structural rearrangements after the ligand interaction (Muraoka et al., 2003; Zhou *et al.*, 2010).

LTTRs DNA Binding Sites

LTTRs usually interact at two sites on the DNA, which are called the activator binding site (ABS) and the repressor binding site (RBS). When the LTTR mediates gene activation, the ABS is generally found near the -35 position, and the RBS is located near the -65 position, with respect to the transcription start site. Because in most cases the TF is divergently transcribed from the regulated genes, the RBS overlaps the promoter for the LTTR (Figure 1-6). However, binding sites have sometimes been found as far downstream as the +350 position of the regulated gene (Porrúa *et al.*, 2007; Viswanathan *et al.*, 2007b). The sequence at RBS is a palindromic T-N11-A (ATC-N9-GAT); although variation in this sequence has been reported (Toledano *et al.*, 1994; Hryniewicz and Kredich, 1995; Grob and Guiney, 1996; Lehnen *et al.*, 2002). In contrast, the ABS sequence is highly variable. The ABS is a secondary site for LTTR

interactions, and studies show that it is involved in activation. On the other hand, the RBS is generally a strong binding site and is required for gene repression (Porrúa *et al.*, 2009). LTTRs usually recognize the dyad symmetry in the binding sequences. However, it has been suggested that the direct repeats within the promoter region of NodD are involved in the binding of NodD to the promoter (Wang and Stacey, 1991; Fisher and Long, 1993).

Mode of Regulation

Mode of Gene Regulation by LTTRs

According to the classical model of gene regulation by LTTRs, the gene encoding the LTTR is constitutively transcribed. Once the LTTR is present in the cellular environment, it binds in the promoter of the "target" gene at the ABS and RBS (Figure 1-6). Next, the two dimers interact with each other, promoting DNA bending. As a result, gene expression is inhibited. This model has been inferred from multiple studies, including those performed with CysB (Hryniewicz and Kredich, 1994), AtzR (Porrúa *et al*., 2010), and OxyR (Toledano *et al.*, 1994). Once an inducer molecule is present in the cellular environment, it interacts in the hinge 3 region of the TF between regulatory domain I and regulatory domain II (Muraoka *et al*., 2003; Parsek *et al.*, 1992; Craven *et al.*, 2009; Tyrrell *et al.*, 1997) leading the LTTR structural rearrangement. The structural rearrangement causes the dimer overlapping the target gene promoter to shift further upstream. This movement results in a relaxation of the DNA bend as a result the RNA polymerase is able to access the promoter which allows transcription initiation from the promoter. This mechanism explains the role of LTTRs as transcriptional activators.

Recently, the structure of ArgP (Zhou *et al*., 2010) was elucidated, which explained at the molecular level the results proposed in the classical model of regulation

discussed above (Figure 1-6). ArgP is a tetramer in solution. In the crystal structure of ArgP, two forms of dimers were observed: one dimerized through DNA-binding domains and the other through regulatory domains. The authors proposed that ArgP interacts with its promoter as a dimer at a high affinity site (RBS). This favors the interaction with a second dimer through the C-terminal regulatory domains of ArgP, and thus it favors tetramerization. The second dimer in the tetramer interacts with the low affinity site (ABS) in the target gene promoter region (extended form of the protein). It was also observed that the DNA binding domains in ArgP (as discussed in Chapter 4) were along one face of the tetramer, and the regulatory domains were present on the other. Within the regulatory domain, a cleft that can accommodate the ligand was observed. However, the effector was not co-crystallized with ArgP. It was suggested that once the ligand binds to the ArgP effector binding doman, it would shift the ArgP binding sites in the ABS and allow the RNAP to access the promoter. These observations are supported by results seen in nitrogen assimilation control protein (NAC) from *Klebsiella pneumoniae*. In NAC, it was suggested that the two forms of the tetramers, compact and extended, were present in the absence of physiologically relevant coinducers. It was proposed that NAC could adopt two forms to recognize different promoter binding sites under different physiological conditions (Rosario *et al.*, 2010).

In a few cases the small molecules that modify the activity of LTTRs transcription factors have been elucidated. In some cases, like that of NodD3 (Mulligan and Long, 1989), LrhA (Lehnen *et al*., 2002), SpvR (Grob and Guiney, 1996), it has been suggested that an abundance of the transcription factor was sufficient for transcriptional activation. These conclusions were based on the target promoter:*lacZ* fusion studies.

Many genes regulated by LTTRs have been identified (For a review, see Schell, 1993; Maddocks and Oyston, 2008). Some representative examples were chosen to illustrate the different modes of regulation discovered.

Negative Gene Regulation by LTTRs

LrhA regulation of the *flhDC* operon

LrhA is a member of a small group within LTTRs which negatively regulates their target genes (Lehnen *et al*., 2002). The *flhDC* operon is involved in flagella biosynthesis. LrhA negatively regulates the expression of these genes by binding to the promoter region. DNase I footprint assays showed that LrhA interacts at -89 to -129 positions, relative to transcription start site by binding to the sequence ATGACTTATACAT (AT-N9-AT). The effector molecule that modulates the binding of LrhA to this promoter have not been identified. A translational fusion of *flhDC* to *lacZ* showed that *flhDC* expression was up regulated in a *lrhA* mutant. Similar upregulation was observed in genes involved in motility, chemotaxis, and flagella biosynthesis. However, EMSA analysis showed that LrhA only interacts with P_{lrhA} and P_{flhD} promoters, and the upregulation of the other genes was indirect.

HexA regulation of virulence genes in *Erwinia carotovora*

HexA is responsible for regulating the expression of several genes involved in the virulence of the plant pathogen *Erw. carotovora*. HexA has high sequence identity (64%) to LrhA from *E. coli*. Unlike classical LTTRs, *hexA* is not associated with the target genes. In *hexA* mutants, an increase in production of pectate lyase, protease, and cellulase enzymes as well as higher motility was observed. EMSA experiments showed that repression is achieved by binding of HexA in the *pelC* promoter, the promoter for pectate lyase (Harris *et al.*, 1998).

Positive Gene Regulation by LTTRs

AtzR regulation of *atzDEF*

This is an example of classical mode of gene regulation by LTTRs. AtzR from *Pseudomonas* sp. ADP acts as an activator for the genes *atzDEF* in response to cyanuric acid and under nitrogen limiting conditions (Porrúa *et al*., 2007). DNase I footprint assays showed that AtzR interacted with a binding region in the P_{atzDEF} at one binding site. This binding site was named as activation binding site (ABS) since it is located around -35 region of the P_{atzDEF}. Three sequences within the ABS were found and named as ABS1, ABS2 and ABS3 In addition, AtzR also bound to a sequence within the divergently oriented P_{atzR}. The site overlapping P_{atzR} was named as repression binding site (RBS) since it is involved in *atzR* auto repression (Porrúa *et al*., 2007). Further hydroxyl radical footprint suggested that ABS and RBS are located on the same face of the helix (Porrúa *et al*., 2010). Based upon *atzD-lacZ* fusion studies, the investigators found that ABS3 situated immediate downstream from the -35 element of P_{atzDEF} was mainly responsible for the transcription repression of this promoter. The sequences at the three sites were GTCG-N6-GGCG-N7-AGTG (the ABS1, ABS2, and ABS3 are underlined). The sequence at the RBS was GTGC-N7-GCAC (Porrúa *et al*., 2010).

As per the proposed mechanism of regulation, under regular conditions, AtzR binds to ABS2 and ABS3 sequences (the extended form of the protein). Once the effector molecule cyanuric acid is added, AtzR binds preferentially to ABS1 and ABS2 (the compact form of the protein). This structural change allows the RNA polymerase access to the promoter, thus enhancing gene transcription (Porrúa *et al*., 2010).

SpvR regulation of *spvABCD*

The *spv* regulon from *Salmonella dublin* consists of *spvR*, encoding LTTR homologue and four structural genes, *spvABCD,* which are up regulated in the stationary phase (Coynault *et al*., 1992). The biological role of the *spvABCD* genes is not known. The *spvR* is located upstream of *spvABCD,* and it is oriented in the same direction as *spvABCD* (Coynault *et al.*, 1992; Fang *et al.*, 1991). The *lacZ* studies suggested that SpvR is required for *spvA* expression (Caldwell and Gulig, 1991; Fang *et al.*, 1991; Krause *et al.*, 1992). DNase I footprint assay showed that SpvR interacts at a 54 bp region spanning from position -54 to -79 at the P_{spvA}. The binding sequence at P_{spvA} is TG**T**GC-N7-GC**A**CA, which is consistent with the LTTR binding consensus sequence at RBS (T-N11-A) (Grob and Guiney, 1996). Based upon presence of two complexes in EMSA studies, it was suggested that SpvR interacts at the two sites. The disappearance of the faster migrating complex at higher protein concentrations suggested that cooperativity might be involved in this interaction (Grob and Guiney, 1996).

Northern blot assay studies suggested that *spvABCD* forms an operon with two transcriptional start sites, at 70 and 98 bp upstream of *spvA* translation start site (Krause *et al*., 1992). The *spvR* gene is monocistronic, while *spvABCD* forms a second transcriptional unit (Abe and Kawahara, 1995).

LTTR Autoregulation

Much of the attention in the research field has been directed towards the target genes regulated by LTTR members rather than to the LTTR autoregulation. It is already accepted that most of the LTTRs are negatively auto regulated whereas LTTRs act as activators for the target genes (Maddocks and Oyston, 2008; Schell, 1993)(Figure 1-6).

Negative Autoregulation

Most of the LTTRs are negatively auto regulated. Some of the examples are described below.

AtzR

As described above, AtzR from *Pseudomonas* sp. strain ADP activates transcription of the *atzDEF* operon (Porrúa *et al*., 2007). AtzR also acts as the repressor at its own promoter. AtzR is divergently transcribed from the target genes (Porrúa *et al*., 2009). DNase I footprint assays suggested that AtzR-His_6 protected the promoter from positions -41 to -14 (RBS), relative to *atzR* transcription start site. In addition, protection was also seen from positions -71 to -42 (ABS, as described previously). Based upon the presence of an alternative sigma factor (σ^N) recognition sequence (CGGCAC-N5-TTGCT), the authors suggested that the P_{atzR} was recognized by RNA polymerase with an alternative sigma factor (σ^N). DNase I footprint studies showed that AtzR competed with the RNA polymerase for binding at the P_{atzR}. *In vitro* transcription assays suggested that AtzR repressed transcription from P_{atzR} in a concentration dependent manner and prevented open complex formation. As a consequence, transcription from P_{atzR} is repressed (Porrúa *et al*., 2009).

The sequence at RBS was found to be GGTGCCG-N5-CGGCACC. Deletion of the ABS relieved AtzR autorepression, and this suggests that ABS has a role in AtzR autorepression. AtzR represents the classical mode of negative autoregulation by LTTRs.

CysB

CysB from *Salmonella typhimurium* positively regulates the cystein biosynthesis genes in response to N-acetyl-L-serine, and is negatively auto regulated (Ostrowsky

and Kredich, 1990; Kredich, 1992). DNase I footprint suggested that CysB protected the *cysB* promoter from positions -10 to +36 (Ostrowsky and Kredich, 1991) overlapping the transcription start site. Hydroxy radical footprint showed that CysB binds the DNA at two divergently oriented half sites, one situated from -8 to -26 positions with the sequence TCAGATATAATGATATAG and another from position +15 to +33 with sequence TTATTATTAAATCGTATTA. These two sequences are separated by around 21 bp corresponding to two helical turns. After adding the ligand N-acetyl-L-serine, no difference in the hypersensitive sites were found indicating the DNA bending was not involved. It was suggested that the addition of the effector simply weakens CysB interaction with the *cysB* promoter. These conclusions were supported by earlier *in vivo* studies using *cysB-lacZ* fusions (Bielinska and Hulanicka, 1986) and *in vitro* transcription assays showed that CysB inhibited the transcription from the *cysB* promoter while the addition of the effector resulted in increased transcription from the promoter (Ostrowsky and Kredich, 1991). While most of the negatively autoregulated LTTRs are believed to be divergently transcribed from the target genes, *cysB* is one of the exceptions to this rule. The CysB regulated genes are not present in the vicinity of *cysB* gene.

Positive Autoregulation

There is a very small group of positively autoregulated LTTRs. In contrast to the negatively autoregulated LTTRs, positively autoregulated LTTRs are not divergently transcribed from the modulated gene. These are usually not located in close proximity of the genes under its regulation.

SpvR

As described earlier, the *SpvR* gene has the same orientation as *spvABCD* and *spvR*. However, the levels of induction of *spvR* are low (3 to 10 fold) when compared to the *spvABCD* genes (100 fold). *SpvR-lacZ* translational fusion studies suggested that spvR is positively autoregulated as a Δ*spvR* mutation results in no expression from *spv* promoter (Abe *et al.*, 1994). The expression of *spvR* gene, alike *spvABCD* operon, is dependent on RpoS, a class II sigma factor (σ^S) (Kowarz *et al.*, 1994; Chen *et al.*, 1995). Class II sigma factor (σ^S) is responsible for recognition of the promoters for the genes expressed in stationary phase and during starvation in *E. coli*. DNase I footprint studies showed that SpvR protected the *spvR* promoter from -23 to -72 containing two spvR binding sites. The region had TGTGC-N7-GGTCA (T-N11-T instead of the conserved T-N11-A) SpvR binding motif (Grob and Guiney, 1996). In addition to these sequences, random PCR mutagenesis experiments and *lacZ* reporter assays suggested that the two palindromic sequences (TNTGCANA) present within this protection region were necessary for optimal recognition of SpvR binding sites. One of the palindromic sequences was present in the spvR recognition motif whereas another was 21bp downstream from the first motif (Grob *et al.*, 1997).

LrhA

LrhA negatively regulates *flhDC* operon involved in flagella biosynthesis in *E. coli*. The genes regulated by LrhA were identified by using DNA microarray studies. LrhA is not encoded in the vicinity of target genes and is present at a different location in genome (Lehnen *et al*., 2002; Bongaerts *et al.*, 1995). *lrhA-lacZ* translational fusion studies showed that *lrhA* is downregulated in *lrhA* mutant strain. LrhA binds to *lrhA* promoter as evidenced by EMSA. DNase I protection assay showed a protected region

from positions -194 to -226 relative to transcription start site (Lehnen *et al*., 2002). The LrhA binding site has AT-N9-AT sequence which is different from the known T-N11-A site. The signal molecule is not identified for LrhA. The positive autoregulation mechanism by LrhA is not understood completely (Lehnen *et al*., 2002).

Project Design and Rationale

For the present study *Lactobacillus brevis* ATCC 367 was considered as the test microorganism. *L. brevis* has been isolated from several flavonoid rich environments such as decaying plants, fruits and grain fermentations (Kandler and Weiss, 1986). These facts indicate the *L. brevis* has the potential to sense and respond to flavonoids. It is proposed that "*L. brevis* ATCC 367 has the ability to sense and respond the flavonoids mainly through LTTRs".

The first objective of this study was to identify the transcription factors classified within the LTTR family that are able to respond to flavonoids. This was achieved through a fluorescence based ligand screening assay (Vedadi *et al*., 2006; Niesen *et al*., 2007). This *in vitro* study was carried out using the purified LTTRs from *L. brevis* and variety of flavonoids with different chemical scaffolds. In order to identify and characterize the promoter region: transcription factor interaction, electrophoretic mobility shift assays (EMSAs) were carried out using different fragments around the transcription factor genomic context. Sequence based DNase I footprint study was used to determine LTTR binding sites.

The second objective of this project was to assess the effects of the identified flavonoids on gene expression. The *in vitro* effects were confirmed *in vivo* with qRT-PCR and *lacZ* reporter assays. The *in vitro* characterization was performed using EMSAs and size exclusion chromatography.

The third objective of this study was to identify the LTTR critical amino acids involved in ligands recognition. For this objective, bioinformatics analyses were carried out. A structural model of the LTTR was constructed and putative amino acids involved in ligands interaction were identified. The involvement of these residues in binding to flavonoids was determined by site directed mutagenesis, and the effects of these mutations on ligand interactions were determined by EMSAs.

Table 1-1. Flavonoids classification: Flavonoids are classified in six classes. Flavonoids structure consists of two benzene rings (A and B) connected through central pyrone (C) ring. The flavonoids members differ based upon the R1, R2, and R3 groups substitution.

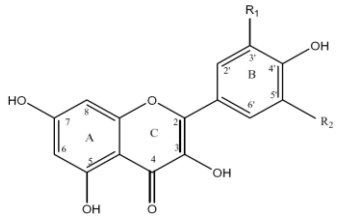
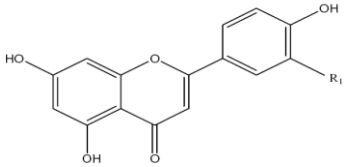
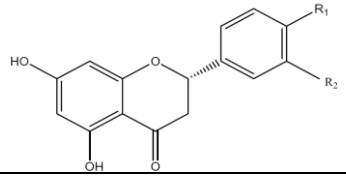
Flavonoid class	Structure	Examples	R1	R2	R3
Flavonol		Isorhamnetin	OMe	H	-
		Kaempferol	H	H	-
		Myricetin	OH	OH	-
		Quercetin	OH	H	-
Flavone		Apigenin	H	-	-
		Luteolin	OH	-	-
Flavanone		Eriodictyol	OH	OH	-
		Hesperetin	OMe	OH	-
		Naringenin	OH	H	-

Table 1-1. Continued

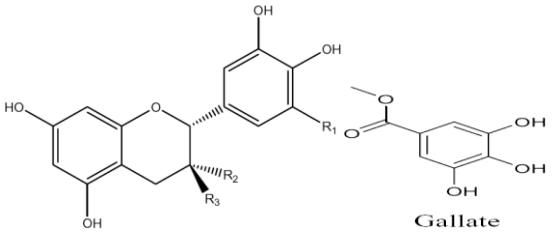
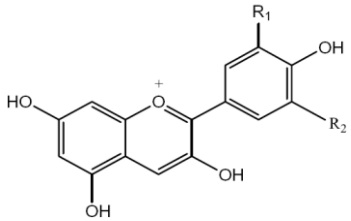
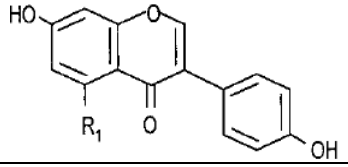
Flavonoid class	Structure	Examples	R1	R2	R3
Flavan-3-ol		(+)-Catechin	H	H	OH
		(+)-Catechin -3-gallate	H	H	Gallate
		(-)-Epicatechin	H	OH	H
		(-)-Epicatechin -3-gallate	H	Gallate	H
		(-)-Epigallocatechin-3-gallate	OH	OH	H
		(-)-Epigallocatechin	OH	Gallate	H
Anthocyanidin		Cyanidin	OH	OH	-
		Delphinidin	OH	OH	-
		Malvidin	OMe	OMe	-
		Pelargonidin	H	H	-
		Peonidin	H	OMe	-
		Petunidin	OH	OMe	-
Isoflavones		Daidzein	H	-	-
		Genistein	OH	-	-

Table 1-2. Major transcription factors families in prokaryotes. Modified from Rodionov (2007).

Family	Examples	DBD[a]	pos[b]	mod[c]	Function	TF #s[d]
AraC	MelR, RhaS, XylR, MarA, SoxS, RhrA	H	C	A	Carbon metabolism, cell wall synthesis, stress responses	6954
ArsR	CadC, CzrA, NmtR, SmtB, ZiaR	w	CR	R	Homeostasis of transition metals (Cd, Co, Zn, Ni, Zn, As, Pb)	982
AsnC	Lrp, BkdR, PutR	w	N	D	Amino acid metabolism	1527
Cro	Cro, CI, CopR, Xre	H	N	R	Bacterial plasmid copy number control	5258
Cold shock	cspB, cspC, cspD, cspE , cspF	CSD	V	A	Low temperature adaptation	607
Crp	Fnr, Dnr, NtcA, PrfA, CooA, HcpR,	w	C	A(R)	Global regulator protein. Catabolite response	891
DeoR	GlpR, AgaR, IolR	w	N	R	Carbohydrates utilization	915
Fis	NtrC, NifA, NorR, FhlA, TyrR, PrpR	H	C	A	Nitrogen, amino acid, and secondary metabolism, flagella (σ54-dependent)	2843
Fur	Zur, Mur, Nur, Irr, PerR	w	N	R(A)	Metal ion homeostasis (Fe, Zn, Mn, Ni), peroxide stress	888
GntR	AraR, ExuR, DgoR, TreR, FadR, HutC, CitR, PdhR, BioR	w	N	R	Carbohydrates, fatty acid and amino acid utilization, biotin metabolism	4293

W-winged helix-turn-helix, H-Helix turn helix, m-miscellaneous, CSD- cold shock domain, C-C terminal, N-N terminal, CR-center, A-Activator, R-repressor, D-dual role. A(R)- mostly activator, R(A)- mostly repressor. [a]DNA Binding Domain, [b]DNA binding domain position, [c]mode of gene regulation, [d]Transcription factors numbers.

Table 1-2. Continued

Family	Examples	DBD[a]	pos[b]	mod[c]	Function	TF #s[d]
IclR	KdgR, PcaR, AllR, MhpR	w	N	R (A)	Sugar, acids and aromatic compounds utilization, secondary metabolism	1122
LacI	GalR, CcpA, CytR, NagR, ScrR, PurR ,ScrR, PurR	H	N	R	Carbohydrates utilization, catabolite repression, purine metabolism	2000
LuxR	RhlR, TraR, ComA, NarP, NarL, FixJ	m	C	A(R)	Quorum sensing, competence, nitrogen oxides metabolism, anaerobic switch	3706
LysR	IlvY, CysB, MetR, CynR, NodD, AmpR, SpvR, CatR	w	N	D	Amino acid biosynthesis, nodulation, antibiotic resistance, virulence, Aromatic compounds utilization	9421
MarR	SlyA, PecS, AdcR, BadR, HucR	w	CR	R	Multiple antibiotic resistance response, Zn uptake	3280
MerR	GlnR, TnrA, SoxR, BmrR,CueR, CadR, PbrR, ZntR	H	N	R	Nitrogen metabolism, response to stress, multidrug efflux, heavy metal resistance (Hg, Cu, Cd, Pb, Zn)	2337
OmpR	ArcA, PhoB, CiaR, ToxR, VirG	w	C	A	Biosynthesis of membrane components, phosphate metabolism, competence, virulence	5010
ROK	NagC, XylR, Mlc	H	N	R	Carbohydrates utilization	1198
RpiR	HexR	H	N	R	Carbohydrates utilization	636
Rrf2	IscR, NsrR, RirA	H	N	R	FeS cluster, iron, nitrogen metabolism	818
TetR	AcrR, QacR, FabR, RutR, BioQ	H	N	R	Antibiotic resistance, fatty acids, pyrimidine, and biotin metabolism	6190

W-winged helix-turn-helix, H-Helix turn helix, m-miscellaneous, CSD- cold shock domain, C-C terminal, N-N terminal, CR-center, A-Activator, R-repressor, D-dual role. A(R)- mostly activator, R(A)- mostly repressor. [a]DNA Binding Domain, [b]DNA binding domain position, [c]mode of gene regulation, [d]transcription factor numbers.

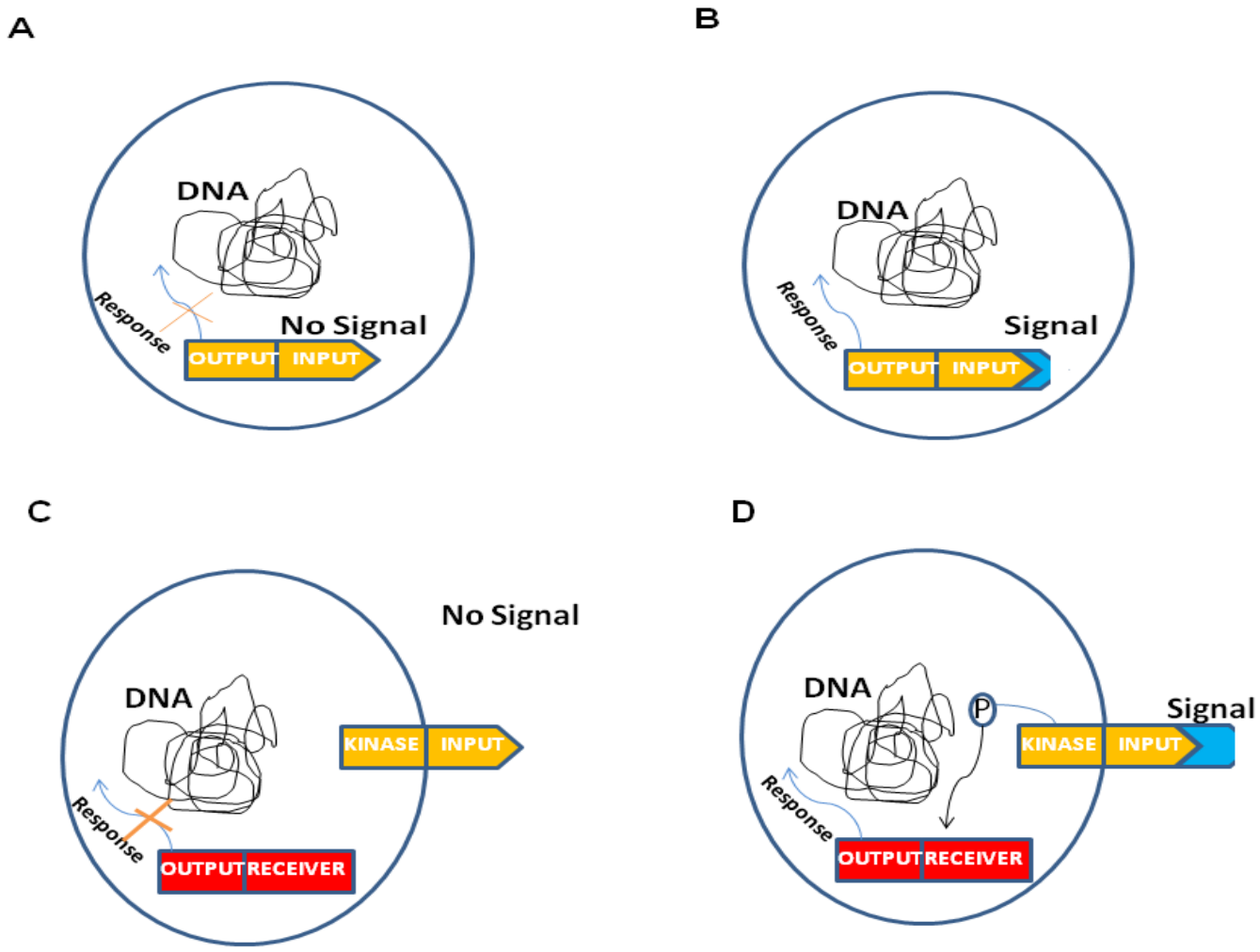


Figure 1-1. One component and two component response regulators: A) In one component response regulators, the input (sensor or EBD) and output (DBD) domains are present in one protein. B) Once a ligand or signal molecule is present in the cellular environment it interacts with the input domain. This interaction activates the output (DBD) domain thus activating or repressing transcription from the promoter. C) In two component response regulators Input (sensor kinase) domain is usually present in the cell membrane. D) Once the ligand or effector molecule is present in the cellular environment the sensor kinase is autophosphorylated which eventually phosphorylates the cytoplasmic output domain (DBD) protein. Phosphorylation of DBD protein activates the DBD, which leads to DBD interaction with the promoter region to activate or repress transcription.

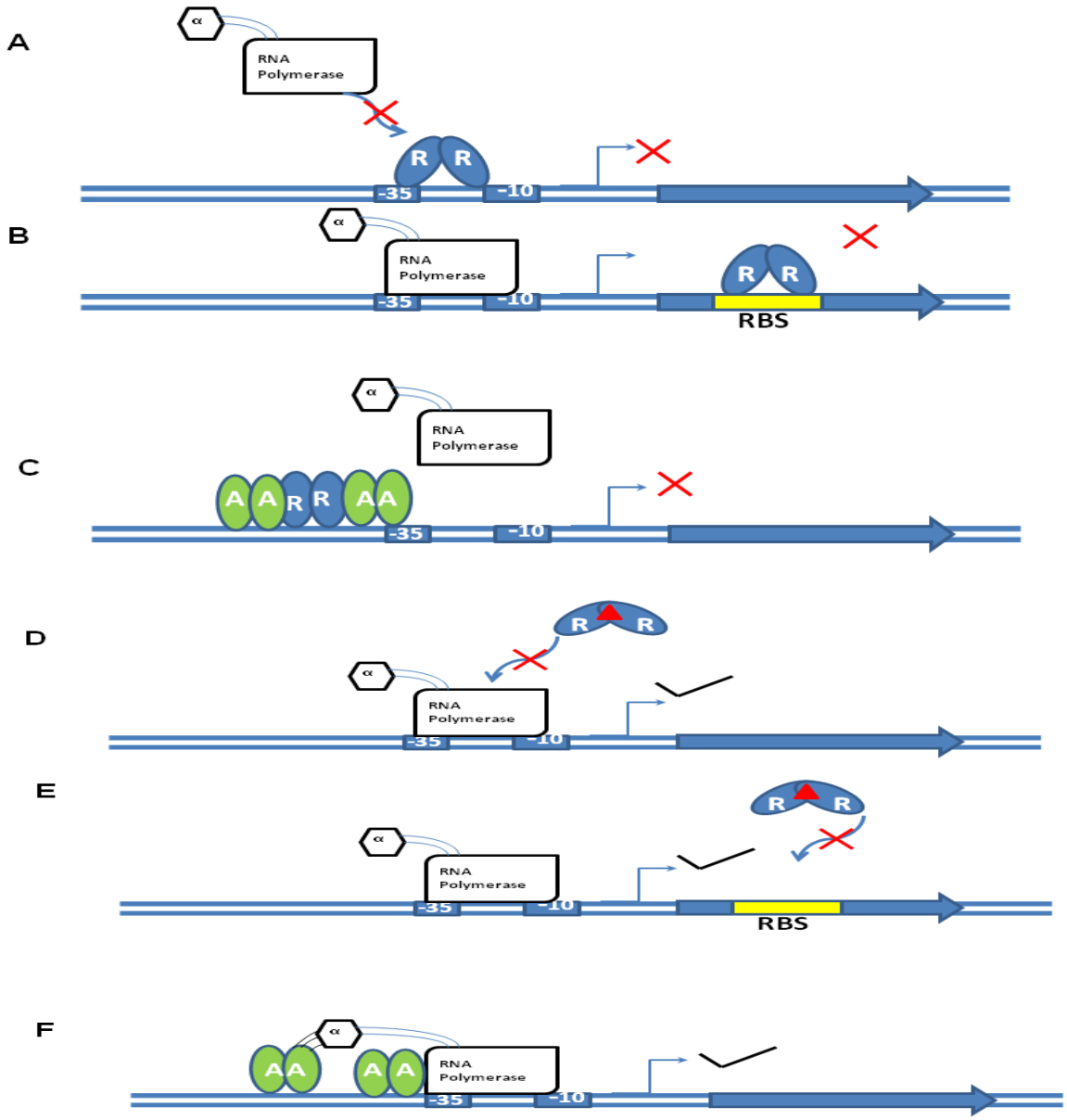


Figure 1-2. Mechanism of transcriptional repression: A) Repressor (sometimes in presence of corepressor) binds at the RNA polymerase binding site in the promoter and hinders RNA polymerase interaction. B) Repressor interaction at a site in the gene/operon blocks transcription elongation, C) The repressor interaction with an activator modulates the activator activity, which affects promoter recognition by RNAP D) The inducer (shown as red triangle) and transcription factor interaction causes a conformational change inhibiting the repressor binding at the promoter region or E) Repressor binding sites (RBS) F) In the absence of active repressor, the interaction of the activator with RNAP stabilizes the interaction of RNAP with the promoter. R-repressor, A-activator, α- α carboxy terminal domain (α-CTD), RBS-repressor binding site.

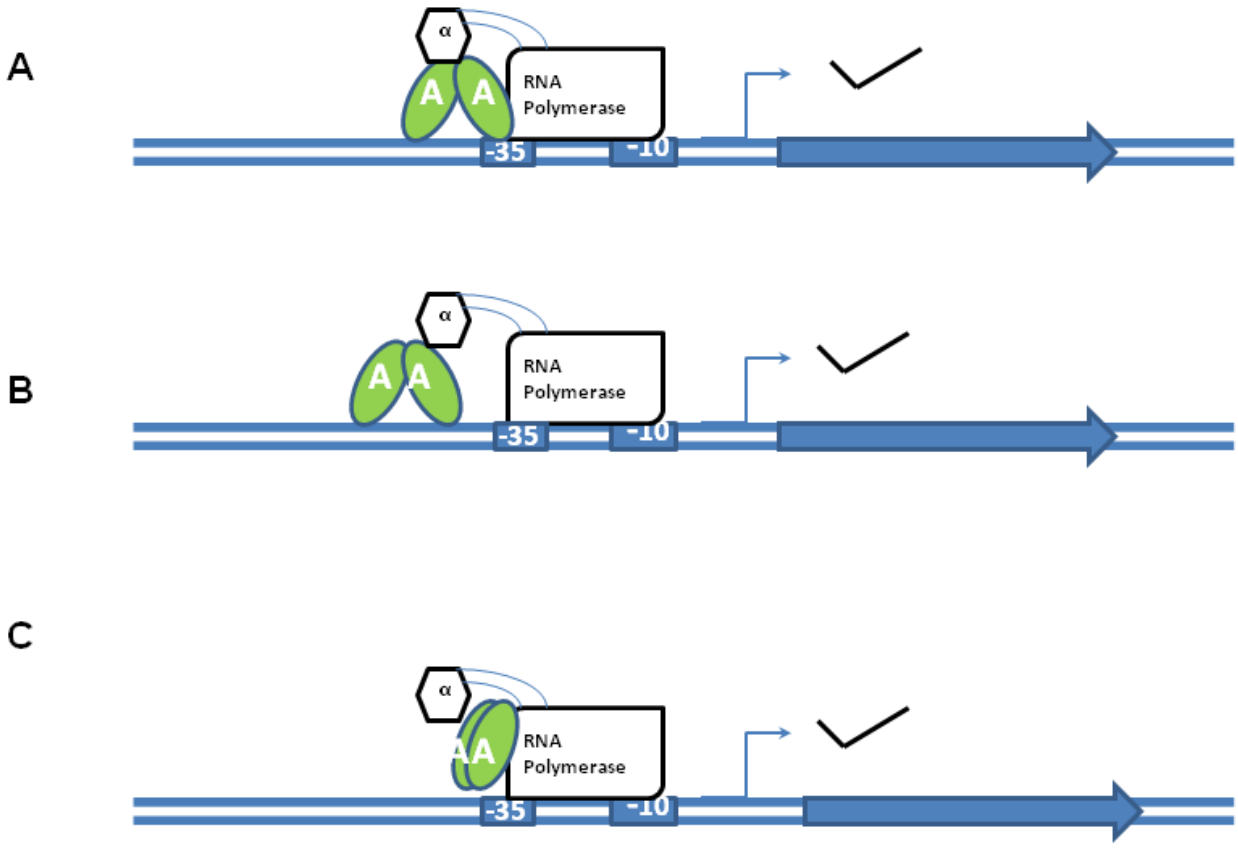


Figure 1-3. Mechanism of transcriptional activation: A) TFs that act as transcriptional activators and bind at sequences adjacent to the -35 region of the promoter (class II promoter) or B) upstream of the -35 region (class I promoter). Transcriptional activation is achieved through the interaction of the TF with RNAP. C) Some transcriptional activators do not bind DNA but rather facilitate recognition of promoter by RNAP.

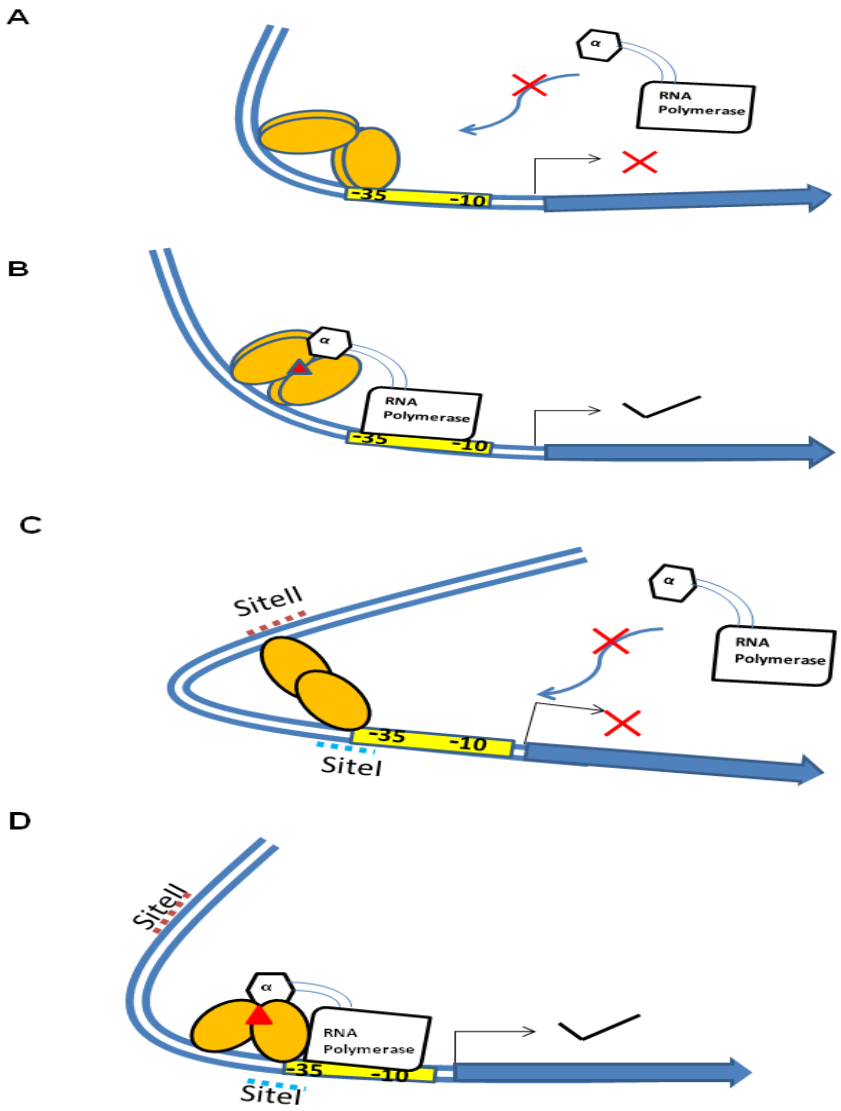
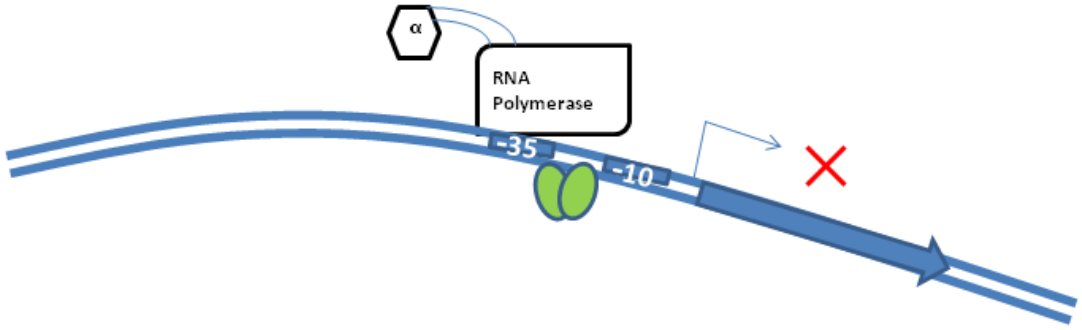


Figure 1-4. Dual role of transcription factors. A) In the absence of an inducer, the repressor overlaps the -35 region (extended form of TF), which blocks RNAP access to the promoter. B) The inducer molecule changes the repressor conformation (compact form of TF) which facilitates the interaction of RNAP with the promoter through a shift in the binding site to a site further upstream of the -35 region. C) Repressor interaction at two sites results in DNA looping which blocks RNA polymerase access to the promoter. D) In the presence of the inducer the repressor conformation is changed which facilitates RNAP interaction with the promoter E) The interaction of the transcription factor at the promoter leads to a conformational change in DNA, which inhibits open complex formation with RNAP. F) In presence of the inducer transcription factor facilitates the open complex formation by reorienting the DNA structure.

E



F

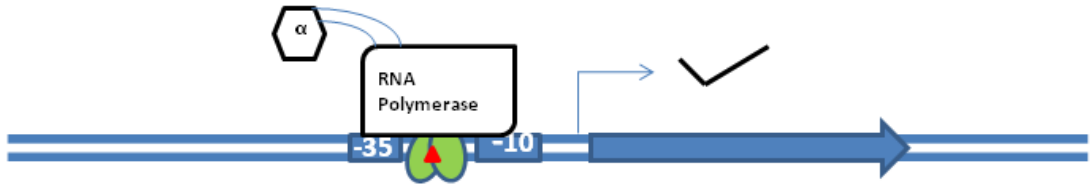


Figure 1-4 continued

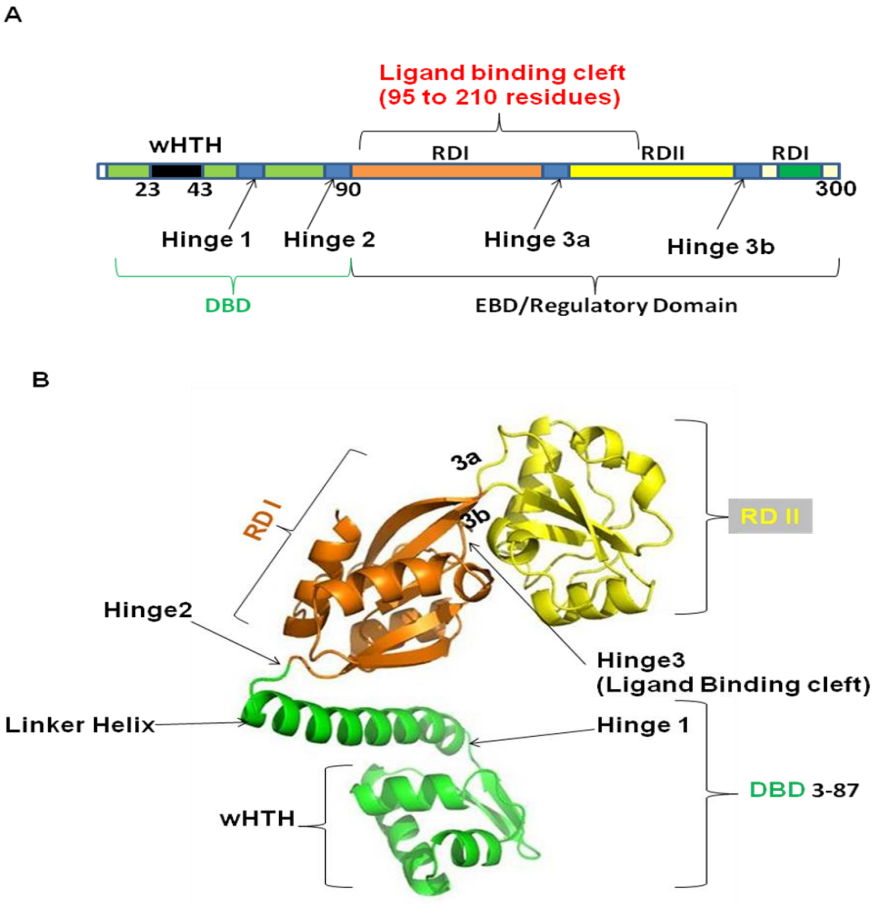


Figure 1-5. LTTR domains arrangement: A) The DBD is formed from the first 90 amino acids (green box). Within this region, 23 to 43 residues form a wHTH DNA binding motif (Black Box). HTH is the most conserved motif among LTTR members. The effector binding domain/ regulatory domain (EBD/RD) is formed from the last 90-300 residues. The RD is divided into two domains: the regulatory domain I (RDI) and the regulatory domain II (RD II). C-terminal amino acids are also part of RD I. C-terminal 40-60 amino acids are also thought to be involved in the DNA interaction (green). Residues 95 to 210 are involved in substrate/ligand interaction. B) The cartoon diagram shows LTTRs domains arrangement. DBD is formed from the first 90 amino acids. wHTH is connected to the regulatory domains through a linker helix. Hinge 1 connects the wHTH to the linker helix, whereas hinge 2 connects the linker helix to regulatory domain I (RDI). Hinge 3 is formed from crossover region between RDI (3a) and RDII (3b) forms the ligand binding cleft.

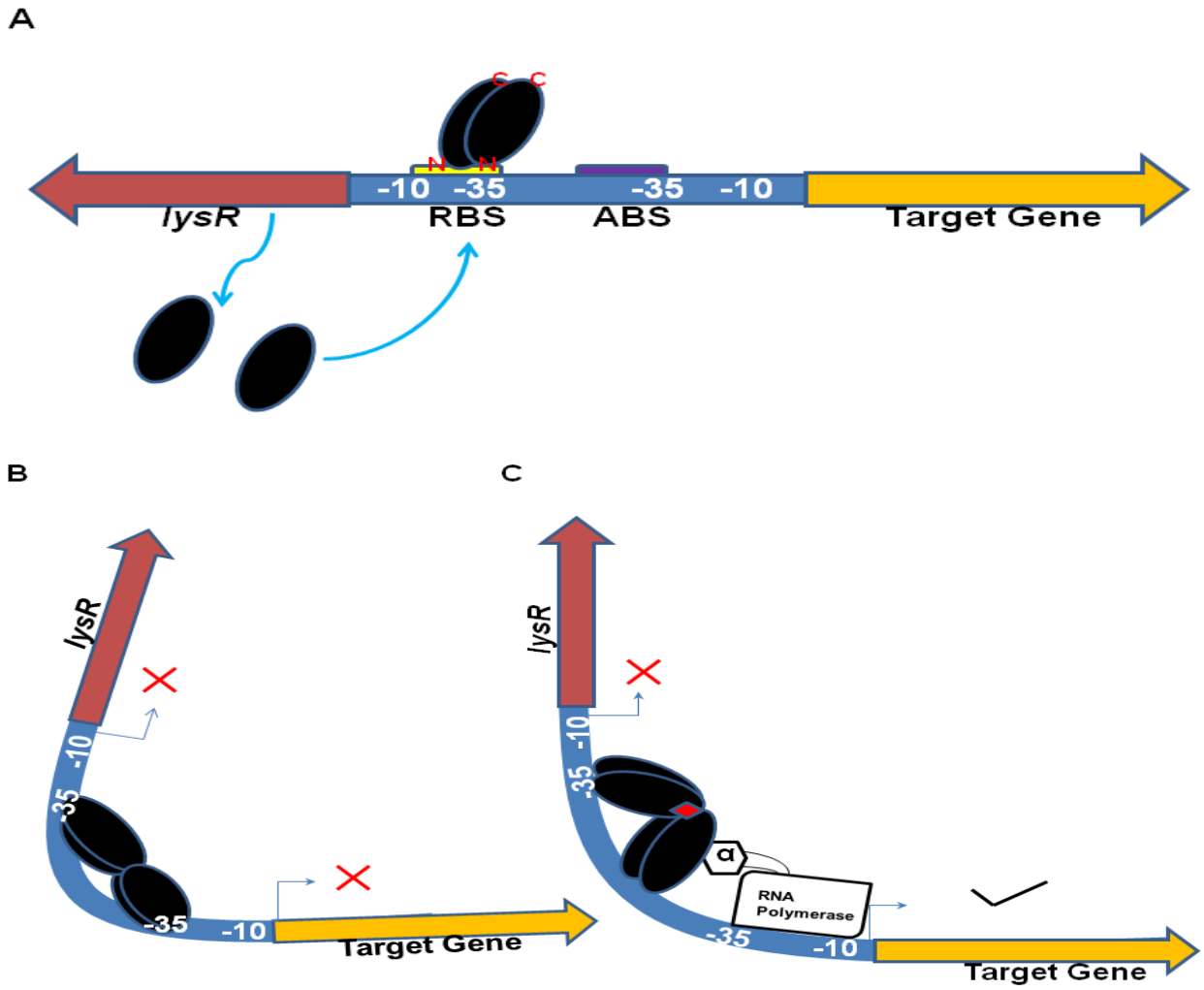


Figure 1-6. Classical mode of gene regulation by LTTRs: A) Classical LTTR members are divergently transcribed from the target genes. A) LTTR dimer first interacts with RBS through an N-terminal DNA binding domain (denoted as N in red). B) LTTR binding at RBS favors tetramerization through the regulatory domain (denoted as C in red).The second dimer in the tetramer interacts with the ABS through the DNA binding domains. This interaction bends the DNA. The LTTR binding at RBS and ABS turns off transcription from the respective promoters. C) When a ligand/ effector molecule is present in the cellular environment, it interacts with the LTTR causing LTTR structural rearrangements. These rearrangements bring the two dimers together (compact form) by relaxing the DNA bend. The structural rearrangements also shift the LTTR binding site at ABS (allows RNAP access to the target gene promoter), and thus transcription from the promoter is enhanced.

CHAPTER 2
MATERIALS AND METHODS

Materials

The oligonucleotides used in this study were from Sigma Aldrich (St. Louis, MO). Biodyne B precut modified nylon membranes and Lightshift chemiluminescent detection kit for Electrophoretic Mobility Shift Assay studies were from Pierce Biotechnology (Rockford, IL). Restriction endonucleases, T4-DNA ligase, DNA size standards, Taq DNA polymerase, Longamp Taq DNA polymerase were from New England Biolabs (Ipswich, MA). Flavonoids, poly (dI-dC), oligonucleotides, chlorophenol red-β-D-galactopyranoside reagent, organic and inorganic analytical grade chemicals were from Sigma-Aldrich (St. Louis, MO). 40 % acrylamide/bis-acrylamide solutions, Bradford protein assay reagents, the empty criterion cassettes were from Bio-Rad (Hercules, CA). The primers used for DNase I footprint were from Applied biosystems (Foster City, CA). The RNA extraction kit was from Ambion (Austin, TX) whereas qRT-PCR reaction reagents and SYPRO orange dye were from Invitrogen (San Diego, CA). Plasmid purification kit (QIAquick), plasmid extraction kit (QIAprep Spin Miniprep), and DNeasy blood and tissue kit were purchased from QIAGEN (Valencia, CA). Pfu turbo polymerase used for site directed mutagenesis was from Stratagene (La Jolla, CA). Ni-NTA superflow resin was purchased from QIAGEN (Valencia, CA). The superose 12 10/300 GL column used for FPLC was from GE healthcare (Uppsala, Sweden) whereas the standards were purchased from Sigma-aldrich (St. Louis, MO). The SDS-PAGE molecular weight standards, Kodak X-ray films were from Fisher (Waltham, MA).

Bacterial Strains and Plasmids

Lactobacillus brevis ATCC367 was obtained from the American Type Culture Collection (ATCC, Manassas, VA). *Bacillus subtilis* M168 was obtained from the *Bacillus* Genetic Stock Center (BGSC, Columbus, OH). *Lactobacillus* strains were grown at 30°C in MRS broth (Difco Laboratories, Detroit, MI). *Escherichia coli* DH5α cells, used to carry and propagate all vectors, were grown in Luria-Bertani medium (Difco) at 37°C, under aerobic conditions.

B.subtilis M168, used for *lacZ* reporter assays, were also grown in Luria-Bertani medium (Difco) at 37°C, under aerobic conditions. When appropriate, media were supplemented with erythromycin, 1 µg/ml, chloramphenicol, 2.5 µg/ml (for *B. subtilis* M168), or ampicillin, 100 µg/ml (for *E. coli*). All antibiotics and chemicals were purchased from Sigma-Aldrich (St. Louis, MO). All strains and plasmids used for this study are shown in Tables 2-1 and 2-2.

Competent Cells Preparation

The *E. coli* competent cells were prepared with $CaCl_2$ and $MgCl_2$ (Sambrook *et al.*, 1989). Briefly, the respective *E. coli* strain was streaked on an LB plate supplemented with 10 mM $MgCl_2$. The plate was incubated at 37 °C overnight. One colony was chosen from the plate for inoculation in 5 ml TYM broth (2 % trypteine, 0.5 % yeast extract, 0.58 % NaCl, and 0.2 % $MgCl_2$). The culture was grown for 2 hr at 37 °C at 250 rpm. The grown culture was transferred to 500 ml of TYM which was incubated at 37 ^{0}C until the OD 600~0.5. The cells were harvested by centrifugation at 3000 rpm for 12 min at 4°C. The pellet was resuspended in 40 ml (for 100 ml of original culture) of TfbI (30 mM K Acetate; 50 mM $MnCl_2$, 100 mM KCl, 10 mM $CaCl_2$, 15 % glycerol). The resuspended cells were incubated on ice for 5-15 min for Rec^+ strains or 60-90 min for Rec^- strain.

The cells were harvested by centrifugation at 3000 rpm for 8 min at 4°C. The pellet was resuspended in 4 ml (for 100 ml of original culture) of TfbII (10 mM MOPS, pH 7; 75 mM $CaCl_2$; 10 mM KCl; 15 % glycerol). The competent cells were aliquoted in 0.2 ml aliquots and stored at -80°C until use.

For *B. subtilis* M168 competent cells preparation, *B. subtilis* M168 was streaked on LB plate. The plate was incubated at 30°C overnight. One colony was inoculated in 2 ml SpC (1X T-base: 1mM $MgSO_4$, 0.5 % glucose, 0.2 % yeast extract, and 0.03 % casamino acids). The culture was incubated at 37 ^{0}C with aeration (250 rpm). After 5 hrs of incubation the culture was transferred to 18 ml of prewarmed SpII (1X T-base, 3.5 mM $MgSO_4$, 0.5 % glucose, 0.1 % yeast extract, and 0.01 % casaminoacids) in 250 ml flask. The culture was incubated at 37 ^{0}C for 90 min. The culture was harvested by centrifugation at 8000 rpm for 3 min at room temperature. The pellet was resuspended in prewarmed 1.6 ml of SpII and 0.4 ml 50 % glycerol. The competent cells were separated in 0.5 ml aliquots and stored at -80 ^{0}C until use.

DNA Manipulations

DNA Amplification

Polymerase chain reaction was used to amplify the gene under study. Quickload Taq 2X Master mix (New England Biolabs, Ipswich, MA) was used for PCR reactions. For amplification of genes larger than 2 kb, longamp Taq 2X mastermix (New England Biolabs, Ipswich, MA) was used. The PCR reaction was carried out as per manufacturer's recommendations. The PCR products were cleaned by QIAquick PCR purification kit (Qiagen, Valencia, CA) as per manufacturer's directions.

Cloning Techniques

The primers were designed to amplify the gene of interest. Extra nucleotides were included at the 5' end of the primers which also included restriction endonuclease sites (Table 2-5). The PCR amplified products were digested with restriction enzymes if required using New England Biolabs (Ipswich, MA) recommended buffers. Similarly, the vector was digested using the same restriction enzymes used for insert digestion. The digested PCR product and vector were purified using QIAquick PCR purification kit (Qiagen). T4 DNA ligase New England Biolabs (Ipswich, MA) was used for liagation. The ligation reaction was carried out as per manufacturer recommended protocol.

For cloning in p15TV-L, the *LVIS0344*, *LVIS0398*, *LVIS0806, LVIS0910*, *LVIS1989*, *LVIS2088* and *LVIS2204* genes were amplified from *L. brevis* ATCC 367 chromosomal DNA using PCR. The primers are listed in Table 2-4. Ligation independent cloning (LIC) infusion reaction was used. Briefly, BD infusion pellet was dissolved with 8.5 µl of predigested p15TV-L plasmid then 2µl infusion/vector mix was mixed with 0.5 µl of insert (PCR target). The infusion: vector: insert mix was incubated at 28 ^{0}C for 30 min and used for transformation.

Transformation

Transformation of bacterial cells with DNA was carried out using the heat shock method (Sambrook *et al*., 1989). The *E.coli* DH5α or *E. coli* XL blue competent cells were used for cloning experiments whereas *E. coli* BL21 competent cells were used for protein expression. An aliquot of the competent cells (80 µl) were mixed with 20 ng to 50 ng reaction sample or target recombinant plasmid and incubated on ice for 10 minutes. The cells were transferred to 37 ^{0}C for 5 min. After incubation the cells were immediately transferred to ice for 2 min. The cells were resuspended in 1ml LB and

incubated at 37 ^{0}C. After 45 min the cells were plated on LB medium supplemented with suitable antibiotic selection marker.

For transformation of *B. subtilis* M168, the *B. subtilis* M168 competent cells were thawed at 37 ^{0}C. An equal volume of SpII-EGTA (0.5 % glucose, 0.09 % $MgSO_4$, 0.1 % yeast extract, 0.01 % casamino acids, 0.002 % EGTA) was added to the competent cells and 20 ng to 50 ng of plasmid DNA. The mix was incubated at 37 ^{0}C for 30 min under aerobic conditions (250 RPM). After 30 min the culture was plated on LB plates using the required selective antibiotic.

DNA Electrophoresis

Chromosomal DNA or DNA fragments were analysed on 1 % agarose (for fragments larger than 1Kb) or. 1.5 % agarose (for fragments smaller than 1Kb). Electrophoresis was performed at 400 mA, 90 V in 1X TAE (40 mM Tris Acetate; 2 mM EDTA, pH 8.5) for 30- 40 min. Before loading on the gel, 5 µl DNA sample was mixed with 1µl of 6X loading dye (New England Biolabs, Ipswich, MA). The 1 kb or 100 bp molecular weight standards (NEB) were used to determine the size of the separated fragments. After electrophoresis, the gel was incubated in ethidium bromide for 10 min. The gel was visualized under UV light by imageQuant 400 (GE Healthcare, Piscataway, NJ) analyzer.

Site Directed Mutagenesis

Site directed mutagenesis was performed using Pfu Turbo DNA polymerase (Stratagene, La Jolla, CA) as per manufacturer's recommendations. Briefly, the reaction mixture used for site directed mutagenesis comprised of 1X reaction buffer (Stratagene, La Jolla, CA). 0.5 µM of each of the primers (sense and antisense), 50 to 100 ng of the template DNA, 2 mM of dNTP mix, 1.5 U of Pfu turbo DNA polymerase, 3 % DMSO.

The PCR conditions were 1-Cycle at 95 ^{0}C for 30 sec, 18 cycles at 95 ^{0}C for 30 sec, 55°C- for 1 min, 68 ^{0}C for 7 min, 1 cycle at 68 ^{0}C for 10 min. The primers used for site directed mutagenesis are shown in Table 2-7. The amplified DNA was subjected to DpnI treatment (20 U for 150 min) for cleavage of hemimethylated template DNA. After DpnI treatment the DNA was transformed into *E. coli* by heat shock transformation method as described previously.

Construction of *lacZ* Fusions

The plasmid pDG1663 (Guérout-Fleury *et al.*, 1996) was used for transcriptional analyses of *LVIS1988* and *LVIS1989* expression. Plasmid pSP01 was constructed by cloning a fragment (amplified with primers P3_Fw and P4_Rv, Table 2-5) containing the complete *LVIS1989* (*kaeR*) sequence, the intergenic region between *LVIS1988-LVIS1989,* and the first 270 nucleotides of *LVIS1988*. Similarly, plasmid pSP02 was constructed using primers P5_Fw and P4_Rv except that only site I was included (Table 2-5). The PCR fragments were cut with *Hin*dIII and *Bam*HI restriction enzymes, and ligated to plasmid pDG1663 digested with the same enzyme. After ligation, samples were transformed into *E. coli* DH5α. Recombinant clones were identified by colony PCR using P9_Fw and 1989_RT_Rv primers. The clones were confirmed by sequencing using the primer LVIS1989_RT_Rv (Table 2-6). Ectopic integration of pSP01 and pSP02 in the *thrC* locus of *B. subtilis* M168 was performed by natural transformation (Petit-Glatron and Chambert, 1992). The resulting strains were labeled SP01 and SP02 (Table 2-1). The ectopic integration of pSP01 and pSP02 was confirmed by PCR using primers P15_Rv and P16_Fw.

Plasmid pLG103 (Belitsky and Sonenshein, 2002) was used for analyses of *LVIS1988* and *kaeR* expression. Transcriptional fusions to *LVIS1988* (to obtain plasmid

pSP05) were constructed by cloning a PCR fragment (using primers P11_Fw and P12_Rv) in the *Bgl*II and *Eco*RI sites. The clones were confirmed by sequencing using primer P9_Fw (Table 2-5). In the plasmids pSP04 and pSP03 the *lacZ* gene was transcriptionally fused to *kaeR*, by cloning fragments that contained either binding site I (using primers P6_Rv and P7_Fw) or both sites (P6_Rv and P8_Fw), respectively. The clones were confirmed by sequencing using primers P9_Fw and P10_Rv (Table 2-5). Strains SP03, SP04, SP05, SP06, and SP07 were constructed by transforming *B. subtilis* with the indicated plasmids followed by ectopic integration in the *amyE* locus.

For heterologous gene expression of *L. brevis* genes in *B. subtilis*, the *LVIS1989* (*kaeR*) gene was amplified (using primers P13_Rv and P14_Fw, Table 2-5) and cloned in the pAX01 plasmid (Härtl *et al.*, 2001) under a xylose inducible promoter. The resulting pSP06 recombinant plasmid was introduced in *B. subtilis* and integrated in the *lacA* locus.

RNA Extraction

Bacterial cells were cultured in MRS broth, in the presence or absence of myricetin, kaempferol or hesperitin (10 µM) when required. The cells were harvested by centrifugation at 4°C, when OD_{600} =0.3 or OD_{600} =1.5 was observed. Total RNA was subsequently isolated using a RiboPure^{TM}-Bacteria kit (Ambion, Austin, TX) in accordance with the manufacturer's protocol. The RNA was quantified by Nanodrop ND-1000 spectrophotometer at 260 nm.

cDNA Synthesis

cDNAs were synthesized with the superscript III first-strand synthesis supermix for qRT-PCR (Invitrogen, San Diego, CA) in accordance with the manufacturer's instructions and stored at -80°C prior to use. Real-time quantitative PCR (qRT-PCR)

was carried out in a Bio-Rad iCycler IQ apparatus, using SYBR® Green ER qPCR SuperMix for iCycler (Invitrogen, SanDiego, CA) in accordance with the manufacturer's recommended protocol. The qRT-PCR cycles conditions were- 1 cycle at 50 ^{0}C for 2 min, 1-cycle at 95 ^{0}C for 10 min, 45 cycles at 95 ^{0}C for 15 sec and 1 min at 60 ^{0}C. The data analysis was carried out by using Bio-Rad iQ5 software.

The primers used in qRT-PCR for *LVIS1986, LVIS1987, LVIS1988 LVIS1989* and for amplification of intergenic regions, are listed in Table 2-6. The *L. brevis* ATCC 367 *rpoD* gene was used as an internal control.

Protein Techniques

Protein Overexpression in *E. coli*

Protein overexpression was carried out as described previously.(Pagliai *et al.*, 2010) Briefly, 2 L of LB supplemented with ampicillin (100 µg/ml) was inoculated with 10 ml overnight grown *E. coli* BL21-Star(DE3) cells (Stratagene, USA) containing the recombinant plasmid. The cells were grown at 37 ^{0}C under aerobic condition (250 RPM) until OD_{600} ~0.6 to 0.8. IPTG (1 mM) was added to the culture which was further incubated at 16 ^{0}C overnight. The cells were harvested by centrifugation at 7500 RPM for 20 min at 4 ^{0}C. The cell pellets were stored at -80 ^{0}C until further use.

Protein Purification

Protein purification was carried out as described previously (Pagliai *et al.*, 2010). The cell pellet was resuspended in binding buffer (500 mM NaCl, 5% glycerol, 50 mM HEPES, pH 7.5, 5 mM imidazole, 0.5 mM TCEP) with 0.5 % NP-40 and 1 mM of PMSF. A French Press was used to lyse the suspension. The lysate was clarified by centrifugation (4 °C, 30 min at 17,000 RPM) and the cell free extract was applied to a metal chelate affinity-column charged with Ni^{2+}. The column was then washed with 250

ml wash buffer (500 mM NaCl, 5 % glycerol, 50 mM HEPES, pH 7.5, 30 mM imidazole, 0.5 mM TCEP). The protein was eluted in elution buffer (binding buffer with 250 mM imidazole). The purified proteins were dialyzed against 10 mM HEPES, pH 7.5, 500 mM NaCl, 2.5% glycerol and 0.5 mM TCEP overnight at 4 ^{0}C. The dialyzed protein was separated in 30 µl aliquots and stored at -80 ^{0}C until use.

Protein Quantification

Protein concentrations were determined using Bio-Rad protein assay kit as per manufacturer's recommendations. This protein quantification is based upon Bradford method (Bradford, 1976). The protein samples were added to 1 ml acidic Bradford reagent (Bio-Rad, Hercules, CA). After 5 min incubation at room temperature, OD at 595 nm was measured by spectrophotometer (UV-1700 pharmaspec, Shimadzu). Bovine gamma-globulin was used as a standard. The relative protein concentration was determined by comparison with a standard curve.

Protein Separation by SDS-PAGE

The cell lysates and Ni-affinity chromatography fractions were assessed for purity by SDS-PAGE. EZ-run SDS-PAGE protein marker was used to determine the polypeptide molecular weights (Fisher, Waltham, MA). 15 µl of protein samples were mixed with 3 µl of 5X loading dye (10 % w/v SDS; 10 mM β-mercaptoethanol; 20 % v/v glycerol; 0.2 M Tris-HCl, pH 6.8; 0.05 % w/v bromophenol blue). The protein samples were denatured by heating at 90 ^{0}C for 5 min. After 5 min the samples and the molecular weight marker were separated on 12 % SDS polyacrylamide gel by electrophoresis (150 V, 40 to 60 min) in 1X running buffer (25 mM Tris HCl, 200 mM glycine, 0.1 % w/v SDS). Gels were stained with coomassie blue and analyzed by imageQuant 400 (GE) analyzer.

Size-Exclusion Chromatography

Protein samples were prepared using 10 mM HEPES, pH 7.5, 500 mM NaCl, 25 μM KaeR, and where indicated, 10 μM kaempferol or myricetin were added. After 20 minutes of incubation on ice, samples were injected onto a prepacked Superose 12 10/300 GL gel filtration column (GE Healthcare, Sweden) connected to an LCC-501 plus (Pharmacia Biotech Inc., Piscataway,NJ) equilibrated with 10 mM HEPES, pH 7.5, 500 mM NaCl. Filtration was carried out at 4°C, using a flow rate of 0.5 ml/min. Eluted proteins were monitored continuously for absorbance at 280 nm using a UV-M II monitor (Pharmacia Biotech Inc., Piscataway, NJ). Blue dextran 2000 was used to determine the void volume of the column. A mixture of protein molecular weight standards, containing IgG (150 kDa), bovine serum albumin (66 kDa), albumin (45 kDa), trypsinogen (24 kDa), cytochrome C (12.4 kDa), and vitamin B12 (1.36 kDa) was also applied to the column under similar conditions. The elution volume and molecular mass of each protein standard was then used to generate a standard curve from which the molecular weight of eluted proteins was determined.

Small Molecule Screening by Differential Scanning Fluorometry

Purified proteins were subjected to screening against apigenin, naringenin, naringin, quercetin, epicatechin, acacetin, luteolin, chrysin, hesperitin, kaempferol and myricetin at a final concentration of 60 μM as previously described (Vedadi *et al.*, 2006; Niesen *et al.*, 2007). Proteins were diluted to a final concentration of 10 μM, using 100 mM Hepes, pH 7 and 150 mM NaCl. 25 μl aliquots of a protein solution were then prepared with each chemical compound, and placed into duplicate 96 well plates (Bio-Rad, Hercules, CA). Samples were then heated from 25 ºC to 80 ºC at a rate of 1 ºC per minute. A real-time PCR device (iCycler IQ, Biorad) was used to monitor protein

unfolding, by measuring the increase in fluorescence of the fluorophore SYPRO Orange (Invitrogen, San Diego, CA). Fluorescence intensities were plotted against temperature for each sample well. Transition curves were fit using the Boltzmann equation using the Origin 8 software (Northhampton, MA). The midpoint of each transition was calculated and compared to the midpoint of the reference sample. If the difference between them was greater than 2.0 ºC, the corresponding compound was considered to be a "hit" and the experiment was repeated to confirm the effect in a dose dependent manner.

Electrophoretic Mobility Shift Assays (EMSAs)

EMSA analysis of KaeR or mutant KaeR was performed using proteins purified and concentrated according to the procedures described above. Different fragments containing the putative binding sites for KaeR (summarized in chapter 3 Figure 3-4A) were generated by PCR using biotin labeled (5'-end) primers (Table 2-3), and subsequently purified using QIAquick spin columns (Qiagen, Valencia, CA). Different probes contained sequences in the surroundings of the intergenic region between *LVIS1988* and *LVIS1989* (Primers E8_Fw to E9_Rv) as well as a putative promoter downstream *LVIS1989* ($P_{LVIS1990}$, generated with primers E10_Fw and E11_Rv; $P_{LVIS1991}$ generated with primers E12_Fw and E13_Rv Table 2-3). The binding buffer for EMSA assays was optimized by the addition of 2.5 mM $MgCl_2$ and 0.5 mM $CaCl_2$. These were found to improve KaeR binding to $F_{kae2,}$ suggesting Ca^{2+} and Mg^{2+} ions play a role during the formation of a strong complex (see Appendix A). The interaction was also found to be very sensitive to pH in both the binding and running buffer. To maintain the stability of the complex, binding buffer was prepared at pH 6.7, while the running buffer was kept at pH 7.5 (see Appendix A). The optimized reaction mix for EMSA (20 μl)

contained 60 pg of a 5'-labelled DNA fragment, 10 mM Tris-HCl, pH 6.8, 150 mM KCl, 0.5 mM EDTA, 0.2 mM DTT, 0.5 mM $CaCl_2$, 2.5 mM $MgCl_2$, 0.1% Triton X100, 25 ng/μl Poly(dI-dC) nonspecific competitor DNA, purified KaeR protein (0-500 nM) and ligand (0-20 μM) unless otherwise specified. After incubation for 30 min at 37°C, samples were separated on 5 % acrylamide/bis-acrylamide non denaturing gels, in 0.5X Tris borate-EDTA buffer, pH 7.5 (TBE, pH 7.5). Electrophoresis was performed at 4 °C in 0.5x TBE as a running buffer. The DNA was then transferred from the polyacrylamide gel to a Biodyne B Nylon Membrane (Pierce Biotechnology, Rockford, IL) by electroblotting at 380 mAmps in 0.5x TBE. Transferred DNA was cross-linked by UV and biotin labeled DNA was detected using a horseradish peroxidase/Super Signal Detection System (Pierce Biotechnology, Rockford, IL). Membranes were exposed to Kodak X-ray film (Fisher, Waltham, MA).

DNase I Footprinting

DNase I footprint assay was carried out at the Plant and Microbe Genomics facility, Ohio State University, Columbus, as described previously by Zianni *et al.*, (2006). Briefly, 5'-VIC or 5'-FAM labeled probes were generated by PCR, using primers E3_Rv and D1_Fw (Table 2-3). The reaction containing 2.5 ng/ul labeled probe, 8 μg of KaeR, 10 mM Tris-HCl, pH 6.8, 150 mM KCl, 0.5 mM EDTA, 0.2 mM DTT, 0.5 mM $CaCl_2$, 2.5 mM $MgCl_2$, 0.1 % Triton X100, 25 ng/μl Poly (dI-dC) nonspecific competitor DNA, and 0.006 U of DNase I (New England Biolabs, Ipswich, MA) was incubated for 20 min at 37 ºC. A digestion reaction without KaeR was included as a control. The reaction was terminated by the addition of 10 mM EDTA, pH 8.0. The digested DNA and sequencing reaction products were analyzed with a 3730 DNA analyzer, and the

protected regions were identified with GeneMarker (Soft genetics) as described earlier (Zianni *et al*., 2006).

β-galactosidase Assays

Cells were grown at 37°C under aerobic conditions (250 RPM) in LB broth in the presence and absence of 160 μM kaempferol or myricetin. Cell samples were taken at various time points during growth phase. Cells were washed twice with 0.9 % NaCl and permeabilized with 1 % toluene in Z-buffer (60 mM Na_2HPO_4, 40 mM NaH_2PO_4, 10 mM KCl, 1 mM $MgSO_4$, 50 mM β-mercapthoethanol) (Miller, 1972). β-galactosidase activity was assayed by following the catalytic hydrolysis of the chlorophenol red-β-D-galactopyranoside (CRPG) substrate (Sigma Aldrich, USA). Absorbance at 570 nm was read continuously using a Synergy HT 96-well plate reader (Bio-Tek Instruments Inc., Winooski, VT) and normalized to initial cell density. β-galactosidase activity is expressed in Miller units. Assays were performed in duplicates at least three times.

The Tools Used for Bioinformatics Studies

The tools used for bioinformatics studies are listed in Table 2-8.

Table 2-1. Strains used in this study.

Name	Relevant genotype/phenotype[a]	Source
E. coli DH5α	φ80 d*lacZ*ΔM15Δ(*lacZYA-argF*)U169 *recA1 endA1 hsdR17* (rk^-. mk^+) *supE44 thi-1 gyrA relA1*	Laboratory stock
E. coli BL21 (DE3)	*E. coli B F⁻ dcm ompT hsdS(rB⁻ mB⁻) gal*	Stratagene
B. subtilis M168	*trpC2*	BGSC[1]
L. brevis ATCC 367	Wild Type	ATCC[2]
B. subtilis SP01	M168 Δ*thrC*: [*LVIS1988* (-387 to +973)-*lacZ*] Em^r.	This work
B. subtilis SP02	M168 Δ*thrC*: [*LVIS1988* (-135 to +973)-*lacZ*] Em^r.	This work
B. subtilis SP03	M168 Δ*amyE*: [LVIS1989 (-400 to +14)-*lacZ*] Cm^r, Δ*lacA*: empty pAX01. Em^r.	This work
B. subtilis SP04	M168 Δ*amyE*: [LVIS1989 (-400 to +14)-*lacZ*] Cm^r, Δ*lacA*: [XylR-P_{XylA}-*LVIS1989* (+1 to +973)] Em^r.	This work
B. subtilis SP05	M168 Δ*amyE*: [LVIS1989 (-182 to +14)-*lacZ*] Cm^r, Δ*lacA*: [XylR-P_{XylA}-*LVIS1989* (+1 to +973)] Em^r.	This work
B. subtilis SP06	M168 Δ*amyE*: [LVIS1988 (-399 to +16)-*lacZ*] Cm^r, Δ*lacA*: [XylR-P_{XylA}-*LVIS1989* (+1 to +973)] Em^r.	This work
B. subtilis SP07	M168 Δ*amyE*: [LVIS1988 (-399 to +16)-*lacZ*] Cm^r, Δ*lacA*: empty pAX01. Em^r.	This work
B. subtilis SP15	M168 Δ*thrC,* empty pDG1663. Em^r	This work
B. subtilis SP16	M168 Δ*amyE*: empty pLG103 Cm^r, Δ*lacA*: empty pAX01 Em^r.	This work

[a]The positions indicated are relative to *LVIS1989* translation start codon. [1]*Bacillus* Genetic Stock Center, [2]American Type Culture Collection

Table 2-2. Plasmids used in this study.

Name	Relevant genotype/Phenotype[a]	Source
p15TV-L	Expression vector for purification of proteins by nickel affinity cromatography. Ap^r.	Addgene plasmid 26093
pAX01	*B. subtilis* vector for ectopic integration into *lacA* site containing *XylR-*P_{XylA} cassette. Ap^r,Em^r.	Härtl *et al*., 2001
pLG103	*B. subtilis* vector for integration into *amyE* site containing two promoterless reporters, *gusA* and *lacZ* in opposite orientaions. Ap^r, Cm^r.	Belitsky and sonenshein, 2002
pDG1663	*B. subtilis* vector for ectopic integration into *thrC* site containing *E. coli spoVG-lacZ*. Ap^r, Em^r.	Guérout-Fleury *et al*., 1996
pSP01	*LVIS1988-lacZ* transcriptional fusion in pDG1663 carrying the *L. brevis* sequence from -387 to +973. Ap^r, Em^r.	This work
pSP02	*LVIS1988-lacZ* transcriptional fusion in pDG1663 carrying the *L. brevis* sequence from -135 to +973. Ap^r , Em^r.	This work
pSP03	*LVIS1989-lacZ* transcriptional fusion in pLG103 carrying the *L. brevis* sequence from -400 to +14. Ap^r; Cm^r.	This work
pSP04	*LVIS1989-lacZ* transcriptional fusion in pLG103 carrying the *L. brevis* sequence from -182 to +14. Ap^r, Cm^r.	This work
pSP05	*LVIS1988-lacZ* transcriptional fusion in pLG103 carrying the *L. brevis* sequence from -399 to +16. Ap^r, Cm^r.	This work
pSP06	*L. brevis LVIS1989* from +1 to +973 cloned in pAX01. Ap^r , Em^r.	This work
pLB01	p15TV-L His6-LVIS0344. Ap^r.	This work
pLB02	p15TV-L His6-LVIS0398. Ap^r.	This work
pLB03	p15TV-L His6-LVIS0806. Ap^r.	This work
pLB05	p15TV-L His6-LVIS0910. Ap^r.	This work
pLB06	p15TV-L His6-LVIS1989. Ap^r.	This work
pLB07	p15TV-L His6-LVIS2088. Ap^r.	This work
pLB08	p15TV-L His6-LVIS2204. Ap^r.	This work

[a]The positions indicated are relative to *LVIS1989* translation start codon.

Table 2-3. Primers used for electrophoretic mobility shift assays (EMSAs) and DNase I footprint assays.

Oligonucleotide Name	Oligonucleotide Sequence (5'-3')	Pos[b]	Purpose	Mod[c](5')
E1_Rv	TGTTGCATGTTCCGTAATCG	+122*	EMSA	
E2_Fw	TTGGCATGGGTTTTTAGCTC	-176 *	EMSA	Biotin
E3_Rv	TGTGATGAAACTGTAGAAATCGTC	+33*	DNase I footprint	VIC
E7_Rv	TGTGATGAAACTGTAGAAATCGTC	+33*	EMSA	
E4_Fw	TTCCCAGGGTTTGGTAATCA	-400*	EMSA	Biotin
E6_Rv	CCCATGCCAACCAGTACCATGAGACGAA	-167*	EMSA	
D1_Fw	GGGTAGGTGCGTCAGTAACC	-463*	DNase I footprint	FAM
E8_Fw	GCCTTCAACGTTATTAAACATCATTG	-312*	EMSA	Biotin
E9_Rv	TTAGTCGAAAGTTGAGTTGTTTGCA	-40*	EMSA	
E10_Fw	TCAGTTGCCCCCTTTATGAC	+745*	EMSA	Biotin
E11_Rv	CCATTCCCACACAATTTTCC	+1106*	EMSA	
E12_Fw	GCGTGCTACATCTGAGCGTA	+1398*	EMSA	Biotin
E13_Rv	CCAGAGCCTTTTGAAACCAA	+1694*	EMSA	

[b]Position, [c]Modification. *The positions indicated are relative to *LVIS1989* translation start codon.

Table 2-4. Primers used to clone the genes for protein expression.

Oligonucleotide Name	Oligonucleotide Sequence (5'-3')[Y]	Position
LVIS0344_Fw	*ttgtatttccagggcat*GTGAATTTTCGACAGTTAGAATATTTC	+1[#]
LVIS0344_Rv	*caagcttcgtcatcatc*AGTAACTGTCAATGAAGCGCTG	+856[#]
LVIS0398_Fw	*ttgtatttccagggc*ATGCTAGATAAACGATACGAAAC	+1[#]
LVIS0398_Rv	*caagcttcgtcatcatc*AGTTACCAGGTTGTCGTAAAG	+886[#]
LVIS0806_Fw	*ttgtatttccagggc*ATGAAAACGAAACAGGAAAGTATC	+1[#]
LVIS0806_Rv	*caagcttcgtcatcatc*AATTTAATTGGTTCCGAAGTTGTTTG	+964[#]
LVIS0910_Fw	*ttgtatttccagggc*ATGTCGCAGCGTGCAGTG	+1[#]
LVIS0910_Rv	*caagcttcgtcatcatc*AAAATTGATGAGTGTCAAAATACC	+817[#]
LVIS1989_Fw	*ttgtatttccagggc*ATGACTATCGACGATTTCTAC	+1 *
LVIS1989_Rv	*caagcttcgtcatca*TCAGTCCGTTAGATAATGAATGAAC	+882*
LVIS2088_Fw	*ttgtatttccagggc*ATGCTACCCTTTGCTTATCG	+1[#]
LVIS2088_Rv	*caagcttcgtcatcatc*AGTTATTGCGGGGATGTTCC	+907[#]
LVIS2204_Fw	*ttgtatttccagggc*ATGAATACAAAAGATTTGGATTATTTC	+1[#]
LVIS2204_Rv	*caagcttcgtcatcatc*ATCCTAGCCGCAAACTCTCTC	+886[#]

[#]The positions indicated are relative to respective genes translation start codons. [Y]The extra bases added for cloning in p15TV-L are shown in italics.

Table 2-5. Primers used to clone the genes for *lacZ* reporter assays and sequencing reactions.

Oligonucleotide Name	Oligonucleotide Sequence (5'-3')[α]	Pos[b]	Purpose	Restriction Enzyme
P3_Fw	aaaaaggatccGTAATCATGATGGAACA	-387 *	*LVIS1988:LVIS1989* cloning in pDG1663 (Site I and II)	*BamHI*
P4_Rv	aaaaaaagcttGGGAATCCTCCAATAAAT	+973 *	*LVIS1988:LVIS1989* cloning in pDG1663	*HindIII*
P5_Fw	tttttggatccATAAGGATCATTTACCAT	-135 *	*LVIS1988:LVIS1989* cloning in pDG1663 (Site I)	*BamHI*
P6_Rv	ggccatagatcTCGTCGATAGTCATTGAAGAACCT	+14 *	*LVIS1988: LVIS1989* cloning in pLG103	*BglII*
P7_Fw	ggccatgaattcTACTGGTTGGCATGGGTTTT	-182 *	*LVIS1988: LVIS1989* cloning in pLG103 (Site I)	*EcoRI*
P8_Fw	ggccatgaattcTTCCCAGGGTTTGGTAATCA	-400 *	*LVIS1988:LVIS1989* cloning in pLG103 (Site I and II)	*EcoRI*
P11_Fw	ctctagagatctTCCCAGGGTTTGGTAATCATGATGG	-399 *	*LVIS1988: LVIS1989* cloning in pLG103 (Site I and II)	*BglII*
P12_Rv	ggtccggaattcAATCGTCGATAGTCATTGAAGAACC	+16 *	*LVIS1988: LVIS1989* cloning in pLG103	*EcoRI*
P13_Rv	aaaaaaccgcggGGGAATCCTCCAATAAATG	+973 *	*LVIS1989* cloning in pAX01	*SacII*
P14_Fw	gaaatgggatccATGACTATCGACGATTTCTACAG	+1 *	*LVIS1989* cloning in pAX01	*BamHI*
P15_Rv	CAGGTTATCTGTACCCGCCGGA	+189 [¥]	pDG1663 ectopic integration confirmation	
P16_Fw	TAACACTCAGTCCCGGTTCC	+590 *	pDG1663 ectopic integration confirmation	
P9_Fw	AGCGCCATTCGCCATTCAGGCT	+769 [$]	Sequencing	
P10_Rv	CATAAGGGACTCCTCATTAAG	+3 [Δ]	Sequencing	

[b]Position. The positions indicated are relative to [*]*LVIS1989*, [¥] *thrB* translation start codons, [$]*lacZ* and [Δ]*gusA,* start codons in pLG103.
[α]Extra nucleotides used are shown in small letters and restriction sites used for cloning are underlined.

Table 2-6. Primers used for quantitative real time PCR (qRT-PCR).

Name	Sequence (5'-3')	Position[b]
LVIS1986_RT_Fw	GACACCCCCATTCAAACAAT	+133[#]
LVIS1986_RT_Rv	AATCTGGGCCTTGGTGATCT	+294[#]
LVIS1987_RT_Fw	TTAAAGCGATGGCCGACTAC	+164[#]
LVIS1987_RT_Rv	AGCCCGCGCAATTAGATTAT	+323[#]
LVIS1988_RT_Fw	CGATGTAATGTGGACCGTGA	+1227[#]
LVIS1988_RT_Rv	CACCAACTTCGATGACATGC	+1377[#]
LVIS1989_RT_Fw	GGCGTGGTACGATTCAGATT	+269[#]
LVIS1989_RT_Rv	TGACTTTGAACGTCGCGTAG.	+458[#]
L1986.1987_Rv	AAGGTACTGGGTGCATTTGG	-2104[*]
L1986.1987_Fw	GTTACTGCCACCCGTGACC	-2263*
L1987.1988_Rv	TGAACGCGCGAAGTTTAAG	-1533*
L1987.1988_Fw	GTTTCAACGCCTGGCTTATC	-1693*
L1988.1989_Rv	TGTGATGAAACTGTAGAAATCGTC	+33*
L1988.1989_Fw	GCTCGGCCAATACTTTTCGT	-160*
rpoD_Fw	ATTCCCGTTCATATGGTGGA	+640[#]
rpoD_Rv	GAACCTTTTCCGTTGGCATA	+769[#]

[b]The positions indicated are relative to [*]*LVIS1989*, [#]respective genes translation start codons.

Table 2-7. Primers used for site directed mutagenesis.

Oligonucleotide Name	Oligonucleotide Primers (5'-3')	Position[b]
1989_P100A_Fw	ATTCAGATTGGCAGTGTGGCCGTCATGGCTCAGTACGGT	+280*
1989_P100A_Rv	ACCGTACTGAGCCATGACGGCCACACTGCCAATCTGAAT	+318*
1989_Q153A_Fw	CTACGCGACGTTCAAAGTGCACAACTGAACCACTCACAG	+439*
1989_Q153A_Rv	CTGTGAGTGGTTCAGTTGTGCACTTTGAACGTCGCGTAG	+477*
1989_V205A_Fw	CTCAGTCCCGGTTCCGGTGCTTATGAACGAATCAGCGAG	+595*
1989_V205A_Rv	CTCGCTGATTCGTTCATAAGCACCGGAACCGGGACTGAG	+633*
1989_R148A_Fw	TTTGATCTCGGTATTCTAGCCGACGTTCAAAGTCAACAA	+424*
1989_R148A_Rv	TTGTTGACTTTGAACGTCGGCTAGAATACCGAGATCAAA	+462*
1989_E127A_Fw	ATTAACTTTTCTTTGGCAGCATTAGAGGGGGCTGATCTC	+361*
1989_E127A_Rv	GAGATCAGCCCCCTCTAATGCTGCCAAAGAAAAGTTAAT	+399*
1989_I229A_Fw	CGTTTTTCCACACCGCATGCTGAAACACTGCTGGCAATG	+667*
1989_I229A_Rv	CATTGCCAGCAGTGTTTCAGCATGCGGTGTGGAAAAACG	+705*

[b]The positions indicated are relative to *_LVIS1989_ translation start codon.

Table 2-8. List of bioinformatics tools used.

Bioinformatics tool	Purpose	Website	References
PHYRE	KaeR structure modeling	http://www.sbg.bio.ic.ac.uk/~phyre/	Kelley LA and Sternberg MJE, 2009
SWISS-MODEL	KaeR structure modeling	http://swissmodel.expasy.org/workspace/	Benkert et al., 2011
PDB RCSB	Identification of ligand bound homologous structure	http://www.pdb.org/pdb/home/home.do	Berman et al., 2000
PyMOL	Structural alignment	http://www.pymol.org/	Warren L. Delano, DeLano Scientific LLC, San Carlos, CA, USA
ClustalW	Sequence alignment	http://www.ebi.ac.uk/Tools/msa/clustalw2/	Larkin et al., 2007
Ortholog neighborhood	Comparative genomics	http://img.jgi.doe.gov/	Mavromatis et al., 2009

CHAPTER 3
LVIS1989 (KAER), A LYSR TYPE TRANSCRIPTIONAL ACTIVATOR, UP REGULATES *LVIS1986*, *LVIS1987*, *LVIS1988* AND *LVIS1989* IN RESPONSE TO KAEMPFEROL

Introduction

Flavonoids are plant derived compounds that constitute a significant component of the human diet. They are believed to possess both anti-carcinogenic and cardio protective properties (Selma *et al*., 2009). Studies have revealed, however, that complex flavonoids are not absorbed by the gastrointestinal tract. They are instead broken down by gut microbiota into smaller compounds with higher biological activity (Spencer, 2003; Hein *et al.*, 2008; Verzelloni *et al.*, 2011). Some of the microorganisms capable of cleaving flavonoids to produce aliphatic organic acids include *Eub. oxidoreducens*, commonly found in the bovine rumen, and the human fecal microbes *Eub. ramulus and Clostridium orbiscindens* (Winter *et al*., 1989; Krumholz and Bryant, 1986; Braune *et al*., 2001). The degradation and transformation of these compounds by other beneficial members of the microbiota such as *Lactobacillus* however, have been scarcely explored (Rodríguez *et al.*, 2009).

Lactobacillus species are widely used throughout the food industry, for both their fermentation and probiotic properties (Lebeer *et al.*, 2008; Gobbetti *et al.*, 2010). Many species of *Lactobacillus* have been isolated from flavonoid rich environments. *L. brevis* is frequently found in decaying plants and fruit fermentations (Kandler and weiss, 1986), where exposure to complex flavonoids and their organic acid constituents is inevitable. Consequently, I hypothesized that *L. brevis* may have developed the ability to sense and respond to these types of molecules, as a result of their continued exposure to them. More specifically, we are interested in the flavonoids present in the food intake,

and their role as modifiers of signal pathways in beneficial microbes that may in turn affect their interaction with the host.

Recently, several microbial transcription factors with the ability to mediate transcriptional activation in response to flavonoid exposure have been identified. These transcription factors have been discovered in different families, including TetR (LmrA, YxaF and TtgR; Hirooka *et al.*, 2007; Teran et.al., 2006), MarR (YetL; Hirooka *et al.*, 2009), LysR (NodD; (Rossen *et al.*, 1985) and most recently, LuxR (RhlR; Vandeputte *et al*., 2010). In most cases, they are involved in the regulation of genes due to the antimicrobial activity of flavonoids.

The mechanisms through which LTTRs modulate gene expression are diverse. In *Rhizobium leguminosarum* NodD, a LysR type transcriptional regulator (LTTR) activates the expression of nodulation genes in response to the flavonoid naringenin (Rossen *et al*., 1985). One of the most intriguing aspects of LTTR members, however, is their ability to mediate transcriptional activation through DNA bending and differential interactions with the RNA polymerase. This type of mechanism has been documented in several LTTR members including CrgA, NodD, ArgP, BenM and CatM; however, they are all unique in the outcomes (Zhou *et al*., 2010; Craven *et al*., 2009; Chen *et al.*, 2005; Sainsbury *et al.*, 2009). For example in *R. leguminosarum* (Chen *et al.*, 2005), in the absence of naringenin, NodD, binds to the *nodA* promoter region as a tetramer, promoting the DNA to bend. The addition of naringenin acted to sharpen this bend, inducing the expression of *nodA* and related genes;(Chen *et al*., 2005).

The present study was undertaken to identify and characterize a new LTTR regulator KaeR (LVIS1989) in *L. brevis,* capable of mediating a positive response to the

flavonoid kaempferol. It was determined that *kaeR* (*LVIS1989*) is also positively autoregulated by the presence of the ligand.

Results

The LysR Member, KaeR (LVIS1989) Interacts With Flavonoids *In Vitro*

A fluorescence based, small molecule screening assay was used to identify transcription factors that bind flavonoids in *L. brevis*. This technique is based upon a principle that when a ligand binds to a protein there will be a shift in its melting temperature. The SYPRO orange dye (Invitrogen, San Diego, CA) was used for these studies which changes its emission properties with the extent of protein unfolding. The more stable the protein the more will be the shift in its melting temperature after ligand addition (Figure 3-1). The flavonoids exhibiting at least 2 ^{0}C difference in the melting temperature were considered as a positive hit.

The genome of *L. brevis* was searched for LysR type transcription regulators that are classified within the COG0583 (Cluster of Orthologous Genes). It was determined that *L. brevis* encodes for seven proteins classified within COG0583: LVIS0344, LVIS0398, LVIS0806, LVIS0910, LVIS1989, LVIS2088 and LVIS2204. Figure 3-2 shows the genomic environment of *L. brevis* all LTTRs.

All genes were cloned and the recombinant proteins were purified with the exception of LVIS0344 and LVIS0910, due to their poor solubility under the conditions tested. To establish the midtransition temperature (Tm), purified proteins for LVIS0398, LVIS0806, LVIS1989, LVIS2088 and LVIS2204 were subjected to thermal denaturation. The effect of flavonoids on the thermal stability of each protein was tested at 60 μM. Apigenin, naringenin, quercetin, acacetin, luteolin and chrysin had no effect on any of the proteins tested while, myricetin, kaempferol and hesperitin increased the thermal

stability of LVIS1989 (Table 3-1). Although myricetin, kaempferol and quercetin are flavon-3-ols that share a similar chemical scaffold, confirmation assays revealed that only myricetin and kaempferol bind to LVIS1989 in a dose dependent manner (50 μM - 1 mM). Quercetin was not found to have stabilization effects on LVIS1989 (Table 3-1).

Flavonoids are Involved in the Upregulation of *LVIS1986*, *LVIS1987*, *LVIS1988*, and *LVIS1989* mRNAs

To determine the *in vivo* effect of these chemicals on *L. brevis*, the genomic environment of *LVIS1989* was analyzed to uncover putative genes regulated by this protein. In general, LTTR members are encoded divergently from the genes under their control. LVIS1989 is encoded on the plus strand, and shares high sequence identity with *Oenococcus oeni* PSU-1 (47% identity; GI: 116491191), *L. plantarum* WCFSI (41% identity, GI: 28379389) and *L. rhamnosus* (37% identity, GI: 258509654). None of the genes have been previously characterized.

Upstream of *LVIS1989*, on the minus strand, three genes are encoded in an "operon like" structure. *LVIS1988* and *LVIS1987* have high sequence identity with 3-polyprenyl-4-hydroxybenzoate decarboxylases (*ubiD* and *ubiX*, respectively). The third gene in the predicted operon, *LVIS1986*, encodes a conserved hypothetical protein composed of 136 amino acids. Interestingly, this operon structure is only conserved in *O. oeni* PSU-1 (NC_008528) (Fig. 3-3 A) where in *Pediococcus pentosaceous* ATCC 25745 (NC_008525), *L. gasseri* JV-V03 (NZ_ACG01000006) and *L. sakei* 23K (NC_007576), the genes homologous to *LVIS1986, LVIS1987, LVIS1988* and *LVIS 1989* may compose a single transcriptional unit. In contrast, partial associations are found in *L. plantarum* WCFSI (NC_004567) and *L. rhamnosus* GG (NC_013198). Downstream of *LVIS1989*, in the plus strain, *LVIS1990* encodes for an N-

formylmethionyl-tRNA deformylase (*def 2*). Comparative genomics performed using the Ortholog Neighborhood Viewer at the JGI Integrated Microbial Genomes website (http://img.jgi.doe.gov/), revealed that def2 is not phylogenetically linked to *LVIS1986_1987_1988_1989*. It was then tested if *LVIS1986_1987_1988* comprised a single transcriptional unit. qRT-PCR experiments were performed using primers designed around the intergenic regions of genes *LVIS1988*, *LVIS1987* and *LVIS1986*. (Table 2-6 indicated as L1986.1987 Fw and Rv, L1987.1988 Fw and Rv, L1988.L1989 Fw and Rv). The results suggested that if *LVIS1986_1987_1988* indeed formed a single transcriptional unit.

Based on these results, the expression levels of *LVIS1986, LVIS1987, LVIS1988* and *LVIS1989* were determined in the presence or absence of myricetin, hesperitin and kaempferol. *L. brevis* cells were grown in MRS broth with 10 µM of each flavonoid, then collected in both the early exponential phase (OD_{600}= 0.3) and late stationary phase (OD_{600}= 1.5). The mRNA levels were determined by qRT-PCR (Figure 3-3B and C). In exponential phase, kaempferol increased the expression of all four genes by 2.5-5 fold (Figure 3-3B), while myricetin was only found to induce the expression of *LVIS1989* (by 4 fold). Hesperitin was found to be a weak inducer of *LVIS1986* (2 fold change), but had no effect on the other genes measured (data not shown). Interestingly, *LVIS1986, LVIS1987, and LVIS1988* were induced in stationary phase cells grown in the presence of kaempferol (~2.5 fold) or myricetin (~2 fold) (Fig. 3-3 C). These results indicate that both flavonoids have the potential to induce gene expression, though the uptake or binding capabilities of kaempferol may occur with higher affinity. Based on these results, LVIS1989 was given the name KaeR, for kaempferol responsive protein.

Identification of KaeR Binding Region

Since most of the LTTRs are divergently transcribed from the promoter regulated by them (Schell, 1993, Maddocks and Oyston, 2008) the putative promoter region of KaeR F_{kae1} was tested using EMSA assays. Putative promoter regions found downstream of *kaeR were* also tested ($P_{LVIS1990}$ *and* $P_{LVIS1991}$). Under standard conditions, KaeR interacted weakly with F_{kae1} (Figure 3-4A and 3-6C) but not with $P_{LVIS1990}$, $P_{LVIS1991}$ (data not shown). A variety of primer combinations was utilized to obtain DNA fragments of various lengths to determine the region of binding that would result in stable DNA: KaeR complexes. The best results were obtained using fragment F_{kae2} which comprised the intergenic region between *LVIS1988* and *LVIS1989,* as well as the first 280 nucleotides of *LVIS1988*.

EMSA analysis of F_{kae2} and KaeR, revealed their binding to be concentration dependent; this was made evident by a correlated increase of high molecular weight oligomers, with increasing concentrations of KaeR (Figure 3-4B). Using KaeR at 100 nM, F_{kae2} was observed to form only a small complex with the protein. With 400 nM KaeR however, a larger, slower migrating complex was obtained, suggesting that KaeR oligomerizes cooperatively with DNA. The multiple intermediate complexes observed when using 250 nM KaeR (evidenced as a smear), suggest that *in vitro* the complexes are unstable.

KaeR Interacts With Two Binding Sites Separated by 280 bp

DNase I footprint assays revealed the presence of two binding sites, separated by 280 bp. Binding site I protected the 41 nucleotide sequence from +2 to -39 (from the *kaeR* translation start codon, Figure 3-5A), while binding site II protected a 40 nucleotide sequence, from -314 to - 353 (Figure 3-5B). KaeR binding site II is located

upstream from binding site I, and lies within the encoding sequence for *LVIS1988*. One imperfect inverted repeat cGATTt-N5-tAATCc was identified from -337 to -353 (Figure 3-5C). Previous studies have reported LTTR binding at two distinct binding sites, namely the Activator Binding Site (ABS) and the Repressor Binding Site (RBS). Within the RBS, it has been proposed that a palindromic sequence (T-N11-A) is involved with gene activation or repression (Wilson *et al*., 1995). Binding site I contains an imperfect inverted repeat (TTATgCCtaaatGGaATAA), and overlaps a putative -10 sequence in the promoter region of *kaeR* (Figure 3-5C). Of note is the presence of the conserved cytosine and guanine (CC-N5-GG) which may mediate contact with KaeR, as shown for CatR (Parsek *et al.*, 1992) and TrpI (Chang and Crawford, 1991; Gao and Gussin, 1991).

The Two Protected Regions are Required for the Formation of Higher Order Complexes *In Vitro*

To determine if the two protected sequences identified by DNase footprint are required for KaeR binding, multiple primer combinations were used to generate ~200 bp long probes containing binding site I (-176 to +33, F_{kae1}) or binding site II (-400 to -167, F_{kae3}) (Figure 3-6). Their binding was then compared to a probe containing both protected regions (-400 to +33, F_{kae2}) (Figure 3-6A). A very weak modification in the electrophoretic mobility was found when using site I (F_{kae1}, Figure 3-6C), while no significant increase in complex formation was observed when site II was considered (F_{kae3}) (Figure 3-6D). The use of a longer fragment (-176 to +122, F_{kae5}) containing the complete intergenic region (but only binding site I) showed an increase in the stability of the KaeR-DNA complex (Figure 3-6B). To determine if additional binding sequences were present in that region, a fragment within the intergenic region, ending just outside

site I and site II, was amplified (-312 to -40, F_{kae4}). KaeR binding was not observed in the presence or absence of the ligand (data not shown). These results indicate that binding site I may be the primary binding site, while binding site II may only be required to stabilize the KaeR oligomer on DNA.

KaeR Activity is Modulated by Kaempferol *In Vitro*

The effect of kaempferol on the ability of KaeR to bind DNA was assessed by EMSA. For this experiment, 20 µM kaempferol was added to increasing concentrations of KaeR. Modifications in the mobility of the KaeR:F_{kae2} complex were observed. When kaempferol was present in excess of 500 fold, the KaeR: F_{kae2} complex was able to form a stable intermediate complex (Figure 3-6A). At lower ratios however (400 fold excess), kaempferol did not affect the migration of the complex. Unfortunately, higher concentrations of the kaempferol could not be tested due to its poor solubility in buffer.

The differences observed in the complex migration, may result from variations in the oligomeric state of KaeR, induced by interactions with the ligand. To explore this possibility, the oligomeric state of KaeR was assessed by gel filtration experiments. The majority of KaeR eluted as a monomer, (apparent MW = 40 kD) however the observed sizes were slightly larger than the expected monomeric size of 34 kDa. Additionally, a small percentage of tetramer-like species (apparent MW=160 kDa) was detected, and a small peak with an estimated MW of 5 kDa was also observed (Fig. 3-7A). No extra bands, however, were obtained on SDS-PAGE gels (Figure 3-7B). An increase in tetramer-like species was observed following the addition of kaempferol or myricetin (10 µM) while the “non binding” flavonoids such as hesperitin and naringenin had no effect. The presence of higher oligomeric forms (>300 kDa) was also detected, evidenced by the increased concentrations observed in the void volume. Marginal increment of high

order oligomers in the presence of ligands has previously been reported for AtzR, although the main product was the tetrameric form in all cases (Porrúa *et al*., 2007). Here, the binding of kaempferol induced a major shift from the monomer to the tetramer-like species of KaeR (Figure 3-7).

Taken together, the increased stability of tetramers in the presence of kaempferol, correlates with the variations observed in the F_{kae2}:KaeR complex migration. The compact state of the tetramer may have mechanistic implications, such as the DNA bending mechanism previously shown for other LTTR proteins (Muraoka *et al*., 2003; Zhou *et al*., 2010; Porrúa *et al*., 2007; Chen *et al*., 2005; Sainsbury *et al*., 2009).

Based on these results, additional EMSA analysis was carried out to determine if the binding of KaeR to site I (F_{kae1} and F_{kae5}) or site II (F_{kae3}), could be improved by the addition of the ligand. As shown in Figure 3-6C, at low protein concentration (250 nM) the complex F_{kae1}:KaeR migration is very weak and kaempferol marginally modified the migration. The addition of kaempferol to the F_{kae3}:KaeR or to the F_{kae5}:KaeR complex had no effect on the stability of the complex (Figure 3-6 B and D). These results indicate that kaempferol-induced variations of the F_{kae2}:KaeR complex, depend on the presence of both binding sites.

In Vivo Assessment of KaeR's Binding Sites on *LVIS1988* Expression

To gain functional insight on the KaeR binding sites, transcriptional fusions to the *lacZ* reporter gene were constructed and the regulation studies were conducted using *B. subtilis* as a surrogate host. Of note, *B. subtilis* M168 does not contain genes with high similarity to the genes under study.

To determine if each binding site was necessary for expression of the *LVIS1988* gene, the promoter region of *LVIS1988* was fused to the *lacZ* gene, using the

integrative plasmid pDG1663 (Guérout-Fleury *et al*., 1996). The complete *kaeR* gene sequence was cloned in *cis*, to maintain the same genomic context found in *L. brevis* (Figure 3-8). Strain SP01 contains both KaeR binding sites, while strain SP02 contains only binding site I. A higher basal expression level was observed in SP01 (54 ± 0.9 Miller Units) when compared to SP02 (17 ± 2.7 Miller units). Interestingly, induction was observed in both transcriptional fusions when kaempferol, or the structurally related myricetin, was added to the medium. In the presence of kaempferol, induction was increased by 2.5 fold in strain SP02, while a 3.4 fold increase was observed in strain SP01. These results indicate that although binding site II is not essential for induction of *LVIS1988*, expression levels are enhanced when it is present.

To determine if KaeR is required for expression of *LVIS1988,* β-galactosidase activities of two strains carrying the *LVIS1988-kaeR* intergenic region (containing binding sites I and II) in absence (strain SP07) or presence of *kaeR* (strain SP06), were compared. The *kaeR* gene was cloned in *trans* as a chromosomal insertion at the *lacA* locus and expressed from a xylose inducible promoter (strain SP06). Similar low expression levels were obtained in SP07 (17 ± 5.0 Miller units) and SP06 (13 ± 0.9 Miller units). The addition of kaempferol resulted in a 2 fold increase in β-galactosidase activity in SP06 (26 ± 0.1 Miller units), while no changes were observed in SP07. Although expression of *LVIS1988* was induced in SP06 with *kaeR* positioned in *trans,* the activity values were lower than those obtained when *kaeR* was expressed in *cis* (SP01, 183 ± 1.6 Miller units). These results indicate that both the location and expression levels of KaeR are important for optimal induction of the *LVIS1988* gene.

KaeR is Positively Autoregulated

To determine if KaeR mediates its own regulation, the same fragment used in SP07 (containing KaeR binding site II and the intergenic region between *LVIS1988* and *LVIS1989*) was cloned in the opposite orientation in pLG103 (Belitsky and Sonenshein, 2002) and integrated into the *amyE* locus (strain SP03). The resulting basal expression level was higher (48 ± 2.2 Miller units) than that observed for *LVIS1988* (SP07, 17 ± 5.0 Miller units) (Figure 3-8). The expression of the *kaeR* promoter was then assessed when *kaeR* was cloned *in trans* (as a chromosomal insertion at the *lacA* locus, strain SP04). In the SP04 strain, basal level expression of *kaeR* did not modify expression from the P_{kaeR} promoter (44 ± 3 Miller units). Furthermore, expression levels were not significantly affected by the addition of 0.5 % xylose, and similar β-galactosidase activity was observed when only binding site I was present (strain SP05, Figure 3-8).

Expression levels of the *kaeR* promoter were assessed in both the presence and absence of kaempferol or myricetin using strains SP04 and SP05. In both strains, the expression of the reporter gene was increased from 1.9 to 2.9 fold, following the addition of kaempferol, although the higher levels were seen only in SP05, where only binding site I was present. A similar ratio was obtained using myricetin as the inducer (Figure 3-8). These results indicate that activation of *kaeR* transcription is dependent on the presence of KaeR and the inducer molecule.

Discussion and Conclusion

In terrestrial environments, plant derived flavonoids serve as important signaling molecules for microbial gene expression. Upon interaction of these compounds with specific microorganisms such as *Rhizobium,* plant nodulation is induced (Brencic and Winans, 2005). However, their effects on the regulatory networks of other

microorganisms commonly isolated from environments that are rich in these plant-derived compounds, like *Lactobacillus,* are largely unknown. Though flavonoids represent a significant component of the human diet, the specific effect of flavonoids and their derivatives within the host signal transduction pathways, are just emerging. We hypothesized that flavonoids may serve as signal molecules that are indirectly involved with the modulation of probiotic traits. *In vitro* screening was conducted against a small library of flavonoids, using purified transcription factors from the LTTR family. KaeR was successfully identified as a transcriptional regulator that binds kaempferol as an effector molecule. Since this is the first study at the molecular level to identify a *Lactobacillus* species responsive to flavonoids, this work was directed towards understanding the mechanisms of transcriptional regulation by this new member of the LysR family.

The LTTR family is one of the largest in prokaryotes (Pareja *et al*., 2006), yet the mechanisms behind their regulatory actions are largely unknown. Members of the LTTR family regulate transcription of a diverse population of genes, including those involved with antibiotic resistance, sporulation, virulence, DNA replication and nodulation (Maddocks and Oyston, 2008). While LTTR members are abundant in most bacterial genomes, only seven members are known to be encoded in *L. brevis* (5% of the transcription factors). These proteins share low sequence identity and most genes under their regulation have unknown functions. KaeR shows the typical genomic organization found in transcription factors that belong to the LTTR family. It is encoded divergently of three genes: *LVIS1988, LVIS1987* and *LVIS1986*. The first two genes (*LVIS1988* and *LVIS1987*) have high sequence identity to 3-polyprenyl-4-

hydroxybenzoate decarboxylase, while *LVIS1986* has no homologs with known function. These genes, including *kaeR*, were upregulated by the addition of kaempferol to the growth media. Although the biochemistry of these genes is unknown, they may be involved in kaempferol modification or degradation.

Several mechanisms of LTTR-mediated gene regulation have previously been described (Schell, 1993; Maddocks and Oyston, 2008). Of these, the large majority involve transcriptional activation of the regulated genes, while the LTTR is negatively autoregulated. Furthermore, the mode of regulation is most frequently determined by the location and affinity of the DNA binding sequences. Based on the results obtained *in vivo* with *LacZ* fusions, KaeR is positively autoregulated. Very few examples have previously been reported with this mode of regulation. YtxR from *Yersinia enterocolitica* (Axler-DiPerte *et al*., 2006) controls the expression of the *ytxAB* genes potentially involved in the production of an enterotoxin. Although *ytxR* is also positively autoregulated, it has been proposed that a small molecule might not be required, since increased expression of YtxR is sufficient to activate transcription (Axler-DiPerte *et al*., 2006). These results indicate that the mechanism of KaeR is different from that of YtxR in the sense that KaeR is positively autoregulated upon interaction with the inducer molecule.

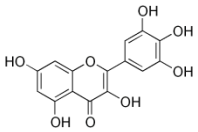
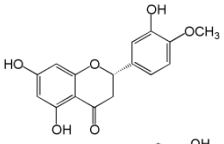
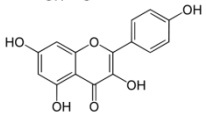
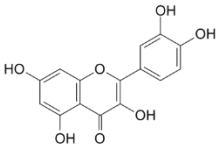
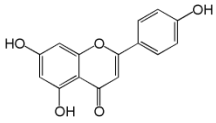
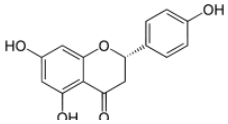
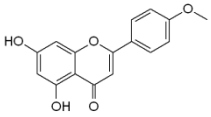
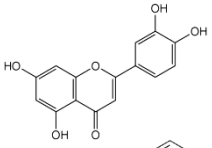
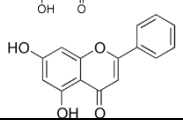
The operon structure of KaeR is commonly found in other LTTR members that are negatively autoregulated. Such is the case with NodD, the only other LysR member previously studied for regulation by flavonoids (Peters *et al.*, 1986; Redmond *et al.*, 1986; Firmin *et al.*, 1986; Rossen *et al*., 1985). Interestingly, similarities between KaeR and NodD also include the presence of overlapping promoters, both proteins are

divergently encoded by the genes they regulate, and each is capable of responding to flavonoids. KaeR, however, does not share sequence identity to NodD, and the mechanism of regulation seems to differ significantly between the two proteins. Previous studies have shown NodD is expressed constitutively in *Rhizobium leguminosarum,* and it is negatively autoregulated by flavonoids (Rossen *et al*., 1985).

According to the classical model of transcriptional activation, LTTRs usually bind DNA as oligomers (tetramers in general) in absence of a signal molecule. The “sliding dimer” model of regulation has been proposed in AtzR in which two subunits bind the RBS while two subunits bind the ABS, provoking a sharp bend in the DNA (Porrua *et al*., 2007). Subsequent binding of cyanuric acid would relax the DNA bend, facilitating the interaction with RNAP, and thus leading the enhanced expression from the promoter. Here I showed that KaeR binds two sequences. Site I is located -39 to +2 bp upstream of *kaeR* (*LVIS1989*) while site II is located within *LVIS1988*, -314 bp to -353 bp upstream of site I. *In vivo* I observed that while both sites are required for optimal expression, induction by kaempferol can be attained by the presence of site I alone. There are few previously documented studies where LTTR binding sites were located within the up-regulated genes. LadA from *Myxococcus xanthus*, binds in the *dev* operon between +319 and +376 (Viswanathan *et al.*, 2007b). Similarly, MetR from *Salmonella typhimurium* binds in *metF* from +62 to +85 (Cowan *et al.*, 1993). MetR binding to the downstream sequences also requires binding to the -95 to -50 sequences. In both LadA and MetR, binding results in gene activation. It has been proposed that MetR could act as an anti-repressor to destabilize MetJ binding though steric hindrance by either directly contacting the RNA polymerase or through DNA bending (Cowan *et al*., 1993).

A similar hypothesis was formulated for LadA based on the fact that negative regulatory elements were found between +219 and +280 in the *dev* operon (Viswanathan *et al.*, 2007a). Although I have not determined the direct binding of another transcription factor to the F_{kae2} fragment, it is possible that KaeR act through a similar mechanism of DNA bending followed by activation of gene expression.

Table 3-1. Stabilization effect of ligand binding. The thermal stabilization of 10 μM KaeR using 60 μM ligand was evaluated by fluorometry.

Chemical	Structure	ΔTm[1] (°C)
Myricetin		23.2
Hesperitin		2.6
Kaempferol		2.5
Quercetin		0.2
Apigenin		0.2
Naringenine		0
Acacetin		0.2
Luteolin		0
Chrysin		0

[1]Delta temperature was calculated as the difference in the transition temperature between the protein in the absence and presence of given ligand.

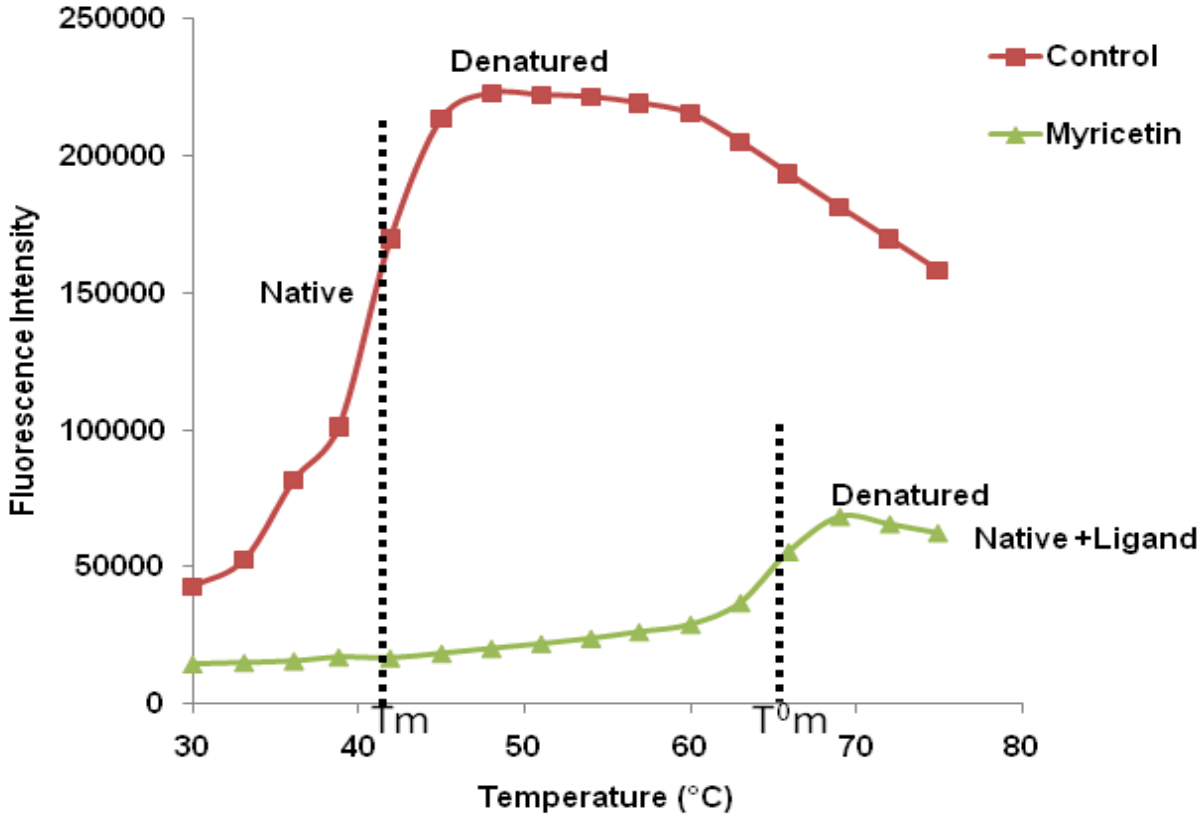


Figure 3-1. Fluorescence based ligand screening assay. The melting curve for KaeR native protein is shown by the red line whereas the melting curve of the KaeR with the ligand (myricetin) is indicated by the green line. The vertical line represents the respective midtransition temperatures. Tm- transition temperature of unfolding in the absence of ligand, T^0m- transition temperature of unfolding in the presence of ligand. Hits were determined by formula T^0m-Tm = > 2 ^{0}C.

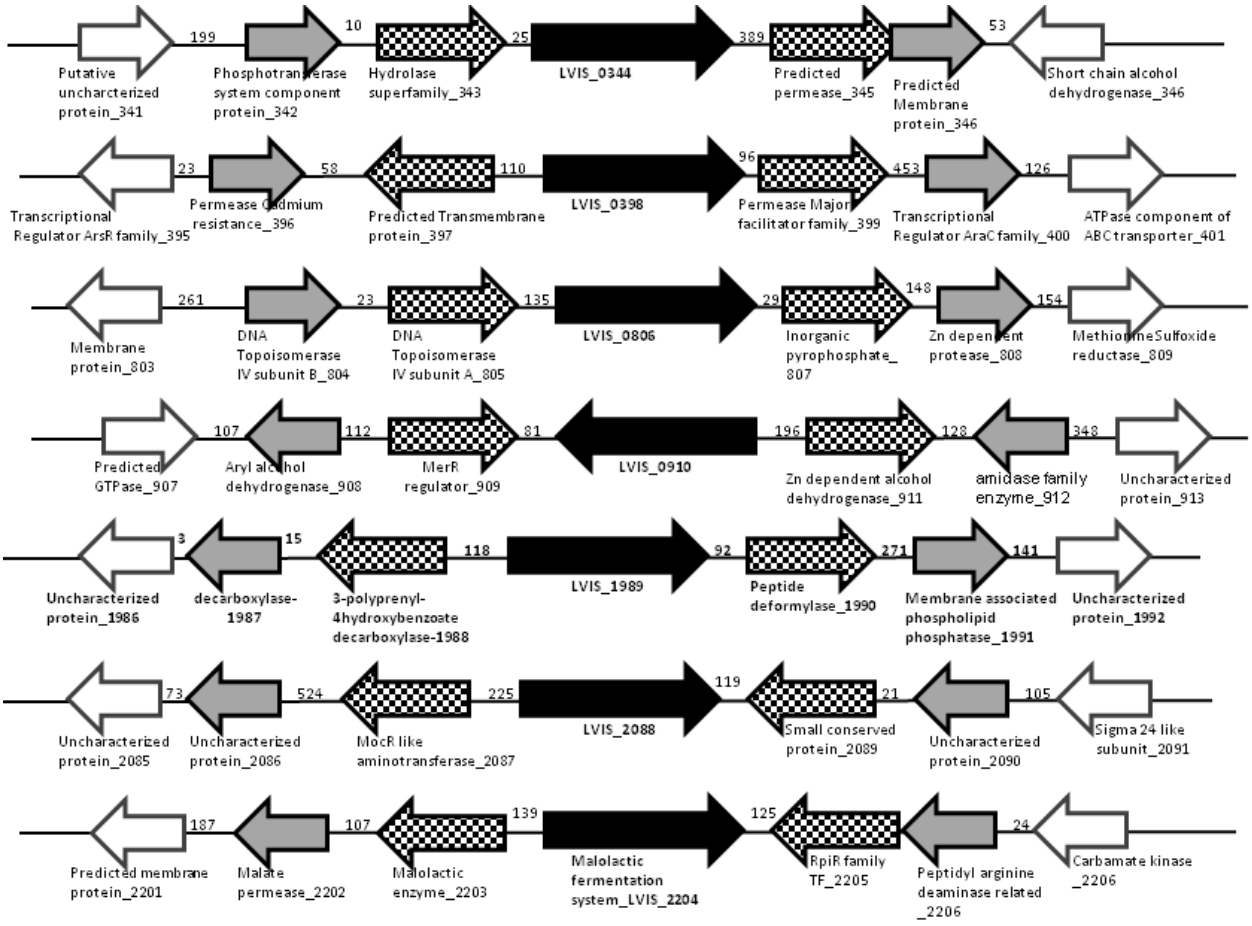


Figure 3-2. Genomic environment of the *L. brevis* genes encoding LysR type transcriptional regulators. The genes encoding LTTRs are shown in black. The numbers represent the number of bp in the predicted intergenic regions.

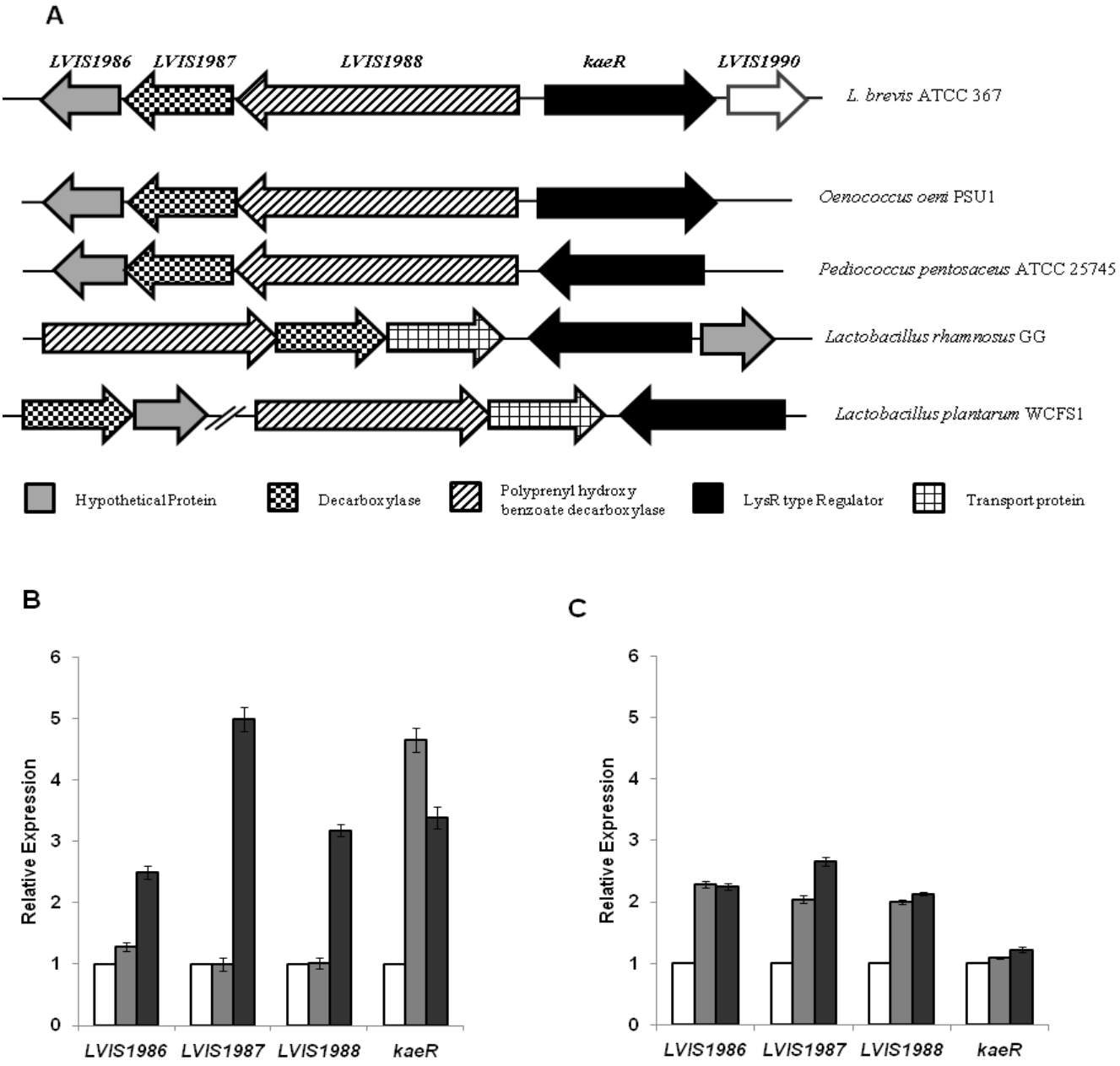
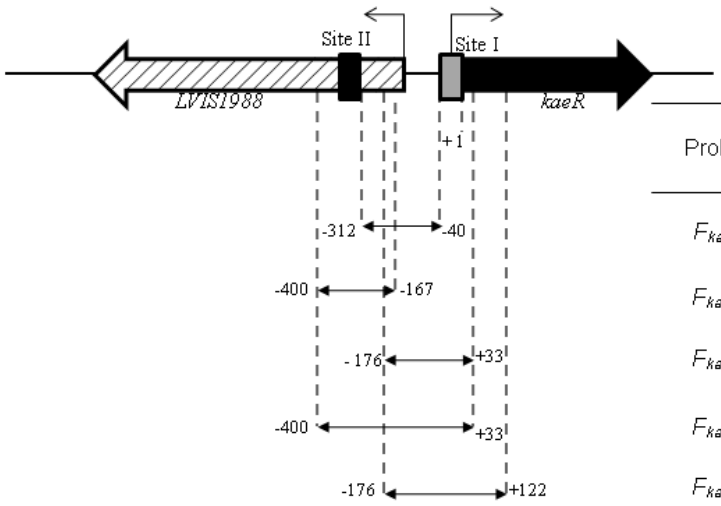


Figure 3-3. Identification of the genes regulated by LVIS1989 (KaeR). (A) The genomic environment of *LVIS1989* was extracted and compared to the closest ortholog neighborhood regions. (B and C) Effect of addition of flavonoids on mRNA levels of genes encoded divergently to *LVIS1989*. *L. brevis* cells were grown in MRS broth in absence (control, white bars) and presence of 10 μM myricetin (grey bars) or kaempferol (black bars) and cells were collected during exponential phase (B) or stationary phase (C). RNA extractions and qRT-PCR was performed as described under “Materials and methods”. The amplification values obtained were corrected with those obtained using *rpoD* as internal control. The values shown are relative to those observed for the same gene in cells grown in absence of flavonoids.

A



Probe	Primers used	KaeR:F_{kaeR} complex
F_{kae4}	E8,E9	-
F_{kae3}	E4,E6	-
F_{kae1}	E2,E7	±
F_{kae2}	E4,E7	+++
F_{kae5}	E2,E1	++

B

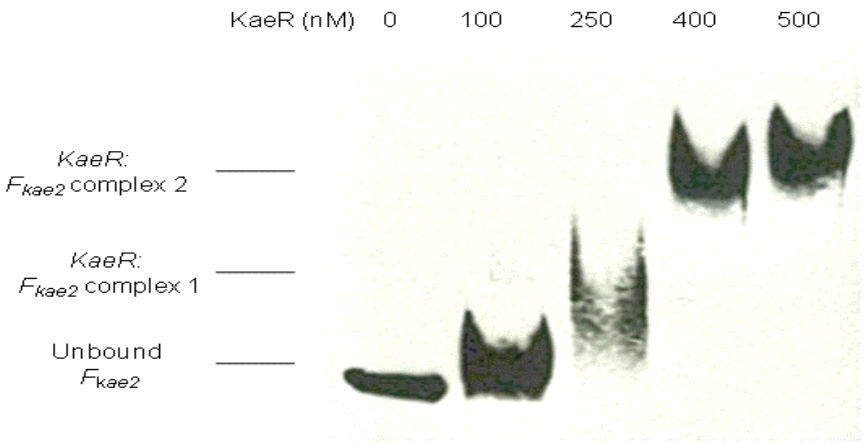


Figure 3-4. Identification of the KaeR binding region. (A) To determine the DNA region for KaeR binding various DNA fragments were obtained by combining different primers in the *LVIS1988-LVIS1989* intergenic region. EMSA results are summarized in the inset Table. (+) Positive binding of KaeR using 250 nM of pure protein; (-) Negative binding using up to 500 nM of purified KaeR protein. (B) EMSA assays performed using increasing concentrations of KaeR (100-500 nM, as indicated) with fragment F_{kae2}.

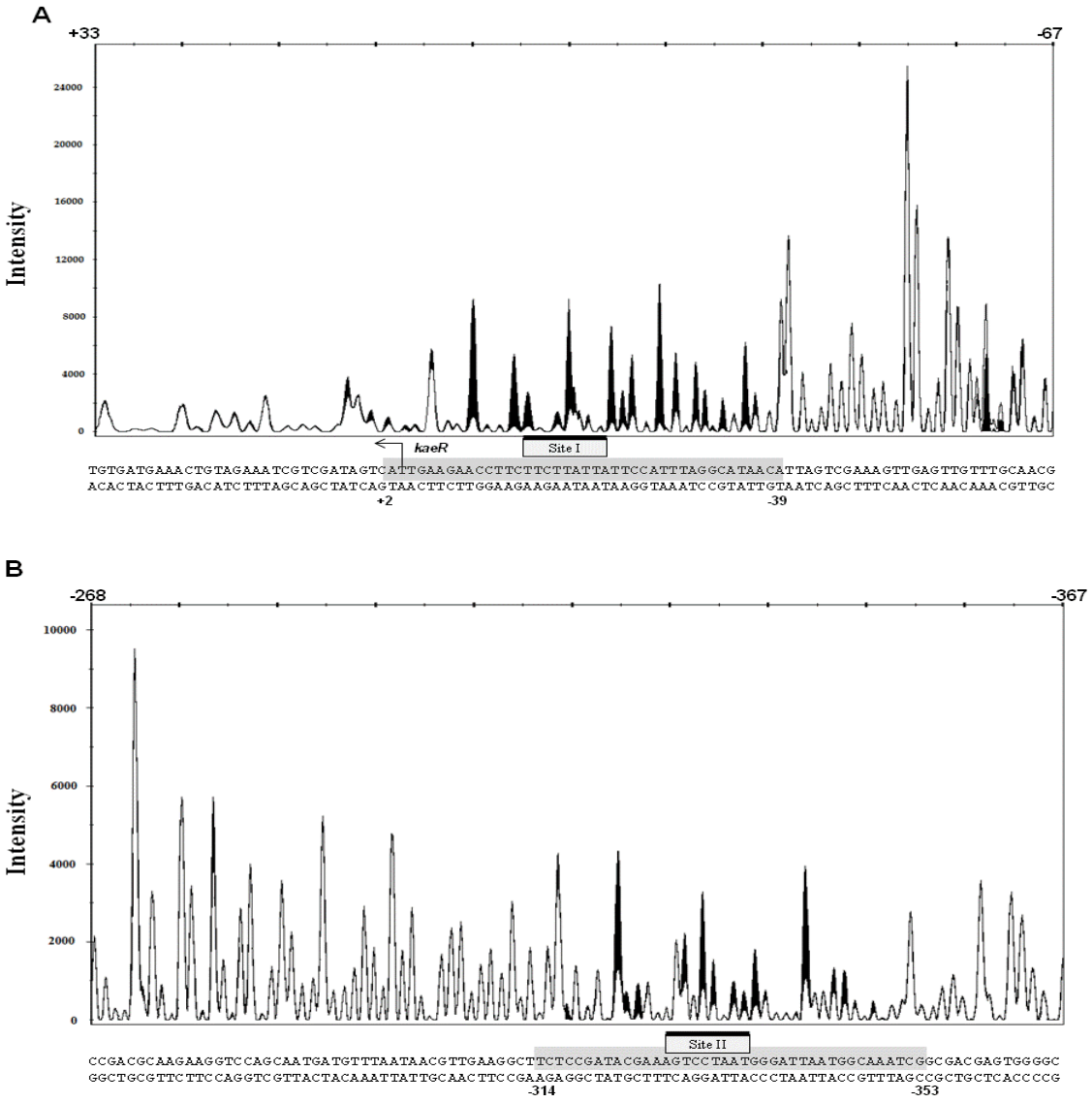


Figure 3-5. Identification of KaeR binding sites. DNase I footprinting assays identified (A) a protected site I located upstream KaeR (residues +2 to -39, from the translation start point of KaeR); and (B) protected site II located within *LVIS1988* (residues -314 to -353, from the translation start point of *kaeR*). The electropherogram shows a fragment of the digested probe in absence (black) or presence (white) of KaeR highlighting the protected region. The reaction mixture was treated as described in “Experimental procedures” using a probe generated by using the primers D1_Fw and E3_Rv shown in Table 2-4. The nucleotide sequence protected by KaeR is shown in the bottom of the panels. (C) Analysis of the intergenic region between the divergently transcribed genes *LVIS1988* and *LVIS1989*. The protected sites I and II are boxed. Predicted Shine Dalgarno sequence (SD) and -10 and -35 of the P_{KaeR} is indicated.

C

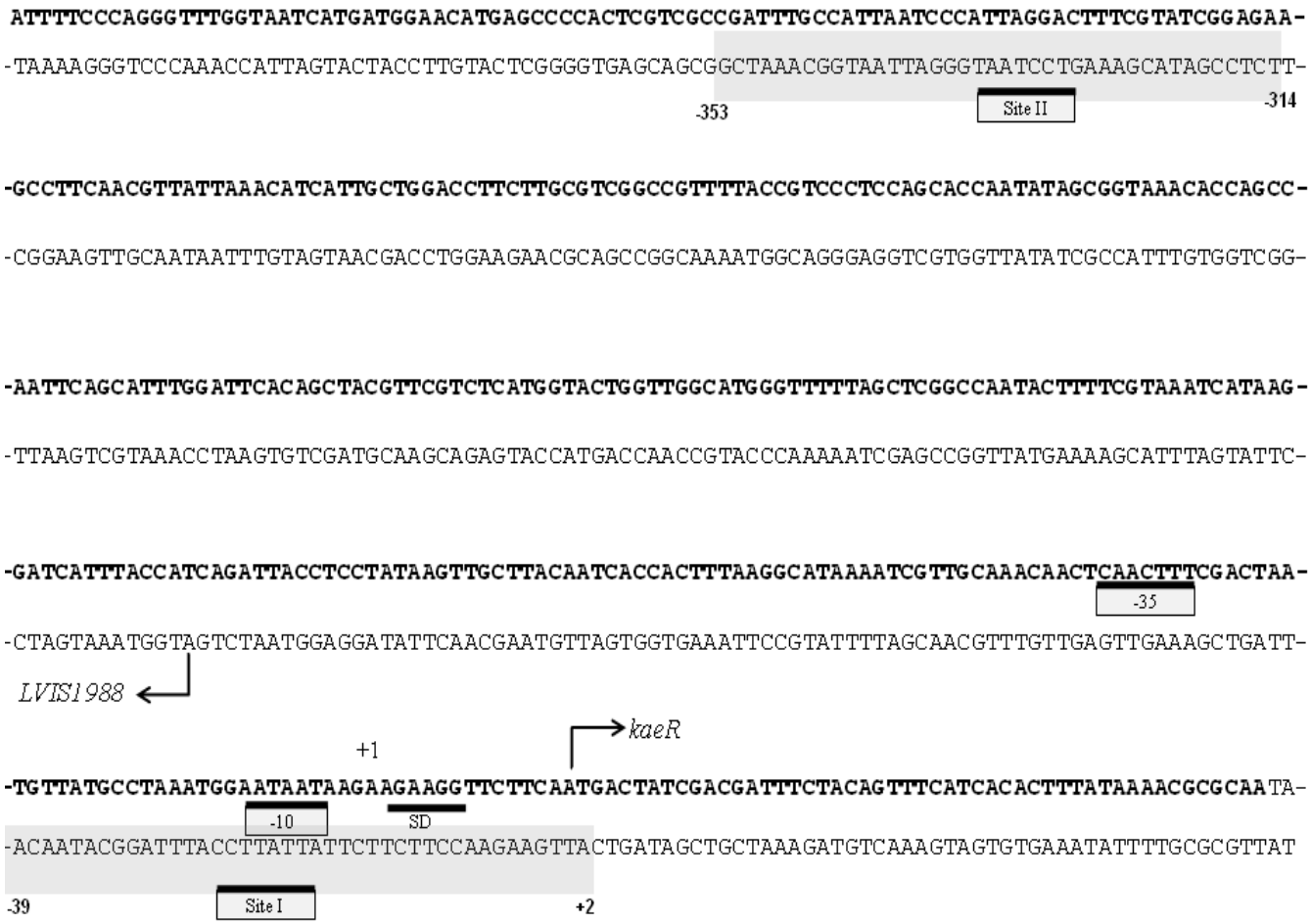


Figure 3-5. Continued

.

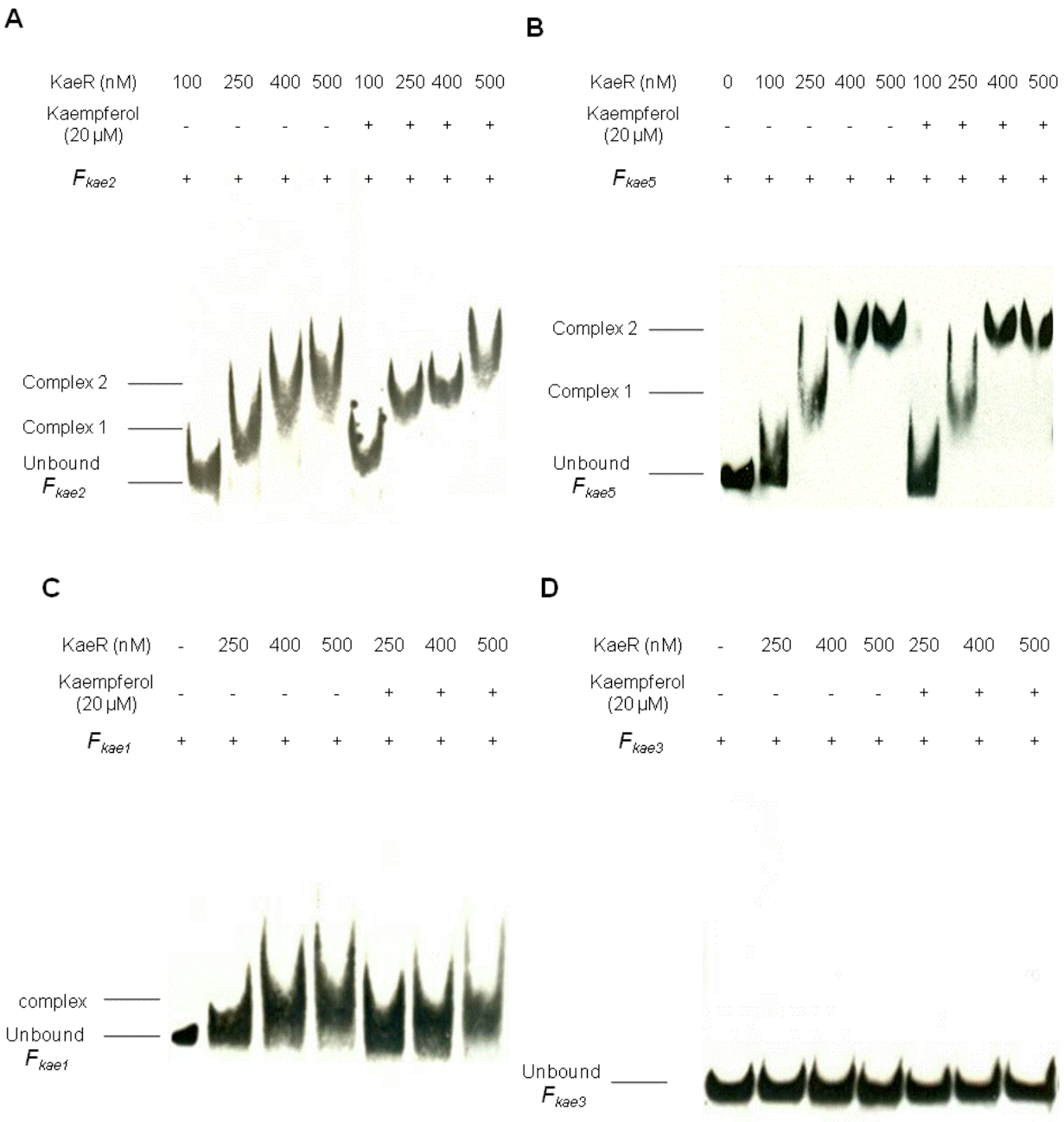


Figure 3-6. Effect of kaempferol on KaeR binding to the different binding sites. EMSA assays were performed using (A) *Fkae2* (containing binding site I and II), (B) *Fkae5* (containing site I), (C) *Fkae1* (containing site I) and (D) *Fkae3* (containing site II) fragments. The concentration of kaempferol was maintained at 20 µM. The presence or absence of each key component in the reaction mix is indicated on top of each panel. The full binding conditions are described under “Materials and methods”.

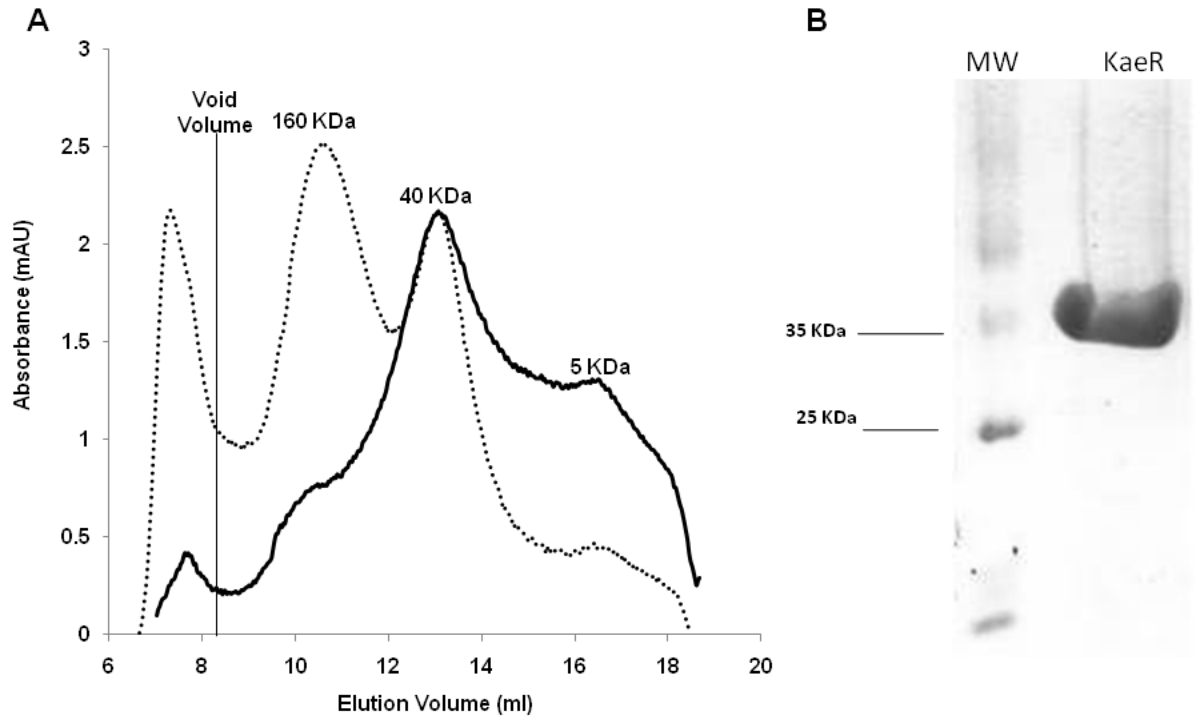
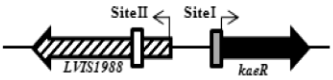
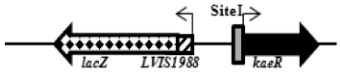
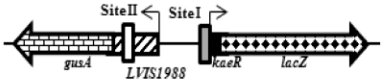
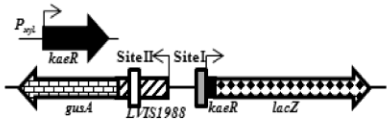
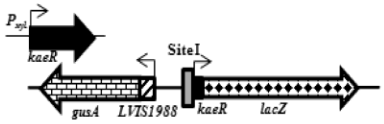
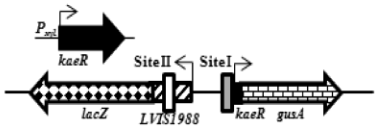


Figure 3-7. (A) Gel filtration analysis of KaeR oligomerization in presence of kaempferol. KaeR preincubated in absence (continuous line) or presence of 10 μM kaempferol (dotted line) was passed through a Superose 12 column, as described in “Materials and methods”. (B) SDS-PAGE showing the homogeneity of KaeR after purification by nickel affinity chromatography.

Strain	Genotype	β-galactosidase activity[a]					
		LB	LB + Xyl[b]	LB + Kae[c]	LB + Myr[d]	Fold induction Kae[c]	Fold induction Myr[d]
L. brevis		NA	NA	NA	NA	NA	NA
SP01		54 ± 0.9	NA	183 ± 1.6	122 ± 0.9	3.4	2.3
SP02		17 ± 2.7	NA	42 ± 5.0	46 ± 2.7	2.5	2.7
SP03		48 ± 2.2	47 ± 1.5	53 ± 3.0	ND	1.1	ND
SP04		44 ± 3.0	47 ± 6.3	89 ± 6.4	103 ± 4.6	1.9	2.2
SP05		36 ± 1.2	34 ± 3.0	98 ± 13.4	104 ± 0.2	2.9	3.1
SP06		19 ± 0.1	13 ± 0.9	26 ± 0.1	ND	2	ND
SP07		ND	17 ± 5.0	17 ± 3.0	ND	1	ND

[a]Enzyme activity was assayed and expressed in Miller units as described in materials and methods. [b]Xylose, [c]kaempferol, [d]myricetin, NA-not applicable, ND-not determined.

Figure 3-8. *In vivo* assessment of KaeR's binding sites. The analyses were performed using *B.subtilis* as a surrogate strain. For each strain used, a graphical representation is made that indicate the presence of site I or II, the location of *kaeR* (in *cis* or integrated at the *lacA* locus) and the orientation of the fusion. Cells were grown to early stationary phase and β-galactosidase assays were performed as indicated in "Materials and methods". NA-not applicable, ND-not determined.

CHAPTER 4
IDENTIFICATION OF KAER AMINO ACIDS INVOLVED IN LIGAND INTERACTION

Introduction

The objective of this study was to identify the KaeR amino acids involved in the kaempferol interaction and oligomerization. CbnR from *Ralstonia eutropha* NH9 was the first full length LTTR structure crystallized in 2003 (Muraoka *et al*., 2003). Though LTTRs are the most abundant family in prokaryotes, only four LTTR structures have been determined, namely, CbnR (Muraoka *et al*., 2003), CrgA (Sainsbury *et al.*, 2009), ArgP (Zhou *et al*., 2010), and TsaR (Monferrer *et al.*, 2010). It is thought that the low solubility of these proteins and the high flexibility of the helix-turn-helix DBDs under crystallization conditions are the reasons for the low success obtained with this family of proteins. No effector has been co-crystallized with proteins either. More protein structures are available for effector binding domains (EBDs) of LTTRs. Some examples include OxyR (Choi *et al.*, 2001), CysB (Tyrrell *et al.*, 1997), BenM (Craven *et al*., 2009), and CbI (Stec *et al.*, 2006).

Based upon ligand bound EBD crystal structures it became evident that the C terminal residues are involved in ligand interaction. This was also confirmed by genetic studies as assessed in CbI (Stec et al., 2006), BenM (Craven *et al*., 2009), OxyR (Choi *et al*., 2001), CysB (Lochowska *et al*., 2001), and ArgP (Zhou *et al*., 2010). Mutation in the C-terminal domain residues caused altered transcriptional activity from the respective promoters. Also, the genetic analysis and bioinformatics studies suggested that the N-terminal 60-70 residues are involved in DNA binding (Schell, 1993; Lochowska *et al*., 2001). As described previously, the DNA binding helix turn helix spans from the 23 to 43 amino acid residues. The C-terminal domain is divided into two

sub domains regulatory domain I (RD I) and regulatory domain II (RD II) (Schell, 1993; Tyrrell *et al*., 1997). The co-inducer binding cleft is present between these two domains. It has also been shown that the last 60 amino acids in the C- terminal of NahR are necessary for optimal DNA interaction (Schell *et al*., 1990). Similarly, in the ArgP crystal structure, the DBD interaction with the extreme C-terminal residues of the effector binding domains was observed supporting the possible role of the C-terminal residues in DNA interaction (Zhou *et al*., 2010).

In CbnR, two forms of the protein were observed, a compact form and another extended form. The CbnR dimers are comprised of the two subunits, a compact form subunit and extended form subunit that interact through the linker helices (Figure 4-1). In the crystal structure, CbnR was dimer of dimers in which the DBDs were aligned along one face of the tetramer whereas the regulatory domains were located on the opposite side. The regulatory domains in the dimers were arranged in an anti parallel fashion. The regulatory domains face had a 50 Å cleft to accommodate a ligand. Similar tetrameric arrangement was observed in the case of ArgP (Figure 4-2). In addition to the two forms of the monomers within a dimer, two forms of dimers were observed in ArgP. The two dimers differed in the dimerization domains, one dimerized through DNA binding domains whereas another through the regulatory domains (Zhou *et al*., 2010) (Figure 4-3). The TsaR structure confirmed that the two types of the monomers might be a common theme in the LTTR family. The dimers were formed from a compact and an extended subunits of the protein which dimerized through the antiparallel linker helices. However, the tetrameric TsaR structure was different from that of CbnR and ArgP. In the TsaR tetramer, the regulatory domains were arranged in antiparellel

fashion but the DBDs were arranged diagonally. The TsaR tetramer had a flat square shape rather than a diamond shape which was evident in CbnR and ArgP tetramers. Based upon the flexible regulatory domain and the lack of extensive interface contacts among C-terminal domains it was suggested that the structure represented the active ligand bound “open form” in contrast to the inactive “closed form” structure found in CbnR (Monferrer *et al.*, 2010).

Thus, most of the crystallized LTTR structures suggested that the dimers can be composed of open and closed subunits whereas the DBDs and RDs are arranged to favor the classical model of gene regulation by LTTRs. However, another protein CrgA is the only exception to this rule. It is ring shaped where only the “compact” form of the monomers is found (Sainsbury *et al*., 2009).

Since the structure for KaeR has not been solved, a bioinformatics approach was used to construct a structural model. The EBD of the modeled KaeR structure was aligned to other proteins crystallized with a ligand and putative residues involved in flavonoid interaction were identified. The amino acid residues were modified to alanine by site directed mutagenesis. The mutant proteins were expressed in *E. coli* and purified by using Ni affinity chromatography as described in materials and methods. The purified proteins were used for size exclusion chromatography and EMSA studies to assess the effect of the mutations on oligomerization, DNA interaction and in the response to flavonoids.

Results

KaeR Structure Modeling-

The protein structure prediction was performed by using the web interface programs SWISS-MODEL workspace (Benkert *et al*., 2011) and PHYRE

(www.sbg.bio.ic.ac.uk/~phyre, Protein Homology Analogy Recognition Engine, Imperial College, London) (Kelley, 2009) (Table 4-1). The search in PHYRE program was able to identify a template protein, PDB 1IXC. This protein corresponds to full length structure of CbnR from *Ralstonia eutropha* NH9 (Muraoka *et al*., 2003). KaeR had a 15 % sequence identity to CbnR. The predicted structure was confirmed by using another program, SWISS-MODEL workspace (automated mode query), which has a different algorithm to predict the structure. In PHYRE, the C-terminal domains were also modeled. The program identified CynR (PDB # 3HFU) as a EBD structure with the highest probability and sequence identity (100 % and 23 % respectively). Table 4-1 summarizes the modeling results obtained.

In the KaeR modeled structure, the DBD is comprised of 3-88 residues whereas 89-293 residues are part of C-terminal regulatory domains. A predicted wHTH domain is connected to the regulatory domain through linker helix formed from 60 to 88 amino acids. The regulatory domain I (RD I) is formed by amino acids 89 to 167 and 268 to 293 whereas regulatory domain II (RD II) is formed by residues 168 to 267. The regulatory domain I is comprised of 5 beta strands and 4 alpha helices and regulatory domain II is comprised of 3 beta strands and 5 alpha helices (Figure 4-4). To compare the two models, the respective PDB files were downloaded and aligned using PyMOL. The two models were highly similar (Figure 4-5 A). Although KaeR was modeled to the CbnR full length protein, the structure of CynR had a small molecule in the predicted binding pocket.

Based on these results further analysis was carried out using the CynR (PDB # 3HFU) structure as the template.

Identification of Residues that Modulate KaeR Activity

CynR from *E. coli* is involved in cellular response to azide (Singer et al., unpublished). Although the molecular structure of CynR is publically available in the RCSB server, the analysis of the structure of CynR has not been published yet. I performed the analysis on the CynR ligand binding pocket. As expected the azide molecule is located in a cleft formed between regulatory domain I and II (Figure 4-5 C). Within the same cleft another small molecule, ethanediol was found. This compound is commonly used as a protein stabilizer during the crystallization conditions. Although, it is not expected that ethanediol has a biological role in CynR, it may be required to stabilize the protein.A structural alignment between KaeR and CynR was performed (Figure 4-5A). The ligand binding site overlapped well. In contrast to CynR, where two binding sites were observed, a large continuous binding site was present in KaeR (Figure 4-5B). These differences may be due to a bigger ligand (i.e. kaempferol) can be accommodated in the KaeR cavity.

To identify the residues in KaeR that mediate binding with kaempferol, the residues involved in ligand binding in CynR were analyzed. The ligand explorer tool at the RCSB was used. Figure 4-6 shows the residues involved in azide binding. Threonine 100 and 200 are responsible for hydrogen bonds with the ligand. T200 is the major residue responsible for the hydrogen contacts with the N1, N2 and N3 in azide. T100 forms one hydrogen bond through a water molecule (bridged hydrogen bond). In the structural alignment of CynR and KaeR, the CynR T100 and T200 correspond to the P100 and V205 in KaeR. The additional residues R148 and Q153 were identified within 4 Å radius of azide in KaeR. Ethanediol interacted with the P99, E126 and I224 residues in CynR. In KaeR these residues correspond to P100, E127 and I229 (Figure 4-7).

Next, I used site directed mutagenesis to determine if the residues identified in the putative ligand binding pocket of KaeR are involved in the modulation of KaeR activity. To this end, residues P100, E127, R148, Q153, V205 and I229 (Figure 4-8) were changed to alanine. The effects of mutations were assessed on the DNA binding capabilities as well as on the oligomeric form of the protein.

Effects of Mutations on the DNA Binding Properties of KaeR-

As described earlier (chapter 3), KaeR binds two sequences within the promoter region of *LVIS1988 / KaeR*. *In vitro*, using EMSA assays, it was observed that KaeR binding to F_{kae2} results in an intermediate complex at low protein concentration (250 nM) (complex 1) and a large complex (complex 2) at concentrations higher than 400 nM. The effects of mutations on KaeR were first assessed on the ability of the proteins to bind to the DNA (Figure 4-9). At protein concentrations of 100 nM, the P100A, V205A, Q153A E127A and I229A mutants showed a weak DNA binding while no binding was observed for the wild type KaeR at this concentration. All mutants tested were able to form complex 1 at concentration similar to the wild type protein. It was observed that mutations on P100A and Q153A dramatically affected the formation of complex 2. Mutations on V205A, E127A and I229A resulted in a formation of weak complex 2. The mutation on R148A had no effect on the formation of higher order complex.

The effect of kaempferol was then tested on each mutant. The V205A KaeR mutant formed a very weak complex 2 that it was disrupted in presence of kaempferol. All other mutations failed to modify their binding to DNA in presence of the ligand.

Effect of Mutations on the Oligomeric State of KaeR

Based on the defects on complex 2 formation displayed by the mutants in EMSA assays, the purified proteins were subjected to size exclusion chromatography studies.

As previously described (chapter 3), KaeR eluted as a monomer (40 kDa observed molecular weight). The addition of kaempferol resulted in the formation of tetramers (160 kD observed molecular weight). All mutations performed in the hinge 3 displayed changes in the oligomeric state of KaeR. Mutations on P100A, V205A resulted in elution as tetramers that were not changed by the addition of kaempferol. KaeR mutated in Q153A, I229A and R148A eluted as dimers (71 kDa). Addition of kaempferol favored the oligomerization of the proteins into larger complexes (216 kDa) (Figure 4-10). The major species of KaeR E127A were tetrameric (160 kDa) whereas a smaller peak (34 kDa) that may correspond to the monomeric form was also observed. The addition of 10 μM kaempferol did not modify the elution profile.

Conclusions

The signal transduction mechanisms underlying transcriptional activation are poorly understood. It involves inter-domain communication upon binding, of a ligand. Some research groups have proposed that the linker helix (between DBD and EBD) is involved in signal transduction (Lu *et al*., 2010). Our results indicate that in KaeR the flexibility of the EBD may mediate this process.

In the modeled KaeR it was observed that the hinge 3 region in EBD (between RDI and RDII) is also involved in dimerization. To determine if the same residues that mediate dimerization are involved in ligand binding, I conducted site directed mutagenesis. Interestingly, the mutants displayed changes in the oligomeric state of the proteins and impaired DNA binding. In CynR, E126 (equivalent to KaeR E127) is located in the dimerization interface (Figure 4-11) and in the close proximity of the ethanediol molecule. The presence of a negatively charged residue in this interface may contribute to the flexibility in the hinge region. In KaeR, the removal of this residue

(E127 mutant) resulted in stable tetramers in solution that were able to oligomerize cooperatively on DNA. The addition of kaempferol did not modify the migration of the F_{kae2}:E127A complex or the oligomeric status of the protein in solution.

Based on these results, I propose that the hinge 3 region is the location of ligand interaction. The presence of the ligand would promote protein oligomerization by sequentially combining different forms of KaeR monomers (in open and closed conformations). This proposal is based on the observation that all residues mutated in the dimerization interface resulted in stable dimers or tetramers with DNA binding defects. Further experiments with additional mutants will be required to test this hypothesis.

Table 4-1. Summary of the KaeR homology modeling results.

Template	PDB #	E-value	% Seq[a] Identity[a]	Ligand	% Prob[b]
CbnR	1IXC	0	15		100
CatM Variant (R156H)	3GLB	0	20	(2Z,4Z)-HEXA-2,4-Dienedioic acid, Glycerol, Sulphate	100
BenM Variant (R156H, T157S)	2H99	0	18	Acetate, Chloride, Glycerol, Sulphate	100
CysB	1AL3	1.40E-45	17	Sulphate	100
DntR (Y110S, F111V)	2UYE	4.00E-38	15	Glycerol, thiocyanate	100
PA0477	2ESN	3.00E-37	19		100
CynR	2HXR	3.80E-32	23		100
CynR	3HFU	ND[c]	23	Azide, ethanediol	ND

[a]Sequence, [b]probability, [c]not determined.

Table 4-2. Summary of KaeR mutation studies

	DNA binding assays			Oligomeric state	
Protein	Complex 1	Complex 2	Modification by Kaempferol	(-) ligand)	(+) ligand
KaeR (WT)	+	+++	+	Monomer	Tetramer
P100A	+	±	-	Tetramer	Tetramer
V205A	+	++	+	Tetramer	Tetramer
R148A	+	+++	±	Dimer	Octomer
Q153A	+	±	±	Dimer	Octomer
E127A	+	+	-	Monomer and Tetramer	Monomer and Tetramer
I229A	+	++	±	Dimer	Octomer

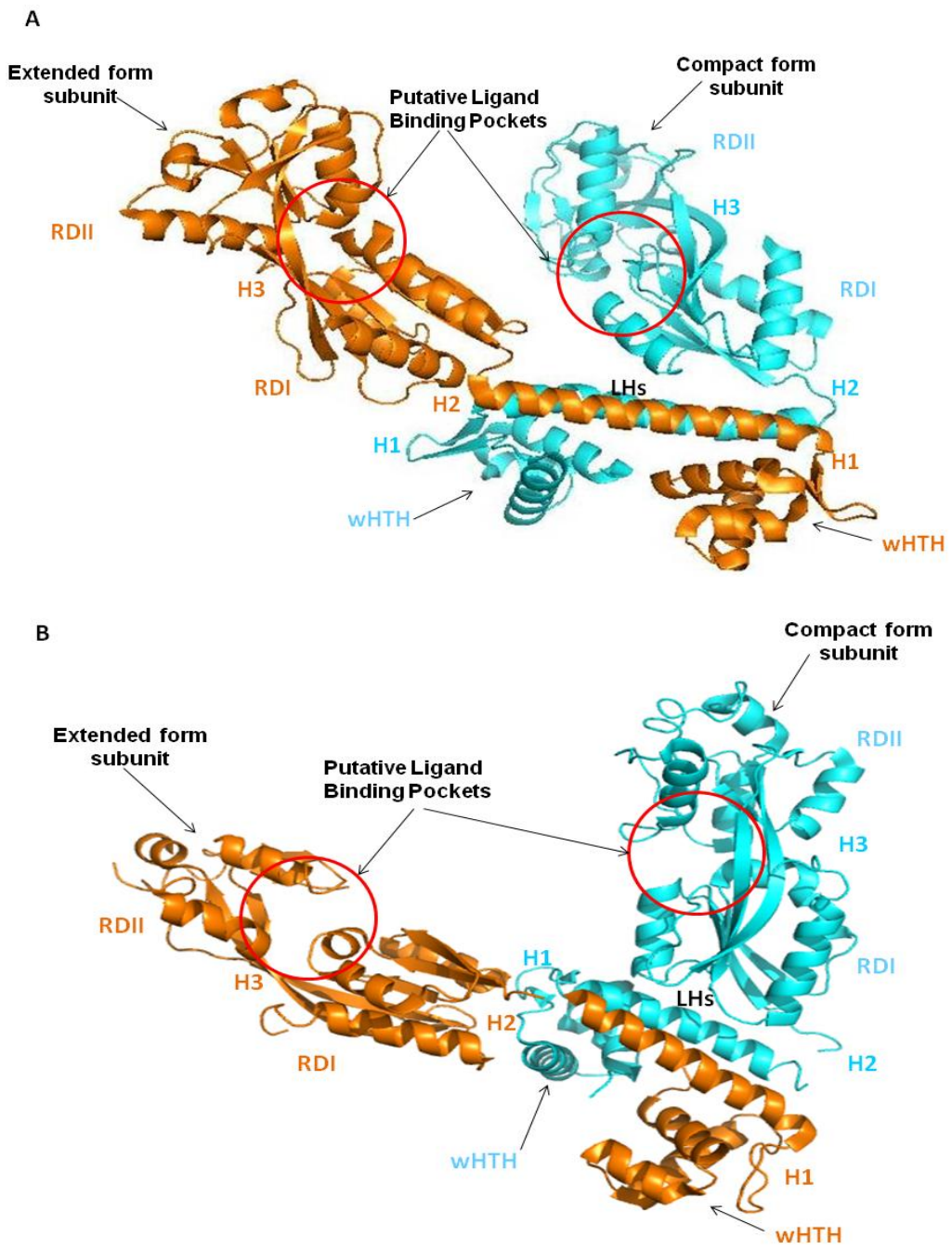


Figure 4-1. CbnR and ArgP dimers: A) CbnR dimer (PDB 1IXC). B) ArgP dimer (PDB 3ISP).The extended subunit is shown in orange color whereas the compact subunit is indicated in cyan color. The putative ligand binding pockets are indicated by a red circle. The ligand binding region is present in a cleft formed around hinge 3. RDI: regulatory domain I, RDII: regulatory domain II, wHTH: winged Helix-Turn-Helix, H1: Hinge1, H2: Hinge 2, H3: Hinge 3, LHs: Linker helices.

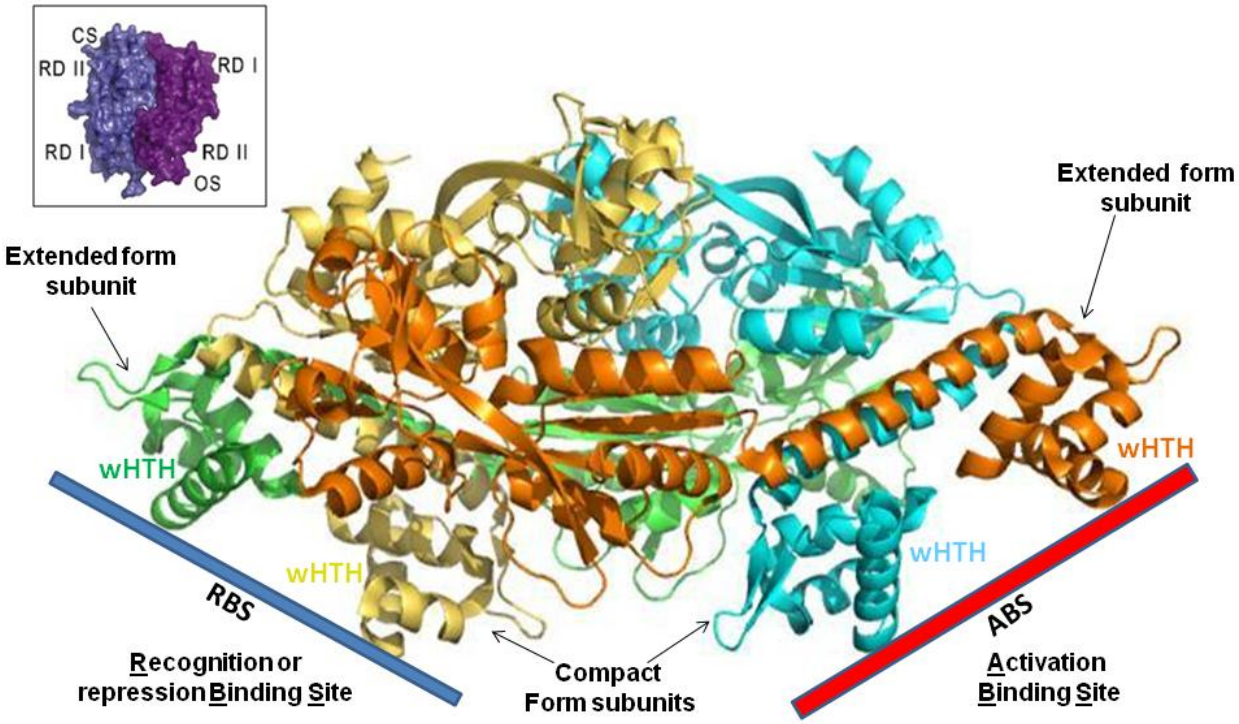


Figure 4-2. CbnR tetramer (PDB # 1IZ1): The DNA binding domains are present along one face of the tetramer. Corresponding DNA recognition motifs are indicated as wHTH (Winged Helix-Turn-Helix). Two DBDs are involved in ABS recognition whereas the remaining two recognize the RBS. The regulatory domains in each dimer are arranged in an antiparallel fashion. The ArgP tetramer regulatory domains arrangement is shown in inset. RBS: Recognition or Repression Binding Site, ABS: Activator Binding Site, wHTH: winged Helix-Turn-Helix,CS: Closed subunit protein (compact form subunit), OS: Open Subunit protein (Extended form subunit)[1].

[1]Inset figure reprinted with permission from Elsevier. Zhou, X., Lou, Z., Fu, S., Yang, A.,Shen, H., Li, Z., Feng, Y., Bartlam, M., Wang, H., and Rao, Z. (2010) Crystal structure of ArgP from *Mycobacterium tuberculosis* confirms two distinct conformations of full-length LysR transcriptional regulators and reveals its function in DNA binding and transcriptional regulation. *J Mol Biol* **396**: 1012-1024.

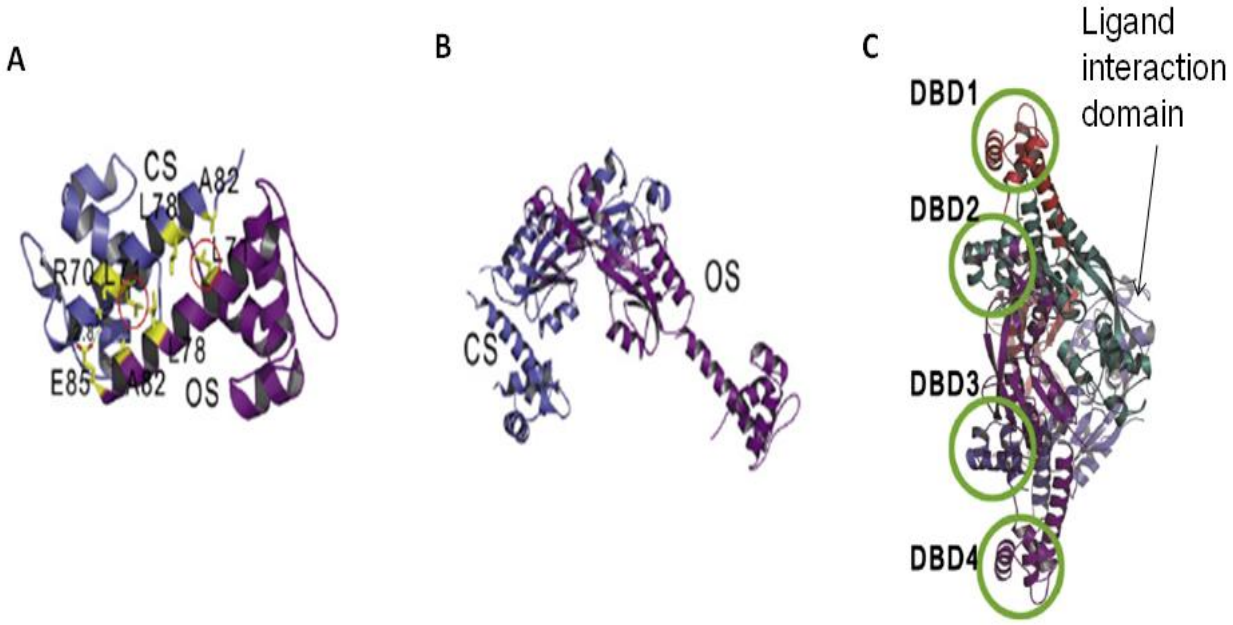


Figure 4-3. ArgP tetramer (PDB # 3ISP): A) DNA binding domain dimer and B) Regulatory domain dimer C) ArgP tetramer. The DNA binding domains are indicated by circles. CS-Closed subunit protein (Compact form), OS-Open Subunit protein (Extended form), DBD1-DNA Binding Domain 1, DBD2- DNA Binding Domain 2, DBD3- DNA Binding Domain 3, DBD4-DNA Binding Domain 4[2].

[2]Reprinted with permission from Elsevier. Zhou, X., Lou, Z., Fu, S., Yang, A., Shen, H., Li, Z., Feng, Y., Bartlam, M., Wang, H., and Rao, Z. (2010) Crystal structure of ArgP from *Mycobacterium tuberculosis* confirms two distinct conformations of full-length LysR transcriptional regulators and reveals its function in DNA binding and transcriptional regulation. *J Mol Biol* **396**: 1012-1024.

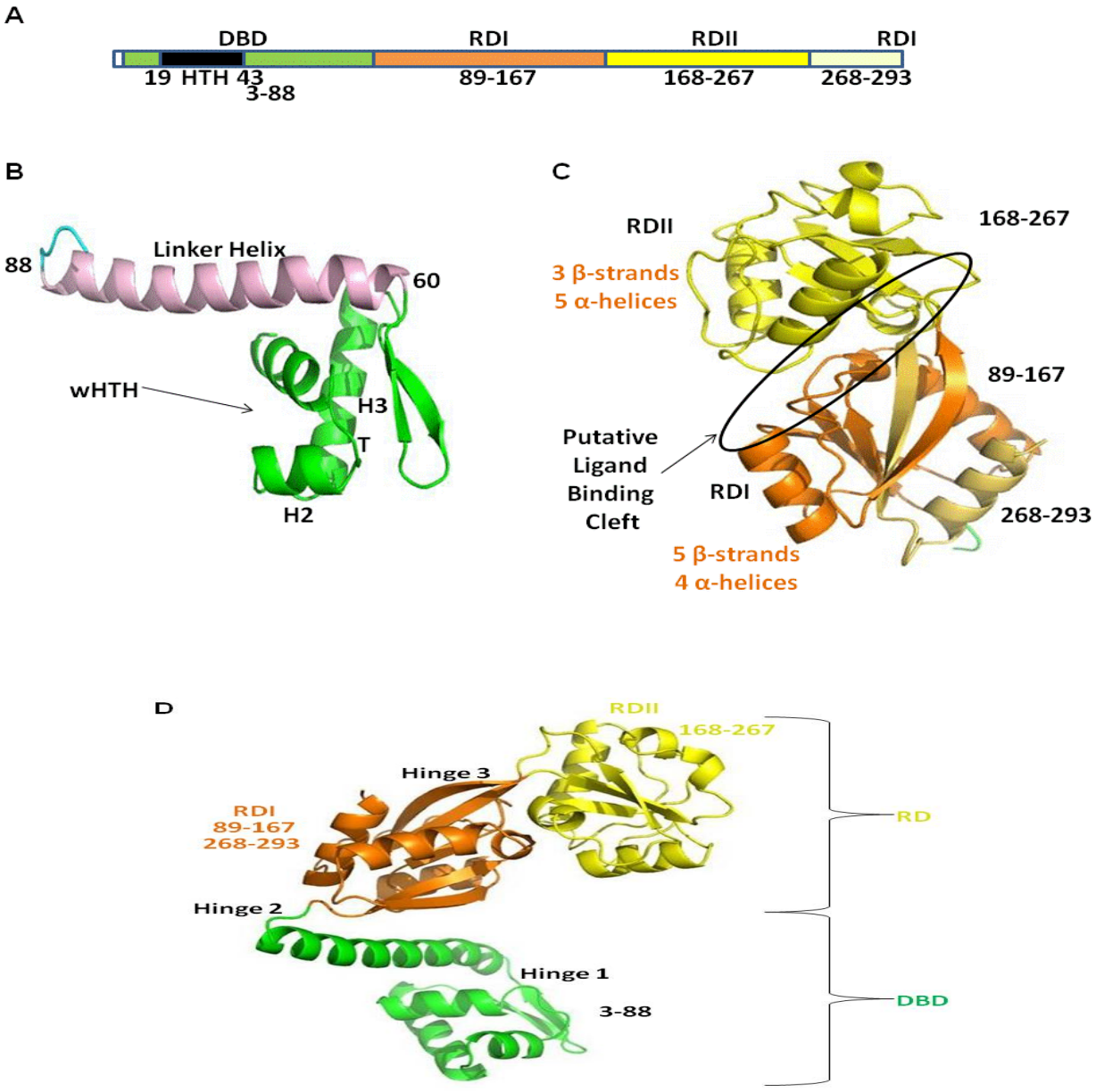


Figure 4-4. KaeR domains arrangement: DNA Binding Domain (DBD) is indicated by a green box, RDI is shown in orange and light yellow. RDII is shown in yellow. B) The KaeR DNA Binding Domain (DBD). Helices involved in DNA recognition are indicated as H2 and H3. C) KaeR regulatory domains arrangement. The putative ligand binding pocket is shown by a black circle. D) the full length KaeR modeled structure. H2-Helix 2, H3-Helix 3, wHTH-winged Helix-Turn-Helix, Regulatory domain I, RD-regulatory domain, RDII-regulatory domain II, DBD-DNA Binding Domain.

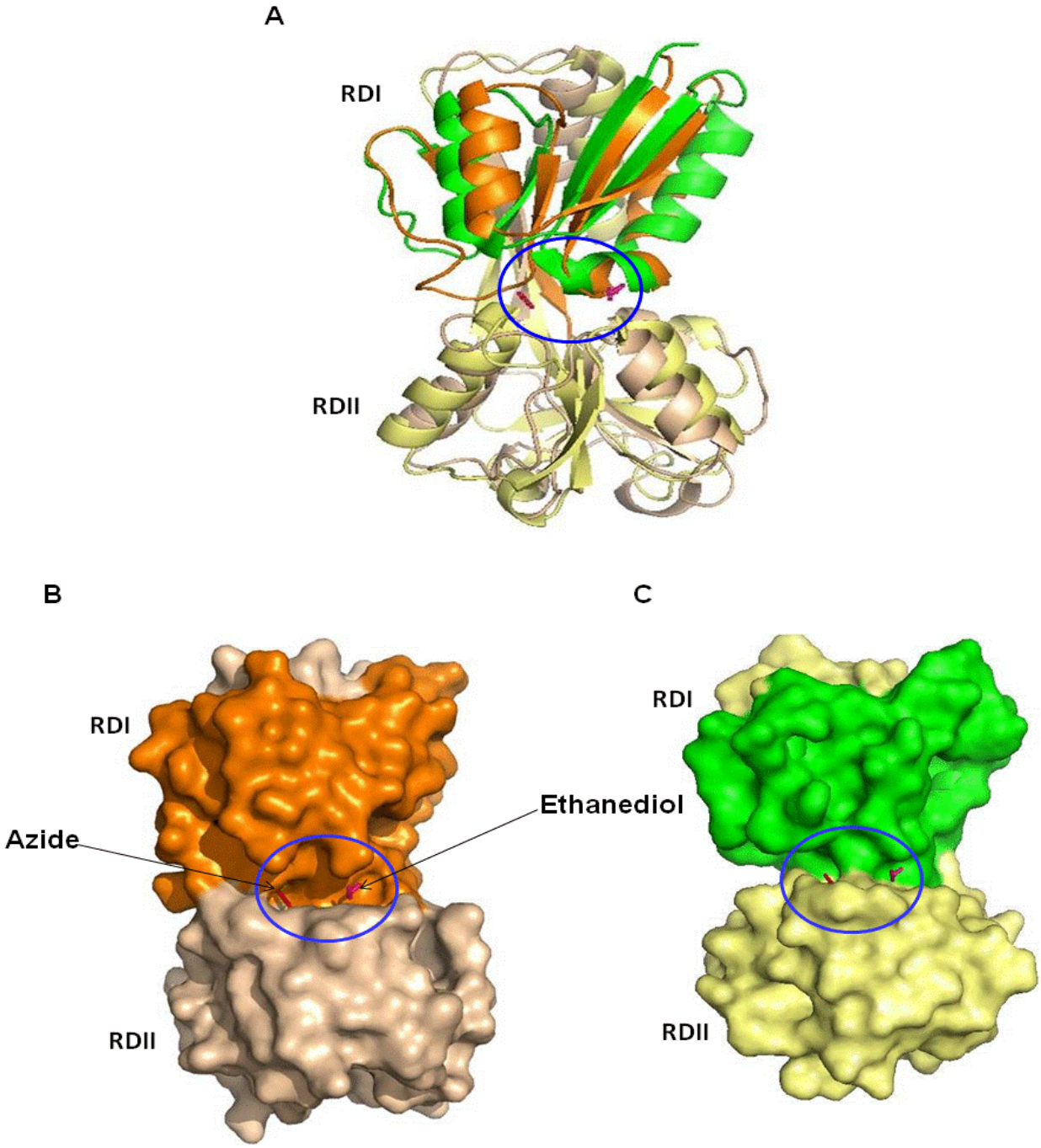


Figure 4-5. The effector binding domain of CynR (PDB # 3HFU) was aligned to the structural model of KaeR. A) Structural alignment of KaeR (RDI- orange and RDII- wheat color) and CynR (RDI-green, RDII-yellow orange). Surface representations of B) KaeR and C) CynR shows the accessibility of the ligand binding pockets. Azide is shown in red and ethanediol is indicated in pink. The ligand binding pocket is indicated by blue circles.

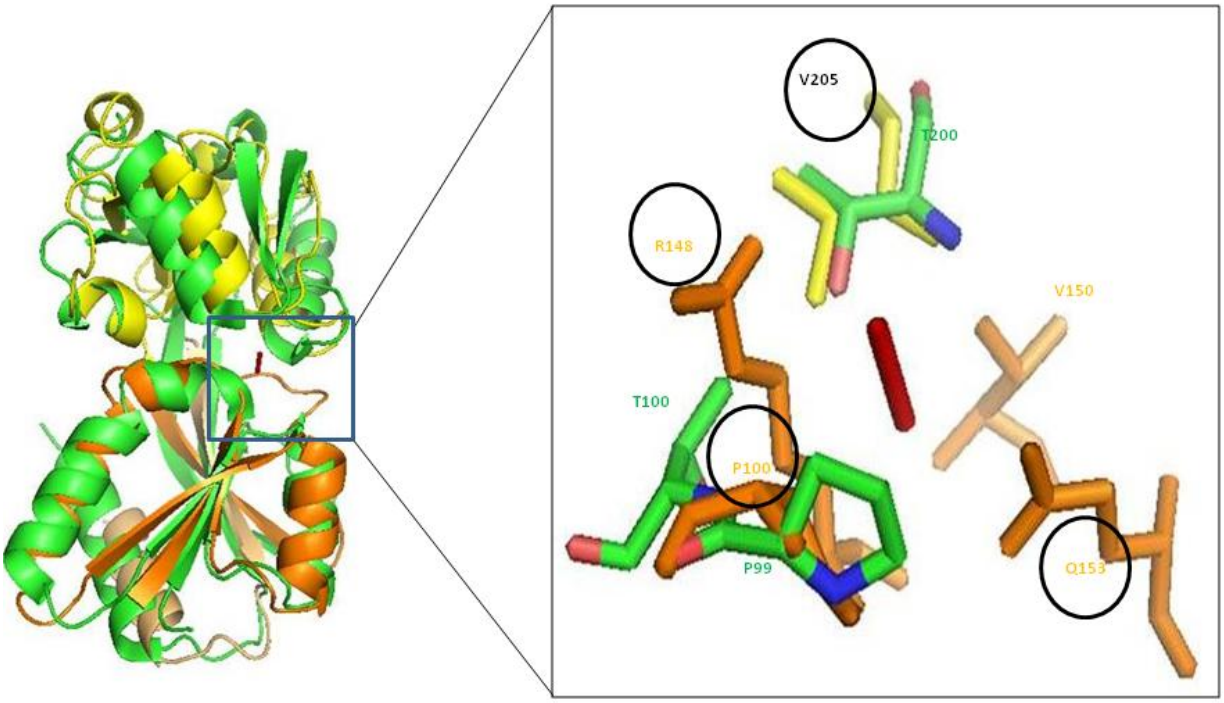


Figure 4-6. Analysis of the azide ligand binding pocket in CynR. The zoomed square depicts the residues in CynR (green) involved in azide (red) binding. The aligned residues in KaeR within a 4Å distance from azide are circled.

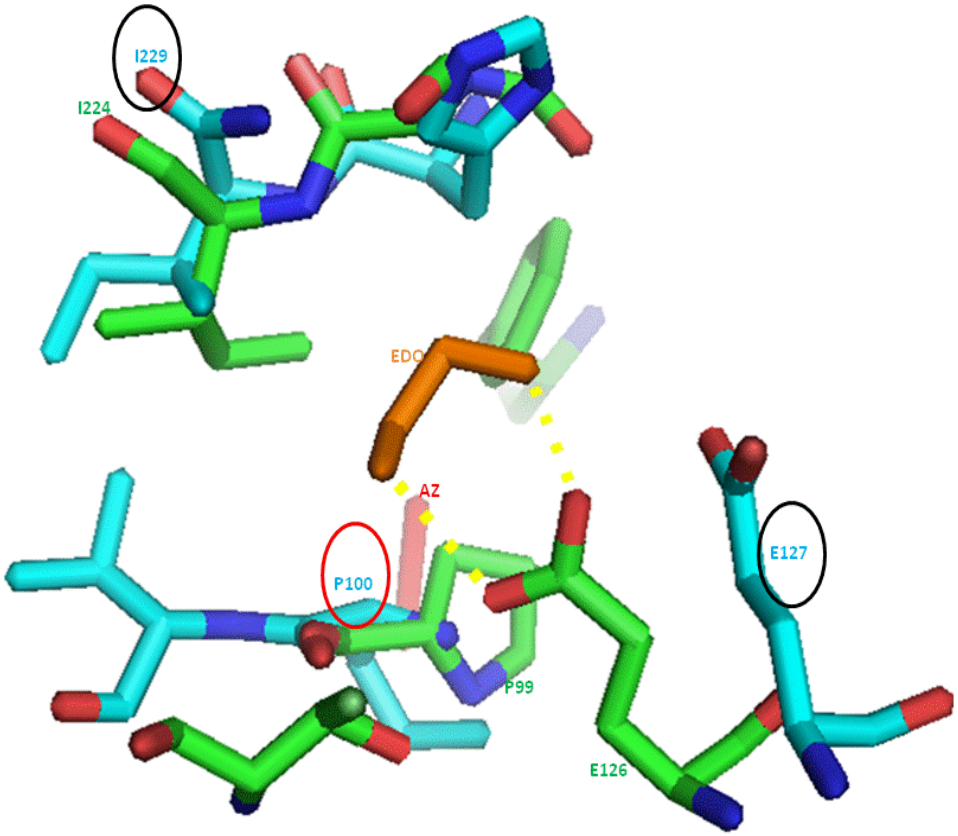


Figure 4-7. Residues involved in ethanediol binding. Structural alignment of CynR (green) and KaeR (cyan) EBDs. The corresponding residues in KaeR are circled. EDO: ethanediol, AZ: azide.

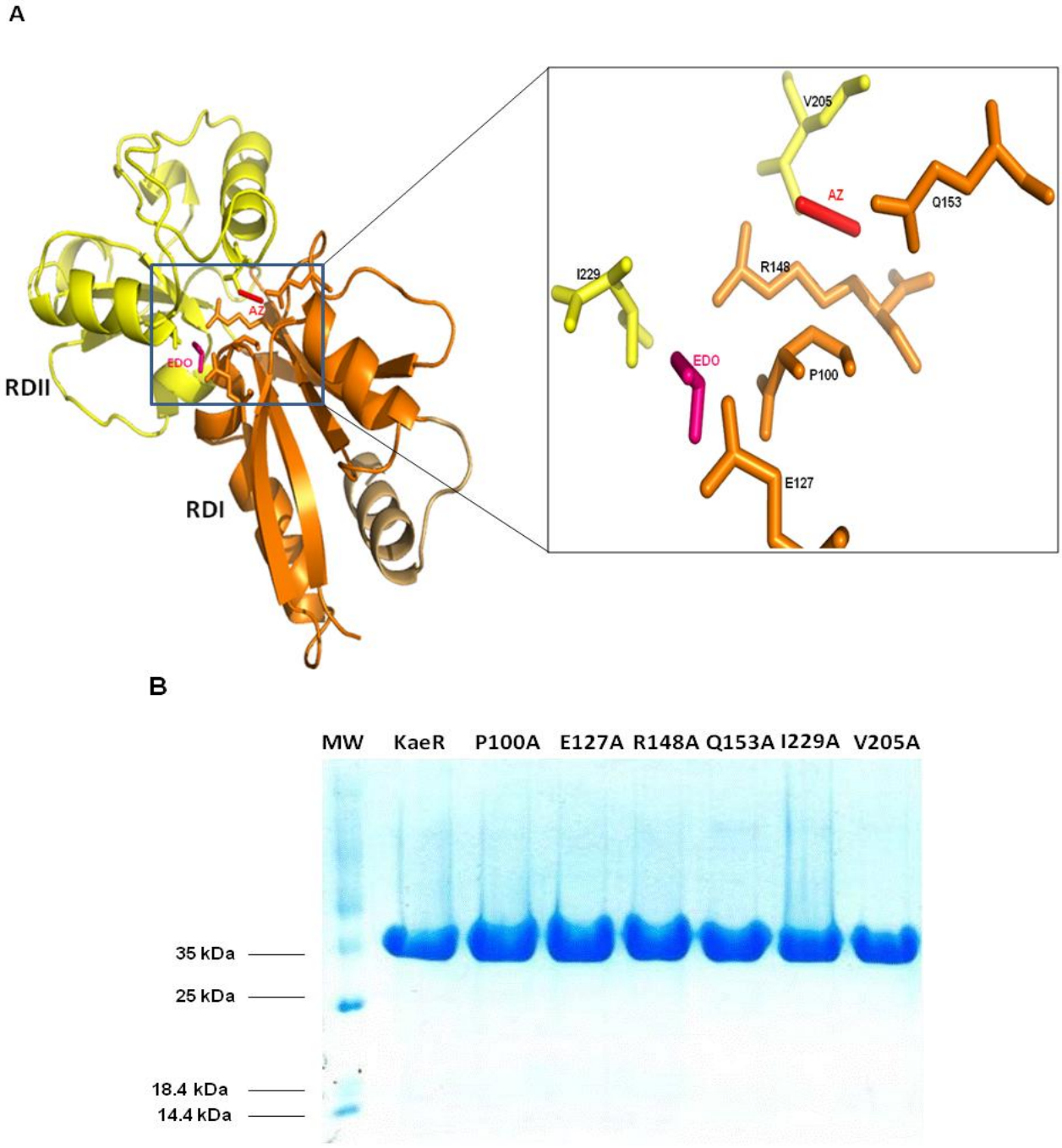


Figure 4-8. Summary of the putative KaeR residues involved in the ligand binding. A) Cartoon diagram of the RDs in KaeR showing the putative amino acids that contact the ligand (indicated as sticks). B) SDS-PAGE gel of the wild type and mutant forms of KaeR. MW- Molecular weight marker. EDO: ethanediol, AZ: azide.

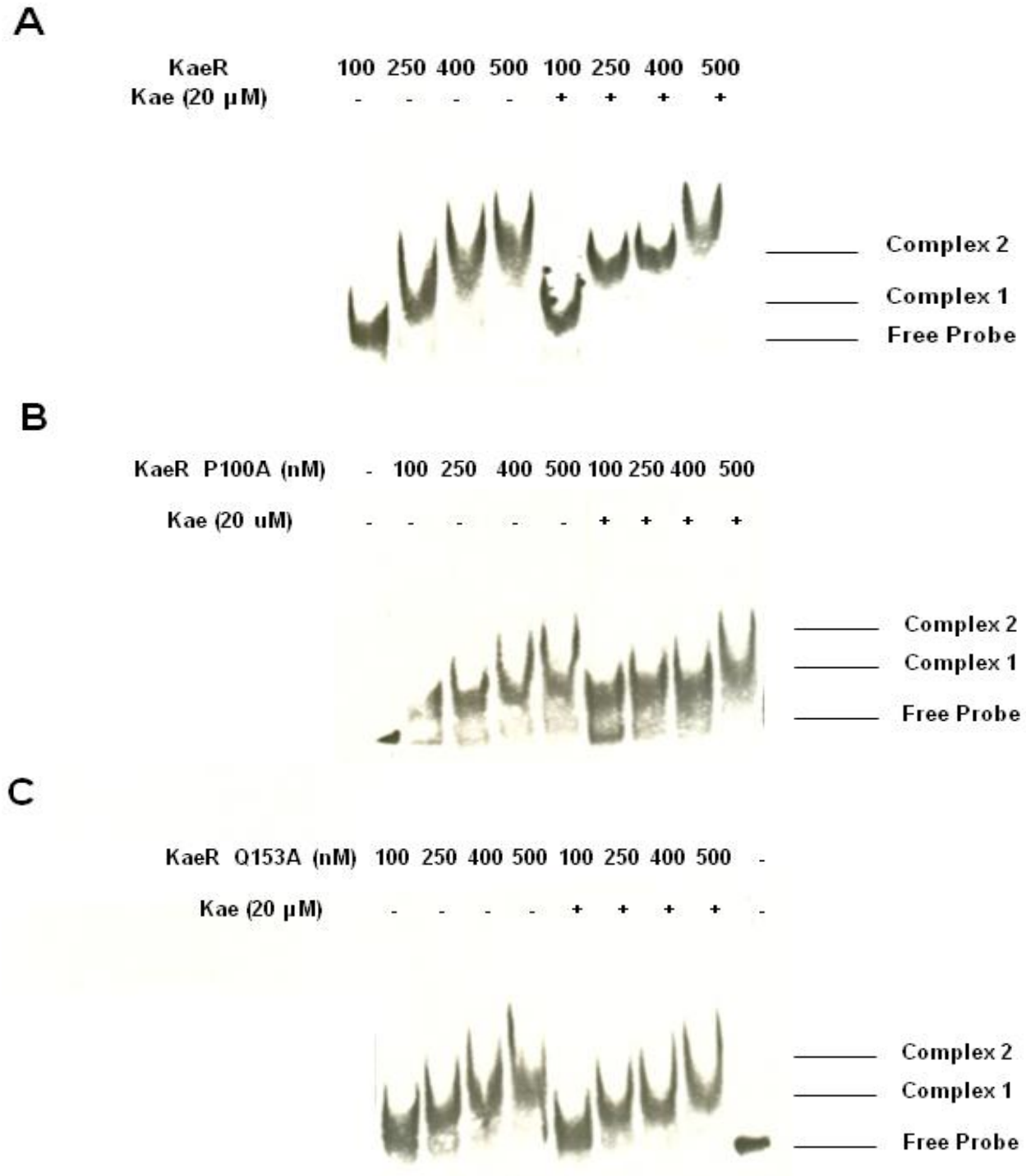
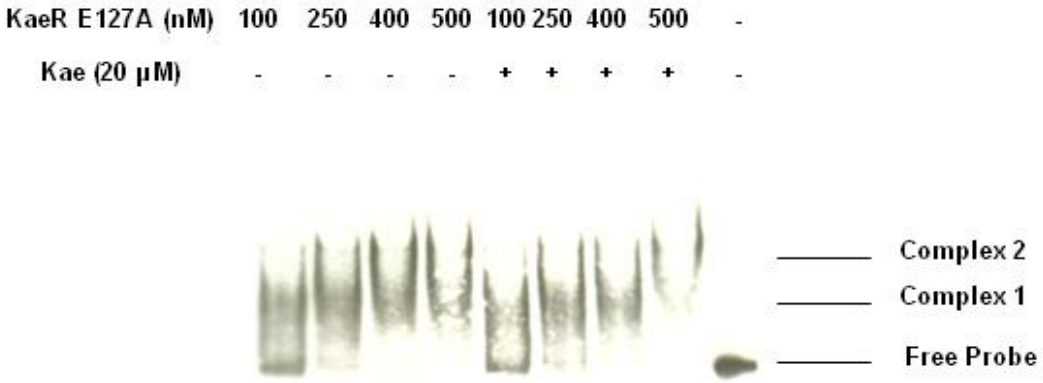
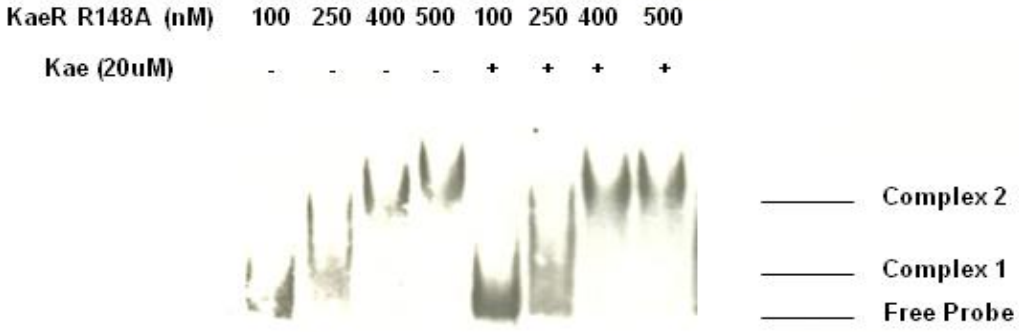


Figure 4-9. Electrophoretic mobility shift assays (EMSAs) to assess kaempferol effect on KaeR mutants. A) KaeR, B) P100A,C) Q153A D) E127A, E) R148A, F) V205A, and G) I229A.Kaempferol was dissolved in DMSO. F_{kae2} probe was used. kae- kaempferol.

D



E



F

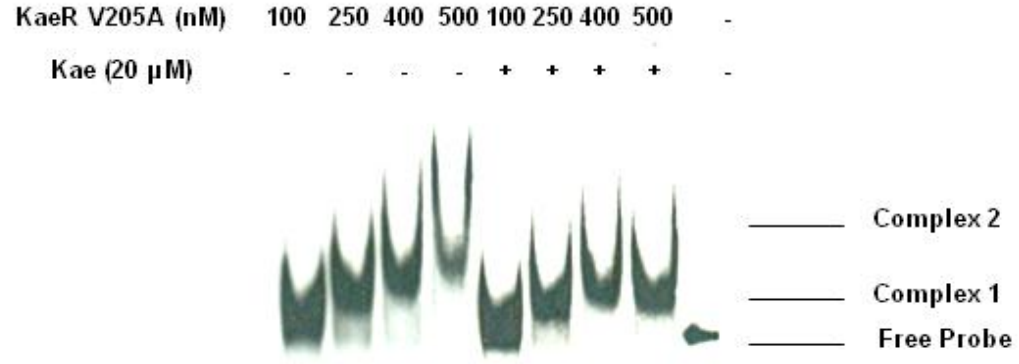


Figure 4-9. Continued

G

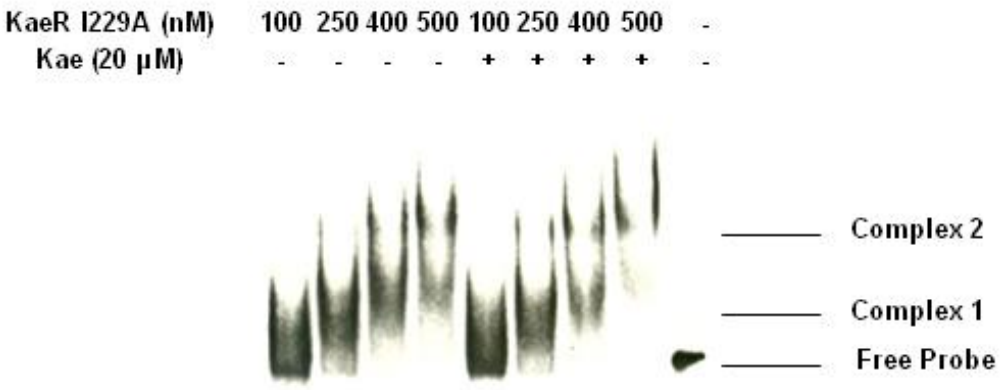


Figure 4-9. Continued

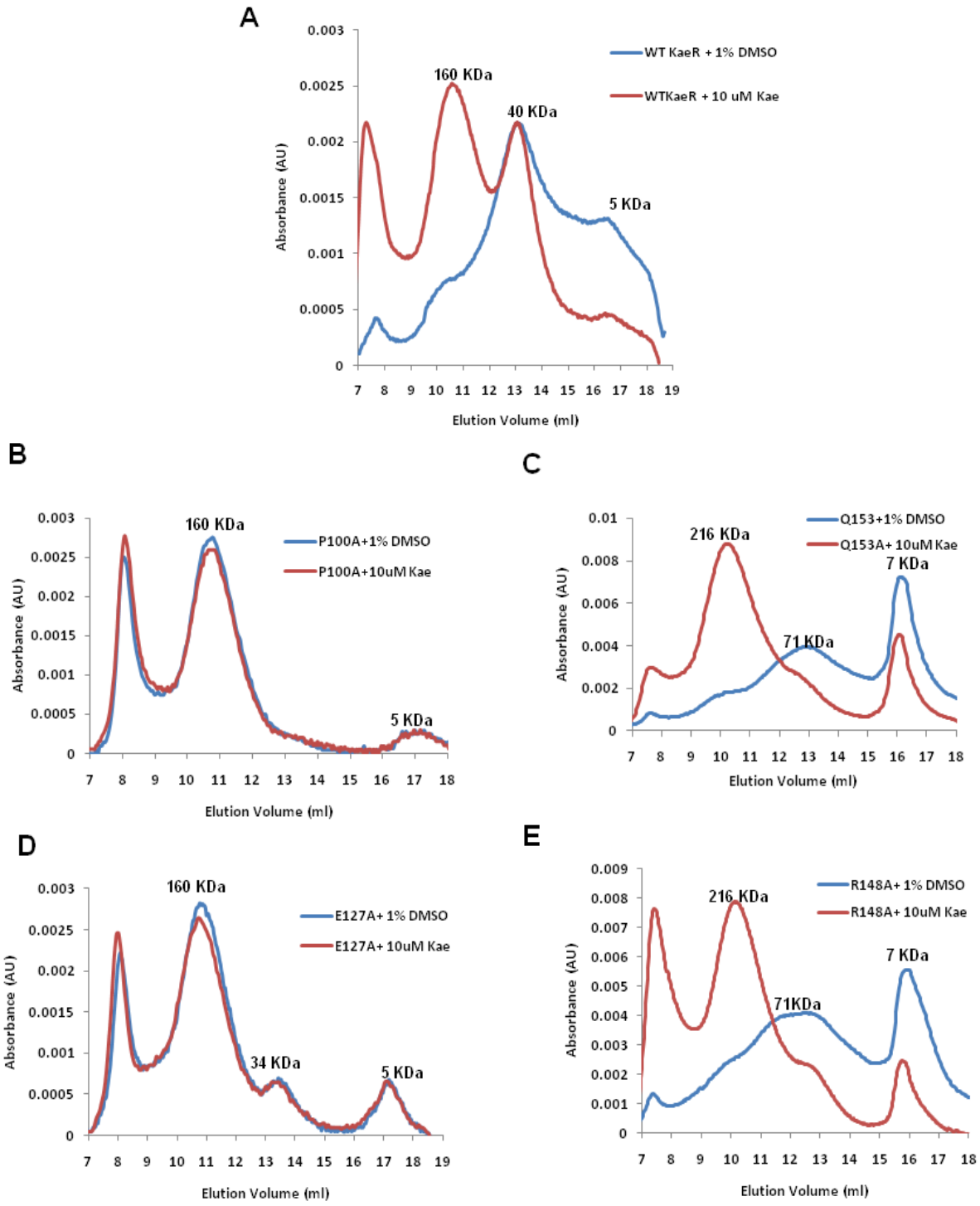
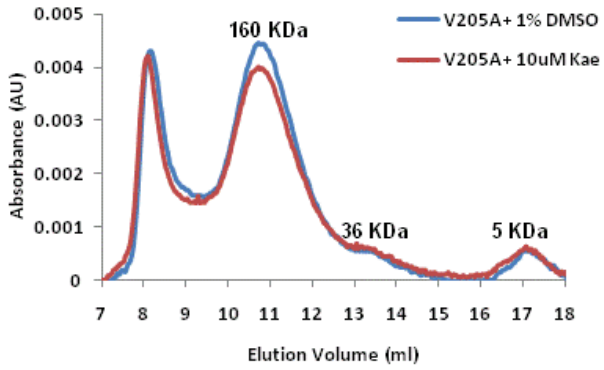


Figure 4-10. Size exclusion chromotography to assess kaempferol effect on KaeR mutants. Elution profiles for A) KaeR, B) P100A,C) Q153A D) E127A, E) R148A, F) V205A, and G) I229A with (red) and without Kaempferol (blue).kae- kaempferol.Kaempferol was dissolved in DMSO.

F



G

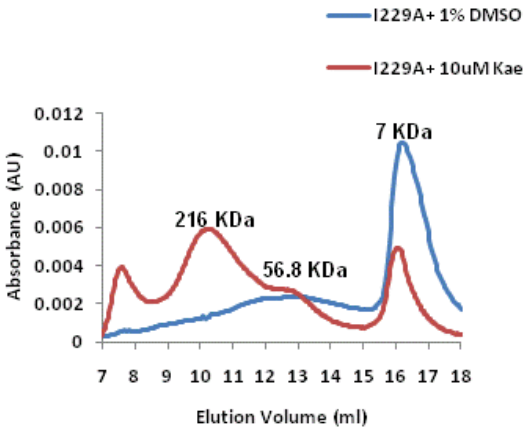


Figure 4-10. Continued

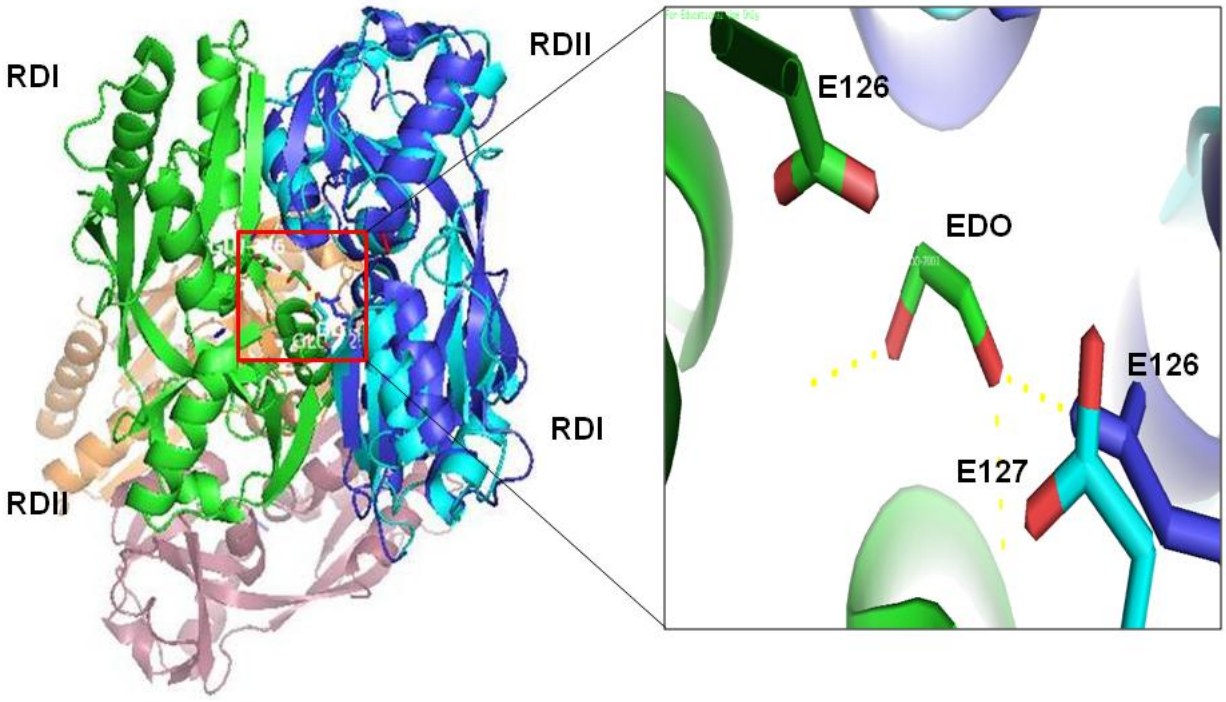


Figure 4-11. CynR dimerization domain within the CynR tetramer. The KaeR effector binding domain (Cyan) was aligned with the CynR effector binding domain (Blue). The CynR residue E126 (blue) is aligned with KaeR E127 (cyan). The aligned residues are indicated to the right in inset. In CynR, the E126 residues from two monomers (green and blue) might interact with ethanediol bringing the two monomers together. EDO-Ethanediol, RDI-Regulatory domain I, RDII-Regulatory domain II.

CHAPTER 5
SUMMARY AND CONCLUSIONS

Summary of Findings

Flavonoids are synthesized in plants and are shown to protect plants from adverse environmental conditions. Since *L. brevis* is found in a flavonoid rich environment, it was hypothesized that *L. brevis* may have the potential to respond to the flavonoids. The overall objective of this study was to characterize a transcription factor that responds to the flavonoids in *L. brevis*.

To identify transcription factors that may interact with flavonoids, an *in vitro* fluorescence based ligand screening assay was used. Only the protein LVIS1989 (KaeR), a LTTR, could interact with the flavonoids, myricetin, hesperetin and kaempferol. Bioinformatics studies suggested that the divergently encoded genes *LVIS1986*, *LVIS1987*, *LVIS1988* may be under the control of *kaeR* transcription regulation. The expression of *LVIS1986*, *LVIS1987*, *LVIS1988* genes in the presence of kaempferol was tested, and it was found that the genes are upregulated as a single transcriptional unit. The binding region of KaeR was located in the intergenic region between the *LVIS1988* and *LVIS1989 (kaeR)* genes. EMSA experiments using different concentrations of KaeR suggested that KaeR tetramerized cooperatively on DNA. Further, DNase I footprint studies identified two distinctive KaeR interaction sites, one in the intergenic region between *LVIS1988* and *LVIS1989* (-39 to +2) and another within *LVIS1988* (-314 to –353, from *kaeR* translational start point). Further experiments using EMSA showed that both binding sites are required for KaeR binding. The effect of flavonoids was tested on the KaeR: P_{kaeR} complex. It was found that kaempferol stabilized the KaeR: P_{kaeR} complex.

The effect of kaempferol on the oligomeric state of KaeR was tested by size exclusion chromatography. KaeR eluted as a monomer (observed MW = 40 kDa) while the addition of kaempferol promoted oligomerization and the formation of a tetramer. The regulatory implications of the KaeR binding sites were assessed *in vivo* using *B. subtilis* as a surrogate system. The intergenic region of *LVIS1988* and *kaeR* were transcriptionally fused to *lacZ*. The *in vivo* studies suggested that KaeR acts as an activator for *LVIS1988,* while *kaeR* is positively autoregulated. These results indicate that KaeR belongs to a small and poorly understood family of LTTRs that are positively autoregulated in the presence of a ligand.

Bioinformatic analysis was carried out to identify putative amino acids involved in ligand interaction. A structural model of KaeR was constructed and compared to another LTTR structure crystallized with small signal molecules. The analysis suggested that P100A, E127A, R148A, Q153A, V205A, I229A residues located in the hinge 3 region may be involved in kaempferol binding. Site-directed mutagenesis followed by size-exclusion chromatography experiments showed that all mutations modified the oligomeric state of the protein. EMSA experiments showed that mutations P100A, E127A, Q153A and I229A modified the ability to form complex 2 and to interact with kaempferol. Interestingly, although KaeR V205A was not able to form a stable complex 2, it was still able to interact with kaempferol in EMSA assays. These results indicate that the hinge 3 region in the protein is important in the formation dimers and tetramers as well as in the binding of the ligand.

In summary, I identified a new member of the LTTR family that is positively modulated by flavonoids. The results described are significant since those represent an

important step in the identification of food components that may modulate probiotic properties in *Lactobacillus* species.

Future Direction

Current efforts are directed to determine the molecular mechanism of KaeR positive autoregulation by flavonoids in *L. brevis*. To this end, the important nucleotides involved in KaeR: DNA binding are being analyzed by performing site directed mutagenesis in the promoter region of *LVIS1988* and *kaeR*. The result will allow in the identification of specific nucleotides involved in binding and those important for activation by kaempferol.The effect of KaeR protein mutations (P100A, E127A, R148A, Q153A, V205A, and I229A) on the regulation of both *kaeR and LVIS1988* will be analyzed by *lacZ* transcriptional fusion studies. It is expected that the results will uncover the mechanism of signal transduction upon binding of a ligand in LTTR family.

APPENDIX
OPTIMIZATION OF KAER BINDING CONDITIONS

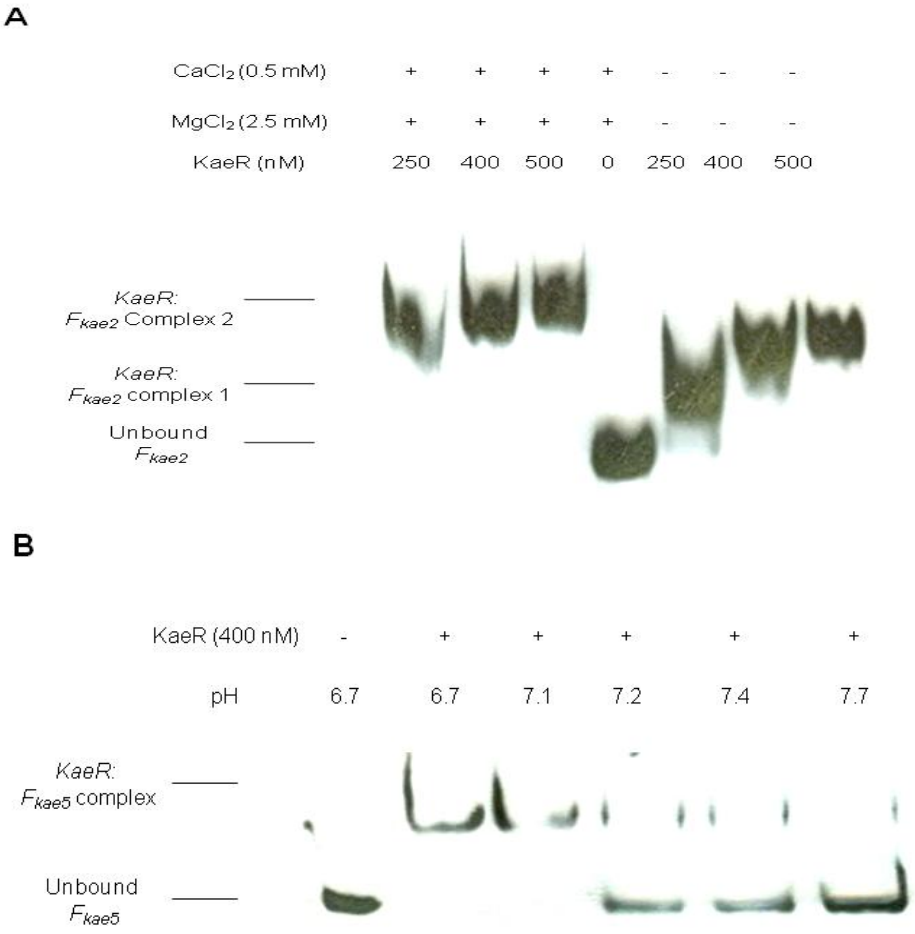


Figure A-1. Optimization of the KaeR binding conditions. Binding of KaeR to F_{kae2} was optimized by (A) adding 0.5 mM $CaCl_2$ and 2.5 mM $MgCl_2$ or (B) modifying the pH of the reaction mix as indicated on top of each panel.

LIST OF REFERENCES

Abe, A., and Kawahara, K. (1995) Transcriptional regulation and promoter sequence of the spvR gene of virulence plasmid pKDSC50 in *Salmonella choleraesuis* serovar Choleraesuis. *FEMS Microbiol Lett* **129**: 225-230.

Abe, A., Matsui, H., Danbara, H., Tanaka, K., Takahashi, H. and Kawahara, K. (1994) Regulation of spvR gene expression of Salmonella virulence plasmid pKDSC50 in Salmonella choleraesuis serovar Choleraesuis. *Mol Microbiol* **12**: 779-787.

Aura, A. M.(2005) In vitro digestion models for dietary phenolic compounds. Thesis, VTT, Espoo, Finland,

Axler-DiPerte, G. L., Miller, V. L., and Darwin, A. J. (2006) YtxR, a conserved LysR-like regulator that induces expression of genes encoding a putative ADP-ribosyltransferase toxin homologue in *Yersinia enterocolitica*. *J Bacteriol* **188**: 8033-8043.

Bais, H. P., Vepachedu, R., Gilroy, S., Callaway, R. M., and Vivanco, J. M. (2003) Allelopathy and exotic plant invasion: from molecules and genes to species interactions. *Science* **301**: 1377-1380.

Baliga, M. S. and Katiyar, S. K., (2006) Chemoprevention of photocarcinogenesis by selected dietary botanicals. *Photochem Photobiol Sci* **5**: 243-253.

Belitsky, B. R. and Sonenshein, A. L. (2002) GabR, a member of a novel protein family, regulates the utilization of γ-aminobutyrate in *Bacillus subtilis*. *Mol Microbiol* **45**: 569-583.

Benkert P, Biasini M, Schwede T. (2011). Toward the estimation of the absolute quality of individual protein structure models. *Bioinformatics*, **27**:343-50.

Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., Shindyalov, I. N., and Bourne, P. E. (2000) The Protein Data Bank. *Nucleic Acids Res* **28**: 235-242.

Bielinska, A., and Hulanicka, D. (1986) Regulation of the cysB gene expression in Escherichia coli. *Acta Biochim Pol* **33**: 133-137.

Bongaerts, J., Zoske, S., Weidner, U., and Unden, G. (1995) Transcriptional regulation of the proton translocating NADH dehydrogenase genes (nuoA-N) of Escherichia coli by electron acceptors, electron donors and gene regulators. *Mol Microbiol* **16**: 521-534.

Bowater, L., Fairhurst, S. A., Just, V. J., and Bornemann, S. (2004) *Bacillus subtilis* YxaG is a novel Fe-containing quercetin 2,3-dioxygenase. *FEBS Lett* **557**: 45-48.

Bradford, M. M. (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem* **72**: 248-254.

Braune, A., Gutschow, M., Engst, W., and Blaut, M. (2001) Degradation of quercetin and luteolin by *Eubacterium ramulus*. *Appl Environ Microbiol* **67**: 5558-5567.

Brencic, A., and Winans, S. C. (2005) Detection of and response to signals involved in host-microbe interactions by plant-associated bacteria. *Microbiol Mol Biol Rev* **69**: 155-194.

Broughton, W. J., and Perret, X. (1999) Genealogy of legume-*Rhizobium* symbioses. *Curr Opin Plant Biol* **2**: 305-311.

Brown, N. L., Stoyanov, J. V., Kidd, S. P., and Hobman, J. L. (2003) The MerR family of transcriptional regulators. *FEMS Microbiol Rev* **27**: 145-163.

Browning, D. F., and Busby, S. J. W. (2004) The regulation of bacterial transcription initiation. *Nat Rev Micro* **2**: 57-65.

Caldwell, A. L., and Gulig, P. A. (1991) The *Salmonella typhimurium* virulence plasmid encodes a positive regulator of a plasmid-encoded virulence gene. *J. Bacteriol.* **173**: 7176–7185.

Chang, M., and Crawford, I. P. (1991) In vitro determination of the effect of indoleglycerol phosphate on the interaction of purified TrpI protein with its DNA-binding sites. *J Bacteriol* **173**: 1590-1597.

Chen, C., Buchmeier, N., Libby, S., Fang, F., Krause, M. and Guiney, D. (1995) Central regulatory role for the RpoS sigma factor in expression of Salmonella dublin plasmid virulence genes. *J. Bacteriol.* **177**: 5303-5309.

Chen, X. C., Feng, J., Hou, B. H., Li, F. Q., Li, Q., and Hong, G. F. (2005) Modulating DNA bending affects NodD-mediated transcriptional control in *Rhizobium leguminosarum*. *Nucleic Acids Res* **33**: 2540-2548.

Choi, H.-J., Kim, S.-J., Mukhopadhyay, P., Cho, S., Woo, J.-R., Storz, G., and Ryu, S.-E. (2001) Structural Basis of the Redox Switch in the OxyR Transcription Factor. *Cell* **105**: 103-113.

Chugani, S.A., Parsek, M.R., Hershberger, C.D., Murakami, K., Ishihama, A., Chakrabarty, A.M. (1997) Activation of the *catBCA* promoter: probing the interaction of CatR and RNA polymerase through in vitro transcription. *J. Bacteriol* **179**: 2221–2227.

Chun, J., Kim, G. M., Lee, K. W., Choi, I. D., Kwon, G.-H., Park, J.-Y., Jeong, S.-J., Kim, J.-S., and Kim, J. H. (2007) Conversion of isoflavone glucosides to aglycones in soymilk by fermentation with Lactic Acid Bacteria. *J Food Sci* **72**: M39-M44.

Close, D. C., and McArthur, C. (2002) Rethinking the role of many plant phenolics: protection from photodamage Not herbivores? *Oikos* **99**: 166-172.

Cowan, J. M., Urbanowski, M. L., Talmi, M., and G. V. Stauffer, (1993) Regulation of the *Salmonella typhimurium metF* gene by the MetR protein. *J Bacteriol* **175**: 5862-5866.

Coynault, C., Robbe-Saule, V., Popoff M. Y., and Norel, F. (1992) Growth phase and SpvR regulation of transcription of *Salmonella typhimurium spvABC* virulence genes. *Microb Pathog* **13**: 133-143.

Craven, S. H., Ezezika, O. C., Haddad, S., Hall, R. A., Momany, C., and Neidle, E. L. (2009) Inducer responses of BenM, a LysR-type transcriptional regulator from *Acinetobacter baylyi* ADP1. *Mol Microbiol* **72**: 881-894.

Davis, E. O., and Johnston, A. W. B. (1990) Analysis of three nodD genes in *Rhizobium leguminosarum* biovar phaseoli; *nodD 1* is preceded by *nolE*, a gene whose product is secreted from the cytoplasm. *Mol Microbiol* **4**: 921-932.

Di Cagno, R., Mazzacane, F., Rizzello, C. G., Vincentini, O., Silano, M., Giuliani, G., De Angelis, M., and Gobbetti, M. (2010) Synthesis of isoflavone aglycones and equol in soy milks fermented by food-related lactic acid bacteria and their effect on human intestinal Caco-2 Cells. *J Agricult Food Chem* **58**: 10338-10346.

Downie, J. A., Knight, C. D., Johnston, A. W. B., and Rossen, L. (1985) Identification of genes and gene products involved in the nodulation of peas by *Rhizobium leguminosarum*. *MGG* **198**: 255-262.

Fang, F. C., Krause, M., Roudier, C., Fierer, J., and Guiney, D. G. (1991) Growth regulation of a *Salmonella* plasmid gene essential for virulence. *J Bacteriol* **173**: 6783-6789.

Finn, R. D., Mistry, J., Schuster-Böckler, B., Griffiths-Jones, S., Hollich, V., Lassmann, T., Moxon, S., Marshall, M., Khanna, A., Durbin, R., Eddy, S. R., Sonnhammer, E. L. L., and Bateman, A. (2006) Pfam: clans, web tools and services. *Nucleic Acids Res* **34**: D247-D251.

Firmin, J. L., Wilson, K. E., Rossen, L., and Johnston, A. W. B. (1986) Flavonoid activation of nodulation genes in *Rhizobium* reversed by other compounds present in plants. *Nature* **324**: 90-92.

Fisher, R. F., and Long, S. R. (1993) Interactions of NodD at the nod Box: NodD binds to two distinct sites on the same face of the helix and induces a bend in the DNA. *J Mol Biol* **233**: 336-348.

Fraga, C. G., Litterio, M. C., Prince, P. D., Calabró V., Piotrkowski B., Galleano M. (2011) Cocoa flavanols: effects on vascular nitric oxide and blood pressure. *J Clin Biochem Nutr* **48**: 63-67.

Fritsch, P.S., Urbanowski, M.L., Stauffer, G.V. (2000) Role of the RNA polymerase α subunits in MetR-dependent activation of *metE* and *metH*: Important residues in the C-terminal domain and orientation requirements within RNA polymerase. *J. Bacteriol.* **182**: 5539–5550.

Gao, J., and Gussin, G. N. (1991) Mutations in TrpI binding site II that differentially affect activation of the trpBA promoter of *Pseudomonas aeruginosa*. *EMBO J* **10**: 4137-4144.

Gobbetti, M., Cagno, R. D., and De Angelis, M. (2010) Functional Microorganisms for Functional Food Quality. *Crit Rev Food Sci Nutr* **50**: 716 - 727.

Griffiths, L. A., and Smith, G. E. (1972a) Metabolism of apigenin and related compounds in the rat. Metabolite formation in vivo and by the intestinal microflora in vitro. *Biochem J* **128**: 901-911.

Griffiths, L. A., and Smith, G. E. (1972b) Metabolism of myricetin and related compounds in the rat. Metabolite formation in vivo and by the intestinal microflora in vitro. *Biochem J* **130**: 141-151.

Grob, P., and Guiney, D. G. (1996) In vitro binding of the *Salmonella dublin* virulence plasmid regulatory protein SpvR to the promoter regions of *spvA* and *spvR*. *J Bacteriol* **178**: 1813-1820.

Grob, P., Kahn, D., and Guiney, D. G. (1997) Mutational characterization of promoter regions recognized by the Salmonella dublin virulence plasmid regulatory protein SpvR. *J Bacteriol* **179**: 5398-5406.

Guérout-Fleury, A.-M., Frandsen, N., and Stragier, P. (1996) Plasmids for ectopic integration in *Bacillus subtilis*. *Gene* **180**: 57-61.

Harris, S. J., Shih, Y.-L., Bentley, S. D., and Salmond, G. P. C. (1998) The *hexA* gene of *Erwinia carotovora* encodes a LysR homologue and regulates motility and the expression of multiple virulence determinants. *Mol Microbiol* **28**: 705-717.

He, B., and Zalkin, H. (1992) Repression of *Escherichia coli* purB is by a transcriptional roadblock mechanism. *J Bacteriol* **174**: 7121-7127.

Heldwein, E. E. Z. and Brennan, R. G. (2001) Crystal structure of the transcription activator BmrR bound to DNA and a drug. *Nature* **409**: 378-382.

Hein, E.-M., Rose, K., van't Slot, G., Friedrich , A. W.,and Humpf, H.-U. (2008) Deconjugation and degradation of flavonol glycosides by pig cecal microbiota characterized by fluorescence in situ hybridization (FISH). *J Agricult Food Chem* **56**: 2281-2290.

Hertog, M. G. L., Feskens, E. J. M., Kromhout, D., Hollman, P. C. H., and Katan, M. B. (1993) Dietary antioxidant flavonoids and risk of coronary heart disease: the Zutphen elderly study. *Lancet* **342**: 1007-1011.

Hirooka, K., Danjo, Y., Hanano, Y., Kunikane, S., Matsuoka, H., Tojo, S., and Fujita, Y. (2009) Regulation of the *Bacillus subtilis d*ivergent *yetL* and *yetM* genes by a transcriptional repressor, YetL, in response to flavonoids. *J Bacteriol* **191**: 3685-3697.

Hirooka, K., Kunikane, S., Matsuoka, H., Yoshida, K.-I., Kumamoto, K., Tojo, S., and Fujita, Y. (2007) Dual regulation of the *Bacillus subtilis* regulon comprising the *lmrAB* and *yxaGH* operons and *yxaF* gene by two transcriptional repressors, LmrA and YxaF, in response to flavonoids. *J Bacteriol* **189**: 5170-5182.

Hollman, P. C. H., and Katan, M. B. (1999) Dietary flavonoids: intake, health effects and bioavailability. *Food ChemToxicol* **37**: 937-942.

Hryniewicz, M. M., and Kredich, N. M. (1994) Stoichiometry of binding of CysB to the *cysJIH*, *cysK*, and *cysP* promoter regions of Salmonella typhimurium. *J Bacteriol* **176**: 3673-3682.

Hryniewicz, M. M., and Kredich, N. M. (1995) Hydroxyl radical footprints and half-site arrangements of binding sites for the CysB transcriptional activator of *Salmonella typhimurium*. *J Bacteriol* **177**: 2343-2353.

Härtl, B., Wehrl, W., Wiegert, T., Homuth, G., and Schumann, W. (2001) Development of a new integration site within the *Bacillus subtilis* chromosome and construction of compatible expression cassettes. *J Bacteriol* **183**: 2696-2699.

Jorgensen, C., and Dandanell, G. (1999) Isolation and characterization of mutations in the *Escherichia coli* regulatory protein XapR. *J. Bacteriol.* **181**: 4397-4403.

Jäger, A., and Saaby, L. (2011) Flavonoids and the CNS. *Molecules* **16**: 1471-1485.

Kandler, O., and Weiss, N. (1986) Regular, non-sporing Gram-positive rods, p. 1208–1234. In: Sneath, P.H.A., Mair, N., Sharpe, M.E., and Holt, J.G. (ed.), Bergey's manual of systematic bacteriology, vol. 2. William and Wilkins, Baltimore.

Kelley L A.,Sternberg M J. (2009) Protein structure prediction on the Web: a case study using the Phyre server. *Nature protocols* **4**: 363--371.

Knekt, P., Järvinen, R., Seppänen, R., Heliövaara, M., Teppo, L., Pukkala, E., and Aromaa, A. (1997) Dietary flavonoids and the risk of lung cancer and other malignant neoplasms. *Am J Epidemiol* **146**: 223-230.

Kowarz, L., Coynault, C., Robbe-Saule, V., and Norel, F. (1994) The *Salmonella typhimurium katF* (*rpoS*) gene: cloning, nucleotide sequence, and regulation of *spvR* and *spvABCD* virulence plasmid genes. *J Bacteriol* **176**: 6852-6860.

Krause, M., Fang, F. C., and Guiney, D. G. (1992) Regulation of plasmid virulence gene expression in *Salmonella dublin* involves an unusual operon structure. *J Bacteriol* **174**: 4482-4489.

Kredich, N. M., (1992) The molecular basis for positive regulation of *cys* promoters in *Salmonella typhimurium* and *Escherichia coli*. *Mol Microbiol* **6**: 2747-2753.

Krumholz, L. R., and Bryant, M. P. (1986) *Eubacterium oxidoreducens* sp. nov. requiring H_2 or formate to degrade gallate, pyrogallol, phloroglucinol and quercetin. *Arch Microbiol* **144**: 8-14.

Kummerfeld, S. K., and Teichmann, S. A. (2006) DBD: a transcription factor prediction database. *Nucleic Acids Research* **34**: D74-D81.

Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R, Thompson JD, Gibson TJ, Higgins DG. (2007). Clustal W and Clustal X version 2.0. *Bioinformatics* **23**: 2947-2948.

Leal-Sánchez, M. V., Ruiz-Barba, J. L., Sánchez, A. H., Rejano, L., Jiménez-Díaz, R., and Garrido, A. (2003) Fermentation profile and optimization of green olive fermentation using *Lactobacillus plantarum* LPCO10 as a starter culture. *Food Microbiol* **20**: 421-430.

Lebeer, S., Vanderleyden, J., and De Keersmaecker, S. C. (2008) Genes and Molecules of Lactobacilli Supporting Probiotic Action. *Microbiol Mol Biol Rev* **72**: 728-764.

Lehnen, D., Blumer, C., Polen, T., Wackwitz, B., Wendisch, V. F. and Unden, G. (2002) LrhA as a new transcriptional key regulator of flagella, motility and chemotaxis genes in *Escherichia coli*. *Mol Microbiol* **45**: 521-532.

Le Marchand, L., (2002) Cancer preventive effects of flavonoids-a review. *Biomed Pharmacother* **56**: 296-301.

Li, J., Ou-Lee, T. M., Raba, R., Amundson, R. G., and Last, R. L. (1993) Arabidopsis flavonoid mutants are hypersensitive to UV-B irradiation. *Plant Cell* **5**: 171-179.

Lochowska, A., Iwanicka-Nowicka, R., Plochocka, D., and Hryniewicz, M. M. (2001) Functional dissection of the LysR-type CysB Transcriptional Regulator. *J Biol Chem* **276**: 2098-2107.

Lorca, G. L., Ezersky, A., Lunin, V. V., Walker, J. R., Altamentova, S., Evdokimova, E., Vedadi, M., Bochkarev, A., and Savchenko, A. (2007) Glyoxylate and pyruvate are antagonistic effectors of the *Escherichia coli* IclR transcriptional regulator. *J Biol Chem* **282**: 16476-16491.

López de Felipe, F., Curiel, J. A., and Muñoz, R. (2010) Improvement of the fermentation performance of *Lactobacillus plantarum* by the flavanol catechin is uncoupled from its degradation. *J Appl Microbiol* **109**: 687-697.

Lu,D., Fillet,S., Meng,C., Alguel,Y., Kloppsteck,P., Bergeron,J., Krell,T., Gallegos,M.T., Ramos,J. and Zhang,X. (2010) Crystal structure of TtgV in complex with its DNA operator reveals a general model for cooperative DNA binding of tetrameric gene regulators. *Genes Dev* **24**: 2556–2565.

Maddocks, S. E., and Oyston, P. C. (2008) Structure and function of the LysR-type transcriptional regulator (LTTR) family proteins. *Microbiology* **154**: 3609-3623.

Mantena, S. K., Meeran, S. M., Elmets, C. A., and Katiyar, S. K. (2005) Orally administered green tea polyphenols prevent ultraviolet radiation-induced skin cancer in mice through activation of cytotoxic T cells and inhibition of angiogenesis in tumors. *J Nutr* **135**: 2871-2877.

Mathesius, U., Mulders, S., Gao, M., Teplitski, M., Caetano-Anollés, G., Rolfe, B. G., and Bauer, W. D. (2003) Extensive and specific responses of a eukaryote to bacterial quorum-sensing signals. *PNAS* **100**: 1444-1449.

Mavromatis K, Chu K, Ivanova N, Hooper SD, Markowitz VM, et al. (2009) Gene Context Analysis in the Integrated Microbial Genomes (IMG) Data Management System. *PLoS ONE* 4(11): e7979. doi:10.1371/journal.pone.0007979

Middleton, E., Kandaswami, C., and Theoharides, T. C. (2000) The effects of plant flavonoids on mammalian cells:implications for inflammation, heart disease, and cancer. *Pharmacol Rev* **52**: 673-751.

Miller, A.D.W., Ebright, R., and Rothman-Denes, L. (1997) RNA polymerase beta′ subunit: a target for DNA binding-independent activation. *Science* **275**: 1655–1657.

Miller, J.H. (1972). Experiments in Molecular Genetics. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York, 466 pp.

Minezaki, Y., Homma, K., and Nishikawa, K. (2006) Genome-wide survey of transcription factors in prokaryotes reveals many bacteria-specific families not found in archaea. *DNA Res* **12**: 269-280.

Molina-Henares, A. J., Krell, T., Eugenia Guazzaroni, M., Segura, A., and Ramos, J. L. (2006) Members of the IclR family of bacterial transcriptional regulators function as activators and/or repressors. *FEMS Microbiol Rev* **30**: 157-186.

Monferrer, D., Tralau, T., Kertesz, M. A., Dix, I., Sola, M., and Uson, I. (2010) Structural studies on the full-length LysR-type regulator TsaR from *Comamonas testosteroni* T-2 reveal a novel open conformation of the tetrameric LTTR fold. *Mol Microbiol* **75**: 1199-1214.

Moreno-Campuzano, S., Janga, S. C., and Perez-Rueda, E. (2006) Identification and analysis of DNA-binding transcription factors in *Bacillus subtilis* and other Firmicutes- a genomic approach. *BMC Genomics* **7**: 147.

Mulligan, J. T., andLong, S. R. (1989) A family of activator genes regulates expression of *Rhizobium meliloti* nodulation genes. *Genetics* **122**: 7-18.

Muraoka, S., Okumura, R., Ogawa, N., Nonaka, T., Miyashita, K., and Senda, T. (2003) Crystal structure of a full-length LysR-type transcriptional regulator, CbnR: unusual combination of two subunit forms and molecular bases for causing and changing DNA bend. *J Mol Biol* **328**: 555-566.

Müller-Hill, B. The Lac Operon: A Short History of a Genetic Paradigm. (Walter de Gruyter, New York, 1996).

Müller-Hill, B., (1998) Some repressors of bacterial transcription. *Curr Opin Microbiol* **1**: 145-151.

Niesen, F. H., Berglund, H., and Vedadi, M. (2007) The use of differential scanning fluorimetry to detect ligand interactions that promote protein stability. *Nat Protocols* **2**: 2212-2221.

Ostrowski, J., and Kredich, N. M. (1990) In vitro interactions of CysB protein with the *cysJIH* promoter of *Salmonella typhimurium*: inhibitory effects of sulfide. *J Bacteriol* **172**: 779-785.

Ostrowski, J., and Kredich, N.M. (1991) Negative autoregulation of *cysB* in *Salmonella typhimurium: in vitro* interactions of CysB protein with the *cysB* promoter. *J. Bacteriol* **173**, 2212–2218.

Pareja, E., Pareja-Tobes, P., Manrique, M., Pareja-Tobes, E., Bonal, J., and Tobes, R. (2006) ExtraTrain: a database of extragenic regions and transcriptional information in prokaryotic organisms. *BMC Microbiology* **6**: 29.

Parsek, M. R., Shinabarger, D. L., Rothmel, R. K., and Chakrabarty, A. M. (1992) Roles of CatR and cis,cis-muconate in activation of the *catBC* operon, which is involved in benzoate degradation in *Pseudomonas putida*. *J Bacteriol* **174**: 7798-7806.

Peters, N., Frost, J.,and Long, S. (1986) A plant flavone, luteolin, induces expression of *Rhizobium* meliloti nodulation genes. *Science* **233**: 977-980.

Petit-Glatron, M.-F., and Chambert, R. (1992) Peptide carrier potentiality of *Bacillus subtilis* levansucrase. *J Gen Microbiol* **138**: 1089-1095.

Pham, T. T., and Shah, N. P. (2009) Performance of Starter in Yogurt Supplemented with Soy Protein Isolate and Biotransformation of Isoflavones during Storage Period. *J Food Sci* **74**: M190-M195.

Porrúa, O., García-González, V., Santero, E., Shingler, V., and Govantes, F. (2009) Activation and repression of a σ^N-dependent promoter naturally lacking upstream activation sequences. *Mol Microbiol* **73**: 419-433.

Porrúa, O., García-Jaramillo, M., Santero, E., and Govantes, F. (2007) The LysR-type regulator AtzR binding site: DNA sequences involved in activation, repression and cyanuric acid-dependent repositioning. *Mol Microbiol* **66**: 410-427.

Porrúa, O., Platero, A. I., Santero, E., Del Solar, G., and Govantes, F. (2010) Complex interplay between the LysR-type regulator AtzR and its binding site mediates *atzDEF* activation in response to two distinct signals. *Mol Microbiol* **76**: 331-347.

Puupponen-Pimiä, R., Nohynek, L., Meier, C., Kähkönen, M., Heinonen, M., Hopia, A. and Oksman-Caldentey, K. M. (2001) Antimicrobial properties of phenolic compounds from berries. *J Appl Microbiol* **90**: 494-507.

Pérez-Rueda, E., and Collado-Vides, J. (2000) The repertoire of DNA-binding transcriptional regulators in *Escherichia coli* K-12. *Nucleic Acids Res* **28**: 1838-1847.

Rechner, A. R., Smith, M. A., Kuhnle, G., Gibson, G. R., Debnam, E. S., Srai, S. K., Moore, K. P., and Rice-Evans, C. A. (2004) Colonic metabolism of dietary polyphenols: influence of structure on microbial fermentation products. *Free Radic Biol Med* **36**: 212-225.

Redmond, J. W., Batley, M., Djordjevic, M. A., Innes, R. W., Kuempel, P. L., and Rolfe, B. G. (1986) Flavones induce expression of nodulation genes in *Rhizobium*. *Nature* **323**: 632-635.

Rice-Evans, C., Miller, N., and Paganga, G. (1997) Antioxidant properties of phenolic compounds. *Trends Plant Sci* **2**: 152-159.

Rodionov, D. A. (2007) Comparative genomic reconstruction of transcriptional regulatory networks in bacteria. *Chem Rev* **107**: 3467-3497.

Rodríguez, H., Curiel, J. A., Landete, J. M., de las Rivas, B., de Felipe, F. L., Gómez-Cordovés, C.,. Mancheño, J. M, and Muñoz, R. (2009) Food phenolics and lactic acid bacteria. *Int J Food Microbiol* **132**: 79-90.

Rosario, C. J., Frisch, R. L., and Bender, R. A. (2010) The LysR-type nitrogen assimilation control protein forms complexes with both long and short DNA binding sites in the absence of coeffectors. *J Bacteriol* **192**: 4827-4833.

Rossen, L., Shearman, C. A., Johnston, A. W. B., and Downie, J. A. (1985) The *nodD* gene of *Rhizobium leguminosarum* is autoregulatory and in the presence of plant exudate induces the *nodA,B,C* genes. *EMBO J* **4**: 3369-3373.

Sainsbury, S., Lane, L. A., Ren, J., Gilbert, R. J., Saunders, N. J., Robinson, C. V., Stuart, D. I., and Owens, R. J. (2009) The structure of CrgA from *Neisseria meningitidis* reveals a new octameric assembly state for LysR transcriptional regulators. *Nucleic Acids Res* **37**: 4545-4558.

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

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

Schell, M. A., Brown, P. H., and Raju, S. (1990) Use of saturation mutagenesis to localize probable functional domains in the NahR protein, a LysR-type transcription activator. *J Biol Chem* **265**: 3844-3850.

Schleif, R. (2003) AraC protein: A love–hate relationship. *BioEssays* **25**: 274-282.

Schneider, H., Schwiertz, A., Collins, M. D. and Blaut, M. (1999) Anaerobic transformation of quercetin-3-glucoside by bacteria from the human intestinal tract. *Arch Microbiol* **171**: 81-91-91.

Schoefer, L., Mohan, R., Braune, A., Birringer, M., and Blaut, M. (2002) Anaerobic C-ring cleavage of genistein and daidzein by *Eubacterium ramulus*. *FEMS Microbiol Lett* **208**: 197-202.

Selma, M. V., Espín, J. C., and Tomás-Barberán, F. A. (2009) Interaction between phenolics and gut microbiota: role in human health. *J Agricult Food Chem* **57**: 6485-6501.

Shin, M., Kang, S., Hyun, S.-J., Fujita, N., Ishihama, A., Valentin-Hansen, P., and Choy, H. E. (2001) Repression of *deoP2* in Escherichia coli by CytR: conversion of a transcription activator into a repressor. *EMBO J* **20**: 5392-5399.

Singer, A. U., Evdokimova, E., Kagan, O., Cuff, M.E., Edwards, A.M., Joachimiak, A., Savchenko, A. Crystal structure of the ligand binding domain of *E. coli* CynR with its specific effector azide.

Spencer, J. P. (2003) Metabolism of Tea Flavonoids in the Gastrointestinal Tract. *J Nutr* **133**: 3255S-3261S.

Stec, E., Witkowska-Zimny, M., Hryniewicz, M. M., Neumann, P., Wilkinson, A. J., Brzozowski, A. M., Verma, C. S., Zaim, J., Wysocki, S., and Bujacz, G. D. (2006) Structural basis of the sulphate starvation response in *E. coli* crystal structure and mutational analysis of the cofactor-binding domain of the Cbl transcriptional regulator. *J Mol Biol* **364**: 309-322.

Swanson, J. A., Mulligan, J. T., and Long, S. R. (1993) Regulation of *syrM* and *nodD3* in *Rhizobium meliloti*. *Genetics* **134**: 435-444.

Tattini, M., Galardi, C., Pinelli, P., Massai, R., Remorini, D., and Agati, G. (2004) Differential accumulation of flavonoids and hydroxycinnamates in leaves of *Ligustrum vulgare* under excess light and drought stress. *New Phytol* **163**: 547-561.

Teran, W., Felipe, A., Segura, A., Rojas, A., Ramos, J.-L., and. Gallegos, M.-T (2003) Antibiotic-dependent induction of *Pseudomonas putida* DOT-T1E TtgABC efflux pump is mediated by the drug binding repressor TtgR. *Antimicrob Agents Chemother* **47**: 3067-3072.

Teran, W., Krell, T. , Ramos, J. L., and Gallegos, M. T. (2006) Effector-repressor interactions, binding of a single effector molecule to the operator-bound TtgR homodimer mediates derepression. *J Biol Chem* **281**: 7102-7109.

Toledano, M. B., Kullik, I., Trinh, F., Baird, P. T., Schneider, T. D., 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.

Tyrrell, R., Verschueren, K. H. G. , Dodson, E. J., Murshudov, G. N., Addy, C., and Wilkinson, A. J. (1997) The structure of the cofactor-binding fragment of the LysR family member, CysB: a familiar fold with a surprising subunit arrangement. *Structure* **5**: 1017-1032.

Tzounis, X., Vulevic, J., Kuhnle, G. G., George, T., Leonczak, J., Gibson, G. R., Kwik-Uribe, C., and Spencer, J. P. (2008) Flavanol monomer-induced changes to the human faecal microflora. *Br J Nutr* **99**: 782-792.

Ulrich, L. E., Koonin, E. V., and Zhulin, I. B. (2005) One-component systems dominate signal transduction in prokaryotes. *Trends Microbiol* **13**: 52-56.

van Rhijn, P., and Vanderleyden, J. (1995) The *Rhizobium*-plant symbiosis. *Microbiol Rev* **59**: 124-142.

Vandeputte, O. M., Kiendrebeogo, M., Rajaonson, S., Diallo, B., Mol, A., El Jaziri, M., and Baucher, M. (2010) Identification of catechin as one of the flavonoids from *Combretum albiflorum* bark extract that reduces the production of quorum-sensing-controlled virulence factors in *Pseudomonas aeruginosa* PAO1. *Appl Environ Microbiol* **76**: 243-253.

Vedadi, M., Niesen, F. H., Allali-Hassani, A., Fedorov, O. Y., Finerty, P. J., Wasney, G. A., Yeung, R., Arrowsmith, C., Ball, L. J., Berglund, H., Hui, R., Marsden, B. D., Nordlund, P., Sundstrom, M., Weigelt, J., and Edwards, A. M. (2006) Chemical screening methods to identify ligands that promote protein stability, protein crystallization, and structure determination. *PNAS* **103**: 15835-15840.

Verzelloni, E., Pellacani, C., Tagliazucchi, D., Tagliaferri, S., Calani, L., Costa, L. G., Brighenti, F., Borges, G., Crozier, A., Conte, A., and Del Rio, D. (2011) Antiglycative and neuroprotective activity of colon-derived polyphenol catabolites. *Mol Nutr Food Res* **55**: S35-S43.

Viswanathan, P., Murphy, K., Julien, B., Garza, A. G., and Kroos, L. (2007a) Regulation of *dev*, an operon that includes genes essential for *Myxococcus xanthus* development and CRISPR-associated genes and repeats. *J Bacteriol* **189**: 3738-3750.

Viswanathan, P., Ueki, T., Inouye, S., and Kroos, L. (2007b) Combinatorial regulation of genes essential for *Myxococcus xanthus* development involves a response regulator and a LysR-type regulator. *PNAS* **104**: 7969-7974.

Vitaglione, P., Donnarumma, G., Napolitano, A., Galvano, F., Gallo, A., Scalfi, L., and Fogliano, V. (2007) Protocatechuic acid is the major human metabolite of cyanidin-glucosides. *J Nutr* **137**: 2043-2048.

Wang, S. P., and Stacey, G. (1991) Studies of the *Bradyrhizobium japonicum nodD1* promoter: a repeated structure for the nod box. *J Bacteriol* **173**: 3356-3365.

Wang, Z.-Y., Huang, M.-T., Ferraro, T., Wong, C.-Q., Lou, Y.-R., Reuhl, K., Iatropoulos, M., Yang, C. S., and Conney, A. H. (1992) Inhibitory Effect of Green Tea in the Drinking Water on Tumorigenesis by Ultraviolet Light and 12-O-Tetradecanoylphorbol-13-acetate in the Skin of SKH-1 Mice. *Cancer Res* **52**: 1162-1170.

Warren L. DeLano "The PyMOL Molecular Graphics System."DeLano Scientific LLC, San Carlos, CA, USA. http://www.pymol.org

Weber, H., Polen, T., Heuveling, J., Wendisch, V.F., andHengge, R. (2005) Genome-wide analysis of the general stress response network in *Escherichia coli*: sigmaS-dependent genes, promoters, and sigma factor selectivity.*J Bacteriol* **187:** 1591–1603.

Wilson, D., Madera, M., Vogel, C., Chothia, C., and Gough, J. (2007) The SUPERFAMILY database in 2007: families and functions. *Nucleic Acids Res* **35**: D308-D313.

Wilson, R. L., Urbanowski, M. L., and Stauffer, G. V. (1995) DNA binding sites of the LysR-type regulator GcvA in the *gcv* and *gcvA* control regions of *Escherichia coli*. *J Bacteriol* **177**: 4940-4946.

Winter, J., Moore, L. H., Dowell, V. R., and Bokkenheuser, V. D. (1989) C-ring cleavage of flavonoids by human intestinal bacteria. *Appl Environ Microbiol* **55**: 1203-1208.

Zheng, D., Constantinidou, C., Hobman, J. L., and Minchin, S. D. (2004) Identification of the CRP regulon using in vitro and in vivo transcriptional profiling. *Nucleic Acids Res* **32**: 5874-5893.

Zhou, X., Lou, Z., Fu, S., Yang, A., Shen, H., Li, Z., Feng, Y., Bartlam, M., Wang, H., and Rao, Z. (2010) Crystal structure of ArgP from *Mycobacterium tuberculosis* confirms two distinct conformations of full-length LysR transcriptional regulators and reveals its function in DNA binding and transcriptional regulation. *J Mol Biol* **396**: 1012-1024.

Zianni, M., Tessanne, K., Merighi, M., Laguna, R., and Tabita, F. (2006) Identification of the DNA Bases of a DNase I footprint by the use of dye primer sequencing on an automated capillary DNA analysis instrument. *J Biomol Tech* **17**: 103-113.

BIOGRAPHICAL SKETCH

Santosh Gurunathrao Pande received his Bachelor of Engineering (B.E.) in chemical engineering from North Maharashtra University, Jalgaon, India in May 2002 and Master of Technology (M. Tech.) in biotechnology from Anna University, Guindy, Chennai, India in June 2006. He was a recipient of a two year fellowship from the Department of Biotechnology (DBT), Government of India to pursue M. Tech. program (July 2004 to June 2006). Subsequently, he was selected through campus placement to work for Dr. Reddy's Laboratories Ltd., Hyderabad, India as a trainee (July 2006 to June 2007). In August 2007, he was admitted to the graduate program at the University of Florida, Gainesville where he was a recipient of the graduate alumni fellowship (August 2007 to August 2011). In January of 2008, he joined Dr. Graciela Lorca's laboratory where he was exposed to many fascinating aspects of bacterial transcriptional gene regulation. In particular, he worked on the characterization of the LysR family transcriptional factor from *L. brevis*. During his PhD program, Santosh attended different conferences and meetings including the American Society for Microbiology (ASM) 3rd conference on beneficial microbes (October 25-29, 2010, Miami, Florida).

In December 2008, he married Rupali Morkhande. Upon completion of his PhD in August 2011, Santosh plans to pursue research training in academia (Postdoctoral training) after which he wants to return to India to pursue research career in industry/ academia.