

Biochemical characterization of a temperature responsive two-component regulatory system from the Antarctic methanogen, *Methanococcoides burtonii*

Author:

Najnin, Tahria

Publication Date:

2016

DOI:

<https://doi.org/10.26190/unsworks/18963>

License:

<https://creativecommons.org/licenses/by-nc-nd/3.0/au/>

Link to license to see what you are allowed to do with this resource.

Downloaded from <http://hdl.handle.net/1959.4/56049> in <https://unsworks.unsw.edu.au> on 2024-04-30

**Biochemical characterization of a temperature
responsive two-component regulatory system
from the Antarctic methanogen,
*Methanococcoides burtonii***

Tahria Najnin

Thesis submitted in fulfilment of the requirements for the degree of
Doctor of Philosophy (Ph.D.)



School of Biotechnology and Biomolecular Sciences
University of New South Wales
Sydney, Australia

March 2016

PLEASE TYPE

THE UNIVERSITY OF NEW SOUTH WALES
Thesis/Dissertation Sheet

Surname or Family name: **Najnin**

First name: **Tahria**

Other name/s:

Abbreviation for degree as given in the University calendar: **PhD**

School: **Biotechnology and Biomolecular Sciences**

Faculty: **Faculty of Science**

Title: **Biochemical characterization of a temperature responsive two-component regulatory system from the Antarctic methanogen, *Methanococcoides burtonii*.**

Abstract 350 words maximum: (PLEASE TYPE)

Temperature is a critical environmental factor that influences microbial growth. A few studies have shown that two component regulatory systems (TCSs) in bacteria may regulate temperature-dependent gene expression. TCSs typically consist of a sensor histidine kinase (SK) that responds to an extracellular signal by phosphorylating a cytoplasmic response regulator (RR) that in turn regulates gene expression. Although TCSs are present in *Archaea* they are poorly studied.

Methanococcoides burtonii is a psychrophilic methanogen, isolated from cold (1-2 °C) methane saturated anaerobic bottom waters of Ace Lake, Antarctica. The genome sequence contains one RR, LtrR, which contains a DNA binding output domain and the gene forms an operon with a SK, LtrK. Comparative proteomic and transcriptomic analyses of *M. burtonii* identified higher abundance of LtrR and/or LtrK at low (4°C) compared to high (23 °C) temperature suggesting that this TCS may be involved in regulating global gene expression in response to growth temperature.

The aim of this dissertation was to study the activity and stability properties of this TCS and determine whether it exhibits temperature responsive activity. To achieve this, a truncated version of LtrK lacking its N-terminal transmembrane domains was purified. The recombinant protein was able to autophosphorylate and perform phosphotransfer to LtrR as well as dephosphorylate LtrR. Analysis of the effect of temperature on LtrK found high activity at 0 °C, optimal activity at 10 °C (half-life ~ 2.8 days), reduced activity at 15 °C, and very low activity at 30 °C (half-life ~24 min). The analyses of thermal stability of LtrK showed that LtrK unfolds irreversibly with unfolding occurring at low temperatures implying that LtrK is thermolabile. The fact that the protein is intrinsically unstable and performs phosphotransfer reactions optimally at low temperatures suggests that LtrK is temperature responsive and that its cytoplasmic domain is the 'temperature sensing' module. The temperature-responsive properties of LtrK are likely to be important physiologically as *M. burtonii* is capable of growth from -2.5 to 29 °C; growth characteristics that align well with the biochemical properties of LtrK. These findings shed light on a temperature sensing TCS in a psychrophile for the first time.

Declaration relating to disposition of project thesis/dissertation

I hereby grant to the University of New South Wales or its agents the right to archive and to make available my thesis or dissertation in whole or in part in the University libraries in all forms of media, now or here after known, subject to the provisions of the Copyright Act 1968. I retain all property rights, such as patent rights. I also retain the right to use in future works (such as articles or books) all or part of this thesis or dissertation.

I also authorise University Microfilms to use the 350 word abstract of my thesis in Dissertation Abstracts International (this is applicable to doctoral theses only).

Signature

Witness Signature

Date

The University recognises that there may be exceptional circumstances requiring restrictions on copying or conditions on use. Requests for restriction for a period of up to 2 years must be made in writing. Requests for a longer period of restriction may be considered in exceptional circumstances and require the approval of the Dean of Graduate Research.

FOR OFFICE USE ONLY

Date of completion of requirements for Award:

Originality Statement

'I, Tahria Najnin, hereby declare that this submission is my own work and to the best of my knowledge it contains no materials previously published or written by another person, or substantial proportions of material which have been accepted for the award of any other degree or diploma at UNSW or any other educational institution, except where due acknowledgement is made in the thesis. Any contribution made to the research by others, with whom I have worked at UNSW or elsewhere, is explicitly acknowledged in the thesis. I also declare that the intellectual content of this thesis is the product of my own work, except to the extent that assistance from others in the project's design and conception or in style, presentation and linguistic expression is acknowledged.'

COPYRIGHT STATEMENT

'I hereby grant the University of New South Wales or its agents the right to archive and to make available my thesis or dissertation in whole or part in the University libraries in all forms of media, now or here after known, subject to the provisions of the Copyright Act 1968. I retain all proprietary rights, such as patent rights. I also retain the right to use in future works (such as articles or books) all or part of this thesis or dissertation.

I also authorise University Microfilms to use the 350 word abstract of my thesis in Dissertation Abstract International (this is applicable to doctoral theses only).

I have either used no substantial portions of copyright material in my thesis or I have obtained permission to use copyright material; where permission has not been granted I have applied/will apply for a partial restriction of the digital copy of my thesis or dissertation.'

AUTHENTICITY STATEMENT

'I certify that the Library deposit digital copy is a direct equivalent of the final officially approved version of my thesis. No emendation of content has occurred and if there are any minor variations in formatting, they are the result of the conversion to digital format.'

Abstract

Temperature is a critical environmental factor that influences microbial growth and biogeochemical cycling on the planet. Despite the importance of environmental microorganisms to global ecosystem health, little is known about their molecular mechanisms of adaptation to the cold, including how they sense temperature. Two component regulatory systems (TCSs) are well characterized systems that enable bacteria to sense environmental signals and regulate gene expression accordingly. TCSs typically consist of a sensor histidine kinase protein that responds to an extracellular signal by phosphorylating a cytoplasmic response regulator that in turn regulates gene expression. Although TCSs are extensively studied in *Bacteria* they are also abundant in *Archaea* especially in the genomes of psychrophilic methanogens, but they remain poorly studied.

Methanococcoides butonii is a psychrophilic methanogen, isolated from cold (1-2 °C) methane saturated anaerobic bottom waters of Ace Lake, Antarctica. From genome sequence analyses, a total of 45 TCSs were identified in *M. burtonii*. One response regulator, LtrR was found to have a DNA binding output domain and form an operon with a predicted sensor kinase, LtrK. Comparative proteomic and transcriptomic analyses of *M. burtonii* identified higher abundance of the TCS at low (4°C) growth temperature compared to high (23 °C) temperature suggesting that it may be involved in regulating global gene expression in response to growth temperature.

The biochemical characterization of the LtrK/LtrR TCS was performed to determine its activity and function. To achieve this, a truncated version of LtrK (lacking its N-terminal transmembrane domains) was purified and found to possess both kinase and phosphatase activities. LtrK can autophosphorylate itself, transfer the phosphoryl group to LtrR, and dephosphorylate LtrR that was previously phosphorylated. Point mutation studies of histidine residues in LtrK showed that H367 is the site of phosphorylation and two other histidine residues (H443 and H448) play significant roles in the autophosphorylation reaction. Mutation analyses of conserved aspartate residues in LtrR revealed D98 as a possible phosphoryl group receiver residue.

Analyses of the effects of temperature on LtrK activity demonstrated that it had highest kinase and phosphatase (on phosphorylated LtrR) activity at 10 °C, and remained very stable at this temperature (half-life of inactivation of ~2.8 days). The

activity was also remarkably high at 0 °C (~50% of maximum) and rapidly fell away at temperatures above 15 °C with little activity at 30 °C (half-life of inactivation, ~24 min). Analysis of phosphotransfer activity of LtrK on LtrR also showed very high activity of LtrK at 0°C compared to 25°C. The thermal stability of LtrK was examined using differential scanning calorimetry which showed that LtrK unfolds irreversibly with unfolding occurring at very low temperatures (*i.e.* ~20 °C), consistent with its low temperature shifted activity profile. The fact that the protein is intrinsically unstable and performs phosphorylation reactions optimally at low temperatures suggests that LtrK is temperature-responsive and that its cytoplasmic domain is the ‘temperature sensing’ module. The temperature-responsive properties of LtrK aligns well with the growth profile of *M. burtonii* that grows at 1-2 °C in the environment and is capable of growth from -2.5 to 29 °C in the laboratory. Therefore, the temperature-dependent properties of LtrK are likely to have a physiological significance in *M. burtonii*. The thermodynamic activation parameters of LtrK stability and activity revealed that the enzyme has kinetic properties typical of a psychrophilic enzyme. The analyses demonstrated that LtrK has evolved properties to achieve high catalytic activity at low temperatures and thermolability by reducing enthalpy of activation. This is only the second biochemical study for a TCS from *Archaea*, and the first for any psychrophile, and provides insight into the properties of a temperature responsive TCS in *Archaea*.

List of Publications

Journal Articles

Najnin T., Siddiqui K.S., Taha, Elkaid N., Kornfeld G., Curmi P.M.G. and Cavicchioli R. (2016) Characterization of a temperature-responsive two component regulatory system from the Antarctic archaeon, *Methanococcoides burtonii*. Scientific Reports, 6: 24278.

Taha, Siddiqui K. S., Campanaro S., **Najnin T.**, Deshpande N., Williams T.J., Aldrich-Wright J., Wilkins M., Curmi P. M. G., and Cavicchioli R. (2016) Single TRAM domain RNA-binding proteins in *Archaea*: functional insight from Ctr3 from the Antarctic methanogen *Methanococcoides burtonii*. Environmental microbiology doi:10.1111/1462-2920.13229.

Conference Abstract

Najnin T., Siddiqui K.S., Curmi PMG., Cavicchioli R. (2014) Characterization of a temperature responsive two-component regulatory system in the Antarctic methanogen, *Methanococcoides burtonii*. Protein Science, 23: 109-110.

Acknowledgements

First of all I am grateful to The Almighty Allah for giving me strength and patience to deal with the Ph.D. stress and complete this project.

I would like to express my sincere thanks to my supervisor, Prof. Rick Cavicchioli for providing me with all the necessary facilities. Without his support, guidance, valuable suggestions and editorial assistance I would not be able to complete this project.

I would especially like to thank Prof. Paul Curmi for his continuous support and valuable suggestions during the project.

I am very grateful to Dr. Khawar Sohail Siddiqui for his enormous support and valuable suggestions. I will never forget your assistance.

Very special thanks to Geoff and Ned for assisting me in the radiation facility and making always time for me when I was in trouble.

I also thank Dr. Jodi Richards and Sohaila for guiding me in the lab at the beginning of my candidature. Thanks to Tim for the useful suggestions and advice. Very special thanks to Taha for the useful discussions and moral support. Thanks to rest of the lab members – Benny, Yan, Peter who were wonderful friends to work with.

My heartiest thanks to my husband- Mokarrom, without his support and encouragement I would not be able to go through this journey. And last but not least, I am extremely grateful to my parents for inspiring confidence in me and sacrificing their comfort for my bright future.

Contents

Declaration	ii
Abstract	iii
List of Publications	v
Acknowledgements	vi
List of Abbreviations	xiii
1 Introduction	
1.1 <i>Methanococcoides burtonii</i>	2
1.1.1 Genomic and proteomic studies on <i>M. burtonii</i>	4
1.1.2 Specific protein studies	7
1.1.2.1 Elongation factor 2 (EF-2)	8
1.1.2.2 RNA polymerase subunit E/F (MbRpoE/F)	8
1.1.2.3 Chaperonin	9
1.1.2.4 Cold responsive TRAM domain protein (Ctr)	10
1.2 Two component system (TCS)	10
1.2.1 An overview of TCS	10
1.2.2 Structure and function of SKs	14
1.2.3 Structure and function of RRs	21
1.2.4 TCS as thermosensor	28
1.2.4.1 Mechanism of temperature signal detection and activation	33
1.2.5 TCS in <i>Archaea</i>	38
1.3 TCS in <i>M. burtonii</i>	39
1.4 Aim of the project	42
1.5 References	44
2 Bioinformatics analyses of the sensor kinase, LtrK and response regulator, LtrR	
2.1 Abstract	54
2.2 Introduction	55

2.3	Experimental procedures	57
2.4	Results	59
2.4.1	Protein sequence analyses of LtrK and LtrR	59
2.4.2	Domains annotation in LtrK	60
2.4.3	Conserved sequence motifs in LtrK	62
2.4.4	Analyses of His residues – the possible phosphorylation site in LtrK ...	64
2.4.5	Domains in LtrR	66
2.4.6	Analyses on Asp residues- the possible phosphorylation site in LtrR ..	68
2.4.7	3D structural homology modelling	69
2.4.7.1	3D structural homology model of LtrK	69
2.4.7.2	3D structural homology model of LtrR	72
2.4.8	Phylogenetic relationship analyses	73
2.5	Discussion	76
2.6	References	80

3 Purification and biophysical characterization of the sensor kinase, LtrK

3.1	Abstract	85
3.2	Introduction	86
3.3	Experimental procedures	90
3.3.1	Cloning	90
3.3.2	<i>E. coli</i> competent cells and transformation	90
3.3.3	Overexpression of recombinant LtrK	91
3.3.4	Purification of recombinant LtrK protein	92
3.3.5	GST tag cleavage from purified GST-LtrK	93
3.3.6	SDS-PAGE analyses	93
3.3.7	Protein identification	94
3.3.8	Native molecular weight determination by gel filtration chromatography.	94
3.3.9	Circular dichroism (CD) spectroscopy analyses	95
3.3.10	Differential scanning calorimetry	96
3.4	Results	97

3.4.1	Expression and purification of LtrK as soluble protein	97
3.4.2	Evaluation of oligomeric state showed LtrK to be a dimer	99
3.4.3	Circular dichroism spectroscopy	100
3.4.4	Differential Scanning Calorimetry (DSC)	101
3.5	Discussion	103
3.6	References	107

4 Purification and biophysical characterization of the response regulator, LtrR

4.1	Abstract	111
4.2	Introduction	112
4.3	Experimental procedures	114
4.3.1	Cloning	114
4.3.2	Overexpression and purification of recombinant LtrR	115
4.3.3	Oligomeric state determination by gel filtration chromatography	117
4.3.4	Intrinsic tyrosine fluorescence analysis.....	117
4.3.5	Circular dichroism (CD) analyses	118
4.3.5.1	Far-ultraviolet (UV) CD spectra	118
4.3.5.2	Near- UV CD spectra	118
4.3.6	Differential scanning calorimetry (DSC)	118
4.3.7	DNA binding assay of LtrR.....	119
4.3.7.1	Growth of <i>M. burtonii</i> culture and genomic DNA extraction	119
4.3.7.2	Fragmentation of whole genomic DNA	120
4.3.7.3	LtrR phosphorylation	121
4.3.7.4	Overexpression and purification of N-terminal HTH domain deleted LtrR*	121
4.3.7.5	Binding of whole genomic DNA with phosphorylated LtrR	122
4.4	Results	123
4.4.1	Expression and purification of soluble LtrR	123
4.4.2	Evaluation of oligomeric state showed LtrR to be a monomer.....	125
4.4.3	Protein folding study using intrinsic tyrosine fluorescence analyses..	126
4.4.4	Analyses of LtrR secondary structure using CD spectroscopy.....	127

4.4.5	Stability analysis of the purified LtrR using DSC.....	129
4.4.6	DNA binding analyses using <i>M. burtonii</i> genomic DNA.....	131
4.5	Discussion	132
4.6	References	137

5 Analyses of phosphorylation activities of LtrK and LtrR

5.1	Abstract	141
5.2	Introduction	142
5.3	Experimental procedures	144
5.3.1	Autophosphorylation assay using Pro-Q Diamond phosphoprotein gel stain	144
5.3.2	Native-PAGE analyses	145
5.3.3	Isothermal titration calorimetry (ITC)	146
5.3.4	Autophosphorylation assay using radioactive [γ - ³² P]-ATP	147
5.3.5	Phosphotransfer assay using radioactive [γ - ³² P]-ATP	148
5.3.6	Dephosphorylation assays of LtrR-P	149
5.3.7	Site directed mutagenesis of LtrK and LtrR genes and purification of the mutant proteins	149
5.4	Results	151
5.4.1	Autophosphorylation activity of LtrK identified by Pro-Q diamond staining protocol.....	151
5.4.2	Analyses of phosphotransfer reaction between LtrK and LtrR using ITC 153	
5.4.3	Autophosphorylation and phosphotransfer assays using radioactive [γ - ³² P]-ATP.....	155
5.4.4	Dephosphorylation of LtrR in the absence and presence of LtrK	158
5.4.5	Effect of histidine residue mutations on LtrK phosphorylation activities 159	
5.4.6	Effect of aspartate residue mutations on LtrR phosphorylation	161
5.5	Discussion	162
5.6	References	168

6 Analyses of temperature-responsive activity of LtrK

6.1 Abstract	173
6.2 Introduction	174
6.3 Experimental Procedures	176
6.3.1 Kinase, phosphotransfer and phosphatase activities analyses at various temperatures.....	176
6.3.2 Half-life of thermal inactivation assay	177
6.3.3 Differential scanning calorimetry (DSC) at various scan rates	178
6.4 Results	179
6.4.1 Kinase and phosphatase activity of LtrK is high at low temperature .	179
6.4.2 LtrK is more stable at 10 °C compared to 30 °C	181
6.4.3 The thermal unfolding of LtrK is scan-rate dependent	182
6.5 Discussion	184
6.6 References	190

7 Activation thermodynamics of LtrK stability and activity

7.1 Abstract	196
7.2 Introduction	197
7.3 Experimental procedures	198
7.3.1 Thermodynamic activation parameters for LtrK stability.....	198
7.3.2 Kinetic parameters for LtrK autophosphorylation	199
7.3.3 Thermodynamic activation parameters for LtrK autophosphorylation.....	200
7.4 Results	201
7.4.1 Activation Thermodynamics of LtrK stability.....	201
7.4.2 Thermodynamics of LtrK activity	202
7.4.2.1 Kinetic parameters for LtrK autophosphorylation	202
7.4.2.2 Activation energy for LtrK autophosphorylation reaction	203
7.5 Discussion	204
7.6 References	209

8 Conclusions and Future Work

8.1 General discussion	212
8.1.1 Unique organization of HTH domain in the DNA-binding RRs of methanogens	212
8.1.2 Presence of periplasmic CHASE domain in LtrK	213
8.1.3 Plausible role of membrane in regulating the function of transmembrane LtrK	214
8.1.4 LtrK is the third only biochemically studied thermosensing SK	215
8.2 Future work	217
8.2.1 Identification of genes regulated by LtrK/LtrR TCS	217
8.2.2 Analyses of the regulation of LtrK/LtrR in <i>M. burtonii</i>	219
8.2.3 Possible biotechnological application of LtrK	220
8.3 Concluding remarks	221
8.4 References	221

Appendix1 Protein sequences used for phylogenetic tree construction 224

Appendix2 Plasmid sequences 233

Appendix 3 Primer sequences and PCR cycling parameters 237

List of Abbreviations

³² P	Radioactive isotope of phosphorus
[γ 32 -P]ATP	Adenosine triphosphate with a 32-phosphate radioactive isotope in the gamma phosphate position
%	Percent
°C	Degrees Celsius
3'	3 prime DNA end
5'	5 prime DNA end
μ g	Microgram
μ L	Microliter
μ M	Micromolar
×	Times
APS	Ammonium persulfate
ATP	Adenosine triphosphate
bp	Base pairs
BLAST	Basic local alignment and search tool
BSA	Bovine serum albumin
CD	Circular dichroism
ChIP	Chromatin immunoprecipitation
Csp	Cold-shock protein
Ctr	Cold responsive TRAM domain protein
DNA	Deoxyribonucleic acid
DTT	Dithiothreitol
EDTA	Ethylene diamine-tetraacetic acid
EF2	Elongation factor 2
EMSA	Electrophoretic mobility shift assay
h	Hour
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
His ₆	Hexa-histidine tag
HTH	Helix-turn-helix motif
IMG	Integrated Microbial Genome database
IPTG	Isopropyl β -D-1-thiogalactopyranoside
K _{av}	Partition coefficient
kb	Kilo-base pairs
kd	Dissociation constant

KDa	Kilo-dalton
L	Liter
LB	Luria-Bertani
LtrK	Low temperature-responsive sensor kinase
LtrR	Low temperature-responsive response regulator
m	Meter
M	Molar
mg	Milligram
min	Minute
mL	Milliliter
mM	Millimolar
mRNA	Messenger RNA
MRW	Mean residue weight
ng	Nanogram
nm	Nanometer
OD	Optical density
PCR	Polymerase chain reaction
PDB	Protein database
pmol	Picomole
RpoE/F	RNA polymerase subunit E/F
RR	Response regulator
s	Second
SDS-PAGE	Sodium dodecyl polyacrylamide gel electrophoresis
SEC	Size-exclusion chromatography
SK	Sensor kinase
TCEP	Tris(2-carboxyethyl)phosphine
TCS	Two-component system
TEMED	Tetramethylethylenediamine
TMD	Transmembrane domain
UV	Ultraviolet
V ₀	Void volume
v/v	Volume per volume
WT	Wild type
w/v	Weight per volume

Chapter 1

Introduction

Archaea represent a domain of life that was initially classified as bacteria and recognized as a separate domain only a few decades ago (Woese *et al.*, 1990; Cavicchioli, 2011). The rapidly increasing knowledge and understanding has made it clear that archaea coexist with bacteria and eukaryotes in both aquatic and terrestrial environments (Cavicchioli, 2011). However, the largest proportion and greatest diversity of archaea live in cold environment, *i.e.* alpine and polar habitat, deep ocean, terrestrial and ocean surface (Cavicchioli, 2006). The function of the psychrophilic archaea in the global biogeochemical cycle is diverse, including a critical role in the nitrification of soil and ocean water mediated by ammonia oxidation (Leininger *et al.*, 2006; Cavicchioli *et al.*, 2007) and the cycling of simple carbon compound via methanogenesis and anaerobic oxidation of methane (Cavicchioli, 2006). Though psychrophilic archaea represent a significant fraction of the cold biomass they have long been understudied. Advances in understanding the biology of psychrophilic archaea have primarily occurred through the studies on Antarctic methanogen, *Methanococcoides burtonii* (Cavicchioli, 2006) - the first formally characterized archaeal psychrophile (Franzmann *et al.*, 1992). The complete genome of *M. burtonii* had been sequenced and the genomic data assisted to establish a range of molecular characteristics of cold adaptation in psychrophiles (Allen *et al.*, 2009). In addition, comparative proteomic and transcriptomic data (at 4 vs. 23 °C) analyses identified the genes, including the genes of two component system (TCS) that may play a critical role in the mechanism of cold adaptation (Goodchild *et al.*, 2004b; Campanaro *et al.*, 2011). TCSs are well characterized regulatory systems in microorganisms that enable the cell to sense environmental signals and regulate gene expression accordingly. TCSs are also

reported to be involved in sensing change in environmental temperature (Smirnova *et al.*, 2001; Klinkert and Narberhaus, 2009; Sengupta and Garrity, 2013).

This chapter reviews the features (*i.e.* habitat, growth condition) of the psychrophilic methanogen *M. burtonii* with a brief description of the currently available genomics, proteomics and transcriptomics data. The proteins that are characterized in *M. burtonii* to date are also briefly described. As the theme of the thesis is focused on TCS in *M. burtonii*, emphasis is given to describe the common features of TCSs and molecular basis of thermosensing TCSs.

1.1 *Methanococoides burtonii*

M. burtonii is the first formally characterized psychrophilic methanogen that was isolated from anaerobic, methane and hydrogen sulphide rich depths of Ace Lake, which is located in the Vestfold Hills region, close to Davis Station, in the Australian Antarctic Territory (Franzmann *et al.*, 1992). Ace Lake is a saline meromictic lake that has a maximum reported depth of 25 m and is covered with ice for 10 - 12 months each year (Laybourn-Parry and Bell, 2014). The deep water of the lake is permanently anoxic and possesses an anaerobic community including methanogenic archaea, sulphate-reducing and sulphur-oxidizing bacteria (Laybourn-Parry and Bell, 2014). *M. burtonii* was isolated from 24m depth of the lake and was named after the limnologist Harry R. Burton, who discovered methane in Ace Lake (Franzmann *et al.*, 1992).

In general, methanogens are strict anaerobes and able to metabolise numerous substrates to methane, generating cellular energy and producing biomass. Depending on the mode of metabolism, methanogens can be assigned to one of three types: CO₂ reducing (use CO₂ + H₂ as substrate), acetoclastic (use acetate + H₂ as substrate) and

methylotrophic (use simple methylated carbon compounds, *i.e.* methylamines and methanol) (Garcia *et al.*, 2000). *M. burtonii* falls into the third group and utilizes only methylamines and methanol for methanogenesis. The organism is motile and can tolerate temperatures ranging from -2.5 °C to 29 °C (T_{\max}) under the laboratory conditions, with maximum growth rate at 23 °C (Franzmann *et al.*, 1992; Cavicchioli, 2006). The capacity to tolerate higher temperature classifies *M. burtonii* as a eurypsychrophile, as opposed to a stenopsychrophile that is restricted to growth at low temperatures, *e.g.* *Methanogenium frigidum* (T_{\max} = ~18 °C) (Cavicchioli, 2006). While most of the psychrophilic archaea have proven to be difficult to isolate and grow in the laboratory, *M. burtonii* is amenable to laboratory cultivation (Cavicchioli, 2006). Because of the readily cultivable capacity of *M. burtonii* and general scarcity of isolated psychrophilic archaea, *M. burtonii* has become a useful model psychrophilic organism to study molecular mechanism of archaeal cold adaptation (Cavicchioli, 2006).

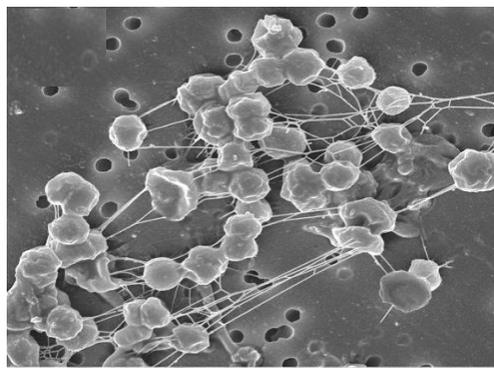


Figure 1.1 Electron micrograph of *M. burtonii* cells

The figure has been adapted from Reid *et al.*, 2006.

The currently available genomics, proteomics and transcriptomics data along with studies on a few specific proteins are discussed in the following sections.

1.1.1 Genomic and proteomic studies on *M. burtonii*

The draft genome sequence of *M. burtonii* was first reported by Saunders *et al.* (2003) and using comparative genomic analyses with methanogens and all available archaeal complete genomes (only 16 were available at the time), genome-wide characteristics of cold adaptation were identified. The key findings included identification of two hypothetical proteins with cold shock domain (CSD) fold and a unique winged-helix DNA binding protein in *M. burtonii* suggesting that they may play significant role in cold adaptation (Saunders *et al.*, 2003). Moreover, protein homology models were generated using the genome sequence of cold adapted *M. burtonii* and *Methanogenium frigidum* and analyses of 1,111 model showed that psychrophilic proteins had a strong tendency to have more glutamine, threonine and hydrophobic residues, and fewer charged residues, in the solvent accessible area that might increase the flexibility with concomitant catalytic efficiency at low temperatures.

To define the key biological processes functioning in *M. burtonii* at cold temperature (4 °C), Goodchild *et al.* (2004a) performed proteomic analysis using liquid chromatography-mass spectrometry and correlated the data with draft genome sequence. The analyses identified 528 proteins including 133 hypothetical proteins. Expression of all components of RNA polymerase, transposases, bacterial-like and eukaryal-like transcriptional regulatory proteins at 4 °C reflected an existing evolutionary and functional complexity in *M. burtonii*. In order to identify the proteins directly associated with cold adaptation, comparative (cold *vs.* high) analyses were required. For comparative proteomic analyses, Goodchild *et al.* (2004b) used two-dimensional gel electrophoresis (2DE) to identify differentially expressed proteins at 4 °C *vs.* 23 °C and following mass spectrometry, 43 differentially abundant proteins were identified. As a

complementary analysis to the 2DE, Goodchild *et al.*, (2005) additionally performed differential proteomic analysis using isotope coded affinity tag (ICAT) chromatography that identified a further 11 proteins (not found in 2DE) as differentially abundant. Moreover, to determine whether secreted proteins play a role in the cold adaptation in *M. burtonii*, proteomic analysis of the secreted proteins at 4 °C and 23 °C was performed by Saunders *et al.*, (2006). Collectively, the proteomic data showed that the cold adaptation of *M. burtonii* involve proteins that are important in transcription, protein folding, protein transport, metabolism, and cellular interactions.

The genome of *M. burtonii* was closed by Allen *et al.* (2009) and a more comprehensive genomic analyses revealed an over-representation of signal transduction genes and genes involved in replication, recombination and repair, polysaccharide biosynthesis genes, transposase genes and cell wall, membrane, and envelope biogenesis genes (Allen *et al.*, 2009). Several of those over-represented genes were bacteria-like which indicated horizontal gene transfer (HGT) from Epsilon- and Deltaproteobacteria (Allen *et al.*, 2009). Transposase genes were also overrepresented in the genome. A previous proteomic study found two transposases that were expressed in cell during growth at 4 °C (Goodchild *et al.*, 2004a) indicating active genome rearrangement events under cold condition. Moreover, it was found that while the codon usage was similar to mesophilic *Methanosarcina* spp., amino acid usage was skewed, with *M. burtonii* preferring those amino acids that were associated with cold adaptation. Overall the analyses revealed the capacity of *M. burtonii* to evolve through genome plasticity including nucleotide skew, horizontal gene transfer, and transposase activity that was assumed to facilitate the cold adaptation (Allen *et al.*, 2009).

After analysing the whole genome sequence a global proteomic analysis of soluble, insoluble and supernatant fractions of *M. burtonii* was performed by Williams *et al.*, (2010) to assess the effect of growth temperature (4 °C vs. 23 °C) and carbon source on cellular protein levels. The study highlighted a large number of specific proteins and functional processes that are involved in the growth of *M. burtonii* at cold temperature. The specific proteins that are abundant at 4 °C compared to 23 °C included TRAM domain proteins which were hypothesized to function as RNA chaperones, DEAD box helicases that are responsible for unwinding of secondary structures in messenger RNA, winged helix domain protein (a predicted transcriptional regulator), isomerases that are involved in protein folding, and surface layer proteins that may assist extensive remodelling of the cell envelope in response to low temperature (Williams *et al.*, 2010).

Subsequently, quantitative proteomics investigations were performed by Williams *et al.*, (2011) to determine the changes in global protein level across the full growth temperature range of *M. burtonii* including upper and lower growth temperature limits; at - 2, 1, 4, 10, 16, 23 and 28 °C. The analyses identified three distinct physiological states across the temperature: cold stressed (at -2 °C), cold adaptation (at 1, 4, 10 and 16 °C) and heat stressed (at 23 and 28 °C). Additionally, it was found that the expressed proteome at 1°C was very similar to that at 4 °C, illustrating that previous proteomics data at 4 °C provides a good approximation of gene expression at environmental temperatures (1-2 °C). The -2 °C (T_{\min}) profile was significantly different and was dominated by oxidative stress proteins. At 28 °C (T_{\max}) the expressed proteome was similar to that at 23°C (T_{opt}) with a high abundance of proteins associated with heat stress. At 10 and 16 °C, the methanogenesis pathways operated efficiently indicating no

stress. Overall, the study provided a great depth of understanding about the molecular mechanisms of adaptation to cold stress, heat stress and regular growth temperatures of *M. burtonii*.

To define the low temperature transcriptional regulation and the arrangement of transcriptional units in *M. burtonii*, microarray for transcriptomics was developed by Camparano *et al.*, (2011). The analyses found that 55 % of the genes are arranged in operons that range in length from 2 to 23 genes. A positive correlation was found between operon length and mRNA abundance. The biological processes overrepresented by operons tend to be those that require high cellular levels of mRNA and efficient synthesis of proteins. Therefore longer operons appear to function to not only coordinate expression of genes with related functions but to facilitate enhanced expression and stability of mRNA. In addition, genes that are involved in maintaining RNA in a suitable state for translation in the cold were upregulated, which indicated that transcriptional regulation rather than translation is responsible for gene regulation in *M. burtonii*.

1.1.2 Specific protein studies

Though comparative proteomic analyses at low vs. high temperature identified a large number of proteins likely to be involved in cold adaptation mechanism in *M. burtonii*, very few of them have yet been characterized. The proteins of *M. burtonii* with functional characterization data are described below.

1.1.2.1 Elongation factor 2 (EF-2)

To define the structural properties and cold adaptation mechanisms of psychrophilic proteins, elongation factor - 2 (EF2) of *M. burtonii* was thoroughly studied. EF-2 is a GTPase involved in the translocation step of ribosome and therefore is an essential factor in protein synthesis. Structural comparison of EF-2 of *M. burtonii* with the phylogenetically closely related thermophile *Methanosarcina thermophila* showed that the psychrophilic EF-2 possesses structural features, that include fewer salt bridges and less-packed hydrophobic cores, indicative of a more flexible and thermolabile protein (Thomas and Cavicchioli, 1998). Comparative biophysical and biochemical characterization showed that EF-2 from *M. burtonii* required less activation energy for GTP hydrolysis and protein unfolding in comparison to its thermophilic counterpart, thereby possessing biochemical properties that should assist function at low temperatures (Thomas and Cavicchioli, 2000; Siddiqui *et al.*, 2002). Moreover, activity and stability analyses in presence of ribosome and intracellular compatible solutes showed stimulated GTPase activity and moderate stabilization of EF-2 of *M. burtonii* indicating that, along with the inherent properties of EF-2, intracellular factors may play critical role in cold adaptation of the protein (Thomas *et al.*, 2001; Siddiqui *et al.*, 2002; Thomas and Cavicchioli, 2002).

1.1.2.2 RNA polymerase subunit E/F (MbRpoE/F)

Proteomic studies showed that *M. burtonii* expresses all 12 components of RNA polymerase (RNAP) at 4 °C (Goodchild *et al.*, 2004a) and the expression of RNAP subunit E is enhanced at 4 °C compared to 23 °C (Goodchild *et al.*, 2004b; Goodchild *et al.*, 2005) indicating the important role played by RNAP in cold adaptation of *M.*

burtonii. To determine the specific transcripts that MbRpoE/F binds to exert cold adaptation in *M. burtonii*, recombinant MbRpoE/F was produced and used to capture specific RNA from *M. burtonii* whole cell extract to determine the RNA targets (De Francisci *et al.*, 2011). It was discovered that only 117 genes (4 % of the total) representing 48 regions of the genome were bound by MbRpoE/F (De Francisci *et al.*, 2011). The bound transcripts represented genes of distinct functional classes: methanogenesis, cofactor biosynthesis, nucleotide metabolism, transcription, translation and import/export (De Francisci *et al.*, 2011). Most of these represented genes are arranged in operons and appear to be arranged in the genome relative to the putative origin of replication to enable effective gene expression (De Francisci *et al.*, 2011). The analyses also identified two motifs as a possible binding site for the enzyme.

1.1.2.3 Chaperonin

Chaperonins are a major subclass of molecular chaperones that assist nascent chain folding and refold unfolded proteins, thereby maintaining the stability of the proteome in extreme environments. *M. burtonii* has three group II chaperonins (MbCpns) that are known to be expressed. The stability, activity and structural properties of these MbCpns had been assessed and the structure of one of the MbCpn has been resolved using X-ray crystallography (Pilak *et al.*, 2011). The chaperonins of *M. burtonii* were found to be predominantly monomeric both *in vitro* and *in vivo* that is different from the other closely related methanogen, *Methanosarcina mazei* chaperonins. Chaperonins from *M. mazei* form ring complexes, indicating that monomeric chaperonin in *M. burtonii* is likely to be required for cold adaptation (Pilak *et al.*, 2011).

1.1.2.1 Cold responsive TRAM domain protein (Ctr)

Ctr proteins contain a single RNA-binding TRAM domain that has no intrinsic catalytic function, and is present in RNA-modifying enzymes and other proteins associated with the translation machinery (Taha *et al.*, 2016). *M. burtonii* expresses three Ctr (Ctr1, Ctr2, Ctr3) which were hypothesized to function as a cold-adapted RNA chaperones (Williams *et al.*, 2010; Williams *et al.*, 2011). Proteomic analyses had found these Ctr to have increased abundance during low temperature growth (4 °C vs. 23 °C) (Williams *et al.*, 2010) with the highest levels at 1 °C and -2 °C (Williams *et al.*, 2011). To analyse the functional properties, one of the Ctr of *M. burtonii*, Ctr3, was purified as a recombinant protein and specific RNA bound by Ctr3 was identified using RNA-sequencing (Taha *et al.*, 2016). The analyses showed that Ctr preferably bind to tRNA and 5S rRNA with a specific motif named 4C_M1 (Taha *et al.*, 2016). The data indicates that binding of Ctr to tRNA and 5S rRNA may facilitate low temperature translation in *M. burtonii*, possibly by promoting low temperature ribosome biogenesis, peptide elongation or by improving the functional capacity of the translational machinery (Taha *et al.*, 2016).

1.2 Two component system (TCS)

1.2.1 An overview of TCS

The fact that unicellular microorganisms are unable to insulate themselves from drastically changing environment, they must be able to adapt to the changes in their surroundings, *i.e.* temperature, pH, nutrients, concentration of ions such as magnesium, sodium. For this, the first consideration is to detect the changes in environmental conditions. Organisms have evolved excellent mechanisms to detect physical and

chemical signals within their own cells and in extracellular environment. One of the most prevalent mechanisms by which microorganism sense, respond and adapt to a range of stresses and environments, is a two-component system (TCS). These TCSs serve as a basic stimulus-response coupling mechanism that regulates a wide variety of cellular behaviours including metabolism, motility, virulence and development (Stock *et al.*, 2000; Gao and Stock, 2009).

The term “two-component system” was first used three decades ago to describe a class of unique regulatory systems that control responses to environmental stimuli found in *Escherichia coli*, *Bacillus subtilis*, *Agrobacterium tumefaciens*, *Klebsiella pneumoniae* and *Salmonella typhimurium* (Nixon *et al.*, 1986). However, to-date TCSs have been found in bacteria, archaea, fungi and plants (Thomason and Kay, 2000). As the name suggests, these systems consist of two components: an input component, the membrane bound sensor histidine kinase (SK) that can sense the environmental stimuli, and an output component, the response regulator (RR) that can elicit a specific cellular response by regulation gene expression (Stock *et al.*, 2000; Gao and Stock, 2009). Upon sensing an environmental signal, the SK becomes autophosphorylated using ATP at a conserved His residue and then transfers the phosphoryl group to a conserved Asp residue in the RR that in turn binds to DNA regulatory sequences affecting gene expression (Stock *et al.*, 1989; Mitrophanov and Groisman, 2008). Finally, the output response of RR is terminated by the hydrolysis of phosphorylated RR (RR~P) (Stock *et al.*, 2000; Huynh and Stewart, 2011). The dephosphorylation of RRs can also be influenced by SKs that also exerts phosphatase activity on the phosphorylated RR (RR~P) as a negative control and accelerate RR~P hydrolysis (Huynh and Stewart, 2011).

Collectively, the chemistry of a typical TCS incorporates three phosphotransfer reactions: autophosphorylation of SKs, phosphotransfer from SKs to RRs, dephosphorylation of RRs, and two phosphoprotein intermediates (SK-P and RR-P) (Fig. 1.2). However, all three reactions require divalent ion, presumably Mg^{2+} (Stock *et al.*, 2000).

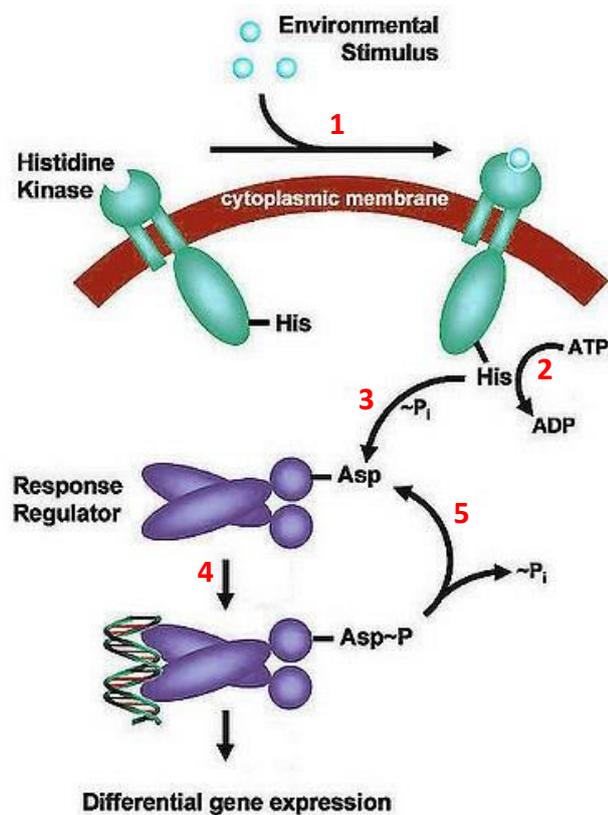


Figure 1.2 Two-component signal transduction system.

The signal transduction pathway begins when an environmental stimulus (1) is sensed by the extracytoplasmic domain of the histidine kinase. This induces ATP dependent autophosphorylation at a conserved His residue (2). The phosphoryl group can then be transferred to a conserved Asp of the response regulator (3). Once phosphorylated, the response regulator may function as a transcriptional regulator (4). After the response the RR gets inactivated by dephosphorylation (5). The figure has been adapted from “flickr” website, <https://www.flickr.com/photos/ajc1/4351324678/>

However, TCSs can be further modified and may include more than two components, which are used to transfer the phosphoryl group from SK several times, called phosphorelay, before it phosphorylates the RR (Fig. 1.3) (Hoch, 2000). For example, *Bacillus subtilis* has a sporulation controlling phosphorelay system that involves four separate proteins KinA, Spo0F, Spo0B, and Spo0A (Eswaramoorthy *et al.*, 2010). In each of phosphorelay, the transfer of the phosphate from His → Asp is conserved (Fig. 1.3) (Hoch, 2000). The phosphorelay allows the system to be controlled at each step and in turn can allow a more tightly controlled system of regulation and response to external and internal signals.

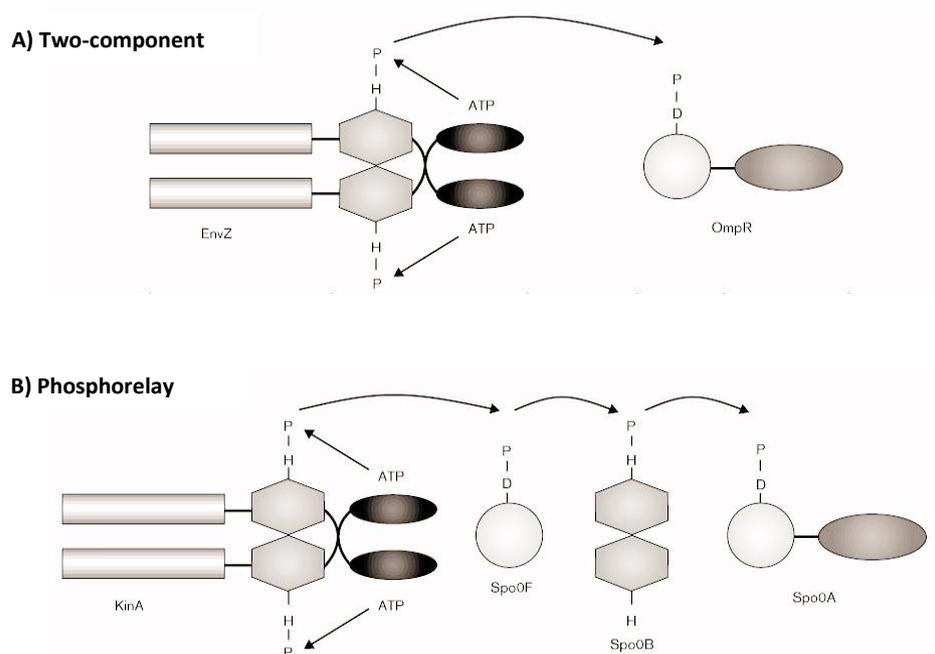


Figure 1.3 Comparison of the pathways involving two-component and phosphorelay

A) The SK EnvZ, that is involved in osmoregulation in *Escherichia coli* autophosphorylates at the histidine (H) residue upon receiving stimuli and transfers the phosphoryl group directly to the RR, OmpR (Cai and Inouye, 2002); B) The sporulation phosphorelay of *B. subtilis* involves the SK KinA and two additional phosphotransferases SpoF and SpoB that transfers the phosphoryl group from KinA to RR, SpoA (Eswaramoorthy *et al.*, 2010). The figure has been modified from Hoch, 2000.

Although TCSs are found in all three domains of life: *Bacteria*, *Archaea* and *Eukarya* (Stock *et al.*, 2000; Wuichet *et al.*, 2010), they are mostly abundant in bacteria, constituting ~1% of the encoded proteins (West and Stock, 2001). The number of TCS possessed by an organism can vary greatly depending on its genome complexity, physiology and the variability of its habitat (Ashby, 2006). For example, *E. coli* encodes 62 TCS proteins (30 SKs and 32 RRs) whereas *Mycoplasma genitalium* encodes none and *Streptococcus pneumoniae* encodes 27 (13 SKs and 14 RRs) (Stock *et al.*, 2000; West and Stock, 2001). In contrast to bacteria, TCSs are less abundant in archaea and eukarya. TCS genes are only present in *Euryarchaeota* and *Thaumarchaeota* and entirely absent in the available genome sequence of *Crenarchaeota*, *Korarchaeota* and *Nanoarchaeota* (Ashby, 2006; Wuichet *et al.*, 2010). Among eukaryotes, TCSs are found primarily in protozoa, fungi, algae and plants (West and Stock, 2001; Galperin, 2010). To-date, TCS genes have not been identified in mammals including human beings (West and Stock, 2001; Gotoh *et al.*, 2010).

1.2.2 Structure and function of SKs

Most SKs are membrane anchored proteins that harbour two domains: a diverse extracytoplasmic sensor with transmembrane helices and a cytoplasmic transmitter (or kinase) domain usually termed as kinase core (Bhate *et al.*, 2015). The cytoplasmic catalytic core is often connected to the transmembrane helices via a linker (Fig. 1.4) (Bhate *et al.*, 2015). The cytoplasmic kinase core is highly conserved and consists of two subdomains: C-terminal catalytic and ATP binding (CA) subdomain (also represented as HATPase_c in Pfam database) and N-terminal dimerization and histidine phosphotransfer (DHp) subdomain (also represented as His Kinase A domain in Pfam

database) (Gao and Stock, 2009). The subdomain CA performs the catalytic reaction by transferring γ -phosphoryl group of ATP to the conserved His residue, which is located in the DHp domain (Fig. 1.4) (Gao and Stock, 2009). Biochemical evidence and structural studies indicate that SKs exist as homodimers in solution and dimerization is confined to DHp domain that form homodimeric anti-parallel four-helical bundle with its two helices (Dutta *et al.*, 1999; Bhate *et al.*, 2015).

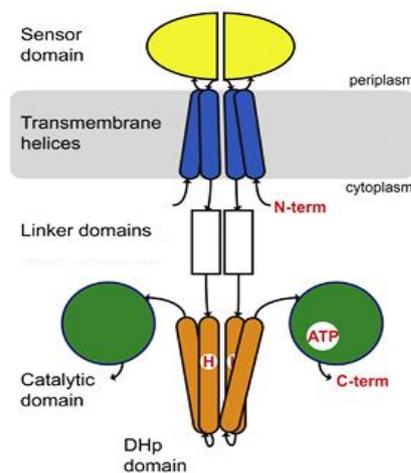


Figure 1.4 Schematic representation of domain organisation of a typical SK

H indicates phosphoryl group acceptor histidine residue in DHp domain. The figure has been adapted from Bhate *et al.*, 2015.

The catalytic core of SK also contains sequence fingers that are highly conserved in the histidine kinase protein family (both in prokaryote and eukaryote) (Stock *et al.*, 2000; Foussard *et al.*, 2001). These amino acid motifs are termed as H, N, G1, F and G2 blocks (Parkinson and Kofoid, 1992). The conserved His, required for autophosphorylation, is the central feature in the H block, whereas the N, G1, F, and G2 blocks define the nucleotide (ATP) binding cleft (Stock *et al.*, 2000). In most HKs, the H block is part of the DHp domain whereas N, G1, F, and G2 are part of CA domain (Stock *et al.*, 2000; Foussard *et al.*, 2001).

In contrast to the cytoplasmic kinase core, periplasmic sensor domains are highly variable in sequence which is indicative of the large array of signals to which SKs can respond (Wolanin *et al.*, 2002). Depending upon the type of stimulus, SKs can be divided into three groups (Fig. 1.5) (Mascher *et al.*, 2006):

- (i) Periplasmic-sensing SKs that use their extracellular input domain for stimulus perception (often small solutes);
- (ii) Intramembrane SKs, that detect their stimuli via their membrane-spanning segments;
- (iii) Cytoplasmic-sensing SKs (either membrane anchored or soluble), that recognize intracellular signals.

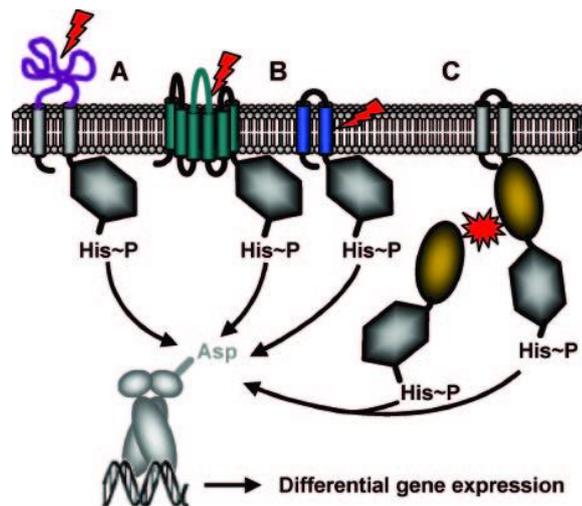


Figure 1.5 Schematic representations of three groups of SKs

A) Periplasmic-sensing SKs; B) Intramembrane SKs; C) Cytoplasmic sensing SKs. The stimulus is shown as red arrow or star, the part of protein involved in stimulus perception is coloured, purple and green in A, blue in B and brown in C. The figure has been adapted from Mascher *et al.*, 2006.

Apart from the extracytoplasmic sensor, cytoplasmic CA and DHp domains SKs may contain additional auxiliary domains and different domain organisation (Stock *et al.*, 2000; Foussard *et al.*, 2001; West and Stock, 2001). Depending on the presence of additional domains, SKs are divided into two groups: orthodox and hybrid (Stock *et al.*,

2000). Orthodox SKs represent the prototypical SKs that are periplasmic membrane receptors and have N-terminal sensing domain and a cytoplasmic catalytic region containing CA and DHp domain, *i.e.* EnvZ from *E. coli* (Fig. 1.6A) (Khorchid *et al.*, 2005). In contrast, hybrid kinases contain multiple phosphodonor and phosphoacceptor sites and use multistep phosphorelay system. For example, ArcB is a hybrid kinase from *E. coli*, which is involved in sensing the oxygen availability and is composed of two N-terminal transmembrane regions followed by a kinase core, a domain similar to the regulatory domain of RRs, and a second His-containing phosphotransfer (HPt) domain at the C-terminal end (Fig 1.6B) (Stock *et al.*, 2000).

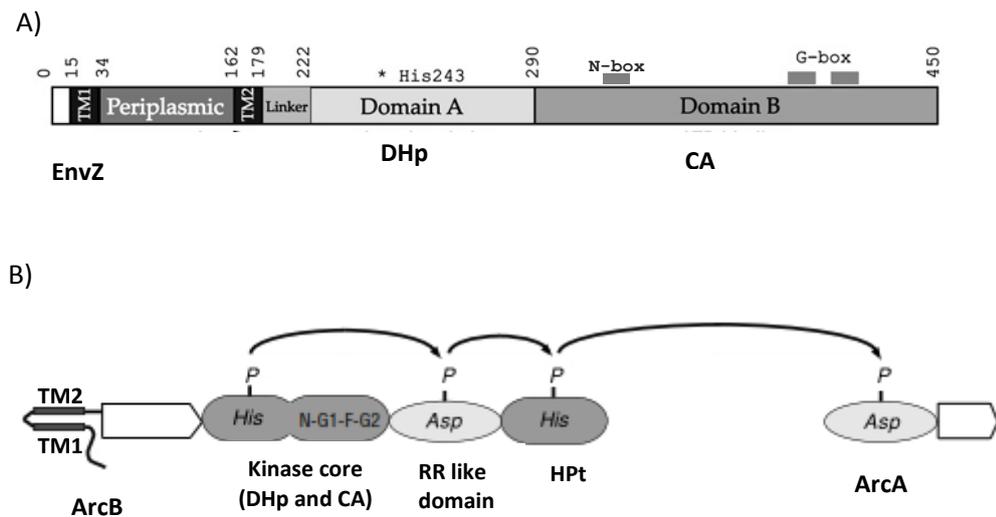


Figure 1.6 Schematic representation of subdomain organisation of orthodox and hybrid kinase

A) EnvZ is 450 amino acids long SK with two transmembrane domain (TM1, TM2), a periplasmic domain, a short linker domain and cytoplasmic DHp and CA domains. The conserved His243 required for phosphorylation is part of conserved H block that is located in DHp domain and conserved N and G blocks (G1, G2) are part of the CA domain (Khorchid *et al.*, 2005); B) The hybrid SK, ArcB has two N terminal transmembrane domains (TM1, TM2), a kinase core with conserved H, N, G and F blocks; a RR like domain with Asp and an HPt domain with another His residue. The phosphoryl group is transferred from His of kinase core to Asp of RR like domain to His of HPt domain from which it is finally transferred to the Asp of RR, ArcA (Stock *et al.*, 2000). The figure has been modified from Stock *et al.*, 2000 and Khorchid *et al.*, 2005.

Most often, SKs function as a dimer and autophosphorylation activity occur in *trans* configuration, where ATP bound to CA domain of one monomer transfers its phosphoryl group to the His of the other monomer (Yang and Inouye, 1991; Dutta *et al.*, 1999). For example, *trans*-autophosphorylation has been demonstrated in CheA (Swanson *et al.*, 1993), EnvZ (Cai and Inouye, 2002) and NRII (Ninfa *et al.*, 1993) of *E. coli*. However, *cis*-autophosphorylation had also been observed in few SKs, *i.e.* VicK of *Streptococcus mutans* (Wang *et al.*, 2013), HK853 of *Thermotoga maritima* (Casino *et al.*, 2009), where each monomer of the dimer catalyse its own phosphorylation. To understand the autophosphorylation mechanism (both *cis* and *trans*), several SKs have been crystallised, although a full length membrane spanning SK structure is not yet available (Casino *et al.*, 2009; Wang *et al.*, 2013; Ferris *et al.*, 2014; Bhate *et al.*, 2015). As a representative of *trans* autophosphorylating SK, the structure of EnvZ of *E. coli* (Ferris *et al.*, 2014) is described in Fig. 1.7. To analyse the autokinase mechanism of EnvZ, crystal structure was resolved for a chimeric protein containing the entire catalytic part of the *E. coli* EnvZ, fused to the sensor HAMP domain of the *Archaeoglobus fulgidus* Af1503 receptor (Ferris *et al.*, 2014). The construct was equivalent to the full cytosolic part of EnvZ. The structure showed that the DHp domain form an elongated helix bundle within a dimer (Fig. 1.7). The C-terminal helix of active CA domain forms a tight helical bundle with the C-terminal DHp helix of its own chain and the N-terminal DHp helix of the other chain in a dimer that allows to position the CA domain bound ATP in close proximity of the conserved His in the other chain of DHp domain (Fig. 1.7). On the other hand, in the inactive state, the C-terminal helix of CA domain is loosely packed with the C-terminal helix of the DHp domain belonging to its own chain.

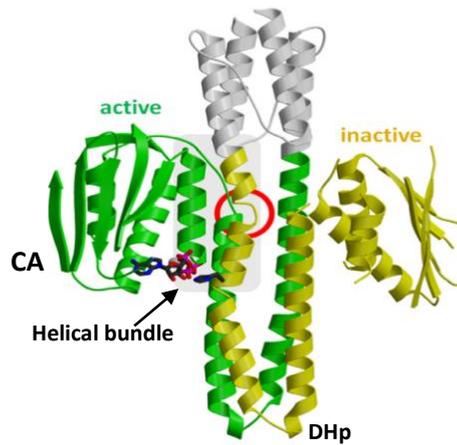


Figure 1.7 The *trans*-autophosphorylation mechanism of EnvZ .

The CA and DHp domain of two chains in a dimer is shown in green and yellow colour, other domains are grayed out. The ATP and conserved His is shown in stick representation. In active state the C terminal helix of ATP bound CA domain of one chain (green) forms helical bundle with N-terminal helix of another chain's DHp domain (yellow). The figure has been adapted from Ferris *et al.*, 2014.

VicK from *S. mutans* is one of the *cis*-autophosphorylating SKs (Wang *et al.*, 2013). After VicK structure analyses a model was suggested for *cis*-autophosphorylation that incorporates DHp helical bending and a CA domain's swing movement for autokinase activation (Fig. 1.8) (Wang *et al.*, 2013).

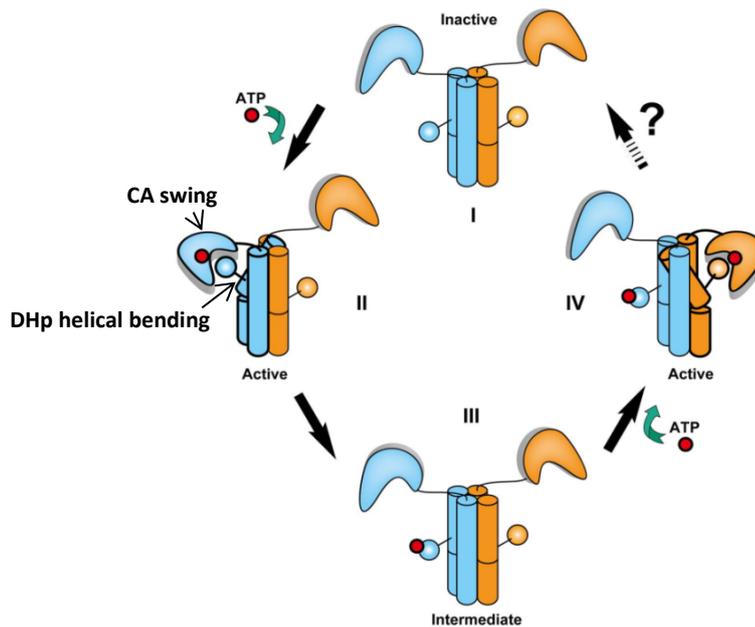


Figure 1.8 Model of VicK *cis*- autophosphorylation mechanism.

The four cylinders in the middle of each state (I–IV) represent the four helical bundle of the DHp domain in VicK dimer, each monomer in a dimer is coloured in blue and orange. The conserved His residue is indicated in sphere and line attached to the four cylinders. The CA domains are shown in an L shape and ATP in a red sphere. The unknown steps for VicK to return to the inactive state are labelled with a question mark. The figure has been adapted from Wang *et al.*, 2013.

The phosphoryl group in SKs is linked to nitrogen atom of the His residue, thus forming a phosphoramidate bond (Stock *et al.*, 2000). This feature makes this class of kinases different from Ser/Thr/Tyr kinases that form a phosphoester bond (Stock *et al.*, 2000; Klumpp and Krieglstein, 2002). The phosphohistidines have a large standard free energy of hydrolysis which makes them highly unstable and suitable for phosphoryl transfer (Stock *et al.*, 2000; Klumpp and Krieglstein, 2002). In addition, many SKs are termed bifunctional, *i.e.*, they are able to switch between two opposing activities and exert both positive and negative control. In positive control, SKs may possess autophosphorylation and phosphotransfer activity towards RR, and in negative control

they may stimulate the fast dephosphorylation of the RR, thus acting as a phosphatase (Huynh and Stewart, 2011). The phosphatase activities of SKs provide an additional means of regulating the level of phosphorylated RR in the cell. Phosphatase activity resides in the DHp domain and in some SKs the conserved His residue is essential for the activity (West and Stock, 2001). One potential dephosphorylation mechanism is reverse transfer where the phosphoryl group of RRs is transferred back to the conserved His of SKs, which has been observed in some SKs including EnvZ of *E. coli* (Zhu *et al.*, 2000), PhoQ of *Salmonella enterica* (Castelli *et al.*, 2000). However not all SKs exhibit such reverse transfer, as some SK mutants that lack the conserved His still retain the phosphatase activity, *i.e.* DesK of *B. subtilis* (Albanesi *et al.*, 2004), NRII of *E. coli* (Kamberov *et al.*, 1994). In contrast, some SKs do not exhibit phosphatase activity thus involve an auxiliary phosphatase protein. For example, the phosphatase activity of SK, NtrB is regulated by an auxiliary protein PII, in the absence of PII, the SK is unable to exhibit phosphatase activity on its RR, NtrC (Kamberov *et al.*, 1994). The SKs involved in chemotaxis are also monofunctional, *i.e.* CheY displays kinase activity on RR, CheA but the dephosphorylation of CheA is regulated by auxiliary phosphatase, CheZ or CheX (Huynh and Stewart, 2011). However, most TCS does not involve auxiliary regulator and the bifunctional activity of SKs is likely to be reciprocally regulated through dynamic conformational changes (Huynh and Stewart, 2011).

1.2.3 Structure and function of RRs

Prototypical RRs are composed of two domains: a conserved N-terminal receiver or regulatory domain (REC), and a variable C-terminal effector domain (Stock *et al.*,

2000; Gao *et al.*, 2007). The REC domain performs three functions (West and Stock, 2001):

- Catalyse phosphotransfer from the His residue of the SK (or HPt, in the case of a hybrid SKs), to a conserved Asp within its own domain (West and Stock, 2001; Gao *et al.*, 2007).
- Catalyse dephosphorylation (Stock *et al.*, 2000; West and Stock, 2001).
- Regulate the activity of associated effector domain in a phosphorylation-dependent manner (West and Stock, 2001; Gao *et al.*, 2007).

The C-terminal effector domain is variable due to the diversity of its output response. The effector domain can be classified into four groups depending on the output response: DNA-binding, RNA-binding, protein-binding and enzymatic domain (Gao *et al.*, 2007). Majority of the RRs (~ 63 %) contains DNA binding domain and function as a transcription factor (Stock *et al.*, 2000; Gao and Stock, 2009). Some RRs (~ 13 %) function as enzymes (Gao and Stock, 2009), *e.g.* the RR of *E. coli* in chemotaxis pathway and CheB, as a methyltransferase (Simms *et al.*, 1985). Few RRs can also function as a RNA binding (~ 1 %) or protein binding (~ 2 %) regulator (Gao *et al.*, 2007), *e.g.* NasR, that mediates nitrate response in *Klebsiella oxytoca*, functions as a RNA binding protein (Chai and Stewart, 1998); CheV, involved in chemotaxis, is a protein binding RR in *B. subtilis* (Galperin, 2006).

The DNA-binding RRs can be further subdivided into three groups based on the homology of their DNA-binding domain: OmpR/PhoB containing winged-helix domain, NarL/FixJ containing four helix domain and NtrC containing ATPase coupled DNA binding domain (West and Stock, 2001). Another RR subfamily has been

characterized lately, namely LytTR that has a unique DNA binding domain containing β fold (Gao and Stock, 2009).

It is noteworthy that variability in the RR common modular architecture has been observed, as some of them consist only of a REC domain without any effector domain, *e.g.* CheY of *E. coli* that controls bacterial motility by intermolecular interactions with flagellar motor (Gao and Stock, 2009); and Spo0F in *B. subtilis*, that function as phosphorylated intermediates in phosphorelay pathways (Varughese, 2005). Such single domain RR comprises 17 % of the bacterial RRs (Galperin, 2010).

Biochemical and structural analyses of the DNA-binding RRs showed that the specific DNA sequences that these proteins recognize, the arrangement of binding sites and the specific mechanism of transcription regulation differ for each RR, even within the same subfamily (Stock *et al.*, 2000). For example, both OmpR and PhoB belong to OmpR/PhoB subfamily that contains winged-helix conformation for the DNA binding domain. The crystal structure of DNA-binding domain of OmpR (Martínez-Hackert and Stock, 1997) and PhoB (Blanco *et al.*, 2002) showed that the recognition helix interacts with the major groove of DNA and flanking loops or wings interact with the minor groove. However, despite having structural similarity, OmpR involves interaction with the α -subunit of RNA polymerase (Martínez-Hackert and Stock, 1997), whereas PhoB interacts with σ^{70} (Blanco *et al.*, 2002).

Although RRs catalyse phosphoryl transfer from the phospho-His of the SKs, some RRs can also use small phosphodonor molecules such as acetyl phosphate, carbamoyl phosphate, imidazole phosphate and phosphoramidate for phosphorylation (Stock *et al.*, 2000). This suggests that some RR can catalyse phosphoryl transfer independently of any assistance from the SKs (Stock *et al.*, 2000). However, the

phosphotransfer to the Asp residue produces a high energy acyl phosphate that is rapidly hydrolysable in both acidic and alkaline conditions, thus making identification of phosphor-Asp technically difficult (Perry *et al.*, 2011). The high energy in the acyl phosphate bond is hypothesized to be used to drive conformational changes in the protein that results in an increased affinity for the recognition sequences of DNA (Perry *et al.*, 2011).

In many RRs the phosphorylation promotes dimerization or higher-order oligomerization. Nearly 50 % of RRs, including the OmpR/PhoB, NarL/FixJ, and LytTR subfamilies, contain only a single DNA binding domain as the effector domain and access to the DNA of the effector domain is blocked by their REC domains (Gao and Stock, 2010). Phosphorylation mediated dimerization of the REC domains induces the repositioning of the N and C-terminal domains and promotes DNA binding and transcription activation (Gao and Stock, 2010). For example, the structures of OmpR/PhoB family members, *T. maritima* DrrB (Robinson *et al.*, 2003) and DrrD (Buckler *et al.*, 2002), showed a monomeric state with different domain arrangements in the inactive state and upon phosphorylation form dimers that release the inhibitory effect of REC domain and allows the DNA binding domain to bind the recognition sequence (Fig. 1.9A) (Gao and Stock, 2009). In contrast, some RRs of NtrC subfamily, which contain an ATPase domain with a helix-turn-helix domain at C-terminus (Stock *et al.*, 2000), form homodimers even in the absence of phosphorylation while phosphorylation alters the mode of dimerization to allow oligomers to be formed (Gao and Stock, 2010). For example, NtrC of *S. enterica* is a dimer in unphosphorylated state where the ATPase domain is inactive (Gao and Stock, 2009). Phosphorylation of REC domains promotes their intermolecular association with neighbouring ATPase

domains resulting in oligomerization to octamer and activation of ATPase (Fig. 1.9B) (Stock *et al.*, 2000; Gao and Stock, 2009). The ATP hydrolysis provides energy for complex formation and activation of transcription (Stock *et al.*, 2000).

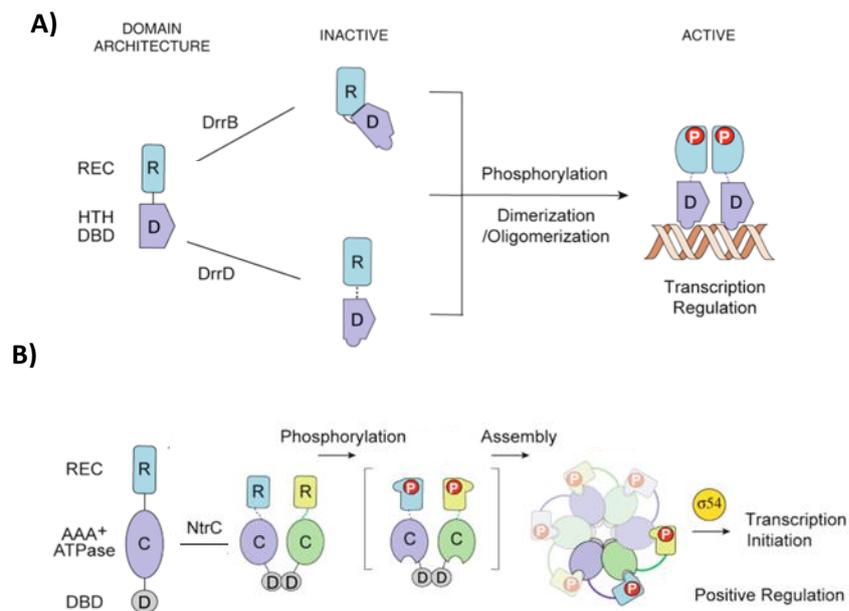


Figure 1.9 Schematic diagram of regulatory mechanism in RR subfamilies

A) OmpR/PhoP subfamily member DrrB and DrrD from *T. maritima* dimerizes upon phosphorylation and binds to DNA; B) NtrC subfamily member NtrC is a dimer in inactive state and phosphorylation induces oligomerization and assembly as octamer; REC domain is denoted as R, helix turn helix DNA binding domain (HTH DBD) is denoted as D and ATPase domain is denoted as C. Red P is phosphoryl group. The figure has been modified from Gao and Stock, 2009.

However, phosphorylation need not necessarily results in activation. Some RRs are dimers in their unphosphorylated state and can bind to DNA, whereas phosphorylation only enhances the DNA binding, *e.g.* *B. subtilis* PhoP (Lejona *et al.*, 2004) and *S. enterica* PmrA (Wösten and Groisman, 1999).

The structural analyses on REC domain of RRs revealed that they adopt a $(\beta\alpha)_5$ topology (Bourret, 2010). This consists of a five stranded parallel β -sheet surrounded by

five α helices (Fig. 1.10A). The REC domain active site contains a cluster of highly conserved acidic residues that includes an Asp at the C-terminal end of $\beta 3$ that is the site of phosphorylation (Bourret, 2010). Structure of the genuine TCS with high resolution has been obtained for the HK853-RR468 complex from *T. maritima* (Casino *et al.*, 2009). The structure revealed that during phosphotransfer, the RR468 sits on the opposite side of the DHp domain dimer of SK, HK583 and align the conserved phosphoryl group acceptor Asp residue with the conserved His of the HK853. In the structure, the $\alpha 1$ of RR468 is inserted between the two DHp domain helices of HK853 and the $\alpha 3$ - $\beta 3$ loop of RR468 protrudes laterally to make a contact with the $\alpha 4$ - $\beta 4$ loop and ATP lid of CA domain of the SK (Fig. 1.10B), which ultimately allow the Asp-His alignment for phosphotransfer reaction.

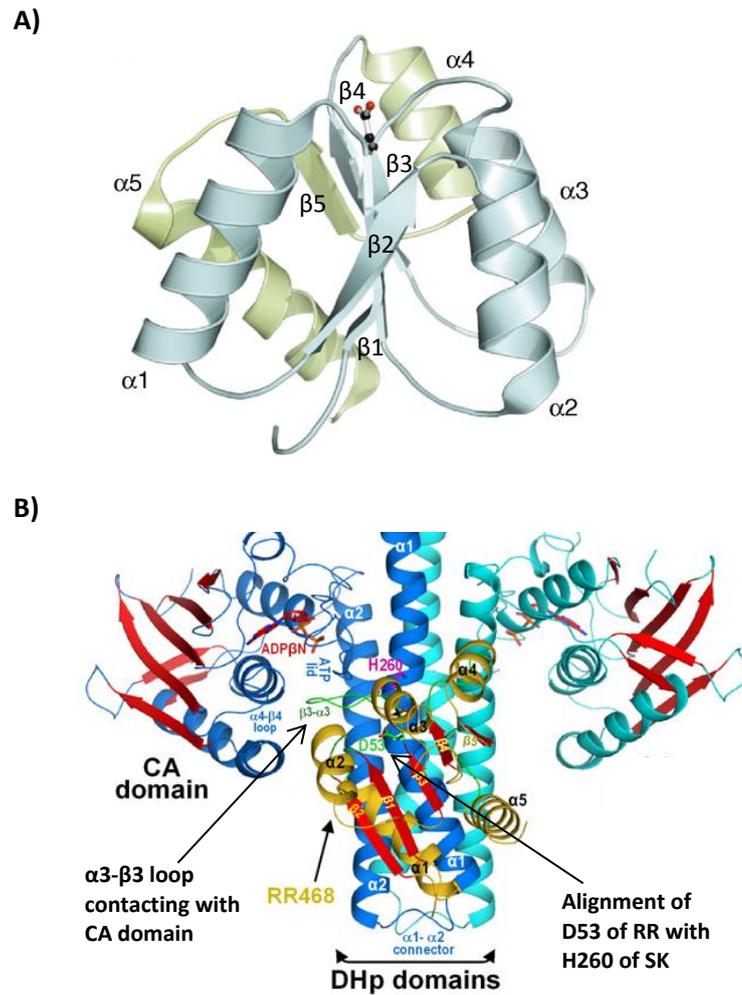


Figure 1.10 Structure of RR and RR-SK complex

A) A ribbon diagram of *E. coli* CheY (PDB: 1FQW) displays the classic $(\alpha\beta)_5$ fold of REC domain (Lee *et al.*, 2001; Gao and Stock, 2009); the site of phosphorylation, Asp at $\beta 3$ end is shown in ball and stick representation; B) HK853-RR468 structure complex (Casino *et al.*, 2009); the helices of each HK853 monomer are coloured blue and cyan and the RR468 molecule is shown in gold, β strands are coloured red in all cases. The side chains of the phosphoacceptor residues H260 of HK853 and D53 of RR468 and the bound ADP molecules are illustrated in stick representation. The figure has been adapted from Lee *et al.*, 2001 and Casino *et al.*, 2009.

The output response of RRs is regulated by either phosphatase activity of SKs or by autophosphatase activity of REC domain of the RR. In an autodephosphorylation process, the phosphate group is transferred to a water molecule in a hydrolysis reaction

(Perry *et al.*, 2011). The autophosphatase activity varies in different RR and as a result the half-life of phosphorylated RRs can range from a few seconds, as in case of CheY of *E. coli* (Hess *et al.*, 1988), to several hours, such as in PhoP (2.5 h) of *B. subtilis* (Liu and Hulett, 1997).

1.2.4 TCS as thermosensor

Temperature is one of the major environmental factors that have profound effects on cellular processes in all organisms. Microorganisms have evolved molecular thermosensing strategies in order to adjust with changes in temperature. One of these strategies includes TCS as a mean of temperature dependent transcription regulation (Klinkert and Narberhaus, 2009).

Thermosensing TCSs are most frequently found in pathogens including animal pathogens: *Edwardsiella tarda* (Chakraborty *et al.*, 2010), *Flavobacterium psychrophilum* (Hesami *et al.*, 2011), *Haemophilus influenza* (Steele *et al.*, 2012), *Bordetella pertussis* (Cotter and Jones, 2003). The TCS functioning as thermosensor in *E. tarda* is PhoP/PhoQ which plays significant role in virulence (Chakraborty *et al.*, 2010). The SK, PhoQ autophosphorylates over a temperature range of 23 °C to 35 °C (virulence promoting temperature) and is inactive below 20 °C or above 37 °C (Chakraborty *et al.*, 2010). The active PhoQ transfers the phosphoryl group to the RR, PhoP and phosphorylated PhoP activates the transcription of *esrB* gene (Chakraborty *et al.*, 2010). EsrB is a RR of another TCS, EsrA/EsrB and phosphorylated EsrB activates the transcription of two clusters of genes that encodes proteins of type III and type VI secretion systems (Srinivasa Rao *et al.*, 2004; Chakraborty *et al.*, 2010). Both type III and type VI systems are associated with virulence. Besides, PhoP was found to be

essential for the expression of 13 proteins in *E. tarda* including zinc metalloprotease Sip1 which is essential for serum resistance and host infection (Guijarro *et al.*, 2015; Zhou *et al.*, 2015). The virulence proteins' secretion was significantly lower at 37 °C compare to 25 °C. Interestingly, it has been observed that when fish are injected with *E. tarda* grown at 37 °C, 90 % of the fish survived whereas 70% of the fish died when they were injected with cells grown at 25 °C (Srinivasa Rao *et al.*, 2004).

Another pathogen, *F. psychrophilum* is responsible for bacterial cold water disease in salmon fish which occurs at temperature below 14 °C (Starliper, 2011) and a thermosensing TCS was found to be associated with its virulence. A temperature sensor SK is identified in this bacterium which is similar to LytS that is involved in cell autolysis and biofilm formation in various bacteria and the expression of this temperature sensing SK is increased ~ 18 fold at 8 °C (optimal temperature at which disease manifests itself) compared to 20 °C (the optimum growth temperature of *F. psychrophilum*) indicating the contribution of thermosensing to the virulence (Hesami *et al.*, 2011; Guijarro *et al.*, 2015).

The thermosensing TCS can also involve more complex phosphorelay mechanism. One such example is BvgS/BvgA, TCS from *B. pertussis*. *B. pertussis* is a causative agent of whooping cough, and involves a complex phosphorelay TCS, BvgS/BvgA. This TCS is involved in virulence gene expression at body temperature (37 °C) in the absence of certain modulating compound, *i.e.* nicotinic acid, sulphate ion (Cotter and Jones, 2003). The SK, BvgS is a multi-domain kinase that transduces signals using four-step His-Asp-His-Asp phosphorelay between its different domains before transferring the phosphoryl group to the RR, BvgA (Cotter and Jones, 2003). This TCS regulates four classes of genes: class I, includes genes encoding toxins, such

as pertussis toxin (PTX), adenylate cyclase toxin (CYA); class II, includes genes encoding adhesins such as filamentous hemagglutinin (FHA), fimbriae (FIM); class III, includes gene encoding an intermediate phase protein, BipA (Stockbauer *et al.*, 2001) and Class IV includes virulence-repressed (*vrg*) genes (Beier and Gross, 2006). Expression of class I genes require high concentration of BvgA~P homodimer while class II genes require relatively small amount of BvgA~P (Beier and Gross, 2006). Class III gene, *bipA* expression is positively controlled by binding of small amount of BvgA~P to the high affinity binding site located upstream of the gene and is negatively controlled by binding of high concentration of BvgA~P to the low affinity binding site within the transcribed region of the gene (Williams *et al.*, 2005). In contrast, class IV genes are only negatively controlled (Fig. 1.11) (Beier and Gross, 2006).

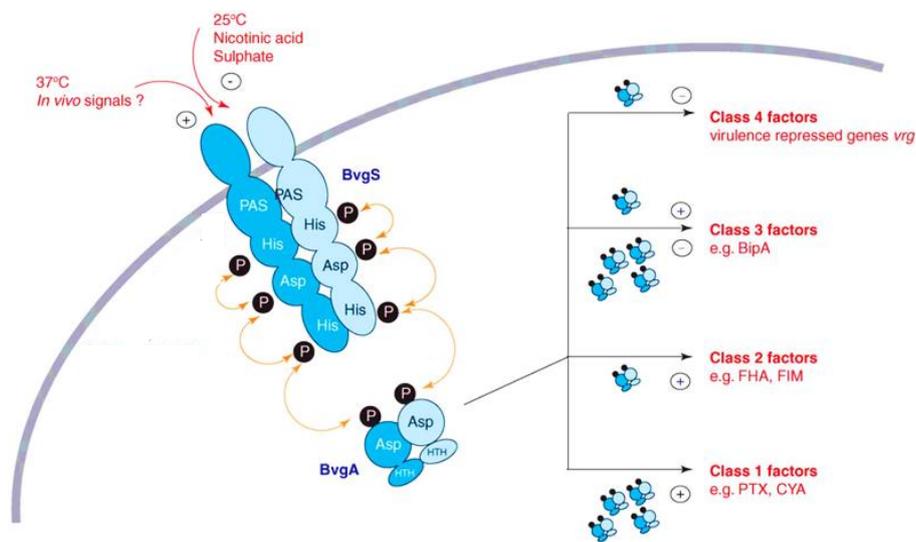


Figure 1.11 Schematic representation of virulence gene control in *B. pertussis* by thermosensing TCS, BvgS/BvgA.

The SK BvgS and RR, BvgA forms dimer. Upon phosphorylation at 37 °C, BvgA positively control the gene expression of class I and class II factor, both positively and negatively controls the expression of class III factors while only negatively controls class IV factors expression. + indicates activation, - indicates repression. The figure has been adapted from Beier and Gross, 2006.

Presence of thermosensing TCS is also detected in opportunistic pathogen *H. influenzae* which is termed as FirRS system where FirR is the SK and FirS is its cognate partner, RR (Steele *et al.*, 2012). Differential gene expression analyses indicated that the transcription of this TCS is 60 to 70 fold higher at 9 °C compared to 37 °C demonstrating its cold responsive activation (Steele *et al.*, 2012).

Thermosensing TCSs are frequently found in food pathogens including *Clostridium botulinum* (Lindström *et al.*, 2012; Mascher *et al.*, 2014), *Listeria monocytogenes* (Chan *et al.*, 2008; Pöntinen *et al.*, 2015) and *Yersinia pseudotuberculosis* (Palonen *et al.*, 2011), *Bacillus cereus* (Diomandé *et al.*, 2014). Differential gene expression analyses revealed TCS RRs to be involved in cold adaptation in foodborne pathogen *L. monocytogenes* that has the capability to grow as low as -0.4 °C (Chan *et al.*, 2007; Chan *et al.*, 2008; Pöntinen *et al.*, 2015). Temperature sensitive TCS involved in cold adaptation in *C. botulinum* nonproteolytic group I strain is CLO3403/CLO3404 (Mascher *et al.*, 2014) and proteolytic group II strain is CBO2306/CBO2307 (Derman *et al.*, 2013) and CBO0366/CBO0365 (Chai and Stewart, 1998; Dahlsten *et al.*, 2014) that have higher expression at cold (~12 °C) compared to high (~ 30 °C) temperatures. The cold tolerance mechanism of food pathogen *Y. pseudotuberculosis* includes TCS CheA/CheY that has significantly higher expression level at 3 °C compared to 28 °C and is important for bacterial growth and motility at 3 °C (Palonen *et al.*, 2011). The Gram-positive spore forming pathogenic bacteria *B. cereus* involves CasK/CasR TCS for its low temperature growth which plays a key role in cell shape during growth (Diomandé *et al.*, 2014).

Temperature responsive TCSs have also been found in plant pathogens: *Agrobacterium tumefaciens* (Jin *et al.*, 1993; Njimona and Lamparter, 2011) and

Pseudomonas syringae (Ullrich *et al.*, 1995; Braun *et al.*, 2007; Braun *et al.*, 2008). Two thermosensing SKs have been found in *A. tumefaciens*: VirA (Jin *et al.*, 1993) and Agp1 (Njimona and Lamparter, 2011). At the temperature (< 32 °C) corresponding to ‘crown gall’ infection, VirA autophosphorylates and transfers the phosphoryl group to the RR, VirG which in turn induces virulence (*vir*) gene expression (Jin *et al.*, 1993). However, above 32 °C VirA undergoes reversible inactivation and as a result cannot induce tumour formation (Jin *et al.*, 1993). The findings for Agp1 were similar to VirA. The kinase activity was found to be optimum at 25 °C whereas above 25 °C the activity starts to decrease and reduces to nil at 40 °C (Njimona and Lamparter, 2011). The thermosensing TCS in soybean pathogen *P. syringae*, is CorS/CorR that is active at 18 °C (the virulence promoting temperature) but is inactive at 28 °C (the optimum growth temperature of the bacterium) and is involved in virulence enhancing phytotoxin coronate synthesis (Ullrich *et al.*, 1995; Braun *et al.*, 2008).

Thermosensing TCSs are also present in environmental bacteria *B. subtilis* (Aguilar *et al.*, 2001; Albanesi *et al.*, 2009), cyanobacteria *Synechocystis* sp. (Suzuki *et al.*, 2000), psychrophilic bacteria *Sphingobacterium antarcticus* (Ray *et al.*, 1994).

Although many TCSs are found to be capable of temperature sensing and play role in cold adaptation, the mechanism of thermosensing by SKs remains poorly understood. The molecular mechanisms of thermosensing of some bacterial SKs are described in the following section.

1.2.4.1 Mechanism of temperature signal detection and activation of thermosensing SKs

The thermosensor CorS in *P. syringae* is involved in phytotoxin (coronate, COR) synthesis and an activation model was proposed for the temperature change mediated activation mechanism for this SK (Braun *et al.*, 2008). At virulence promoting temperature (18 °C), the SK CorS autophosphorylates and subsequently transphosphorylates the RR, CorR that in turn binds the *cma* operon and induce coronamic acid gene expression, a component of COR, leading to large amount of COR production (Braun *et al.*, 2007; Braun *et al.*, 2008). However, at 28 °C the CorS is inactive which makes COR biosynthesis negligible (Braun *et al.*, 2008). CorS is a transmembrane histidine kinase that is composed of six transmembrane domains (TMDs), followed by a catalytic H-box containing the conserved phosphoryl group acceptor, His residue (Smirnova and Ullrich, 2004; Braun *et al.*, 2008). The H-block is predicted to be located in the cytoplasm and partially within a putative seventh TMD (Braun *et al.*, 2008). At the virulence promoting temperature (18 °C), the cytoplasmic portion of the H-block is accessible to autophosphorylation, but at the optimum growth temperature of the bacterium (28 °C), it is assumed to be incorporated into the membrane leading to termination of CorS activity (Fig. 1.12) (Braun *et al.*, 2008; Steinmann and Dersch, 2013). Though these data demonstrate the activation of CorS at cold temperature, the mechanism of sensing change in temperature is still unclear.

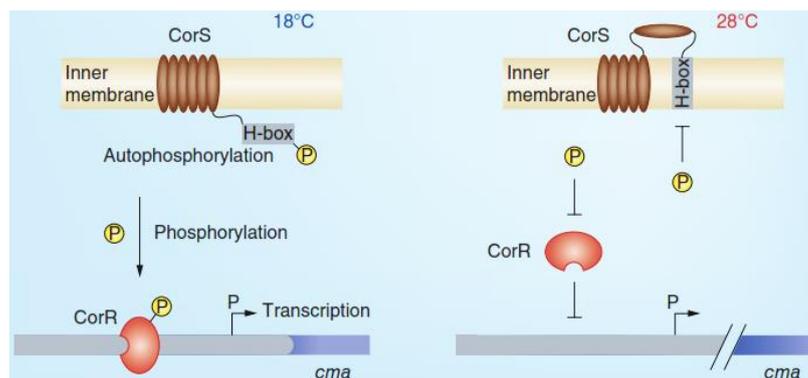


Figure 1.12 Schematic representation of the temperature dependent regulation of thermosensor, CorS

At low temperatures (18 °C), the membrane associated sensor kinase CorS autophosphorylates at a conserved His residue in the H-block and subsequently phosphorylates the RR, CorR, which in turn binds to *cma* operon and induce *cor* gene transcription. At 28°C, the CorS TMDs rearrange that force the H-block to localize in the membrane where the SK is not accessible for autophosphorylation and as a result the phosphotransfer to CorR is abolished. The figure has been adapted from Steinmann and Dersch, 2013.

In some thermosensing SKs the temperature effect had been linked with the physical state of the lipid membrane, *i.e.* Hik33 in *Synechocystis* and DesK in *B. subtilis*. With the change in temperature membrane fluidity and arrangement of the phospholipids in the membrane changes (Vigh *et al.*, 1998), which may trigger the conformational changes in the membrane associated SKs leading to activation. Two SKs have been identified as necessary for the expression of cold-inducible desaturases in *Synechocystis*, Hik33 and Hik19 (Suzuki *et al.*, 2000). The identified RR, phosphorylated by these two SKs is Rer1 that regulates the expression of desaturase, *desB* gene. Through computational analyses, Hik33 is predicted to be a transmembrane protein with two TMDs and Hik19 as cytosolic protein which may function downstream of the membrane bound Hik33. Mutational analyses suggested that Hik33 is activated by reduced membrane fluidity at cold temperature (22 °C), allowing the

autophosphorylation of Hik33 and the subsequent transfer of a phosphate group to Hik19 and finally to Rer1 that in turn induce the desaturase gene expression and restores the membrane fluidity (Suzuki *et al.*, 2000). However, experimental data of Hik33 conformational changes triggered by reduced membrane fluidity is not yet available.

A more direct evidence of membrane fluidity triggering the activity of thermosensors is available for DesK (Cybulski *et al.*, 2002; Mariana and Diego, 2013). DesK is a transmembrane SK with five TMDs (TMD1-TMD5) and forms a thermosensing TCS with RR, DesR in *B. subtilis* (Cybulski *et al.*, 2010). At cold temperatures (< 30 °C), DesK activates the RR DesR, which induces transcription of *des* gene coding the $\Delta 5$ fatty acid desaturase ($\Delta 5$ -Des) (Aguilar *et al.*, 1998). $\Delta 5$ -Des introduces a double bond into saturated lipids producing unsaturated fatty acids (UFAs) (Aguilar *et al.*, 1998). To detect if DesK can sense membrane fluidity, the proportion of branched-chain fatty acids (BCFAs) were varied in the *B. subtilis* cell and the transcription of *des* gene was monitored (Cybulski *et al.*, 2002). BCFAs are essential for maintaining the appropriate fluidity and reduction in the amount of BCFAs result in ordered membrane lipids (or decrease fluidity). The experimental data showed decreased BCFAs induced activation of *des* transcription using a DesK/DesR-dependent mechanism (Cybulski *et al.*, 2002). Moreover, *B. subtilis lipA* mutant, which is unable to synthesize BCFA precursors exhibited increased transcription of the *des* gene (Martin *et al.*, 2009). These physiological and transcriptional data indicated that DesK senses membrane fluidity. As growth temperature decreases, the order of membrane lipid increases (or membrane fluidity decreases), leading to the activation of DesK and

inducing *des* gene expression that in turn increases the membrane disorder by producing UFAs and restoring membrane fluidity.

Experimental data showed that the temperature stimulus is received by the TMDs of DesK (Albanesi *et al.*, 2009) and the temperature sensing motif is held in TMD1 (Cybulski *et al.*, 2010). It is hypothesized that TMD1 detects the change in temperature and transfers the signal to TMD5 which is connected to the DesK catalytic core through a small linker (Cybulski *et al.*, 2010). Crystal structures of DesK catalytic core (DesKC), both in kinase (PDB ID 3EHJ) and phosphatase (PDB ID 3GIE) competent state, provided clear insight into the conformational changes that occur during the enzyme activity (Albanesi *et al.*, 2009). The structure displays a homodimeric structure where each monomer consists of an N-terminal antiparallel 2-helix hairpin that is connected to the C-terminal ATP binding (CA) domain (Albanesi *et al.*, 2009). The helical hairpins of two monomers dimerize to form four-helix bundle (4-HB) in DHp domain (Albanesi *et al.*, 2009). The structure of phosphatase competent state has a more compact and rigid conformation than the kinase-competent state. The phosphatase competent DesKC involves tighter interaction of DHp-CA domains and formation of two helix-coiled coil (2-HCC) that is predicted to extend all the way from the DHp domain through the entire linker region toward the N-terminus up to the C-terminus of TMD5 helix (Fig. 1.13) (Saita *et al.*, 2015). Models were generated by based on the X-ray structure of DesKC of both kinase and phosphatase competent state and the amino acid sequences of TMD5 which demonstrated that the stabilization or destabilization of the 2-HCC directly regulates the activity of DesK (Saita *et al.*, 2015). Stabilization of the 2-HCC allows the phosphatase state while preventing any autokinase activity, whereas the reverse happens upon 2-HCC destabilization. Based on this model and

biochemical studies it has been proposed that at cold temperature the membrane becomes thicker and more structured which imposes a stress on the 2-HCC and induces helical rotation resulting in the conversion of 2-HCC from the stabilised phosphatase state (Fig. 1.13A) into the destabilised autokinase-competent state (Fig. 1.13 B) (Saita *et al.*, 2015).

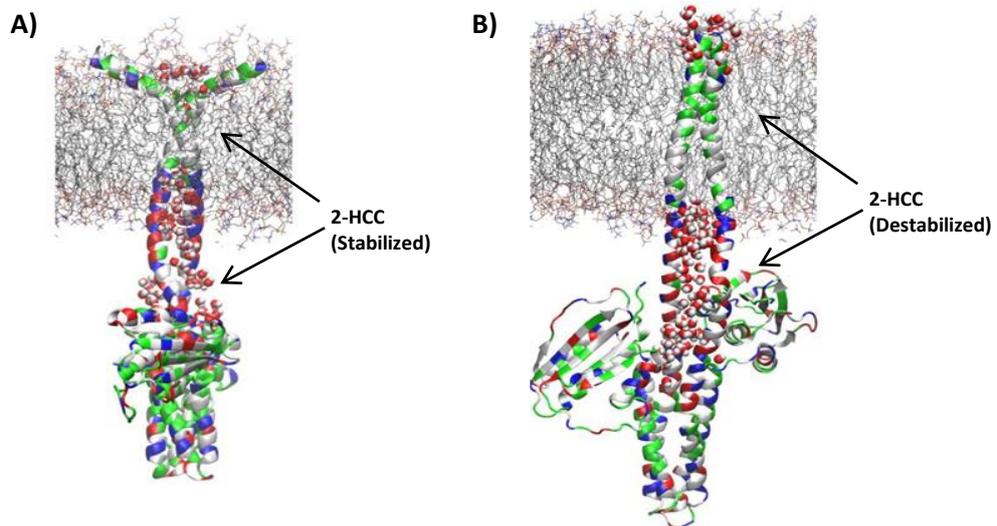


Figure 1.13 Models of DesK in the phosphatase and kinase state

A) Stabilized 2-HCC and more compact DHp-CA domain structure in the phosphatase state; B) Destabilized 2-HCC structure in the kinase state. After destabilization the 2-HCC gets highly hydrated. The water molecules are showed in ball representation. The figure has been adapted from Saita *et al.*, 2015.

In summary, these few studies described above give an indication of the importance of domain organisation and membrane fluidity for activation of the thermosensing SKs. However, with this limited experimental data it is difficult to get an insight into the general thermosensing mechanism in SKs.

1.2.5 TCS in Archaea

TCS genes are less abundant in archaea compared to bacteria and are found only in ~ 50 % of the archaeal genomes (Wuichet *et al.*, 2010). A total of 2,516 TCS genes have been identified in the 165 complete genome sequences of archaea and another 3,156 TCS genes are identified in the 90 draft genome sequences (Ulrich and Zhulin, 2010). TCS genes are found in 124 of the 190 genomes representing *Euryarchaeota*, five of the nine genomes representing *Thaumarchaeota*, one of 52 genomes representing *Crenarchaeota* and none in *Korarchaeota* and *Nanoarchaeota* (Ulrich and Zhulin, 2010; Ortet *et al.*, 2015). The archaeon encoding the highest number of TCS genes is a mesophilic methanogen from *Euryarchaeota* phylum, *Methanospirillum hungatei* encoding 78 SK and 53 RR genes (Ortet *et al.*, 2015).

Like bacteria the typical archaeal SKs are comprised of cytoplasmic dimerization DHP and catalytic CA domains. The majority of the archaeal SKs also contain a cytoplasmic sensor GAF and PAS or PAC domain that can bind a diverse set of intracellular small regulatory molecules including cAMP, cGMP, heme, FAD (Ashby, 2006) whereas few contain extracellular ligand binding domains between the two transmembrane regions, *i.e.* Cache, CHASE4, SBP (Ashby, 2006).

In contrast, almost half of the archaeal RRs are orphans containing only a single REC domain (Ashby, 2006; Galperin, 2010). In fact, in some archaea the stand-alone REC domain comprises 90% - 100% of all RRs, *e.g.* *Methanocella paludicola*, *Methanococcus aeolicus* (Jenal and Galperin, 2009; Galperin, 2010). Many RRs are comprised of a sensor GAF or PAS domain along with REC domain and very few contain a DNA-binding effector domain coupled with REC domain (Ashby, 2006).

Though TCS genes are frequently found in many sequenced archaeal genomes, only a single TCS has been biochemically characterized to-date, FilI/FilR system from *Methanosaeta harundinacea* (Li *et al.*, 2014). The FilI/FilR is the only TCS encoded by the archaeon and involves a transmembrane SK, FilI and two RRs, FilR1 and FilR2, of which FilR1 carries a REC domain combining with a DNA binding domain and FilR2 carries a single REC domain (Li *et al.*, 2014). Experimental data have shown that FilI can autophosphorylate in the presence of ATP and transfer the phosphoryl group to both FilR1 and FilR2 implying that these three proteins can make a TCS phosphorelay (Li *et al.*, 2014). Analyses on DNA binding ability of FilR1 determined that the regulator is associated with key genes involved in acetoclastic methanogenesis, which demonstrated that the FilI/FilR TCS regulates the methanogenesis in *M. harundinacea* (Li *et al.*, 2014).

While there is no report on temperature-responsive TCS in archaea, global transcriptomic and proteomic analyses in few psychrophilic methanogens, *i.e.* *M. burtonii* (Goodchild *et al.*, 2004b) and *Methanlobus psychrophilus* (Chen *et al.*, 2012), have detected abundance of TCSs at low compared to high temperature indicating a possibility of the presence of temperature-responsive TCS in psychrophilic archaea. However, because of the lack of experimental data there is very limited understanding of the functional properties of these differentially expressed TCSs in archaea.

1.3 TCS in *M. burtonii*

Genome sequence analyses showed an overrepresentation of TCS in *M. burtonii*, encoding 30 SK genes and 14 RR genes (Fig. 1.14) (Allen *et al.*, 2009; Barakat *et al.*, 2009, Ortet *et al.*, 2015) compared to many other methanogens. The number of TCS

genes possessed by an organism, is related to the complexity of the genome that is reflected by the number of regulatory genes present in the genome, its physiology and its ability to adapt to diverse environmental conditions, and thus is employed as a measure of adaptive potential of the organism, termed as IQ (Galperin, 2005; Ashby, 2006). *M. burtonii* has a high IQ (=124) compared with other methanogens, *i.e.* *Methanosarcina mazei* (IQ = 88), *M. barkeri* (IQ = 70) (Galperin *et al.*, 2010), which infers a greater requirement for complex internal regulation in *M. burtonii* (Ashby, 2006).

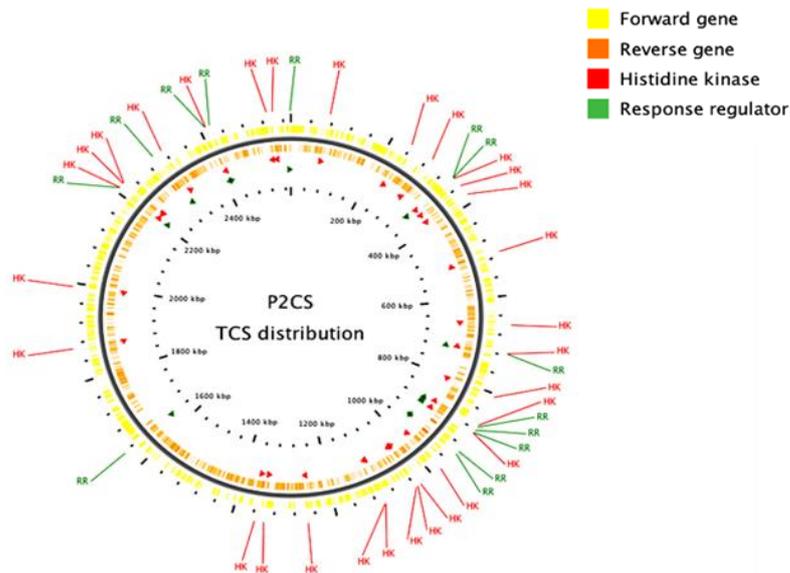


Figure 1.14 Overrepresentation of TCS genes in *M. burtonii* genome

M. burtonii genome size is ~2.6 Mbp and it contains total 44 TCS components. The figure is adopted from P2CS database, a database of TCS proteins (Ortet *et al.*, 2015).

M. burtonii encodes 30 SK genes of which, one is chemotaxis histidine kinase CheA, 25 are orthodox (or classic) and 4 are hybrid type. Eight of the orthodox SKs consist of a predicted extracytoplasmic sensor domain and are probably involved in environmental signalling (Galperin, 2005; Barakat *et al.*, 2009). In addition, most of the orthodox SKs (18 out of 25) contain PAS domain, which is a sensor domain located

intracellularly but may monitor changes in both external and internal environment including light, redox potential, oxygen, small ligands, and overall energy level of a cell (Taylor and Zhulin, 1999). So, the SKs containing PAS domain in *M. burtonii* may be involved in sensing cellular level of O₂, NO, CO and other molecules and this could be especially important for the strictly anaerobic *M. burtonii* which might have to cope with higher solubility of O₂ at low temperature (Allen *et al.*, 2009). Similarly, the four hybrid SKs also contain PAS domain in addition to a RR like REC domain that may be involved in more complex phosphorelay system (Allen *et al.*, 2009; Barakat *et al.*, 2009).

Of the 14 RRs encoded in *M. burtonii* genome, eight (57 %) consist of a standalone REC domain and two consist the sensor PAS domain along with REC domain (Allen *et al.*, 2009). These ten RRs that do not contain any output effector domain are likely to be involved in protein-protein interaction in *M. burtonii* (Allen *et al.*, 2009). Of the remaining four, one RR consists of an enzymatic effector domain, methyltransferase, which is predicted to be involved in chemotaxis pathway, one RR contains a DNA binding winged helix-turn-helix (HTH) domain and the remaining two have REC with unidentified domain organisation (Allen *et al.*, 2009; Galperin, 2010).

In archaeal species, presence of RR with DNA binding domain is uncommon (Galperin, 2010), therefore having RR with DNA binding HTH domain (locus tag-Mbur_0695), is a unique characteristic of *M. burtonii*. In addition, genome sequence analyses showed Mbur_0695 forms an operon like structure with an upstream SK gene, Mbur_0694, which is one of the eight orthodox SKs that have extracytoplasmic sensor domain (Allen *et al.*, 2009). Transcriptomic analyses further confirmed that these two genes, Mbur_0695 and Mbur_0694, are expressed in a single operon (Campanaro *et al.*,

2011; Taha *et al.*, 2016). Comparative proteomic analyses at low vs. high temperature identified higher abundance of the RR, Mbur_0695 at low (4 °C) temperature (Goodchild *et al.*, 2004b). Subsequent transcriptomic analyses also showed abundance of Mbur_0695 transcript (Campanaro *et al.*, 2011) and both Mbur_0695 (RR) and Mbur_0694 (SK) transcripts (Taha *et al.*, 2016) at 4 °C in comparison with 23 °C indicates low temperature regulation of both of these RR and SK. These genomics, proteomics and transcriptomics data strongly suggest that the RR, Mbur_0695 and SK, Mbur_0694 may form a temperature responsive TCS in *M. burtonii* and play a critical role in cold adaptation. In view of the cold temperature regulation, Mbur_0695 is referred as the low temperature responsive response regulator- LtrR and Mbur_0694 is referred as low temperature responsive sensor kinase- LtrK.

1.4 Aim of the project

The psychrophilic archaea have evolved specific traits to adapt to permanently cold environments in order to survive and proliferate and thus constitute a significant proportion of biomass in the earth's low-temperature ecosystems. *M. burtonii* is the most intensively studied psychrophilic archaea. Comparative proteomics and transcriptomics analyses have provided valuable insight into the processes important for cold adaptation in *M. burtonii* and have identified a particular TCS, LtrK/LtrR that may play a vital role in cold adaptation.

As described in section 1.2, TCSs are the most prevalent signalling cascades found in microorganisms and are frequently found to function as a cellular thermometer by sensing change in environmental temperature especially in bacterial pathogens. TCSs are extensively studied in bacteria while very little is known about archaeal TCSs. There

is just a single biochemical study of TCS from archaea, FilI/FilR system from mesophilic *M. harundinacea* (Li *et al.*, 2014). Therefore, given the lack of experimental data regarding TCS in archaea and specific importance of TCS in *M. burtonii* cold adaptation, there is a compelling reason to experimentally characterize the SK, LtrK and RR, LtrR from *M. burtonii*.

The aim of the project was to characterize LtrK (SK) and LtrR (RR) of *M. burtonii* and assess their capacity to function as a thermosensor to enable temperature responsive gene regulation in *M. burtonii*. The aim of this project was achieved through following key steps:

- Bioinformatics analyses of the LtrK and LtrR genes to get an insight into the protein domain organisation, structural properties and sequence similarities/dissimilarities with bacterial SKs and RRs (described in Chapter 2)
- Expression and purification of both recombinant LtrK and LtrR and their biophysical characterization to understand the structural properties of the purified proteins (described in Chapter 3 and 4)
- Assessment of kinase activity of LtrK (autophosphorylation of LtrK and phosphotransfer to LtrR) and dephosphorylation of LtrR (described in Chapter 5).
- Mutational analyses of both LtrK and LtrR to assess if they possess structural properties common to bacterial SKs and RRs (described in Chapter 5).
- Analyses of the temperature dependency of activity and stability of LtrK and LtrR in order to assess their thermosensing capacity (described in Chapter 6).

- Determination of the activation thermodynamic parameters of LtrK stability and activity to assess if the enzyme possesses properties common to other cold-adapted enzymes (described in Chapter 7).

Finally, an overview of the main conclusion is presented in Chapter 8 with some suggestions for the future research.

1.5 References

Aguilar, P.S. *et al.* (1998) A *Bacillus subtilis* gene induced by cold shock encodes a membrane phospholipid desaturase. *J. Bacteriol.* 180: 2194-2200.

Aguilar, P.S. *et al.* (2001) Molecular basis of thermosensing: a two-component signal transduction thermometer in *Bacillus subtilis*. *EMBO J.* 20: 1681-1691.

Albanesi, D. *et al.* (2004) The membrane fluidity sensor DesK of *Bacillus subtilis* controls the signal decay of its cognate response regulator. *J. Bacteriol.* 186: 2655-2663.

Albanesi, D. *et al.* (2009) Structural plasticity and catalysis regulation of a thermosensor histidine kinase. *Proc. Natl. Acad. Sci. USA.* 106: 16185-16190.

Allen, M.A. *et al.* (2009) The genome sequence of the psychrophilic archaeon, *Methanococoides burtonii*: the role of genome evolution in cold adaptation. *ISME J.* 3: 1012-1035.

Ashby, M.K. (2006) Distribution, structure and diversity of “bacterial” genes encoding two-component proteins in the *Euryarchaeota*. *Archaea* 2: 11-30.

Barakat, M. *et al.* (2009) P2CS: a two-component system resource for prokaryotic signal transduction research. *BMC Genomics* 10: 1.

Beier, D., and Gross, R. (2006) Regulation of bacterial virulence by two-component systems. *Curr. Opin. Microbiol.* 9: 143-152.

Bhate, M.P. *et al.* (2015) Signal transduction in histidine kinases: insights from new structures. *Structure* 23: 981-994.

Blanco, A.G. *et al.* (2002) Tandem DNA recognition by PhoB, a two-component signal transduction transcriptional activator. *Structure* 10: 701-713.

- Bourret, R.B. (2010) Receiver domain structure and function in response regulator proteins. *Curr. Opin. Microbiol.* 13: 142-149.
- Braun, Y. *et al.* (2008) Impact of temperature on the regulation of coronatine biosynthesis in *Pseudomonas syringae*. *Pseudomonas syringae Pathovars and Related Pathogens*: 159-165.
- Braun, Y. *et al.* (2007) A temperature-sensing histidine kinase—function, genetics, and membrane topology. *Methods Enzymol.* 423: 222-249.
- Buckler, D.R. *et al.* (2002) Evidence of intradomain and interdomain flexibility in an OmpR/PhoB homolog from *Thermotoga maritima*. *Structure* 10: 153-164.
- Cai, S.J., and Inouye, M. (2002) EnvZ-OmpR interaction and osmoregulation in *Escherichia coli*. *J. Biol. Chem.* 277: 24155-24161.
- Campanaro, S. *et al.* (2011) Temperature-dependent global gene expression in the Antarctic archaeon *Methanococcoides burtonii*. *Environ. Microbiol.* 13: 2018-2038.
- Casino, P. *et al.* (2009) Structural insight into partner specificity and phosphoryl transfer in two-component signal transduction. *Cell* 139: 325-336.
- Castelli, M.a.E. *et al.* (2000) The phosphatase activity is the target for Mg²⁺ regulation of the sensor protein PhoQ in *Salmonella*. *J. Biol. Chem.* 275: 22948-22954.
- Cavicchioli, R. (2006) Cold-adapted archaea. *Nat. Rev. Microbiol.* 4: 331-343.
- Cavicchioli, R. (2011) Archaea—timeline of the third domain. *Nat. Rev. Microbiol.* 9: 51-61.
- Cavicchioli, R. *et al.* (2007) Metagenomic studies reveal the critical and wide ranging ecological importance of uncultivated archaea: the role of ammonia oxidizers. *BioEssays* 29: 11-14.
- Chai, W., and Stewart, V. (1998) NasR, a novel RNA-binding protein, mediates nitrate-responsive transcription antitermination of the *Klebsiella oxytoca* M5al nasF operon leader in vitro. *J. Mol. Biol.* 283: 339-351.
- Chakraborty, S. *et al.* (2010) Temperature and Mg²⁺ sensing by a novel PhoP-PhoQ two-component system for regulation of virulence in *Edwardsiella tarda*. *J. Biol. Chem.* 285: 38876-38888.
- Chan, Y.C. *et al.* (2007) Microarray-based characterization of the *Listeria monocytogenes* cold regulon in log-and stationary-phase cells. *Appl. Environ. Microbiol.* 73: 6484-6498.

- Chan, Y.C. *et al.* (2008) Contributions of two-component regulatory systems, alternative factors, and negative regulators to *Listeria monocytogenes* cold adaptation and cold growth. *J. Food Prot.* 71: 420-425.
- Chen, Z. *et al.* (2012) The genome and transcriptome of a newly described psychrophilic archaeon, *Methanobolus psychrophilus* R15, reveal its cold adaptive characteristics. *Environ. Microbiol. Rep.* 4: 633-641.
- Cotter, P.A., and Jones, A.M. (2003) Phosphorelay control of virulence gene expression in *Bordetella*. *Trends Microbiol.* 11: 367-373.
- Cybulski, L.E. *et al.* (2010) Membrane thickness cue for cold sensing in a bacterium. *Curr. Biol.* 20: 1539-1544.
- Cybulski, L.E. *et al.* (2002) Mechanism of membrane fluidity optimization: isothermal control of the *Bacillus subtilis* acyl-lipid desaturase. *Mol. Microbiol.* 45: 1379-1388.
- Dahlsten, E. *et al.* (2014) The cold-induced two-component system CBO0366/CBO0365 regulates metabolic pathways with novel roles in group I *Clostridium botulinum* ATCC 3502 cold tolerance. *Appl. Environ. Microbiol.* 80: 306-319.
- De Francisci, D. *et al.* (2011) The RNA polymerase subunits E/F from the Antarctic archaeon *Methanococcoides burtonii* bind to specific species of mRNA. *Environ Microbiol* 13: 2039-2055.
- Derman, Y. *et al.* (2013) The two-component system CBO2306/CBO2307 is important for cold adaptation of *Clostridium botulinum* ATCC 3502. *Int. J. Food Microbiol.* 167: 87-91.
- Diomandé, S.E. *et al.* (2014) The CasKR two-component system is required for the growth of mesophilic and psychrotolerant *Bacillus cereus* strains at low temperatures. *Appl. Environ. Microbiol.* 80: 2493-2503.
- Dutta, R. *et al.* (1999) Histidine kinases: diversity of domain organization. *Mol. Microbiol.* 34: 633-640.
- Eswaramoorthy, P. *et al.* (2010) The threshold level of the sensor histidine kinase KinA governs entry into sporulation in *Bacillus subtilis*. *J. Bacteriol.* 192: 3870-3882.
- Ferris, H.U. *et al.* (2014) Crystallographic snapshot of the *Escherichia coli* EnvZ histidine kinase in an active conformation. *J. Struct. Biol.* 186: 376-379.
- Foussard, M. *et al.* (2001) The molecular puzzle of two-component signaling cascades. *Microbes Infect.* 3: 417-424.
- Franzmann, P. *et al.* (1992) A methanogenic archaeon from Ace Lake, Antarctica: *Methanococcoides burtonii* sp. nov. *Syst. Appl. Microbiol.* 15: 573-581.

- Galperin, M.Y. (2005) A census of membrane-bound and intracellular signal transduction proteins in bacteria: bacterial IQ, extroverts and introverts. *BMC Microbiol.* 5: 35.
- Galperin, M.Y. (2006) Structural classification of bacterial response regulators: diversity of output domains and domain combinations. *J. Bacteriol.* 188: 4169-4182.
- Galperin, M.Y. (2010) Diversity of structure and function of response regulator output domains. *Curr. Opin. Microbiol.* 13: 150-159.
- Galperin, M.Y. *et al.* (2010) Interplay of heritage and habitat in the distribution of bacterial signal transduction systems. *Mol. Biosyst.* 6: 721-728.
- Gao, R., and Stock, A.M. (2009) Biological insights from structures of two-component proteins. *Annu. Rev. Microbiol.* 63: 133.
- Gao, R., and Stock, A.M. (2010) Molecular strategies for phosphorylation-mediated regulation of response regulator activity. *Curr. Opin. Microbiol.* 13: 160-167.
- Gao, R. *et al.* (2007) Bacterial response regulators: versatile regulatory strategies from common domains. *Trends Biochem. Sci.* 32: 225-234.
- Garcia, J.-L. *et al.* (2000) Taxonomic, phylogenetic, and ecological diversity of methanogenic Archaea. *Anaerobe* 6: 205-226.
- Goodchild, A. *et al.* (2004a) Biology of the cold adapted archaeon, *Methanococoides burtonii* determined by proteomics using liquid chromatography-tandem mass spectrometry. *J. Proteome Res.* 3: 1164-1176.
- Goodchild, A. *et al.* (2005) Cold adaptation of the Antarctic archaeon, *Methanococoides burtonii* assessed by proteomics using ICAT. *J. Proteome Res.* 4: 473-480.
- Goodchild, A. *et al.* (2004b) A proteomic determination of cold adaptation in the Antarctic archaeon, *Methanococoides burtonii*. *Mol. Microbiol.* 53: 309-321.
- Gotoh, Y. *et al.* (2010) Two-component signal transduction as potential drug targets in pathogenic bacteria. *Curr. Opin. Microbiol.* 13: 232-239.
- Guijarro, J.A. *et al.* (2015) Temperature-dependent expression of virulence genes in fish-pathogenic bacteria. *Frontiers in microbiology* 6.
- Hesami, S. *et al.* (2011) Identification of cold-temperature-regulated genes in *Flavobacterium psychrophilum*. *Appl. Environ. Microbiol.* 77: 1593-1600.
- Hess, J.F. *et al.* (1988) Phosphorylation of three proteins in the signaling pathway of bacterial chemotaxis. *Cell* 53: 79-87.

- Hoch, J.A. (2000) Two-component and phosphorelay signal transduction. *Curr. Opin. Microbiol.* 3: 165-170.
- Huynh, T.N., and Stewart, V. (2011) Negative control in two-component signal transduction by transmitter phosphatase activity. *Mol. Microbiol.* 82: 275-286.
- Jenal, U., and Galperin, M.Y. (2009) Single domain response regulators: molecular switches with emerging roles in cell organization and dynamics. *Curr. Opin. Microbiol.* 12: 152-160.
- Jin, S. *et al.* (1993) The regulatory VirA protein of *Agrobacterium tumefaciens* does not function at elevated temperatures. *J. Bacteriol.* 175: 6830-6835.
- Kamberov, E.S. *et al.* (1994) Effect of mutations in *Escherichia coli* glnL (ntrB), encoding nitrogen regulator II (NRII or NtrB), on the phosphatase activity involved in bacterial nitrogen regulation. *J. Biol. Chem.* 269: 28294-28299.
- Khorchid, A. *et al.* (2005) Structural characterization of *Escherichia coli* sensor histidine kinase EnvZ: the periplasmic C-terminal core domain is critical for homodimerization. *Biochem. J.* 385: 255-264.
- Klinkert, B., and Narberhaus, F. (2009) Microbial thermosensors. *Cellular and molecular life sciences* 66: 2661-2676.
- Klumpp, S., and Krieglstein, J. (2002) Phosphorylation and dephosphorylation of histidine residues in proteins. *Eur. J. Biochem.* 269: 1067-1071.
- Laybourn-Parry, J., and Bell, E.M. (2014) Ace Lake: three decades of research on a meromictic, Antarctic lake. *Polar Biol.* 37: 1685-1699.
- Lee, S.-Y. *et al.* (2001) Crystal structure of activated CheY comparison with other activated receiver domains. *J. Biol. Chem.* 276: 16425-16431.
- Leininger, S. *et al.* (2006) Archaea predominate among ammonia-oxidizing prokaryotes in soils. *Nature* 442: 806-809.
- Li, J. *et al.* (2014) Characterization of an archaeal two-component system that regulates methanogenesis in *Methanosaeta harundinacea*. *PloS one* 9: e95502. doi:95510.91371/journal.pone.0095502.
- Lindström, M. *et al.* (2012) Involvement of two-component system CBO0366/CBO0365 in the cold shock response and growth of group I (proteolytic) *Clostridium botulinum* ATCC 3502 at low temperatures. *Appl. Environ. Microbiol.* 78: 5466-5470.
- Liu, W., and Hulett, F.M. (1997) *Bacillus subtilis* PhoP binds to the phoB tandem promoter exclusively within the phosphate starvation-inducible promoter. *J. Bacteriol.* 179: 6302-6310.

Mariana, M., and Diego, d.M. (2013) Regulation of *Bacillus subtilis* DesK thermosensor by lipids. *Biochem. J.* 451: 269-275.

Martin, N. *et al.* (2009) A lipA (yutB) mutant, encoding lipoic acid synthase, provides insight into the interplay between branched-chain and unsaturated fatty acid biosynthesis in *Bacillus subtilis*. *J. Bacteriol.* 191: 7447-7455.

Martínez-Hackert, E., and Stock, A.M. (1997) The DNA-binding domain of OmpR: crystal structures of a winged helix transcription factor. *Structure* 5: 109-124.

Mascher, G. *et al.* (2014) The CLO3403/CLO3404 two-component system of *Clostridium botulinum* E1 Beluga is important for cold shock response and growth at low temperatures. *Appl. Environ. Microbiol.* 80: 399-407.

Mascher, T. *et al.* (2006) Stimulus perception in bacterial signal-transducing histidine kinases. *Microbiol. Mol. Biol. Rev.* 70: 910-938.

Ninfa, E.G. *et al.* (1993) Mechanism of autophosphorylation of *Escherichia coli* nitrogen regulator II (NR II or NtrB): trans-phosphorylation between subunits. *J. Bacteriol.* 175: 7024-7032.

Nixon, B.T. *et al.* (1986) Two-component regulatory systems responsive to environmental stimuli share strongly conserved domains with the nitrogen assimilation regulatory genes ntrB and ntrC. *Proc. Natl. Acad. Sci. USA.* 83: 7850-7854.

Njimon, I., and Lamparter, T. (2011) Temperature effects on *Agrobacterium* phytochrome Agp1. *PloS one* 6: e25977.

Ortet, P. *et al.* (2015) P2CS: updates of the prokaryotic two-component systems database. *Nucleic Acids Res.* 43: D536-D541.

Palonen, E. *et al.* (2011) Expression of signal transduction system encoding genes of *Yersinia pseudotuberculosis* IP32953 at 28°C and 3°C. *PLoS ONE* 6: e25063. doi:25010.21371/journal.pone.0025063.

Parkinson, J.S., and Kofoed, E.C. (1992) Communication modules in bacterial signaling proteins. *Annu. Rev. Genet.* 26: 71-112.

Perry, J. *et al.* (2011) Receptor domains of two-component signal transduction systems. *Molecular BioSystems* 7: 1388-1398.

Pilak, O. *et al.* (2011) Chaperonins from an Antarctic archaeon are predominantly monomeric: crystal structure of an open state monomer. *Environ Microbiol* 13: 2232-2249.

Pöntinen, A. *et al.* (2015) Two-component-system histidine kinases involved in growth of *Listeria monocytogenes* EGD-e at Low Temperatures. *Appl. Environ. Microbiol.* 81: 3994-4004.

- Ray, M. *et al.* (1994) Phosphorylation of membrane proteins in response to temperature in an Antarctic *Pseudomonas syringae*. *Microbiology* 140: 3217-3223.
- Reid, I. *et al.* (2006) Terrestrial models for extraterrestrial life: methanogens and halophiles at Martian temperatures. *International Journal of Astrobiology* 5: 89-97.
- Robinson, V.L. *et al.* (2003) Structural analysis of the domain interface in DrrB, a response regulator of the OmpR/PhoB subfamily. *J. Bacteriol.* 185: 4186-4194.
- Saita, E. *et al.* (2015) A coiled coil switch mediates cold sensing by the thermosensory protein DesK. *Mol. Microbiol.* 98: 258-271.
- Saunders, N.F. *et al.* (2003) Mechanisms of thermal adaptation revealed from the genomes of the Antarctic Archaea *Methanogenium frigidum* and *Methanococoides burtonii*. *Genome Res.* 13: 1580-1588.
- Sengupta, P., and Garrity, P. (2013) Sensing temperature. *Curr. Biol.* 23: R304-R307.
- Siddiqui, K. *et al.* (2002) Thermodynamic activation properties of elongation factor 2 (EF-2) proteins from psychrotolerant and thermophilic Archaea. *Extremophiles* 6: 143-150.
- Simms, S. *et al.* (1985) Multiple forms of the CheB methylsterase in bacterial chemosensing. *J. Biol. Chem.* 260: 10161-10168.
- Smirnova, A. *et al.* (2001) Thermoregulated expression of virulence factors in plant-associated bacteria. *Arch. Microbiol.* 176: 393-399.
- Smirnova, A.V., and Ullrich, M.S. (2004) Topological and deletion analysis of CorS, a *Pseudomonas syringae* sensor kinase. *Microbiology* 150: 2715-2726.
- Srinivasa Rao, P. *et al.* (2004) Use of proteomics to identify novel virulence determinants that are required for *Edwardsiella tarda* pathogenesis. *Mol. Microbiol.* 53: 573-586.
- Starliper, C.E. (2011) Bacterial coldwater disease of fishes caused by *Flavobacterium psychrophilum*. *J. Adv. Res.* 2: 97-108.
- Steele, K.H. *et al.* (2012) Characterization of a ferrous iron-responsive two-component system in nontypeable *Haemophilus influenzae*. *J. Bacteriol.* 194: 6162-6173.
- Steinmann, R., and Dersch, P. (2013) Thermosensing to adjust bacterial virulence in a fluctuating environment. *Future Microbiol.* 8: 85-105.
- Stock, A.M. *et al.* (2000) Two-component signal transduction. *Annu. Rev. Biochem.* 69: 183-215.

Stockbauer, K.E. *et al.* (2001) Identification and characterization of BipA, a *Bordetella* Bvg-intermediate phase protein. *Mol. Microbiol.* 39: 65-78.

Suzuki, I. *et al.* (2000) The pathway for perception and transduction of low-temperature signals in *Synechocystis*. *EMBO J.* 19: 1327-1334.

Swanson, R.V. *et al.* (1993) Intermolecular complementation of the kinase activity of CheA. *Mol. Microbiol.* 8: 435-441.

Taha *et al.* (2016) Single TRAM domain RNA-binding proteins in Archaea: functional insight from Ctr3 from the Antarctic methanogen *Methanococcoides burtonii*. *Environ. Microbiol.* doi: 10.1111/1462-2920.13229

Taylor, B.L., and Zhulin, I.B. (1999) PAS domains: internal sensors of oxygen, redox potential, and light. *Microbiol. Mol. Biol. Rev.* 63: 479-506.

Thomas, T., and Cavicchioli, R. (1998) Archaeal cold-adapted proteins: structural and evolutionary analysis of the elongation factor 2 proteins from psychrophilic, mesophilic and thermophilic methanogens 1. *FEBS Lett.* 439: 281-286.

Thomas, T., and Cavicchioli, R. (2000) Effect of temperature on stability and activity of elongation factor 2 proteins from Antarctic and thermophilic methanogens. *J. Bacteriol.* 182: 1328-1332.

Thomas, T., and Cavicchioli, R. (2002) Cold adaptation of archaeal elongation factor 2 (EF-2) proteins. *Curr. Protein Pept. Sci.* 3: 223-230.

Thomas, T. *et al.* (2001) Effects of ribosomes and intracellular solutes on activities and stabilities of elongation factor 2 proteins from psychrotolerant and thermophilic methanogens. *J. Bacteriol.* 183: 1974-1982.

Thomason, P., and Kay, R. (2000) Eukaryotic signal transduction via histidine-aspartate phosphorelay. *J. Cell Sci.* 113: 3141-3150.

Ullrich, M. *et al.* (1995) A modified two-component regulatory system is involved in temperature-dependent biosynthesis of the *Pseudomonas syringae* phytotoxin coronatine. *J. Bacteriol.* 177: 6160-6169.

Ulrich, L.E., and Zhulin, I.B. (2010) The MiST2 database: a comprehensive genomics resource on microbial signal transduction. *Nucleic Acids Res.:* D401-407.

Varughese, K.I. (2005) Conformational changes of Spo0F along the phosphotransfer pathway. *J. Bacteriol.* 187: 8221-8227.

Vigh, L. *et al.* (1998) Does the membrane's physical state control the expression of heat shock and other genes? *Trends Biochem. Sci.* 23: 369-374.

- Wang, C. *et al.* (2013) Mechanistic insights revealed by the crystal structure of a histidine kinase with signal transducer and sensor domains. *PLoS Biol.* 11: e1001493.
- West, A.H., and Stock, A.M. (2001) Histidine kinases and response regulator proteins in two-component signaling systems. *Trends Biochem. Sci.* 26: 369-376.
- Williams, C.L. *et al.* (2005) BvgA functions as both an activator and a repressor to control Bvgi phase expression of bipA in *Bordetella pertussis*. *Mol. Microbiol.* 56: 175-188.
- Williams, T.J. *et al.* (2010) Global proteomic analysis of the insoluble, soluble, and supernatant fractions of the psychrophilic archaeon *Methanococcoides burtonii* part I: the effect of growth temperature. *J. Proteome Res.* 9: 640-652.
- Williams, T.J. *et al.* (2011) Defining the response of a microorganism to temperatures that span its complete growth temperature range (− 2 °C to 28 °C) using multiplex quantitative proteomics. *Environ. Microbiol.* 13: 2186-2203.
- Woese, C.R. *et al.* (1990) Towards a natural system of organisms: proposal for the domains Archaea, Bacteria, and Eucarya. *Proc. Natl. Acad. Sci. USA.* 87: 4576-4579.
- Wolanin, P.M. *et al.* (2002) Histidine protein kinases: key signal transducers outside the animal kingdom. *Genome Biol.* 3: 3013.
- Wuichet, K. *et al.* (2010) Evolution and phyletic distribution of two-component signal transduction systems. *Curr. Opin. Microbiol.* 13: 219-225.
- Yang, Y., and Inouye, M. (1991) Intermolecular complementation between two defective mutant signal-transducing receptors of *Escherichia coli*. *Proc. Natl. Acad. Sci. USA.* 88: 11057-11061.
- Zhou, Z.-j. *et al.* (2015) *Edwardsiella tarda* Sip1: a serum-induced zinc metalloprotease that is essential to serum resistance and host infection. *Vet. Microbiol.* 177: 332-340.
- Zhu, Y. *et al.* (2000) Phosphatase activity of histidine kinase EnvZ without kinase catalytic domain. *Proc. Natl. Acad. Sci. USA.* 97: 7808-7813.

Chapter 2

Bioinformatics analyses of the sensor kinase, LtrK and response regulator, LtrR

2.1 Abstract

In silico analyses were performed on the putative TCS SK and RR of *M. burtonii*, LtrK and LtrR respectively, using the amino acid sequence. The analyses revealed that LtrK is a large (67 kDa) protein with two transmembrane domains and contains a distinct ligand binding CHASE domain in the periplasmic area and typical SK domains, HisKA and HATPase in the cytoplasmic area while LtrR is relatively small protein (19 kDa) that possesses domains typical of a RR including a helix turn helix (HTH) domain and a receiver domain (REC). The HTH domain is at the N-terminus of LtrR, whereas the vast majority of RRs in *Bacteria* have the HTH domain at the C-terminus. Further analyses on archaeal RR domain organisation revealed that all methanogens have evolved RRs that have distinct domain architecture to bacteria. Multiple sequence alignment of LtrK with bacterial SKs identified the conserved H, N, F, G1, G2 and G3 blocks in LtrK that are diagnostic of TCS SKs. The 3D structure homology model of both LtrK and LtrR gave an insight into the proteins tertiary structure including the possible active site residues. And finally the evolutionary relationship analyses indicated that the closely related SKs and RRs to LtrK and LtrR are from methanogens and both proteins (LtrK and LtrR) have evolved together. Overall the data suggest that LtrK and LtrR possess typical characteristics of a TCS SK and RR, respectively with some unique features.

2.2 Introduction

In silico analyses of protein sequences using bioinformatics have proven useful for predicting the functional and structural properties of proteins, including the identification of domains and their organisation, the presence of signature motifs and the presence of functionally or structurally important amino acid residues. By performing bioinformatics analyses putative SK and RR genes can be identified, which can subsequently be validated using biochemical experiments. For example, the FilI/FilR, from *M. harundinacea* was initially identified as a putative TCS by *in silico* analyses before being experimentally characterized (Li *et al.*, 2014).

Identification of the domains is the first step to recognize a putative SK and RR from the protein sequence. Typically a TCS SK is comprised of an extracellular domain combining with a cytoplasmic dimerization (HisKA) and catalytic domains (HATPase_c) whereas a RR is comprised of a receiver (REC) domain coupled with an effector domain (Gao and Stock, 2009). Key features of SKs include the presence of conserved residues named as H, N, G1, G2, G3 and F block (Parkinson and Kofoid, 1992) and presence of transmembrane domains. However, few bacterial SKs are only cytoplasmic proteins that do not contain any transmembrane segments, *i.e.* CheA from *E. coli*, KinA from *B. subtilis* (Mascher *et al.*, 2006).

The SKs and RRs also have distinct structural features. For example, SKs function as a dimer and form a helical bundle using its cytoplasmic dimerization domain (Bhate *et al.*, 2015) whereas RRs structure shows ($\alpha\beta$)₅ topology (Bourret, 2010) (described in Chapter 1). For the identification of the structural features 3D homology modelling using the protein sequence has proven to be useful. For instance, 3D homology model of EF-2 from *M. burtonii* revealed structural features including fewer

salt bridges and less packed hydrophobic core that are predicted to be important for its structural flexibility (Thomas and Cavicchioli, 1998), whereas 3D structure model of Ctr3 gave valuable insight into the possible key amino acid residues involved in RNA binding (Taha *et al.*, 2016).

TCS genes are abundant in bacterial genomes and for this reason most of the current experimental data on SKs and RRs are from *Bacteria* while only one TCS, FilI/FilR of *M. harundinacea*, had been studied from *Archaea* (Li *et al.*, 2014). However, no biochemical analyses have been performed on TCS from psychrophiles (*Bacteria* or *Archaea*). The absence of any well-documented study on specific psychrophilic TCS promoted the bioinformatics analyses to be performed on the predicted SK (LtrK) and RR (LtrR) of psychrophilic *M. burtonii* to get a general view of the protein's physiochemical properties, domain organisation, tertiary structure and evolution. The analyses were also performed in order to complement the subsequent experimental studies of both these proteins that are described in this dissertation. This chapter describes the *in silico* analyses of both proteins including comparing and contrasting the protein sequences with bacterial TCSs, assessment of the phylogenetic relationship with bacterial and archaeal TCSs and analysis of the structural features using 3D homology modelling.

2.3 Experimental procedures

Table 2.1 lists the tools used for the bioinformatics analyses of both LtrK and LtrR proteins:

Table 2.1: List of bioinformatics tools used for protein analyses

Tools	Purpose	Reference
ProtParam	Physico-chemical properties	(Walker, 2005)
Pfam database	Identification of domains	(Finn <i>et al.</i> , 2014)
ClustalW	Multiple sequence alignment	(Goujon <i>et al.</i> , 2010)
TMHMM Server v. 2.0	Transmembrane domain prediction	(Sonnhammer <i>et al.</i> , 1998)
I-TASSER	3D structural modelling	(Zhang, 2008)
Mega v 6.0	Phylogenetic tree construction	(Tamura <i>et al.</i> , 2013)
IMG	Archaeal genome sequence analyses	(Markowitz <i>et al.</i> , 2012)
PyMOL	3D structure model analyses	Molecular Graphics System, Version 1.5 Schrödinger, LLC

All the analyses except phylogenetic tree construction were performed using the amino acid sequences as shown below:

Sequence for LtrK:

```
MFATKSRWNVIIVVMVTMLFISSILFLGIRANTEVEDLFVEMFTREQTSQAQQISTGITTFLNEKITMLEIISRNHNSIPDDNFN
TIFSIVYNESEGFHAIEYINSTGIIVSGYPEENVVPGYDLYEKDKAAVFEQIKETGETHITNPLTTLEGKLAIFVWIPFYLENGS
FEGAIVAVIEIDEITKQTIKTEYESGYIYLIDDNAKLLYDSSDDHLVQGNYFNIINEPDHRRLLDIIGMQTEGLSGSGQFSERNT
NGFLEEKIISYVPVNWYNQQWSVGVVTPWYVGS LIQSVYVKQGLFAMISIAFILFISAFIALILLSWNKTL EDEVKSKTSEL
ERSNDY LQNANKK LKELDR LKND FVSMVSHELKTPLTAMKTSSEFLRESECSREIKEEMLDLIIRNIDRQARMVDDLLDISR
IESGKMKFTPEDVNIKEIIEISLHNVTKHAKDKSIKIMVNCPPDDVPAISTDKDKLIRIFVNLLTNAIKFTPEEGEVTVIVEDHED
YLQTSIKDNGIGILEEKRDKIFDKFYQVDSTATRKAGGTGLGLAIKGIIDGQGGKIYLESETGKGSTFTFRLPKELKEDDFTE
IEGT
```

Sequence for LtrR:

```
MEREKLEKDIMRLLLEQQPEISAKGIADRLSVSEDIVTRTIESLSDRHKVLIVDDEPDAVIATKRALEADGYNVIEANNSTM
AFDALKSDIPDVILLDVMMPDMDGFEVCRRLKEEPLYENIPVIMLTAKGEINDKVEGLDIGADDYMTKPFNLKELKARIKT
VLRRTQD
```

To construct the phylogenetic tree for LtrK, bacterial SKs with ≥ 40 % and archaeal SKs with ≥ 38 % identity and for LtrR bacterial RRs with ≥ 50 % and archaeal RRs with ≥ 46 % identity were selected. The sequences were identified from Blastp (Madden, 2013) search, using cytoplasmic domain amino acid sequence of LtrK (C-tr 224 amino acids) or receiver domain (REC) sequence of LtrR (C-tr 120 amino acids) as query. The selected protein sequences are given in Appendix 1. The protein (LtrK or LtrR) with the selected sequences (SK or RR) were subsequently aligned using the alignment tool MUSCLE of Mega software (Edgar, 2004). With the alignment file the phylogenetic tree was constructed using Neighbour-joining method where the site coverage cut-off of and bootstrap values were set to 80 % and 10,000 respectively.

2.4 Results:

2.4.1 Protein sequence analyses of LtrK and LtrR

The physical and chemical properties of LtrK and LtrR were predicted from their sequence using ProtParam tool (Table 2.2). The data indicated that LtrK is a large protein with 67.3 kDa molecular weight compared to LtrR that has a molecular weight of 19.2 kDa.

Table 2.2: Physical and chemical parameters of LtrK and LtrR using ProtParam

Protein	Length of amino acids	Molecular weight (kDa)	pI	Extinction co-efficient (M ⁻¹ cm ⁻¹)
LtrK	592	67.3	4.8	61435
LtrR	170	19.2	4.6	4470

Amino acid composition of each protein was also analysed using the same tool (Table 2.3, 2.4).

Table 2.3 Amino acid composition of LtrK

Amino acid	Number of residues	Percentage (%)
Ala (A)	23	3.9
Arg (R)	19	3.2
Asn (N)	32	5.4
Asp (D)	37	6.2
Cys (C)	2	0.3
Gln (Q)	18	3.0
Glu (E)	56	9.5
Gly (G)	35	5.9
His (H)	9	1.5
Ile (I)	68	11.5
Leu (L)	49	8.3
Lys (K)	41	6.9
Met (M)	13	2.2
Phe (F)	31	5.2
Pro (P)	14	2.4
Ser (S)	43	7.3
Thr (T)	42	7.1

Trp (W)	6	1.0
Tyr (Y)	19	3.2
Val (V)	35	5.9

Table 2.4 Amino acid composition of LtrR

Amino acid	Number of residues	Percentage (%)
Ala (A)	12	7.1
Arg (R)	11	6.5
Asn (N)	6	3.5
Asp (D)	19	11.2
Cys (C)	1	0.6
Gln (Q)	3	1.8
Glu (E)	17	10.0
Gly (G)	7	4.1
His (H)	1	0.6
Ile (I)	15	8.8
Leu (L)	17	10.0
Lys (K)	13	7.6
Met (M)	8	4.7
Phe (F)	3	1.8
Pro (P)	7	4.1
Ser (S)	6	3.5
Thr (T)	9	5.3
Trp (W)	0	0.0
Tyr (Y)	3	1.8
Val (V)	12	7.1

The amino acid composition reveals that the majority of the amino acid residues in LtrK are hydrophobic (Ile, 11.5 % and Leu, 8.3 %) and in LtrR are acidic (Asp, 11.2 % and Glu, 10 %).

2.4.2 Domains annotation in LtrK

To identify protein domains and their organisation, LtrK protein sequence was analysed using Pfam database. The search found the following three domains:

1. CHASE (cyclases/histidine kinases associated sensory extracellular), from residue 80 to 216

2. HisKA (histidine kinase A), from residue 357 to 423 and
3. HATPase (histidine kinase like ATPase), from residue 479 to 580.

As, many TCS histidine kinases function as membrane receptor and have hydrophobic N-terminal domain that traverse the membrane (Hoch and Silhavy, 1995), presence of TMDs in LtrK was analysed using TMHMM tool which revealed two TMDs at N-terminal each consisting of 22 amino acids (Fig. 2.1).

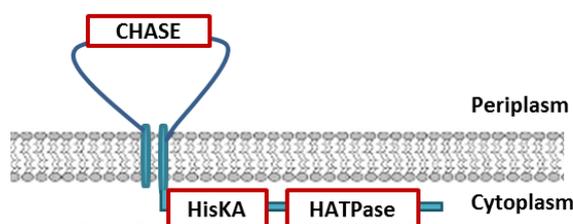


Figure 2.1 Schematic representations of different domains of LtrK

Collectively, Pfam and TMHMM analyses predicted that LtrK has two TMDs, a periplasmic CHASE and cytoplasmic HisKA and HATPase domains (Fig. 2.1). The extracellular CHASE domain is exclusively found in histidine kinases and adenylyl or diguanylate cyclases, serine/threonine protein kinases or methyl-accepting chemotaxis proteins from lower eukaryotes, plants, *Bacteria* and *Archaea* (Zhulin *et al.*, 2003). In contrast, HisKA and HATPase are the typical SK domains. HisKA is a dimerization and histidine phosphotransfer domain and contains a conserved histidine that is the site of phosphorylation whereas HATPase is a catalytic and ATP-binding domain (Willett and Kirby, 2012).

2.4.3 Conserved sequence motifs in LtrK

A significant characteristic feature of SKs is the presence of short amino acid sequence motifs referred as H, N, G1, G2, G3 and F blocks (Parkinson and Kofoid, 1992; Dutta and Inouye, 2000). The position and sequence identity of these blocks are summarized in Fig. 2.2.

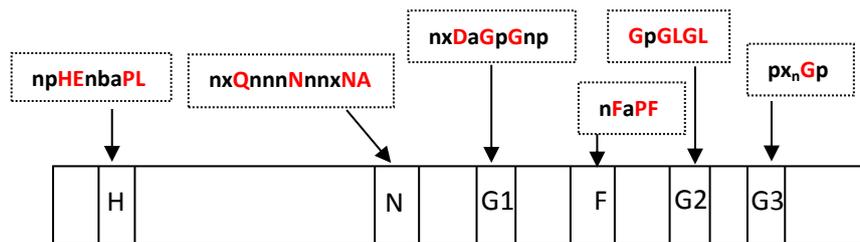


Figure 2.2 Conserved motifs in SK

Six sequence motifs of SK are shown in the dotted boxes. Upper case Red letters indicate amino acids present in 70% of SK sequences at that position. ‘n’, ‘p’, ‘b’ and ‘a’ refer to positions where at least 50% of the amino acids belong to the same family : n- nonpolar (I, L, M, V), p- polar (A, G, P, S, T), b-basic (H, K, R) and a-acidic or amide (D, E, N, Q) residues. ‘x’ represents non conserved residue positions (Parkinson and Kofoid, 1992; Dutta and Inouye, 2000).

To search for the conserved motifs in LtrK, the cytoplasmic catalytic domain (C-terminal 234 amino acids) was compared with bacterial SKs that have crystal structure available and shared at least 35 % amino acid sequence identity with LtrK; *i.e.* HK853, pdb: 3DGE (Marina *et al.*, 2005) from *T. maritima*; VicK, pdb: 4I5s (Wang *et al.*, 2013) from *Streptococcus mutans* and WalK, pdb: 3SL2 (Celikel *et al.*, 2012) from *B. subtilis*.

F	IFDKF
G2	GTGLGL
G3	ETGKGS

conserved H and N residues of H and N blocks are shown in red bold letter, conserved motif in H block is underlined

On the other hand, block N, G1, F, G2 and G3 are located in the C-terminal half of the cytoplasmic domain (HATPase domain), which forms the catalytic core. Block N contains a conserved Asn residue which is N480 in LtrK. The blocks G1, G2 and G3 are Gly rich motifs whereas block F is Phe rich motif and both of these blocks are involved in nucleotide (ATP) binding and positioning, near the conserved His of H block (Parkinson and Kofoid, 1992).

2.4.4 Analyses of His residues – the possible phosphorylation site in LtrK

In a typical SK, the H block His serve as the site of phosphorylation which indicates H367 in LtrK to be the phosphoryl group acceptor during the autophosphorylation reaction (Stock *et al.*, 1989; Parkinson and Kofoid, 1992). But some histidine residues outside of the H block have been shown to function in phosphorylation reactions in some SKs, *i.e.* E. coli NarX and NarQ (Cavicchioli *et al.*, 1995), thus without the experimental validation, H367 could not be confirmed as the only essential His residue for autophosphorylation reaction. Apart from the conserved H block H367, there are three His residues in the cytoplasmic domain of LtrK: H443, H448 and H502. To check if these residues are conserved in the related methanogens, multiple sequence alignment was performed (Fig. 2.4).

2.4.5 Domains in LtrR

Domain search using Pfam on LtrR protein sequence found two domains:

1. N-terminal HTH (helix turn helix) DNA-binding domain and
2. C-terminal REC (response receiver) domain

REC domain is a response regulatory receiver domain that is typically found in RRs (Stock *et al.*, 2000). However, DNA binding domain is rarely found in RR of *Archaea* though it is frequently present in *Bacteria* (Galperin, 2006). Presence of DNA binding domain was searched in the archaeal genome sequence of IMG (Integrated Microbial Genomes) portal (<http://img.jgi.doe.gov/>). It was found that very few RR of *Halobacteria*, *Methanococci* and *Methanobacteria* have HTH domain (Table 2.6).

Table 2.6 Presence of HTH domain in *Archaea*

<i>Archaea</i> (number of genome sequence)	Presence of HTH (Yes/No)
<i>Euryarchaeota</i> (368)	
<i>Halobacteria</i>	Y
<i>Methanococci</i>	Y
<i>Methanomicrobia</i>	Y
<i>Methanobacteria</i>	N
<i>Methanopyri</i>	N
<i>Archaeoglobi</i>	N
<i>Methanopyri</i>	N
<i>Thermococci</i>	N
<i>Thermoplasmata</i>	N
<i>Aenigmarchaeota</i> (3)	N
<i>Aigarchaeota</i> (18)	N
<i>Crenarchaeota</i> (138)	N
<i>Diapherotrites</i> (3)	N
<i>Korarchaeota</i> (1)	N
<i>Nanoarchaeota</i> (11)	N
<i>Nanohaloarchaeota</i> (4)	N

<i>Parvarchaeota</i> (2)	N
<i>Thaumarchaeota</i> (40)	N
<i>Geoarchaeota</i> (1)	N
Unclassified (7)	N

Y indicates present; N indicates not present

Interestingly, the HTH domain organization in methanogens is also different from *Bacteria* and is located at N-terminal while the vast majority of bacterial RRs contain the DNA binding domain at C-terminal. In contrast to methanogen, ~ 90 % of halophilic RRs have C-terminal DNA binding domain similar to *Bacteria* (Table 2.7).

Table 2.7 Organization of HTH Domain in Archaea

<i>Archaea</i> Halophilic	N-terminal	C-terminal
<i>Halonotius</i> sp. J07HN		J07HN4v2_01675, J07HN4v2_00780, J07HN4v2_02997, J07HN6v2_01307
<i>Halorubrum aidingense</i>		C461_00607
<i>Halorubrum lacusprofundi</i>		Hlac_3274
<i>Halobiforma lacisalsi</i>		C445_12891
<i>Halopiger djelfamassiliensis</i>		Ga0036365_12412, Ga0036365_12651, Ga0036367_103616
<i>Haloterrigena turkmenica</i>		Htur_0128, Htur_0348, Htur_0709
<i>Haloquadratum walsbyi</i>		Hqrw_3103
<i>Haloferax volcanii</i>		HVO_1357, HVO_B0272
<i>Haloarcula marismortui</i>	rrnB0301	pNG7159, pNG7223
<i>Halopiger xanaduensis</i>		Halxa_3599
<i>Halovivax ruber</i>		Halru_0454
<i>Halobacterium</i> sp. DL1	HalDL1_2148, HalDL1_1858	HalDL1_2066
<i>Halostagnicola larsenii</i>		Halla_0473
<i>Halalkalicoccus jeotgali</i>	HacjB3_0162	HacjB3_04505, HacjB3_00310, HacjB3_18828
<i>Halorhabdus utahensis</i>		Huta_1740, Huta_0317
<i>Halohasta litchfieldiae</i> tADL		halTADL_2543
<i>Natrinema pallidum</i>		C487_17780
<i>Natrinema pellirubrum</i>		C488_09871
<i>Natrinema versiforme</i>		C489_01256
<i>Natronolimnobius innermongolicus</i>		C493_05780
<i>Natronorubrum tibetense</i>		C496_05962
<i>Natrialba hulunbeirensis</i>		C483_14902
<i>Natrialba magadii</i>	Nmag_2044	Nmag_1012, Nmag_2147, Nmag_2692, C500_06936
<i>Natrialba taiwanensis</i>		C484_14808
<i>Natrialba aegyptia</i>		C480_05466
<i>Natrialba chahannaoensis</i>		C482_19154
<i>Natronobacterium gregoryi</i>		Natgr_1657

<i>Natronococcus amylolyticus</i>		C491_04490
<i>Natronomonas pharaonis</i> Gabara		NP0654A
Psychrophilic and Mesophilic Methanogens		
<i>Methanococcoides methylutens</i>	Ga0070849_114114	
<i>Methanohalophilus mahii</i>	Mmah_0822	
<i>Methanolobus psychrophilus</i>	Mpsy_1252	
<i>Methanosarcina barkeri</i>	Mbar_A2896	
<i>Methanosarcina mazei</i>	MM2351	
<i>Methanococcoides burtonii</i>	Mbur_0695 (LtrK)	

2.4.6 Analyses on Asp residues- the possible phosphorylation site in LtrR

Typically, during phosphotransfer reaction the phosphoryl group acceptor is an Asp residue that resides in the REC domain of RR (Stock *et al.*, 1989). The REC domain of LtrR has 14 Asp residues. To find out the possible site of phosphorylation in LtrR, the protein sequence was aligned with the bacterial RRs that have the crystal structure available in addition to sharing > 45 % amino acid sequence identity with LtrR such as PhoP (50 % identity) of *B. subtilis* and PhoB (46 % identity) of *E. coli*. The active site residues in both structures are D10 and D53 (Solà *et al.*, 1999; Birck *et al.*, 2003). Multiple sequence alignment showed that these active site residues in PhoP and PhoB are conserved in LtrR as D55 and D98 (Fig. 2.5). This indicates that these residues could be the part of the active site. The probability of D54 to be the active site residue cannot be excluded as it is also conserved and located close to D55. Based on these analyses, all three conserved residues (D54, D55, D98) were mutated to elucidate their function (described in Chapter 5).

domain of LtrK. The structure of the cytoplasmic domain was superimposable with the cytoplasmic domain structure of VicK which indicated that the cytoplasmic part of LtrK structure is more reliable. The cytoplasmic part included antiparallel helices of HisKA domain and α/β sandwich fold for the catalytic HATPase (Fig. 2.6).

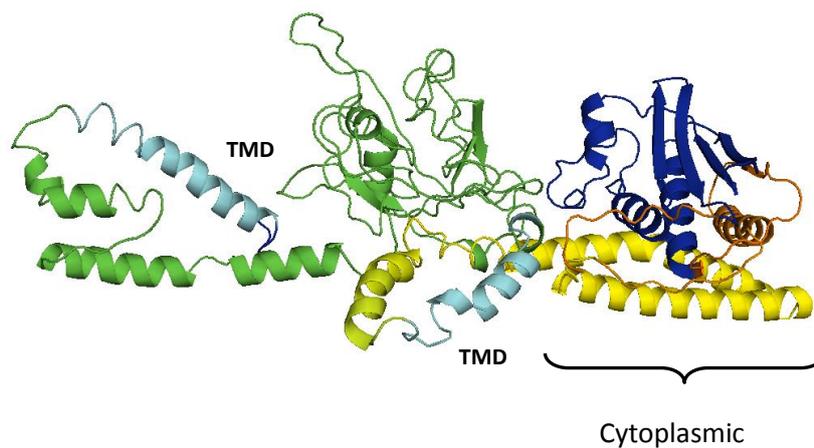


Figure 2.6 LtrK tertiary structure model

LtrK homology model predicted by I-TASSER. Different regions are shown in different color: two TMDs in cyan, periplasmic domain in green, cytoplasmic HisKA in yellow and HATPase in blue.

It had been found that during autophosphorylation reaction the H block acidic residue (D/E) makes a catalytic triad by forming hydrogen bond and salt bridge with the Arg and conserved Asn of N block respectively (Casino *et al.*, 2014). These catalytic triad residues are also conserved in LtrK. The conserved His (H367) and the catalytic triad residues R476-E368-N480 are shown in Fig. 2.7. Moreover during autophosphorylation reaction, the G2 block and F block residues form an ATP lid where the ATP binds and come in close contact with the H block residues (Marina *et al.*, 2005; Casino *et al.*, 2014). This observation suggests that in the structural model of the

cytoplasmic domain of LtrK G2 and F block may form an ATP lid between $\alpha 5$ and $\alpha 6$ (Fig. 2.7).

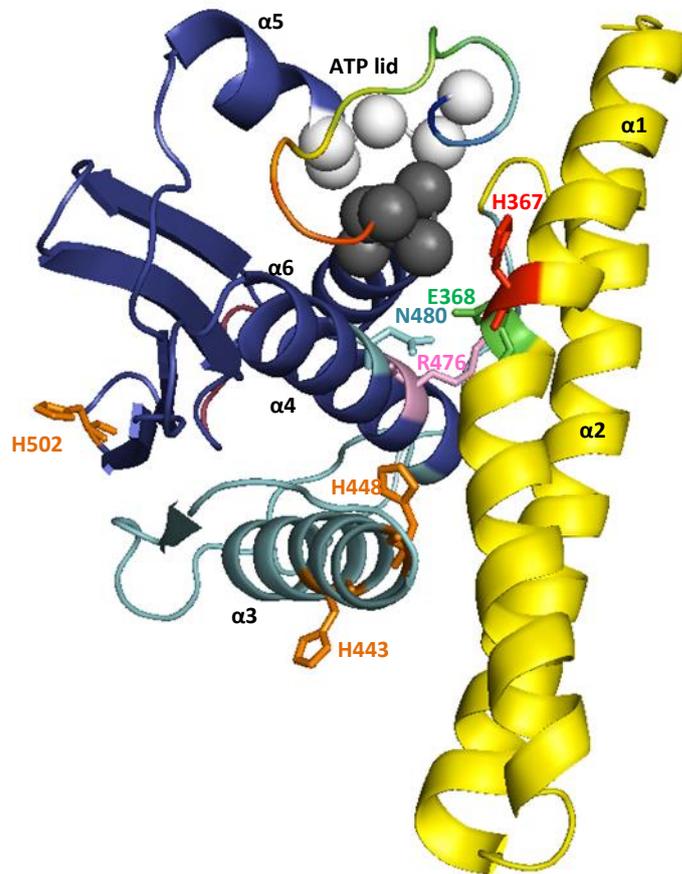


Figure 2.7 Cytoplasmic domain structure of the LtrK homology model

The HisKA domain (yellow) includes the $\alpha 1$ and $\alpha 2$ helices. The $\alpha 1$ helix contains the conserved H367 (red) and E368 (green) residues of the H block. The $\alpha 4$ helix of HATPase domain (blue) contains the conserved N480 (cyan) and R476 (purple) residues of the N block. A catalytic triad involved in autophosphorylation (Casino *et al.*, 2009; Casino *et al.*, 2014) is formed by R476 (purple), E368 (green) and N480 (cyan). The $\alpha 3$ helix (between the HisKA and HATPase domains) contains the additional histidine residues H443 and H448 (magenta). The G (black) and F (white) block residues are shown in ball representation that forms the ATP lid (rainbow).

As described in section 2.3.4, there are three non-conserved His residues in the cytoplasmic domain of LtrK: H443, H448 and H502. In the structure model, H443 and

H448 are in the $\alpha 3$ helix that is positioned between the HisKA and HATPase domains and connected to the HATPase domain via a β -sheet to the $\alpha 4$ helix containing the N block whereas H502 is located in a loop between β -sheets of HATPase domain (Fig. 2.7).

2.4.7.2 3D structural homology model of LtrR

The model generated by I-TASSER for LtrR best matched with PhyR (pdb: 4G97), a stress response regulator from *Brucella abortus* (unpublished data), with highest TM score of 0.86 and lowest RMSD of 2.01. Typically, REC domain structure shows $(\beta\alpha)_5$ topology that preserves the conserved active site residues at the C-terminal end of β strands (Bourret, 2010). For instance, the C-terminal end of $\beta 1$ contains two conserved Asp residues which are involved in Mg^{2+} binding during the phosphotransfer reaction and C-terminal end of $\beta 3$ contains highly conserved Asp residue that is the site of phosphorylation (Bourret, 2010). The $\beta 4$ strand ends with a highly conserved Thr or Lys residue that interacts with the phosphoryl group during the phosphotransfer reaction whereas $\beta 5$ ends with a conserved Lys residue that is important for the phosphorylation mediated conformational changes (Bourret, 2010). All these conserved residues are also present at the same location in LtrR structure, Mg^{2+} binding two Asp residues: D54 and D55 at $\beta 1$ end; phosphoacceptor Asp residue: D98 at $\beta 3$ end; phosphoryl group interacting residue: T128 at $\beta 4$ end and conformational change inducing residue: K150 at $\beta 5$ end (Fig. 2.8). But to confirm the role of these residues in LtrR, experimental validation is required including mutation of the conserved C-terminal residues of each β strand and phosphotransfer reaction assays (described in Chapter 5).

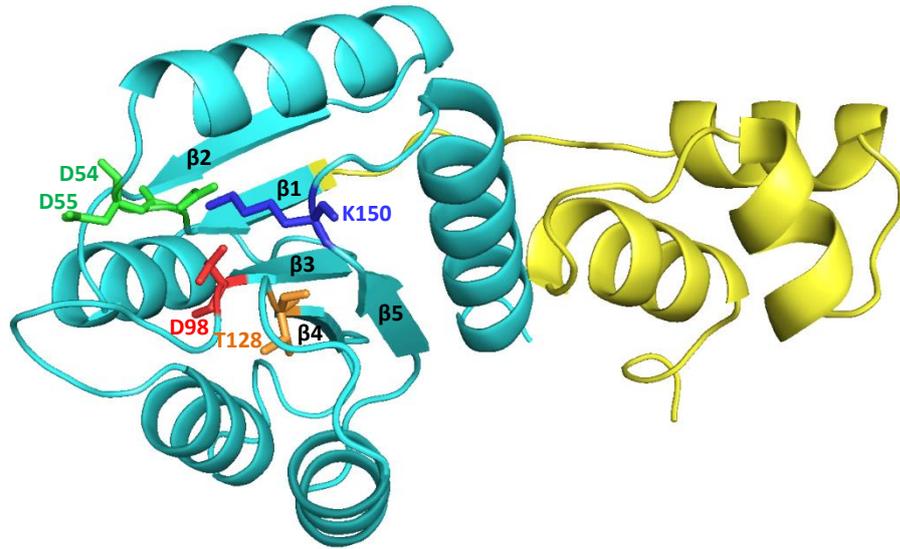


Figure 2.8 LtrR tertiary structure model

The N-terminal HTH domain is shown in yellow and the C-terminal REC domain is shown in cyan. The conserved D54 and D55 (both green) at the end of $\beta 1$, site of phosphorylation D98 (red) at the end of $\beta 3$, conserved T128 at the end of $\beta 4$ and conserved K150 at the end of $\beta 5$.

2.4.8 Phylogenetic relationship analyses

Due to the diversity of the amino acid sequence of periplasmic domain in SKs, only the cytoplasmic domain amino acid sequence (both from bacteria and archaea) was selected to construct a phylogenetic tree for LtrK. The bootstrap consensus tree of LtrK was inferred from 10,000 replicates (Felsenstein, 1985) and taken to represent the evolutionary history of the taxa analysed.

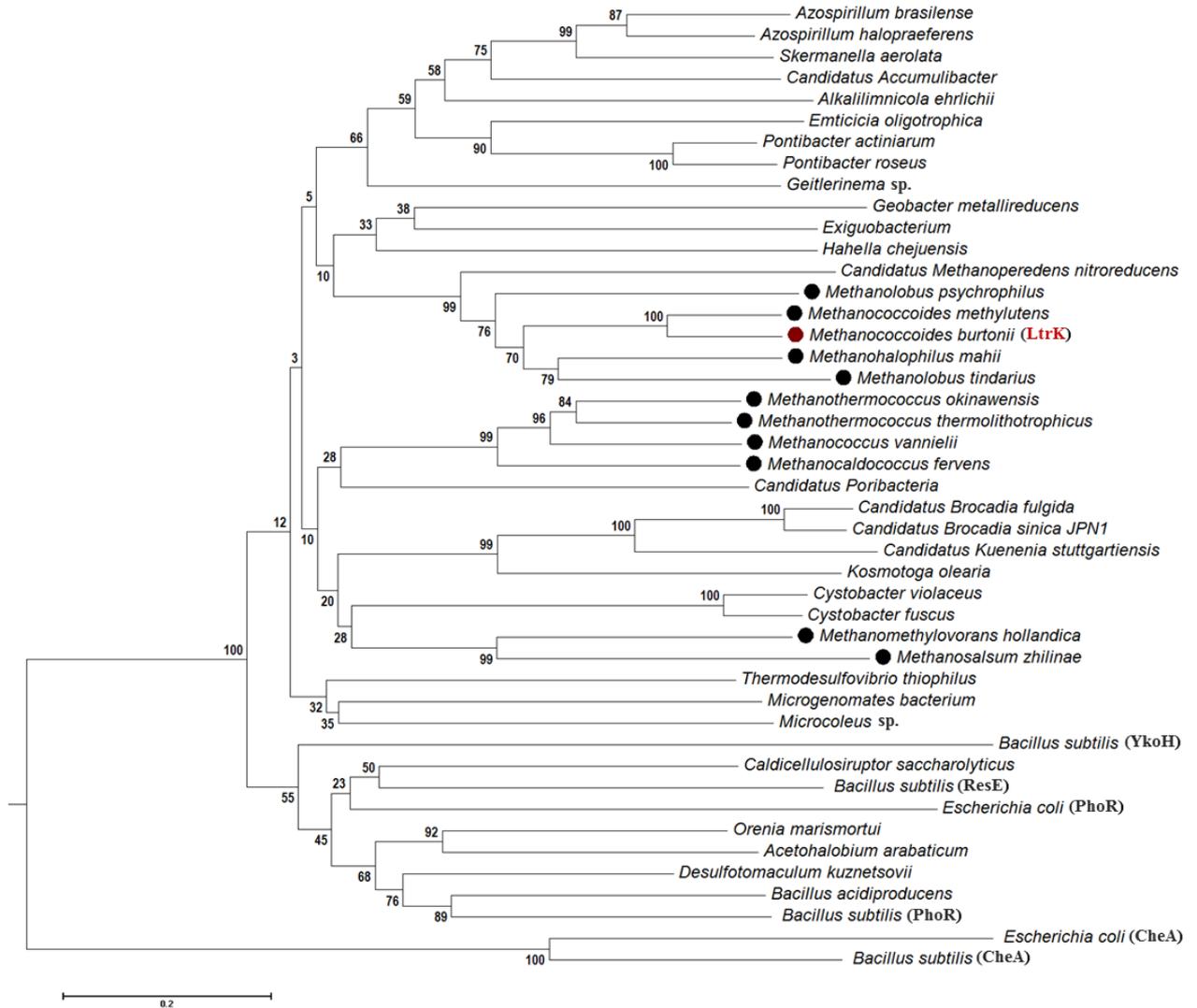


Figure 2.9 Phylogenetic tree of LtrK

Amino acid neighbor-joining (NJ) tree highlighting the SKs from methanogens (●); LtrK is shown in red symbol.

Few bacterial SKs with known function (i.e CheA, involved in chemotaxis (Hess *et al.*, 1988); PhoR, involved in phosphate homeostasis in cell (Makino *et al.*, 1989; Shi and Hulett, 1999)) were also included in the tree to see if LtrK clusters with any functionally known HK. But LtrK clustered with the methanogens. In fact, functionally similar proteins, PhoR of *E. coli* and *B. subtilis* have clustered separately indicating the

sequence of HKs have diverged significantly through horizontal gene transfer during the course of evolution.

The tree showed that the group containing *Candidatus Methanoperedens nitroreducens*, *Methanlobus psychrophilus*, *Methanococcoides methylutens*, *M. burtonii*, *Methanohalophilu mahii* and *Methanlobus tindarius* has high bootstrap value (> 75) which indicates the SKs of these methanogens are close to each other. The most closely related SK to LtrK, is from *M. methylutens* and the second closest one is from *M. psychrophilus*. There is another branch where a set of methanogens (*Methanothermococcus okinawensis*, *Methanothermococcus thermolithotrophicus*, *Methanococcus vannieli*, *Methanocaldococcus fervens*) have clustered together but due to poor bootstrap value between the two sub-branches of methanogens it is difficult to conclude if these SKs are closely related. Overall the tree indicates that the SKs within a genus or closely related genera are quite similar, but across genera the sequences have changed.

On the other hand, the tree for LtrR (Fig. 2.10) was generated using the receiver domain amino acid sequences of archaeal and bacterial RRs. The pattern of the constructed tree was similar to the LtrK tree; the methanogens clustered together while the bacterial RRs generated a separate cluster. Similar to LtrK, the most closely related RR to LtrR was from *M. methylutens*. Because of the low bootstrap value it's difficult to indicate the second closest RR. However, it is indicative that the RRs within the methanogens are similar. In addition, the similar pattern within both the trees indicates that LtrK and LtrR have evolved together.

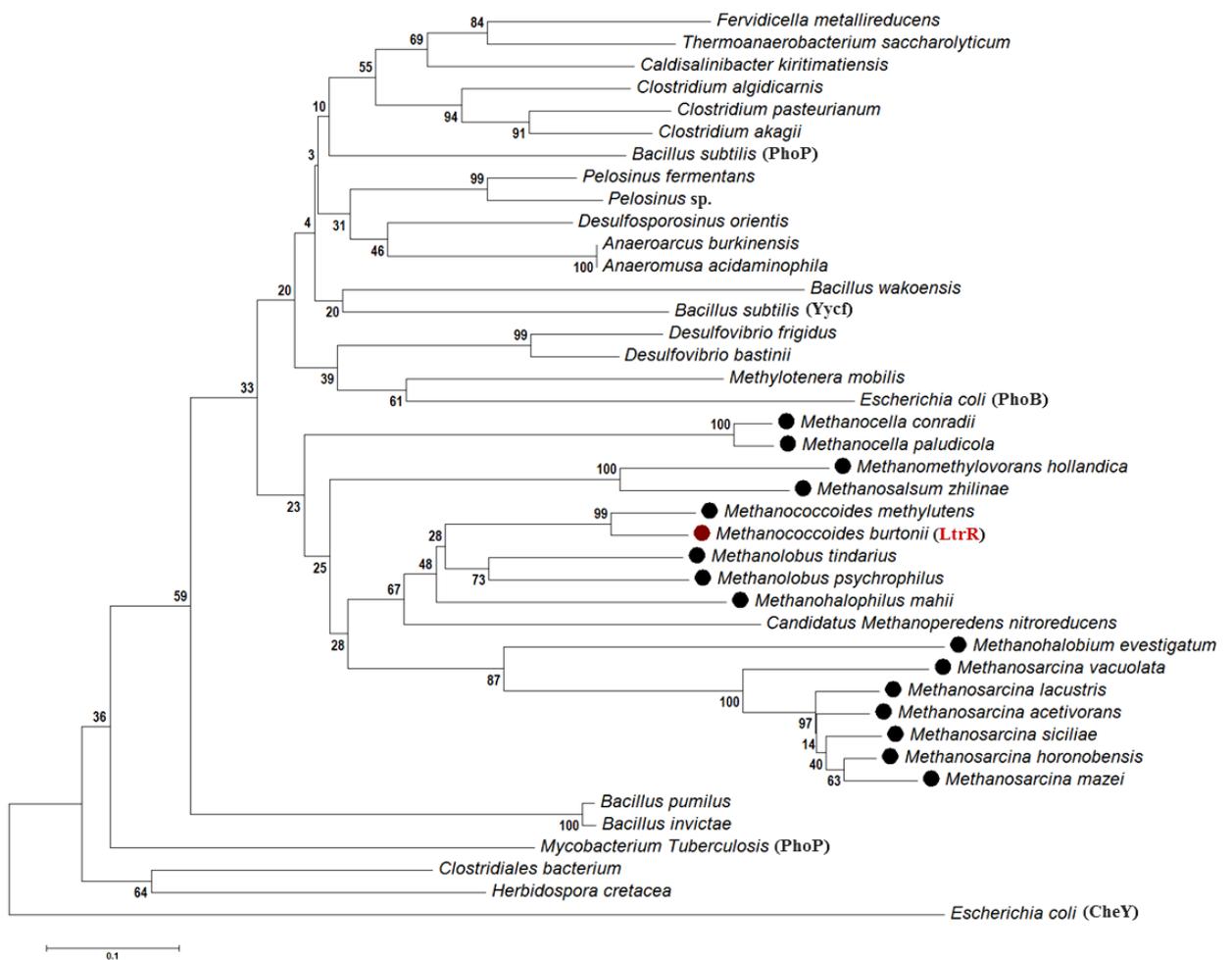


Figure 2.10 Phylogenetic tree of LtrR

Amino acid neighbor-joining (NJ) tree highlighting the RRs from methanogens (●); LtrR of *M. butonii* is shown in red symbol.

2.5 Discussion

The key features of LtrK include: presence of two TMDs with a long periplasmic region containing a CHASE domain (Fig. 2.1) and the presence of conserved characteristic blocks of bacterial SKs (H, N, G1, F, G2, and G3) (Table 2.2) in the cytoplasmic domain.

CHASE domain is predicted to bind diverse low molecular weight ligands, such as the cytokinin-like adenine derivatives or peptides that trigger developmental changes (Anantharaman and Aravind, 2001; Heyl *et al.*, 2007). In *Archaea* a specific type of CHASE domain- CHASE4 has been identified that is only present in SKs, including methanogens such as *Methanosarcina acetivorans* (Zhulin *et al.*, 2003) and *M. harundinacea* (Li *et al.*, 2014). The secondary structure of CHASE4 contains a $\alpha\beta$ fold with an extensive loop in the middle portion of the domain (Zhulin *et al.*, 2003). However, LtrK does not possess the signature sequences of CHASE4 and may therefore represent an additional class within *Archaea*.

The conserved block search revealed the universally conserved H block His in LtrK as H367 (Fig. 2.3). The H block His is reported to be the phosphoryl group acceptor during autophosphorylation reaction which indicates H367 as the site of phosphorylation in LtrK (Stock *et al.*, 1989; Parkinson and Kofoid, 1992). In addition, analyses on three other His residues of cytoplasmic domain of LtrK revealed that they are not conserved in closely related methanogens (Fig. 2.4) which further implies H367 to be the site of phosphorylation. Another conserved block of SK, N block contains a conserved Asn which is N480 in LtrK (Table 2.5). Mutation at conserved Asn residue in bacterial SKs have shown loss of autophosphorylation activity (Kanamaru *et al.*, 1989; Yang and Inouye, 1993) which indicates that N480 in LtrK may play a critical role for the kinase reaction. Furthermore, the typical E/DxxT/N motif of HisKA subfamily is conserved as ELKT adjacent to H367 in LtrK (Table 2.4). It has been found that the conserved E/D of the motif is essential for the kinase activity and T/N is essential for the phosphatase activity (Huynh *et al.*, 2010; Willett and Kirby, 2012). This infers that

the Glu (E368) and Thr (T371) of LtrK may play significant roles in kinase and phosphatase activity respectively.

The structure model of LtrK gives an insight about the conserved residues essential for the kinase activity. However, as the periplasmic domain's sequence shares almost no similarity with the available bacterial SK structures, the N-terminal domain of the predicted structure is not reliable which renders the LtrK structure rather incomplete. The cytoplasmic domain (containing HisKA and HATPase domain) of the predicted 3D homology model aligned completely with the bacterial SK (VicK) structure and exhibited antiparallel helices for the dimerization domain of HisKA domain and α/β sandwich fold for the catalytic HATPase domain that are typical of SK structures (Fig. 2.6). A structural guided functional analyses of HK853 of *T. maritima* and EnvZ of *E. coli* showed that the conserved acidic residue (Asp/Glu) of the H block, whose carboxylic side chain is bound to the δ N of the conserved His, act as a general base and extracts a proton from the histidine to induce its nucleophilic attack on ATP γ -phosphate (Casino *et al.*, 2014). The catalytic triad made by this H block Asp/Glu with the Arg and Asn of N block assist to achieve the correct deposition and polarization of Asp/Glu to act as a general base (Casino *et al.*, 2014). The H block acidic residue in LtrK is Glu 368 and the N block residues that may form the catalytic triad are Arg476 and Asn480 (Fig. 2.7). It would be interesting to mutate these residues to investigate their role in the kinase activity of LtrK. Overall analyses indicated that the essential residues for the autophosphorylation reaction are preserved at their corresponding position in the LtrK structure thus indicating that the molecular basis of kinase reaction could be similar to bacterial SKs.

In contrast, the hallmark of LtrR includes: the presence of HTH domain and its N-terminal positioning (Table 2.7). It was interesting to find that only three classes of *Archaea* possess HTH domain: *Halobacteria*, *Methanococci* and *Methanomicrobia* (Table 2.6). Moreover, the HTH domain is located at N-terminal in methanogens which is opposite to *Bacteria* whereas in *Haloarchaea* it is mostly located at C-terminal (Table 2.7) similar to bacteria. It is noteworthy that *Haloarchaea* also tend to possess cold shock proteins (Csps), a class of small nucleic acid binding proteins that is the hallmark of *Bacteria*, which is not produced in most of the archaeal species including methanogens (Giaquinto *et al.*, 2007). These data indicate that within the *Archaea*, psychrophilic and mesophilic methanogens have evolved a class of RRs that have distinct domain architecture to *Bacteria*, and that *Haloarchaea* are likely to have acquired most of their RR and Csp genes from *Bacteria*.

The multiple sequence alignment of LtrR with bacterial RR that share the highest sequence identity (PhoP of *B. subtilis* and PhoB of *E. coli*) indicated D55 and D98 to be the possible active site residue in LtrR (Fig. 2.5). Interestingly, the 3D structure homology model of LtrR revealed that D55 and D98 are located at C-terminal end of $\beta 1$ and $\beta 3$ respectively (Fig. 2.7). Structural study on RRs showed that Asp at $\beta 1$ and $\beta 3$ end are active site residues and are involved in Mg^{2+} binding and phosphorylation (Bourret, 2010). So collectively the multiple sequence alignment and homology model data strongly indicate that D55 and D98 are essential residues for the phosphorylation of LtrR.

The phylogenetic tree analyses of LtrK and LtrR showed that both of the proteins are most closely related to the TCS of *M. methylutens* and probably to other methanogens but distinct from bacterial ones indicating the vertical inheritance of the

gene within closely related methanogens. Moreover, similar pattern of both trees indicated the co-evolution of LtrK and LtrR. As a matter of fact, both genes are expressed in an operon in *M. burtonii* genome (Campanaro *et al.*, 2011) which further supports the co-evolutionary relationship.

In conclusion, the dissection of the sequences of LtrK and LtrR and its comparison with other TCSs has not only demonstrated that both proteins conforms to the hallmarks of other TCSs, but also contains unique features, *i.e.* presence of CHASE domain in LtrK, N-tr positioning of HTH domain in LtrR. In addition, 3D structure homology modelling allowed greater insights into the positioning of conserved residues and possible identification of the active site residues. Experimental validation of the active site residues were validated in the Chapter 5.

2.6 References

- Anantharaman, V., and Aravind, L. (2001) The CHASE domain: a predicted ligand-binding module in plant cytokinin receptors and other eukaryotic and bacterial receptors. *Trends Biochem. Sci.* 26: 579-582.
- Bhate, M.P. *et al.* (2015) Signal transduction in histidine kinases: insights from new structures. *Structure* 23: 981-994.
- Birck, C. *et al.* (2003) The crystal structure of the phosphorylation domain in PhoP reveals a functional tandem association mediated by an asymmetric interface. *J. Bacteriol.* 185: 254-261.
- Bourret, R.B. (2010) Receiver domain structure and function in response regulator proteins. *Curr. Opin. Microbiol.* 13: 142-149.
- Campanaro, S. *et al.* (2011) Temperature-dependent global gene expression in the Antarctic archaeon *Methanococcoides burtonii*. *Environ. Microbiol.* 13: 2018-2038.
- Casino, P. *et al.* (2014) Visualizing autophosphorylation in histidine kinases. *Nat. Commun.* 5: 3258.

- Cavicchioli, R. *et al.* (1995) The NarX and NarQ sensor-transmitter proteins of *Escherichia coli* each require two conserved histidines for nitrate-dependent signal transduction to NarL. *J. Bacteriol.* 177: 2416-2424.
- Celikel, R. *et al.* (2012) ATP forms a stable complex with the essential histidine kinase WalK (YycG) domain. *Acta Crystallogr. D. Biol. Crystallogr.* 68: 839-845.
- Dutta, R., and Inouye, M. (2000) GHKL, an emergent ATPase/kinase superfamily. *Trends Biochem. Sci.* 25: 24-28.
- Edgar, R.C. (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* 32: 1792-1797.
- Felsenstein, J. (1985) Confidence limits on phylogenies: an approach using the bootstrap. *Evolution*: 783-791.
- Finn, R.D. *et al.* (2014) Pfam: the protein families database. *Nucleic Acids Res.* 42: D222-D230.
- Galperin, M.Y. (2006) Structural classification of bacterial response regulators: diversity of output domains and domain combinations. *J. Bacteriol.* 188: 4169-4182.
- Gao, R., and Stock, A.M. (2009) Biological insights from structures of two-component proteins. *Annu. Rev. Microbiol.* 63: 133.
- Giaquinto, L. *et al.* (2007) Structure and function of cold shock proteins in archaea. *J. Bacteriol.* 189: 5738-5748.
- Goujon, M. *et al.* (2010) A new bioinformatics analysis tools framework at EMBL–EBI. *Nucleic Acids Res.* 38: W695-W699.
- Hess, J.F. *et al.* (1988) Phosphorylation of three proteins in the signaling pathway of bacterial chemotaxis. *Cell* 53: 79-87.
- Heyl, A. *et al.* (2007) Evolutionary proteomics identifies amino acids essential for ligand-binding of the cytokinin receptor CHASE domain. *BMC Evol. Biol.* 7: 62.
- Hoch, J.A., and Silhavy, T.J. (1995) Two-component signal transduction. ASM press Washington, DC.
- Huynh, T.N. *et al.* (2010) Conserved mechanism for sensor phosphatase control of two-component signaling revealed in the nitrate sensor NarX. *Proc. Natl. Acad. Sci. USA.* 107: 21140-21145.
- Kanamaru, K. *et al.* (1989) Signal transduction and osmoregulation in *Escherichia coli*. A single amino acid change in the protein kinase, EnvZ, results in loss of its phosphorylation and dephosphorylation abilities with respect to the activator protein, OmpR. *J. Biol. Chem.* 264: 21633-21637.

- Li, J. *et al.* (2014) Characterization of an archaeal two-component system that regulates methanogenesis in *Methanosaeta harundinacea*. *PloS one* 9: e95502. doi:95510.91371/journal.pone.0095502.
- Madden, T. (2013) The BLAST sequence analysis tool. *The NCBI Handbook*.
- Makino, K. *et al.* (1989) Signal transduction in the phosphate regulon of *Escherichia coli* involves phosphotransfer between PhoR and PhoB proteins. *J. Mol. Biol.* 210: 551-559.
- Marina, A. *et al.* (2005) Structure of the entire cytoplasmic portion of a sensor histidine-kinase protein. *EMBO J.* 24: 4247-4259.
- Markowitz, V.M. *et al.* (2012) IMG: the integrated microbial genomes database and comparative analysis system. *Nucleic Acids Res.* 40: D115-D122.
- Mascher, T. *et al.* (2006) Stimulus perception in bacterial signal-transducing histidine kinases. *Microbiol. Mol. Biol. Rev.* 70: 910-938.
- Parkinson, J.S., and Kofoid, E.C. (1992) Communication modules in bacterial signaling proteins. *Annu. Rev. Genet.* 26: 71-112.
- Shi, L., and Hulett, F.M. (1999) The cytoplasmic kinase domain of PhoR is sufficient for the low phosphate-inducible expression of Pho regulon genes in *Bacillus subtilis*. *Mol. Microbiol.* 31: 211-222.
- Solà, M. *et al.* (1999) Three-dimensional crystal structure of the transcription factor PhoB receiver domain. *J. Mol. Biol.* 285: 675-687.
- Sonnhammer, E.L. *et al.* (1998) A hidden Markov model for predicting transmembrane helices in protein sequences. In *Ismb* 6: 175-182.
- Stock, A.M. *et al.* (2000) Two-component signal transduction. *Annu. Rev. Biochem.* 69: 183-215.
- Stock, J. *et al.* (1989) Protein phosphorylation and regulation of adaptive responses in bacteria. *Microbiol. Rev.* 53: 450.
- Taha *et al.* (2016) Single TRAM domain RNA-binding proteins in Archaea: functional insight from Ctr3 from the Antarctic methanogen *Methanococcoides burtonii*. *Environ. Microbiol.* doi: 10.1111/1462-2920.13229.
- Tamura, K. *et al.* (2013) MEGA6: molecular evolutionary genetics analysis version 6.0. *Mol. Biol. Evol.* 30: 2725-2729.
- Thomas, T., and Cavicchioli, R. (1998) Archaeal cold-adapted proteins: structural and evolutionary analysis of the elongation factor 2 proteins from psychrophilic, mesophilic and thermophilic methanogens 1. *FEBS Lett.* 439: 281-286.

Walker, J.M. (2005) *The proteomics protocols handbook*: Springer.

Wang, C. *et al.* (2013) Mechanistic insights revealed by the crystal structure of a histidine kinase with signal transducer and sensor domains. *PLoS Biol.* 11: e1001493.

Willett, J.W., and Kirby, J.R. (2012) Genetic and biochemical dissection of a HisKA domain identifies residues required exclusively for kinase and phosphatase activities. *PLoS Genet* 8(11): e1003084. doi:10.1371/journal.pgen.1003084 e1003084.

Yang, Y., and Inouye, M. (1993) Requirement of both kinase and phosphatase activities of an Escherichia coli receptor (Taz1) for ligand-dependent signal transduction. *J. Mol. Biol.* 231: 335-342.

Zhang, Y. (2008) I-TASSER server for protein 3D structure prediction. *BMC Bioinformatics* 9: 40.

Zhulin, I.B. *et al.* (2003) Common extracellular sensory domains in transmembrane receptors for diverse signal transduction pathways in bacteria and archaea. *J. Bacteriol.* 185: 285-294.

Chapter 3

Purification and biophysical characterization of the sensor kinase, LtrK

3.1 Abstract

The functionally annotated sensor kinase (SK) of psychrophilic methanogen *M. burtonii*, LtrK is predicted to contain an N-terminal periplasmic domain with two transmembrane domains (TMDs) and a C-terminal catalytic cytoplasmic domain. Due to the insoluble nature of TMDs, the cytoplasmic catalytic domain (C-terminal 222 amino acids) of LtrK was overexpressed in *E. coli* with an N-terminal GST tag. Despite the absence of the TMDs, LtrK formed insoluble inclusion bodies. Soluble protein was obtained by a protein unfolding/refolding technique using ionic sarkosyl and non-ionic Triton X-100 detergents during purification. Size exclusion chromatography analyses revealed the native oligomeric state of LtrK as dimer and the circular dichroism data indicated the protein had secondary α -helical structure. The native protein structure stability was assessed by differential scanning calorimetry technique which showed that the thermal unfolding of the protein was irreversible.

3.2 Introduction

Psychrophilic enzymes have weak thermal stability which makes them difficult to purify in large-scale, as oppose to a mesophilic or thermophilic enzymes that are thermally stable. Proteins can be purified directly from the native organism by large-scale fermentation or as a heterologous protein by expressing in a mesophilic host, typically in *E. coli*. The main drawback of purifying native protein by large scale fermentation of the psychrophilic organism is the low production level of the wild strain (Feller *et al.*, 1998; Cavicchioli *et al.*, 2006). In contrast, the main concern of heterologous protein expression is if the psychrophilic protein that is normally synthesized near 0 °C, will be properly folded in the mesophilic host (Feller *et al.*, 1998). Fortunately, many psychrophilic enzymes have been successfully overexpressed and purified in *E. coli* in its native folded state using the available protein purification method that were originally developed for mesophilic and thermophilic proteins, with some cell culture growth condition optimisation. For example, psychrophilic α - amylase that is synthesized at 0 ± 2 °C in *Alteromonas haloplanctis*, was overexpressed in *E. coli* with its genuine folding properties (identical with the native protein) at 18 °C but not at 37 °C (Feller *et al.*, 1998); EF-2 from *M. burtonii* was purified with high yield when overexpressed in *E. coli* at 14 °C compared to at 23 or 30 °C (Cavicchioli *et al.*, 2006). However, reduction of growth temperature can greatly reduce the growth rate and thus the rate of protein synthesis. For instance, a shift from 37 to 10 °C greatly reduces the growth rate of *E. coli* and the cells are in cold shock. So the temperature of culture growth needs to be optimised in such a way so that a compromise can be reached between the stability of the psychrophilic enzyme and the mesophilic growth rate. But some psychrophilic enzymes are very sensitive to autolytic degradation, thermal

inactivation, and molecular aggregation even during expression at lower temperatures in *E. coli*, *i.e.* psychrophilic Atlantic cod trypsin, was found to be sensitive to autolysis even at 18 °C during expression in *E. coli* (Jónsdóttir *et al.*, 2004).

Moreover, proteolytic degradation by proteases in *E. coli* cell, presence of rare codons in psychrophilic genes and unavailability of the cognate tRNA in *E. coli* are another obstacles in achieving high expression level of recombinant proteins (Thomas and Cavicchioli, 2000; Cavicchioli *et al.*, 2006). Many different *E. coli* strains have been developed to overcome these issues, *i.e.* *E. coli* strain BL21 lacks protease enzymes (Gräslund *et al.*, 2008), *E. coli* strain Rosetta encodes tRNA genes for rarely used codons (Novy *et al.*, 2001). The strain BL21 was proven to be useful to purify psychrophilic EF-2 (Thomas and Cavicchioli, 2000) and RNA polymerase E/F (De Francisci *et al.*, 2011) of *M. burtonii* while *E. coli* strain Rosetta was found to be suitable to express recombinant Ctr3 (Taha *et al.*, 2016) and chaperonin (Pilak *et al.*, 2011) protein of *M. burtonii*. Another approach to overcome the problem of biased codon usage of the non-native host is codon optimisation. In this technique the gene to be expressed in *E. coli* is completely synthesized in which the coding sequences are modified to eliminate rare codons. Synthesized codon optimised genes are now frequently used for heterologous expression and found to significantly improve the translational efficiency of the gene in non-native host (Gustafsson *et al.*, 2004). For instance, the psychrophilic *M. burtonii* protein, Ctr3 was expressed in *E. coli* and purified with high yield by using a synthetic clone obtained from DNA 2.0, containing codon optimised *ctr3* gene (Taha *et al.*, 2016).

Recombinant proteins are mostly produced as a fusion to an affinity tag as tags significantly aid in protein purification and seldom affect the biochemical activity of the

expressed protein. Besides, some affinity tags can enhance the solubility of the partner proteins, *i.e.* glutathione-S-transferase (GST), maltose binding protein (MBP); but none of these tags work universally with every partner protein (Esposito and Chatterjee, 2006). However, GST tag has been used to produce some N-terminal TMDs truncated SKs with high solubility, *e.g.* cytoplasmic domain of PhoR and DesK of *B. subtilis* had been purified with GST-tag as a soluble protein (Shi and Hulett, 1999; Albanesi *et al.*, 2004).

As cold-adapted enzymes tend to unfold at low to moderate temperatures, it is important to check the structural stability of the purified protein through biophysical studies. A number of biophysical techniques are available for determining protein stability. Spectrophotometric techniques such as circular dichroism in the near UV range (near UV-CD) and fluorescence (FS) provide a measure of the structural changes of specific amino acid residues, *i.e.* tyrosine or tryptophan for FS and aromatic residues for near UV-CD, and are suitable for studying protein's tertiary structure (Cavicchioli *et al.*, 2006). CD in the far-UV range (far UV-CD) is useful for examining protein's secondary structure. All these methods can also be used to follow the structural (secondary or tertiary) transition (unfolding/refolding) to check reversibility of the protein (Cavicchioli *et al.*, 2006). However, FS and near-UV CD have limited use for evaluating the stability of the whole protein as they depend on the presence of specific residues and far UV-CD is not suitable for determining the thermal transition of the protein since it gives an evaluation of the secondary structure and doesn't reflect the transition affecting the tertiary structure (Gerday *et al.*, 1997). A useful method for understanding protein unfolding/refolding transition is differential scanning calorimetry (DSC). DSC is a calorimetric technique that measures the apparent molar heat capacity

(Cp) of a protein as a function of temperature (T) (Cavicchioli *et al.*, 2006). DSC also enables the calculation of enthalpy changes which can be used to determine kinetic and thermodynamic parameters of unfolding (Cavicchioli *et al.*, 2006) and had been used to study temperature induced unfolding of many psychrophilic proteins. For example, thermodynamic activation parameters for protein stability had been determined using DSC for psychrophilic enzyme, EF-2 (Thomas and Cavicchioli, 2002), RNA polymerase E/F (De Francisci *et al.*, 2011), and Ctr3 (Taha *et al.*, 2016) of *M. burtonii*.

TCS SKs are typically a transmembrane protein with N-terminal transmembrane domains (TMDs) (Wolanin *et al.*, 2002). Transmembrane proteins are difficult to purify because of their insoluble nature due to the presence of large hydrophobic surfaces. In *Bacteria*, many SKs are found to have activity without their N-terminal TMDs (Makino *et al.*, 1989; Liu and Hulett, 1997; Shi and Hulett, 1999). In addition, kinase activity analyses on Fill, the only studied SK in *Archaea*, *M. harundinacea*, showed that only the cytoplasmic histidine kinase domain is necessary for catalytic activity (Li *et al.*, 2014).

LtrK is a protein from psychrophilic *M. burtonii* and has two predicted TMDs in its N-terminal region (described in Chapter 2). The fact that SKs can exhibit catalytic activity without TMDs indicated that LtrK may also possess enzymatic activity without its N-terminal TMDs. Thus, only the cytoplasmic catalytic domain of LtrK was attempted for purification. This chapter described the overexpression and purification protocol of the cytoplasmic domain of LtrK along with characterization of its biophysical properties. These data provided valuable information on protein solubility, oligomeric state and protein's secondary and tertiary structure that formed the basis for

conducting studies on kinase function analyses of LtrK described in the following chapters (Chapter 4 and 5).

3.3 Experimental procedures

3.3.1 Cloning

The cytoplasmic region of LtrK gene was commercially synthesized and subsequently cloned in an expression vector pReceiver-B03 from GeneCopoeia, Maryland, USA. The gene was codon optimised for expression in *E. coli*. The plasmid contained N-terminal glutathione S-transferase (GST) tag and an ampicillin resistance gene. There was a tobacco etch virus (TEV) protease cleavage site between the GST tag and LtrK coding gene to remove the fusion tag afterwards. The plasmid map is shown in Appendix 2.

3.3.2 *E. coli* competent cells and transformation

E. coli cells were streaked on LB agar plate and incubated overnight at 37 °C. A single colony was inoculated into 20 ml of Luria Broth (LB; 5g l⁻¹ yeast extract, 10g l⁻¹ Tryptone, 10g l⁻¹ NaCl) and cultured overnight at 37°C at 250 rpm. On the following day, 100 µl of overnight culture was inoculated in freshly made 20 ml LB media and was allowed to grow at 37 °C to an OD₆₀₀ of 0.4 – 0.6. The cells were collected by centrifugation at 6000 × g, 4 °C for 10 min. The cell pellet was kept on ice and resuspended initially in 1 ml 0.1 M CaCl₂. A further 9 ml of cold 0.1 M CaCl₂ was added and the cell suspension was incubated for 20 min on ice before being harvested by centrifugation at 6000 × g, 4 °C for 10 min. The cell pellet was resuspended with 1

mL of cold 0.1M CaCl₂ containing 20% (v/v) glycerol and was stored in 50 ml aliquots at -80°C until further use. Two *E. coli* cell strains, TOP10 and BL21 (DE3) were successfully made competent with this method.

For transformation, a volume of 2 µl of pReceiver-B03 plasmid containing LtrK gene was added to 50 µl of competent *E. coli* strain BL21 (DE3) cells that were already thawed on ice. The cells and plasmid were mixed by tapping the tube gently, and incubated on ice for 20 min. The cells were then heat shocked at 42 °C for 45 s, and placed immediately on ice for 5 min. One ml of fresh LB media was added to the tube and the cells were grown at 37 °C for 1 h with shaking at 250 rpm. The transformed cells were then centrifuged for 1 min at 21,000 × g and most of the supernatant was removed leaving approximately 100 µl to resuspend the cell pellet. Finally, the resuspended cells were spread on a LB agar plate containing 100 µg ml⁻¹ of ampicillin (Sigma) and the plate was incubated at 37 °C overnight.

3.3.3 Overexpression of recombinant LtrK

A single colony of transformed cells was picked and inoculated into 20 mL of LB media containing ampicillin (100µg ml⁻¹) and cultured at 37 °C and 280 rpm overnight. Next day, 2 L of the same media was inoculated with the overnight culture at a ratio of 100 to 1. The culture was grown in 2.5 L -capacity flask at 37 °C and 280 rpm. At OD₆₀₀ of 0.7, the protein expression was induced using Isopropyl β-D-1-thiogalactopyranoside (IPTG; Gold Biotechnology) to a final concentration of 0.4 mM. The culture was incubated for 4 h before being harvested by centrifugation at 11,000 × g for 30 min.

3.3.4 Purification of recombinant LtrK protein

The bacterial cell pellet from two a litre culture was thawed on ice and suspended in 10 ml of sonication buffer (50 mM HEPES, pH 8.0, 50 mM KCl and 20 % glycerol). Lysozyme solution was added to the cell suspension at a final concentration of 100 $\mu\text{g ml}^{-1}$ and the cells were incubated on ice for 15 min. Just before sonication, DTT (Promega) and N-Lauroylsarcosine sodium salt (Sigma-Aldrich) was added to a final concentration of 5 mM and 1.5% respectively. The cells were then sonicated on ice using a Branson digital sonifier with 0.5 s pulse at 50 % amplitude for 1 min. Finally, the supernatant containing soluble proteins was collected by centrifugation at 23,000 x g for 30 min at 4 °C. Four ml of 10 % Triton X-100 was added to the supernatant and the final volume was made to 20 ml with sonication buffer to give the final concentration of Triton X-100 and N-Lauroylsarcosine sodium salt 2 % and 0.7 % respectively. The supernatant was incubated on ice for 1 h and then applied to a 1 ml bed volume Pierce Glutathione Spin Column (Thermo Scientific) that was pre-equilibrated with sonication buffer and incubated for 1 h with agitation. The column was washed with 50 column volumes (CV) of sonication buffer to remove the unbound proteins and finally, the GST-LtrK was eluted with elution buffer (50mM HEPES pH 8.0, 50mM KCl, 20 % glycerol and 20 mM reduced L-Glutathione). The eluted protein was identified by 12 % sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) and concentrated using Amicon Ultra centrifugal filtration units with a 3.5 kDa cut-off (Millipore) according to manufacturer's protocol.

3.3.5 GST tag cleavage from purified GST-LtrK

The GST tag was cleaved by incubating overnight GST-LtrK with Tev protease (Thermo Fisher Scientific) in a w/w ratio of 100:1 in sonication buffer containing 1 mM tris-(2-carboxyethyl)phosphine (TCEP, Gold Biotech) and 0.5 mM EDTA at 4 °C with gentle shaking. Next day, the protein solution was incubated in a 1 ml bed volume Pierce Glutathione Spin Column for 3 h at 4 °C with rotation to allow the cleaved GST tag to bind the beads and the GST-free LtrK was finally collected in the flowthrough. Effective removal of the GST tag was determined by visualizing bands and its corresponding molecular weight using the protein molecular weight standards by 12% SDS-PAGE.

3.3.6 SDS-PAGE analyses

The fractions containing the protein of interest were identified by SDS-PAGE. For PAGE analysis the protein samples were prepared by mixing with 4 x NuPAGE LDS sample buffer (Invitrogen) in 4:1 followed by denaturation at 95 °C for 4 min. The denatured samples were run on the gel using SDS running buffer containing 15.1 g l⁻¹ of Tris, 72 g l⁻¹ of glycine (Ajax Finechem) and 5 g l⁻¹ of SDS. The composition of resolving gel was 12 % acrylamide (Bio-RAD), 0.375 M Tris, 1 % SDS, 0.05 % ammonium persulfate (APS, Sigma), 0.5 % Tetramethylethylenediamine (TEMED, Bio-RAD) and the stacking gel was- 4 % acrylamide, 0.126 M Tris, 0.1 % SDS, 0.05 % APS and 0.5 % TEMED. The gel was stained with Coomassie blue solution (10 % v/v acetic acid, 0.006 % w/v Coomassie Brilliant blue R-250 (CalbioChem)) for at least 30 min with gentle shaking, and then de-stained with de-staining solution (20 % methanol, 10 % Acetic acid) for 1 h.

3.3.7 Protein identification

To identify the protein, the band of interest on the SDS-PAGE was excised with a scalpel and subjected to liquid chromatography–tandem mass spectrometry (LC-MS/MS) analysis at Bioanalytical Mass Spectrometry Facility (BMSF), UNSW.

3.3.8 Native molecular weight determination by gel filtration chromatography.

HiLoad 26/600 Superdex 75 prep grade gel filtration column (GE Healthcare) was equilibrated with 50mM HEPES pH 8.0, 50mM KCl and 20 % glycerol buffer using AKTA purification system (GE Healthcare). Five different molecular weight markers were passed through the column and the elution volume (V_e) was recorded (Table 3.1). The elution volume of blue dextran was regarded as the void volume (V_0). From these data the partition coefficient (K_{av}) was calculated using following equation (Laurent and Killander, 1964):

$$K_{av} = \frac{V_e - V_0}{V_t - V_0} \quad \dots\dots\dots (1)$$

Where K_{av} = partition coefficient, V_e = elution volume, V_0 = void volume and V_t = total column volume (330 ml). A standard curve was generated by plotting K_{av} against the log of protein marker's logarithm of molecular weights (M_r).

Both GST-LtrK and tag free LtrK were passed through the gel filtration column and the V_e was recorded. Using the V_e value the K_{av} was calculated for both GST-LtrK and LtrK from equation 1. Finally, the K_{av} was used to determine M_r from the standard curve, from which the molecular weight was calculated as antilog (M_r).

3.3.9 Circular dichroism (CD) spectroscopy analyses

Circular dichroism was performed on a temperature controlled CD-spectrometer (Chirascan plus) under constant nitrogen flow. The far-UV CD spectra were recorded using 0.1 cm path length quartz cuvette and a protein concentration of 0.35 mg ml⁻¹ in 10 mM HEPES, pH 8.0, 50 mM KCl, 5 % glycerol and 1 mM TCEP buffer. Measurements were carried out over a range of 205 – 260 nm at 4 °C and were averaged over 50 scans.

The secondary structure melting was performed from 4 – 98 °C, with an excitation wavelength of 222 nm with 1 °C interval. The raw data obtained was subsequently converted to the mean residue ellipticity $[\theta]_{\text{mrw}}$ (deg cm² dmol⁻¹) based on following equation (Kelly *et al.*, 2005),

$$[\theta]_{\text{mrw}} = \frac{(\text{MRW} \times \theta_{\text{obs}})}{10 \times d \times c} \quad \dots\dots\dots (2)$$

Where, θ_{obs} = observed ellipticity in degrees; d = path length (=0.1 cm); c = concentration of the protein (=0.35 mg ml⁻¹) and MRW = molar residual weight. The MRW was calculated from following equation

MRW = molecular mass / (N-1); where N = total number of amino acids (Kelly *et al.*, 2005). For LtrK, N = 271 and the molar mass was 30,779 g mol⁻¹.

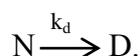
3.3.10 Differential scanning calorimetry

Purified LtrK protein was concentrated to 2.5 mg ml⁻¹ in 50 mM HEPES, pH 8.0, 50 mM KCl, 10 % glycerol, 0.5 mM ATP (NEB) and 5 mM TCEP buffer using Amicon Ultra centrifugal filter device (3 kDa cut-off). The filtrate was used as a reference sample for DSC measurement. The protein sample and the reference buffer were

degassed with stirring under vacuum for 30 min at 4 °C. Thermogram of the protein sample was recorded using a Nano DSC machine (TA instrument) with a cell volume of 0.3 ml and a pressure of 3 atm. Prior to scanning, the cells of the microcalorimeter were prepared by repeated washing with Milli-Q water. Attention was paid during loading the protein sample and buffer reference into the cells to avoid the introduction of air bubble. For each run the calorimeter was set to perform a series of scans with identical parameters; temperature range: 4 to 90 °C, scan rate: 1 °C min⁻¹.

Data analysis was performed with the NanoAnalyze Software v 3.4.0. Buffer baseline was subtracted from the sample data and the data were normalised with respect to the protein concentration. The deconvolution of the thermal transition curves was performed with different transition models incorporated in the software and the temperature (T_m) corresponding to the maximum heat capacity (C_p) was determined.

In order to further analyse the denaturation process, a simple kinetic model for the thermal transition of the LtrK was used, where the native protein- N undergoes an endothermic and irreversible step to a denatured state- D with a first-order rate constant k_d (Thomas and Cavicchioli, 2000; Cavicchioli *et al.*, 2006).



According to this model, the rate constant of the irreversible thermal inactivation (k_d) at a given temperature can be calculated by the following relationship (Sanchez-Ruiz *et al.*, 1988; Thomas and Cavicchioli, 2000),

$$k_d = \frac{v \times C_p}{Q - Q_t} \dots\dots\dots (3)$$

Where, v = scan rate; C_p = Heat capacity at temperature t; Q_t = heat evolved at a given temperature t, and Q = total heat of the process.

The calculated values of k_d (as $\ln k_d$) for LtrK were plotted against the inverse of the absolute temperature ($1/T$). According to the Arrhenius equation, $k = Ae^{(-E_a/RT)}$, the slope (m) of the best linear fit to the data corresponds to $-E_a/R$, with E_a being the activation energy of the process and R being the universal gas constant ($= 8.314 \text{ J mol}^{-1} \text{ K}^{-1}$) (Siddiqui *et al.*, 2002; Siddiqui and Cavicchioli, 2006).

3.4 Results

3.4.1 Expression and purification of LtrK as soluble protein

To obtain a soluble form of LtrK, the N-terminal region encoding the TMDs was deleted and the gene encoding from residue 323 to 592 was synthesized with an N-terminal GST tag and overexpressed in *E. coli*. The insoluble fraction (which represented the majority of the protein) was solubilised using Triton X-100 and N-lauroylsarcosine sodium salt (described in Experimental procedures). The soluble protein was bound to 1 ml glutathione-agarose column and eluted by 20 mM reduced glutathione. SDS-PAGE showed a protein band corresponding to a molecular weight of ~59 kDa as expected from the amino acid sequence (Fig. 3.1, lane 4). GST tag was subsequently cleaved by Tev protease in the glutathione-agarose column and LtrK (~31 kDa) was collected in the flowthrough fraction (Fig. 3.1, lane 6). The bound GST tag (27 kDa) was also eluted by buffer containing 20 mM glutathione (Fig. 3.1, lane 7).

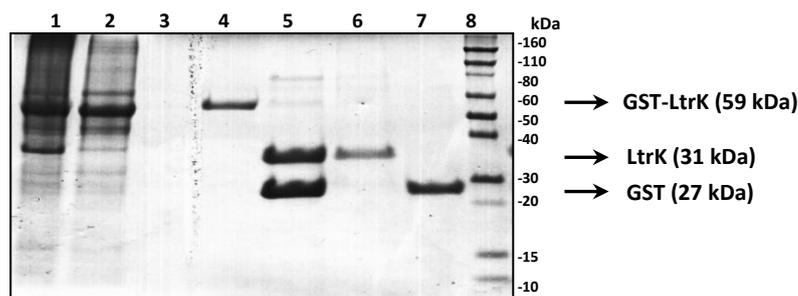


Figure 3.1 Purification of overexpressed LtrK

SDS-PAGE of different fractions of LtrK purification steps. Lanes 1, soluble fraction of sonicated lysate; 2 & 3, flowthrough fractions; 4, eluted GST-LtrK by 20 mM glutathione; 5, Tev protease treated GST tag cleaved LtrK; 6, tag-free LtrK collected with flowthrough; 7, Cleaved GST tag; 8, protein standard marker

The purified LtrK was further analysed by LC-MS/MS to determine the identity of the protein. The snapshot of the result including the list of most abundant proteins is given in Fig. 3.2. The protein hit with highest score was for the histidine kinase from *M. burtonii* which confirmed the identity of LtrK.

MASCOT SCIENCE Mascot Search Results

```

User :
Email :
Search title : TN_270514_2_r.RAW
MS data file : D:\Data\Sydney\2014\May\28\TN\TN_270514_2_r.RAW
Database : NCBInr 26_10_13 (33055681 sequences; 11532217697 residues)
Timestamp : 27 May 2014 at 23:34:13 GMT
Enzyme : Trypsin
Variable modifications : Acrylamide \(C\), Carbamidomethyl \(C\), Oxidation \(M\)
Mass values : Monoisotopic
Protein Mass : Unrestricted
Peptide Mass Tolerance : ± 4 ppm
Fragment Mass Tolerance: ± 0.4 Da
Max Missed Cleavages : 1
Instrument type : ESI-TRAP
Number of queries : 1895
Protein hits : gi|91772717 external sensor signal transduction histidine kinase [Methanococcoides burtonii DSM 6242]
               gi|595706 glutathione S-transferase [unidentified cloning vector]
               gi|387895650 phospholipid-binding domain protein [Pseudomonas fluorescens A506]
               gi|387892422 OmpA family lipoprotein [Pseudomonas fluorescens A506]
               gi|387896416 phosphate ABC transporter substrate-binding protein [Pseudomonas fluorescens A506]
               gi|70729111 succinyl-CoA synthetase subunit beta [Pseudomonas protegens Pf-5]
               gi|518658790 putative outer membrane porin A protein [Pseudomonas]
               gi|435476 cytokeratin 9 [Homo sapiens]
               gi|3201828 major outer membrane lipoprotein I [Pseudomonas oleovorans]
               gi|7331218 keratin 1 [Homo sapiens]
               gi|229587620 osmotically inducible protein c [Pseudomonas fluorescens SBW25]

```

Figure 3.2: LC-MS/MS result for purified LtrK

Protein hit list of LC-MS/MS analyses; red box shows the hit with highest score to be the histidine kinase from *M. burtonii*.

3.4.2 Evaluation of oligomeric state showed LtrK to be a dimer

To evaluate the oligomeric state, an equimolar mixture of GST-LtrK and LtrK was analysed using SEC. The column was calibrated using the standards of known molecular weight markers and a standard curve was generated by plotting their partition coefficient (K_{av}) against the log of their molecular weights (M_r) (Table 3.1).

Table 3.1 Molecular weight markers and their respective M_r , V_e and K_{av}

Protein marker	Molecular weight,	Log (Mwt),	Elution volume,	K_{av}
	Mwt (kDa)	M_r	V_e (ml)	
Blue dextran	100	-	100 (=V ₀)	-
Conalbumin	75	1.9	125	0.12
Ovalbumin	45	1.7	142	0.18
Carbonic anhydrase	29	1.5	165	0.28
Lysozyme	14	1.0	230	0.57
Vitamin B ₁₂	1.3	0.11	270	0.74

The V_e recorded for GST-LtrK and tag free LtrK were 110 and 145 ml respectively. The calculated K_{av} , for GST-LtrK was 0.043 and LtrK was 0.2 which were used to determine molecular weight from the standard curve (Fig. 3.3).

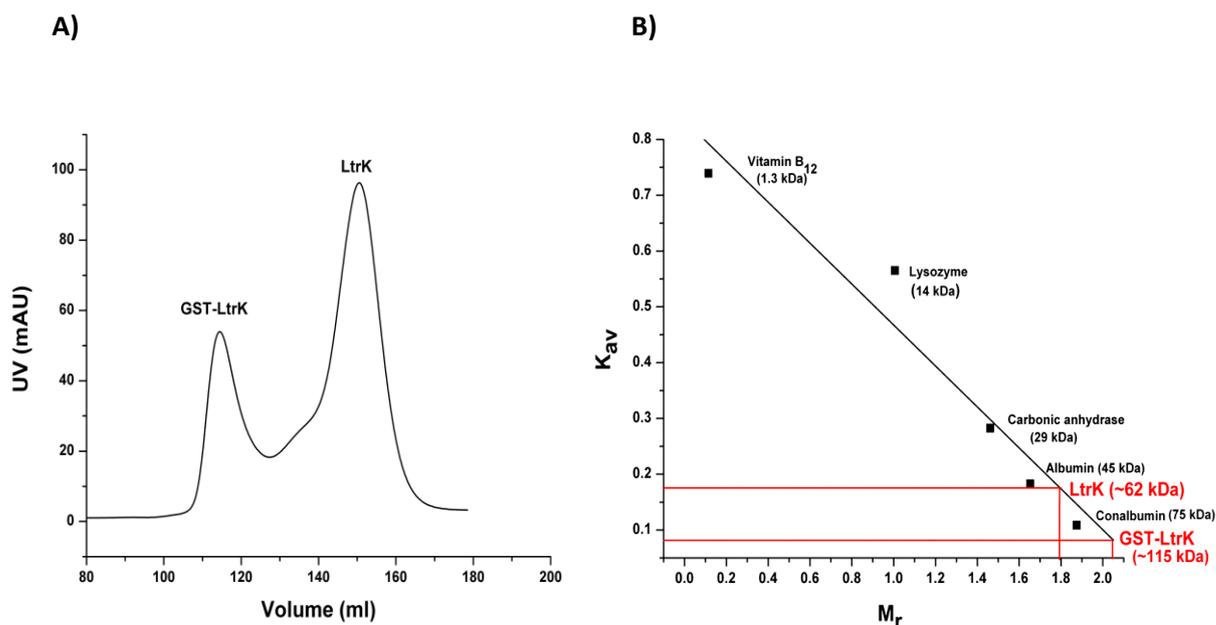


Figure 3.3: Analytical size-exclusion chromatography of GST-LtrK and LtrK

A) A typical chromatogram for GST-LtrK and LtrK protein mixture, B) The calibration graph used for the determination of the molecular weight of the recombinant proteins. The square and the red lines refer to protein standards and recombinant proteins respectively.

The expected molecular weight for GST-LtrK and LtrK based on the amino acid sequence were ~59 and ~31 kDa respectively. Estimated molecular weight of GST-LtrK and LtrK from the standard curve was ~115 and ~62 kDa which correspond closely to the weight of GST-LtrK and LtrK dimer respectively (Fig. 3.3). This indicated that both GST-LtrK and LtrK were purified as a dimer.

3.4.3 Circular dichroism spectroscopy

Far-UV CD spectra analyses were performed in order to determine the presence of secondary structure in LtrK (Siddiqui and Cavicchioli, 2006). The spectra of LtrK at 4 °C displayed two peaks at 205 and 222 nm (Fig. 3.4A) which are characteristics of α -helix (Kelly *et al.*, 2005) confirming the presence of α -helical secondary structure in

LtrK. The melting of secondary structure was performed in a temperature range of 4 °C – 98 °C at 222 nm. At scan rate 1 °C min⁻¹ the estimated melting temperature (T_m) was ~ 47 °C (Fig. 3.4B). However, the spectra of the protein before and after the first temperature melt were not superimposable indicating that LtrK unfolds irreversibly (Fig. 3.4A).

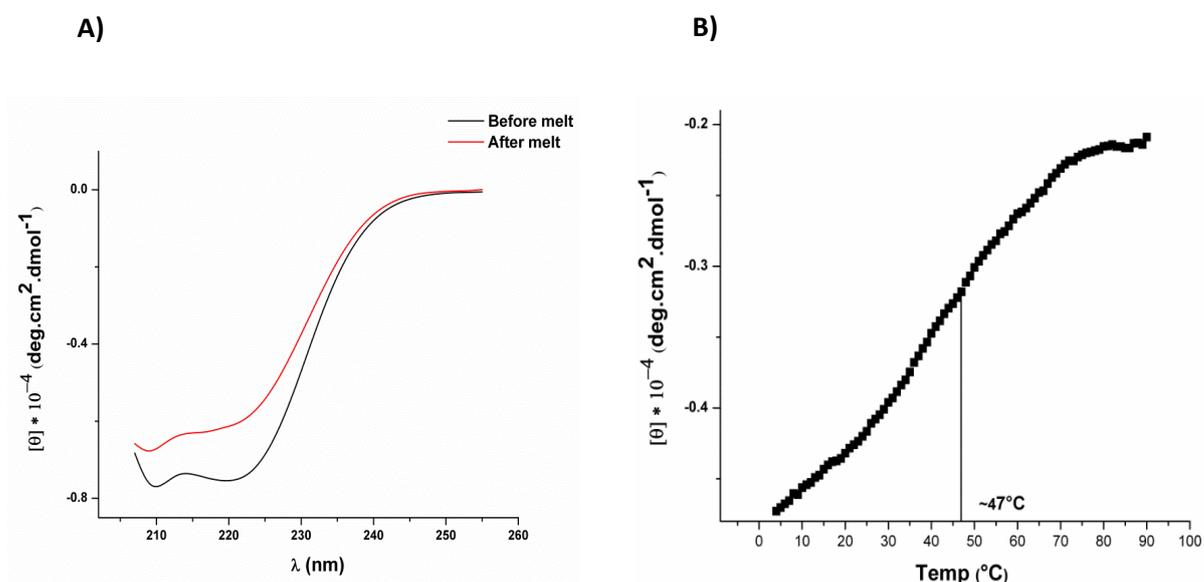


Figure 3.4: Circular dichroism analyses of LtrK

A) Far-UV CD spectra of LtrK at 4 °C both before (black) and after (red) thermal unfolding; B) Secondary structure melting profile at 222 nm showing a T_m 47 °C

3.4.4 Differential Scanning Calorimetry (DSC)

The thermostability of purified LtrK was examined by DSC. At a scan rate 1 °C min⁻¹ the temperature corresponding to the maximum heat capacity (T_m) was 63 °C. Rescanning of the protein sample after it was heated beyond the transition peak and then cooled to 4 °C resulted in no further increases in heat capacity. This indicated that the

thermal unfolding of LtrK is an irreversible process. This finding is consistent with the Far-UV CD spectra scan result (Fig. 3.4).

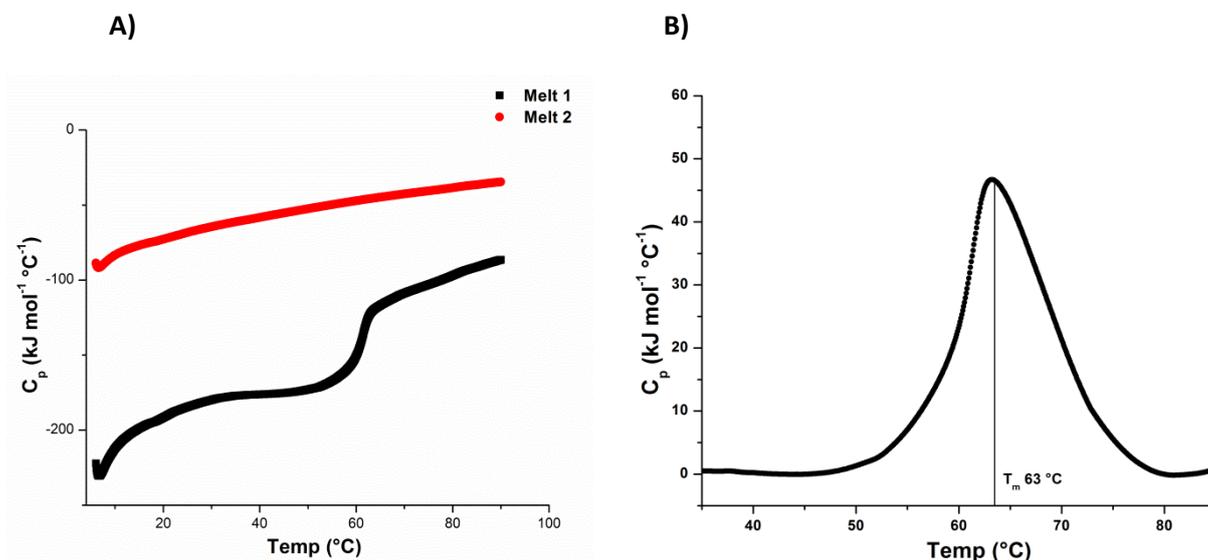


Figure 3.5: DSC curve of LtrK

A) Irreversible thermal unfolding of LtrK before baseline correction. Melt 1 (black) is the first scan of the protein sample in a temperature range of 4 – 90 °C and melt 2 (red) is the rescan of the same sample after cooling it down to 4 °C. B) Thermal unfolding of LtrK (after baseline correction) showing T_m to be 63 °C.

Arrhenius plot (Fig. 3.7) was generated using the DSC data for the determination of activation energy (E_a) required for the denaturation of LtrK. From the $\ln K_d$ vs. $1/T$ graph the slope (m) of the best linear fit line was determined to be 32.1 (Fig. 3.7) which was used to calculate the E_a as $m \times R = (32.1 \times 1000K) \times 8.314 \text{ J mol}^{-1} \text{ K}^{-1} = 267 \text{ kJ mol}^{-1} \text{ K}^{-1}$

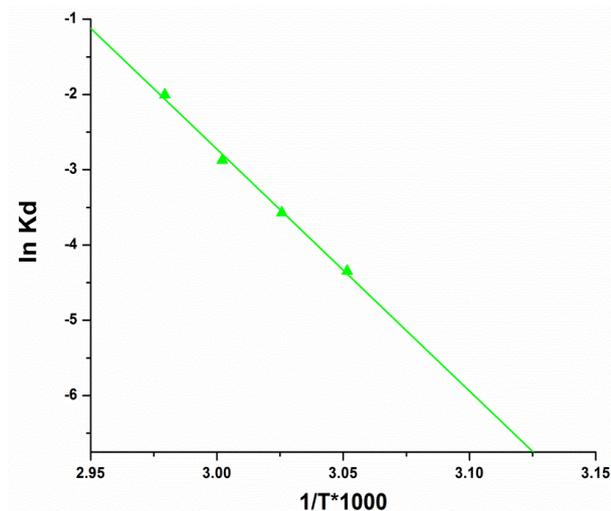


Figure 3.6: Arrhenius plot for thermal denaturation of LtrK

The green line represents the linear fit to the data, which was used to calculate the activation energy of the reaction.

3.5 Discussion

The GST-LtrK protein, which was overexpressed in *E. coli* at 37 °C, was expressed with the insoluble inclusion bodies and was later solubilised from the inclusion bodies by using protein unfolding/refolding method. In spite of lacking hydrophobic TMDs and having GST-tag at the N-terminus which tends to improve protein solubility (Esposito and Chatterjee, 2006) the kinase was expressed as an insoluble protein. Typically, protein solubility can be influenced by a number of factors including pH, ionic strength, temperature and the presence of various solvent additives (Kramer *et al.*, 2012). Previous studies on *M. burtonii* proteins, EF-2, RNA polymerase E/F, chaperonin and Ctr3, showed that temperature shift of *E. coli* culture growth from 37 °C to ~15 °C during induction of expression (using IPTG) enhances protein expression and stability in *E. coli* (Thomas and Cavicchioli, 2000; De Francisci *et al.*, 2011; Pilak *et al.*, 2011; Taha *et al.*, 2016). However, in this study similar approach was

attempted to induce the LtrK expression with IPTG at 16 °C but it was not successful and GST-LtrK remained insoluble (data not shown). Some studies have shown that GST in *E. coli* is a poor solubility enhancer (Esposito and Chatterjee, 2006). Maltose binding protein (MBP) and thoredoxin (Trx) have been found to be better solubility enhancer for some protein in *E. coli* (Esposito and Chatterjee, 2006). In fact, a psychrophilic trypsin from Atlantic cod was unstable and sensitive to molecular aggregation during recombinant expression even at low temperature, 18 °C (Jónsdóttir *et al.*, 2004). Expression of this protein as a fusion partner to Trx in *E. coli* resulted a stable and highly purified trypsin (Jónsdóttir *et al.*, 2004). It would be interesting to find out if LtrK can express as a soluble protein with a MBP or Trx fusion partner.

Nevertheless, the LtrK was successfully solubilised from the inclusion bodies by protein denaturing-refolding technique using sarkosyl and Triton X-100. Sarkosyl is an anionic detergent that can disrupt the hydrophobic interactions present in the protein aggregates and improve solubility (Park *et al.*, 2011). But as sarkosyl molecules encapsulate proteins (including the affinity tag), presence of sarkosyl can significantly reduce the binding affinity of GST tag to glutathione (Tao *et al.*, 2010). This problem can be overcome by introducing a non-ionic detergent such as Triton X-100 followed by sarkosyl treatment. Addition of Triton X-100 initiates the formation of micelle structures that incorporate sarkosyl molecules from the solution and decrease the apparent concentration of sarkosyl surrounding GST tagged protein, facilitating proper protein refolding (Tao *et al.*, 2010; Park *et al.*, 2011). This method has been proven to be useful to recover several GST-fusion proteins from insoluble aggregates (Park *et al.*, 2011). The LtrK unfolding/refolding was carried out on ice with a 2:1 volume ratio of Triton X-100 and sarkosyl and this method proved effective to retrieve the soluble LtrK

from inclusion bodies. Moreover, the far-UV CD spectra of the soluble LtrK showed the characteristic α -helical peaks (Fig. 3.4A) indicating the protein is properly folded.

SEC studies revealed the oligomeric state of LtrK to be dimeric (Fig. 3.3). Dimerization has been found to be a key factor for autophosphorylation, as SKs catalyse the phosphorylation of a conserved histidine residue via trans-autophosphorylation mechanism where one monomer phosphorylates the second monomer within a dimer (Ninfa *et al.*, 1993; Surette *et al.*, 1996; Dutta *et al.*, 1999; Robinson *et al.*, 2000). It is not yet clear whether activation of a SK induces a monomer to dimer transition. But some recombinant SKs were found to form dimer without the TMDs *in vitro* and many SKs had been identified to exist as a dimer or higher oligomer *in vivo* irrespective of ligand binding or signal sensing. For example, the purified cytoplasmic domain of EnvZ of *E. coli* forms dimer (Hidaka *et al.*, 1997); KdpD, the turgor pressure sensor in *E. coli*, also forms dimer in the cell membrane and the dimeric state doesn't change upon phosphorylation (Heermann *et al.*, 1998); whereas VirA, the transmembrane SK of *A. tumefaciens* involved in virulence, was found to exist as a dimer *in vivo* and the dimerization is independent of any signal sensing or ligand binding (Pan *et al.*, 1993). Therefore, the dimeric state of the cytoplasmic LtrK is indicative of the *in vitro* dimerization of the protein regardless of the presence of any ligand or environmental signal.

The structural stability of purified LtrK was analysed by using far UV-CD spectrometry and DSC techniques. The far-UV CD data indicated that the protein has secondary structure that unfolds irreversibly (Fig. 3.4). The analyses on native protein unfolding by DSC indicated that the thermal unfolding of LtrK is irreversible (Fig. 3.5). So collectively the CD and DSC data indicated that purified cytoplasmic domain of

LtrK is properly folded with secondary and tertiary structure and the thermal unfolding is irreversible. A couple of large multi-domain proteins of *M. burtonii* studied to-date unfold irreversibly, *i.e.* EF2 (Thomas and Cavicchioli, 2000), RNA polymerase E/F (De Francisci *et al.*, 2011). As hydrophobic groups of these proteins get exposed to the aqueous solvent during unfolding, it initiates irreversible aggregation of the protein which results in the irreversible thermal unfolding profile. Moreover, the T_m of irreversibly unfolding proteins is typically scan rate dependent and the T_m shifts to a higher temperature when heated with higher rate as the protein gets less time at any specific temperature (Lepock *et al.*, 1992; Thomas and Cavicchioli, 2000). The scan-rate dependency of LtrK's T_m was analysed in Chapter 6.

The E_a for LtrK thermal denaturation was determined to be $267 \text{ kJ mol}^{-1} \text{ K}^{-1}$ from Arrhenius plot (Fig. 3.6). Typically, the activation energy required for the unfolding of a psychrophilic protein is significantly lower than the mesophilic and thermophilic homologues indicating their less stable structure (Siddiqui *et al.*, 2002; Siddiqui and Cavicchioli, 2006). For example, the activation energy required for thermal unfolding of psychrophilic EF2 of *M. burtonii* is $203 \text{ kJ mol}^{-1} \text{ K}^{-1}$, which is significantly lower than that for thermophilic *M. thermophila* ($351 \text{ kJ mol}^{-1} \text{ K}^{-1}$) (Thomas and Cavicchioli, 2000). It would be interesting to compare the activation thermodynamics of thermal denaturation of LtrK with similar SKs from thermophiles such as *M. thermophila*.

In summary, the cytoplasmic catalytic domain of SK, LtrK was successfully purified as a soluble protein in a proper folded dimeric state which allowed further experiments to be performed on the protein including phosphorylation and phosphotransfer assays described in Chapter 5 and 6.

3.6 References

- Albanesi, D. *et al.* (2004) The membrane fluidity sensor DesK of *Bacillus subtilis* controls the signal decay of its cognate response regulator. *J. Bacteriol.* 186: 2655-2663.
- Allen, M.A. *et al.* (2009) The genome sequence of the psychrophilic archaeon, *Methanococcoides burtonii*: the role of genome evolution in cold adaptation. *ISME J.* 3: 1012-1035.
- Cavicchioli, R. *et al.* (2006) Proteins from Psychrophiles. *Methods Microbiol.* 35: 395-436.
- De Francisci, D. *et al.* (2011) The RNA polymerase subunits E/F from the Antarctic archaeon *Methanococcoides burtonii* bind to specific species of mRNA. *Environ. Microbiol.* 13: 2039-2055.
- Dutta, R. *et al.* (1999) Histidine kinases: diversity of domain organization. *Mol. Microbiol.* 34: 633-640.
- Esposito, D., and Chatterjee, D.K. (2006) Enhancement of soluble protein expression through the use of fusion tags. *Curr. Opin. Biotechnol.* 17: 353-358.
- Feller, G. *et al.* (1998) Expression of psychrophilic genes in mesophilic hosts: assessment of the folding state of a recombinant α -amylase. *Appl. Environ. Microbiol.* 64: 1163-1165.
- Gerday, C. *et al.* (1997) Psychrophilic enzymes: a thermodynamic challenge. *Biochim. Biophys. Acta-Protein Structure and Molecular Enzymology* 1342: 119-131.
- Gräslund, S. *et al.* (2008) Protein production and purification. *Nat. Meth.* 5: 135-146.
- Gustafsson, C. *et al.* (2004) Codon bias and heterologous protein expression. *Trends Biotechnol.* 22: 346-353.
- Heermann, R. *et al.* (1998) The turgor sensor KdpD of *Escherichia coli* is a homodimer. *Biochim. Biophys. Acta-Biomembranes* 1415: 114-124.
- Hidaka, Y. *et al.* (1997) Demonstration of dimer formation of the cytoplasmic domain of a transmembrane osmosensor protein, EnvZ, of *Escherichia coli* using Ni-histidine tag affinity chromatography. *FEBS Lett.* 400: 238-242.
- Jónsdóttir, G. *et al.* (2004) Recombinant cold-adapted trypsin I from Atlantic cod—expression, purification, and identification. *Protein Expr. Purif.* 33: 110-122.
- Kelly, S.M. *et al.* (2005) How to study proteins by circular dichroism. *Biochim. Biophys. Acta-Proteins and Proteomics* 1751: 119-139.

Kramer, R.M. *et al.* (2012) Toward a molecular understanding of protein solubility: increased negative surface charge correlates with increased solubility. *Biophys. J.* 102: 1907-1915.

Laurent, T.C., and Killander, J. (1964) A theory of gel filtration and its experimental verification. *J. Chromatogr. A* 14: 317-330.

Lepock, J.R. *et al.* (1992) Influence of transition rates and scan rate on kinetic simulations of differential scanning calorimetry profiles of reversible and irreversible protein denaturation. *Biochemistry* 31: 12706-12712.

Li, J. *et al.* (2014) Characterization of an archaeal two-component system that regulates methanogenesis in *Methanosaeta harundinacea*. *PLoS one* 9.

Liu, W., and Hulett, F.M. (1997) *Bacillus subtilis* PhoP binds to the phoB tandem promoter exclusively within the phosphate starvation-inducible promoter. *J. Bacteriol.* 179: 6302-6310.

Makino, K. *et al.* (1989) Signal transduction in the phosphate regulon of *Escherichia coli* involves phosphotransfer between PhoR and PhoB proteins. *J. Mol. Biol.* 210: 551-559.

Ninfa, E.G. *et al.* (1993) Mechanism of autophosphorylation of *Escherichia coli* nitrogen regulator II (NRII or NtrB): trans-phosphorylation between subunits. *J. Bacteriol.* 175: 7024-7032.

Novy, R. *et al.* (2001) Overcoming the codon bias of *E. coli* for enhanced protein expression. *Innovations* 12: 1-3.

Pan, S.Q. *et al.* (1993) Preformed dimeric state of the sensor protein VirA is involved in plant-*Agrobacterium* signal transduction. *Proc. Natl. Acad. Sci. USA.* 90: 9939-9943.

Park, D.-W. *et al.* (2011) Improved recovery of active GST-fusion proteins from insoluble aggregates: solubilization and purification conditions using PKM2 and HtrA2 as model proteins. *BMB reports* 44: 279-284.

Pilak, O. *et al.* (2011) Chaperonins from an Antarctic archaeon are predominantly monomeric: crystal structure of an open state monomer. *Environ. Microbiol.* 13: 2232-2249.

Robinson, V.L. *et al.* (2000) A tale of two components: a novel kinase and a regulatory switch. *Nat. Struct. Mol. Biol.* 7: 626-633.

Sanchez-Ruiz, J.M. *et al.* (1988) Differential scanning calorimetry of the irreversible thermal denaturation of thermolysin. *Biochemistry* 27: 1648-1652.

- Shi, L., and Hulett, F.M. (1999) The cytoplasmic kinase domain of PhoR is sufficient for the low phosphate-inducible expression of Pho regulon genes in *Bacillus subtilis*. *Mol. Microbiol.* 31: 211-222.
- Siddiqui, K. *et al.* (2002) Thermodynamic activation properties of elongation factor 2 (EF-2) proteins from psychrotolerant and thermophilic Archaea. *Extremophiles* 6: 143-150.
- Siddiqui, K.S., and Cavicchioli, R. (2006) Cold-adapted enzymes. *Annu. Rev. Biochem.* 75: 403-433.
- Surette, M.G. *et al.* (1996) Dimerization is required for the activity of the protein histidine kinase CheA that mediates signal transduction in bacterial chemotaxis. *J. Biol. Chem.* 271: 939-945.
- Taha *et al.* (2016) Single TRAM domain RNA-binding proteins in Archaea: functional insight from Ctr3 from the Antarctic methanogen *Methanococcoides burtonii*. *Environ Microbiol.*
- Tao, H. *et al.* (2010) Purifying natively folded proteins from inclusion bodies using sarkosyl, Triton X-100, and CHAPS. *BioTechniques* 48: 61-64.
- Thomas, T., and Cavicchioli, R. (2000) Effect of temperature on stability and activity of elongation factor 2 proteins from Antarctic and thermophilic methanogens. *J. Bacteriol.* 182: 1328-1332.
- Thomas, T., and Cavicchioli, R. (2002) Cold adaptation of archaeal elongation factor 2 (EF-2) proteins. *Curr. Protein Pept. Sci.* 3: 223-230.
- Wolanin, P.M. *et al.* (2002) Histidine protein kinases: key signal transducers outside the animal kingdom. *Genome Biol.* 3: 3013.

Chapter 4

Purification and biophysical characterization of the response regulator, LtrR

4.1 Abstract

LtrR is a functionally annotated RR with a DNA binding output domain that forms an operon with SK, LtrK in *M. burtonii*. LtrR was successfully overexpressed and purified as a soluble monomeric protein with N-terminal His₆ tag. Far UV-CD data indicated the presence of α -helical secondary structure in LtrR that unfolds reversibly. The tyrosine fluorescence spectroscopy and near-UV CD analyses indicated that the purified protein is properly folded with tertiary structure. In addition, DSC analyses revealed that the thermal unfolding of LtrR is reversible with a melting temperature (T_m) of 57 °C. The DNA binding affinity of the purified LtrR was assessed using affinity chromatography but proved unsuccessful. The research provided basic information about the biophysical properties of recombinantly expressed LtrR.

4.2 Introduction

Unlike SKs, RRs are cytoplasmic proteins typically containing a DNA binding and a receiver (REC) domain (described in Chapter 1). Recombinant RRs (bacterial and archaeal) had been expressed in *E. coli* as stable full length proteins unlike SKs that are mostly expressed as TMDs truncated protein.

Recombinant proteins are produced as a fusion to an affinity tag so that they can be purified from their crude biological source using an affinity technique. The most commonly used affinity tags include GST, MBP and polyhistidine tags. Some affinity tags, *e.g.* GST and MBP, are also able to enhance solubility of the partner proteins to which they are attached (describe in Chapter 3, Section 3.2). In contrast, polyhistidine tag typically does not alter the solubility properties of the target protein and yet most frequently used because of having smaller size (< 1 kDa) compared to other available tags, *e.g.* GST (>26 kDa) and MBP (~ 42.5 kDa), which rarely affects the biological or biochemical activity of the fused protein (Gräslund *et al.*, 2008). Moreover, the polyhistidine-tagged proteins can be purified using a relatively simple protocol using immobilized metal ion (typically Ni²⁺) affinity chromatography (Gräslund *et al.*, 2008). Apart from the affinity tag, factors influencing the recombinant protein expression includes bacterial strain type, culture condition, codon usage bias of *E. coli* which were describe in Chapter 3 Section 3.2.

To study the autophosphorylation and phosphotransfer activities *in vitro* both SKs and RRs are required to be purified. The assessment of autophosphorylation activity involves incubation of SKs with ATP while the phosphotransfer activity involves incubation of phosphorylated SKs with RRs. It is often observed that SKs and RRs are

purified as a fusion partner of two different tags so that after phosphotransfer reaction phosphorylated RRs can be separated from the SKs to study its stability and DNA binding property. For example, to characterize DesK/DesR TCS of *B. subtilis* the SK, DesK was expressed with hexa-histidine tag (His₆) and the RR, DesR was expressed with GST tag (Albanesi *et al.*, 2004). Expressing with two different tags had allowed the isolation of GST-DesR after incubating with His₆-DesK for phosphotransfer reaction using glutathione-based affinity purification (Albanesi *et al.*, 2004). Similar technique was applied for another TCS of *B. subtilis*, PhoR/PhoP purification, but this time the SK, PhoR had GST tag while the RR, PhoP had His₁₀ tag (Liu and Hulett, 1997).

Most of the RRs play transcriptional regulatory role using its DNA-binding domain. The DNA-binding ability of the RRs can be assessed by different biochemical techniques including electrophoretic mobility shift assay (EMSA), chromatin immunoprecipitation followed by sequencing (ChIP-Seq), *in vitro* DNA affinity purification and sequencing (IDAP-Seq). EMSA is a common electrophoresis technique used to study protein-DNA interaction but this method requires DNA probe to which the target protein is expected to bind. Therefore, this method doesn't apply to those proteins for which the target DNA binding sequence is completely unknown. ChIP-seq is an *in vivo* technique which involves cross-linking of the protein with its associated DNA in the cell followed by cell lyses and selective immunoprecipitation of DNA-protein complex using a protein-specific antibody. The bound DNA is then purified and sequenced. The success of ChIP assay typically relies on the antibodies to specifically identify the target protein. In contrast, IDAP-Seq is an *in vitro* analog of ChIP-Seq in

which purified His-tagged protein is mixed with genomic DNA, the DNA bound to protein is isolated by affinity purification, and the bound DNA is analysed by high throughput DNA sequencing. This approach has been used for few transcriptional regulators to analyse DNA binding (Château *et al.*, 2013; Smith and Grossman, 2015). The similar approach had been successfully used to study RNA binding of RNA pol E/F and Ctr3 of *M. burtonii* in which the His-tagged protein was incubated with whole genomic RNA instead of DNA (De Francisci *et al.*, 2011; Taha *et al.*, 2016).

The functionally annotated RR of *M. burtonii*, LtrR is a small (~20kDa) protein that possesses an N-terminal DNA binding HTH domain and a C-terminal receiver, REC domain (described in Chapter 2, section 2.3.1). This chapter describes the overexpression and purification procedure of recombinant LtrR with an N-terminal His₆ tag and the biophysical characteristics of purified LtrR including the oligomeric state, secondary and tertiary structural properties and thermal unfolding/refolding transition. Different spectrophotometric techniques (*i.e.* far and near-UV CD, FS), and calorimetric technique (DSC) were used for the biophysical characterization which were described in Chapter 3 (section 3.2). This chapter also describes an attempt to determine the DNA-binding affinity of LtrR by using whole genomic DNA binding assay.

4.3 Experimental procedures

4.3.1 Cloning

The LtrR gene was commercially synthesized and subsequently cloned in an expression vector pJexpress404 from DNA 2.0 California, USA. The gene was codon

optimised for expression in *E. coli* and the plasmid contained N-terminal His₆ tag and an ampicillin resistance gene. The plasmid map is shown in Appendix 2.

4.3.2 Overexpression and purification of recombinant LtrR

E. coli competent cell preparation and transformation was performed as described for LtrK in Chapter 3 Section 3.3.2. Recombinant LtrR was overexpressed in competent *E. coli* strain BL21 (DE3). The transformed cells were grown as described for recombinant LtrK overexpression in Chapter 3 Section 3.3.2. At an OD₆₀₀ of 0.7 the protein expression was induced using 1 mM IPTG. The culture was incubated for 4 h before being harvested by centrifugation at 11,000 × *g* for 30 min.

The cell pellet was suspended in 40 mL of lyses buffer (20 mM HEPES pH 7.4, 250 mM NaCl and 2 tablets of complete EDTA free protease inhibitor (Roche)). The cell suspension was subsequently passed through a French cell press and a pressure of 1240 psi was applied to lyse the cells. The crude cell lysate was centrifuged at 10,000 × *g* for 30 min at 4 °C. The supernatant was collected and filtered through a 0.45 µm syringe filter (Millipore). The clear lysate was loaded onto a nickel charged 5ml HiTrap affinity column which was pre-equilibrated with buffer A (20 mM HEPES pH 7.4, 500 mM NaCl). The protein was eluted from the column fitted to an AKTA protein purification system using a step gradient of buffer B (20 mM HEPES pH 7.4, 500 mM NaCl and 1 M imidazole). The fractions containing recombinant protein were identified by SDS-PAGE analysis, pooled and concentrated to a volume of 8 ml using Amicon Ultra centrifugal filtration units with a 3.5 kDa cut-off.

The collected protein was further purified in a HiLoad 26/600 Superdex 75 prep grade gel filtration column which was pre-equilibrated with gel filtration (GF) buffer (20 mM HEPES pH 7.4, 250 mM NaCl, 1 mM TCEP) and the fractions containing the protein were identified using 12% SDS-PAGE and concentrated as described above. The whole purification procedure is schematically represented in Fig. 4.1.

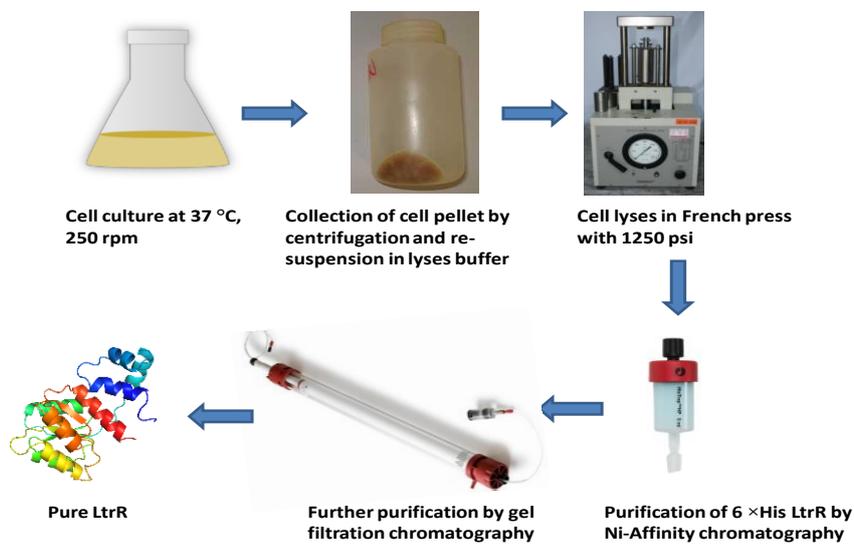


Figure 4.1 Schematic representation of LtrR purification

The protein concentration was measured using Bradford assay (Bradford, 1976), using the Bio-Rad Protein Assay Kit as instructed by the manufacturer. The protein identity was confirmed using LC-MS/MS as described for LtrK in Chapter 3 section 3.2.6.

4.3.3 Oligomeric state determination by gel filtration chromatography

HiLoad 26/600 Superdex 75 prep grade gel filtration column was equilibrated with 20 mM HEPES pH 7.4, 500 mM NaCl and 2 mM TCEP buffer. The protocol of recording elution volume using protein standards and the calculation were as described for LtrK in Chapter 3 section 3.2.8.

4.3.4 Intrinsic tyrosine fluorescence analysis

Fluorescence spectrometry was performed with a luminescence spectrometer LS50B (Perkin Elmer) that was connected to a peripheral peltier temperature controller unit. A volume of 200 μL of protein at a final concentration of 1mg mL^{-1} was loaded into a clean quartz cuvette of 1 cm path length and placed into the LS50B. The fluorescence spectral data were acquired with 5 nm slit width for both excitation and emission, with a scan rate of 100 nm min^{-1} , at an excitation wavelength of 275 nm and an emission wavelength range of 290 – 350 nm. All fluorescence experiments were performed thrice, and averaged automatically by the instrument software FL WinLab. The melting of the protein was performed from 4 – 94 $^{\circ}\text{C}$, with an excitation and emission wavelength of 275 nm and 308 nm respectively, with a temperature interval of 1 $^{\circ}\text{C}$, and a 5 nm slit width. The control buffer run without protein was performed under identical conditions and subtracted from the experimental run.

4.3.5 Circular dichroism (CD) analyses

Both far-UV and near-UV CD analyses of LtrR were performed on a temperature controlled CD-spectropolarimeter (Jasco model J-810) under constant nitrogen flow. The raw data obtained was converted to the mean residue ellipticity, $[\theta]_{\text{mrw}}$, in $\text{deg cm}^2 \text{dmol}^{-1}$ as for LtrR described in Chapter 3 section 3.2.7.

4.3.5.1 Far-ultraviolet (UV) CD spectra

The far-UV CD spectra were recorded using a 0.1 cm path length quartz cuvette and a protein concentration of 0.05 mg ml^{-1} in 15 mM HEPES, pH 7.5 buffer containing 100 mM NaCl. All measurements were carried out at 4 °C over a range of 203 – 260 nm and were averaged over 50 scans. The secondary structure melting was performed from 4 – 98 °C, at an excitation wavelength of 222 nm with 1 °C temperature interval.

4.3.5.2 Near- UV CD spectra

The near-UV CD spectra were recorded using 1 cm path length quartz cuvette and a protein concentration of 1.2 mg ml^{-1} in 15 mM HEPES, pH 7.5 buffer containing 100 mM NaCl. The measurements were carried out at 4 °C as described for the far-UV CD.

4.3.6 Differential scanning calorimetry (DSC)

DSC of LtrR was performed using a protein concentration of 2.5 mg ml^{-1} in 20 mM HEPES, pH 7.4 buffer containing 250 mM NaCl and 3 mM TCEP. The scanning protocol and data analyses procedure were the same as for LtrK described in Chapter 3 section 3.2.8.

4.3.7 DNA binding assay of LtrR

4.3.7.1 Growth of *M. burtonii* culture and genomic DNA extraction

M. burtonii culture was grown in liquid modified methanogen growth media (MFM). The media was prepared by adding 0.335 g of KCl, 6 g of MgCl₂·6H₂O, 1 g of MgSO₄·7H₂O, 0.25 g of NH₄Cl, 0.14 g of CaCl₂·2H₂O, 23.37 g of NaCl, 2 mg of Fe(NH₄)₂(SO₄)₂·6H₂O, 1 mg of resazurin (Sigma), 5 g of trimethylamine HCl (Sigma), 2 g of yeast extract, 10 ml of 100 × vitamin stock solution (containing 2 mg l⁻¹ biotin, 2 mg l⁻¹ folic acid, 2 mg l⁻¹ pyridoxine HCl, 10 mg l⁻¹ thiamine HCl, 5 mg l⁻¹ riboflavin, 5 mg l⁻¹ nicotinic acid, 5 mg l⁻¹ DL-Ca pantothenate, 0.1 mg l⁻¹ vitamin B12, 5 mg l⁻¹ p-aminobenzoic acid, and 5 mg l⁻¹ lipoic acid.), 10 ml of 100 × mineral stock solution (containing 1.5 g l⁻¹ nitrilo-triacetic acid, 3 g l⁻¹ MgSO₄·7H₂O, 0.5 g l⁻¹ MnSO₄·7H₂O, 1 g l⁻¹ NaCl, 0.1 g l⁻¹ FeSO₄·7H₂O, 0.1 g l⁻¹ CoSO₄, 0.1 g l⁻¹ CaCl₂·2H₂O, 1 g l⁻¹ ZnSO₄, 0.01 g l⁻¹ CuSO₄·5H₂O, 0.01 g l⁻¹ AlK(SO₄)₂, 0.01 g l⁻¹ H₃BO₃, and 0.01 g l⁻¹ Na₂MoO₄·2H₂O.), 0.1 g of sodium acetate and 0.14 g of K₂HPO₄ in 800 ml of milli-Q water. The solution was bubbled with N₂ gas for 15 min followed by 80% N₂:20% CO₂ gas for another 15 min with continuous stirring to purge the oxygen content within the culture to create an anaerobic culture condition. Then, 0.5 g of cysteine·HCl (pre-dissolved in 1ml dH₂O), and 2.52 g of Na₂CO₃ was added to the solution and the volume was adjusted to 1 L. The solution was left to stir, with continuous bubbling with 80% N₂:20% CO₂ gas, for 30 min. The pH of MFM was adjusted to 6.8 using 10M HCl and the media was dispensed into clean serum bottles (120 ml capacity) as 100mL of culture per serum bottle. The serum bottles containing 100 ml MFM media were flushed with 80% N₂:20% CO₂ gas for 15 min and plugged with a butyl rubber stopper

and sealed with aluminium crimp seal. The head space of each bottle was flushed with the same gas again using an inlet and outlet needle for ~10 min. Finally the media containing bottles were autoclaved, and swirled overnight to dissolve the precipitated salts. A volume of 1 ml of anaerobic 2.5 % Na₂S was then injected into each MFM culture bottle and swirled for another two hours before injecting 1mL of *M. burtonii* culture into each bottle. The cultures were then incubated at 23 °C without shaking until they were ready for harvesting.

M. burtonii cells (100 ml) were harvested at late logarithmic phase (optical density at 620 nm of 0.3) by centrifugation at 4000 x g for 30 min. The genomic DNA was extracted by using genomic DNA purification kit (Thermo Fisher Scientific) according to the manufacturer's protocol. The DNA concentration was measured using a Nanodrop spectrometer (Thermo Fisher Scientific).

4.3.7.2 Fragmentation of whole genomic DNA

Four samples of 50 µg of genomic DNA were each dissolved in buffer containing 20 mM HEPES, pH 8.0, 50 mM KCl and sheared to get 200 – 500 bp DNA fragments by sonication using a Branson 250 microtip sonicator in five 30-s cycles on ice with 0.5 s pulse at 50 % amplitude. 1 µg of sonicated DNA was subjected to gel electrophoresis using 1 % (w/v) agarose gel to determine the length of the fragmented DNA. The gel was prepared by melting 1mg of agarose (Sigma) in 100 ml of TAE buffer containing 40 mM Tris acetate, pH 8.2 and 1 mM EDTA; 1 µl of 0.5 mg ml⁻¹ ethidium bromide (Bio-Rad) solution was added to the gel just before pouring on the casting tray. The fragmented DNA was mixed with the gel loading dye (Thermo Fisher Scientific) and

run on the gel using running buffer (TAE). The DNA band was visualized using gel imaging and documentation (Gel-Doc) system (Bio-Rad).

4.3.7.3 LtrR phosphorylation

Radioactive phosphorylation of His₆ LtrR by purified LtrK using radioactive [γ ³²P]-ATP showed that LtrR can get phosphorylated by LtrR (described in Chapter 5). Therefore, for DNA binding assay purified His₆ LtrR was phosphorylated using recombinant cytoplasmic GST-LtrK and ATP (NEB). 100 μ l of glutathione-agarose beads, Glutathione Sepharose 4B (GE healthcare), were washed with phosphorylation buffer (P buffer; 50 mM HEPES, pH 8.0, 50 mM KCl, 5 mM MgCl₂, 2 mM TCEP and 20 % glycerol) in a 1 ml gravity flow column (Bio-Rad) and incubated with 1.5 mg of LtrK protein at room temperature for 30 min. The unbound proteins were washed off and 200 μ l of P buffer containing 10 mM ATP was added to the beads and incubated for 30 min to conduct autophosphorylation of LtrK. The beads were thoroughly washed with P buffer to get rid of free ATP. Then, 5 mg of LtrR in 200 μ l of binding buffer (50 mM HEPES pH 8.0, 50 mM KCl, 5 mM MgCl₂) was added to the column to allow phosphotransfer reaction and the flowthrough containing phosphorylated LtrR (LtrR-P) was collected. A further 200 μ l of the same buffer was passed through the column and added to the previous flow through.

4.3.7.4 Overexpression and purification of N-terminal HTH domain deleted LtrR*

To perform a control experiment for DNA-binding assay, the HTH domain deleted LtrR* with a His₆ fusion tag at N-terminal region was purified. This mutated

LtrR* doesn't have the DNA binding ability as it lacks the HTH domain and thus can be used as a negative control for DNA binding assay. The deletion mutation was carried out using Phusion Site - Directed Mutagenesis kit (Thermo Fisher Scientific). The wild type plasmid construct containing commercially synthesized LtrR gene, pJexpress404, served as the template for the mutagenesis reaction. The primer sequences and details of PCR cycle is presented in Appendix 3. The PCR amplified DNA was circularized by incubating 20 ng of PCR product with 0.5 μ l of T4 DNA ligase (provided with the kit) for 5 min. The circularized plasmid was cloned into a cloning strain, *E. coli* DH5a. The deletion mutation was confirmed by DNA sequence analysis. The HTH domain deleted LtrR mutant gene was overexpressed and the mutant protein was purified as described for wild type in this chapter.

4.3.7.5 Binding of whole genomic DNA with phosphorylated LtrR

The collected LtrR-P was bound to 150 μ l of Talon Co⁺ affinity resin (Clontech) pre-equilibrated with binding buffer in the 1 ml gravity flow column. In parallel, the HTH domain deleted LtrR* (5 mg) was also bound to 150 μ l of Talon Co⁺ resin in a separate 1 ml gravity flow column. Both LtrR-P and LtrR* was incubated with 50 μ g of fragmented DNA in 200 μ l of binding buffer for 10 min. The unbound DNA was washed off using the same buffer and the protein was incubated again for 10 min with another lot of 50 μ g fragmented DNA in the column. With four separate incubation step, a total of 200 μ g fragmented DNA was passed through the column bound LtrR-P. Finally, the unbound DNA was washed off using 5 ml of binding buffer and the DNA - LtrR-P complex was eluted with elution buffer (50 mM HEPES, pH 8.0, 50 mM KCl

and 100mM imidazole). The eluate was collected and the resin was washed twice with 200 µl of elution buffer to recover all of the protein-DNA complex.

The DNA bound to the LtrR-P was extracted using phenol-chloroform isolation method. Equal volume of 25:24:1 phenol:chloroform: isoamylalcohol (v/v/v) was mixed with the eluted solution, vortexed and centrifuged at $16,000 \times g$ for 5 min. The upper aqueous phase was carefully transferred to a fresh tube and the extraction process was repeated two more times. Next, the DNA was precipitated by adding 0.1 volume of sodium acetate and 2.5 volume of ice cold ethanol and incubated at $-20 \text{ }^{\circ}\text{C}$ o/n. On the following day, the DNA pellet was collected by centrifugation at $16,000 \times g$ for 10 min, washed with 70% ethanol and dissolved in 25 µl of nuclease free water (Invitrogen). The DNA concentration was measured using Nanodrop spectrophotometer.

4.4 Results

4.4.1 Expression and purification of soluble LtrR

The recombinant plasmid, pJexpress404 containing LtrR gene with N-terminal His₆ tag was cloned into *E. coli* expression strain BL21 (DE3). Cells harbouring the recombinant plasmid were harvested after induction with IPTG for 3 h at $37 \text{ }^{\circ}\text{C}$. After cell lysis the overexpressed LtrR protein was in the soluble fraction. The affinity purification was carried out by binding the protein to nickel charged 5 ml HiTrap affinity column where His₆-LtrR showed different binding affinities; about 20 % of the protein was eluted at 20 mM imidazole and 80 % was eluted at 100 mM imidazole wash (Fig. 4.2A). The latter exhibited greater homogeneity as determined by 12% SDS-PAGE with Coomassie brilliant blue staining (Fig. 4.2B).

The protein was further purified using gel filtration chromatography where LtrR was eluted from the column as a single peak (Fig. 4.2C) which was again confirmed by 12% SDS-PAGE (Fig. 4.2D). The protein was purified with high yield and after purification ~60 mg of recombinant LtrR was recovered from 1L of culture.

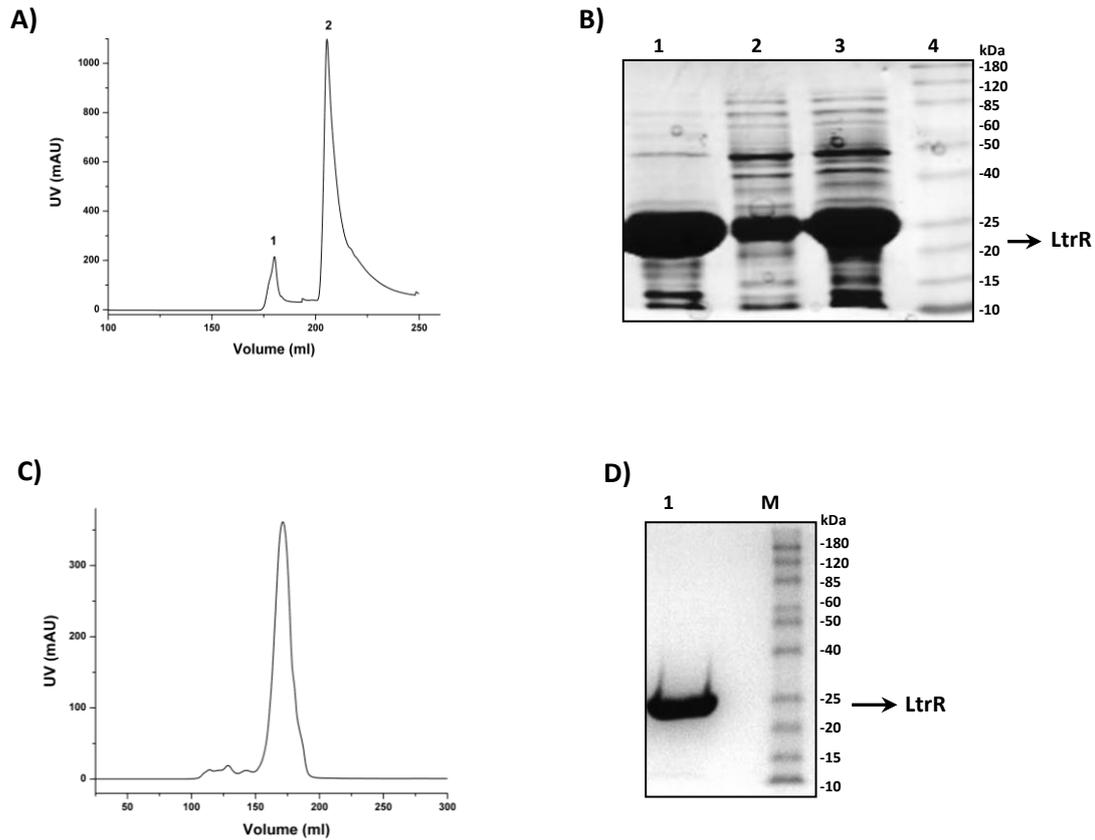


Figure 4.2 Purification of recombinant LtrR

A) LtrR was eluted from nickel affinity column using a step gradient of buffer B containing 1M imidazole. Peak 1 and 2 indicates the protein eluted at 20 mM and 100 mM imidazole concentration respectively; B) SDS-PAGE of different peak fractions. Lanes: 1, eluted His₆ LtrR at peak 2 with 100 mM imidazole; 2, flowthrough fraction of 20 mM imidazole wash (peak 1); 3, soluble fraction of sonicated lysate; 4, protein molecular weight standard. C) Gel filtration chromatography of LtrR; D) SDS-PAGE of eluted GF purified LtrR. Lanes: 1, eluted pure LtrR; M: protein molecular weight standard

The identity of LtrR was confirmed by LC-MS/MS (Fig. 4.3). The snapshot of the result including the list of most abundant proteins is given in Fig. 4.3. The protein hit with highest score was for the response regulator receiver from *M. burtonii* which confirmed the identity of LtrR.

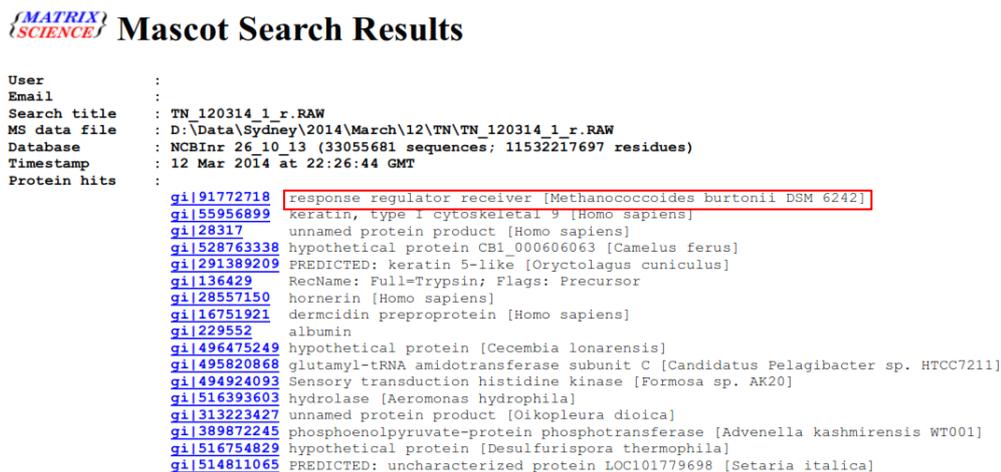


Figure 4.3: LC-MS/MS result for purified LtrR

Protein hit list of LC-MS/MS analyses; red box shows the hit with highest score to be the response regulator from *M. burtonii*.

4.4.2 Evaluation of oligomeric state showed LtrR to be a monomer

The oligomeric state of LtrR was analysed using SEC standard curve. The elution volume (V_e) of LtrR was 175 ml from which partition coefficient (K_{av}) was calculated to be 0.33. The expected molecular weight of His₆ LtrR based on the amino acid sequence was 20.2 kDa and the determined molecular weight from the standard curve was ~ 21 kDa (Fig. 4.4) which correspond closely to the weight of LtrR monomer. This indicated that the recombinant LtrR was purified as a monomer.

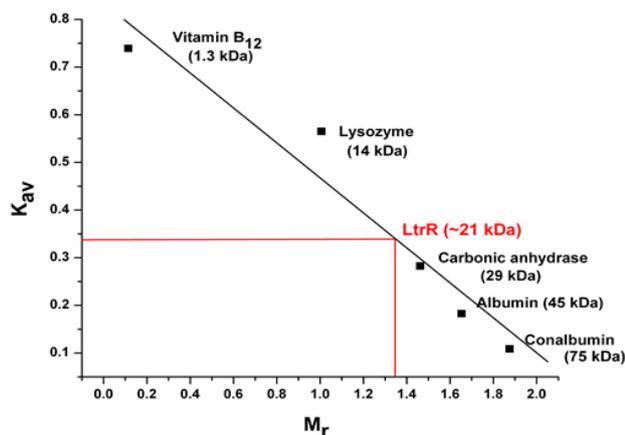


Figure 4.4 Calibration graph for LtrR oligomeric state determination

The square data points and the red line refer to protein standards and recombinant protein LtrR respectively

4.4.3 Protein folding study using intrinsic tyrosine fluorescence analyses

Bioinformatics analyses on the amino acid composition showed that LtrR has only three tyrosines and no tryptophan (described in Chapter 2, Table 2.4). Because of the presence of three tyrosines and no tryptophan, the fluorescence spectrometry was carried out at 276 nm, the maximum absorption wavelength of tyrosine (Kawahara *et al.*, 1962). The fluorescence emission spectra of the recombinant LtrR showed maximum emission at around 308 nm (Fig. 4.5A). The fluorescence intensity decreased upon unfolding of LtrR at 98 °C and after cooling back to 4 °C the refolded protein showed greater intensity than the folded LtrR (Fig. 4.5A) indicating the tertiary structures of the earlier-folded and refolded LtrR are not identical.

The melting temperature (T_m) of the protein was calculated from the fluorescence intensity vs. temperature plot assuming a two-state model (folded state \rightleftharpoons unfolded state), where the midpoint of the transition between the folded and unfolded form

represents the T_m of the protein (Suh and Savizky, 2011). The thermal melting profile of LtrR from 4 to 98 °C showed the T_m to be ~52 °C (Fig. 4.5B). After the first melt the protein was cooled back to 4 °C and the subsequent second melting profile was the same as the first melt (Fig. 4.5B). The identical first and second melts indicated that the thermal unfolding of LtrR is reversible.

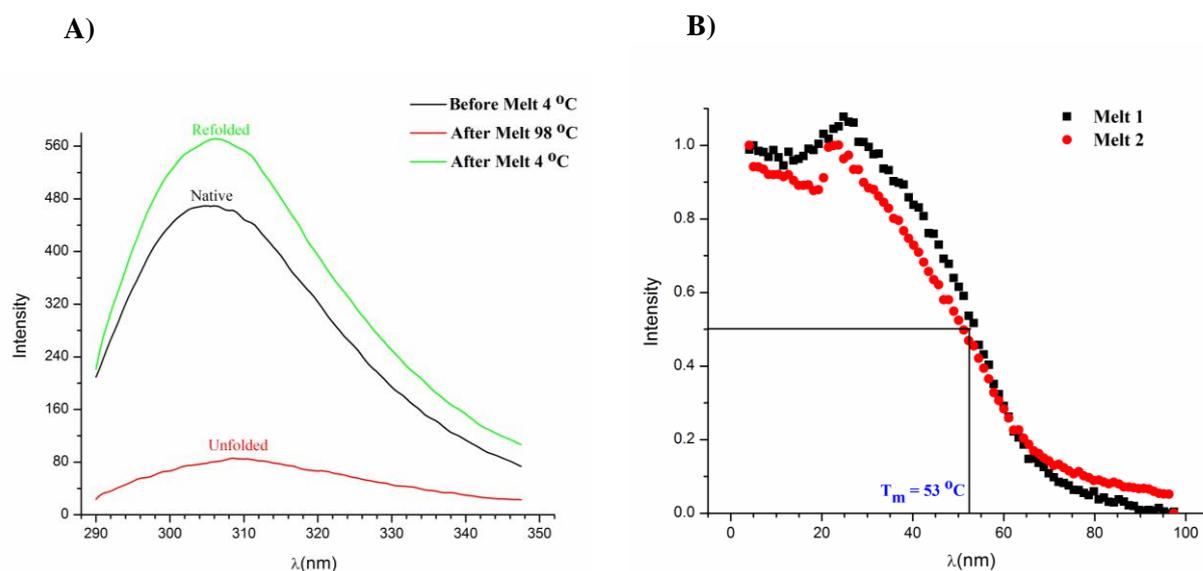


Figure 4.5 Tyrosine intrinsic fluorescence of LtrR

A) Fluorescence intensity of tyrosine regarding native (black; at 4 °C), unfolded (red; at 98 °C) and refolded state (green; at 4 °C); B) Plot of protein's reversible thermal unfolding showing T_m of 53 °C, Melt 1 (black) shows the first melt from 4 to 98 °C with 1 °C min^{-1} scan rate and Melt 2 (red) shows the second melt of the refolded protein with the same temperature range and scan rate

4.4.4 Analyses of LtrR secondary structure using CD spectroscopy

Both far- and near-UV CD analyses were performed in order to determine if the purified LtrR was properly folded with secondary and tertiary structure. The far-UV CD spectra at 4 °C displayed two negative peaks, one at around 208 nm and another at 222 nm, which are the characteristic of α -helix (Kelly *et al.*, 2005). The protein was

unfolded by increasing the temperature to 94 °C and then cooled back to 4 °C and melted again to determine if the unfolding is reversible. The spectra taken before and after thermal melt (Fig. 4.6A) as well as the first and second thermal melt profile of LtrR (Fig. 4.6B) were superimposable. These data indicated that the unfolding of the LtrR secondary structure is reversible. The estimated T_m from the melt profile was ~ 57 °C (Fig. 4.6B).

The near-UV CD spectra, in the region of 260 – 320 nm, arise from the aromatic amino acids and provides fingerprint of the tertiary structure of the protein. Typically, each aromatic amino acid has individual wavelength profile. For example, Tyr shows a peak between 275 and 282 nm with a shoulder at longer wavelengths and Phe shows a sharp band between 255 and 270 nm (Kelly *et al.*, 2005). LtrR has three Tyr and three Phe. The near-UV CD spectra of LtrR showed a relatively sharp peak between 260 and 270 nm which is characteristic of Phe. In addition, the spectra showed three small peaks with shoulder between 272 and 285 nm which denoted three Tyr residues (Fig. 4.6C).

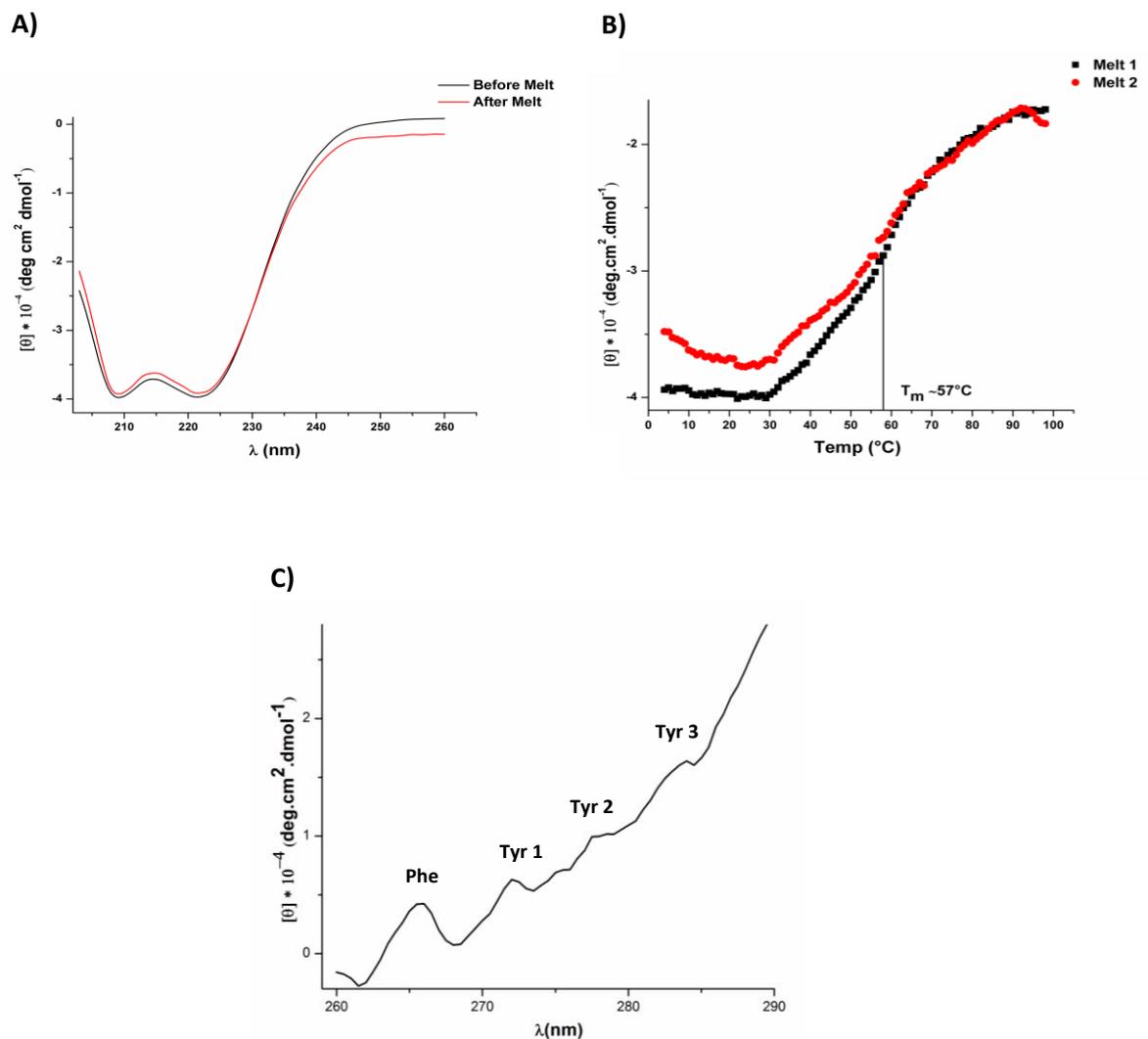


Figure 4.6 CD spectroscopy analyses of LtrR

A) Far-UV CD spectra of native (black) and refolded (red) protein; B) Temperature melt of LtrR showing similar first (black) and second (red) melt profile with T_m 57 °C; C) Near-UV CD spectra of LtrR showing characteristic peaks of phenylalanine (Phe) and tyrosine (Tyr).

4.4.5 Stability analysis of the purified LtrR using DSC

Far-UV CD analyses already showed that the secondary structure of LtrR unfolds reversibly but it does not reflect the unfolding of the tertiary structure. To analyse the overall stability and unfolding of LtrR DSC was performed. At a scan rate 1 °C min⁻¹ the temperature corresponding to the maximum heat capacity (T_m) was 57 °C (Fig. 4.7).

Rescanning of the protein sample after it was heated beyond the transition peak and then cooled to 4 °C resulted in the identical DSC profile (Fig. 4.7A). Rounds of melting resulted in the same melt profile which indicated that the thermal unfolding of LtrR tertiary structure was a reversible process.

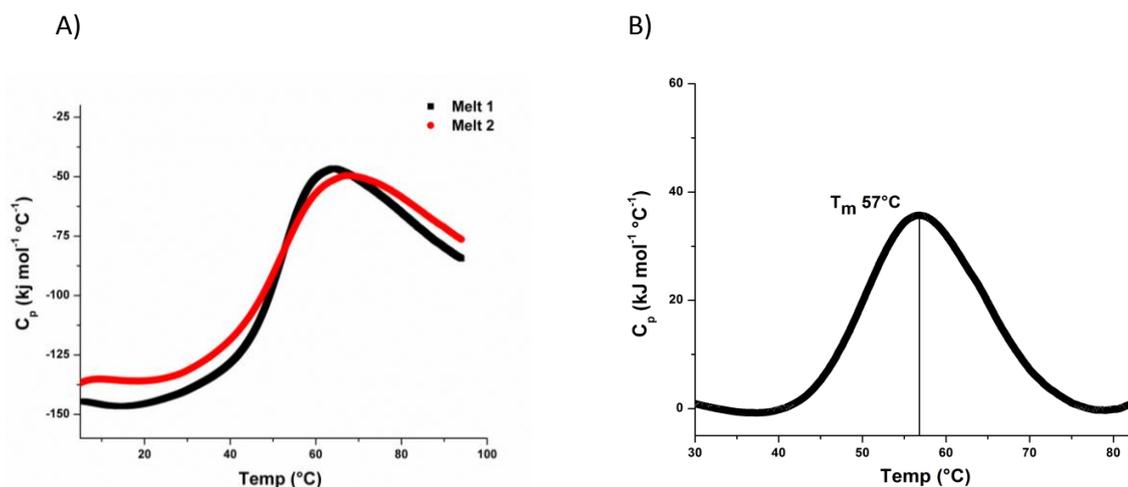


Figure 4.7 DSC profile of LtrR

A) Melt 1 (black) is the first scan of the protein sample in a temperature range of 4 – 90 °C at a scan rate of 1°C min⁻¹ and melt 2 (red) is the rescan of the same sample after cooling it down to 4 °C; B) DSC plot (after baseline correction) showing the T_m to be 57 °C.

To determine the activation energy (E_a) required for LtrR thermal denaturation, Arrhenius plot was generated using the DSC data (Fig. 4.8). The calculated E_a for LtrR thermal denaturation was 192.2 kJ mol⁻¹ K⁻¹.

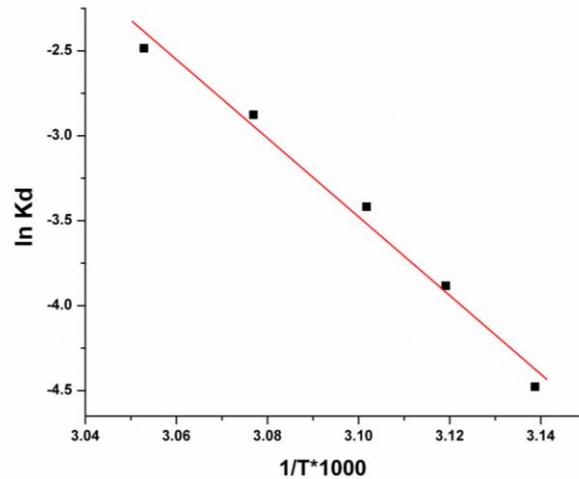


Figure 4.8 Arrhenius plot for LtrR thermal denaturation

Red line is the best linear fit to the data with a slope of -23.12.

4.4.6 DNA binding analyses using *M. burtonii* genomic DNA

The purified His₆ LtrR was phosphorylated by GST-LtrK-P bound to glutathione-agarose beads in a gravity flow column (described in the Experimental procedure section). To perform the DNA binding assay *M. burtonii* genomic DNA was isolated and fragmented to 200–500bp by sonication. DNA fragmentation was necessary to perform high throughput DNA sequencing afterwards that required the DNA to be of 200–500bp length. For the binding assay, the His₆ LtrR-P was bound to the Talon Co⁺ resin affinity column and incubated with fragmented genomic DNA to allow the binding. A negative control experiment was performed in parallel where the mutant LtrR* that doesn't have the DNA binding HTH domain was incubated with the same amount of DNA. The unbound DNA was washed off from both of the column and the DNA-LtrR-P or LtrR* (from negative control) was eluted, the bound DNA was extracted from both sample and the concentration of extracted DNA was measured and

compared. Theoretically, the concentration should be higher for the extracted DNA from DNA-LtrR-P elute than the negative control (DNA extracted from LtrR* elute). But the DNA concentration obtained from both sample was similar which indicated that the DNA binding experiment was not successful.

4.5 Discussion

The response regulator, LtrR, was purified with His₆ tag as a soluble protein (Fig. 4.2) and the protein identity was determined by LC-MS/MS (Fig. 4.3). Using SEC, the oligomeric state of the purified LtrR was found to be a monomer (Fig. 4.4). The monomeric state of LtrR is similar to many bacterial RRs that are expressed as a monomer, *e.g.* OmpR (Barbieri *et al.*, 2013), PhoB (Fiedler and Weiss, 1995) and NarL (Maris *et al.*, 2002) of *E. coli*, DrrB of *T. maritima* (Robinson *et al.*, 2003) but is different from the RRs that are expressed as a dimer, *e.g.* ArcA of *E. coli* (Jeon *et al.*, 2001), DesR of *B. subtilis* (Najle *et al.*, 2009). But majority of the monomeric RRs, including OmpR, PhoB, NarL, DrrB, bind to DNA as a dimer and the dimerization occurs upon phosphorylation by their cognate SKs (described in Chapter 1). As LtrR contains a DNA binding HTH domain and expressed as a monomer, it is likely to form a dimer upon phosphorylation by LtrK. But to determine this, the oligomeric state of phosphorylated LtrR is required to be examined by SEC.

The biophysical properties of the purified LtrR were investigated by tyrosine fluorescence spectrometry (TFS), CD and DSC analyses. The TFS showed maximum emission of LtrR at 308 nm and the fluorescence intensity decreased upon unfolding of the protein (Fig. 4.5A). Usually, the emission maximum of Tyr depends on the H-bonds

it makes with its neighbouring residues with concomitant transfer of energy and reduction in emission intensity (Munishkina and Fink, 2007). During unfolding, the emission intensity decreases due to breaking of H-bonds and as a result, the unfolding of protein without Trp is accompanied by changes in the intensity but not in the wavelength of emission (Munishkina and Fink, 2007). For this reason, upon unfolding of LtrR at 98 °C the fluorescence intensity decreased while the wavelength of emission (308 nm) remained unchanged (Fig. 4.5A). In addition, when the denatured LtrR was refolded by cooling back to 4 °C, the protein showed greater intensity than its earlier-folded form (Fig. 4.5A) which indicates the Tyr residues of the refolded protein have less H-bonds than in the earlier-folded protein. In contrast, the far-UV CD analyses showed identical spectra for earlier-folded and refolded LtrR (Fig. 4.6A) indicating the secondary structure, containing α -helix and β -sheet, unfolds reversibly. From the 3D structure homology model of LtrR (described in Chapter 2) it was observed that two of the three Tyr are located on the surface loop while one is located inside in the middle of a β -strand in the structure (not shown). It is therefore possible that after unfolding when the protein is refolded back the surface loops containing Tyr (one or both) become buried slightly inside that may result the less H-bond of Tyr and higher emission intensity of the refolded LtrR than the earlier-folded protein, while the refolded secondary structure (including the β -strand containing Tyr) remained identical to the earlier-folded LtrR resulting in identical far-UV CD spectra for folded and refolded LtrR.

The melting profile of LtrR obtained from TFS and DSC showed the thermal unfolding of the protein is reversible (Fig. 4.5B, 4.7A). This indicates the

unfolding/refolding transition of LtrR is different from its cognate SK, LtrK which unfolds irreversibly (described in Chapter 3). Typically the large multi-domain protein unfolds irreversibly whereas small single domain proteins tend to unfold reversibly. For example, EF2 (~80.5 kDa) and RNA polymerase E/F (~40 kDa) of *M. burtonii* are large multi-domain proteins that unfold irreversibly (Thomas and Cavicchioli, 2000; De Francisci *et al.*, 2011) while Ctr3 (~7 kDa) is a small single domain protein, thus unfolds reversibly (Taha *et al.*, 2016). LtrK and LtrR are both multi-domain proteins but LtrR (~20 kDa) is smaller than purified cytoplasmic LtrK (~ 31 kDa). Moreover, the oligomeric state of LtrR is monomer whereas LtrK is a dimer. Few monomeric multi-domain proteins had been found to unfold reversibly, *e.g.* aldolase (~35 kDa) from *Staphylococcus aureus* (Rudolph *et al.*, 1992), α -amylase (~50 kDa) from the Antarctic bacterium *A. haloplanctis* (Feller *et al.*, 1999). Perhaps the small monomeric state allows the LtrR to unfold reversibly.

The far-UV CD spectroscopy determined that LtrR had secondary structure that unfolds reversibly (Fig. 4.6AB) and the near-UV CD spectra showed the characteristic peaks of Phe and Tyr residues which indicated the presence of tertiary structure in LtrR (Fig. 4.7C). Interestingly, the T_m of LtrR's secondary structure melt (57 °C) (Fig. 4.6B) is exactly the same as the T_m of tertiary structure melt (DSC melt, Fig. 4.7B) which indicates that the secondary and tertiary structure of LtrK unfolds concurrently.

The DSC profile showed that the unfolding of LtrR initiates at ~40 °C (Fig. 4.7B). A common characteristic of cold-adaptor proteins is weak conformational stability which promotes them to unfold at relatively low temperature in comparison to equivalent proteins from mesophile and thermophiles (Feller and Gerday, 2003). This

thermosensitive property allows the psychrophilic proteins to achieve maximal activity at low temperatures (Feller and Gerday, 2003). The thermolability of psychrophilic proteins is demonstrated by the low activation free energy (E_a) of unfolding (Siddiqui *et al.*, 2002). The E_a of LtrR thermal denaturation was calculated to be $192 \text{ kJ mol}^{-1} \text{ K}^{-1}$ (Fig. 4.8). To determine if the E_a of LtrR denaturation is lower than its thermophilic counterpart, thermal unfolding of similar RRs from thermophilic archaea, e.g. *M. thermophile*, or bacteria, e.g. *T. maritima* is required to be assessed.

The DNA binding ability of phosphorylated LtrR was analysed by DNA affinity purification technique where DNA binding was carried out with His₆ LtrR-P bound to the Co²⁺ resin in a column. Co²⁺ was preferably used for the affinity purification because it exhibits more specific interaction with His tags than other metals, *i.e.* Ni²⁺, Cu²⁺, resulting in less nonspecific interaction. Along with the His₆ tag the negatively charged DNA can also bind with divalent cations including Mn²⁺, Co²⁺, Ni²⁺ (Granot and Kearns, 1982). For this reason, before carrying the DNA binding experiment various amount of LtrR was incubated with the Co²⁺ resin to find out the protein weight to resin volume ratio at which all Co²⁺ ions are saturated with His₆ tag of LtrR to prevent non-specific protein or DNA binding (data not shown) and it was found that 3mg LtrR per 100 μl resin prevents non-specific binding. To compare the experimental data a negative control experiment was performed where DNA-binding incompetent LtrR* was incubated with DNA. It was found that the amount of DNA bound by LtrR-P and LtrR* was almost same which indicates the experiment was unsuccessful. It is noteworthy that the HTH domain is the N-terminal domain in LtrR and the His₆ tag to which LtrR was fused was also at the N-terminus. No cleavage site or linker region was

incorporated between His₆ and HTH sequence during the gene synthesis. It is therefore possible that binding of HTH domain adjacent His₆ to Co²⁺ resin interrupted the interaction between DNA and HTH domain of LtrR. However, LtrR gene with a C-terminal His₆ tag sequence and a linker region between C-terminal REC domain of LtrR and His₆ was synthesized and attempted to overexpress in *E. coli* strain BL21 (DE3) (data not shown). The linker region, was designed to contain ten codon sequences for glycine and serine residues (GSSGGSGSSG) to separate the REC domain of LtrR from the tag so that phosphoryl group incorporation in the REC domain during phosphotransfer reaction was less likely to be hindered. But the newly synthesized gene containing the C-terminal tag and linker did not overexpress after induction with IPTG and as a result the protein could not be purified. It is possible that the protein was expressed in extremely low level that could not be detected on the SDS-polyacrylamide gel or the protein didn't express due to the toxicity of the target protein or because of the presence of mutation in the T7 RNA polymerase promoter region in the plasmid (pJexpress404). In some occasion, the low-level of T7 RNA polymerase expression by the strain can also affect the expression of the target protein (Gräslund *et al.*, 2008). To overcome this problem the gene can be cloned in other expression vector, *i.e.* vector with arabinose promoter system which is more tightly regulated, and also can be expressed in different *E. coli* strain, *i.e.* Rosetta (DE3).

In summary, this chapter reported the successful overexpression and purification protocol of recombinant LtrR. In addition, the biophysical aspects of the purified LtrR investigated by tyrosine fluorescence spectrometry, CD and DSC indicated that the protein is properly folded into its secondary and tertiary structure and the thermal

unfolding is reversible. The DNA binding analyses using *in vitro* DNA affinity purification technique failed to assess the target DNA that LtrR regulates, which is likely because of having His₆ tag adjacent to the DNA binding domain at the N-terminal region. This data strongly recommends LtrR to be expressed with the C-terminal tag to analyse the DNA binding ability.

4.6 References

- Albanesi, D. *et al.* (2004) The membrane fluidity sensor DesK of *Bacillus subtilis* controls the signal decay of its cognate response regulator. *J. Bacteriol.* 186: 2655-2663.
- Barbieri, C.M. *et al.* (2013) Comprehensive analysis of OmpR phosphorylation, dimerization, and DNA binding supports a canonical model for activation. *J. Mol. Biol.* 425: 1612-1626.
- 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.
- Château, A. *et al.* (2013) Identification of CodY targets in *Bacillus anthracis* by genome-wide *in vitro* binding analysis. *J. Bacteriol.* 195: 1204-1213.
- De Francisci, D. *et al.* (2011) The RNA polymerase subunits E/F from the Antarctic archaeon *Methanococoides burtonii* bind to specific species of mRNA. *Environ. Microbiol.* 13: 2039-2055.
- Feller, G., and Gerday, C. (2003) Psychrophilic enzymes: hot topics in cold adaptation. *Nat. Rev. Microbiol.* 1: 200-208.
- Feller, G. *et al.* (1999) Thermodynamic stability of a cold-active α -amylase from the Antarctic bacterium *Alteromonas haloplanctis*. *Biochemistry* 38: 4613-4619.
- Fiedler, U., and Weiss, V. (1995) A common switch in activation of the response regulators NtrC and PhoB: phosphorylation induces dimerization of the receiver modules. *EMBO J.* 14: 3696.
- Granot, J., and Kearns, D.R. (1982) Interactions of DNA with divalent metal ions. III. Extent of metal binding: experiment and theory. *Biopolymers* 21: 219-232.

- Gräslund, S. *et al.* (2008) Protein production and purification. *Nat. Meth.* 5: 135-146.
- Jeon, Y. *et al.* (2001) Multimerization of phosphorylated and non-phosphorylated ArcA is necessary for the response regulator function of the Arc two-component signal transduction system. *J. Biol. Chem.* 276: 40873-40879.
- Kawahara, F.S. *et al.* (1962) The preparation and properties of crystalline Δ^5 -3-ketosteroid isomerase. *J. Biol. Chem.* 237: 1500-1506.
- Kelly, S.M. *et al.* (2005) How to study proteins by circular dichroism. *Biochim. Biophys. Acta-Proteins and Proteomics* 1751: 119-139.
- Liu, W., and Hulett, F.M. (1997) *Bacillus subtilis* PhoP binds to the phoB tandem promoter exclusively within the phosphate starvation-inducible promoter. *J. Bacteriol.* 179: 6302-6310.
- Maris, A.E. *et al.* (2002) Dimerization allows DNA target site recognition by the NarL response regulator. *Nat. Struct. Mol. Biol.* 9: 771-778.
- Munishkina, L.A., and Fink, A.L. (2007) Fluorescence as a method to reveal structures and membrane-interactions of amyloidogenic proteins. *Biochim. Biophys. Acta-Biomembranes* 1768: 1862-1885.
- Najle, S.R. *et al.* (2009) Oligomerization of *Bacillus subtilis* DesR is required for fine tuning regulation of membrane fluidity. *Biochim. Biophys. Acta-General Subjects* 1790: 1238-1243.
- Robinson, V.L. *et al.* (2003) Structural analysis of the domain interface in DrrB, a response regulator of the OmpR/PhoB subfamily. *J. Bacteriol.* 185: 4186-4194.
- Rudolph, R. *et al.* (1992) Reversible unfolding and refolding behavior of a monomeric aldolase from *Staphylococcus aureus*. *Protein Sci.* 1: 654-666.
- Siddiqui, K. *et al.* (2002) Thermodynamic activation properties of elongation factor 2 (EF-2) proteins from psychrotolerant and thermophilic Archaea. *Extremophiles* 6: 143-150.
- Smith, J.L., and Grossman, A.D. (2015) In vitro whole genome dna binding analysis of the bacterial replication initiator and transcription factor DnaA. *PLoS Genet* 11: e1005258.
- Taha *et al.* (2016) Single TRAM domain RNA-binding proteins in Archaea: functional insight from Ctr3 from the Antarctic methanogen *Methanococcoides burtonii*. *Environ. Microbiol.* doi: 10.1111/1462-2920.13229

Thomas, T., and Cavicchioli, R. (2000) Effect of temperature on stability and activity of elongation factor 2 proteins from Antarctic and thermophilic methanogens. *J. Bacteriol.* 182: 1328-1332.

Chapter 5

Analyses of phosphorylation activities of LtrK and LtrR

5.1 Abstract

The phosphorylation activity analyses revealed that the purified cytoplasmic LtrK (lacking N-terminal TMDs) is a bifunctional enzyme possessing both kinase and phosphatase activity. LtrK can autophosphorylate, transfer the phosphoryl group to LtrR, and also dephosphorylate LtrR that was previously phosphorylated. The phosphorylated LtrR (LtrR-P) had a half-life of approximately 2.2 h, which was reduced significantly (60 to 720 fold) by addition of LtrK. LtrK phosphatase activity involves transferring the phosphoryl group of LtrR-P back to itself in a reverse phosphotransfer reaction. Site directed mutagenesis was performed to make amino acid replacements of histidine and aspartate residues in LtrK and LtrR respectively, in order to assess their role in phosphorylation activities. The analyses identified conserved H block H367 of LtrK as the essential residue for both kinase and phosphatase activity. In addition, two other histidine residues, H443 and H448, were identified to play important role in the autophosphorylation reaction. The mutational analysis also determined D98 as the phosphoryl group receiver residue in LtrR during its phosphorylation by LtrK. Overall the data experimentally validate that LtrK and LtrR possess properties typical of SKs and RRs, respectively.

5.2 Introduction

Typically, all SKs display autophosphorylation and phosphotransfer activity while many SKs also exert phosphatase activity on the phosphorylated RRs (RR~P) (Stock *et al.*, 2000; Gao and Stock, 2009). Majority of the purified cytoplasmic domain of SKs (c-SKs) showed autophosphorylation and phosphotransfer (to RRs) activities at room temperature *in vitro* (Jin *et al.*, 1990; Schröder *et al.*, 1994; Liu and Hulett, 1997; Albanesi *et al.*, 2004) with few exceptions; *e.g.* the turgor pressure sensor KpdE becomes inactive when expressed without the N-terminal TMDs (Jung and Altendorf, 1998). Also, many bacterial c-SKs are capable of dephosphorylating RR~P that demonstrates their phosphatase activity (described in Chapter 1) (Shi *et al.*, 1999; Castelli *et al.*, 2000; Zhu *et al.*, 2000).

Another key feature of TCSs is the presence of conserved histidine and aspartate residue in SKs and RRs respectively, which are the phosphoryl group acceptor during the kinase and phosphotransfer reactions (Stock *et al.*, 2000). All bacterial SKs have a conserved histidine in H block located in the cytoplasmic kinase domain whereas RRs have a conserved aspartate residue at the C-terminal end of β 3 strand in their structure, that is the site of phosphorylation (Gao and Stock, 2009). However, there is no experimental study on the active site residues on archaeal SKs and RRs. The only biochemically studied TCS of archaea, FilI/FilR system (Li *et al.*, 2014) had not been assessed to identify the active site residues. In order to assess the role of a putative active site residue in SKs and RRs, typically the target residue is substituted by using site-directed mutagenesis method and the resultant variant is assessed for the activity (Cavicchioli *et al.*, 1995; Chen *et al.*, 2003; Willett and Kirby, 2012). Site-directed mutagenesis technique involves introduction of a mutation in the DNA sequence, most often just a single nucleotide base change (point

mutation), to change the codon sequence that results in substitution of the target amino acid. However, it is important to choose replacement residues that are unlikely to affect protein structure dramatically, otherwise the loss of protein activity could be a result of conformational changes (Bordo and Argos, 1991).

To assess the SKs and RRs activity *in vitro* the phosphorylation state of the protein is required to be determined and several techniques are available to serve that purpose including Pro-Q Diamond staining, isothermal titration calorimetry (ITC) and radioactive kinase assay. Pro-Q Diamond (Invitrogen) is a staining reagent that can selectively stain phosphoproteins in a polyacrylamide gel. This reagent is frequently used in phosphoproteomic studies to qualitatively detect the phosphoproteins on gel (Agrawal and Thelen, 2005; Stasyk *et al.*, 2005). In contrast, ITC is a quantitative technique that determines the heat absorbed or generated during a chemical reaction and can be used to study enzyme kinetics (Siddiqui *et al.*, 2014). This technique had been used to study, the binding of SKs with different ligands (Busch *et al.*, 2007), binding affinity of RRs for their target DNA sequences (Barbieri *et al.*, 2013) and also the interaction between different SK and RR domains (Schmöe *et al.*, 2011).

Nevertheless, the most popular method to detect phosphorylation activity of the proteins (or kinases) is *in vitro* radioactive kinase assay. This technique analyzes the kinase activity by autoradiography or scintillation counting using radioactive ATP ($\gamma^{32}\text{P}$ -ATP) and has been frequently used to detect kinase, phosphotransfer and phosphatase activities of SKs and RRs (Schröder *et al.*, 1994; Liu and Hulett, 1997; Albanesi *et al.*, 2004).

The putative TCS proteins, LtrK (SK) and LtrR (RR) were purified as a dimer and monomer respectively in a soluble, properly folded state (described in Chapter 3 and 4).

Bioinformatics analyses on these proteins identified the presence of a universally conserved H-block histidine (H367) and three non-conserved histidine (H443, H448 and H502) residues in the cytoplasmic domain of LtrK and three putative active site aspartate (D54, D55 and D98) residues in LtrR (described in Chapter 2). In order to determine the enzymatic activity of purified proteins the kinase, phosphotransfer and phosphatase activities were assessed at room temperature. In addition, the functional significance of all histidine residues of cytoplasmic domain in LtrK and three putative active site aspartate residues in LtrR was determined by using site-directed mutagenesis. The analyses provided valuable data on both of the proteins phosphorylation activity, stability of the phosphorylated proteins and important amino acid residues required for the activity.

5.3 Experimental procedures

5.3.1 Autophosphorylation assay using Pro-Q Diamond phosphoprotein gel stain

Purified LtrK (1 mg ml^{-1}) was incubated with 1 mM ATP (NEB) in 50 μl phosphorylation buffer (P buffer; 50 mM HEPES, pH 8.0, 50 mM KCl, 5 mM MgCl_2 , 2 mM TCEP and 20 % glycerol). At 30 s, 10 min and 30 min intervals, 10 μl of samples were withdrawn, mixed with 3 μl of $4 \times$ NuPAGE LDS sample loading buffer (106 mM Tris HCl, 141 mM Tris Base, 2% lithium dodecyl sulfate, 10% glycerol, 0.51 mM EDTA, 0.22 mM SERVA blue G250, 0.175 mM phenol red pH 8.5; Thermo Fisher Scientific) and subjected to 12% SDS-PAGE. A negative control, LtrK without ATP, was also run on the gel along with the phosphorylated sample. PeppermintStick™ phosphoprotein molecular weight standards (Invitrogen), a mixture of phosphorylated and non-phosphorylated

proteins, were run as marker proteins. The gel was stained with Pro-Q Diamond phosphoprotein gel stain using the manufacturer's protocol and phosphorylated protein bands on the gel were visualized using the phosphorimager FLA-5000 (Fuji Photo Film Co, Ltd.). After photography, the gel was further stained with SYPRO Ruby protein gel stain (Invitrogen) for visualization of total proteins and the gel image of total protein was compared to the gel image of phosphorylated proteins in order to determine the phosphorylation levels relative to the total amount of protein. The whole process is shown in the following flow chart (Fig. 5.1).

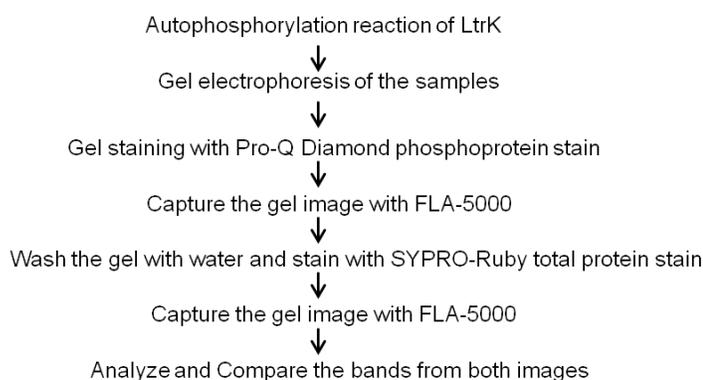


Figure 5.1 Flowchart showing the autophosphorylation assay protocol using Pro-Q Diamond stain

5.3.2 Native-PAGE analyses

LtrK was incubated in P buffer to a final concentration of 1 mg ml^{-1} both in the presence and absence of 1 mM ATP for various periods of time and $10 \text{ }\mu\text{l}$ of sample was run on a prepacked NativePAGE Novex 4–16% Bis-Tris Gel (Invitrogen) according to the manufacturer's protocol. The gel was then washed with water and stained with Coomassie Blue and destained with destaining solution as described in Chapter 3 section 3.2.5.

5.3.3 Isothermal titration calorimetry (ITC)

The phosphotransfer reaction, from phosphorylated LtrK to LtrR was assessed using an isothermal titration calorimeter MicroCal iTC₂₀₀ (GE Healthcare). To prepare the kinase for ITC experiment, LtrK was autophosphorylated by incubating 0.5 mM of protein with 1 mM ATP (NEB) for 30 min in P buffer and the free ATP was removed by using 5 ml HiTrap Desalting column (GE healthcare). Desalting was performed using AKTA purification system by injecting the protein solution in the desalting column, which was pre-equilibrated with P buffer, and LtrK-P was eluted using same buffer with 1 ml min⁻¹ flow rate.

Prior to ITC experiment, the syringe and cell were washed with distilled water twice and the reference cell was filled with milli-Q water. The cell was then filled with 83 nM of phosphorylated LtrK-P and the syringe with the 2 mM of substrate, LtrR. ITC was performed in single injection mode (SIM) at 25 °C with constant stirring at 750 rpm and an initial delay of 60 s. To start the reaction, the substrate solution (18 µl) in the syringe was injected in the cell over 45 s. The change in thermal power, P (µcal s⁻¹) vs. time was monitored until the trace returned to the original baseline. To check the non-enzymatic heat generation (heat of dilution, HOD), the buffer solution (without substrate) was injected into the cell containing enzyme and the HOD was subtracted from the original experimental data. The interval used for calculating enzymatic rates starts from the point of maximum exothermic heat flow (corresponding to the maximum reaction rate at high substrate concentration) up to the point where the trace returns to the baseline (corresponds to the rate at initial substrate concentration) (Siddiqui *et al.*, 2014). The calculation was carried out using Origin software supplied with the MicroCal iTC200 instrument. The integration

of area under the curve gave an apparent reaction enthalpy (ΔH_{app}) from which the rate of enzymatic reaction was calculated as follows (Todd and Gomez, 2001; Ertan *et al.*, 2012):

$$v = \left(\frac{1}{V} \times \Delta H_{app} \right) \frac{dQ}{dt}$$

Where, v = reaction rate; V = cell volume (330 μ l), dQ/dt = rate of heat generated during the reaction. After processing the data with Origin software it was exported to the Enzyme Kinetics Module 1.1 linked to SigmaPlot 8.0 for kinetic analysis where the reaction rate (v) was plotted as a function of instantaneous substrate concentration, $[S]$ and the data were fitted to the monophasic Hills equation (Tracy and Hummel, 2004; Siddiqui *et al.*, 2014):

$$v = \frac{V_{max} [S]^n}{K_m^n + [S]^n}$$

Where, V_{max} is the maximum rate or velocity, K_m is the enzyme-substrate affinity constant and n is the Hill's coefficient. From the kinetic plot V_{max} and K_m was determined and the turnover rate (k_{cat}) was calculated as $V_{max} / [E]$, where $[E]$ is enzyme concentration (83 nM) used in ITC cell.

5.3.4 Autophosphorylation assay using radioactive $[\gamma\text{-}^{32}\text{P}]\text{-ATP}$

Autophosphorylation assay was carried out at room temperature in P buffer containing 0.2 $\mu\text{Ci } \mu\text{l}^{-1}$ of $[\gamma\text{-}^{32}\text{P}]\text{-ATP}$ (specific activity 3000 Ci mmol^{-1} , Perkin Elmer). 1 μg of both GST-LtrK and LtrK (no tag) were incubated with $[\gamma\text{-}^{32}\text{P}]\text{-ATP}$ in 10 μl of P buffer and the reaction was terminated by adding 4 \times NuPAGE LDS sample loading buffer

(Thermo Fisher Scientific) in 2:1 (v/v) ratio followed by heating at 95 °C for 3 min and the samples were run on any KD mini-PROTEAN TGX precast gradient gel (Bio-Rad). The gel was then dried in a gel drier (Model 583, Bio-Rad) at 80 °C for 20 min and exposed overnight to the radioactivity imaging plate (IP) (Fujifilm BAS-IP MS 2025) in a cassette (Fujifilm BAS IP cassette). On the following day, the radioactivity of protein resolved on gel was visualized from the IP by autoradiography using phosphorimager FLA5000.

5.3.5 Phosphotransfer assay using radioactive [γ -³²P]-ATP

For phosphotransfer assay, 20 μ l of glutathione-agarose beads, Glutathione Sepharose 4B (GE healthcare), were washed with P buffer in a 1 ml gravity flow column (Bio-Rad) and incubated with 15 μ g of LtrK protein at room temperature for 30 min. The unbound proteins were washed off and 10 μ Ci of [γ -³²P]-ATP was added to the beads. The autophosphorylation was conducted in the column at room temperature for 30 min. The beads were thoroughly washed with P buffer to get rid of free ATP. Then, LtrR (50 μ g) in 50 μ l of P buffer was added to the column and the flowthrough was collected. A further 50 μ l of buffer was passed through the column and added to the previous flow through. Finally, 10 μ l (5 μ g) of samples were mixed with 3 μ l of 4 \times NuPAGE LDS sample loading buffer containing 100 mM EDTA and subjected to SDS-PAGE. The radioactivity was determined by autoradiography using phosphorimager FLA5000 as for LtrK autophosphorylation.

5.3.6 Dephosphorylation assays of LtrR-P

To determine autophosphatase activity of LtrR, LtrR-P was incubated in P buffer at the room temperature (25 °C) and at various time intervals, 10 µl (~10 µg) of samples were added to 3 µl of sample loading buffer containing 100 mM EDTA and loaded onto any KD mini-PROTEAN TGX precast gradient gel followed by electrophoresis.

To assess the dephosphorylation of LtrR-P in presence of LtrK, LtrR-P was incubated with LtrK in P buffer at 2:1 molar ratio at room temperature. A time course assessment was performed with 10 µl (~10 µg) aliquots added to 3 µl of SDS sample buffer containing 100 mM EDTA, and the samples were subjected to SDS-PAGE followed by autoradiography. ³²P incorporation (activity) was quantified as band intensity using phosphorimager, calculated as a percentage of the total radioactivity of the respective autoradiogram and plotted against time. The LtrR-P dephosphorylation data in absence of LtrK was fitted to the monophasic exponential plot using “ExpDecay” fitting tool from OriginPro software and the dephosphorylation data in presence of LtrK was fitted to biphasic exponential plot using “ExpDecay2” fitting tool of the same software. The value of first-order rate constant (k) was obtained from the fitted exponential plot from which the half-life ($t_{1/2}$) of LtrR-P was determined as $\ln 2/k$.

5.3.7 Site directed mutagenesis of LtrK and LtrR genes and purification of the mutant proteins

Amino acid replacements for histidine residues in LtrK and aspartate residues in LtrR were generated by site-directed mutagenesis using a Phusion Site-Directed Mutagenesis kit (Thermo Scientific) according to the manufacturer's protocol. The cytoplasmic catalytic

region of LtrK contained four histidine residues: H367, H443, H448 and H502, which were substituted with arginine and/or alanine. The receiver domain (REC) of LtrR contained three aspartate residues: D54, D55 and D98 that were replaced with asparagine. The wild type plasmid constructs containing commercially synthesized LtrK and LtrR gene, pReceiver-B03-LtrK and pJexpress404 respectively, served as the template for the mutagenesis reaction. Primers were designed containing the desired mutation sequence and PCR reactions were carried out according to the kit's manual. The primer sequences and details of PCR cycle is given in Appendix 3. The PCR amplified DNA was circularized by incubating 50 ng of PCR product with 0.5 μ l of T4 DNA ligase (provided with the kit) for 5 min. The circularized plasmid was cloned into a cloning strain, *E. coli* DH5 α . All site-directed mutations were confirmed by DNA sequence analysis.

All LtrK and LtrR mutant genes were over expressed and the proteins were purified as described for their wild type in Chapters 3 and 4.

5.4 Results

5.4.1 Autophosphorylation activity of LtrK identified by Pro-Q diamond staining protocol

To analyse the autophosphorylation, LtrK was incubated with ATP at room temperature and the samples were run on a SDS-polyacrylamide gel along with the PeppermintStick™ phosphoprotein molecular weight standards. PeppermintStick™ contained two phosphoprotein molecular weight standards along with four non-phosphorylated standards and as a result it served both as a positive and negative control. After staining the gel with Pro-Q Diamond staining solution, only phosphoproteins of PeppermintStick™ standard were visualized as dark black bands while the non-phosphorylated protein remained undetectable (Fig. 5.2A, Lane 5). A single band was observed for LtrK incubated with ATP along with the two phosphoprotein standards which indicated that cytoplasmic LtrK was phosphorylated by ATP (Fig. 5.2A). But, surprisingly the control sample, LtrK incubated in P buffer without ATP, also showed a band (Fig. 5.2A, Lane 4).

Total protein staining with SYPRO Ruby showed almost same band intensity profile for all samples including control indicating that equal amounts of protein had been loaded in all lanes (Fig. 5.2B). In addition, six bands for phosphoprotein and non-phosphorylated molecular weight standards were observed (Fig. 5.2B, lane 5) that confirmed the efficiency of Pro-Q Diamond solution to stain phosphoproteins only. This indicated that the band observed in the control sample after Pro-Q Diamond staining is indeed the band for phosphorylated LtrK though the sample did not contain any ATP.

The positive band in the control sample could result from LtrK being phosphorylated during overexpression in *E. coli* using the ATP available inside the cell. To investigate that, native-PAGE was run with LtrK after incubating with or without ATP. Native-PAGE analyse the protein in its folded state and separate the proteins based upon their charge and mass. In the gel, the sample of LtrK without ATP showed two closely spaced bands (Fig. 5.2C, lane 4) that were higher up the gel than the positive control, LtrK incubated with ATP (Fig. 5.2C, lanes 1-3). Various factors can influence the electrophoretic mobility of the proteins during native-PAGE including charge-to-mass ratio, physical shape and size of the proteins. Besides, as no denaturants are used in native-PAGE the subunit interactions within a multimeric protein are generally retained. The oligomeric state of the purified cytoplasmic LtrK was determined to be a dimer (described in Chapter 3). Typically in a dimeric SK each monomer retains a phosphoryl group after autophosphorylation (West and Stock, 2001). The two bands of purified LtrK in the gel may indicate the presence of both phosphorylated (LtrK-P) and non-phosphorylated LtrK where LtrK-P might be partially phosphorylated with just one monomer within a dimer containing the phosphoryl group or it might be involved in a protein-protein interaction with non-phosphorylated LtrK that may result two bands higher up the gel than the positive control. The two bands may also reflect the presence of two LtrK isoforms with slightly different molecular weight. Hence the result is inconclusive and does not explain the reason of getting positive band in the negative control during Pro-Q Diamond staining.

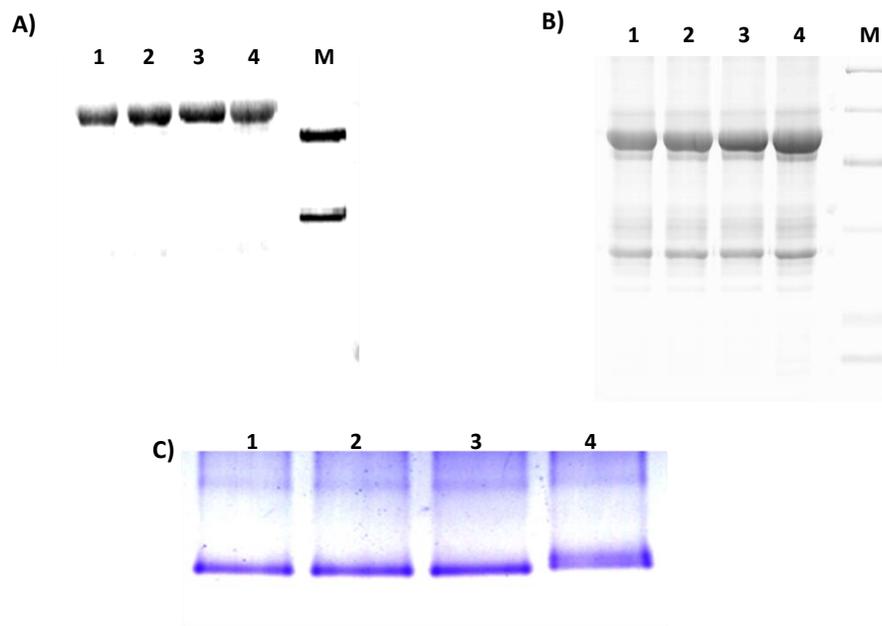


Figure 5.2 Autophosphorylation assay of LtrK using Pro-Q Diamond staining

A) LtrK (1 mg ml^{-1}) was incubated with 1 mM ATP and $8 \mu\text{l}$ samples were withdrawn at various time intervals, added to $2 \mu\text{l}$ of SDS loading buffer, run on SDS-PAGE and was stained with Pro-Q Diamond phosphoprotein stain, lanes 1, 30 s; 2, 10 min; 3, 20 min; 4, LtrK without ATP; M, PeppermintStick molecular weight standard) ; B) Same gel (in panel A) was stained with SYPRO Ruby stain; C) Native-PAGE analyses of LtrK after incubating with 1 mM ATP and samples were collected at various time intervals; lanes 1; 20 min; 2, 10 min; 3, 30 s; lane 4 contains LtrK without ATP

5.4.2 Analyses of phosphotransfer reaction between LtrK and LtrR using ITC

ITC measures the thermal power generated by the enzymatic conversion of substrate into the product and was used to assess the phosphotransfer reaction from LtrK-P to LtrR. The experiment was performed using a single injection mode which delivered the substrate in a single pulse and the reaction between enzyme and substrate generated an exothermic peak (Fig.5.3A). In this experiment, the enzyme was autophosphorylated kinase (LtrK-P) and the substrate was the regulator, LtrR. The ITC data was exported to SigmaPlot 8.0

software for kinetics analyses and the sigmoidal kinetic data fitted biphasic Hill model (Fig. 5.3B) (Siddiqui *et al.*, 2014).

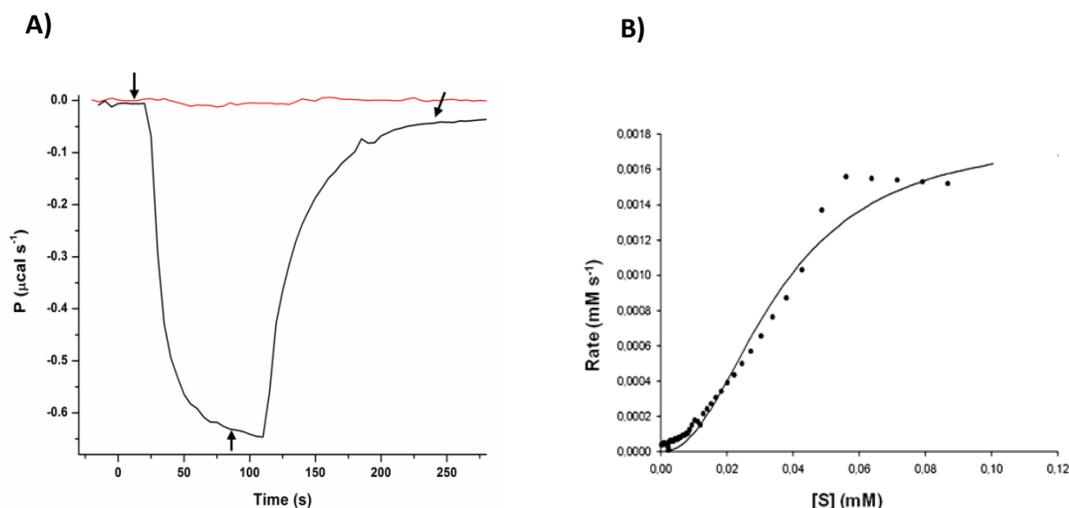


Figure 5.3 ITC profile of LtrK enzymatic reaction kinetics

A) Heat flow (thermal power, P) as a function of time. The red top curve represents heat of dilution (HOD) of buffer in absence of LtrR and lower exothermic curve is for enzymatic reaction after substrate (LtrR) injection. HOD is subtracted from the experimental curve. Points correspond to substrate injection (down arrow), high substrate concentration (up arrow) and low substrate concentration (diagonal arrow); B) Kinetic plot of experimental data and line of best fit to Hill's biphasic equation.

The kinetic values generated from the model are given in Table 5.1. The kinetic plot displayed positive cooperativity which has the Hill coefficient (n) of 2.2. The calculated k_{cat} was $\sim 1300 \text{ min}^{-1}$ which implied the number of substrate molecules that was turned over into product per enzyme molecule per minute; K_m value was 0.035 mM indicating the affinity of the substrate (LtrR) for the enzyme (LtrK) molecule.

Table 5.1 Kinetic values of the phosphotransfer reaction

k_{cat}	$\sim 1300 \text{ min}^{-1}$
K_m	$\sim 0.035 \text{ mM}$
n	2.2

The data suggested that the reaction took place between LtrK-P and LtrR as addition of the substrate, LtrR to the enzyme (LtrK-P) generated an exothermic peak while no peak was observed after adding buffer containing no LtrR (Fig. 5.3A). The reaction producing the exothermic peak is likely to be a phosphotransfer reaction, as phosphorylated SKs tend to transfer the phosphoryl group upon incubating with their cognate RRs generating the product RR-P. However, many TCS SKs exhibit phosphatase activity on its phosphorylated RRs (Keener and Kustu, 1988; Lois *et al.*, 1993; Huynh *et al.*, 2010). In addition, phosphorylated RRs can also show auto-dephosphorylation activity (Hess *et al.*, 1988; Keener and Kustu, 1988; Liu and Hulett, 1997). It is therefore possible that during ITC reaction after producing LtrR-P by transferring the phosphoryl group, LtrK initiated its phosphatase activity on the phosphorylated RR (LtrR-P) resulting a dephosphorylation reaction. So, the observed exothermic peak could be a result of parallel kinase and phosphatase reaction of LtrK on LtrR.

5.4.3 Autophosphorylation and phosphotransfer assays using radioactive [γ - ^{32}P]-ATP

Pro-Q Diamond staining and ITC experiments were indicative of LtrK autophosphorylation and LtrR phosphotrasfer activity. But a positive band in the negative

control during Pro-Q Diamond staining and inability of ITC experiment to distinguish between phosphotrasfer and phosphatase reaction lead to perform radioactive kinase assay to precisely determine the autophosphorylation and phosphotransfer activities of the proteins.

To determine the autophosphorylation reaction, both GST-LtrK and LtrK (without tag) were incubated with $[\gamma\text{-}^{32}\text{P}]\text{-ATP}$ in P buffer. Both proteins were successfully phosphorylated (Fig. 5.4A) and the reaction was rapid, reaching a steady state for GST-LtrK in 30 min at room temperature (Fig. 5.4B). Both proteins were also incubated with ATP in absence of Mg^{2+} which showed no phosphorylation on the autoradiograph (data not shown) indicating that Mg^{2+} is essential for the autophosphorylation reaction.

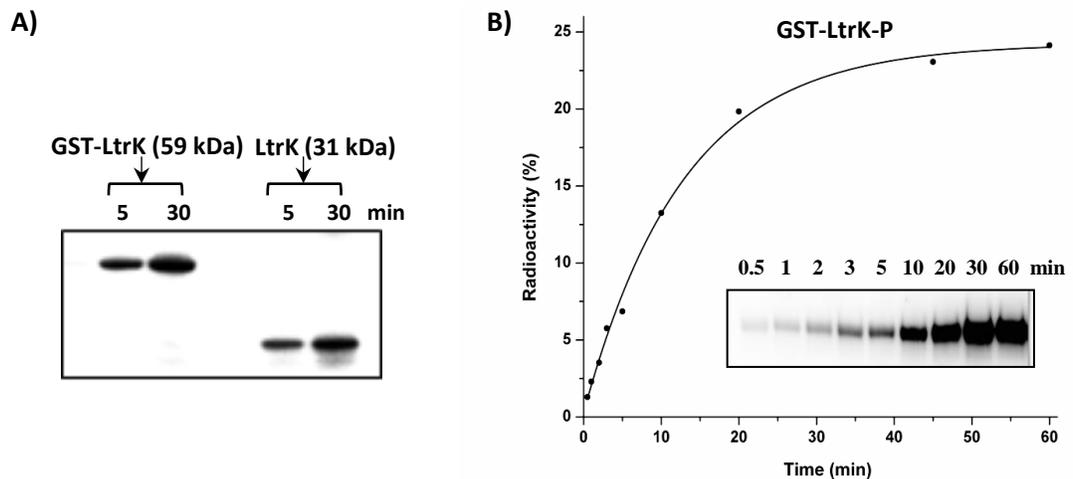


Figure 5.4 Time course of LtrK autophosphorylation

A) LtrK fused to GST (GST-LtrK) and LtrK (1 μg) were incubated with $[\gamma\text{-}^{32}\text{P}]\text{-ATP}$ at room temperature. At indicated times, 10 μl samples were added to 5 μl of SDS-polyacrylamide gel running buffer, heated at 95 $^{\circ}\text{C}$ for 3 min and 3 μl of each sample was analysed by gel-phosphorimaging; B) Time course of autophosphorylation (inset). Plot showing incorporation of GST-LtrK with $[\gamma\text{-}^{32}\text{P}]\text{-ATP}$ over a 60 min incubation at room temperature, the solid line is the exponential fit curve with a rate constant 0.08.

The result also indicated that GST tag did not interfere with the kinase activity of LtrK which allowed rest of the experiments to be performed using GST-LtrK and also enabled the phosphotransfer experiments to be performed with GST-LtrK immobilized on a glutathione agarose column (described below).

To investigate whether the phosphoryl group can be transferred from phosphorylated LtrK to LtrR, the LtrK was autophosphorylated by incubating with [γ - 32 P]-ATP for 30 min and LtrR was added to the reaction to allow phosphotransfer. The samples were then collected at various time intervals and subjected to SDS-PAGE followed by autoradiography. However, the autoradiograph did not show any phosphorylated LtrR (data not shown). Previous studies showed that, many SKs exert phosphatase activity on its phosphorylated RRs and the presence of ATP can dramatically stimulate the dephosphorylation reaction (Igo *et al.*, 1989; Liu and Hulett, 1997). Based on this study, GST-LtrK was bound to the glutathione-agarose column and phosphorylated by incubating with [γ - 32 P]-ATP and the free ATP was washed out thoroughly with P buffer. Subsequently, LtrR was passed through the column without any incubation; the flowthrough was collected and subjected to SDS-PAGE followed by autoradiography. The autoradiograph clearly showed radioactive band for LtrR-P (Fig. 5.5, lane 2) indicating phosphoryl group was successfully transferred from GST-LtrK-P to LtrR (Fig. 5.5). It also indicated that, the presence of ATP restricts the phosphotransfer reaction.

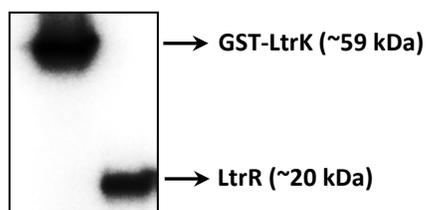


Figure 5.5 Phosphotransfer from GST-LtrK to LtrR

GST-LtrK (15 μ g) bound to glutathione agarose beads in a gravity-flow column was phosphorylated with [γ - 32 P]-ATP for 30 min at room temperature, free [γ - 32 P]-ATP washed off, LtrR (60 μ g) passed through the column and LtrR-P collected in the flowthrough. Aliquots of LtrR-P (10 μ l) were added to 3 μ l of 100 mM EDTA containing SDS-polyacrylamide gel running buffer, and 3 μ l of each mixture were analyzed by gel-phosphorimaging.

5.4.4 Dephosphorylation of LtrR in the absence and presence of LtrK

To assay the phosphatase activity of LtrK and autophosphatase activity of LtrR, LtrR was phosphorylated by GST-LtrK-P in the column as described above and the phosphorylated LtrR-P was incubated in P buffer with and without GST-LtrK. A slow dephosphorylation due to autophosphatase activity of LtrR-P was observed with a half-life of approximately 2.2 h (Fig. 5.6A). In contrast, in the presence of GST-LtrK, dephosphorylation was rapid and the reaction was biphasic with a half-life for LtrR-P of 12 s in the first phase and 2.4 min in the second phase (Fig. 5.6B).

The autoradiograph for LtrR-P dephosphorylation in presence of GST-LtrK also showed radioactive band for GST-LtrK (Fig. 5.6B) indicating GST-LtrK being phosphorylated though the sample did not contain any radioactive [γ - 32 P]-ATP. Moreover, the loss in 32 P signal for LtrR-P was matched by an increasing signal for GST-LtrK-P (Fig. 5.6B) which indicates a reverse-phosphotransfer reaction from LtrR-P to GST-LtrK.

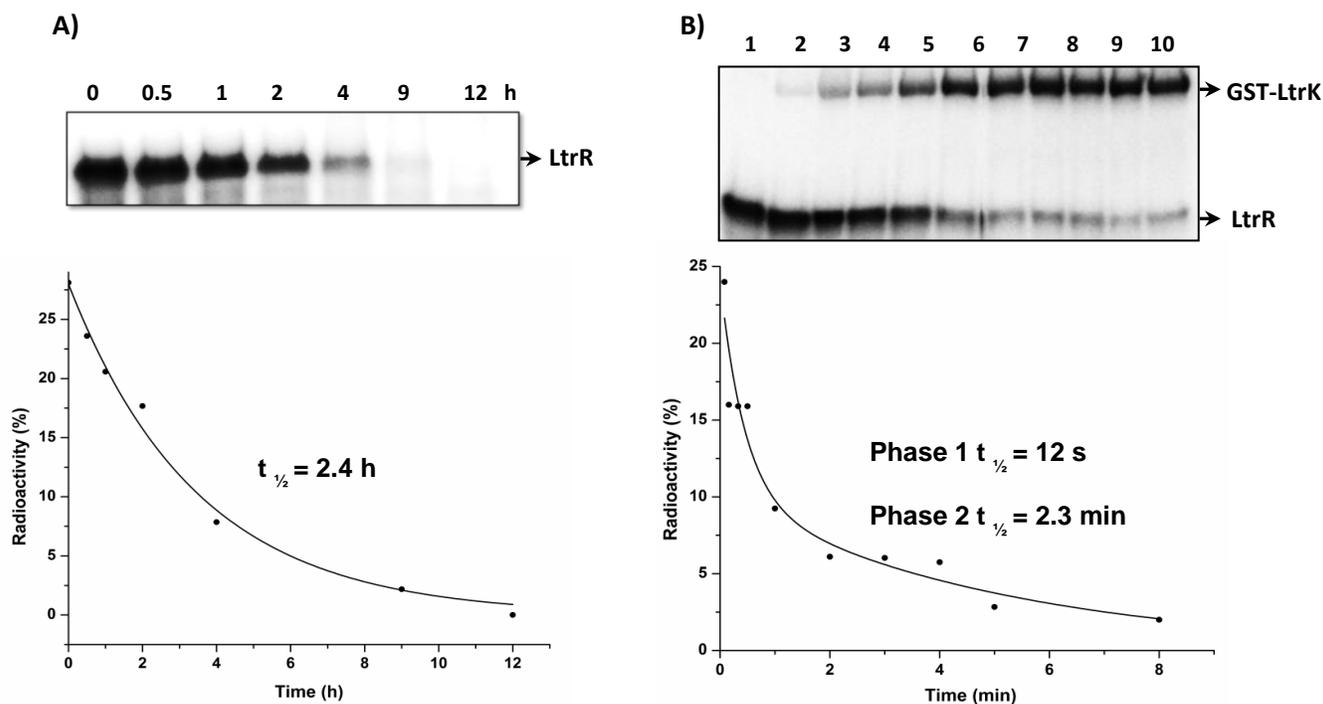


Figure 5.6 Dephosphorylation assays of LtrR-P.

A) LtrR-P generated from phosphotransfer was incubated at room temperature and retention of γ - 32 P assessed over time. The graph shows the exponential fit to the data with rate constant (k) of 0.29 from which the $t_{1/2}$ is calculated as $\ln 2/k$; B) LtrK phosphatase activity. As for panel A except LtrR-P was incubated with LtrK. The solid line represents the two exponential fit curve with k_1 3.45 in phase 1 and k_2 0.3 in phase 2 from which the $t_{1/2}$ calculated as $\ln 2/k$.

5.4.5 Effect of histidine residue replacement on LtrK phosphorylation activities

To determine the active site histidine residue of LtrK, site directed mutagenesis was performed on all four histidine residues (H367, H443, H448 and H502) of LtrK cytoplasmic domain. These residues of LtrK were first replaced with arginine as it has similar charge properties to histidine with a different pKa value and thus were expected to

have less effect on the protein conformation. With single point mutation, mutant variants GST-LtrK H367R, H443R, H448R, H502R and with dual point mutation GST-LtrK H443R/H448R were obtained. All the mutant variants were incubated with [γ - 32 P]-ATP for 20 min and the autoradiograph showed that only GST-LtrR H367R mutant was unable to phosphorylate (Fig. 5.7A, lane 2) while other mutant variants showed phosphorylated bands but with different intensity (Fig. 5.7A). Mutation at H502 had the least impact on autophosphorylation (Fig. 5.7A, lane 6), although the extent of phosphorylation was less than the wild-type. In contrast, mutations at H443 and H448 greatly reduced the autophosphorylation (Fig. 5.7A, lane 4, 5) with double mutant H443R/H448R (Fig. 5.7A, lane 3) exhibiting even lower activity than the single mutants.

However, in spite of having similar charge properties arginine may still disrupt the LtrK structural conformation as it has large side-chain that is able to form multiple H-bonds and ionic interactions. Therefore, to verify the observed effect of histidine to arginine mutation on LtrK phosphorylation, histidine residues were also mutated to alanine. Alanine has a small, non-polar side-chain, and is less likely to change protein conformation. The autophosphorylation activity analyses on the mutant protein showed that GST-LtrK H367A had no activity (Fig. 5.7B, lane 2) while H443A/H443A had very little activity (Fig. 5.7B, lane 3) which was the same as for H367R and H443R/H448R mutants.

GST-LtrK mutants were also evaluated for the phosphatase activity and incubated with LtrR-P for 30 min at room temperature. It appeared that only LtrK with H367 mutation (H367R) was unable to dephosphorylate LtrR-P (Fig. 5.7C, lane 2) while other mutant proteins, H443R, H502R and H443/448R exhibited phosphatase activity similar to the wild type (Fig. 5.7C, lanes 3-5). Thus, only the conserved H367 appeared to be required

for the phosphatase activity. In addition, all mutant LtrK except H367R showed reverse phosphotransfer reaction similar to wild type protein indicating the importance of H367 for reverse phosphotransfer reaction as well.

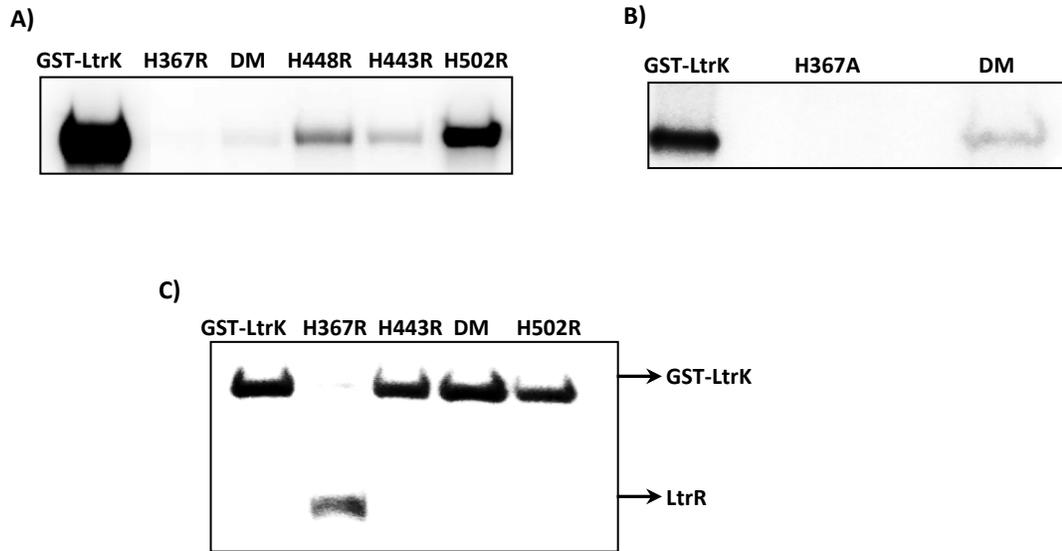


Figure 5.7 Effect of mutations on LtrK phosphorylation activities

To assess the phosphorylation state of proteins, samples were electrophoresed on a SDS-polyacrylamide gel, and autoradiography performed by phosphorimaging (panels A-C).

A) Autophosphorylation of GST-LtrK mutant proteins. Proteins (7 μ g) were incubated with $[\gamma\text{-}^{32}\text{P}]\text{-ATP}$ at room temperature for 30 min. Histidine residues were replaced with arginine, including the double mutant (DM) H443R/H448R; B) As for panel A except histidine residues were replaced with alanine including the double mutant (DM) H443A/H448A; C) Phosphatase activity of GST-LtrK mutant proteins. LtrR-P was incubated with wild-type (lane 1) and mutant proteins H367R (lane 2), H443R (lane 3), DM: H443R/H448R (lane 4), H502R (lane 5) for 30 min at room temperature

5.4.6 Effect of aspartate residue replacement on LtrR phosphorylation

In LtrR three aspartate residues, D54, D55 and D98, were targeted for mutation. All the mutants were passed through the glutathione agarose column bound with GST-LtrK-P

and analyzed via autoradiography. The autoradiograph showed that only the mutant variant LtrR D98N was unable to become phosphorylated, whereas the other mutants, D54N and D55N were phosphorylated by the kinase similar to the wild-type (Fig. 5.8). The data indicated D98 to be the likely phosphoryl group accepting aspartate residue of LtrR.

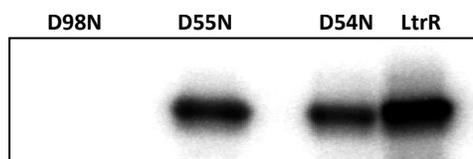


Figure 5.8 Effect of mutations on LtrR phosphorylation

Aspartate (D) residues were replaced with asparagine (N). LtrR (80 μ g) wild-type and mutant variants were phosphorylated by GST-LtrK-P (20 μ g) immobilized on a column. 10 μ l of samples were collected and mixed with 3 μ l sample loading buffer containing 100 mM EDTA and were subjected to SDS-PAGE followed by autoradiography.

5.5 Discussion

Initially Pro-Q Diamond staining protocol was used to identify the autophosphorylation activity of LtrK. Though the stained gel showed positive band of phosphorylated LtrK after incubating LtrK with ATP, positive band for the control sample, LtrK without ATP, (Fig. 5.2A) makes the result inconclusive. The native-PAGE analysis suggested that LtrK might become partially phosphorylated during overexpression in *E. coli* and this makes the Pro-Q Diamond staining approach inapplicable to identify the *in vitro* phosphorylation activity of LtrK. To assess the phosphotransfer reaction ITC was performed and the results provided evidence of kinetic reaction between LtrK and LtrR with high catalysis rate, k_{cat} (Table 5.1). But many SKs are found to be bifunctional enzyme exerting both kinase and phosphatase activity which indicates that the result seen by ITC

could be a consequence of parallel kinase and phosphatase reaction where LtrK transfers its phosphoryl group to LtrR using the kinase activity and also dephosphorylate the LtrR-P using the phosphatase activity. As ITC can only measure the thermal power generated by an enzymatic reaction and cannot differentiate between different types of reactions, it is not a suitable approach to specifically identify the phosphotransfer reaction from LtrK to LtrR.

In comparison to Pro-Q Diamond staining and ITC technique the use of radioactive [γ - 32 P]-ATP provided a straightforward method to determine the activity of LtrK and LtrR. The *in vitro* kinase assays using radioactive [γ - 32 P]-ATP identified that the purified cytoplasmic LtrK retained both autophosphorylation (Fig. 5.4) and phosphatase (Fig. 5.6B) activities as has been observed for many bacterial c-SKs (Makino *et al.*, 1989; Liu and Hulett, 1997; Shi and Hulett, 1999). The purified LtrR catalysed the phosphoryl group transfer from GST-LtrK-P to itself only in the absence of ATP (Fig. 5.5). In many bacterial TCS pathway it had been found that the phosphatase activity of SKs becomes significantly stimulated in the presence of ATP (Igo *et al.*, 1989; Kamberov *et al.*, 1994; Liu and Hulett, 1997). Typically in SKs the ATP binds to the ATP lid which is a loop between G and F block and the lid interacts with N and H block residues during the autophosphorylation reaction (described in Chapter 2, Fig. 2.7). The crystal structure analyses of HK853-RR468, a SK-RR complex, of *T. maritima* showed that the RR, RR468 interacts with the ADP β N (an ATP analog) bound ATP lid of the SK, HK853 and this interaction is suggested to play a key role for the specific recognition of RR-P during the dephosphorylation of RRs by SKs (Casino *et al.*, 2009). Thus addition of ATP might facilitate the interaction between ATP lid and RR and increase the affinity of SKs for RR-P (Casino *et al.*, 2009). It is therefore possible that the phosphotransfer reaction took place after incubating LtrR with LtrK-P but

due to the presence of ATP, LtrK immediately dephosphorylated the phosphotransfer product LtrR-P that resulted in no radioactive LtrR-P band on the gel.

Many TCS RRs have been reported to have widely differing phosphatase activity, with dephosphorylation rates tuned to the time-scale requirements of the specific system (Stock *et al.*, 2000). Dephosphorylation study on LtrR-P showed that the regulator has weak autophosphatase activity with a $t_{1/2}$ of ~ 2.2 h (Fig. 5.6A) which is similar to OmpR in *E. coli* or PhoP in *B. subtilis* that has a $t_{1/2}$ of 1.5 h and 2.5 h respectively (Igo *et al.*, 1989; Liu and Hulett, 1997), but different from RRs that have high autophosphatase activity resulting in a shorter half-life ($t_{1/2}$); *e.g.* CheY of *E. coli* has a $t_{1/2}$ of few seconds, NtrC of *Salmonella typhimurium* has a $t_{1/2}$ 3 to 5 min (Hess *et al.*, 1988; Keener and Kustu, 1988).

Dephosphorylation of RRs is greatly influenced by the phosphatase activities of SKs. Analyses of LtrR-P dephosphorylation in the presence of LtrK showed significantly higher phosphatase activity of LtrK than LtrR autophosphatase activity, reducing the half-life of LtrR-P 60 – 720-fold in a biphasic response (Fig. 5.6B). Concurrently, the phosphoryl-group of LtrR-P was transferred back to LtrK (Fig. 5.6B) in a reverse phosphotransfer reaction. Thus the biphasic response of LtrR dephosphorylation in presence of GST-LtrK reflects the complexity of the reaction as it could be a mixture of slow autophosphatase activity of LtrR, very high phosphatase activity of LtrK and the reverse phosphotransfer activity of LtrK (Fig. 5.6B).

During the phosphatase activity of LtrK, the loss in ^{32}P signal associated with LtrR-P was equivalent to the increase of concomitant ^{32}P signal of LtrK-P (Fig. 5.6B) demonstrating an efficient reverse phosphotransfer reaction. These data indicates that the phosphatase activity of LtrK is similar to the bacterial SKs that utilize reverse

phosphotransfer as a potential mechanism of exerting phosphatase activity on RRs, *e.g.* CheA (Stewart, 1997) and ArcB (Georgellis *et al.*, 1998) of *E. coli*, PhoR of *B. subtilis* (Shi *et al.*, 1999) but different from SKs that require presence of auxiliary protein, *e.g.* NtrB of *E. coli* requires auxiliary protein PII, for the phosphatase activity.

Mutation analyses on the histidine residues of LtrK determined the conserved H block H367 to be the essential residue for both kinase and phosphatase activities (Fig. 5.7). Moreover, it was observed that during phosphatase activity all histidine mutants except H367 reversed back the phosphoryl group from LtrR-P to itself (Fig. 5.7C). This implies that the same histidine residue (H367) utilized in autophosphorylation reaction, to be the likely phospho-acceptor residue during the reverse phosphotransfer reaction. This characteristic of LtrK is similar to PhoR of *B. subtilis* (Shi *et al.*, 1999) and PhoQ of *S. typhimurium* (Castelli *et al.*, 2000) in which the H block histidine residue is essential for both kinase and phosphatase activities and it is also the phosphoryl group accepting residue in both autokinase and reverse phosphotransfer reactions, but contrasting with DesK of *B. subtilis* in which the H block histidine is only essential for kinase activity but not for phosphatase activity though this residue is the phosphoryl group acceptor in both autokinase and reverse phosphotransfer reactions (Albanesi *et al.*, 2004).

The significance of non-conserved histidine residues (H443, H448 and H502) of LtrK was also assessed using site directed mutagenesis method. Mutation at H443 and H448 showed significantly reduced kinase activity while their phosphatase activity remained unaffected (Fig. 5.8), indicating the important role played by these residues during autophosphorylation. Though the H block histidine residue is the site of phosphorylation in a typical SK (Stock *et al.*, 2000), histidine residues outside of the H block have been shown

to function in phosphorylation reactions in some SKs, *e.g.* NarX and NarQ of *E. coli* (Cavicchioli et al., 1995). The essential histidine residues in NarX and NarQ are in the N block and conserved in a subfamily of *E. coli* and *B. subtilis* SK sequences (Cavicchioli et al., 1995). But the non-conserved histidine residues, H443 and H448, of LtrK reside between conserved H and N block (Fig. 5.9A) and are not conserved in the sequences from the most closely related methanogens (described in Chapter 2, section 2.3.4). The protein homology model of the cytoplasmic domain of LtrK (described in Chapter 2, section 2.3.7.1) showed that both H443 and H448 are in the $\alpha 3$ helix that is positioned next to the HisKA (dimerization and phosphotransfer) domain and connected to the HATPase (catalytic and ATP-binding) domain via a β -strand to the $\alpha 4$ helix containing the N block (Fig. 5.9).

During the autophosphorylation reaction, the N (in $\alpha 4$) and G2 (in $\alpha 6$) block residues within the HATPase domain (blue) interact with ATP to position the γ -phosphate near the conserved histidine (H367 in LtrK) in the HisKA domain (yellow) (Casino *et al.*, 2009; Casino *et al.*, 2014). Mutation studies targeting the $\alpha 3$ helix have not previously been reported in the literature. The data for H443 and H448 suggest that the $\alpha 3$ helix in LtrK may function by facilitating interactions with ATP to assist in catalysis leading to autophosphorylation. As these histidine residues are not generally conserved in SKs, their function may be specific to interactions between LtrK and ATP. To confirm the role of these additional histidine residues crystal structures are required for both LtrK and LtrK-ATP complex to explore the residues involved in interaction with ATP and also the conformational changes that occurs due to phosphorylation.

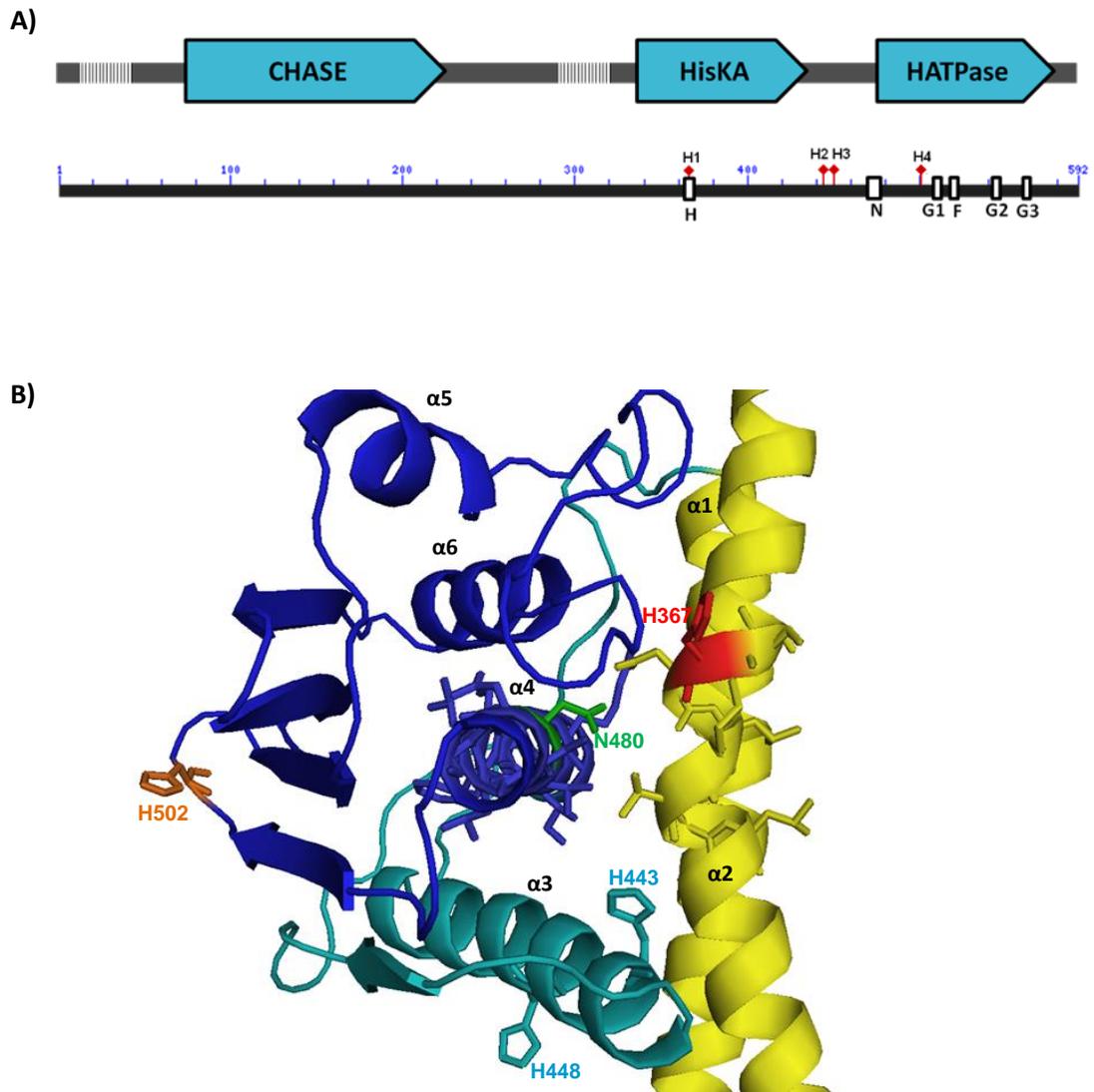


Figure 5.9 Location of the non-conserved histidine residues in LtrK cytoplasmic domain

A) Schematic of LtrK protein domains and sequence motifs drawn to scale. Protein domains identified using Pfam (cyan arrow boxes); predicted TMDs (hatched regions); H, N, G1, F, G2 and G3 blocks (white boxes) diagnostic of TCS SKs (Parkinson and Kofoed, 1992; Dutta and Inouye, 2000); specific histidine residues H367 (H1), H443 (H2), H448 (H3), H502 (H4) of LtrK. B) Close up view of the cytoplasmic domain structure of LtrK. The HisKA domain (yellow) includes the $\alpha 1$ and $\alpha 2$ helices. The $\alpha 1$ helix contains the H block residues (shown in stick representation) including universally conserved H367 (red). The $\alpha 4$ helix of HATPase domain (blue) contains the N block residues (shown in stick representation) including mostly conserved N480 (green). The $\alpha 3$ helix (cyan) between the HisKA and HATPase domains contains the additional histidine residues H443 and H448 (cyan). The third additional H502 (orange) resides in a loop far from N and H box.

Mutation studies on LtrK aspartate residues indicated only D94 to be essential for the phosphorylation of the regulator while mutation at D54 and D55 had no effect (Fig. 5.8). Typically, the active site of RRs is an acid pocket incorporating three aspartate residues; one is phospho-accepting and the rest are involved in Mg^{2+} co-ordination (Bourret, 2010). Structural analyses on RRs showed that the aspartate residues involved in Mg^{2+} binding resides at the C-terminal end of $\beta 1$ stand (Bourret, 2010). The homology tertiary structure model of LtrK (in Chapter 2, Fig. 2.8) showed D54 and D55 are at the $\beta 1$ end indicating D54 and D55 to be the potential Mg^{2+} binding residues. Though the mutation of D54 and D55 with asparagine showed no effect on the LtrR phosphotransfer activity, it does not rule out their possibility of Mg^{2+} co-ordination as asparagine can also coordinate with Mg^{2+} similar to aspartate (Glusker *et al.*, 1999). So, further mutation studies (*i.e.* aspartate to alanine) are required to specifically determine the function of D54 or D55.

Overall the chapter revealed the typical SK and RR properties of LtrK and LtrR respectively (*i.e.* autokinase, phosphotransfer, phosphatase activities) along with few distinct features (*i.e.* participation of non-conserved histidine residues in LtrK kinase activity). The data experimentally established that the LtrK and LtrR have typical properties of a SK and RR respectively and therefore can form a TCS in *M. burtonii* which was fundamental to continue further experiments on temperature dependent activities described in the following chapters.

5.6 References

Agrawal, G. K., & Thelen, J. J. (2005) Development of a simplified, economical polyacrylamide gel staining protocol for phosphoproteins. *Proteomics* 5: 4684-4688.

- Albanesi, D. *et al.* (2004) The membrane fluidity sensor DesK of *Bacillus subtilis* controls the signal decay of its cognate response regulator. *J. Bacteriol.* 186: 2655-2663.
- Barbieri, C.M. *et al.* (2013) Comprehensive analysis of OmpR phosphorylation, dimerization, and DNA binding supports a canonical model for activation. *J. Mol. Biol.* 425: 1612-1626.
- Bordo, D., and Argos, P. (1991) Suggestions for “safe” residue substitutions in site-directed mutagenesis. *J. Mol. Biol.* 217: 721-729.
- Bourret, R.B. (2010) Receiver domain structure and function in response regulator proteins. *Curr. Opin. Microbiol.* 13: 142-149.
- Busch, A. *et al.* (2007) Bacterial sensor kinase TodS interacts with agonistic and antagonistic signals. *Proc. Natl. Acad. Sci. USA* 104: 13774-13779.
- Casino, P. *et al.* (2009) Structural insight into partner specificity and phosphoryl transfer in two-component signal transduction. *Cell* 139: 325-336.
- Casino, P. *et al.* (2014) Visualizing autophosphorylation in histidine kinases. *Nat Commun* 5: 3258.
- Castelli, M.a.E. *et al.* (2000) The phosphatase activity is the target for Mg²⁺ regulation of the sensor protein PhoQ in *Salmonella*. *J. Biol. Chem.* 275: 22948-22954.
- Cavicchioli, R. *et al.* (1995) The NarX and NarQ sensor-transmitter proteins of *Escherichia coli* each require two conserved histidines for nitrate-dependent signal transduction to NarL. *J. Bacteriol.* 177: 2416-2424.
- Chen, Y. *et al.* (2003) Residue R113 is essential for PhoP dimerization and function: a residue buried in the asymmetric PhoP dimer interface determined in the PhoPN three-dimensional crystal structure. *J. Bacteriol.* 185: 262-273.
- Dutta, R., and Inouye, M. (2000) GHKL, an emergent ATPase/kinase superfamily. *Trends Biochem. Sci.* 25: 24-28.
- Ertan, H. *et al.* (2012) Kinetic and thermodynamic characterization of the functional properties of a hybrid versatile peroxidase using isothermal titration calorimetry: Insight into manganese peroxidase activation and lignin peroxidase inhibition. *Biochimie* 94: 1221-1231.
- Gao, R., and Stock, A.M. (2009) Biological insights from structures of two-component proteins. *Annu. Rev. Microbiol.* 63: 133.
- Georgellis, D. *et al.* (1998) Signal decay through a reverse phosphorelay in the Arc two-component signal transduction system. *J. Biol. Chem.* 273: 32864-32869.

- Glusker, J.P. *et al.* (1999) Metal ions in biological system. *The Rigaku Journal* 16: 8-16.
- Hess, J.F. *et al.* (1988) Phosphorylation of three proteins in the signaling pathway of bacterial chemotaxis. *Cell* 53: 79-87.
- Huynh, T.N. *et al.* (2010) Conserved mechanism for sensor phosphatase control of two-component signaling revealed in the nitrate sensor NarX. *Proc. Natl. Acad. Sci. U. S. A.* 107: 21140-21145.
- Igo, M.M. *et al.* (1989) Phosphorylation and dephosphorylation of a bacterial transcriptional activator by a transmembrane receptor. *Genes Dev.* 3: 1725-1734.
- Jin, S. *et al.* (1990) Phosphorylation of the VirG protein of *Agrobacterium tumefaciens* by the autophosphorylated VirA protein: essential role in biological activity of VirG. *J. Bacteriol.* 172: 4945-4950.
- Jung, K., and Altendorf, K. (1998) Truncation of amino acids 12–128 causes deregulation of the phosphatase activity of the sensor kinase KdpD of *Escherichia coli*. *J. Biol. Chem.* 273: 17406-17410.
- Kamberov, E.S. *et al.* (1994) Effect of mutations in *Escherichia coli* glnL (ntrB), encoding nitrogen regulator II (NRII or NtrB), on the phosphatase activity involved in bacterial nitrogen regulation. *J. Biol. Chem.* 269: 28294-28299.
- Keener, J., and Kustu, S. (1988) Protein kinase and phosphoprotein phosphatase activities of nitrogen regulatory proteins NTRB and NTRC of enteric bacteria: roles of the conserved amino-terminal domain of NTRC. *Proc. Natl. Acad. Sci. USA.* 85: 4976-4980.
- Li, J. *et al.* (2014) Characterization of an archaeal two-component system that regulates methanogenesis in *Methanosaeta harundinacea*. *PLoS one* 9: e95502. doi:95510.91371/journal.pone.0095502.
- Liu, W., and Hulett, F.M. (1997) *Bacillus subtilis* PhoP binds to the phoB tandem promoter exclusively within the phosphate starvation-inducible promoter. *J. Bacteriol.* 179: 6302-6310.
- Lois, A.F. *et al.* (1993) Autophosphorylation and phosphatase activities of the oxygen-sensing protein FixL of *Rhizobium meliloti* are coordinately regulated by oxygen. *J. Biol. Chem.* 268: 4370-4375.
- Makino, K. *et al.* (1989) Signal transduction in the phosphate regulon of *Escherichia coli* involves phosphotransfer between PhoR and PhoB proteins. *J. Mol. Biol.* 210: 551-559.
- Parkinson, J.S., and Kofoid, E.C. (1992) Communication modules in bacterial signaling proteins. *Annu. Rev. Genet.* 26: 71-112.

- Schmöe, K. *et al.* (2011) Structural insights into Rcs phosphotransfer: the newly identified RcsD-ABL domain enhances interaction with the response regulator RcsB. *Structure* 19: 577-587.
- Schröder, I. *et al.* (1994) Phosphorylation and dephosphorylation of the NarQ, NarX, and NarL proteins of the nitrate-dependent two-component regulatory system of *Escherichia coli*. *J. Bacteriol.* 176: 4985-4992.
- Shi, L., and Hulett, F.M. (1999) The cytoplasmic kinase domain of PhoR is sufficient for the low phosphate-inducible expression of Pho regulon genes in *Bacillus subtilis*. *Mol. Microbiol.* 31: 211-222.
- Shi, L. *et al.* (1999) Decay of activated *Bacillus subtilis* Pho response regulator, PhoP~ P, involves the PhoR~ P intermediate. *Biochemistry* 38: 10119-10125.
- Siddiqui, K.S. *et al.* (2014) Versatile peroxidase degradation of humic substances: Use of isothermal titration calorimetry to assess kinetics, and applications to industrial wastes. *J. Biotechnol.* 178: 1-11.
- Stewart, R.C. (1997) Kinetic characterization of phosphotransfer between CheA and CheY in the bacterial chemotaxis signal transduction pathway. *Biochemistry* 36: 2030-2040.
- Stock, A.M. *et al.* (2000) Two-component signal transduction. *Annu. Rev. Biochem.* 69: 183-215.
- Todd, M.J., and Gomez, J. (2001) Enzyme kinetics determined using calorimetry: a general assay for enzyme activity. *Anal. Biochem.* 296: 179-187.
- Tracy, T.S., and Hummel, M.A. (2004) Modeling kinetic data from in vitro drug metabolism enzyme experiments. *Drug Metab. Rev.* 36: 231-242.
- West, A.H., and Stock, A.M. (2001) Histidine kinases and response regulator proteins in two-component signaling systems. *Trends Biochem. Sci.* 26: 369-376.
- Willett, J.W., and Kirby, J.R. (2012) Genetic and biochemical dissection of a HisKA domain identifies residues required exclusively for kinase and phosphatase activities. *PLoS Genet* 8: e1003084. doi:10.1371/journal.pgen.1003084 e1003084.
- Zhu, Y. *et al.* (2000) Phosphatase activity of histidine kinase EnvZ without kinase catalytic domain. *Proc. Natl. Acad. Sci. USA.* 97: 7808-7813.

Chapter 6

Analyses of temperature-responsive activity of LtrK

6.1 Abstract

Analyses of the effects of temperature on LtrK activity demonstrated that it had highest kinase and phosphatase activities at 10 °C and remained very stable at this temperature with a half-life of inactivation of ~2.8 days. The activity was also very high at 0 °C (~50 % of maximum) and rapidly fell away at temperatures above 20 °C with little activity at 30 °C (half-life of inactivation, ~24 min). The thermal stability of LtrK was examined using differential scanning calorimetry at different scan rates which showed scan rate dependent irreversible unfolding of LtrK. At the slowest scan rate (0.1 °C min⁻¹) the kinase initiates unfolding after ~20 °C which is consistent with its low temperature shifted activity profile. The data are consistent with the cytoplasmic domain of LtrK having the low temperature responsive property indicating that the LtrK can exhibit low-temperature responsive activity without its N-terminal transmembrane domain. The T_{opt} of the LtrK was low compared to the other psychrophilic enzymes. In addition, the temperature range of the kinase activity matched well with normal environmental temperatures of *M. burtonii* (1 - 2 °C) in Ace Lake, Antarctica and the temperature this psychrophile is capable of growing in the laboratory (-2 to 28 °C). These findings indicate that LtrK may perform temperature-responsive gene regulation in *M. burtonii*.

6.2 Introduction

Psychrophilic archaea have evolved a capacity to grow and survive at low temperature and are found to be abundant in the cold environment, especially in the cold ocean depth, *i.e.* $\sim 10^{28}$ cells has been estimated in oceans (Karner *et al.*, 2001). Cold adaptation mechanism analyses in few psychrophilic archaea, *i.e.* *M. burtonii* and *M. psychrophilus*, have found cold-enhanced expression of TCS genes indicating their involvement in cold responsive gene regulation (Campanaro *et al.*, 2011; Chen *et al.*, 2012). However, no temperature-responsive TCS have been characterized in archaea. In contrast many TCSs have been reported to be linked with thermal responses in bacteria (Klinkert and Narberhaus, 2009; Sengupta and Garrity, 2013) including *A. tumefaciens* (Jin *et al.*, 1993) and *P. syringae* (Braun *et al.*, 2007), *E. coli* (Maeda *et al.*, 1976), *B. subtilis* (Aguilar *et al.*, 2001), *E. tarda* (Chakraborty *et al.*, 2010), *Synechocystis* (Suzuki *et al.*, 2000) (described in Chapter 1).

Typically in a bacterial thermosensing TCSs the membrane bound SKs have the ability to sense the change in temperature and regulate the level of phosphorylated RRs in the cell which in turn regulate the gene expression in a temperature-dependent manner (Steinmann and Dersch, 2013; Saita *et al.*, 2015). The mechanism of thermosensing in many SKs is not clear. It has been observed in few SKs that the TMDs play a crucial role in sensing the temperature change and transferring the signal to the kinase core of the SKs, *e.g.* in DesK of *B. subtilis* (Albanesi *et al.*, 2004) and CorS of *Synechocystis* (Suzuki *et al.*, 2000), while few SKs have also demonstrated the thermosensitivity without the TMDs, *i.e.* VirA of *A. tumefaciens* (Jin *et al.*, 1993).

The mechanism of thermal regulation can be different in different SKs. In a few thermosensing TCSs, change in temperature may promote a change in membrane fluidity which in turn can function to switch the SKs between kinase and phosphatase activity. For example, in *B. subtilis* the cell membrane tends to become rigid at low temperature which promotes the DesK kinase activity, but at high temperature the membrane becomes fluid again which switches the activity of Desk from kinase to phosphatase (Albanesi *et al.*, 2004). Additionally, the thermosensitivity can also be regulated by protein stability. For instance, the thermosensing VirA is properly folded and active as a kinase in *A. tumefaciens* at low temperature but at high temperature (>32 °C) the kinase becomes inactive through protein unfolding (Jin *et al.*, 1993).

The TCS that showed differential gene expression in *M. burtonii* is LtrK/LtrR. Previous proteomic and transcriptomic analyses identified higher abundance of LtrR (Goodchild *et al.*, 2004; Campanaro *et al.*, 2011) and a recent transcriptomic analysis found higher abundance of both LtrK and LtrR in cells grown at low (4°C) vs. high (23°C) temperature (Taha *et al.*, 2016) which indicates that LtrK and LtrR may form a temperature responsive TCS. The cytoplasmic kinase core of LtrK and the full length LtrR had been purified (described in Chapter 3 and 4) and the activity analyses on the purified proteins showed that the cytoplasmic LtrK exerts both kinase activity (on itself and LtrR) and phosphatase activity (on LtrR-P) at room temperature (described in Chapter 5). To assess the thermo-sensitivity of LtrK, the kinase and phosphatase activity were investigated at various temperatures. In addition, the thermal stability of both proteins (LtrK and LtrR) was analyzed by performing DSC. The data gave an insight into the temperature dependent activity of both proteins.

6.3 Experimental Procedures

6.3.1 Kinase, phosphotransfer and phosphatase activities analyses at various temperatures

Autophosphorylation assay was performed by incubating 4 μg of LtrK in 20 μl of P buffer containing 0.2 $\mu\text{Ci } \mu\text{l}^{-1}$ [γ - ^{32}P]-ATP at 0, 5, 10, 15, 20, 25 and 30 $^{\circ}\text{C}$ in a refrigerated water bath equipped with a circulator (Julabo). At 10, 30, 60 and 120 min intervals, 6 μl samples were withdrawn and mixed with 4 μl of 4 \times NuPAGE LDS sample loading buffer. The samples were immediately heated at 95 $^{\circ}\text{C}$ for 3 min and subjected to SDS-PAGE followed by autoradiography. The amount of radioactivity associated with LtrK was calculated from the band intensity using the autoradiogram. The experiment was performed in duplicate and the mean percentage radioactivity, as the percentage of highest band intensity, was plotted against temperature.

To analyze the phosphatase activity at various temperatures, GST-LtrK (25 μg) bound to glutathione beads column was autophosphorylated in the presence of [γ - ^{32}P]-ATP for 30 min at 10 $^{\circ}\text{C}$ in the refrigerated water bath. After 30 min the column was placed at room temperature and free ATP was washed off using P buffer. LtrR (100 μg in 50 μl P buffer) was passed through the column and LtrR-P was collected with the flowthrough. Another 50 μl of buffer was passed through the column and collected with the previous flowthrough to recover all LtrR-P. The LtrR-P (10 μg) sample was incubated with LtrK in a 2:1 molar ratio for 10 min at 0, 5, 10, 15, 20, 25 and 30 $^{\circ}\text{C}$. Reaction was stopped by adding 5 μl of 4 \times NuPAGE LDS sample loading buffer containing 100 mM of EDTA. The samples were subjected to SDS-PAGE and analyzed by autoradiography as described for LtrK kinase activity.

The phosphotransfer activity of LtrK on LtrR was analysed at 0 °C and 25 °C. Two samples of 25 µg of GST-LtrK bound to glutathione beads were phosphorylated in presence of [γ -³²P]-ATP at room temperature in two gravity flow column for 30 min. The free [γ -³²P]-ATP was thoroughly washed off using P buffer. After that, one of the column was placed on ice at 0 °C and the other one was kept at room temperature at 25 °C for 10 min. Ice cold LtrR (100 µg) in 50 µl P buffer was added to the column on ice and an equivalent amount was warmed to room temperature for 5 min and then added to the column kept at room temperature. The flowthrough was immediately collected and 10 µl aliquots were mixed with 3 µl of sample loading buffer containing 100 mM EDTA, and subjected to SDS-PAGE followed by autoradiography.

6.3.2 Half-life of thermal inactivation assay

For determining the half-life ($t_{1/2}$) of thermal inactivation, LtrK was incubated at 10 and 30 °C in 50 µl P buffer to a final concentration of 0.5 mg ml⁻¹. LtrK samples (4 µl) that were already at 10 °C, were collected, at 0, 1, 5, 10, 24, 36, 48, 72 and 96 h and mixed with 3 µl of P buffer containing 0.5 µCi µl⁻¹ [γ -³²P]-ATP. The same procedure was repeated for the samples incubated at 30 °C and aliquots were withdrawn at 0, 5, 15, 30, 60 min and incubated at 10 °C for 2 to 3 min to cool it before mixing with 3 µl of the same [γ -³²P]-ATP containing buffer. The autophosphorylation reaction was carried out at 10 °C for 10 min. The reaction was stopped by adding 5 µl of LDS sample loading buffer followed by heating at 95 °C for 3 min and the samples were subjected to SDS-PAGE followed by autoradiography. The protocol is shown in Fig. 6.1.

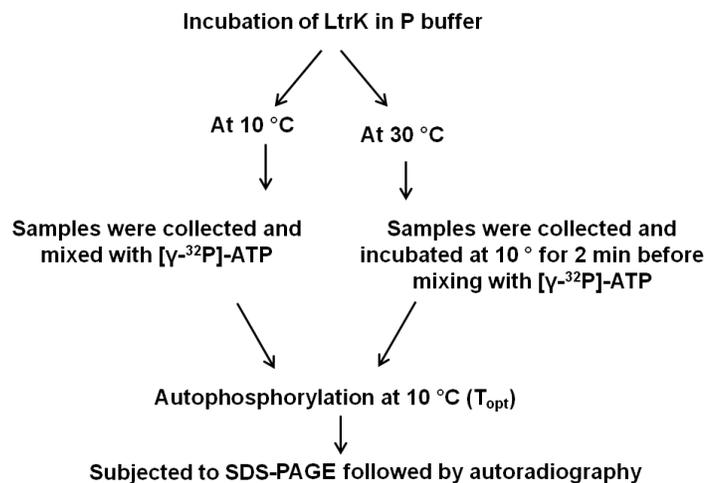


Figure 6.1 Flowchart showing autophosphorylation protocol for inactivation assay

The amount of radioactivity associated with LtrK was calculated from the band intensity using the autoradiogram and the natural log (ln) of activity was plotted against incubation time. The inactivation constant (k_{in}) was directly determined from the slope of the first-order plot of ln activity vs. time and half-life ($t_{1/2}$) was determined as $\ln 2/k_{in}$.

6.3.3 Differential scanning calorimetry (DSC) at various scan rates

DSC was performed as described in Chapter 3, section 3.3.9 at three different scan rates: 1.0, 0.2 and 0.1 °C min⁻¹. Proteins were concentrated to 3 mg ml⁻¹ using 50 mM HEPES pH 8.0, 50 mM KCl, 20 % glycerol, 5 mM MgCl₂, 5 mM TCEP and 0.5 mM ATP for LtrK, and 20 mM HEPES pH 7.5 and 250 mM NaCl for LtrR. For each run the calorimeter was set to perform a series of scans using identical parameters with a temperature range from 4 to 90 °C.

6.4 Results

6.4.1 Kinase and phosphatase activity of LtrK is high at low temperature

To analyze the temperature dependence of the kinase activity, autophosphorylation assay was performed by incubating the enzyme with [γ - ^{32}P]-ATP at 0, 5, 10, 15, 20, 25 and 30 °C. The analyses showed maximum kinase activity at 10 °C (T_{opt}) (Fig. 6.2A). At 0 and 5 °C, low levels of LtrK-P were detected after 30 min of incubation which subsequently increased with increasing incubation time and after 2 h of incubation the level of radioactive LtrK-P almost (~85%) reached the maximum level observed at 10 °C (Fig. 6.2A). This indicated that the enzyme was very active at temperatures as low as 0 °C. The enzyme was also relatively active above 10 °C but after 20 °C the activity rapidly decreased. Extremely low levels of LtrK-P activity was observed at 30 °C which initially slightly increased with extended incubation but after 1 h no more increase in kinase activity was observed (Fig. 6.2A) which indicated that the enzyme lost its activity after 1 h of incubation at 30 °C.

Phosphatase activity of LtrK was also evaluated at different temperatures by incubating LtrK with phosphorylated regulator, LtrR-P at 0, 5, 10, 15, 20, 25 and 30 °C. Analyses of LtrR-P stability at room temperature have shown that LtrR possesses slow autophosphatase activity (describe in Chapter 5) which may increase at higher temperature (*i.e.* 30 °C) due to the kinetic effect of heat on the reaction rate. Moreover, the phosphatase activity of LtrK at room temperature have shown that the enzyme dephosphorylate LtrR-P in a reverse phosphotransfer reaction where loss in ^{32}P signal of LtrR-P is equivalent to the incorporation of ^{32}P with LtrK-P (described in Chapter 5). Therefore to analyse specifically the LtrK phosphatase activity, the radioactivity associated with LtrK instead of

LtrR was plotted. The phosphatase activity profile of LtrK was similar to the kinase profile showing maximum activity at 10 °C (Fig. 6.2B). A moderate phosphatase activity was observed at 0 °C and 5 °C and also above 10 °C. But similar to kinase activity, the phosphatase activity decreased after 20 °C with lowest activity at 30 °C.

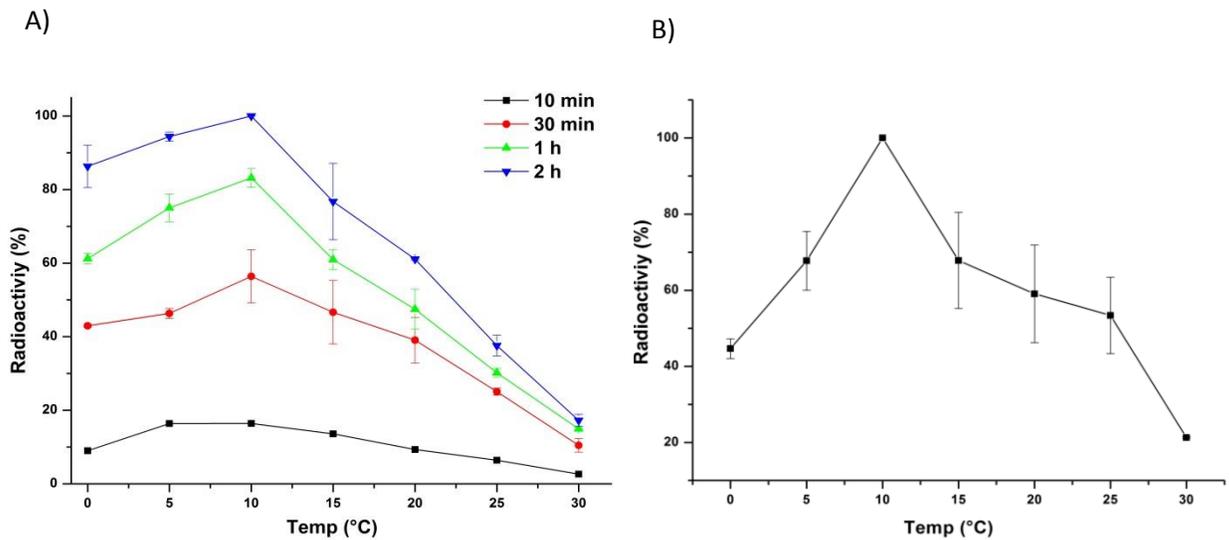


Figure 6.2 Effect of temperature on kinase and phosphatase activities of LtrK

To assess the phosphorylation state of proteins, samples were electrophoresed on a SDS-polyacrylamide gel, and autoradiography performed by phosphorimaging (panels A, B). The radioactivity was directly derived from the band intensity and the mean values for two replicates are plotted. Error bars represent the standard error of the mean. A) Effect of temperature on autophosphorylation. Autophosphorylation was performed at different temperatures (0, 5, 10, 15, 20, 25, 30 °C) with aliquots withdrawn for analysis at different times of incubation (10 min, 0.5 h, 1 h, 2 h). Incorporation is shown as a percentage of the highest band intensity on autoradiograms across all samples (2 h at 10 °C). The mean values for two replicates are plotted for 0.5 h, 1 h and 2 h, and values for a single time course for 10 min; B) Effect of temperature on phosphatase activity. LtrR-P was incubated with LtrK in 2:1 molar ratio for 10 min at different temperatures (0, 5, 10, 15, 20, 25, 30 °C) and the band intensity of LtrK-P plotted as a percentage of the highest band intensity on the autoradiograms (LtrK-P at 10°C).

The assessment of the phosphotransfer activity of LtrK on LtrR involves GST-LtrK autophosphorylation in a glutathione beads column, washing of free $[\gamma\text{-}^{32}\text{P}]\text{-ATP}$ and passing of LtrR through the column. As the protocol involved several steps in a glutathione beads column, the phosphotransfer activity could not be performed in the refrigerated water bath at different temperatures. Instead, the activity was analysed at 0 °C on ice and at room temperature (25 °C). The analyses showed enhanced phosphotransfer activity at cold temperature (Fig. 6.3). The amount of ^{32}P incorporation with LtrR was ~2 fold higher at 0 °C compared to 25 °C, consistent with the higher autophosphorylation and phosphatase activity of LtrK for the same temperatures (Fig. 6.2).

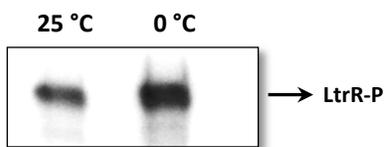


Figure 6.3 Effect of temperature on LtrK phosphotransfer activity.

GST-LtrK (25 μg) bound to glutathione beads was phosphorylated in presence of $[\gamma\text{-}^{32}\text{P}]\text{-ATP}$ at room temperature for 30 min, free $[\gamma\text{-}^{32}\text{P}]\text{-ATP}$ washed off and one column was placed on ice (0 °C) and the other was placed at room temperature (25 °C) for 10 min. Ice cold LtrR (100 μg) was passed through the cold column and an equivalent amount of LtrR was warmed to room temperature for 5 min and then added to the column kept at room temperature. LtrR-P was collected in the flowthrough and 10 μl samples were added to 3 μl of 100 mM EDTA containing sample buffer and analysed by gel-phosphorimaging.

6.4.2 LtrK is more stable at 10 °C compared to 30 °C

The low T_{opt} and temperature-dependent reduction of kinase and phosphatase activity of LtrK at 30 °C was indicative of the enzyme becoming rapidly inactivated at high temperature. To assess further, the half-lives of inactivation ($t_{1/2}$) were determined at 10 °C

(T_{opt}) (Fig. 6.4A) and 30 °C (Fig. 6.4B) by incubating the enzyme at 10 or 30 °C and assaying the autophosphorylation at different time intervals. The results showed that LtrK was very stable at 10 °C with a $t_{1/2}$ of ~ 2.8 d, whereas at 30 °C $t_{1/2}$ was ~ only 24 min. The data is consistent with the result of autophosphorylation at different temperature where LtrK-P activity discontinued after 1 h of incubation at 30 °C (Fig. 6.2A).

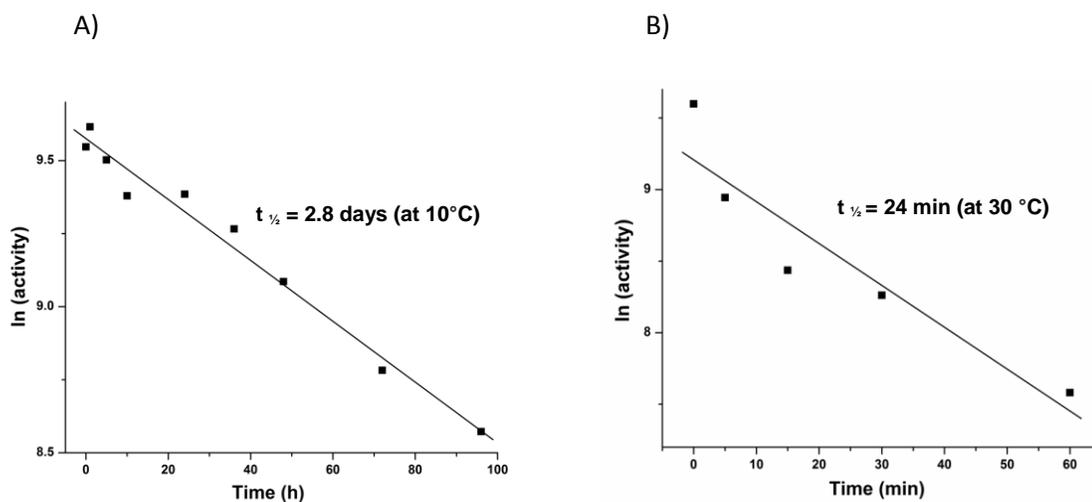


Figure 6.4 Half-life of inactivation at 10 °C and 30 °C

A) Half-life of inactivation at 10 °C. LtrK was incubated at 10°C for up to 4 d and residual autophosphorylation activity was determined by incubating aliquots of the enzyme with [γ - 32 P]-ATP for 10 min at 10 °C (T_{opt}). The natural log (ln) of activity (band intensity) was plotted against incubation time. The straight line represents the linear fit to the data and the slope of the line was used to calculate $t_{1/2}$ (see Experimental procedures); B) Half-life of inactivation at 30 °C. As for panel A, except LtrK was incubated at 30°C.

6.4.3 The thermal unfolding of LtrK is scan-rate dependent

Thermal denaturation of LtrK was already analysed by DSC which showed that the protein unfolds irreversibly with a T_m of 63 °C at 1 °C min^{-1} scan rate (described in Chapter 3, section 3.4.3). In order to assess the thermal instability of the enzyme DSC was

performed at different scan rates. The melt profile demonstrated that the T_m of LtrK greatly depends on the scan rate and shifts towards lower temperature at slower scan rates (42 °C at 0.2 °C min⁻¹ and 33 °C at 0.1 °C min⁻¹, Fig. 6.5A).

The thermal unfolding of LtrR was also examined at different scan rates to determine if the regulator also unfolds in a scan-rate dependent manner. The DSC melt of LtrR already showed reversibly unfolding profile with a T_m of 57 °C (described in Chapter 4). Thermal denaturation at slower scan rate, 0.1 °C min⁻¹, did not change the melting profile and the T_m remained the same (57 °C) (Fig. 6.5B) which indicates that the thermal unfolding of LtrR is scan-rate independent. Moreover, the LtrR starts melting after ~ 40 °C (both with 1.0 °C min⁻¹ and 0.1 °C min⁻¹) while with slowest scan rate (0.1 °C min⁻¹) LtrK starts melting at ~ 20 °C demonstrating that the LtrR is more stable than LtrK.

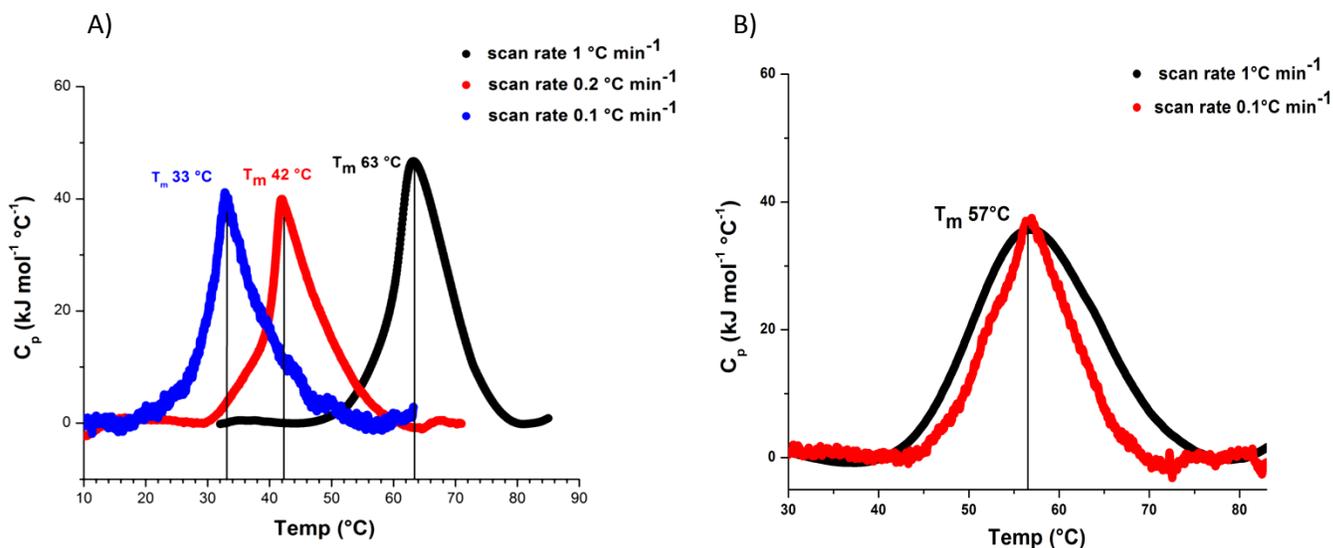


Figure 6.5 Thermal unfolding of LtrK and LtrR at different scan rate

A) DSC plot of LtrK (after baseline correction) with different scan rates illustrating T_m downshift with slower scan rate. B) DSC plot (after baseline correction) with different scan rate illustrating scan rate independent thermal unfolding of LtrR.

6.5 Discussion

The enzyme activity analyses at various temperatures determined that LtrK had high activity (both kinase and phosphatase) at low temperatures with optimum activity at 10 °C and the activity started to decrease after 20 °C with lowest activity at 30 °C (Fig. 6.2). The phosphotransfer activity also showed enhanced activity at low (0 °C) compared to high temperature (25 °C) (Fig. 6.3). The half-life of inactivation assay further showed prolonged half-life of inactivation at 10 °C (2.8 d) compared to at 30 °C (24 min) (Fig. 6.4). These results demonstrated the temperature dependent activity of LtrK.

The scan rate dependent irreversible thermal unfolding profile of LtrK revealed by DSC (Fig. 6.5A) indicated that the enzyme is structurally unstable at higher temperature. At the slowest scan rate (0.1 °C min⁻¹), LtrK started to unfold at ~20 °C while at highest scan rate (1 °C min⁻¹) LtrK started unfolding at 50 °C (Fig. 6.5A). At 0.1 °C min⁻¹ scan rate the machine took ~8.3 h to reach 20 °C whereas at 1 °C min⁻¹ it took 20 min only. So, heating the LtrK up to 20 °C for 8.3 h initiates unfolding whereas heating up to the same temperature for 20 min does not affect the folding properties. Thus if the melting was performed at even slower scan rate (*e.g.* 0.05 °C min⁻¹) the enzyme might have started to unfold at an even lower temperature (*i.e.* < 20 °C). This thermal unstable nature of LtrK also explains the reduced half-life (~24 min) of inactivation at 30 °C (Fig. 6.4). In contrast, the thermal unfolding of LtrR is not scan rate dependent and the regulator is more stable than LtrK with a T_m of 57 °C (Fig. 6.5B). This indicates that only LtrK exhibits temperature dependent activity involving irreversible thermal unfolding.

The T_{opt} (10 °C) of LtrK is much lower than the T_m of thermal unfolding (63 °C, 42 °C and 33 °C at scan rate 1, 0.2 and 0.1 °C min⁻¹ respectively). Typically the T_{opt} of the

enzymes tend to be close to the T_m due to the kinetic effect of temperature on the reaction rate. The reaction rate tends to increase with increasing temperature up until the point at which the enzyme starts to unfold (near T_m) (Feller and Gerday, 2003; Cavicchioli, 2006). But when the active site of the enzyme is more flexible and melts (local unfolding) before the overall protein structure (global unfolding) the enzyme loses its activity at temperatures lower than the T_m . Unlike mesophiles or thermophiles, psychrophilic enzymes tend to inactivate at temperatures lower than the overall protein unfolding temperature (T_m) due to the local unfolding nature (Feller and Gerday, 2003; Georlette *et al.*, 2003; Siddiqui and Cavicchioli, 2006). Therefore, the $T_{opt} < T_m$ indicates that the active-site of LtrK is more thermolabile and unfolds much before than the overall protein structure.

However, the 10 °C T_{opt} of LtrK is low, even for an enzyme from a psychrophile. After examining 33 different proteins purified from psychrophiles (including two from *Archaea*, four from *Eucarya* and 27 from *Bacteria*), the T_{opt} was found to range from 16 – 64 °C with an average of ~ 36 °C ± 12 (Table 6.1). The lowest T_{opt} was found to be 16 °C for an Antarctic marine bacterium DNA ligase (Georlette *et al.*, 2003).

Table 6.1 Temperature optimum values for proteins from psychrophiles.

Organism	Protein	T_{opt}	Reference
<i>Bacteria</i>			
Antarctic bacterium	Citrate synthase	31 °C	Russel <i>et al.</i> , 1998
Antarctic <i>Pseudomonas</i> sp.	Phosphoglycerate Kinase	35°C	Bentahir <i>et al.</i> , 2000
Antarctic <i>Pseudomonas syringae</i>	RNA polymerase	37 °C	Uma <i>et al.</i> , 1999
Antarctic seawater bacterium	Alkaline phosphatase	25 °C	Kobori <i>et al.</i> , 1984
Antarctic seawater bacterium	DNA ligase	16 °C	Georlette <i>et al.</i> , 2003
Arctic sea shore sediment bacterium	Esterase	30 °C	Jeon <i>et al.</i> , 2009a
<i>Bacillus globisporus</i>	Adenylate kinase	35 °C	Bae and Phillips, 2004
<i>Bacillus</i> TA41	Subtilisin	40°C	Davail <i>et al.</i> , 1994
<i>Carnobacterium piscicola</i> BA	β -galactosidase	30 °C	Coombs and Brenchley, 1999
<i>Chlamys islandica</i>	Chlamysin	22 °C	Nilsen <i>et al.</i> , 1999
<i>Colwellia psychrerythraea</i>	Aminopeptidase	39 °C	Huston <i>et al.</i> , 2008

<i>Colwellia psychrerythraea</i>	Phenylalanine hydroxylase	25 °C	Leiros <i>et al.</i> , 2007
<i>Colwellia maris</i>	Malate synthase	45 °C	Watanabe <i>et al.</i> , 2001
<i>Colwellia maris</i>	Isocitratelase	20 °C	Watanabe <i>et al.</i> , 2001
Cold sea sediment sample bacterium	Lipase	25 °C	Jeon <i>et al.</i> , 2009b
<i>Cytophaga</i> sp.	Valine dehydrogenase	20 °C	Oikawa <i>et al.</i> , 2001
<i>Desulfotalea psychrophila</i>	Isocitrate dehydrogenase	45 °C	Fedoy <i>et al.</i> , 2007
Marine psychrophile strain PA-43	Serine protease	58 °C	Irwin <i>et al.</i> , 2001
<i>Micrococcus</i> sp.	Pollulanase (extracellular)	50 °C	Kimura and Horikoshi, 1990
<i>Moraxella</i> sp.	Alcohol dehydrogenase	25 °C	Tsigos <i>et al.</i> , 1998
<i>Pseudoalteromonas haloplanktis</i>	α -amylase	28 °C	D'Amico <i>et al.</i> , 2003
<i>P. haloplanktis</i>	Cellulase	40 °C	Garsoux <i>et al.</i> , 2004
<i>P. haloplanktis</i>	Xylanase	35 °C	Collins <i>et al.</i> , 2003
<i>P. haloplanktis</i>	Pectate lyase (extracellular)	30 °C	Tuyen <i>et al.</i> , 2001
<i>P. haloplanktis</i>	Aspartate amino transferase	64 °C	Birolo <i>et al.</i> , 2000
<i>Pseudomonas</i> strain DY-A	Serine Alkaline protease	40 °C	Zeng <i>et al.</i> , 2003
<i>Sphingomonas paucimobilis</i> (Antarctic marine bacterium)	Extracellular metalloprotease	25 °C	Turkiewicz <i>et al.</i> , 1999
Archaea			
<i>Cenarchaeum symbiosum</i>	DNA polymerase	40 °C	Schleper <i>et al.</i> , 1997
<i>Methanococcoides burtonii</i>	EF2	34 °C	Thomas and Cavicchioli, 2000
Eucarya			
Antarctic fish	Lactate dehydrogenase	50 °C	Coquelle <i>et al.</i> , 2007
Fish liver	Imidase	55 °C	Huang and Yang, 2003
Psychrophilic green alga	Argininosuccinatelyase	37 °C	Loppes <i>et al.</i> , 1996
<i>Sclerotinia borealis</i>	Polygalacturonase (extracellular)	45 °C	Takasawa <i>et al.</i> , 1997

It is noteworthy that though the psychrophiles grow at temperatures near to 0 °C, the T_{opt} of their enzymes was found to be much higher than the environmental growth temperature (Table 6.1). The kinetic effect of heat on the enzyme reaction rate typically results in higher T_{opt} for the enzymes activity (Feller and Gerday, 2003; Cavicchioli, 2006, 2015a). Thus, the higher T_{opt} of the cold-adapted enzymes does not imply the inactivity or greatly reduced activity at the environmental low temperatures that is close to 0 °C. In fact, the primary distinguishing feature of psychrophilic enzymes from its mesophilic

counterparts is the higher catalytic efficiency at low temperatures (Gerday *et al.*, 1997). A similar trend is also observed for the growth rate of psychrophilic organisms. The kinetic effects of temperature on the rate of reactions lead the psychrophilic organisms to grow faster at higher temperatures than their environmental growth temperature, up until the point at which cellular processes become sufficiently compromised *e.g.* T_{opt} of *M. burtonii* growth is 23 °C while the native growth temperature is 1-2 °C (Williams *et al.*, 2011). Therefore, T_{opt} does not necessarily reflect the adaptation or fitness of the organism (Cavicchioli, 2015b). T_{opt} of the enzyme (or organism growth) is only considered as a useful relative value to compare between similar enzymes or organisms (Gerday *et al.*, 1997).

M. burtonii is capable of growth in the laboratory between -2 °C and 28 °C, and has been found to be heat stressed at temperatures between 23 °C and 28 °C, cold stressed at -2 °C, and grows normally between 1–16 °C (Williams *et al.*, 2011). The activity features of LtrK, including high activity between 0 - 10 °C, inactivation at 30 °C and thermal unfolding at lower temperature matches well with the environmental temperatures (1 - 2 °C) of *M. burtonii* in Antarctica and the temperature this organism is capable of growing in the laboratory (-2 to 28 °C) (Fig. 6.7).

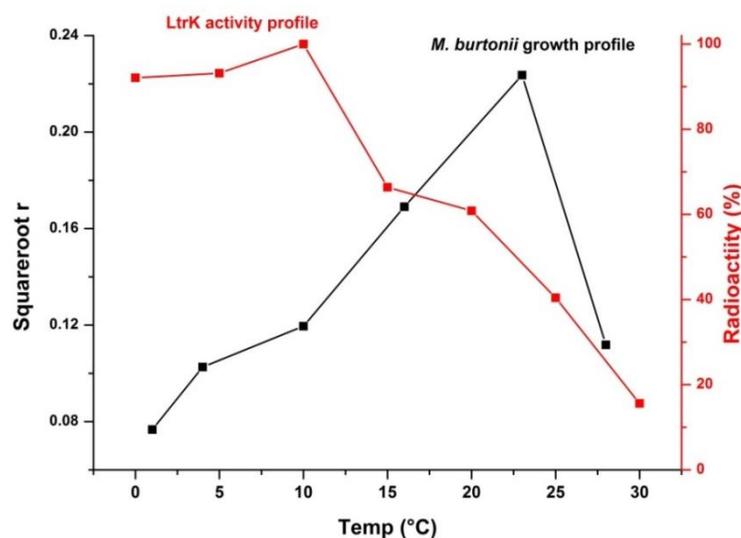


Figure 6.7 Comparison of LtrK activity profile with *M. burtonii* growth profile

LtrK autophosphorylation profile (red) after incubating with ATP for 2 h is aligned with the growth profile of *M. burtonii* (black). The data for *M. burtonii* growth profile (temperature vs. squareroot r) has been derived from Williams *et al.* 2011. “Squareroot r” is the squareroot of the reciprocal of the time taken to increase the initial absorbance at 620 nm by 0.25 absorbance unit.

The psychrophilic enzymes in Table 6.1 reflect the fact that they can have very high activity at high temperatures though they belong to psychrophilic organisms that grow near to 0 °C. However, LtrK is an exception having a lower T_{opt} than enzymes from other psychrophile, which also matches the growth profile of *M. burtonii* (Fig. 6.7). This is indicative of evolutionary selection pressure on the enzyme to have low temperature-dependent properties to complement the growth temperature profile of the organism. It also implies that the LtrK/LtrR TCS may fulfil a physiological role in regulating gene expression in response to growth temperature in *M. burtonii*.

Based on the low T_{opt} , thermolabile DSC profile and correlation of activity with the organism’s growth profile, it is clear that LtrK is a temperature sensitive protein and that it is an intrinsic property. To-date, very little is known about the temperature sensitive SKs and biochemical analyses data for thermosensing of SKs is limited to *B. subtilis*

DesK/DesR TCS (Aguilar *et al.*, 2001) and *A. tumefaciens* VirA/VirG TCS (Jin *et al.*, 1993). In *B. subtilis*, the TCS Desk/DesR stringently controls the Des pathway which responds to a decrease in the growth temperature and preserves the membrane fluidity upon temperature change (Aguilar *et al.*, 1998; Aguilar *et al.*, 2001). DesK is a membrane bound kinase which has five TMDs with small periplamic region in between. At low temperatures (25 °C or lower) when the cell membrane tends to become rigid, DesK phosphorylates the regulator DesR that in turn act as a transcriptional activator for *des* gene, coding for $\Delta 5$ fatty acid desaturase ($\Delta 5$ -des); $\Delta 5$ -des causes unsaturation of fatty acyl chains of membrane phospholipids that promotes the membrane to become fluid again (Aguilar *et al.*, 1998; Aguilar *et al.*, 2001; Albanesi *et al.*, 2004). $\Delta 5$ -des is barely detected at 37 °C and the depression of *des* exclusively occurs at the level of transcription in a promoter dependent manner (Aguilar *et al.*, 1999). Analyses on the molecular mechanism of thermosensing showed that temperature responsive activity is dependent on attachment of DesK to the membrane via its TMDs that can sense the membrane thickening at low temperatures and trigger the kinase (DesK) into an active state by inducing conformational changes (Albanesi *et al.*, 2009; Mariana and Diego, 2013) (described in Chapter 1).

The sensor kinase VirA of *A. tumefaciens* was also found to be temperature responsive and at temperature < 32 °C, it phosphorylates the RR, VirG that regulates the *vir* gene expression and induces tumor formation in plants (Jin *et al.*, 1990; Jin *et al.*, 1993). However, unlike DesK, VirA has two TMDs with a large periplasmic region and was found to be temperature responsive without its TMDs through reversible thermal unfolding of the cytoplasmic catalytic region (Jin *et al.*, 1993).

So, structurally and functionally LtrK is more similar to VirA than DesK. Like VirA, LtrK has two TMDs with large periplasmic domain and shows temperature sensitivity without the TMDs through thermal unfolding of the cytoplasmic kinase core. But the thermal unfolding of LtrK is irreversible while it is reversible in VirA.

In summary, the analyses results are indicative of LtrK being a temperature responsive SK that mediates the thermosensitivity using its irreversible thermal unfolding nature and may regulate the gene expression in response to the *M. burtonii* growth temperature change via regulating the level of phosphorylated LtrR in the cell.

6.6 References

- Aguilar, P.S. *et al.* (1998) A *Bacillus subtilis* gene induced by cold shock encodes a membrane phospholipid desaturase. *J. Bacteriol.* 180: 2194-2200.
- Aguilar, P.S. *et al.* (1999) Transcriptional control of the low-temperature-inducible *des* gene, encoding the $\Delta 5$ desaturase of *Bacillus subtilis*. *J. Bacteriol.* 181: 7028-7033.
- Aguilar, P.S. *et al.* (2001) Molecular basis of thermosensing: a two-component signal transduction thermometer in *Bacillus subtilis*. *EMBO J.* 20: 1681-1691.
- Albanesi, D. *et al.* (2004) The membrane fluidity sensor DesK of *Bacillus subtilis* controls the signal decay of its cognate response regulator. *J. Bacteriol.* 186: 2655-2663.
- Albanesi, D. *et al.* (2009) Structural plasticity and catalysis regulation of a thermosensor histidine kinase. *Proc. Natl. Acad. Sci. USA.* 106: 16185-16190.
- Braun, Y. *et al.* (2007) A temperature-sensing histidine kinase—function, genetics, and membrane topology. *Methods Enzymol.* 423: 222-249.
- Bae, E., and Phillips, G.N. (2004) Structures and analysis of highly homologous psychrophilic, mesophilic, and thermophilic adenylate kinases. *J. Biol. Chem.* 279: 28202-28208.
- Bentahir, M. *et al.* (2000) Structural, kinetic, and calorimetric characterization of the cold-active phosphoglycerate kinase from the Antarctic *Pseudomonas* sp. TACII18. *J. Biol. Chem.* 275: 11147-11153.

- Birolo, L. *et al.* (2000) Aspartate aminotransferase from the Antarctic bacterium *Pseudoalteromonas haloplanktis* TAC 125. *J. Biochem.* 267: 2790-2802.
- Campanaro, S. *et al.* (2011) Temperature-dependent global gene expression in the Antarctic archaeon *Methanococcoides burtonii*. *Environ. Microbiol.* 13: 2018-2038.
- Cavicchioli, R. (2006) Cold-adapted archaea. *Nat. Rev. Microbiol.* 4: 331-343.
- Cavicchioli, R. (2015a) Microbial ecology of Antarctic aquatic systems. *Nat. Rev. Microbiol.* 13: 691-706.
- Cavicchioli, R. (2015b) On the concept of a psychrophile. *ISME J.* 10.1038/ismej.2015.160.
- Chakraborty, S. *et al.* (2010) Temperature and Mg²⁺ sensing by a novel PhoP-PhoQ two-component system for regulation of virulence in *Edwardsiella tarda*. *J. Biol. Chem.* 285: 38876-38888.
- Chen, Z. *et al.* (2012) The genome and transcriptome of a newly described psychrophilic archaeon, *Methanlobus psychrophilus* R15, reveal its cold adaptive characteristics. *Environ. Microbiol. Rep.* 4: 633-641.
- Collins, T. *et al.* (2003) Activity, stability and flexibility in glycosidases adapted to extreme thermal environments. *J. Mol. Biol.* 328: 419-428.
- Coombs, J.M., and Brenchley, J.E. (1999) Biochemical and phylogenetic analyses of a cold-active β -galactosidase from the lactic acid bacterium *Carnobacterium piscicola* BA. *Appl. Environ. Microbiol.* 65: 5443-5450.
- Coquelle, N. *et al.* (2007) Activity, stability and structural studies of lactate dehydrogenases adapted to extreme thermal environments. *J. Mol. Biol.* 374: 547-562.
- D'Amico, S. *et al.* (2003) Temperature adaptation of proteins: engineering mesophilic-like activity and stability in a cold-adapted α -amylase. *J. Mol. Biol.* 332: 981-988.
- Davail, S. *et al.* (1994) Cold adaptation of proteins. Purification, characterization, and sequence of the heat-labile subtilisin from the antarctic psychrophile *Bacillus* TA41. *J. Biol. Chem.* 269: 17448-17453.
- Fedøy, A.-E. *et al.* (2007) Structural and functional properties of isocitrate dehydrogenase from the psychrophilic bacterium *Desulfotalea psychrophila* reveal a cold-active enzyme with an unusual high thermal stability. *J. Mol. Biol.* 372: 130-149.
- Feller, G., and Gerday, C. (2003) Psychrophilic enzymes: hot topics in cold adaptation. *Nat. Rev. Microbiol.* 1: 200-208.

- Garsoux, G. *et al.* (2004) Kinetic and structural optimization to catalysis at low temperatures in a psychrophilic cellulase from the Antarctic bacterium *Pseudoalteromonas haloplanktis*. *Biochem. J* 384: 247-253.
- Georlette, D. *et al.* (2003) Structural and functional adaptations to extreme temperatures in psychrophilic, mesophilic, and thermophilic DNA ligases. *J. Biol. Chem.* 278: 37015-37023.
- Gerday, C. *et al.* (1997) Psychrophilic enzymes: a thermodynamic challenge. *Biochim. Biophys. Acta -Protein Structure and Molecular Enzymology* 1342: 119-131.
- Goodchild, A. *et al.* (2004) A proteomic determination of cold adaptation in the Antarctic archaeon, *Methanococoides burtonii*. *Mol. Microbiol.* 53: 309-321.
- Huang, C.-Y., and Yang, Y.-S. (2003) A novel cold-adapted imidase from fish *Oreochromis niloticus* that catalyzes hydrolysis of maleimide. *Biochem. Biophys. Res. Commun.* 312: 467-472.
- Huston, A.L. *et al.* (2008) Cold adaptation of enzymes: Structural, kinetic and microcalorimetric characterizations of an aminopeptidase from the Arctic psychrophile *Colwellia psychrerythraea* and of human leukotriene A4 hydrolase. *Biochim. Biophys. Acta* 1784: 1865-1872.
- Irwin, J.A. *et al.* (2001) Purification and characterisation of a serine peptidase from the marine psychrophile strain PA-43. *FEMS Microbiol. Lett.* 201: 285-290.
- Jeon, J.H. *et al.* (2009a) Characterization and its potential application of two esterases derived from the arctic sediment metagenome. *Mar. Biotechnol.* 11: 307-316.
- Jeon, J.H. *et al.* (2009b) Cloning and characterization of a new cold-active lipase from a deep-sea sediment metagenome. *Appl. Microbiol. Biotechnol.* 81: 865-874.
- Jin, S. *et al.* (1990) Phosphorylation of the VirG protein of *Agrobacterium tumefaciens* by the autophosphorylated VirA protein: essential role in biological activity of VirG. *J. Bacteriol.* 172: 4945-4950.
- Jin, S. *et al.* (1993) The regulatory VirA protein of *Agrobacterium tumefaciens* does not function at elevated temperatures. *J. Bacteriol.* 175: 6830-6835.
- Kimura, T., and Horikoshi, K. (1990) Characterization of pullulan-hydrolysing enzyme from an alkalopsychrotrophic *Micrococcus* sp. *Appl. Microbiol. Biotechnol.* 34: 52-56.
- Klinkert, B., and Narberhaus, F. (2009) Microbial thermosensors. *Cell. Mol. Life Sci.* 66: 2661-2676.

- Kobori, H. *et al.* (1984) Heat-labile alkaline phosphatase from Antarctic bacteria: rapid 5' end-labeling of nucleic acids. *Proc. Natl. Acad. Sci. USA.* 81: 6691-6695.
- Leiros, H.-K.S. *et al.* (2007) Structure of phenylalanine hydroxylase from *Colwellia psychrerythraea* 34H, a monomeric cold active enzyme with local flexibility around the active site and high overall stability. *J. Biol. Chem.* 282: 21973-21986.
- Li, J. *et al.* (2014) Characterization of an archaeal two-component system that regulates methanogenesis in *Methanosaeta harundinacea*. *PLoS one* 9: e95502. doi:95510.91371/journal.pone.0095502.
- Lopes, R. *et al.* (1996) Effect of temperature on two enzymes from a psychrophilic chloromonas (chlorophyta) 1. *J. Phycol.* 32: 276-278.
- Maeda, K. *et al.* (1976) Effect of temperature on motility and chemotaxis of *Escherichia coli*. *J. Bacteriol.* 127: 1039-1046.
- Mariana, M., and Diego, d.M. (2013) Regulation of *Bacillus subtilis* DesK thermosensor by lipids. *Biochem. J.* 451: 269-275.
- Nilsen, I.W. *et al.* (1999) Protein purification and gene isolation of chlamysin, a cold-active lysozyme-like enzyme with antibacterial activity. *FEBS Lett.* 464: 153-158.
- Oikawa, T. *et al.* (2001) Psychrophilic valine dehydrogenase of the antarctic psychrophile, *Cytophaga* sp. KUC-1. *Eur. J. Biochem.* 268: 4375-4383.
- Russell, R.J. *et al.* (1998) Structural adaptations of the cold-active citrate synthase from an Antarctic bacterium. *Structure* 6: 351-361.
- Saita, E. *et al.* (2015) A coiled coil switch mediates cold sensing by the thermosensory protein DesK. *Mol. Microbiol.* 98: 258-271.
- Schleper, C. *et al.* (1997) Characterization of a DNA polymerase from the uncultivated psychrophilic archaeon *Cenarchaeum symbiosum*. *J. Bacteriol.* 179: 7803-7811.
- Sengupta, P., and Garrity, P. (2013) Sensing temperature. *Curr. Biol.* 23: R304-R307.
- Siddiqui, K.S., and Cavicchioli, R. (2006) Cold-adapted enzymes. *Annu. Rev. Biochem.* 75: 403-433.
- Steinmann, R., and Dersch, P. (2013) Thermosensing to adjust bacterial virulence in a fluctuating environment. *Future Microbiol.* 8: 85-105.
- Suzuki, I. *et al.* (2000) The pathway for perception and transduction of low-temperature signals in *Synechocystis*. *EMBO J.* 19: 1327-1334.

Taha *et al.* (2016) Single TRAM domain RNA-binding proteins in Archaea: functional insight from Ctr3 from the Antarctic methanogen *Methanococcoides burtonii*. *Environ Microbiol.* doi: 10.1111/1462-2920.13229

Takasawa, T. *et al.* (1997) Polygalacturonase isolated from the culture of the psychrophilic fungus *Sclerotinia borealis*. *Can. J. Microbiol.* 43: 417-424.

Thomas, T., and Cavicchioli, R. (2000) Effect of temperature on stability and activity of elongation factor 2 proteins from Antarctic and thermophilic methanogens. *J. Bacteriol.* 182: 1328-1332.

Tsigos, I. *et al.* (1998) Purification and characterization of an alcohol dehydrogenase from the Antarctic psychrophile *Moraxella* sp. TAE123. *Eur. J. Biochem.* 254: 356-362.

Turkiewicz, M. *et al.* (1999) Biosynthesis and properties of an extracellular metalloprotease from the Antarctic marine bacterium *Sphingomonas paucimobilis*. *J. Biotechnol.* 70: 53-60.

Tuyen, H. *et al.* (2001) Cloning of two pectate lyase genes from the marine Antarctic bacterium *Pseudoalteromonas haloplanktis* strain ANT/505 and characterization of the enzymes. *Extremophiles* 5: 35-44.

Uma, S. *et al.* (1999) A RNA polymerase with transcriptional activity at 0 °C from the Antarctic bacterium *Pseudomonas syringae*. *FEBS Lett.* 453: 313-317.

Watanabe, S. *et al.* (2001) Purification and characterization of a cold-adapted isocitrate lyase and a malate synthase from *Colwellia maris*, a psychrophilic bacterium. *Biosci. Biotechnol. Biochem.* 65: 1095-1103.

Williams, T.J. *et al.* (2011) Defining the response of a microorganism to temperatures that span its complete growth temperature range (– 2 °C to 28 °C) using multiplex quantitative proteomics. *Environ. Microbiol.* 13: 2186-2203.

Zeng, R. *et al.* (2003) Cold-active serine alkaline protease from the psychrophilic bacterium *Pseudomonas* strain DY-A: enzyme purification and characterization. *Extremophiles* 7: 335-337.

Chapter 7

Activation thermodynamics of LtrK stability and activity

7.1 Abstract

LtrK is an enzyme with optimum kinase and phosphatase activity at 10 °C and low thermal stability (inactivated at 30 °C). Activation thermodynamics showed that irreversible unfolding of the protein is achieved enthalpically through a decrease in activation enthalpy (ΔH^\ddagger). Additionally, thermodynamic study on LtrK kinase activity indicated that the efficient catalytic activity at lower temperatures (0 °C and 10 °C) is also achieved by reducing ΔH^\ddagger required to reach the activated transition-state (TS^\ddagger) from the ground state. The catalytic efficiency of LtrK was ~1.2 fold higher at 10 °C compared to 0 °C. Overall the data indicated that LtrK has evolved mechanisms to achieve high catalytic activity at low temperatures and thermolability by reducing enthalpy of activation.

7.2 Introduction

Temperature plays a major role in the function of enzymes as all biochemical reactions are strongly temperature dependent (Siddiqui, 2015). For example, the reaction rate of an enzyme can be reduced from 30 to 80 fold when the temperature is decreased from 37 °C to 0 °C (Lonhienne *et al.*, 2000). However, psychrophilic enzymes have evolved strategies to overcome the temperature dependence by synthesizing heat-labile flexible enzymes that possess high specific activity and catalytic efficiency at low temperatures (Gerday *et al.*, 1997; Lonhienne *et al.*, 2000; Siddiqui and Cavicchioli, 2006; Siddiqui, 2015). The most common adaptive feature of cold-adapted enzymes is the temperature independent reaction rate (k_{cat}) and majority of the psychrophilic enzymes achieve this by reducing the activation energy required to reach the activated state from ground state (Siddiqui *et al.*, 2013; Siddiqui, 2016). For example, the activation energy of thermophilic α -amylase is 70 kJ mol⁻¹ whereas for the psychrophilic α -amylase, it is 35 kJ mol⁻¹ and this allows the psychrophilic α -amylase to enhance k_{cat} by 21 fold at 10 °C (D'Amico *et al.*, 2003a). High k_{cat} at low temperature is also achieved by the flexible structure and associated low stability of psychrophilic enzymes which is referred to as activity-stability trade-off (Siddiqui and Cavicchioli, 2006; Siddiqui, 2016). Additionally, to aid substrate binding at low energy cost, the active site of the psychrophilic enzymes tends to be larger and more accessible which result in lower binding affinity for substrate (higher K_m) than that of their thermophilic counterparts (Siddiqui and Cavicchioli, 2006; Siddiqui *et al.*, 2013; Siddiqui, 2016).

The kinase, LtrK of *M. burtonii* exhibited very high kinase activity at temperatures between 0-10 °C indicating the cold-adaptive nature of the protein (described in Chapter 6). LtrK unfolds irreversibly and the unfolding can be initiated at

temperatures as low as ~20 °C which indicates the low thermal stability of the protein. Therefore, like other psychrophilic enzymes, specific properties must have been evolved in LtrK to attain high catalytic efficiency at low temperatures. To investigate that, the activation thermodynamic parameters regarding thermal unfolding and kinase activity of the enzyme were determined. In addition, the kinetic parameters of LtrK autophosphorylation reaction including catalysis rate (k_{cat}) and substrate binding affinity (K_m) were determined at 0 °C and 10 °C (T_{opt}).

7.3 Experimental procedures

7.3.1 Thermodynamic activation parameters for LtrK stability

From the DSC data at three different scan rates (described in Chapter 6, Fig. 6.5A) the rate constant of the irreversible thermal inactivation (k_d) at a given temperature was determined as described in Chapter 3 section 3.3.10. The Arrhenius plots were generated using the data for three different scan rates (1, 0.2 and 0.1 °C min⁻¹) by plotting natural log (ln) of k_d vs. the inverse of absolute temperature (1/T). The slope (m) of the best linear fit line to the data was -31.1, -29.4 and -32.1 for scan rate of 0.1, 0.2 and 1 °C min⁻¹ respectively. The average m (= -30.9) was used to calculate the activation energy (E_a) as $m \times R$, where R is the universal gas constant with a value of 8.314 J mol⁻¹ K⁻¹ (Siddiqui and Cavicchioli, 2006). The thermodynamic data were calculated from the transition state theory (Siddiqui *et al.*, 2002; Siddiqui and Cavicchioli, 2006) using following equations:

$$\Delta H^\ddagger \text{ (enthalpy of activation)} = E_a - RT \quad (1)$$

$$\Delta G^\ddagger \text{ (free energy of activation)} = -RT \ln ((k_d \times h) / (K_B \times T)) \quad (2)$$

$$\Delta S^{\#} \text{ (entropy of activation)} = (\Delta H^{\#} - \Delta G^{\#}) / T \quad (3)$$

Where T = temperature, k_d = rate constant for thermal unfolding, h = Planck constant (6.63×10^{-34} J s), K_B = Boltzman Constant (1.38×10^{-23} J K⁻¹), and R = universal gas constant.

7.3.2 Kinetic parameters for LtrK autophosphorylation

LtrK autophosphorylation activity was determined in the presence of increasing substrate, [γ -³²P]-ATP, concentration at 0 °C and 10 °C. LtrK (0.012 nmol) was incubated with various concentrations of [γ -³²P]-ATP in 10 μ l of P buffer for 10 min at 0 °C and 10 °C and the reaction was stopped by adding 4 μ l of sample buffer followed by heating at 95 °C for 3 min. The samples were subjected to SDS-PAGE and the protein bands were visualized by staining with Coomassie blue solution followed by destaining (described in Chapter 3, section 3.2.5). The radioactivity associated with each protein band was measured using liquid scintillation counter (Tri-Carb 3110 TR, PerkinElmer). To prepare the sample for the scintillation counting, the protein band was excised using a sharp scalpel, dipped into a scintillation vial containing 5 ml of liquid scintillation cocktail (Ultima Gold MV, PerkinElmer) followed by radioactivity measurement in counts per minute, CPM in the scintillation counter. As a control, the background radioactivity was measured by subjecting a small piece of SDS-polyacrylamide gel for the scintillation counting.

To convert the raw CPM to specific radioactivity (curie, Ci), a standard curve was generated using CPM data for a series of known [γ -³²P]-ATP concentrations. A specific amount of [γ -³²P]-ATP (0.01, 0.05, 0.1, 0.25, 0.5, 0.75 and 1 nCi) was added to 5 ml of Ultima Gold MV liquid scintillation cocktail as described for LtrK samples, the

radioactivity of each sample was measured (CPM) and the standard curve of CPM vs. ATP concentration was generated. Using the standard curve, CPM (measured in the experiment), were converted to μCi . The calculated μCi was subsequently converted to nmole by dividing with the specific activity of ^{32}P . The specific activity of ^{32}P on the day of experiment was $3000 \text{ Ci mmol}^{-1}$.

The calculated amount (nmol) was divided by the duration of the reaction (10 min) to get the reaction rate (V) in nmol min^{-1} . The experiment was carried out in triplicate, the data (reaction rate vs. substrate concentrations) were plotted using the Enzyme Kinetics Module1.1 linked to SigmaPlot 8.0 and fitted to the Michaelis-Menten equation from which the maximum reaction rate (V_{max}) and the affinity constant (K_m) at 0 and 10 °C were determined. The turnover number, k_{cat} was calculated as $V_{\text{max}} / [\text{E}]$; where, [E] is the enzyme concentration (0.012 nM) used in the reaction and the catalytic efficiency was calculated as k_{cat}/K_m .

7.3.3 Thermodynamic activation parameters for LtrK autophosphorylation

To calculate the activation energy required for autophosphorylation reaction, Arrhenius plot was constructed by performing the autophosphorylation assay from 0 °C to 10 °C (T_{opt}) with 2 °C interval. One μg of LtrK was incubated with 1 μCi of $[\gamma\text{-}^{32}\text{P}]\text{-ATP}$ in 10 μl P buffer for 10 min and the reaction was stopped as described above. Samples were subjected to SDS-PAGE and the autoradiograph was generated as described in Chapter 5, section 5.3.4. The activity was measured as band intensity using the phosphorimager. The experiment was carried out in triplicate and the mean activity was used to construct an Arrhenius plot, log of autophosphorylation activity vs $1/T$ (Siddiqui *et al.*, 2002; Cavicchioli *et al.*, 2006). The activation energy (E_a) was

calculated as described above. The thermodynamic data were calculated from the equations 1, 2 (by using k_{cat} value in place of k_d), and 3.

7.4 Results

7.4.1 Activation Thermodynamics of LtrK stability

The DSC analyses showed that the thermally induced unfolding of LtrK is irreversible and scan rate dependent (described in Chapter 6, Fig. 6.5A). From the data of LtrK unfolding at different scan rate (0.1, 0.2 and 1 °C min⁻¹), the rate constants of unfolding (k_d) were calculated and Arrhenius plot was generated (Fig. 7.1).

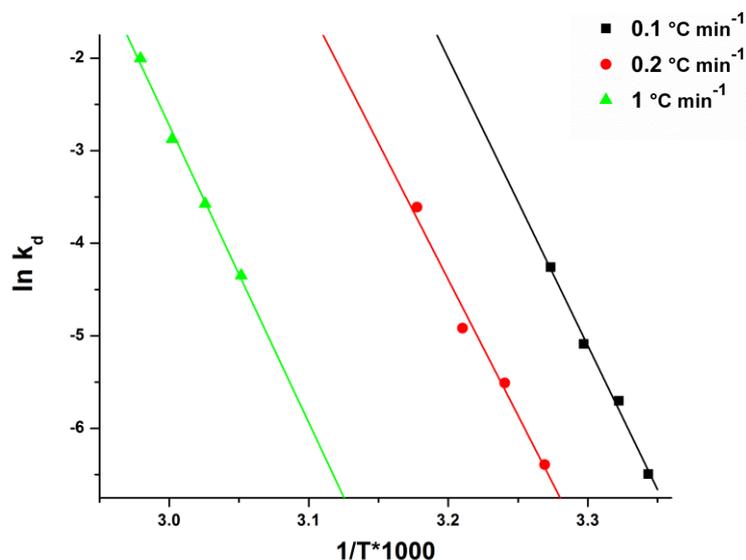


Figure 7.1 Arrhenius plot for the determination of activation energy for LtrK denaturation

The green, red and black line represents the linear fit to the data obtained at 1, 0.2 and 0.1 °C min⁻¹ scan rate respectively.

The activation energy (E_a) was calculated as ~ 256.6 kJ mol⁻¹ K⁻¹ and using the E_a the thermodynamic activation parameters for LtrK irreversible denaturation were

determined over a temperature range of 26 to 42 °C (Table 7.1) which showed decreased value of activation enthalpy (ΔH^\ddagger), activation entropy (ΔS^\ddagger) and free energy of activation (ΔG^\ddagger) at 41.7 °C compared to 26.1°C.

Table 7.1 Thermodynamic activation parameters for LtrK thermal unfolding

Temperature (°C)	ΔH^\ddagger (kJ mol ⁻¹)	ΔG^\ddagger (kJ mol ⁻¹)	ΔS^\ddagger (J mol ⁻¹ K ⁻¹)
26.1	254.1	90.5	547
30.3	254.1	87.3	550
41.7	253.9	86.4	532

7.4.2 Thermodynamics of LtrK activity

7.4.2.1 Kinetic parameters for LtrK autophosphorylation

Kinetic parameters for LtrK autophosphorylation were determined both at 0 °C and 10 °C (T_{opt}) using liquid scintillation counting method. The standard curve of CPM vs. ATP concentration (Fig. 7.2A) was used to calculate the reaction rate, V (nmol min⁻¹) of the LtrK kinase reaction. The enzyme activity with ATP showed typical Michaelis-Menten (MM) kinetics both at 0 °C and 10 °C (Fig. 7.2BC).

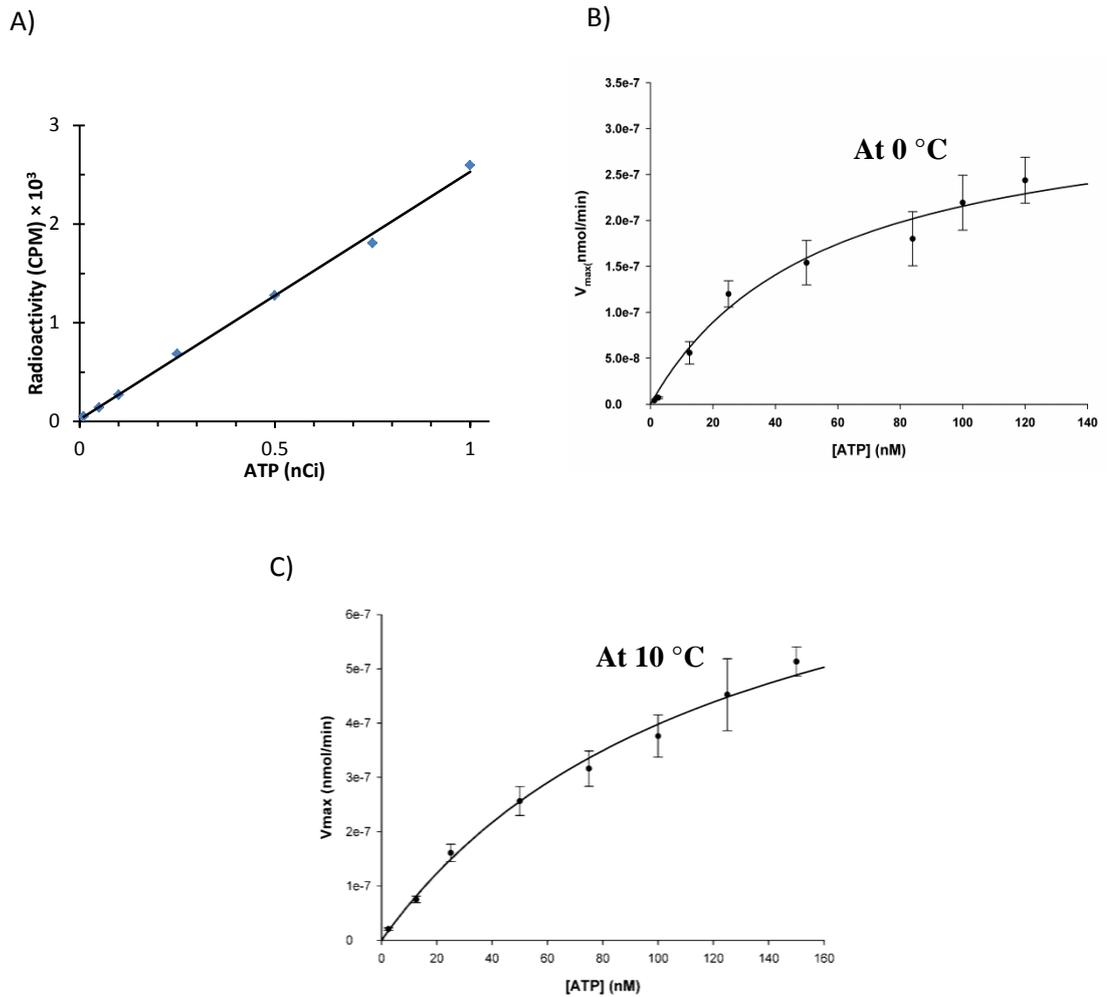


Figure 7.2 Michaelis-Menten kinetics for LtrK autophosphorylation

A) Standard curve, showing the radioactivity (CPM) of different known concentration of $[\gamma^{32}\text{P}]\text{-ATP}$, used to convert CPM value to reaction rate, V (nmol min^{-1}); B) MM plot showing relation between autophosphorylation rate and ATP concentration at 0 °C, and C) at 10 °C

The kinetic parameters including maximum reaction rate (V_{max}), substrate affinity (K_m), turnover number (k_{cat}) and catalytic efficiency (k_{cat}/K_m) determined from the MM plot are given in Table 7.2. All the parameter had higher values at 10 °C compared to 0 °C indicating more efficient catalyses at 10 °C.

Table 7.2 Kinetic parameters of LtrK autophosphorylation activity at 0 °C and 10 °C

Parameters	At 0 °C	At 10°C
V_{\max} (nmol min ⁻¹)	3.3×10^{-7}	9×10^{-7}
K_m (nM)	55	125
k_{cat} (s ⁻¹)	4.6×10^{-7}	1.2×10^{-6}
k_{cat}/K_m (nM ⁻¹ s ⁻¹)	8.4×10^{-9}	1×10^{-8}

7.4.2.2 Activation energy for LtrK autophosphorylation reaction

The activation energy, E_a , for LtrK autophosphorylation was calculated using Arrhenius plot (Fig. 7.3) as 31.3 kJ mol⁻¹.

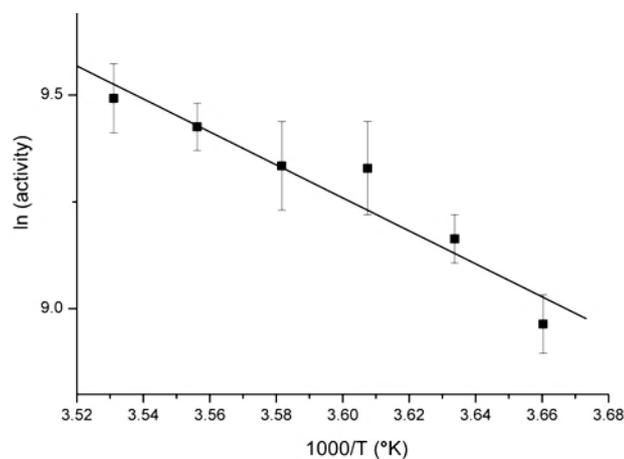


Figure 7.3 Arrhenius plot for the determination of activation energy of autophosphorylation

K=Kelvin, T= Temperature

Using the E_a value, change in activation enthalpy (ΔH^\ddagger), activation entropy (ΔS^\ddagger) and free energy of activation (ΔG^\ddagger) were calculated at 0 °C and 10 °C from equation 1, 2 and 3 (described in Experimental procedure) and the values are presented in Table 7.3.

Table 7.3 Thermodynamic activation parameters at 0 °C and 10 °C of LtrK activity

Parameters	at 0 °C	at 10 °C
ΔH^\ddagger (kJ mol ⁻¹)	29	29
ΔG^\ddagger (kJ mol ⁻¹)	100	101
ΔS^\ddagger (J mol ⁻¹ K ⁻¹)	-259	-255

The results showed that the change in activation entropy (ΔS^\ddagger) to reach the transition state (TS[‡]) from ground-state was negative and slightly increased at 10 °C along with free energy of activation (ΔG^\ddagger) while the activation enthalpy (ΔH^\ddagger) remained unchanged.

7.5 Discussion

The determination of activation thermodynamic parameters of the irreversibly unfolding proteins is based on the assumption that the native protein (F) is in equilibrium with partially unfolded protein in the transition state (TS[‡]) before unfolding irreversibly to completely unfolded protein (U) (Siddiqui *et al.*, 2002; D'Amico *et al.*, 2003b). In the following scheme K is the equilibrium constant between folded (F) and partially unfolded protein and k_d is the first order rate constant for the conversion of transition state to (TS[‡]) completely denatured state (U) (Siddiqui and Cavicchioli, 2006).



The kinetic stability of an enzyme, that reflects how rapidly an enzyme can unfold at a given temperature, can be measured from the enzyme denaturation data. During thermal denaturation of an enzyme, the magnitude of free energy change (ΔG^\ddagger) between

the folded state (F) and transition state (TS[#]) is used to measure the kinetic stability of the enzyme. The magnitude of ΔG^\ddagger reflects how much energy is required for the folded protein to form the partially unfolded TS[#]. The ΔG^\ddagger depends on ΔH^\ddagger and ΔS^\ddagger in accordance with equation 4 (Cavicchioli *et al.*, 2006; Siddiqui and Cavicchioli, 2006) where ΔH^\ddagger is a measure of heat change during unfolding while ΔS^\ddagger is a measure of disorder between the ground-state (folded enzyme) and activated transition-state, TS[#] (partially unfolded form). In order to increase the rate of thermal unfolding, ΔG^\ddagger is required to decrease and the reduction in ΔG^\ddagger can be caused either by decreasing ΔH^\ddagger or by increasing ΔS^\ddagger . Low ΔH^\ddagger implies a reduced number of noncovalent interactions that need to be broken to reach TS[#] and high ΔS^\ddagger implies a higher disordered TS[#] (Siddiqui and Cavicchioli, 2006).

$$\Delta G^\ddagger = \Delta H^\ddagger - T\Delta S^\ddagger \quad (4)$$

From the activation parameters (Table 7.1) it is evident that with increasing temperatures both ΔG^\ddagger and ΔH^\ddagger decreased reflecting that less energy is required to unfold the LtrK. Interestingly, ΔS^\ddagger decreased from 547 J mol⁻¹ K⁻¹ at 26 °C to 532 J mol⁻¹ K⁻¹ at 41.7 °C indicating that at higher temperature the partially unfolded TS[#] is less disordered (more compact) than the folded form. The more compact TS[#] (lower ΔS^\ddagger) relative to fully folded protein may arise indirectly due to the formation of cage-like ordered water molecules around hydrophobic residues which are exposed as a result of protein unfolding (Siddiqui *et al.*, 2002; Siddiqui *et al.*, 2010; Siddiqui, 2016). However, the data indicates the unfolding of LtrK is enthalpy driven as the decrease in ΔG^\ddagger is caused by the reduction in ΔH^\ddagger . Couple of enzyme's thermodynamic activation parameters of unfolding have previously been studied in *M. burtonii*, *i.e.* EF2 and RNA

polymerase E/F and the analyses showed that the unfolding of these proteins is also controlled enthalpically (Siddiqui *et al.*, 2002; De Francisci *et al.*, 2011).

The kinetic parameters of LtrK autophosphorylation both at 0 °C and 10 °C were determined using the MM plot (Fig. 7.2). LtrK showed maximum kinase activity at 10 °C (described in Chapter 6) and the catalytic efficiency was found to be ~1.2 fold higher at 10 °C compared to 0 °C. It has been reported that, cold-active enzymes have evolved towards improved active site dynamics that bind substrate less firmly in the binding site giving rise to less affinity (K_m) between the substrate and enzyme (Siddiqui and Cavicchioli, 2006; Struvay and Feller, 2012). In addition, the cold active enzymes have improved k_{cat} at the expense of K_m leading to increased k_{cat}/K_m value (Struvay and Feller, 2012). Similar trend was observed in LtrK. The catalytic efficiency at 10 °C increased by increasing the turnover rate (higher k_{cat}) while decreasing the substrate affinity (higher K_m) (Table 7.2).

Arrhenius plot was constructed to calculate the activation energy required by LtrK to reach the transition state during autophosphorylation reaction. In a typical enzymatic reaction, enzyme [E] combines with the substrate [S] to form an enzyme-substrate complex [ES] which requires energy (activation-energy) to form the transition-state (TS^\ddagger) (Feller, 2003; Siddiqui and Cavicchioli, 2006). Finally, TS^\ddagger decomposes into the product [P] and the free enzyme is generated.



The activation energy, E_a is the minimum energy requirement that must be met to reach the TS^\ddagger and the higher the activation energy, the harder it is for a reaction to occur and vice versa (Fig. 7.4).

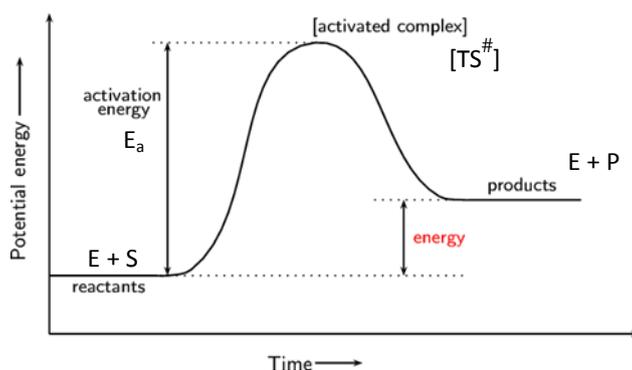


Figure 7.4 Activation energy profile of a reaction

Enzyme (E) combines with the substrate (S) and by using activation energy (E_a) forms an enzyme-substrate complex in the transition state $[TS^\ddagger]$. And finally product (P) and free enzyme (E) are generated.

Enzymes adapted to low-temperatures decrease E_a (or ΔH^\ddagger) in order to attain higher intrinsic activities relative to mesophilic and thermophilic homologues (Siddiqui and Cavicchioli, 2006). The ΔH^\ddagger is low for LtrK autophosphorylation reaction (29 kJ mol^{-1} at $0 \text{ }^\circ\text{C}$ and $10 \text{ }^\circ\text{C}$); even lower than the ΔH^\ddagger required for GTP hydrolysis in psychrophilic enzyme, EF2 of *M. burtonii*, which is 76.0 kJ mol^{-1} at $10 \text{ }^\circ\text{C}$ (Siddiqui *et al.*, 2002; Siddiqui and Cavicchioli, 2006). This further implies the cold-adaptive nature of LtrK.

The activation entropy (ΔS^\ddagger) was negative for the autophosphorylation reaction which implies that the TS^\ddagger is more ordered as compared to the ground-state (ES) (Table 7.3). This indicates that the phosphorylation reaction is enthalpy-driven as negative ΔS^\ddagger will prevent the reaction from proceeding to the formation of TS^\ddagger . The low ΔH^\ddagger is also likely to generate a more flexible structure of the active site of the enzyme facilitating the substrate binding (Siddiqui and Cavicchioli, 2006). The free energy barrier, ΔG^\ddagger increased by 1.4 kJ mol^{-1} from $0 \text{ }^\circ\text{C}$ to $10 \text{ }^\circ\text{C}$ (Table 7.3) reflecting the conformational

changes at 10 °C within the catalytic site of the protein that may require higher energy contribution to form the transition state.

In summary, the thermodynamic activation parameters of LtrK stability and activity revealed that the enzyme contains typical characteristics of a cold-active enzyme including enthalpy driven unfolding and autophosphorylation activity and higher catalysis rate at cold temperatures at the expense of low substrate binding affinity.

7.6 References

Cavicchioli, R. *et al.* (2006) Proteins from Psychrophiles. *Methods Microbiol.* 35: 395-436.

D'Amico, S. *et al.* (2003a) Temperature adaptation of proteins: engineering mesophilic-like activity and stability in a cold-adapted α -amylase. *J. Mol. Biol.* 332: 981-988.

D'Amico, S. *et al.* (2003b) Activity-stability relationships in extremophilic enzymes. *J. Biol. Chem.* 278: 7891-7896.

De Francisci, D. *et al.* (2011) The RNA polymerase subunits E/F from the Antarctic archaeon *Methanococcoides burtonii* bind to specific species of mRNA. *Environ. Microbiol.* 13: 2039-2055.

Feller, G. (2003) Molecular adaptations to cold in psychrophilic enzymes. *Cell Mol. Life Sci.* 60: 648-662.

Gerday, C. *et al.* (1997) Psychrophilic enzymes: a thermodynamic challenge. *Biochim. Biophys. Acta-Protein Structure and Molecular Enzymology* 1342: 119-131.

Lonhienne, T. *et al.* (2000) Psychrophilic enzymes: revisiting the thermodynamic parameters of activation may explain local flexibility. *Biochim. Biophys. Acta-Protein Structure and Molecular Enzymology* 1543: 1-10.

Siddiqui, K. *et al.* (2002) Thermodynamic activation properties of elongation factor 2 (EF-2) proteins from psychrotolerant and thermophilic Archaea. *Extremophiles* 6: 143-150.

Siddiqui, K.S. (2015) Some like it hot, some like it cold: Temperature dependent biotechnological applications and improvements in extremophilic enzymes. *Biotechnol. Adv.* 33: 1912-1922.

Siddiqui, K.S. (2016) Defying activity-stability trade-off in enzymes: taking advantage of entropy to enhance activity and thermostability *Crit. Rev. Biotechnol.* doi: 10.3109/07388551.

Siddiqui, K.S., and Cavicchioli, R. (2006) Cold-adapted enzymes. *Annu. Rev. Biochem.* 75: 403-433.

Siddiqui, K.S. *et al.* (2013) Psychrophiles. *Annu Rev. Earth Planet Sci.* 41: 87-115.

Siddiqui, K.S. *et al.* (2010) A chemically modified α -amylase with a molten-globule state has entropically driven enhanced thermal stability. *Protein Eng. Des. Sel.* 23: 769-780.

Struvay, C., and Feller, G. (2012) Optimization to low temperature activity in psychrophilic enzymes. *Int. J. Mol. Sci.* 13: 11643-11665.

Chapter 8

Conclusions and Future Work

8.1 General discussion

The biochemical characterization of a TCS, LtrK/LtrR from *M. burtonii* and its response to changes in temperature are described in this thesis. The study has revealed some striking features including temperature sensing which are summarized in the following sections.

8.1.1 Unique organization of HTH domain in the DNA-binding RRs of methanogens

It was surprising that the positioning of HTH domain in RRs from methanogens is opposite to that of RRs from bacteria. The HTH domain is located at the C-terminal region in all bacterial RRs and the majority (~90 %) of the DNA-binding RRs encoded in the genomes of haloarchaea whereas it is located at N-terminus in all RRs identified in the genomes of methanogens (described in Chapter 2). This exceptional domain organization begs the question, is there any evolutionary significance of this arrangement? Before searching for an answer to this question it is important to get an insight into the selection of domain combination in a protein during the course of its evolution. The structural domains that perform distinct functions have been shuffled during evolution. These domains can interact with each other to create proteins with different combinations of domains. Nature selects only those domain combinations within a protein that provides a benefit to the survival of an organism (Bourret, 2006). Interestingly, it has been found that if a protein is formed by a combination of domains then these domains typically arrange in one sequential order (Bashton and Chothia, 2002). For example, if two domains of two subfamilies, A and B form a protein it can occur either in AB or BA order and only in ~2 % of the cases they

can occur concurrently in both AB and BA order (Bashton and Chothia, 2002). If the REC domain of RRs is represented as A and DNA-binding domain is represented as B then the bacterial RRs are in order AB, haloarchaeal RRs have both AB (~90 %) and BA (~10 %) order and methanogenic RRs are only found in order BA. The typically conserved domain order suggests that the recombination of the domains has occurred once during the evolution resulting in a specific domain order and that this superfamily arose from the subsequent duplication and divergence of that domain combination (Bashton and Chothia, 2002). Therefore, it is possible that the N-terminal positioning of HTH domain in methanogens is simply a result of a single recombination event that randomly chose the DNA-binding output domain at the N-terminus during the evolution which has since been conserved in the genome sequence of methanogens by gene duplication.

8.1.2 Presence of periplasmic CHASE domain in LtrK

Bioinformatics analyses have identified the presence of a small ligand binding CHASE domain in the periplasmic region of LtrK (described in Chapter 2). The genomes of unicellular microorganisms do not appear to have excess coding capacity (Bourret, 2006) which suggests that if a domain is expressed within a protein it must have some functional significance. If temperature is the primary environmental factor sensed by LtrK, the cytoplasmic domain has already shown the capacity to respond to temperature by means of irreversible thermal unfolding (described in Chapter 6). Thus, the presence of CHASE domain indicates the possibility of LtrK to respond to signals other than temperature. A limited number of studies have shown that few SKs can receive various stimuli through its various domains and hence can act as a multi-sensor, *i.e.* PhoQ of *E. tarda* (Chakraborty *et*

al., 2010), VirA of *A. tumefaciens* (Chang and Winans, 1992). The periplasmic sensor domain of PhoQ can sense temperature and bind Mg^{2+} and antimicrobial peptides at the same time (Chakraborty *et al.*, 2010) whereas in VirA the periplasmic domain interacts with a galactose binding protein, ChvE; the linker region senses the phenolic compounds, *i.e.* acetosyringone, and acidity (Chang and Winans, 1992) and the kinase domain functions as a thermosensor (Jin *et al.*, 1993). This allows PhoQ to respond to changes in temperature, Mg^{2+} and antimicrobial peptide concentration and VirA to sugar concentration, pH and temperature simultaneously. So, it will not be surprising if LtrK functions as a multi-sensor and is capable of sensing stimuli other than temperature.

8.1.3 Plausible role of membrane in regulating the function of transmembrane

LtrK

Though the study on LtrK activity showed that its cytoplasmic domain is inherently thermosensitive and sufficient for kinase and phosphatase activities, the cellular regulation of kinase versus phosphatase activity could not be inferred from the present data. Previously, it was reported that lipid environments play a key role in dictating the signaling state of the sensor (Mariana and Diego, 2013). In addition, few studies on thermosensing SKs have found that structural change in the cellular membrane can play a regulatory role controlling the activity of SKs. For example, change in membrane structure in response to temperature, functions to switch the DesK activity between kinase and phosphatase competent forms in *B. subtilis* (Albanesi *et al.*, 2004; Mendoza, 2014). Similarly, the membrane state is assumed to play a vital role in regulating the kinase versus phosphatase activity of Hik33 in cyanobacteria *Synechocystis* (Suzuki *et al.*, 2000; Mendoza, 2014).

Therefore, it is possible that LtrK activity is regulated by the membrane structure and lipid composition similar to DesK or Hik33. Archaeal lipids are structurally different to bacterial lipids and contain a glycerol-1-phosphate backbone attached to ether-linked isoprenoid moiety, whereas bacterial lipids contain ester-linked fatty acids attached to a glycerol-3-phosphate backbone. However, similar to bacteria, lipid unsaturation increases at low temperature in *M. burtonii* (Nichols *et al.*, 2004). For example, the unsaturated lipids archaeol phosphatidylinositol and hydroxyarchaeol phosphatidylinositol increase from ~14 % during growth at 23 °C to ~28 % at 4 °C (Nichols *et al.*, 2004). The mechanism of producing unsaturated fatty acids in *M. burtonii* does not involve a desaturase, as in the case of *B. subtilis* (Albanesi *et al.*, 2009; Mendoza, 2014) and *Synechocystis* sp. (Suzuki *et al.*, 2000; Mendoza, 2014). However, as a transmembrane sensor LtrK may still be responsive to membrane structural changes and regulate the genes (*i.e.* geranylgeranyl reductase) involved in selective membrane saturation (Nichols *et al.*, 2004). Concurrently, the state of membrane fluidity arising from changes in lipid saturation could possibly regulate kinase versus phosphatase activities of LtrK.

8.1.4 LtrK is only the third biochemically studied thermosensing SK

Membrane receptors functioning as a thermosensor was first reported in *E. coli* in 1976 (Maeda *et al.*, 1976) and the existence of TCS was recognized in mid 1980s (Nixon *et al.*, 1986) and yet surprisingly only few studies have been reported on the molecular basis of thermosensing by TCSs. The biochemical analyses data for thermosensing SKs are limited to *B. subtilis* DesK/DesR that regulates the membrane fluidity by transcriptional regulation of desaturase (*des*) gene (Aguilar *et al.*, 2001) and *A. tumefaciens* VirA/VirG

that induce tumour formation by virulence (*vir*) gene regulation (Jin *et al.*, 1993) (Fig. 8.1). This dissertation presents the third biochemical analyses report for temperature responsive TCS, LtrK/LtrR of *M. burtonii* (Fig. 8.1).

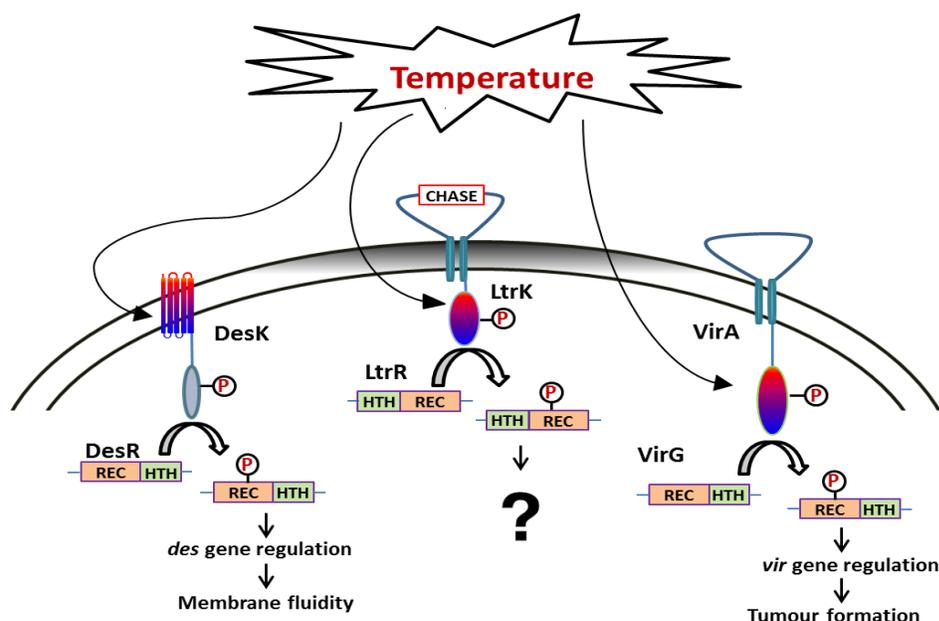


Figure 8.1 Schematic representation of thermosensing by SKs from Archaea and Bacteria.

The figure represents the biochemically analyzed temperature-responsive SKs with their structural features. The temperature sensing domain is highlighted with blue-red gradient color. **DesK/DesR** from *B. subtilis*. DesK has five transmembrane domains (TMD1-TMD5) and senses temperature through its TMDs, thereby regulating the phosphotransfer activities of DesK with DesR. DesR regulates the expression of *des* gene which functions to control membrane fluidity (Aguilar *et al.*, 2001). **LtrK/LtrR** from *M. burtonii*. LtrK has two TMDs with a long extracellular CHASE containing domain but unlike DesK it senses temperature directly through its cytoplasmic domain, which controls its phosphorylation activity with LtrR. The extracellular CHASE domain is also likely to detect an environmental signal other than temperature. LtrR has an N-terminal HTH domain, typical of *Archaea*, which is in the opposite orientation to HTH domains from most *Bacteria*. Also typical of *Archaea*, *M. burtonii* possesses ether-linked isoprenoid lipids attached to a glycerol-1-phosphate backbone (membrane shaded grey), whereas *Bacteria* contain ester-linked fatty acids attached to a glycerol-3-phosphate backbone (shaded white). *M. burtonii* lipid unsaturation increases at low growth temperatures (Nichols *et al.*, 2004). The genes regulated by LtrR have not been determined. **VirA/VirG** from *A. tumefaciens*. Similar to LtrK, VirA has two TMDs and senses temperature via its cytoplasmic domain (Jin *et al.*, 1993). Distinct structural domains of VirA are also capable of directly or indirectly detecting sugars, phenolic compounds and acidity levels (Chang and Winans, 1992). VirG regulates *vir* genes involved in plant tumour formation (Jin *et al.*, 1993).

Few studies carried out to date reflect experimental limitations associated with analysing the thermosensing TCSs. In this regard the major limitation is the lack of experimental procedures to purify the transmembrane SKs as full length soluble proteins. As extracellular sensor domain and TMDs can play a critical role in sensing temperature stimuli, understanding the molecular mechanism of temperature response may require comprehensive *in vitro* analyses of the full length SKs. However sometimes, even the full length protein may not be enough to study the thermal regulation because the protein may require to be tethered to the membrane, as in case of DesK of *B. subtilis*. In that case *in vivo* analyses are necessary that requires genetic modification to be performed in the organism. But for many microorganisms, no molecular tool is available for genetic manipulation, which makes it impossible to perform *in vivo* analyses including the overexpression of native proteins, introduction of mutations and knock out genes. Fortunately, new molecular tools are emerging, *e.g.* techniques to purify membrane proteins, advanced gene transformation protocol in new organisms (*e.g.* psychrophilic bacteria such as *P. haloplanktis*). This would possibly assist in performing comprehensive study on thermosensing TCSs in future.

8.2 Future work

8.2.1 Identification of genes regulated by LtrK/LtrR TCS

Guided by the method (phosphorylation of LtrR by LtrK) established in this study, the identification of the genes regulated by LtrR could be attempted in future. This thesis has already reported an unsuccessful attempt of capturing LtrR bound DNA using the DNA

binding assay (described in Chapter 4). Though the attempt was not successful, it assists in proposing some suggestions that could be useful to identify LtrR regulated DNA in future.

➤ The success of an experiment greatly depends on its design and the protocol of DNA binding assay used in this study included multiple steps which made the experimental setup problematic. Due to high phosphatase activity of LtrK in presence of ATP, the protocol involved steps including phosphorylation of GST-LtrK in the glutathione column, washing free ATP, passing LtrR through the column, binding of LtrR-P in another affinity column, passing of fragmented DNA, thorough washing of free DNA and finally isolate the DNA from protein bound DNA complex. Phosphatase incompetent LtrK would make the protocol simpler as there would be no necessity to perform the phosphorylation of LtrK in glutathione column and wash off the free ATP. By using phosphatase incompetent LtrK, phosphorylation of LtrR could be performed by incubating with LtrK and ATP in a tube followed by removal of LtrK and ATP by binding the phosphorylated LtrR to an affinity column. Studies have shown that mutation at the polar residue adjacent to the universally conserved histidine residue in few SKs disrupts the phosphatase activity while the kinase activity remains unchanged (Huynh *et al.*, 2010; Willett and Kirby, 2012). The polar residue in LtrK adjacent to the conserved histidine residue is T367 (described in Chapter 2). Introduction of a mutation at T367 could be attempted to assess any difference in phosphatase activity of LtrK.

➤ In the unsuccessful attempt of the DNA-binding assay described in Chapter 4, the LtrR-DNA complex was eluted from the Co^{2+} charged affinity column using an elution buffer containing imidazole, and DNA was subsequently purified from the complex using

typical phenol:chloroform DNA extraction method. The steps of the phenol:chloroform DNA extraction protocol are prone to DNA loss. Moreover, phenol contamination in extracted DNA can greatly reduce the DNA quality. A recent study on the characterization of genomic DNA binding ability of transcription factor DnaA in *B. subtilis*, has reported the use of EDTA in elution buffer to elute protein-DNA complex from the affinity column followed by the purification of DNA from the complex using QiaQuick PCR purification kit (Qiagen) (Smith and Grossman, 2015). The use of EDTA in elution buffer might give advantage over imidazole as EDTA would chelate out the metal to which protein-DNA complex is bound, thereby recover majority of protein-DNA complex from the affinity column. Additionally purification of DNA using QiaQuick PCR purification kit would avoid the use of phenol and increase the DNA recovery with high quality.

➤ Many DNA binding RRs have the ability to bind to its own promoter sequence (Bijlsma and Groisman, 2003). Therefore, the DNA-binding ability of the phosphorylated LtrR could initially be checked by incubating with its promoter sequence followed by electrophoretic mobility shift assay on agarose gel. This would also allow the DNA binding conditions to be optimised including buffer composition, incubation period and DNA:protein weight ratio.

8.2.2 Analyses of the regulation of LtrK/LtrR in *M. burtonii*

The current study indicated the temperature-responsive activity of LtrK but further validation is necessary to substantiate the finding. For this purpose, *in vivo* analyses are required to be performed. *M. burtonii* is not cultivatable on solid medium and therefore, genetic manipulation cannot be performed. However, in a very recent study gene

transformation method has been developed for a psychrophilic haloarchaea, *Halorubrum lacusprofundi*, that is isolated from hypersaline Deep Lake in Antarctica (Yan Liao, personal communication). Global proteomic analyses at low versus high temperature in *H. lacusprofundi* did not identify any SK differentially expressed at cold temperatures (Yan Liao, personal communication). As LtrK/LtrR has been found to overexpress at low growth temperature compared to high temperature in *M. burtonii* (Goodchild *et al.*, 2004), expression of this TCS in *H. lacusprofundi* would promote the investigation on the cellular factors that may regulate the high activity of this TCS at low temperature. The function of CHASE domain in LtrK and the influence of cell membrane on LtrK activity could also be investigated. However, *H. lacusprofundi* is extremely halophilic and grows in a hypersaline media which may affect the regulation of expressed LtrK/LtrR. Therefore, analyses of the LtrK/LtrR regulation in a methanogen would be more appropriate. Gene transformation method has already been developed for few methanogens, *e.g.* polyethylene glycol (PEG) mediated gene transformation protocol was developed for *Methanococcus maripaludis* (Tumbula *et al.*, 1994) and a liposome delivery transformation method was developed for *Methanosarcina acetovorans* (Metcalf *et al.*, 1997). LtrK/LtrR TCS could be attempted to express in these methanogens. However in order to confirm the *in vivo* regulation of the LtrK/LtrR, gene manipulation methods are required to be developed for *M. burtonii*.

8.2.3 Possible biotechnological application of LtrK

Due to the temperature regulated activity of LtrK (*i.e.* optimum activity at 10 °C, greatly reduced activity at room temperature, 25 °C and inactive at ≥ 30 °C), the enzyme might have biotechnological applications. Typically, psychrophilic enzymes can provide

economic benefit by being more productive at low temperatures than their mesophilic and thermophilic counterparts. For this reason, many cold adapted enzymes (*e.g.* psychrophilic lipase, protease) are employed in various applications including food and feed industry and cell and molecular biology (Cavicchioli *et al.*, 2011). LtrK might also be useful, especially in a multi-step reaction procedure that requires kinase. As the enzyme has such low T_{opt} , the kinase reaction could be carried out at low temperatures (*i.e.* 10 °C) and then switched off simply by increasing the temperature to moderate temperatures (*i.e.* 30°C). This would allow the reaction to proceed to the next step without requiring separation of the enzyme (LtrK).

8.3 Concluding remarks

The work described in this thesis has characterised a functioning TCS, LtrK/LtrR of *M. burtonii* that has temperature-responsive activity and might play a significant role in cold adaptation by regulating genes in response to growth temperature. The study has established a foundation for the future investigation on thermosensing mechanism and the complete gene regulatory pathway governed by the LtrK/LtrR system. Considering the fact that this is the first report on temperature sensing SKs on psychrophiles (both bacteria and archaea) and second only on archaeal TCS, the present study would eventually help to expand the thermosensing TCS research field in future.

8.4 References

Aguilar, P.S. *et al.* (2001) Molecular basis of thermosensing: a two-component signal transduction thermometer in *Bacillus subtilis*. *EMBO J.* 20: 1681-1691.

- Albanesi, D. *et al.* (2004) The membrane fluidity sensor DesK of *Bacillus subtilis* controls the signal decay of its cognate response regulator. *J. Bacteriol.* 186: 2655-2663.
- Albanesi, D. *et al.* (2009) Structural plasticity and catalysis regulation of a thermosensor histidine kinase. *Proc. Natl. Acad. Sci. USA.* 106: 16185-16190.
- Bashton, M., and Chothia, C. (2002) The geometry of domain combination in proteins. *J. Mol. Biol.* 315: 927-939.
- Bijlsma, J.J., and Groisman, E.A. (2003) Making informed decisions: regulatory interactions between two-component systems. *Trends Microbiol.* 11: 359-366.
- Bourret, R.B. (2006) Census of prokaryotic senses. *J. Bacteriol.* 188: 4165-4168.
- Cavicchioli, R. *et al.* (2011) Biotechnological uses of enzymes from psychrophiles. *Microb. Biotechnol.* 4: 449-460.
- Chakraborty, S. *et al.* (2010) Temperature and Mg²⁺ sensing by a novel PhoP-PhoQ two-component system for regulation of virulence in *Edwardsiella tarda*. *J. Biol. Chem.* 285: 38876-38888.
- Chang, C.-H., and Winans, S.C. (1992) Functional roles assigned to the periplasmic, linker, and receiver domains of the *Agrobacterium tumefaciens* VirA protein. *J. Bacteriol.* 174: 7033-7039.
- Huynh, T.N. *et al.* (2010) Conserved mechanism for sensor phosphatase control of two-component signaling revealed in the nitrate sensor NarX. *Proc. Natl. Acad. Sci. USA.* 107: 21140-21145.
- Jin, S. *et al.* (1993) The regulatory VirA protein of *Agrobacterium tumefaciens* does not function at elevated temperatures. *J. Bacteriol.* 175: 6830-6835.
- Maeda, K. *et al.* (1976) Effect of temperature on motility and chemotaxis of *Escherichia coli*. *J. Bacteriol.* 127: 1039-1046.
- Mariana, M., and Diego, d.M. (2013) Regulation of *Bacillus subtilis* DesK thermosensor by lipids. *Biochem. J.* 451: 269-275.
- Médigue, C. *et al.* (2005) Coping with cold: the genome of the versatile marine Antarctica bacterium *Pseudoalteromonas haloplanktis* TAC125. *Genome Res.* 15: 1325-1335.
- Mendoza, D.d. (2014) Temperature sensing by membranes. *Annu. Rev. Microbiol.* 68: 101-116.
- Metcalf, W.W., *et al.* (1997) A genetic system for Archaea of the genus *Methanosarcina*: liposome-mediated transformation and construction of shuttle vectors. *Proc. Natl. Acad. Sci. USA.* 94: 2626-2631.

Nichols, D.S. *et al.* (2004) Cold adaptation in the Antarctic archaeon *Methanococoides burtonii* involves membrane lipid unsaturation. *J. Bacteriol.* 186: 8508-8515.

Nixon, B.T. *et al.* (1986) Two-component regulatory systems responsive to environmental stimuli share strongly conserved domains with the nitrogen assimilation regulatory genes *ntxB* and *ntxC*. *Proc. Natl. Acad. Sci. USA.* 83: 7850-7854.

Smith, J.L., and Grossman, A.D. (2015) In vitro whole genome dna binding analysis of the bacterial replication initiator and transcription factor DnaA. *PLoS Genet* 11: e1005258.

Suzuki, I. *et al.* (2000) The pathway for perception and transduction of low-temperature signals in *Synechocystis*. *EMBO J.* 19: 1327-1334.

Tumbula, D.L. *et al.* (1994) Transformation of *Methanococcus maripaludis* and identification of a Pst I-like restriction system. *FEMS Microbiol. Lett.* 121: 309-314.

Tutino, M.L. *et al.* (2002) Secretion of α -amylase from *Pseudoalteromonas haloplanktis* TAB23: two different pathways in different hosts. *J. Bacteriol.* 184: 5814-5817.

Willett, J.W., and Kirby, J.R. (2012) Genetic and biochemical dissection of a HisKA domain identifies residues required exclusively for kinase and phosphatase activities. *PLoS Genet* 8: e1003084.

Appendix 1: Protein sequences used for phylogenetic tree construction

A1.1 Bacterial and archaeal SK protein sequences to construct phylogenetic tree for LtrK (in FASTA format)

>Pontibacter actiniarum

SINNELRRKSMELRRATEQLRSANERLKLLDELKDDFLSTVTHELRTPLTSIRALSEILYDNP
PEMERSDQEHFLHTIVKESERLTRLITHVLDLERFESGRQKLNLEPVQLKEVIQESVEALAQLVQEKQIE
LVVDVQHSLPSINADRDRLIQVLVNLISNAIKFCSPENGRIVSAYYIDGNVKVNVVDNGKGDIAE
SQKQVFDKQYQAKHQNIRKPEGSGLGLAISKKIIESHQGRLWVESEPGKGARFSFTLP

>Skermanella aerolata

QLEQKSAELERTSVALRAANEKLELDRLKDDFLSTVTHELRTPLTSIRAF AEILHDNPDI EPDQA
REFLHLIIKESERLTRLINQVLDMAKIEAGEIDWNIETVDLGEIVREATATTAQVFKDKGIALEAHI
PDTVPGVRGDHDRLTQVAMNLLSNAAKFTPGGSGRVVVSVEHSGDGVKVAVSDNGPGLTPND
AEIVFDKFRQVGNTMTDKPQGTGLGLAICKRIVEHLGGRIWVVESSIGHGATFAFVIPAD

>Candidatus Accumulibacter

RTYSRELERKSRELTAATLELQEANERLQELDRVKDDIMSSVTHELRTPLTSIRAFSELLRDDPKM
HLADRERFLGLIVSEAERLTRLINQTLDLAKIESGRADWNGCELDLKEVVEQSAATSQLVREKE
AHVEVDLPDNLPLILADRDLIQVMLNLLSNAVKFLTPGSGRIRVTLRSDDGLEVSADNGPGI
RPEDQQLVFEKFRQVGDTMTAKPSGTGLGLPISRRIVEHFGGQLWVESVPEGATFRFTLP

>Pontibacter roseus

SINNELRRKSLELRRATEQLRSANERLKQLDELKDEFLSTVTHELRTPLTSIRALSEILYDNP
EMEREDQEHFLQTIKESERLTRLITQVLDLERFESGKQKLNLPVQVEEVIQEAQTLEQLVQEKRIEL
VVDVQQNLPVGTGDRDRLMQVMVNLSNAVKFCHPENGRIVSGYCIDGDIKVNVDNGKGDIP
EFHKQIFDKFYQAQNNIRKPEGSGLGLAISKQIIESHQGRLWVESEPGRGAKFSFVLP

>Azospirillum brasilense

QLEHKTAELQRASAALRAANERLTELDRKDEFLSTVTHELRTPLTSIRALTEVLHDNPDI
ELEQRQEFLGLVITESERLTRLINQVLDMAKIEAGEIDWQVRPMDLGAALQAAAATARLFHERGIALAV
TIPHGLPPVRGDHDRLVQVAVNLLSNAAKFTPAGGRASLSVAEAGDALRVSVTDSGPGIAAEHQ
AIVFDRFRQVGDTMTDKPQGTGLGLAISKRIVEHLGGRIWVASAPGKGATFAFTVP

>Azospirillum halopraeferens

QLEHKTAELERASAALRAANDRLKELDRKDEFLSTVTHELRTPLTSIRALSEILHDHPDIEAEQR
QEFLGTIIESERLTRLINQVLDMAKIEAGEVEWTIGPVAIGTVMEQAAAATSQLLRDKGVTLTL
AIPDLPPVEGDADRLVQVAVNLLSNAAKFTPAGGRVELRVVAQEGGVRIDVADDGPGIAPEHQ
GIIFERFRQVGDTMTDKPQGTGLGLAICKRIVEHLRGRIVVDSAPGRGATFSFTVP

>Emticicia oligotrophica

ELKQKSNELKLLTEQLQSANEQLKLTDKLKDDFLATVTHEIRTPLSIKALSEILYDNEDIEHEER
QHFLNTIIESDRLSRLINQVLDLEKFESGKVRIVQESINFQEIINDSLDSIEQLAKEKGVKIKTV
IDKNLPKLIIGDRDRLMQVILNLLSNAVKFTANKGLITITAYNEENNLLFNVIDDGAGIAKEYQELIFD
KFYQAHDQTIKKPKGSGLGLAISKRIELHNGKIWVESEVGGKTKFSFSIP

>Candidatus Brocadia fulgida

LEKKVEERTKQLRQANIALEKANRLKSEFLATMSHELRTPLNAIIGFSEVLRDEVIGPLLEKQEK
LDDIRSSGQHLLGMINNILDFSKIEAGKFEITYEEFSLKEVIDEVLNAISEFSSKKGISIHTQFHPGIP

PIIVDKVKFKQIMFNLLSNAVKFTSEKGRITINAELVEQYVQIGVGDGTGIGIKSEIDKLFKPFQL
DGTYSRRYEGTGLGLILTKHLVELQGGKIWVESEYGKGSTFTFTLPIKPRKND

>Cystobacter violaceus

DLEAQNAMLRLNQLRKEADRIKSSFLGTVSHELRTPLSSIIGYSEMLAEGLVGALNAEQMQFVR
TIEKGNTLLKLSSILDMSQIEAGKVRLAFEWVDVQELVSSLTSVMPQAQRKGLTVKVHLPEV
AQPRVVADREKLSQVVVNLLANAVKFTPGGGQVEVRLSAHGPPGLVGEYRIEVEDTGVGIPL
DQRERIFQSFYQVDDSPTREYGGAGLGLAIVKSYVEGHGGQVSVTSEVGGSCFRVVLPE

>Alkalilimnicola ehrlichii

LEQKSRELEMATDELRRANERLKELDHLKDEFVSMVSHELRTPLTSIRAFGEILLSNPELDAEQR
KEFLQVVVKESERLTRLINQVLDLSKIESGAAEWHLETLDLNQVVQEAADATRQIFHDTRVDLQ
VKAPEQPTIITGDHDLIQLVINLLSNAAKFTDPDNGRVEVSVPVSARKLELRVQDNGPGISEAE
QRKIFDKFHQVSSQAGKPKGSLGLAICKLIMDAHSGDIRVESEPGAGATFICEFP

>Cystobacter fuscus

ELEHQNSQLLRLNQLRKESDRSKSSFLSTVSHELRTPLASIIGYSEMLAEGLVGGLNPEQMVFVRT
IMEKGNTLLKLSSILDMSQIEAGKVRLAFEWVDVREMVESAITSVTPQAQRKGLTLEANLPVGP
QPRVVADREKLRQVVVNLLANAVKFTPSLGRIDVRLSEVGTRESELASAGYHIEVEDTGVGIPENQ
RDRIFQSFYQVDDSPTREYGGAGLGLAIVKSYVEGHGGQVSVRSEVGGSCFRVVLPE

>Candidatus Poribacteria

YLFDDITREKEIDQMKSDFISLVSHELRTPLTSIIGFVSFILDGKAGAINDRQRNSLARVQRQSKRL
AALINDLLDISRIESGRIQMEQKSISLLDIVTQRLEEIRPQADEKSIRLVLTAPESIPDILGDEARMGQ
VFTNLIGNAIKFTPDNGEVNVKVEADGNLLHVEVIDTGPGIPPEERQKIFDKFYQLSDISTRQOQG
SGLGLSISKSIVEAHGGKLWIDDGNQGGKSNFQFVLPVREDD

>Microgenomates bacterium

TLEQLLNVVSTSLDRAQLYENLKIANEKLKELDKLDEFVSLASHELRSPTTIKGSSTILDGYA
GEVSKETREFLTAAYSENDRLIRLVNLLNISRIESGRKFLVTNVDMSKLMKDEVNDMMAAK
ERGLTLTIDKTENLPLVLADEDKIREVLINLLGNAIKFTHQGGITVSAKVDKDMVVTSTVDTGNGI
APEDQDLLFKKFSQVRRGTYSRQTGGTGLGLYISKKIIEGLQGQIWLSTVVGKGTFTFFSLP

>Candidatus Brocadia sinica JPN1

LEKKVEERTKELRQANIELEKANKLKSEFLATMSHELRTPLNAIIGFAEVLRDEVIGHLSKEQKEC
LDDIHSSGQHLLSMINNILDFSKIEAGKFELKYEEFSLEEVIDEVLNAISEFSGKKGISIHTHFHADIP
PIVADKVKFKQIMFNLLSNAVKFTPENGRVTIDAELAEQSVQIGVSDTGIGIKSEIDKLFKPFQL
DSSYSRRYEGTGLGLVLTAKHLVELHGGKIWVESEYEKSTFTFTLP

>Geobacter metallireducens

EDVDRVKTEFISTVSHELKTPLTSIKGSLQFIHDRGEGFGGMEGELLSVCLRNTRDRLIRLVNDILDI
ARIESGRLEFVPEPQDVRTLVTHSIEELAGYALEHDVSVENRCGQDLPPYGDRLVQVLTNLLS
NAIKFSRGERVVVTAERTSNYVAISVGDRGREILWSDRGKLFKQQLDNDSDRRQHGGTGLGL
AICKEIVEHHHGRIFYNEGTTGGNVFTFTVP

>Kosmotoga olearia

KIKELERLKSEFLATMSHELRTPLNSITGFVSMILQGISGEINEEQRKQLSMVHSSSIHLLNLINDIL
DISKIEAGRMSFSKERFNKDVISEVVQNVSLISKKEVKLITEIPDEIPEIYSDKRRVFQILLNLSN
AIKFTTEKGEVRIKCEVDGNLKLKISVSDTGKGIKEEDMETLFEAFRQISGSARRRYQGAGLGLYLS
KKLVTLGGEIWVESEYGKGSTFTFTLPLKSVEENNHEE

>Candidatus Kuenenia stuttgartiensis

NTILEKRVQEKATDLQMANVALEKANRLKSEFLTTMSHELRTPLNAIIGFAEVLRDEIAGSLSKD
QKEYVTDIHSSGHLLDMINNILDLSKIETGKMHLQYEEFCIEDAINDTLTIINASANNKGISVHTN

IQDNTPLLSADKTKFRQILYNLLSNAVKFTPENGGKITINVFQKDNSLQFEIVDTGIGIKPEDKEKLF
EAFHQADASLTREYEGTGLGLHLTKRLEVELHGGKIWAESTFGKGSTFFFILP

>Hahella chejuensis

KRLDILKNEFISTVSHELRTPLTAINGSLALLDSGRLGLLPDQAKQMLHIAYSNCERLLRLINDLLD
IQKIESGGLEFSFERADLRDLIEEALNANRQLAEKASVTLVFATPESPALVKADRDLRQVLDNLI
SNAIKFSPADEVTLTLNAHAHYWEVAVQDQGGVPPPEFRNKIFQKFAQADSSDTRLKGGTGLG
LSISREIVQRHHGQIGYHSEANCGATFYFRL

>Exiguobacterium

EVDRLKTELVSTVSHELRTPLTSIIGFTELLRYRSLQPERVDQYLGMIHQETLRLEDLISNLLDVQR
MESGQQAYDNQPERLHDILAETIAMHQGSSLHHNLTLECGTNLLVYGDRDRLSQLFTNLLHNAI
KYSPEGGKVAVCVTHDQADTIEISVQDKGIGIPESAISKLFDFKRYRVDNSASRKIGGTGLGLAICK
EIAEGHGGSIHVDSSVGEESTFTITLPHYVAED

>Orenia marismortui

IRLEEMRKDFVGNVSHELKTPLTSIKGYVETLLDSKPSDNIYNSFLNIIKEEADRLQRLIEDLLNL
SKIESHSNYKSKEDVDIVKIIDDTILLLESKAKKKDINLVADLNLSAPNIQADQDQLSRLMINLIDN
AIKYTPQGGKVKVSVDQKEEKLILEVEDNGIGIPEDDLPRIFERFYRVDKARSRKLGGTGLGLSIV
KHIVEEHQGNISVESEIGQGTRFSVELP

>Microcoleus sp

KLKELDTLSEFVATASHELRTPLTGMAMSLNLLLETAQQKLSNREQELLHAAVEDVERLRTL
NDLLDLSKIESGRIEFIAVEVELLIEK AISILSVQAQEKQIELTPSIPTDIEPVKADPNKIIWVLINLI
ANALRYTEAGGHIQVSAQQKDDWVYLSVADNGLGIPWEYQAKIFDKFVQVKTDKDVGGSGLG
LAICKEIVKAHSGTIWVDSAPGEGSTFTFTLP

>Desulfotomaculum kuznetsovii

KKLEQMRTEFVANVSHELRTPLTSIKGFLETLLDGAMNDPKTSRQFLEIMSQETERLTRLVDDLL
DLSKIEERRVHRWQPVNLVDIINRVASLFRPQAKEKALTLSLEVPRDLPSVYGDPDMAQVLIN
LLDNAIKYTPPRGSVTIRAMVLEDQLRVEVEDTGIGIPAESLPRIFERFYRVDKARSRELGGFGIGL
AIVKHIIRAHGGKIEVESTPGKGSIFYFTLP

>Geitlerinema sp

KEVDRMKTDFISTVSHELRTPLTSVLGFASIVQEKLEESVFPLVLEGDRKVQKHVRRVRQNLIII
SEAERLTVLINDVLDIAKMEAGKVEWNMTLLSAEELLDRAAATSALIDGKNIEIRREVEPDLP
YGDRDLRIQVVINLLSNAVKFTDEGTITCRVEHRDNYLCLSVQDTGIGIAPEDIDKVFDFKQVG
ETLTDKPKGTGLGLPICQQILEHHGGRIWAESQFGEESTFSFTLP

>Acetohalobium arabaticum

ELRRELIQIRTEFVSNVSHELRTPLTSIKGYVETLLDERDCEPGVRERFLQVIKDETDRLERLITDLL
NLSQLESASDSFDQELVNLNQVIENVLTTVMKADNKGIDLKVDVPVDITGIGKSRGQLERLYIN
LVDNGIKYTSSEGGQVKIKVYEDEDVWSEIIDTGMGIPEEDLPRIFERFYRVDKTRSRLGGTGL
GLSIVKHILERHNGGIEVESKVEEGTKFIFWLPK

>Caldicellulosiruptor saccharolyticus

KLDSMRKQFVANVSHELRTPIITIKTYSETLLDVDNEETKKQFLSVIIKECDRMTRLVSDLLYLSR
LDSEGENILNLEEVNLSLVRVCEKLIHAKKNQTLSCSILQDIVAMVDRDKIEQVLINLISNAV
TYVQEGGQINVLQKEEDKIKIIVKDNNGPIPEEDLPRIFERFYRVDKARSRELGGSGGLGLSIADEI
VKAHGGKILVESKVGSGTTFTVVLP

>Bacillus acidiproducens

IRRLEKIRSEFVANVSHELKTPITSVKGAETLLDGAMYHEETCREFLKIIYEESDRLHRLISDILL
SKIEQHRVFPNTEPLNIKEVILETAGTAKKIAEKQLELILPQQQTVTIEADKDRLQQIILNLLTNAI
AYTPEKGKIEIRLQEREDEVLDIVADTGIGISQKDLPRIFERFYRVDKARSRQSGGTGLGLAIVKHL
AESNHGRINVESEEGKGSVFTVTLPKTQKNPD

>Thermodesulfovibrio thiophilus

KLKEMDKMKSDFSLMAHELRTPLASIKEGTNLLLEGIGGEVSEKQKLLSIISEESNRLIELVNSL
LDLSKMESGMMTFYFTESDITQLINDVIEELEPLAMAKNITLRVDISQNLPHYKIDQERILRVMRN
LIGNAIKFTPDGGHITVSARAVDHGLMVSVRDTGPGIPEKDLEIIFDKFKQTIGSYGKIEGTGLGL
AIVKHIINAHGKGVWAESKLGYGSTFIFVLP

>Methanococcoides methylutens

NKTLEGEVKNKTSELERSNEYLQNANTKLKELDRLKNDFVSMVSHELKTPLTAMKTSSEFMKES
ECNPEVKEEMLDLIIRNIDRQARMVDDLLDISRIESGRMKFNPEDVNINEIVETSLQTVKQAKDK
GINIKVGYPEESLTIKTDKDKLIRVFNLLTNAIKFTPHDGEVEISVTDAGNHIQASIKDNGIGIPPE
KREKIFDKFYQVDSTATRKAGGTGLGLAIKGIIDGQGSIFVESETYKGSTFTFRLPKELKEEDFT
EIE

>Methanobolus psychrophilus

NKTLEREVEKKTRELQESNESLTAANIKLKELDGLKNEFLSMVSHELKTPLTAMKTSSEFLLED
CDPAVRNQLLALIVRNVDRQARLVDDLLDISRIESNRMKFNMEPVSLQEVLDHSIENIHRLSESK
GVALGADISRSLPAVLADRDKLIQIFVNLNNAVKFTQRGGSVTVLAQESGDNIEVIVSDTGIGID
PSHAEHIFDKFYQIDSTSTRTAGGCGLGLAITKGLVEGMNGSIRVESEPGIGSRFIVTLKKA

>Methanohalophilus mahii

NRKLEEEVDKKTQLENSNQLQKLNVRVNEFLSMVSHELKTPLTAIRTSSEFLREEDYDRETQE
EMLDLIIRNIDRQSRMVDDLLDISRIESGRMVFKEEKVDLKEIIDNVIQMMPEMATKHGIAIEREL
DNDPRVKADKDKLLRVFNLLNNAIKFTADRGESIKIKTEETGDFVEISVIDKIGIPENEQEKIFE
KFYQVDSTSRKVGGSGLGLAIKGIIEGHGGSIRVKSEPGVGSFTVFTLRKWE

>Methanobolus tindarius

SKSLEKEVGVKTQELKESNESLRSANTKLKELDRLKTEFLSVVSHELKTPLTAMRTSSEFLRESDE
CDISIRKMLDIIRNIDRQSRMVDDLLDISRIESNRMKFHQETMDISKAIRCSLEMLSSNLKDKNM
NVLVELPDDLSSVYTDKDKLVQVFNLLNNAVKFSKNNGNIRIYAQEDGDFVRISVSDDGIGMS
SDELDKIFDKFYQIDSTSTRKVGGSGGLSIVKGIIEGQGGTISATSEPGNGSTFVFTLRK

>Candidatus Methanoperedens nitroreducens

NKSLELEVARKTSELNESNKLQDANKRLKELDKSKSDFLSMVSHLKTPLTAMKVSSEFLLED
DSKLVTRKELIQIIKNIDRLTRLVNDLLDISMIESGELRFSKEIVDLHDIIDIAVGTVKNQYEKEL
NITTDIPKNLSKINADKDRIVQVFNVLNALRFTPEGGNVEIRACEFEECIEVHVKDDGVGIPPD
KIDKIFDKFYRINTTSTRSHNGAGLGLAITRGIMEGHGGSIRAQSTPAKGSVFILTF

>Methanothermococcus okinawensis

MADEIRKSNEKLLKQAELEKSYNELKELDKLKS DIVAIVSHELRTPLTSIKGYVELVLDGTMGA
ITESQKKCLEIANKNIDRLKRLIDNMLDLSKIEHGELEMHMEINLKVENVVDTLKPLADEKNI
NIYKINDIIMKGDKDRITQVLTNLIENAIKFSVNGKVEIQALKEGNSVHIKIIDNGPGIPKDLDR
IDRFYQVDSPEKRIKGGSGGLGLAVCKSIETHGGTIWVESKLGSGSVFHIPI

>Methanomethylovorans hollandica

KRAEDSLKKYAEDLAKANEELSKVNEELKSLDRMKDEFNSVSHFKTPLTSIRGYSQLILDGTL
GDINQQKKAMDTVIRNSDRLRRLVDSLLYLSRAQVGRMSYYFEKLQLADIINNCINDLLLQAKN
KGIILRTEMEDIPPVRADKDKLTDVLTNLIDNALKFTSEGGLVIVAARKIPNGVHIEVKDTGIGIPA
DHIPHLFQRFYQVDSSTSRRYGGTGLGLYISRTIVEAHGGNIWIESEEGMGTTVHVELPPWQE

>Methanothermococcus thermolithotrophicus

MAEEIRISHEKLLKHAEELEKSYNELKELDKMKSDIIVIVSHELRTPLTSIKGYVELVLDGTIGPITE
SQRKCLEIAEDNIKRLKRLIESMLDLSKIERGELEMNMEEIGIKHFVEKILSSLKPLADEKNINMNH
DVEDIAIKGDKDRIAQVLTNLVENAIKFTPINGNIGVNAFKENEYAHITVTDNGPGIPEKDLCRIFD
RFYQVDSSAKRKKGGSGGLGLAVCKSIVEAHGGSIWVESKHGKGSTFHILLPLNQDE

>Methanococcus vannielii

EMADQINSSNEKLLKHAEELEKSYNELKELDKMKSDIIVIVSHELRTPLTSIKGYVELVLDGTMG
TINESQRKCLQVADDNIIRLRLIESMLDLSKIERGELEMYREAMNVKDTVSDVIEYLTPLATEKN
IKLKQDIKDLLINADKDRITQVFTNLIENAIKFSANESIMIIGKETENGDVHITVKDNGAGIPKDL
EKIFDQFYQVDSSTKRKKGGSGGLGLAVCKSIIQAHGGTIWVESELGRGSTFHIVLPALIYEESSIVE

>Methanosalsum zhilinae

HKKAEDALKKYSNKLSQVNKELWEANKELKSLNQLKNEFISNVSHELKTPLVSIRGYSEIMNGE
NLGPLNDKQKRAVETILRNSERLRRHVDSLMIYSMEQMGKIRYSFDSINIEGIIDDAITDILPQIRD
TGKDLKLEKNVSEDLPEIQGDAQKLTDLLTNLLSNSVKFTPDGGRISINVKQEGDELHIVVEDTGI
GIPKEVIPDIFDRFYQIDSSTKRKFGGTGVGLYICKSIVEGHSGKIWVESDRGKGTSVHVKLPVKN
RPEEKKQ

>Methanocaldococcus fervens

MAEDIRKSHEELKRHTKELKESYEKLKEIDKDKSEIISIVSHELRTPLTSIKGYVELVLDGLMGELN
SNQRRCLEIANNIDRLKRLIDDMLDLSKIESGTIKFDIKDVKIKDMVVDVLSLGPQIKEKNIEIK
CEIEENLTAKVDKDRITQVLINLIENAIKFSVKGVIEIHAFRDNNYAHIIIKDYGPPIPCKDLKIF
DKFYQVNFPIKKGAGLGLAICKSIIAHGGKIWVESELGKGTFSHVLLPIE

>ResE Bacillus subtilis

NVSHELRTPISSLQGYSEAIVDDIASSEEDRKEIAQIIYDESLRMGRLVNDLLDLARMES
GHTGLHYEKINVNEFLEKIIRKFSGVAKEKNIALDHDISLTEEEFMFDEDKMEQVFTNLI
DNALRHTSAGGSVSISVHSVKDGLKIDIKDSGSGIPEEDLPFIFERFYKADKARTRGRAG
TGLGLAIVKNIVEAHNGSITVHSRIDKGTTFSFYIPTKR

>YkoH Bacillus subtilis

ARRIINPIRRLMITMKDIQRDKEFKTISLEGQSNDELYQMGLTFNEMAMMLKEHYDKQQQ
FVQDASHELKTPLTIIESYSSLMKRWGAKKPEVLEESIEAIHSEAVHMKKLTNQLLALAK
SHQGLEVDLKTIDLIIKAAAVMQTLQSVYQRDILLETDKESLLVKADEERIKQLLTILLD
NAIKYSEKPIEMSAGTRNGRPFLSVRDEGIGIPEEHIPHLFERFYRADEARNRKTGGTGL
GLSIAKQIADHEGIELSVKSKPGQGTAVTMQFSEQNGGGR

>Mbur_0694 (LtrK)

NKTLEDEVKSKTSELERSNDYLNANKKLKELDRLKNDVSMVSHLKTPLTAMKTSSEFLRES
ECSREIKEEMLDLIIRNIDRQARMVDDLLDISRIESGKMKFTPEDVNIKEIIEISLHNVTKHAKDKSI
KIMVNCPPDDVPAISTDKDKLIRIFVNLLTNAIKFTPEEGEVTVIVEDHEDYLQTSIKDNGIGILEEK
RDKIFDKFYQVDSTARKAGGTGLGLAIIKGIIDGQGGKIYLESETGKGSTFTFRLPKELKEDDFTE
IEGT

>PhoR Escherichia coli

RLSWWLWVDRSMTPPPGRGSWEPLLYGLHQMQLRNKKRRRELGNLIKFRFRSGAESLPDAV
VLTTEEGGIFWCNGLAQQILGLRWPEDNGQNILNLLRYPEFTQYLKTRDFSRPLNLVLNT
GRHLEIRVMPYTHKQLLMVARDVTQMHQLEGARRNFFANVSHELRTPLTVLQGYLEMNE
QPLEGAVREKALHTMREQTQRMGLVKQLLTLKIEAAPHLLNEKVDVPMMLRVVEREA
QTLKQKQTFTEIDNGLKVSIGNEDQLRSAINLVYNAVNHHTPEGTHITVRWQRPVPHGAE
FSVEDNGPGIAPEHIPRLTERFYRVDKARSRQTGGSGGLGLAIVKHAVNHESRLNIESTV
GKGTRFSFVIPERLIAKNSD

>PhoR *Bacillus subtilis*

SMTSRYKRSIESATNVATELSKGNVDARTYGGYIRRSCLKGHAMNSLAIDLMEMTRTQEM
QRDRLLTVIENIGSGLIMIDGRGFNLVNRSYAKQFHINPNHMLRRLYHDAFEHEEVIQL
VEDIFMTETKCKLLRPLIKIERRYFEVDGVPIMGPDDEWKGIVLVFHDMTETKKLEQMR
KDFVANVSHELKTPITSIKGFTETLLDGAMEDKEALSEFLSIILKESERLQSLVQDLLDL
SKIEQQNFTLSIETFEPKMLGEIETLLKHKADEKGISLHLNVPKDPQYVSGDPYRLKQV
FLNLVNNALTYTPEGGSVAINVKPREKDIQIEVADSGIGIQKEEIPRIFERFYRVDKDRS
RNSGGTGLGLAIVKHLIEAHEGKIDVTSELGRGTVFTVTLKRAAEKSA

>CheA *Escherichia coli*

RSNESTSIRVAVEKVDQLINLVGELVITQSMLAQRSSSELDPVNHGDLITSMGQLQRNARDLQESV
MSIRMMPEYVFSRYPRDLVRLAGKLGKQVELTLVGSSTELDKSLIERIIDPLTHLVRNSLDHGIE
LPEKRLAAGKNSVGNLILSAEHQGGNICIEVTDDGAGLNRERILAKAASQGLTVSENMSDDEVA
MLIFAPGFSTAEQVTDVSGRGGVMDVVKRNQIKMGGHVEIQSKQGTGTTIRILLPLTL

>CheA *Bacillus subtilis*

GGSKTIRVNIDRLDSLMLNLFEEELVIDRGRLEQIAKELEHNETTETVERMTRISGDLQSII
LNMRMPVETVFNRFPRMIRQLQKELNKKIELSIIIGAETELDRTVIDEIGDPLVHLIRNS
IDHGIEAPETRLQKGKPESGKVVLKAYHSGNHVFIEVEDDGAGLNRKKILEKALERGVIT
EKEAETLEDNQIYELIFAPGFSTADQISDISGRGVGLDVVKNKLESLGGSVSVKSAEGQG
SLFSIQLPLTL

A1.2 Bacterial and archaeal RR protein sequences to construct phylogenetic tree for LtrK (in FASTA format)

>*Methanococcoides methylutens*

TREKILIVDDEPDVAVMATRRALEADGYNVIDAGDGAMAFDLLRSDIPDVILLDVMMPDMDGFE
VCKRLKEDPAYRNIPVIMLTAKGEIDDKVGGLDIGADDYVTKPFNLKELKARIKTVLRRRAQE

>*Methanohalophilus mahii*

SRQTILIIDDEPDVAVMATRRALQSEGYNVIEAYNGKTGLEAIEEKTDLILLDVMMPDMDGFEICKQ
LKKDELYNHIPIIMLTAKGEVDDRIEGIETGADDYITKPFNLRELKARIRMLVLRRAQN

>*Methanobolus tindarius*

TRQKVLIVDDEMDALISLKVVALEAEGYNVAEAKDGHAEIDKVHSELPAILLDLMIPGIDGFEVC
RQLKSDDMYRHIPVIMLTARGEIDDKVEGIELGADDYVTKPFNLKELKARVVMVLR

>*Candidatus Methanoperedens nitroreducens*

TRVKILIVDDEKDVVPLKMSLEADNYSVIEAYAGYGAEKARSEIPDLIILLDMLPDMGFEVC
NRLREDPLTELIPIIMLTGKDDISNKIEGLERGADDYITKPFNLSELKARIRTVLRRSR

>*Methanobolus psychrophilus*

REKILIVDEEIDASTALKLALETEGYNVIEALDGYEGIMKAKSENPDILLDIMPGMDGFEVCQL
LKADPASRHIPVIMLTAKGEVDDKVEGLELGADDYVTKPFNLKELKARIRIVLR

>*Methanomethylovorans hollandica*

RLKILIVDDEPDISDLLRSLSNQPYDFVVAINGQQAIEATLRERPDILLDIMPGMDGFEVVEH
LKQSATSANIPVIMISAKTAIEDKVRGMELGIDDYIAKPFDKREVNARIRVMRR

>*Methanocella conradii*

KIMVVDDEPDLVEVVKLILES DGYQVVTALSGQEALDKMEKEMPDLVLLDIIMPKMDGWEVHS
RIKSNPKTHDIPVIMLTAKDQRIDKLIGLHVVRVDDYITKPFGR AELLERIKRVLQ

>Methanosarcina acetivorans

KPKVLIVDDK KENVELMEAYLA VEPYDVITAYGGKEAFQKVKEEKPDILLDVMMP EVNGYEV
CKILKGNPETQFIPVLM LTA LSELEDRI R GIEVGADDFLTKPINRLELKTRVKSLLR

>Methanosarcina lacustris

KPKILIVDDM KENVELMEAYLA VEPYEVVSASGGKEALQKVKNENPDILLDVMMP ELNGYEV C
KILKENPETQFIPVLM LTA LSELEDRI R GIEVGADDFLTKPINRLELKTRVKSLLR

>Methanocella paludicola

KIMVVDDEPDLLEVV KLILES DGYQVVTANSGQEALDKIEKEMPDLVLLDIIMPRMDGWEVFSRI
KGNTKTHDIPVIMLTAKDQRIDKLIGLHVVRVDDYITKPFGR AELLERIKRVL

>Methanosarcina siciliae

KPKVLIVDDK KENVELMEAYLA VEPYDVITAYGGKEALQKVR YENPDIVLLDVMMP EVNGYEV
CKILKENPETQFIPVLM LTA LSELEDRI K GIEVGADDFLTKPINRIELKTRVKSLLR

>Methanosalsum zhilinae

RIKILVDDDEPDI VELIKLSFARQPYDFIAAYS GEEAVEKAREELPDILLDIMMPGMDGYEVVEQ
LKQIEDTREIPVIMVSAKTDVDDKVRGMELGIDDYISKPFDKREMNARIKMVMKR

>Methanohalobium evestigatum

SESKILIVDDEPEL VELLADYLDGYNTIAAYNGQEAINIESDNIDVILLDVMMPDINGFEVCTHIK
NDDSLNYIPVLMITALS DHDNKIHLGDSGADDFLTKPVDGEELNARVRSALR

>Methanosarcina horonobensis

KPKVLIVDDM KENVELMEAYLS VEPYEVICAYGGKEALQKV EEEKPDIVLLDVMMP EVNGYEV
CKILKENPETQFIPVLM LTA LSELEDRI K GIEVGADDFLTKPINRLELKTRVKSLLR

>Methanosarcina mazei

RPKVLIVDDM RENVELMEAYLA VEPYKVFCAYGGKEALRIV DKEKPDIVLLDVMMP EVNGYEV
CKILKENPETQFIPVLM LTA LSELEDRI K GIEVGADDFLTKPINRLELKTRVKSLLR

>Methanosarcina vacuolata

MLIVDDVLENIELIEAYLS VEPYDLITANS GKEAIQKLKEEKLDMILLDIMMPEISGYEVCKIIKKD
PETQFIPVLM LTA LSEIEDRI K GIEAGADDFLTKPINRLELKTRVKSLLR

>Fervidicella metallireducens

KILIIDDELHIVELLKYNLEANGYKV IYELNGKNGFEAAVEKRPDLILLDIMLPEDMGDFVCKALK
REKELENIPIIMLTAKSEEFDKILGLEIGAEDYITKPF SVRELLARIKVVLRRN

>Thermoanaerobacterium saccharolyticum

HTILVIEDEAHILELLRYNLEAQGYNVVLADNGKDGFEKCKETNPDLVLLDLMLPDIDGIDVCKK
LKSDEHLKNIPIIMLTAKSEEV DKILGLELGADDYITKPF SIRELLARIKVVLRRSK

>Pelosinus fermentans

ILVVDDEPSIVELIEFNLQKTGYHVLKAENGHEALQLVRTNKPDLIILDLMPDIDGIEVCRRLKGQ
QETAAPIIMLTAKNEEV DKIVGLELGADDYMTKPF SPRELMARVKA VLRSSH

>Clostridiales bacterium

RILVIDDDSKITAF LRRSLMYEGYHVLEASDGYRGLQLARENPPDLVILDVMMPGMDGWEVCRR
LREVSNPILMLTARDEIADRVRGLDVGADDYLVKPF ALEELLARVRALLRRYQ

>Clostridium pasteurianum

KILIVDDEEHIIELIKFNLENGYKVICANN GND AVKLAKSEIPQLILLDLMLPGLDGYDVCREIRK
DPSISTVPVIMITAKSEIDKILGLELGADDYITKPF SIREMMARVKAMLRRAK

>Clostridium algidicarnis
SKEKILIVDDEEHIIIPELLKFNLEKSGYEVLVAMDGNNNAISKAKSEIPKLILLDLMLPGMDGYDVCK
EIRKDNLTAHIPIIMLTAKGEELDKIIGLELGADDYITKPFVRELVARVKAVLRRVN

>Desulfovibrio frigidus
KILVVEDHHDTELLKYNLTSSGFVVVTAMDGLKGLALAKSELPLILLDIMLPEIDGLEVCRRLLK
QEAAQTQHIPPVMLTAKGEEVDRVVGLELGADDYIVKPFSPRELVLRIKAVLRRSAD

>Caldsalinibacter kiritimatiensis
KILVVDDEEHIVELINFNLEKHGYKVVTAFNGKDAFKVKEESPDIVLDLMLPEIDGIDVCKML
KKDSDETEKIPIIMLTARSEETDKVLGLELGADDYVTKPFVRELVARVKAVLRRSSD

>Pelosinus sp
VLIVDDEPTIAELIEFNHKAAGFQVLKADNGHTALQLVRSKPDILDLMLPGIDGMEVCRRLLK
QHTASIPHIIMLTAKNEEVVKIVGLELGADDYMTKPFSPRELVARVKAVLRRSH

>Clostridium akagii
KILVVDDEEHIVELIKFNLETNGFKVVSASNGIDALKLAKHEVPQLVLLDLMLPGMDGYDVCREI
RRDQSISTMPVIMITAKGEELDKILGLELGADDYITKPFVRELVARVKAVLRRTR

>Bacillus wakoensis
AKSKVLVADDDPNVCEIIRLYFQENNFVIEANNGRQALTLFASDKPDIILLDIMPPIDIDGYEVC
EIRKKADIPHIIMLSAKAEVDRVLGLEIGADDYVTKPFSPREILARIKAIFRRT

>Bacillus pumilus
ILIVDDELDMLELIRSFLQRQGFHVITANNGTGALHQLEKEAIDLVLVLDIMPPDMDGFEVCQRIR
QTSQIPILFLTARSYEEDRIKLEIGADDYIMKPFSLRELAARIETTLRR

>Desulfosporosinus orientis
ILVVDDEEPILELLRFNLEKEGYLVCLAKDGQEALDLVEKEQPDLVVLDDLMLPGMDGLEVCRR
RFIPKYQQIPIIMLTAKGEVIDKVLGLELGADDYMTKPFSPRELIARIKARLRR

>Bacillus invictae
ILIVDDELDMLELIGSFLQRQGFNVITANNGTGALHQLEKEAIDLVLVLDIMPPDMDGFEVCQRIR
QTSQIPILFLTARSYEEDRIKLEIGADDYIMKPFSLRELAARIETTLRR

>Anaeroarcus burkinensis
ILIVDDEASIRELLRFNLQKEGYVVSEAADGLEALQLIKQQRPDVLLDLMLPGMDGLEVCRR
GQQGTAGIPIIMLTAKNEEVVKIVGLELGADDYQTKPFGLRELMARVKAVLRRSQ

>Anaeromusa acidaminophila
ILIVDDEASIRELLRFNLQKEGYVVSEAADGLEALQLIKQQRPDVLLDLMLPGMDGLEVCRR
GQQGTAGIPIIMLTAKNEEVVKIVGLELGADDYQTKPFGLRELMARVKAVLRRSQ

>Herbidospira cretacea
RVLVVDDEPEVRTAIARALRVEGHRVASAVDGVVTALAAIAGNAPDVVVLVDMMPDMDGLEVC
RRLRKNDRTPVLMMLTALDGVGDRVAGLDAGADDYLVKPFALAEELFARVRALLRRA

>Desulfovibrio bastinii
KILVVEDNSDTLELLKYNLTSSGYSVVVTAMDGHKALEQARNEKPDILILLDLMLPGLDGLEVCRR
LKQEAGIQHIPVIMLTAKGEEVDRVVGLELGADDYVVKPFSPRELVLRIKAVLRRSSE

>Methylobacterium mobilis
ILVIEDEPAIQELLALNITQAGHNAIKALSCEIALDLMRETVPDLILLDWMLPGMNGLELARKLKS
DTYTKNIPHIIMLTARGEEDKVRGLEVGADDYVTKPFSPRELNARIKAVLRR

>PhoP Bacillus subtilis
KILVVDDEESIVTLLQYNLERSGYDVITASDGEEALKKAETEKPDIVLDVMLPKLDGIEVCCKQL
RQQLMFPILMLTAKDEEFDKVLGLELGADDYMTKPFSPREVNARVKAILRRSE

>Phob Escherichia coli

RRILVVEDEAPIREMVCVLEQNGFQPVVEAEDYDSAVNQLNEPWPDLILLDWMLPGGSGIQFIKH
LKRESMTRDIPVVMLTARGEEDRVRGLETGADDYITKPFSPKELVARIKAVMRR

>Yycf Bacillus subtilis

KILVVDDEKPIADILEFNLRKEGYEVHCAHDGNEAVEMVEELQPDLILLDIMLPNKDGVVEVCREV
RKKYDMPHMLTAKDSEIDKVIGLEIGADDYVTKPFSTRELLARVKANLRR

>Phop Mycobacterium tuberculosis

RVLVVDDEANIVELLSVSLKFQGFVYATATNGAQALDRARETRPDAVILDVXXPGXDGFVLR
LRADGIDAPALFLTARDSLQDKIAGLTLGGDDYVTKPFSLEEVVARLRVILRR

>Mbur 0695

TRHKVLIVDDEPDAVIATKRALEADGYNVIEANNGTMAFDALKSDIPDVILLDVMMPDMDGFV
CRRLKEEPLYENIPVIMLTAKGEINDKVEGLDIGADDYMTKPFNLKELKARIKTVLRRTQD

>CheY Escherichia coli

MADKELKFLVDDFSTMRRIVRNLLKELGFNNVEEAEDGVDALNKLQAGGYGFVISDWNMPN
MDGLELLKTIRADGAMSALPVMVTAEAKKENIIAAAQAGASGYVVKPFTAATLEEKLNKIFEK
LGM

Appendix 2: Plasmid sequences

A2.1 Plasmid sequence of pReceiverBo3_LtrK: The highlighted capital sequence represents the cytoplasmic LtrK gene

```
gcttcgaaggaattcggtagcATGAACGAACCGGATCATCGGCGTCTGGATATAATCGGCATGCAAACC
GAAGGCCTGAGCGGCAGCGCCAGTTTAGCGAACGTAATACGAATGGCTTCCTTGAAGAGA
AGATAATTAGCTATGTGCCGGTGAACCTGGTACAATCAACAGTGGAGCGTGGGCGTGGTGAC
CCCCCTGTGGTATGTGGGCAGCCTGATTAGAGCGTGTATGTGAAACAGGGCCTGTTTGCGA
TGATTAGCATTGCGTTTATTCTGTTTATTAGCGCGTTCATCGCGCTGATTCTGCTGAGCTGGA
ATAAAACCTGGAGGATGAAGTGAAAAGCAAACCAGCGAACTGGAACGTAGCAATGATTA
CCTGCAGAACGCGAATAAGAAATTGAAAGAACTTGATCGTTGAAAAATGATTTTGTGAGCA
TGGTGAGCCATGAGCTGAAAACCCCGCTGACCGCCATGAAAACGAGCAGCGAATTTCTGCG
TGAAAGCGAATGCAGCCGTGAAATAAAGGAAGAAATGCTGGATCTGATCATAACGTAACATA
GATCGTCAGGCGCGTATGGTGGATGATCTGTTGGATATTAGCCGTATTGAGAGCGGGAAAAT
GAAATTCACCCCGAAGATGTAAACATCAAAGAGATAATAGAGATTAGCCTGCATAATGTT
ACCAACATGCGAAGGATAAGAGCATTAAAGATTATGGTTAATTGCCCGACGATGTGCCAG
CGATTTGACCGATAAAGATAAACTGATCCGTATTTTTGTGAACCTGCTGACCAATGCGATC
AAATTTACCCGAGAAGAAGGCGAAGTGACCCTGATTGTGGAAGATCATGAAGATTACTTAC
AGACCAGCATAAAAGACAATGGCATAGGCATTTTGGAGGAAAAGCGGGATAAAATCTTCGA
CAAATTCTATCAGGTTGATAGCACCGCGACCCGTAAGCAGGCGGCACCGGCCTGGGCCTG
GCGATCATCAAGGGGATCATTGATGGCCAGGGTGGCAAAATCTACCTGGAATCGGAAACCG
GCAAAGGCAGTACCTTTACCTTTCTGCTGCCGAAGGAACTGAAAGAGGACGATTTACCGAG
ATTGAAGGCACCTAGctcgagcaccaccaccaccactaatgtaaftaagtggcgcttcctagcctgataaaacagaatttgctggc
ggcagtagcgcggtggtccacctgaccccatgccgaactcagaagtgaacgccgtagcggcgatgtagtggtggggtctcccatcgagagtag
ggaactgccagcatcaataaaacgaaaggctcagtcgaaagactgggcttctgtttatctgtttgtcggggaacgctcctcctgagtaggacaat
ccgccgggagcggattggaactgcaagcaacggccggagggtggcgggcaggacgcccgcataaactgccagcatcaataagcagaa
ggccatcctgacggatggccttttgcgtttctacaactctttgttttttctaataatcaaatatgatccgctgagcaataactgataacccttgg
ggcctctaaacgggcttgagggttttgcgtgaaaggaggaactatccgattggcgaatgggacgcccctgtagcggcgcatgaagcggcgcg
ggtgtggtggttacgcgcagcgtgaccgctacactgccagcgccttagcggcctcttctgctttctccctccttctcggcactgctccggcttcc
cccgtcaagctctaaatcgggggctcccttaggggtccgatttagtgccttacggcacctcaccaccccaaaaactgattagggatggttcacgtagtg
ggccatgccctgatagacggttttccctttgacgttgagtcacgtctttaaagtgactctgttccaaactggaacaactcaaccctatctcg
gtctatctttgattataaggattttccgatttggcctattggttaaaaatgagctgatttaacaaaatgaacgcgaatttaacaaaatgaacttata
caatttctgcccacgatggcatgagattatcaaaaaggatcttcaactatgactctttaaataaaaatgaagtttaaatcaatcaagtatatatgata
aactgtgctgacagttaccatgcttaacagtgaggcacctatctcagcagatctgtctatttctcattcaccatagttgctgactccccgctgtagataac
tacgatacgggagggttaccatctgcccagtgctgcaatgataccgcgagaccacgctcaccggctccagattatcagcaataaacagccagc
cggaaaggccgagcgcagaagtgtcctgcaactttatccgctccatccagcttataattggtccgggagctagagtaagtagttccaggttaata
gttggcgaacgttggcattgctacagcagcgtggtgtcacgctcgtctgttggatggctcattcagctccggttcccaacgatcaaggcaggttac
atgatccccatgttgcataaaagcggtagctcttccgctccgatcgttgcagaagtaagttggccagctgtatcactcatggttatggcagca
ctgcataattcttactgtcatgccatccgtaagatgctttctgtgactggtgagtactcaaccaagtcattctgagaatagtgtagcggcgaccgagttgc
tcttggccggcgtcaatacgggataataccgcgccacatagcagaacttfaaaagtctcatcattggaaaacgctcttccggggcgaactcctcaagat
cttaccgctgttgagatccagttcagatgtaaccactcgtgacccaactgatctcagcacttttactttaccagcgttctgggtgagcaaaaacaggaa
ggcaaatgccgcaaaaagggaataaggcgacacggaatgtgaaactcactcttcttttcaatcatgattgaagcatttatcagggttattgtct
catgagcggatacatattgaaatgatttagaaaaataaacaataggatgacaaaatecccttaacgtgagtttctgtccactgagcgtcacacccgt
agaaaagatcaaaagatcttctgagatccttttctgcgcgtaactgtgctgttcaaaacaaaaaaccaccgctaccagcgggtgttgttggcgatc
aagagctaccaactcttttccgaaggtaactggcttcagcagagcgcagataccaactgtccttctagttagccgtagttaggccaccactcaaga
actctgtagcaccgctacatacctcgtctgtaactctgttaccagtgctgctccagtgccgataagctgtgcttaccgggttggactcaagacgat
agttaccggataaggcgcagcggctgggctgaacggggggtcgtgcacacagcccagctggagcgaacgacctacaccgaactgagatacctaca
gctgtgagctatgaaaagcggcaccgcttcccgaaggagaaaagcggacaggtatccggtaaagcggcagggcggcaacggagagcgcacgagg
gagctccagggggaaacgcctggtatctttagctctgctgggttccaccctctgactgagcgtcattttgtgatgctcgtcagggggcgaggc
```

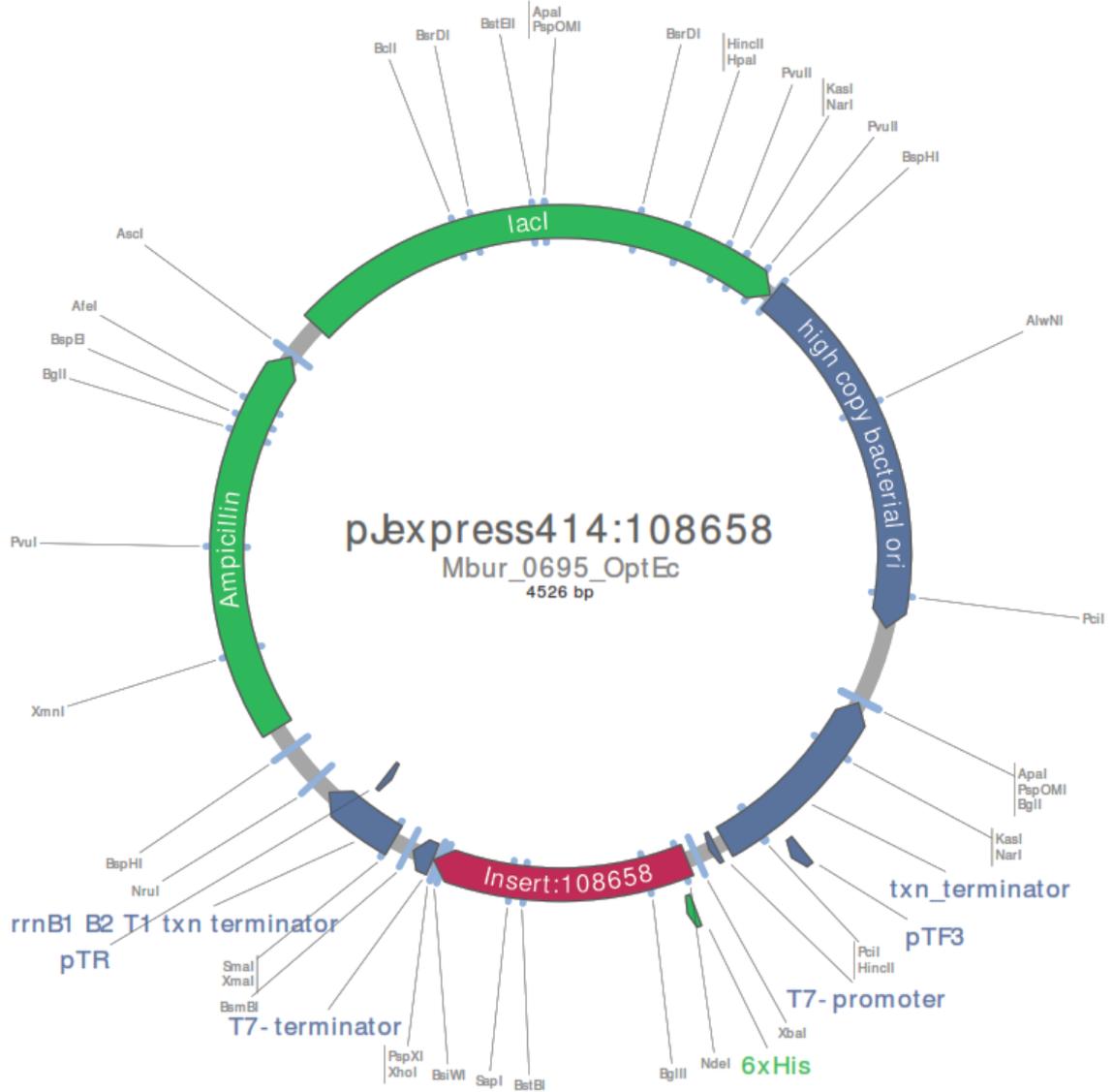
ctatgaaaaacgccagcaacggcctttttacggttcctggccttttctgacacatgttcttctgcttatcccctgattctgtgataacc
gtattaccgctttgagtgagctgataccgctcggcgaccggaacgaccgagcgcagcagtgagcaggaagcggagagcgcctgatgc
ggatftttctcttacgcatctgtgctgatttcacaccgcatatattggtgactctcagtaaatctgctctgatgccgcatagtaagccagataactccg
ctatcgtactgactgggcatggctgcgccccgacaccgccaacaccgctgacgcgcctgacgggctgtctgctccggcatccgcttacaga
caagctgtgaccgtctccgggagctgcatgtgtcagaggtttaccgctcaccgaaacgcgcgaggcagctgaggtaagctcatcagcgtgctc
tgcagcgaattcacagatgtctgctgttcatccgctcagcgtctgtgatttccagaagcgttaattgtctgcttctgataaaggcggccatgtaagg
cggttttctgtttgctactgatgcctccgtgaagggggatttctgttcatgggggtaatgataccgatgaacgagagaggatgctcacgatacgggt
tactgatgatgaacatgccgggtactggaacgtgtgagggtaaacactggcggatggatgcggcgggaccagagaaaaatcactcagggtaaat
ccagcgtctcgtaatacagatgtagggttccacagggtagccagcagcatctgcgatgcagatccggaacataatggtgcaggcgcgtgactccg
cgttccagactttacgaaacggaaaccgaagaccattcatgttggctcaggtcgcagacgtttgcagcagcagctcgttccagctcgcgtat
cgggtgattcattctgtaaccagtaaggcaacccccgccagctgcccggctcctcaacgacagggcagcagatcgtatgctatgccccgcgccacc
ggaaaggagctgactgggtgaaggctctcaaggcagctgctcagatccggctcaatgagtgagtaacttacattaattgctgtgctcactgcc
cgctttcagctgggaaacctgtcgtgccagctgcattaatgaatggccaacgcgcgggagagggcgtttgctgattggcgccagggtggttttct
ttaccagtgagagggcaacagctgattgcccctcaccgctggcctgagagagttgcagcaagcggccacgctggttccccagcagggcga
atcctgtttgatgggtggtaacggcgatataacatgagctgtctcggatcgtcgtatcccactaccgagatgctccgaccaacgcgcagcccgact
cggtaatggcgcgcatggcccagcgcctatctgctgttggcaaccagcagcagtgaggaaacgatcccctattcagcattgcatggtttgtaaaa
ccggacatggcactccagctccttcccgttccgctatcgggtgaattgattgcagtgagatattatgccagccagccagacgcagacgcgcgaga
cagaactaatggcccgtaaacgcgcgatttctggtgacctaatgcgaccagatgctccaccccagctcgcgtaccgtcttcatgggagaaaaat
actgtttaggggtgctgctcagagacatcaagaataacgccggaacattagtcagcagcagcttccacagcaatggcatcctgtcattccagcggatg
ttaatgatcagcccactgacgcgttgcgcgagaagattgtgcaccgcttctacagcttcgacgcgcttcttaccatcagaccaccagcgtgg
caccagttgacggcgcgagatttaacgccgcgacaatttgcgacggcgcgtgcagggccagactggagggtggcaacgccaatcagcaacgactg
ttgcccgccagttgtgtgccacgcggttgggaatgaattcagctccgcatcgcgcttccacttttcccgcgttttcgagaaacgtgctgcccgtg
ttaccacgcgggaaacggctgataagagacaccggcactctgcgacatctataacgttactggtttcacattcaccacctgaattgactctctcc
ggcgctatcatgccataccgcaaggttttgcgccattcgtatgttccgggactcgcgctctccctatgcgactcctgattaggaagcagccca
gtagtaggtgagccgttgagcaccgccgcaaggaatgtgcatgcaaggagatggcggccaacagctccccggccagggcctgcccacca
taccacgccgaaacaagcgtcatgagcccgaagtggcgagcccgatctcccatcgggtgatgctggcgatagggccagcaaccgacactgt
ggcgccggatgcccggccagatgctccggcgtagaggatcagatcagatcgcgacccgcgaaattaacgactcactataggggaattgtgag
cgataacaattcccctctagaataattttgttaacttaagaaggagatatacatatgcccctatactaggtattggaataaaggccttgcgaacc
cactcacttctttgaaatcttgaagaaaaatgaagagcatttctatgagcgcgatgaagggtgataaatggcgaacaaaaagttgaattgggttg
gatttcccaatcttcttattatattgatggtgatgtaaaftaacagctatgcccacatagcttatatagctgacaagcacaacatggtgggtggtg
caaaagagcgtgcagagatttcaatgcttgaaggagcgggttggataftagatacgggtttcagaaatgcatatagtaaaagacttgaactctcaaagt
gattttctagcaagctactgaaatgctgaaatggttaagatcgtttatgataaaacatatttaaatggtgatcatgtaaccatcctgacttcatgtt
gacgctctgatgttgtttatacatggaccatgtgcctggtgcttcccaaatagttgtttaaaaaacgtattgaagctatccacaaaatgataagt
acttgaatccagcaagtatatagcatggccttgcagggctggcaagccacgttgggtggtggcgaccatcctccaaaatcgatcgttccgcgttcg
ggcgaaaacctgtatttccagggc

A2.2 Plasmid sequence of pJexpress404_LtrR: The highlighted capital sequence represents the LtrR gene

ctggctggctggcataaatatctcactcgaatcaaatcagccgatagcggaaacgggaaggcactggagtgccatgtccggtttcaacaaccatgc
aaatgctgaatgaggcactgftcccactgcgatgctggttccaacgatcagatggcgtggcgcaatgcgcgcaattaccgagtcgggctgcgc
gttggctcggatctcggtagtgggatacagcgaaccgaagatagctcatgttatatcccggcgttaaccacatcaaacagatttccgctgctggg
gcaaacagcgtggaccgttctgcaactctcagggccagcgggtgaagggaatcagctgttccagctcactggtgaaaagaaaaaccacc
tggcgccaatacgaaccgctctccccgcgcttggccgattcattatcagcgtggcacgacaggtttcccactgaaagcgggcaagtactca
tgacaaaatccctaacgtgattacgcgcgctgttccactgagcgtcagaccccgtagaaaaatcaaaaggatcttctgagatcctttttctgcgc
gtaatctgctgcttcaacaaaaaacaccgctaccagcgggtgttgggttccggatcaagagctaccaactcttttccgaaagtaactggttccagc
agagcagatcaaaaactgttcttctagttagccgtagttagccaccactcaagaactctgtagcaccgctacatacctcgtctgtaaatcctgt
taccagtgctgctccagtggtgataagctgtcttaccgggtggactcaagcagatgtagcggataagggcagcggctgggctgaacgggg
ggtctgacacagcccagcttgagcgaacgacctaccgaaactgagatacctacagcgtgagctatgagaaagcggccacgctcccgaaggga
gaaagcggacaggtatccgtaagcggcagggctggaacaggaagcgcagcagaggagcttccagggggaacgcctggtatcttatagtcctgt
cggggttcgccactctgactgagcgtcattttgtgatgctcagggggggcggagcctatgaaaaacgccagcaacgcggccttttacggttc

tgcccttttctggccttttctcacatgttcttctcgttatccccctgattctgtggataaccgtattaccgctttgagtgagctgataccgctcgcgcgag
ccgaacgaccgagcgcagcagtgagcaggaagcggaggcagagtagggaactgccaggcatcaaactaagcagaagggccctgacg
gatggccttttgcgtttctacaaactcttctgtgtttaaagcagcggcagcttaagctcgggccccctgggcggttctgataacgagtaacgttaac
cgcaataacgtaaaaaccgcttcggcgggtttttatggggggagtttaggaaaagagcatttgcagaatattaaaggcgcctgtcactttgctgat
atatgagaattatataacctataaatgagaaaaagcaacgcactttaataagatacgttcttttcgattgatgaacacctataataaactattcatctatta
tttatgatttttctatatacaatatttctagttttaaagagaatfaagaaaataaatctcgaataataaagggaatacagttttgatatacaaaattatacat
gtcaacgataatacaaaaataatacaaaactataagatgttatcagattattatcatttagaataaattttgtcgcacctccgcaaaataatacagactcat
ataggggaattgtgagcggataacaattcccctctagaataattttgttaacttttagggaggtaaaacatagcaccatcatcaccaccac **ATGGAA**
CGCGAAAACTGGAGAAAGACATTATGCGCTTGCTGGAGCAGCAGCCGGAGATCTCTGCTA
AGGGTATTGCCGACCGCCTGAGCGTGAGCGAGGATATCGTTACGCGTACCATTGAGAGCCTG
TCCGATACCCGTCATAAAGTCTTGATCGTGGACGATGAGCCGGATGCGGTCATCGCAACGAA
GCGTGCGCTGGAAGCGGACGGCTACAACGTTATTGAGGCCAATAACGGTACTATGGCATTG
ACGCGCTGAAAAGCGATATCCCGGATGTTATTCTGCTGGACGTGATGATGCCGGACATGGAT
GGTTTCGAAGTGTGCCGTCGTTTGAAAGAAGACCGCTGTATGAGAATATCCCGGTCATCAT
GCTGACCGGAAGGGTGAATCAATGACAAAGTAGAGGGCCTGGATATTGGCGCAGACGAC
TACATGACCAACCATTC AACCTGAAAGAACTGAAGGCTCGTATTAAGACCGTTCTGCGCCG
TACGCAAGATTA ACTCGAGccccctagcataacccttggggccttaaacgggtcttgagggggtttttgcccttgagacgcgtcaat
cgagtctgactcaagggcgaccccccaattagcccgggcgaaaggccagcttctgactgagccttctgtttattgatcgctggcagttccctactc
tcgatggggagtcaccactaccatggcgctacggcgttctactctgagttcggcatggggtcaggtgggaccaccgctactgcccaggc
aaacaaggggtgtatgagccatattcaggtataaatgggctcgcgataatgtcagaattggttaattggtgtaacactgaccctattgtttttctaa
tacattcaaatatgatccgctcatgagacaataaccctgataaatgcttcaafaattgaaaaaggaagaatagattcaacattccgtcgcctta
ttccctttttgcgccattttccctctgttttctcaccagaaacgctggtgaaagtaaaagatgctgaagatcagttgggtgcacgagtggttacatc
aactggatctcaacagcggtaagatcctgagagtttcccccgaagaacgtttccaatgatgagcacttttaagttctgctatgtggcgcggtattatcc
cgtattgacgccggcaagagcaactggcgcgcatacactattctcagaatgacttgggtgagttaccagtcacagaaaagcatcttacggatg
gcatgacagtaagagaattatgcagtgctccataacctgagtgataaactcggccaacttactctgacaacgatcggaggaccgaaggagctaa
ccgctttttgcacaacatggggatcatgtaactcgccttgatcgttgggaaccggagctgaatgaagccatacaaacgacgagcgtgacaccacgat
gcctgtagcagtggaacaacgttgcgcaactataactggcgaacttactctagcttcccggcaacaattaagactggatggaggcggataaaa
gttgacaggaccacttctcgcctcggccttccggctggctggtttattgctgataaatccggagccggtagcgtggttctcgcggtatcatcgacgcgt
ggggccagatgtaagccctcccgtatcgtatctacacgacggggagtcaggcaactatggatgaacgaaatagacagatcgtgagatagggtg
cctcactgattaagcattgtaagcggcgcgccatcgaatggcgcaaaccttctcgggtatggcatgatagcgccggaagagatcaattcagggtg
gtgaatafgaaaccagtaacgttatacagatgctgcagagatgccggtgtctctatcagaccgttcccgcgtggtgaaccaggccaccacgttctcgc
aaaacgcgggaaaaagtgaagcggcagtgccggagctgaattacattccaaccgctggcacaacaactggcggcaaacagctgtgctgattg
gcgttgcacctcagctcggcctgcacgcgccgtcgcaattgtcgcggcgattaaatctcgcgccgatcaactgggtgccagcgtggtggtcgtgat
ggtagaacgaagcggcgtcgaagcctgtaagcggcgggtgcacaatcttctcgcgcaacgcgtcagtgggctgatcattaactatccgctggatgacca
ggatgccattgctgtgaaagctgcctgactaatgttccggcgttatttctgatctctgaccagacaccatcaacagattattttctccatgaggacg
gtacgcgactggcggtggagcatctgtcgcattgggtcaccagcaaatcgcgctgttagcgggcccaatagttctgtctcggcgcgtctcgcg

A2.3 Plasmid map of pJexpress404_LtrR



Appendix 3: Primer sequences and PCR cycling parameters

Table A1: Primer sequences for mutagenesis reactions

Name	Oligonucleotide sequence from 5' to 3'
H367R_for	ATGGTGAGCCGTGAGCTGAAAACC
H367R_rev	GCTGACAAAATCATTTCCTCAAACGATC
H367A_for	ATGGTGAGCGCGGAGCTGAAAACC
H443R_for	TTAGCCTGCGTAATGTTACCAAACATG
H443R_rev	TCTCTATTATCTCTTTGATGTTTACATCTTCC
H448R_for	TTAGCCTGCATAATGTTACCAAACGTGCGA
H502R_for	GATTGTGGAAGATCGTGAAGATTACTTAC
H502R_rev	ACGGTCACTTCGCCTTCTTCTGG
H443/448R_for	TTAGCCTGCGTAATGTTACCAAACGTGCGA
H443/448A_for	ACCAAAGCGGCGAAGGATAAG
H443/448A_rev	AACATTCGCCAGGCTAATCTCTAT
D54N_for	TTGATCGTGAACGATGAGCCG
D54N_rev	GACCTTATGACGGGTATCGGACA
D55N_for	TTGATCGTGGACAATGAGCCG
D98N_for	TTATTCTGCTGAACGTGATGATGC
D98N_rev	CATCCGGGATATCGCTTTTCA
LtrR*_for	ACCCGTCATAAGGTCTTGATCG
LtrR*_rev	CATGTGGTGGTGATGATGGTG

Table A2: PCR cycling parameters for mutagenesis reactions

Mutation site	Substituted with	Primer pair for PCR reaction	PCR reaction cycle
In LtrK : H367	R	Forward : H367R_for Reverse: H367R_rev	98 °C – 30s × 1 cycle; 98 °C- 10s, 68 °C- 30s, 72 °C- 3 min × 25 cycle; 72 °C- 1 min × 1 cycle
	A	Forward : H367A_for Reverse: H367R_rev	

H443	R	Forward : H443R_for Reverse: H443R_rev	98 °C – 30s × 1 cycle; 98 °C- 10s, 65 °C- 30s, 72 °C- 3 min × 25 cycle; 72 °C- 1 min × 1 cycle
H448	R	Forward : H448R_for Reverse: H443R_rev	98 °C – 30s × 1 cycle; 98 °C- 10s, 65 °C- 30s, 72 °C- 3 min × 25 cycle; 72 °C- 1 min × 1 cycle
H502	R	Forward : H502R_for Reverse: H502R_rev	98 °C – 30s × 1 cycle; 98 °C- 10s, 64 °C- 30s, 72 °C- 3 min × 25 cycle; 72 °C- 1 min × 1 cycle
H443 + H448	R	Forward : H443/448R_for Reverse: H502R_rev	98 °C – 30s × 1 cycle; 98 °C- 10s, 65 °C- 30s, 72 °C- 3 min × 25 cycle; 72 °C- 1 min × 1 cycle
	A	Forward : H443/448A_for Reverse: H443/448A_rev	
In LtrR: D54	N	Forward : D54N_for Reverse: D54N_rev	98 °C – 30s × 1 cycle; 98 °C- 10s, 68 °C- 30s, 72 °C- 2 min × 25 cycle; 72 °C- 1 min × 1 cycle
D55	N	Forward : D55N_for Reverse: D54N_rev	
D98	N	Forward : D98N_for Reverse: D98N_rev	
LtrR* (HTH domain deleted)	-	Forward : LtrR*_for Reverse: LtrR*_rev	98 °C – 30s × 1 cycle; 98 °C- 10s, 67 °C- 30s, 72 °C- 2 min × 25 cycle; 72 °C- 1 min × 1 cycle

H = Histidine, R = Arginine, A = Alanine, D = Aspartate, N = Asparagine