Structural and mechanistic insights into the Rel protein of Mycobacterium tuberculosis

Singal, Bharti

2020

Singal, B. (2020). Structural and mechanistic insights into the Rel protein of Mycobacterium tuberculosis. Doctoral thesis, Nanyang Technological University, Singapore.

<https://hdl.handle.net/10356/142669>

<https://doi.org/10.32657/10356/142669>

This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License (CC BY-NC 4.0).

Downloaded on 26 Apr 2023 18:58:47 SGT

**NANYANG
TECHNOLOGICAL
UNIVERSITY**

SINGAPORE

**STRUCTURAL AND MECHANISTIC INSIGHTS
INTO THE REL PROTEIN OF *MYCOBACTERIUM*
*TUBERCULOSIS***

**SINGAL BHARTI
SCHOOL OF BIOLOGICAL SCIENCES
2020**

**STRUCTURAL AND MECHANISTIC INSIGHTS
INTO THE REL PROTEIN OF *MYCOBACTERIUM*
*TUBERCULOSIS***

SINGAL BHARTI

School of Biological Sciences

A thesis submitted to the Nanyang Technological University in
partial fulfillment of the requirement for the degree of Doctor of
Philosophy

2020

Statement of Originality

I hereby certify that the work embodied in this thesis is the result of original research done by me except where otherwise stated in this thesis. The thesis work has not been submitted for a degree or professional qualification to any other university or institution. I declare that this thesis is written by myself and is free of plagiarism and of sufficient grammatical clarity to be examined. I confirm that the investigations were conducted in accord with the ethics policies and integrity standards of Nanyang Technological University and that the research data are presented honestly and without prejudice.

13/01/2020

Date

Singal Bharti

Supervisor Declaration Statement

I have reviewed the content and presentation style of this thesis and declare it of sufficient grammatical clarity to be examined. To the best of my knowledge, the thesis is free of plagiarism and the research and writing are those of the candidate's except as acknowledged in the Author Attribution Statement. I confirm that the investigations were conducted in accord with the ethics policies and integrity standards of Nanyang Technological University and that the research data are presented honestly and without prejudice.

13/01/20
.....
Date

.....
Prof. Dr. Gerhard Grüber

Authorship Attribution Statement

This thesis contains content from one paper published in *The FEBS Letters* and the two manuscripts in preparation where I will be the co-author.

Chapters 3.2 is published as Singal, B., Balakrishna, A.M., Nartey, W., Manimekalai, M.S.S., Jeyakanthan, J. and Grüber, G. (2017) Crystallographic and solution structure of the N-terminal domain of the Rel protein from *Mycobacterium tuberculosis*. *FEBS Lett*, 591, 2323-2337.

The contributions of the co-authors are as follows:

- Prof. Gerhard Grüber provided the initial project direction and the direction of the manuscript.
- I, together with Prof. Gerhard Grüber, Dr. Asha Balakrishna, Dr. Wilson Nartey, Dr. Malathy Sony Subramanian Manimekalai wrote the manuscript. Dr. Jeyaraman Jeyakanthan and Prof. Gerhard Grüber helped in editing the manuscript.
- Cloning of the *MtRel* NTD was done by Dr. Nartey. Isolation of the proteins and purification of the protein was done by me. Purified protein required for all the experiments was provided by me.
- The HPLC experiments were performed by me and Dr. Nartey. All the protein samples used in the experiment at several time points were provided by me.
- Samples for SAXS experiments were prepared by me whereas SAXS data collection was done by Dr. Manimekalai. The SAXS data were analyzed by Dr. Manimekalai and me.
- Crystallization set up was performed by me and Dr. Asha Balakrishna. The crystals' freezing, shipment, remote data collection at Taiwan synchrotron from BII facility at Biopolis was done by me and Dr. Balakrishna. The data were processed by Dr. Balakrishna, whereas model building, refinement, structure calculations and validation were performed by me under the guidance of Dr. Balakrishna.

The contributions to the unpublished Chapter 3.3.3 are as follows:

- The *MtRel* TGS was cloned, expressed and purified to high purity by me.
- I produced and purified the single and double labelled proteins for all NMR spectroscopy studies. The samples were also prepared by me.
- The NMR data was collected and processed by Dr. Shin Joon. The 2D prediction and 3D structural calculations were also performed by him.
- Sample preparation for tRNA titration was done by me whereas the data was collected, processed and analyzed by Dr. Shin.

The contributions to the unpublished Chapter 3.4 are as follows:

- Prof. Gerhard Grüber provided the initial project direction and the direction of the manuscript.
- I, together with Prof. Gerhard Grüber and Dr. Shin wrote the manuscript.
- The *MtRel* ACT construct was cloned, expressed, produced and purified by me.
- I have generated the point mutants of *MtRel* ACT: -H12A, -L15A, -A23S and -N27A. Also, I have done the expression, production and purification of the mutant forms of *MtRel* ACT.
- The ^{15}N labeled, $^{15}\text{N}^{13}\text{C}$ labeled, selectively labeled with alanine (A), valine (V) and arginine (R), and the samples titrated with amino acids and RNA were produced and purified by me. The NMR data collection was done by Dr. Shin whereas the assignment of chemical shifts for secondary structure prediction was done under the guidance of Dr. Shin. The 3D structure calculations and validation was completely done by Dr. Shin.
- The SAXS experimental data was collected by Dr. Manimekalai and the analysis was done together with her.

13/01/2020

.....
Date

.....
Singal Bharti

Acknowledgments

First and foremost, I would like to express my sincere gratitude to my advisor Prof. Dr. Gerhard Grüber for his continuous support and encouragement in the four years of my PhD. Without his immense knowledge, guidance and support, this thesis would have not been possible. His approach to see scientific problems from different directions has greatly influenced my scientific thought process. I am thankful to him for giving me an opportunity to join his group and providing this learning experience.

I would also like to thank Ministry of Education, Singapore for the funding support provided to Prof. Gerhard Grüber in the form of grant, MoE Tier 1 (Rg137/15), which aided in the studies performed in this thesis. I would also like to acknowledge NTU research scholarship provided as research stipend in the course of my PhD which helped me to support myself in Singapore.

I would like to thank Dr. Wilson Nartey for training me in molecular biology techniques. I would like to thank Dr. Arvind Kumar for devoting his time in reading this thesis and discussing small to big issues related to PhD life and project and motivating me to dream big. I am extremely thankful to Dr. Asha M. Balakrishna for sharing her extraordinary knowledge with me while guiding through protein crystallization and structure determination. I learnt a lot from her while doing data collection and processing remotely at the National Synchrotron Radiation Research Centre (NSRRC, Hsinchu, Taiwan). I am ever grateful to Dr. Dharendra Singh with whom I spent numerous hours in discussing science and life. His invaluable suggestions helped me during my PhD confirmation.

I am also grateful to Dr. Malathy Sony S. Manimekalai for her invaluable knowledge during our discussion, and her guidance in collecting and processing the SAXS data for the *MtRel* protein. I would further like to thank Dr. Neelagandan Kamariah for guiding me through basic strategies to do effective literature survey and Dr. Saw Wuan Geok for her enthusiasm in teaching me protein purification strategies and giving feedback on my thesis. I am grateful to Dr. Priya Raghunathan for being there in stressful times. In addition, I gratefully acknowledge Dr. Shin Joon for providing the NMR support in the two chapters of my thesis.

Furthermore, I owe special thanks to Dr. Ardina Grüber, and my other colleagues, Ms Chong Shi Min Sherilyn and Ms. Chui Faan for keeping the pleasant

working environment in the lab. I would like to express my gratitude to Dr. Ankita Pan and Dr. Nebojša Bogdanović for their support, constant encouragement and for being my happy space. I would also like to thank my friends from NTU and NUS for sharing their life experiences and making me understand different cultures from different parts of the world.

I am sincerely grateful to my mother Seema Gupta and my sister Niharika Singal for always believing in me and providing selfless support, reinforcement and unconditional love. I am also grateful to my loving husband Yashvir Singh Tomar who has brought so much love and encouragement in my life. Last but not the least, I would like to express my eternal gratitude to my late father Ashok Kumar, whose teachings has inspired me to accept every challenge courageously and to never stop dreaming. I dedicate all my success to these four pillars of my life and graciously accept all my failures as a part of the learning experience.

Dedicated to My parents, My sister and My husband

Table of Contents

Title Page	I
Statement of Originality	II
Supervisor Declaration Statement	III
Authorship Attribution Statement	IV
Table of Contents	IX
List of Figures	XIII
List of Tables	XVI
Summary	XVII
1 Introduction	1
1.1 Overview of tuberculosis	2
1.2 Etiology of TB – from the transmission to pathogenesis	5
1.3 Persistence and LTBI	8
1.4 The stringent response	10
1.5 Effector of stringent response - (p)ppGpp	11
1.6 <i>In vivo</i> concentration fluctuations of (p)ppGpp leads to dormancy	12
1.7 Pleiotropic effects of (p)ppGpp	13
1.7.1 Regulation of transcription via RNA polymerase and the non-template DNA binding proteins	13
1.7.1.1 Binding sites of (p)ppGpp on RNAP	14
1.7.2 Indirect regulation of transcription	16
1.7.3 (p)ppGpp and DNA replication	16
1.7.4 (p)ppGpp and its interaction with translation factors	16
1.7.5 (p)ppGpp and virulence	17
1.8 Regulator of (p)ppGpp – RelA/SpoT homolog superfamily proteins	17
1.9 Mechanism behind the accumulation of stringent factor	19
1.10 Understanding of stringent effector regulation by RelA and SpoT	21
1.11 Metabolism of (p)ppGpp by bifunctional Rel protein- <i>SeRel</i>	24
1.12 Significance to study stringent response in <i>Mtb</i>	26
1.13 Discovery of <i>MtRel</i> and its significance in persistent bacteria	27
1.14 Organization of N-terminal catalytic domains	29
1.15 Organization of C-terminal regulatory domains	30
1.16 Role of oligomerization by CTDs in regulation of activity	33
1.17 Goals of the thesis	34

2	Materials and Methods	36
2.1	Materials	37
2.1.1	Chemicals	37
2.1.2	Molecular biology materials	37
2.1.2.1	Enzymes and molecular biology kits	37
2.1.2.2	Primers, synthetic oligonucleotides and plasmids	38
2.1.2.3	<i>E. coli</i> strains	38
2.1.3	Chromatography materials	39
2.1.4	Other instrumentation	39
2.1.5	Computer software	40
2.2	Methods	41
2.2.1	Generation of the plasmid pET9d: <i>MtRel</i> , and pET9d: <i>MtRel</i> ₁₋₃₉₄	41
2.2.2	Generation of the plasmid pET-9d: <i>MtRel</i> ₆₆₀₋₇₃₈	41
2.2.3	Site-directed mutagenesis of <i>MtRel</i> ₆₆₀₋₇₃₈	44
2.2.4	Generation of the <i>MtRel</i> ₃₉₅₋₇₃₈ construct	47
2.2.5	Generation of the <i>MtRel</i> ₃₉₄₋₄₅₉ construct	47
2.2.6	Generation of the <i>MtRel</i> ₁₋₄₅₉ construct using mutagenesis	49
2.2.7	Recombinant expression of <i>MtRel</i> and truncated constructs	50
2.2.7.1	Expression of <i>MtRel</i> , <i>MtRel</i> ₁₋₃₉₄ , <i>MtRel</i> ₆₆₀₋₇₃₈ and <i>MtRel</i> ₃₉₅₋₇₃₈ in <i>E. coli</i> BL21 (DE3) strain	50
2.2.7.2	Expression of <i>MtRel</i> , <i>MtRel</i> ₃₉₄₋₄₅₉ , <i>MtRel</i> ₁₋₄₅₉ and <i>MtRel</i> ₆₆₀₋₇₃₈ wild type and its mutants in <i>E. coli</i> C41 (DE3) strain	50
2.2.7.3	Expression of <i>MtRel</i> in <i>E. coli</i> BL21 GroEL-ES strain	51
2.2.8	Isotopic labelling of proteins	51
2.2.8.1	Minimal Media (M9) preparation	51
2.2.9	Production of recombinant <i>MtRel</i> and truncated constructs	52
2.2.9.1	Induction test	52
2.2.9.2	Solubility test	53
2.2.10	Large scale production of recombinant <i>MtRel</i> and truncated constructs	53
2.2.10.1	Production of <i>MtRel</i> , <i>MtRel</i> ₁₋₃₉₄ , <i>MtRel</i> ₃₉₄₋₄₅₉ , <i>MtRel</i> ₁₋₄₅₉ , <i>MtRel</i> ₃₉₅₋₇₃₈ and <i>MtRel</i> ₆₆₀₋₇₃₈ in <i>E. coli</i> BL21 DE3 or C41 DE3 cell strains	53
2.2.10.2	Production of <i>MtRel</i> in <i>E. coli</i> BL21 GroEL-ES cell strain	54
2.2.11	Purification of <i>MtRel</i> constructs	54
2.2.12	Analysis of synthetase activity of <i>MtRel</i> and constructs using luminescence assay	55

2.2.13	Activity analysis of the <i>MtRel</i> ₁₋₃₉₄ using HPLC	57
2.2.14	Crystallization of <i>MtRel</i> ₁₋₃₉₄	58
2.2.14.1	Data collection and structure determination of <i>MtRel</i> ₁₋₃₉₄	59
2.2.14.2	Crystal structure determination of <i>MtRel</i> ₁₋₃₉₄	60
2.2.15	Small angle X-ray scattering for low resolution studies	61
2.2.15.1	SAXS data collection	61
2.2.15.2	SAXS data analysis	62
2.2.16	Introduction to nuclear magnetic resonance (NMR) spectroscopy and its use in obtaining protein structure	63
2.2.16.1	Preparation of NMR samples	64
2.2.16.2	NMR data collection and processing	65
2.2.16.3	Resonance assignment	66
3	Results	67
3.1	Purification and biophysical characterization of <i>MtRel</i>	68
3.1.1	Introduction	69
3.1.2	Expression and purification of <i>MtRel</i>	69
3.1.2.1	Optimization of buffer to improve solubility using cells from <i>E. coli</i> BL21(DE3)	69
3.1.2.2	Optimization of the expression system to improve solubility	71
3.1.3	Existence of <i>MtRel</i> in multi-oligomeric forms in solution	73
3.1.3.1	Solution structure of the <i>MtRel</i> using SAXS	73
3.1.4	Determination of enzymatically active <i>MtRel</i>	77
3.2	Crystallographic and solution structure of <i>MtRel</i> NTD	79
3.2.1	Introduction	80
3.2.2	Production and purification of <i>MtRel</i> NTD	80
3.2.3	Confirming the presence of enzymatically active protein in solution	81
3.2.3.1	HPLC studies of catalytically active <i>MtRel</i> NTD	81
3.2.3.2	Luminescence based detection of synthetic activity of <i>MtRel</i> NTD	83
3.2.4	Crystallization and data collection of <i>MtRel</i> NTD crystals	84
3.2.5	Crystal Structure of <i>MtRel</i> NTD	87
3.2.6	Intermolecular analysis of <i>MtRel</i> NTD	89
3.2.7	Solution structure studies by SAXS	90
3.2.7.1	Dimeric solution structure of the <i>MtRel</i> NTD	90
3.2.7.2	SAXS studies of <i>MtRel</i> NTD with substrates	94

3.3	Deciphering the structural features of the tRNA binding regulatory domain, the <i>MtRel</i> TGS	97
3.3.1	Introduction	98
3.3.2	Cloning, generation and purification of <i>MtRel</i> CTD, <i>MtRel</i> TGS and <i>MtRel</i> NTD-TGS	98
3.3.3	NMR solution structure determination for <i>MtRel</i> TGS	102
3.3.4	Molecular interaction between <i>MtRel</i> TGS with deacylated tRNA	105
3.3.5	Understanding the effect of the TGS domain on synthetase activity	108
3.4	Unravelling the allosteric feedback mechanism exhibited by the <i>MtRel</i> ACT	109
3.4.1	Introduction	110
3.4.2	Cloning, production and purification of <i>MtRel</i> ACT	110
3.4.3	NMR Solution structure of dimeric <i>MtRel</i> ACT	112
3.4.4	Dimeric interface of <i>MtRel</i> ACT	116
3.4.5	Molecular interaction between <i>MtRel</i> ACT and single stranded RNA	117
3.4.6	<i>MtRel</i> ACT selectively interacts with branched amino acid	118
3.4.7	Low-resolution solution structure of the <i>MtRel</i> ACT	123
3.4.8	Characterization of the dimeric interface and amino acid binding cleft	126
3.4.9	NMR analysis of the <i>MtRel</i> ACT mutants H12A, L15A, A23S and N28A	128
4	Discussion	130
4.1	<i>MtRel</i> prefers forming homo-dimer in solution	131
4.2	<i>MtRel</i> NTD shows species specific differences to the non-mycobacterial bifunctional long RSH	133
4.3	The significance of understanding regulatory domains of <i>MtRel</i>	138
4.4	Novel <i>MtRel</i> TGS exhibits conserved structure across RSH homologs and interacts with tRNA	141
4.5	<i>MtRel</i> ACT exists as dimer in solution	143
4.6	The <i>MtRel</i> ACT dimer shows biological relevance understood on the basis of the <i>EcRelA</i> complex with stalled ribosomal complex	145
4.7	<i>MtRel</i> ACT dimer interacts with branched chain amino acid	146
5	Conclusions	151
6	References	157
7	Appendix	174

List of Figures

Figure 1.1: Estimated TB incidences in 2018.	3
Figure 1.2: Etiology of TB.	6
Figure 1.3: Dynamics of granuloma formation.	8
Figure 1.4: Chemical structure of GDP and GTP.	11
Figure 1.5: Chemical structure of ppGpp and pppGpp.	12
Figure 1.6: Pleiotropic effects of (p)ppGpp.	14
Figure 1.7: Regulation of RNA polymerase activity by (p)ppGpp.	15
Figure 1.8: Schematic representation of the RelA/SpoT homologs (RSH) domain structure.	18
Figure 1.9: Mechanism of stringent response by <i>E. coli</i> RelA.	20
Figure 1.10: Model of RelA activation.	21
Figure 1.11: Generalized schematic showing biosynthetic pathway of ppGpp.	23
Figure 1.12: The structure of the complete ribosome 70S from <i>E. coli</i> .	24
Figure 1.13: Crystal structure of Rel from <i>S. equisimilis</i> (SeRel).	25
Figure 1.14: Functional characterization of the MtRel.	27
Figure 1.15: Sequence alignment of TGS domain from mycobacterial and non-mycobacterial species.	31
Figure 1.16: Sequence alignment of ACT domains from mycobacterial and non-mycobacterial species.	32
Figure 2.1: Schematic diagram of site-directed mutagenesis.	46
Figure 2.2: Schematic diagram of protein purification protocol.	55
Figure 2.3: Schematic diagram showing luminescence based AMP glo Assay.	56
Figure 2.4: Flow chart for determining the activity of MtRel ₁₋₃₉₄ .	57
Figure 2.5: Schematic diagram of crystallization optimization strategies.	59
Figure 2.6: Scheme used to solve the crystal structure of MtRel ₁₋₃₉₄ .	60
Figure 3.1: Purification of MtRel from <i>E. coli</i> BL21 (DE3).	71
Figure 3.2: Purification of MtRel from <i>E. coli</i> GroEL-ES and C41 (DE3) strains.	72
Figure 3.3: SAXS data analysis of MtRel.	73
Figure 3.4: Normalized Kratky plot of MtRel.	74
Figure 3.5: <i>Ab-initio</i> envelope reconstruction of MtRel.	75
Figure 3.6: Luminescence based AMP detection assay to check the synthetase activity of MtRel.	78
Figure 3.7: Purification of MtRel NTD.	81
Figure 3.8: IP-RP HPLC used to confirm the presence of enzymatically active MtRel NTD.	83
Figure 3.9: AMP generation as a result of (p)ppGpp synthetase activity by MtRel and MtRel NTD.	84
Figure 3.10: Initial crystal of MtRel NTD and diffraction pattern.	85
Figure 3.11: Screening of MtRel NTD crystals.	85
Figure 3.12: Optimization of MtRel NTD crystals obtained in Wizard Screen 4 #13.	86
Figure 3.13: Final optimized MtRel NTD crystal in Wizard Screen 4 #13.	86
Figure 3.14: Representation of one asymmetric unit of MtRel NTD crystals.	88
Figure 3.15: A monomer of MtRel NTD.	89
Figure 3.16: Dimer formation by MtRel NTD.	90
Figure 3.17: SAXS studies of MtRel NTD and Pair-distance distribution P(r) function of MtRel NTD generated after merging the three datasets.	91
Figure 3.18: Normalized Kratky plot of MtRel NTD.	93

Figure 3.19: Solution shape of dimer <i>MtRel</i> NTD and SAXS data analysis of crystal dimer fitting to the solution shape of <i>MtRel</i> NTD.	94
Figure 3.20: SAXS studies of <i>MtRel</i> NTD in the absence and presence of substrates.	95
Figure 3.21: PCR amplification of <i>MtRel</i> CTD and <i>MtRel</i> TGS.	99
Figure 3.22: PCR amplification and induction test of <i>MtRel</i> NTD-TGS.	100
Figure 3.23: Purification of <i>MtRel</i> CTD and <i>MtRel</i> TGS.	101
Figure 3.24: Purification of the <i>MtRel</i> NTD-TGS from <i>E. coli</i> C41 (DE3)	102
Figure 3.25: $^1\text{H}^{15}\text{N}$ HSQC spectrum of <i>MtRel</i> TGS.	103
Figure 3.26: NMR spectroscopic secondary structure characterization of <i>MtRel</i> TGS.	104
Figure 3.27: NMR solution structure of the <i>MtRel</i> TGS.	105
Figure 3.28: $^1\text{H}^{15}\text{N}$ -TROSY-HSQC experimental analysis of <i>MtRel</i> TGS titration with deacylated tRNA.	106
Figure 3.29: Molecular interactions of <i>MtRel</i> TGS with tRNA at a molar ratio of 1:2.	107
Figure 3.30: AMP generation as a result of (p)ppGpp synthetase activity by <i>MtRel</i> , <i>MtRel</i> NTD and <i>MtRel</i> NTD-TGS.	108
Figure 3.31: PCR amplification and induction test of <i>MtRel</i> ACT.	111
Figure 3.32: Purification of the <i>MtRel</i> ACT from <i>E. coli</i> C41 (DE3).	112
Figure 3.33: $^1\text{H}^{15}\text{N}$ HSQC spectrum of <i>MtRel</i> ACT.	113
Figure 3.34: Assigned $^1\text{H}^{15}\text{N}$ HSQC- spectrum of <i>MtRel</i> ACT.	114
Figure 3.35: NMR spectroscopic characterization of <i>MtRel</i> ACT.	115
Figure 3.36: NMR solution structure of <i>MtRel</i> ACT.	115
Figure 3.37: Molecular contact between each monomer unit of <i>MtRel</i> ACT.	116
Figure 3.38: A ribbon representation of the molecular contact between each monomeric unit of <i>MtRel</i> ACT based on NOE data.	117
Figure 3.39: <i>MtRel</i> titration with ssRNA:	118
Figure 3.40: Molecular interaction of <i>MtRel</i> ACT and branched chain amino acids based on NMR titration experiments.	119
Figure 3.41: Successive titration of <i>MtRel</i> ACT with valine for identification of interacting residues.	119
Figure 3.42: A plot of changes of peak intensity based on the ratio (I_1/I_0) at a molar ratio of 1:10.	120
Figure 3.43: 2D $^1\text{H}^{15}\text{N}$ HSQC experimental analysis of <i>MtRel</i> ACT titration with valine at an 1:50 molar ratio.	121
Figure 3.44: Mapping of the valine binding surface and conformational changes using NMR data.	122
Figure 3.45: SAXS studies of <i>MtRel</i> ACT.	124
Figure 3.46: Purification of WT <i>MtRel</i> ACT and mutants eluting at (A) 150 mM and (B) 350 mM imidazole.	127
Figure 3.48: Dynamic light scattering analysis of <i>MtRel</i> ACT and the mutant L15A.	128
Figure 3.49: Overlaid $^1\text{H}^{15}\text{N}$ HSQC spectra of WT (red) and mutants (H12A, L15A, A23S and N28A).	129
Figure 4.1: Domain architecture of <i>MtRel</i> .	131
Figure 4.2: Oligomeric state of <i>MtRel</i> in solution.	133
Figure 4.3: Sequence alignment between <i>MtRel</i> NTD, <i>EcRelA</i> NTD, <i>EcSpoT</i> NTD and <i>SeRel</i> NTD and structural alignment between <i>SeRel</i> NTD and <i>MtRel</i> NTD [224].	135

Figure 4.4: A structural alignment between the nucleotide-bound <i>SeRel</i> NTD monomer (PDB: 1VJ7) and <i>MtRel</i> NTD (color coded as in Fig 4.3B) [224].	136
Figure 4.5: Intermolecular interactions between two monomers of <i>MtRel</i> NTD [224].	137
Figure 4.6: Overlapping structure of <i>MtRel</i> NTD and <i>SeRel</i> NTD showing the C-terminal region of the synthetase domain [224].	138
Figure 4.7: Comparison of <i>MtRel</i> and <i>MsRel</i> constructs for the synthetase activity. A pictorial representation of the deletion constructs studied from <i>MtRel</i> and <i>MsRel</i> to understand the effect on catalytic activity.	140
Figure 4.8: Structure of <i>EcRelA</i> TGS domain interacting with deacylated tRNA [161].	141
Figure 4.9: Sequence and structure alignment of <i>MtRel</i> TGS with its homologs [161].	142
Figure 4.10: Sequence alignment of the ACT domain of mycobacterial and non-mycobacterial RSH family proteins with <i>E. coli</i> 3-phosphoglycerate dehydrogenase (<i>EcSerA</i>) and <i>P. furiosus</i> Leucine receptor protein A (<i>PfLrpA</i>).	143
Figure 4.11: Surface representation of dimeric interface based on the NMR data and determined solution structure.	144
Figure 4.12: Inter and intramolecular interactions of the ACT domain of RSHs.	145
Figure 4.13: Structural alignment of <i>MtRel</i> ACT and <i>EcRelA</i> ACT from the cryo EM structure PDB 5IQR.	146
Figure 4.14: Structural comparison of <i>MtRel</i> ACT to a serine bound <i>EcSerA</i> ACT dimer.	147
Figure 4.15: Structural comparison of <i>MtRel</i> ACT with the ACT domain of <i>E. coli</i> .	148
Figure 4.16: The schematic presenting the role of <i>MtRel</i> and its regulatory domain in nutrient starved condition.	149
Figure 5.1: Overall insights into <i>MtRel</i> .	156

List of Tables

Table 2.1: Primers for generation of <i>MtRel</i> and <i>MtRel</i> ₁₋₃₉₄ .	41
Table 2.2: Primers for generation of <i>MtRel</i> ₆₆₀₋₇₃₈ .	41
Table 2.3: Reaction mix for PCR.	42
Table 2.4: PCR cycling program.	42
Table 2.5: Reaction mix for ligation of DNA fragments	43
Table 2.6: Primers for generation of <i>MtRel</i> ₆₆₀₋₇₃₈ mutants.	45
Table 2.7: Reaction mix for site directed mutagenesis PCR.	45
Table 2.8: PCR cycling program for complete plasmid.	45
Table 2.9: Reaction mix for <i>DpnI</i> digestion.	46
Table 2.10: Primers used for the generation of <i>MtRel</i> ₃₉₅₋₇₃₈ .	47
Table 2.11: Primers used for the generation of <i>MtRel</i> ₃₉₄₋₄₅₉ .	48
Table 2.12: Reaction mix used for <i>MtRel</i> ₃₉₄₋₄₅₉ .	48
Table 2.13: PCR cycling program to amplify <i>MtRel</i> ₃₉₄₋₄₅₉ .	48
Table 2.14: Primers used for the generation of <i>MtRel</i> ₁₋₄₅₉	49
Table 2.12: Reaction mix used for <i>MtRel</i> ₃₉₄₋₄₅₉	49
Table 2.13: PCR cycling program to amplify <i>MtRel</i> ₃₉₄₋₄₅₉	49
Table 2.14: Buffers used for the purification of <i>MtRel</i> and constructs	54
Table 3.1: Data collection and scattering derived parameters for <i>MtRel</i>	77
Table 3.2: Data processing, refinement and validation statistics of <i>MtRel</i> NTD crystals at the synchrotron.	87
Table 3.3: Data collection and scattering derived parameters for <i>MtRel</i> NTD	93
Table 3.4: SAXS data of <i>MtRel</i> NTD in the presence and absence of nucleotide substrates.	96
Table 3.5: Data collection and scattering derived parameters for <i>MtRel</i>	126
Supplementary Table S1: MS+MS/MS data of the <i>MtRel</i> band of lane 2 in Figure 3.2D inset.	175

Summary

Mycobacterium tuberculosis (*Mtb*) is known to survive as slow-growing bacteria. However, under in-hostile conditions inside the human host, it can survive as persister. The persistent bacteria can stay dormant for months to years and hence doesn't show any progression from latent infection to active disease development. During latent infection, *Mtb* is challenged by the host immune system in multiple ways which includes, hypoxia, acidification and nutrient starvation. To survive through these host responses, *Mtb* responds in an orchestrated manner. The elaborate machinery to withstand nutrient starvation enables the pathogen to survive, fatty acid limitation, carbon, iron and amino acid starvation. Amino acid starvation leads to stringent response which is understood in many bacteria including *Mtb*. However, the mechanism behind the regulation of stringent response independent to ribosomal complex have not been understood clearly. The stringent response is capable of changing gene expression at a global scale by means of stringent effector molecules which are a hyper-phosphorylated form of guanosine, ((p)ppGpp). The accumulation of (p)ppGpp is crucial to trigger stringent response which is regulated by proteins of the RelA/SpoT Homologous (RSH) family. *Mtb* shows a single functional copy of the *relA* gene which encodes the bifunctional protein *MtRel*, responsible to synthesize and hydrolyse the effector molecules. The *MtRel* is the only regulator of (p)ppGpp and can serve as a potential target for therapeutic interventions against latent tuberculosis infection as shown by *relA* knockout studies *in-vitro* and *in-vivo* as well as for drug resistant TB.

In this thesis, the *MtRel* protein has been characterized structurally and mechanistically. The work has been subdivided into the framework that can be summarised as follows:

- Purification and biophysical characterization of the *MtRel*. The small angle X-ray scattering (SAXS) studies highlighted the coexistence of monomer-dimer-trimer of the catalytic active protein. The limitation to concentrate the protein without aggregation encouraged to pursue domain-wise characterization of *MtRel*.
- Structure determination and identification of potential epitopes in the enzymatically active recombinant fragment of the *MtRel* catalytic domains. The catalytic N-terminal domains (NTDs), the hydrolase and the synthetase, are called as the *MtRel* NTD (1-394 amino acid residues). The purified catalytically active recombinant *MtRel* NTD was found to be a dimer in solution as determined by SAXS studies. The

atomic resolution structure was obtained by the X-ray crystallography agreed with the SAXS studies. The structure showed for the first time the arrangement of the mycobacterial hydrolase and synthetase domain in the absence of any substrate highlighting the unique features of binding sites in comparison to non-mycobacterial Rel proteins. SAXS studies in the presence of substrates further revealed unaltered overall shape of *MtRel* NTD in solution.

- Characterization of the *MtRel* C-terminal regulatory domains. The C-terminal of the *MtRel* (394-738) in which structural domains, ACT and TGS were studied by NMR spectroscopy. The recombinant protein-fragments *MtRel* CTD (394-738), *MtRel* TGS (394-459), *MtRel* NTD-TGS (1-459) and *MtRel* ACT (660-738) were cloned, produced and purified. The recombinant *MtRel* CTD fragment behaved like the wild type *MtRel* upon purification, possibly due to its intrinsically disordered linker of 200 amino acids joining TGS and ACT. The characterization of C-terminal domains was done individually in the absence of N-terminal domains and linker region.
- Understanding the structure of the *MtRel* TGS domain and its interaction with unacylated tRNA (utRNA). The structure was determined by NMR spectroscopy and further studied for its conserved features required to interact with utRNA. The titration-based mapping of residues involved in tRNA binding were identified and compared with the homologs of different species. The *MtRel* NTD in the presence of TGS domain showed almost similar synthetic activity as of *MtRel* NTD and approximately 3 fold higher than that of *MtRel* showing auto-inhibition by rest of the CTD segment.
- High quality structure determination of the *MtRel* ACT and regulation by branched chain amino acid, valine. The *MtRel* ACT structure was determined as a dimer and the oligomeric state was further confirmed using SAXS. The crucial dimeric interface residues were identified and the mutants were studied to understand the role of these residues in dimerization. The NMR spectroscopy titration experiments with branched chain amino acids revealed the specific binding of valine to the *MtRel* ACT and showed structural rearrangement upon valine binding.

The project provided the opportunity to study the structural domains of *MtRel* and thus insights were gained into the functional regulation of *MtRel*. The *MtRel* NTD showed the presence of specific pocket for synthetase domain and better resolution of

the residues at the C-terminus of synthetase domain, providing new epitopes for drug discovery. The structure of the *MtRel* TGS and titration with utRNA answered the long-argued question of interaction of the TGS domain to utRNA in the absence of ribosomal machinery using biophysical methods. The studies unfolded the mechanism of branched chain amino acids in allosterically regulating the catalytic activities of *MtRel* and hence acting as a sensor to determine the amino acid concentration. Overall, the work has guided in identifying species-specific epitopes in various domains and regulation by regulatory domains which can aid in structure aided drug discovery and therapeutic development.

1 Introduction

1.1 Overview of tuberculosis

Mtb was first recognized by Robert Koch in 1882 as the causative agent of tuberculosis (TB) [1]. In 2019, TB still prevails as a principal cause of death in low- and middle-income countries. Every year TB claims about 1.5 million deaths globally and has infected 10 million new cases in the year 2018 [2]. Synergistically with HIV infection, the TB epidemic has been further intensified and the disease comes in the top 10 diseases which leads to death [2]. Despite all the efforts to control the disease, the global burden of TB is not declining significantly anymore. The emergence of drug-resistant TB is among many reasons why TB is posing a public health crisis. The End TB Strategy (2015-2035) initiated by the World Health Organization (WHO) aims to reach milestone set for year 2025 to observe a 10% decline per year in the number of global TB incidences. The current annual decline rate which is 2% should get accelerated to accomplish the bigger goals. For this purpose, the WHO has proposed new plans to accelerate the research and development. The priorities are set to bring new vaccines/drugs to target the adult pulmonary-, drug-resistant and drug-tolerant TB. According to WHO, research initiatives are needed to treat the 1/4th of the world population which can develop active TB at any point in their lives [2]. Other goals focus on the development and use of rapid diagnostics in the TB care centres and developing regimes that are shorter and simpler to eradicate TB [2].

TB is prevalent in most of the world, especially in the developing countries. As per WHO 2019 TB report, the largest number of cases have originated from countries in the western pacific region and south-east Asia which is followed by African regions. TB has been linked to the socio-economic status of the countries. As per statistics, most of the high-income countries have reported 10 cases per 100,000 people in comparison to 150-400 cases which are reported in high TB burden countries [2]. In the countries like the Central African Republic, the democratic People's Republic of Korea, Mozambique, Lesotho, South Africa, The Philippines and Namibia, 500 cases per 100,000 populations have been reported in the previous year (Fig. 1.1) [2]. The emergency to eradicate TB thus strengthens the motivation behind research and development to win the global war against this disease.

Figure 1.1: Estimated TB incidences in 2018. The annual number of TB incidences relative to the population size is shown. The figure is reprinted from Global Tuberculosis Report 2019, WHO [2].

The mortality rate due to TB has come down since 1990 with the introduction of standard TB regime which comprises of isoniazid, rifampicin, pyrazinamide and ethambutol [3]. Despite, there is no single milestone in the TB treatment regime that has reduced the regime duration any further. The emergence of drug-resistant strains against these standard regimes are calling for attention which also includes the resistance posed by phenotypic changes in the bacteria during asymptomatic TB infection [4]. The multidrug-resistant (MDR) TB is caused by bacteria which becomes resistant to the most potent first-line drugs of TB, isoniazid and/or rifampicin. In addition to isoniazid and rifampicin, there are bacterial strains that become resistant to second-line drugs as well such as fluoroquinolones, kanamycin, and capreomycin [5]. The infection caused by such bacterial strains is known as extremely drug resistant (XDR) TB [6]. The drug resistance can be observed in patients described as “new cases” who have never received any anti-TB treatment earlier, defined as primary drug resistance and also in patients who were previously treated with a TB therapy, known as acquired drug resistance. Poor treatment and/or non-compliance by the patients are the major reasons behind acquired drug resistance which is often linked to poor diagnostics/treatment and socio-economic status, respectively [5]. On the other hand, new cases arise due to transmission of *Mtb* from infected to healthy individual [2].

So far, several studies and successful formulations have been developed for the treatment of actively replicating *Mtb*. However, it seems that *Mtb* has devised several ways to escape such drugs either by genetic mutations or by surviving as dormant bacteria. The dormancy of *Mtb* can be associated with one of the critical stages of *Mtb* etiology i.e. latent TB (LTB) infection [7]. *Mtb* during latency show adaptations in terms of morphology and metabolism. During this stage, dormant *Mtb* bacteria also known as “persisters” become less susceptible to the available drugs which target the processes associated with replicating bacteria [8, 9]. To understand this further we can take examples of the drugs that are used in the drug-resistant TB regime.

The regime for MDR/RR(rifampicin resistant)-TB identifies drugs in Group A (levofloxacin/moxifloxacin, bedaquiline, linezolid), B (clofazimine, cycloserine/terizidone) and C (ethambutol, delamanid, pyrazinamide, imipenem–cilastatin, meropenem, amikacin (streptomycin), ethionamide/prothionamide, p-aminosalicylic acid) [10]. DNA synthesis of *Mtb* is targeted by fluoroquinolones such as moxifloxacin and gatifloxacin and thus prevents growth of *Mtb* [11]. Clofazimine is a repurposed drug for MDR-TB for which the mechanism of action has not been understood clearly [12]. Although, it has been proposed to either interact with the respiratory chain to produce reactive oxygen species, H_2O_2 , and superoxides or interact with membrane phospholipids and interfere with potassium transport, to eventually interfere with ATP production [12, 13]. There are drugs like streptomycin, kanamycin, capreomycin, and amikacin which target the protein synthesis by obstructing the ribosome [14]. Ethionamide arrests cell wall biosynthesis during division by affecting the synthesis of mycolic acids and arabinogalactan [15]. Linezolid is another drug prescribed for MDR-TB which binds and blocks tRNA in the peptidyltransferase center (PTC) on the 50S ribosomal subunit [16]. Recently, with the rigorous efforts of the research community, two new regimens has been added to the WHO guidelines for the treatment of the MDR/RR-TB [17]. First bedaquiline (BDQ) containing all oral standard regime, where the injectables have been replaced by BDQ while reducing the duration of the regime [18]. Another novel regime suggested in the report involves BDQ, pretomanid and linezolid, for the treatment of XDR, intolerant or non-responsive pulmonary MDR-TB [19]. All these drugs have virtually no or little effect on persisters, which serve as a reservoir of infection and can thrive in the occurrence of immunodeficiency [20, 21]. In addition to that, the promising drug BDQ shows signs of resistance [22]. BDQ in combination with pyrazinamide (PZA) has shown tremendous

potency, whereas BDQ with clofazimine was identified as a potent threat because of acquired cross-resistance in certain strains of mycobacteria [23-26].

As per the latest guidelines by the WHO for the TB preventive treatment for the latently infected population, TB regimes includes either isoniazid and/or rifapentine/rifampicin [10]. The therapy includes identification of the high risk population (people living with HIV or household contacts irrespective of HIV status) and ruling out the possibility of active TB disease [10]. The use of proper diagnostics (either tuberculin skin test or interferon-gamma release assay) is crucial to eliminate any risk of mis-treatment, otherwise it can further lead to drug resistance to isoniazid/rifampicin [10]. It has been observed that the treatment is 36 months long in the high TB transmission settings, pregnant women with HIV or women who have recently given birth in TB-burden regions [27]. The above facts highlight the essentiality of understanding mycobacterial adaptability in-depth to unravel the endogenous mechanisms which will aid in therapeutic interventions and improving TB-preventive therapy. The patients with LTB infection are generally positive for interferon-gamma release assay. A therapy which can target these mycobacterial persisters will be highly beneficial for drug resistant as well as latent TB. In this context, the mechanisms involved in the LTB infection has recently been re-validated for its role in the persistence and drug tolerance and provide a suitable target space for antimicrobial drug discovery. First and foremost, it is important to apprehend the etiology of TB to understand LTB infection which comes into play during the pathogenesis of *Mtb*.

1.2 Etiology of TB – from the transmission to pathogenesis

Mtb is an aerobic to facultative anaerobe which is surrounded by a cell envelope comprising of outermost layer which is also called capsule, relatively impermeable cell-wall and a plasma membrane made of glucans, proteins, mycolic acids, peptidoglycans, arabinogalactans, polysaccharides, unique glycolipids, and phospholipids [28]. It is visualized as red rod-shaped bacilli by acid fast stain such as Ziehl-Neelsen stain. The transmission of *Mtb* occurs when an airborne droplet of 1-5 micron in diameter makes its way to the respiratory tract of a healthy person. These droplets are generated when a person carrying pulmonary or laryngeal TB infection coughs, sneezes, or sings. Before entering the respiratory tract of a healthy person, these infectious droplet nuclei remain suspended in the air for several hours (Fig. 1.2). Once it makes its way to the mouth or

nasal passages of the receiver, droplet nuclei traverse to upper respiratory bronchi to reach the pulmonary alveoli in the lungs (Fig. 1.2).

Figure 1.2: Etiology of TB. *Mtb* transmits through the air from an infected person to a healthy person and develops primary infection. The infection is contained by cell mediated immunity. When bacteria develop ways to still survive as asymptomatic infection in granulomas, the stage is called latent TB infection. Upon immune-suppression by risk factors, the bacteria get reactivated. The host act as a new source of TB infection upon cavitation in the apical portion of lungs. The reactivated TB can be cured with proper treatment or can lead to death if left untreated. The figure is adapted and modified from [29].

The probability of transmission depends on the following four key factors:

1. **Susceptibility** – the liability of getting affected by the infectious droplets depends on the immune status of the individual exposed to the infection.
2. **Infectiousness** - This factor is directly proportional to the number of droplets carrying tubercle bacilli that are released by the person into the air. Higher the release of the tubercle bacilli, the higher is the infectiousness of that patient.
3. **Environment** - The environmental factors that enhance the chances of *Mtb* transmission includes more concentration of infectious droplet nuclei in the air, enclosed small spaces, inadequate ventilation, improper specimen handling and

positive air pressure in the patient's room which leads to the flow of *Mtb* to other areas.

4. Exposure – The closer the proximity, the more frequent the exposure and the longer the duration of exposure, the higher the risk of transmission of *Mtb*.

Once bacilli reach the pulmonary alveoli of the lungs after violating all the above factors, they first encounter the innate immune response of the host which involves the recruitment of inflammatory cells to the lungs. Alveolar macrophages and dendritic cells act as the prime defense cells against the bacilli and internalizes the pathogen via receptor-mediated phagocytosis [30-34]. The internalized pathogen gets captured by phagosome which starts maturing and enables the defense mechanism by means of acidification, increase in reactive oxygen species (ROS) and reactive nitrogen species (RNS) [35, 36]. However, the engulfing alveolar macrophages, which are programmed to destroy pathogens, becomes a favorable niche for bacilli. *Mtb* escapes the combat and flourish inside the phagosomes of alveolar macrophages by disrupting the acidification of phagosome [37-40]. The growing bacilli inside the phagosomes disrupt the phagosome membrane with the help of its secretion system which leads to the rupturing of the phagosome and release of bacterial products into the macrophage cytosol [41-43]. The cytokines produced by the infected macrophages initiate the recruitment of additional immune cells to the site of infection [44, 45]. The recruitment of these lymphocytes and fibroblasts along with the bacteria leads to the formation of a structure called granuloma as shown in mice model (Fig. 1.3) [46-48]. The granulomas which are understood as the protective shield for the host to contain the infection in the cellular prison can be exploited by *Mtb* to sustain as dormant bacteria [49]. The bacteria have the ability to become dormant by undergoing a non-replicating low metabolically active state under such unfavorable conditions. The adaptive immune response tries to control the infection by sterilizing these granulomas as shown in minipigs [50], but mycobacterium has devised several ways to escape sterilization and develop as LTBI (LTBI) as studied in non-human primate models [51, 52] (Fig. 1.2). The bacterium stays in the granuloma stage for months or years without leading to illness during LTBI (Fig. 1.3) as studied with the help of the Cornell model [53, 54]. The *Mtb* re-aligns its metabolism to assure survival through acid stress, hypoxic conditions, nutritional stress, ROS intermediates and also by showing heat shock response during acute infection, etc. [55-61]. These conditions are known as stress stimuli in which bacteria exhibits protective responses by changes in its transcriptional machinery [56].

The immunocompromised conditions such as HIV infection, lead to the spread of TB infection by reactivation of bacteria in the granulomas (Fig. 1.2 and 1.3). The reactivated infection can either be cured by TB regime or the person can become another source of transmission. The etiology highlights the importance of understanding “persisters” and thereby LTBI which helps the bacteria to survive for years inside the host.

Figure 1.3: Dynamics of granuloma formation. Upon local inflammatory infiltrate of *Mtb* the following may occur (i) formation of granuloma by immune cells, (ii) balanced inflammation such that the tissue damage is minimal and *Mtb* growth is controlled, or (iii) granuloma necrosis leading to endobronchial transmission and reactivation of the disease. These granulomas represent stages of the *Mtb* life cycle. They either show controlled growth or exhibit metabolic adaptation upon stringent response within the lesion. *Mtb* can also demonstrate relapse and thus spread to the next host following granuloma disruption and cavitation. The figure is recreated based on figure 1 from [62].

1.3 Persistence and LTBI

TB can be categorized as drug-susceptible, drug-tolerant and drug-resistant TB. The treatment regime is for 6 to 9 months for a patient infected with drug-susceptible TB whereas the drug-resistant strains of *Mtb* take more than 18 months for the treatment.

Such long treatment, despite being less effective, is costly and more toxic to the TB patients [6]. In general, bacteria become drug-resistant when they either acquire new genes through horizontal gene transfer or carry certain genetic mutations which make it survive irrespective of the dosage of the drug [63]. In *Mtb*, there is no report stating the case of horizontal gene transfer whereas genetic mutations have been seen as the major cause of drug resistance [64]. On the other hand, persistent bacteria are identified as the bacteria exhibiting replication shutdown and switch in metabolic pathways during antibiotic pressure [63, 65]. The persistent bacteria can also become drug tolerant upon prolonged exposure of antibiotics. The induced drug tolerance cannot be passed onto the descendants. It has been reported that cessation of antibiotic pressure leads to growth of drug-susceptible bacteria. Therefore, it becomes difficult to sterilize the infection with the cocktail of drugs which can only target actively replicating bacteria. The TB regimes should address the dynamic bacterial population co-existing as drug tolerant, drug-resistant and drug-susceptible forms [65-70]. Upon failure of sterilization, these non-replicating bacteria lead to latent infection identified by the presence of an asymptomatic lesion inside the host and believed to prolong the TB regime [7, 71]. These persistent bacteria serve as a potential danger to reactivate the disease years later. Patients with LTBI cannot spread TB infection to others but at some time of their lives, 5-10% of them will develop TB disease within five years of initial infection [72]. The risk is higher in the people having a weak immune system especially due to HIV infection and diabetes [7]. The concerns are even more for those who have been infected by someone with MDR/XDR TB because on activation under any circumstances, these people will develop drug resistant TB, not regular TB disease [73]. To understand and target persistent bacteria, LTBI has drawn the attention of several researchers. The latent phenotype and its mechanisms are linked to persistent bacteria and are believed to be understood by studying mechanisms involved in restricted growth and metabolism. Among various complex mechanisms involved in LTBI, the stringent response in *Mtb* is known as a global response which leads to pleiotropic changes in the pathogen under nutrient starvation [74]. In order to understand the causes of drug tolerance and the mechanisms to alleviate the defence mechanism exhibited by pathogen during persistence, it is prerequisite to answer the fundamental questions of *Mtb* pathology, identify pathways involved in providing this luxury to mycobacterial system and finally finding ways to unravel the immune-response mediated metabolic shifts in the pathogen leading to its success.

To develop an understanding of stringent response as a characteristic of persistent bacteria, and further characterize the essential regulator of stringent response, following sections describe the current knowledge from *E. coli* and other bacteria. While building the understanding of stringent response from model organisms, the significance of stringent response in *Mtb* is described followed by discussion on studies highlighting the regulator of stringent response, *Mtb* Rel protein (*MtRel*) as an attractive target for therapeutics development and what is known from literature about its individual domains. The identification of research gaps in the current knowledge of stringent response in *Mtb* thus helped in putting forward the research goals of this thesis to characterize *MtRel*.

1.4 The stringent response

It is a common phenomenon, among humans and bacteria, to adapt to their surroundings. Bacteria to some extent have out-mastered this skill due to their less complex genome and short generation time. Whenever a microorganism encounters an environment of limited resources or any condition which is not favourable for its existence, it shows complex and highly efficient responses in an orchestrated manner. The elaborate pathways exhibited by microorganism during nutrient limitation or other stress stimuli enables them to survive the host induced hostile environment. In order to do so, the bacterium reduces its protein synthesis with the down regulation of rRNA and tRNA synthesis during nutrient limitation [75]. One such mechanism which occurs during amino acid starvation is known as the stringent response. Such mechanism is displayed by bacteria of both gram-negative and -positive groups. The stringent response was first observed by Sands and Roberts in *Escherichia coli* (*E. coli*) cells undergoing amino acid starvation [76]. They noticed an exponential decrease in the accumulation of stable RNA in the amino acid starved cells and named it as ‘stringent response’. The response was mapped to a single locus *relA* earlier named as RC gene as it was responsible in RNA synthesis control. The mutations in *relA* abolished the phenotype of stringent response and thus named as ‘relaxed response’ phenotype. To understand this phenotype and find out the regulators responsible, Gallant and Cashel identified hyper-phosphorylated forms of guanosine as magic spots which accumulated in response to starvation [77]. Extensive research has been carried out to understand the regulatory pathways associated to stringent response and the processes which regulated the regulators of stringent response. However, the components of the mycobacterial

stringent response have not been fully understood, which should be addressed to target persisters. These components of the stringent response can be first understood by taking examples from model organisms such as *E. coli* and other pathogens. The first component of stringent response is the stringent effector molecule.

1.5 Effector of stringent response - (p)ppGpp

The hyperphosphorylated forms of guanosine named as tetra- and penta-phosphate guanosine, ppGpp (I) and pppGpp (II) were first identified and named as magic spots [77]. These magic spots come under the category of bacterial nucleotide-based second messengers. They are responsible to initiate a signal cascade in response to stimuli such as environmental changes or intracellular conditions. Over the years since the discovery of (p)ppGpp, various bacteria have been studied to show the stringent response which is mediated by the concentration of hyperphosphorylated guanosine nucleotides which can be derivative of GDP (ppGpp) or GTP (pppGpp), collectively referred as (p)ppGpp (Fig. 1.4).

Figure 1.4: Chemical structure of GDP and GTP. (A) GDP showing diphosphate moiety at the 5' carbon of the ribose ring. (B) GTP showing the triphosphate group at the 5' carbon of the ribose ring.

When the pyrophosphate group, ppi, from ATP transfers to 3' carbon of the ribose moiety of GDP or GTP, it leads to the formation of ppGpp or pppGpp (Fig. 1.5).

Figure 1.5: Chemical structure of ppGpp and pppGpp. (A) The structure of ppGpp showing the diphosphate group at the 3' carbon of the GDP (B) The chemical structure of pppGpp showing the diphosphate group at the 3' carbon of the GTP. The phosphate groups at the 3' carbon of the ribose sugar moiety is accepted from ATP as pyrophosphate.

These small nucleotide-based second messengers also termed as alarmones are responsible to elicit reprogramming of gene expression at a global scale in order to enable organism to survive in stringent conditions. (p)ppGpp mediates these changes by direct or indirect interactions with RNA polymerase and enzymes involved in controlling cell growth. In response to these changes, rRNA, tRNA, protein synthesis gets downregulated so as to escape the amino acid starvation and various other nutritional stresses like fatty acid starvation, carbon, nitrogen, and phosphate starvation. The enzyme pppGpp-5'- γ -phosphohydrolase (GppA) has been reported in *E. coli* which aids in converting penta-phosphate to tetraphosphate moiety as the latter is preferred effector molecule for stringent response in *E. coli* [78-80].

1.6 *In vivo* concentration fluctuations of (p)ppGpp leads to dormancy

For *E. coli* the basal level of (p)ppGpp during the log phase is in the low steady-state which approximates to about 20 pmol/A450 [81]. As soon as the amino acid starvation is induced, the level of ppGpp starts increasing and peaks in about 5-15 mins after starvation and gradually achieves a new steady state of 10- to 20 folds higher concentration than the original level [81]. The increase in (p)ppGpp is related to the drop of GTP and becomes equal to the new steady state of (p)ppGpp.

1.7 Pleiotropic effects of (p)ppGpp

Several studies have been carried out to understand the global effect of (p)ppGpp during stringent response. The effect of (p)ppGpp was observed at all levels of transcription, translation and post translation. The (p)ppGpp targets have been studied extensively in gram negative and gram-positive bacteria where (p)ppGpp is shown to be involved in direct or indirect regulation of transcription, respectively. Before further characterizing the regulator of (p)ppGpp, it will be good to have an understanding of the targets of (p)ppGpp. Some of the cellular targets of (p)ppGpp are described in the following subsections.

1.7.1 Regulation of transcription via RNA polymerase and the non-template DNA binding proteins

During stress conditions such as nutrient starvation, it is crucial for the bacteria to allocate the cellular resources for the production of cell survival components. (p)ppGpp has been studied to play a key role in altering the gene expression by modulating RNA polymerase (RNAP) activity. It has also been shown that during amino acid starvation, (p)ppGpp inhibits the promoters responsible for the expression of rRNA and ribosomal protein genes whereas activates the expression of genes involved in amino acid biosynthesis by activating their promoters selectively (Fig. 1.6) [82, 83].

This can be understood by the specific discriminator sequence between -10 element and the transcription start site of *rrn* genes and genes of amino acid biosynthesis. The former was shown to have GC-rich discriminator and the latter harbours AT-rich discriminator elements. The *rrn* genes form unstable open complexes with RNAP which is further destabilized by (p)ppGpp [84]. In contrast, the AT-rich discriminator of amino acid promoters stabilizes the binding upon interacting with 1.2-region of sigma factors [85]. The (p)ppGpp binding favours and increases the rate of formation of open complexes by these promoters.

Figure 1.6: Pleiotropic effects of (p)ppGpp. RelA/SpoT Homologs (defined in section 1.8) sense the stringent stimuli and synthesizes (p)ppGpp with the help of active synthetase domain. (p)ppGpp interacts with RNA polymerase and orchestrates the gene expression by upregulating the expression of genes involved in stress response and downregulating genes of components of replication, transcriptional and translational machinery. The figure is recreated based on figure 1 of [86].

1.7.1.1 Binding sites of (p)ppGpp on RNAP

To understand the interaction of (p)ppGpp with the RNAP structural studies were carried out in *Thermus thermophilus* (*Tth*) and *E. coli* [87]. According to the study in *Tth*RNAP, the ppGpp binds to the active centre of the RNAP and directly interferes with the nucleotide coordination to RNAP. The findings from the *Tth*RNAP didn't hold true with the biochemical and mutagenesis experiments performed *in-vitro* in RNAP from *E. coli* [88]. It was also highlighted that the RNAP holoenzyme bound to ω subunit should be used for bacterial RNAP complex preparation to study the effect of (p)ppGpp [89]. The structural studies on *E. coli* RNAP highlighted the significance of ω subunit

for the (p)ppGpp binding to RNAP and thereby its inhibition [80, 90]. Overall, the results show that the binding mode of ppGpp in *E. coli* and *T. thermophilus* are not appropriate for comparison as the conserved motif required for the binding of (p)ppGpp differs in species-specific manner in α -, β - and γ -proteobacteria from actinobacteria and thermophiles [90, 91]. Moreover, the site 2 of (p)ppGpp in *Ec*RNAP is DksA dependent. DksA, a non-template DNA binding protein which acts as a global transcriptional regulator and directly interacts with RNAP along with ω subunit to elicit a stringent response along with ppGpp or independently otherwise, investigated both *in-vivo* and *in-vitro* [82, 89, 92, 93]. The structural studies on DksA:ppGpp:RNAP complex have confirmed the DksA mediated effect of ppGpp in transcription inhibition [94, 95]. The studies with DksA in the presence of ppGpp have shown that the binding site of ppGpp as proposed in *T. thermophilus* was created in *E. coli* RNAP in the presence of DksA and had a more profound effect on stringent response than the other site of ppGpp proposed far away from the secondary channel of *E. coli* RNAP (Fig. 1.7) [95].

Figure 1.7: Regulation of RNA polymerase activity by (p)ppGpp. The *E. coli* RNA polymerase (α 1-slate blue, β - pale blue, β' - teal and ω – grey subunits) is shown with the two binding sites of (p)ppGpp (red and green spheres) shown in structural and biochemical studies. The DksA provides the interaction site for the (p)ppGpp near the secondary channel of RNAP whereas the site 1 allosterically regulates the RNAP activity by ratcheting. The model was generated by using PDB 4JKR (RNAP:ppGpp) and 1TJL (DksA).

1.7.2 Indirect regulation of transcription

In *B. subtilis*, these findings were shown to bring alteration in nucleotide triphosphate pool thereby changing the promoter preferences by RNAP which are dependent on GTP levels for initiation [96]. However, the mechanism behind reprogramming of gene expression by identification of NTPs still need to be understood during stringent response.

Guanylate kinase is also a target of (p)ppGpp in *Streptomyces coelicolor* (a member of Actinobacteria). However, (p)ppGpp does not inhibit the guanylate kinase homolog in *Mtb* [83, 97]. The other direct targets of (p)ppGpp such as hypoxanthine-guanine phosphoribosyltransferase (HprT) and inosine-5'-monophosphate (IMP) dehydrogenase from purine biosynthesis pathway as identified in *B. subtilis* needs validation in *Mtb* [96, 98].

1.7.3 (p)ppGpp and DNA replication

The association of stringent response and inhibition of DNA replication was understood in *E. coli*, *B. subtilis* and other bacterial species in order to maintain genome integrity during starvation [99-101]. The regulation of DNA replication by ppGpp was studied as a decrease in transcription of the replication-initiation protein-coding gene, *dnaA* [102]. However, recently, studies in *E. coli* and *B. subtilis* investigated the role of ppGpp in inhibiting the replication elongation step catalyzed by primase, DnaG [103, 104]. The replication machinery is suspended during nutrient limitation as a result of inhibition of DnaG by the (p)ppGpp. Such arrested replication forks in general are disrupted by the recruitment of DNA repair protein, RecA at replication forks, however, (p)ppGpp interaction to the primase limits this mechanism [103]. Further differential activity of ppGpp and pppGpp on DnaG were studied and showed that ppGpp is a stronger regulator than pppGpp in *E. coli* unlike as in *B. subtilis* [104].

1.7.4 (p)ppGpp and its interaction with translation factors

The translation repression by (p)ppGpp happens at a transcriptional level where it suppresses the expression of genes coding for rRNAs, tRNAs and ribosomal proteins as well as at translational level where it interferes with GTPases such as initiation and elongation factors [105, 106]. The initiation factor IF2 has been reported to interact with ppGpp which completely disrupts the binding of fMet-tRNA^{fMet} to the 30S pre-initiation complex for the initiation of translation [107]. In these studies, IF2 is demonstrated as a

metabolic sensor and regulator which either inhibits translation upon interaction with ppGpp or continues translation by interacting with GTP [106]. On the other hand, the binding of ppGpp with elongation factor thermo-unstable (EF-Tu) restricts the exchange of EF-Tu:GDP in the presence and absence of elongation factors thermos-stable (EF-Ts) [108]. Studies have also reported that ppGpp interaction with EF-Tu decreases the missense error and increases “proofreading flow” and hence aids in translational accuracy and fidelity during starvation conditions [109-112]. The aforementioned results were further validated for the GTPase inhibition upon the interaction of ppGpp to IF2 and translocase EF-G by using thermodynamic studies [113].

1.7.5 (p)ppGpp and virulence

Pathogenic bacteria take advantage of stringent response to survive in adverse conditions related to nutritional stress during infection. The virulence can be indirectly linked to the physiological adaptations of pathogenic bacteria. The studies of complex systems as in *E. coli*, *Streptococcus dysgalactiae* subsp. *equisimilis* (*S. equisimilis*), and *Salmonella typhimurium* (*S. typhimurium*) have unveiled how the coordination among biological regulators during stringent response leads to virulence [114-118]. In Enterohemorrhagic, *E. coli* virulence is associated with cell adherence to the intestinal epithelial cells leading to the development of the locus of enterocyte effacement (LEE). Studies have shown that in concert with DksA, ppGpp levels activates LEE expression and increases bacterial adherence [117]. In another strain of *E. coli*, which is uropathogenic, ppGpp together with DksA regulates the *fim* operon for the fimbriae formation [119]. Both these strains like the non-pathogenic laboratory strain of *E. coli* K-12 encodes a RelA and a SpoT enzyme (defined in the next section) for (p)ppGpp regulation [114].

The (p)ppGpp based direct and indirect regulation of cellular targets further validates that the concentration fluctuation of (p)ppGpp elicit global stringent response. The understanding of (p)ppGpp concentration regulator will be interesting to study and further characterization will aid in targeting the stringent response.

1.8 Regulator of (p)ppGpp – RelA/SpoT homolog superfamily proteins

The metabolic regulator of (p)ppGpp is categorized as a special family of enzymes, RelA/SpoT Homolog Superfamily (RSH) which is solely responsible for the

accumulation of (p)ppGpp. The RSH family members can be broadly classified into the following categories depending upon their functionality:

- Monofunctional (p)ppGpp Synthetases – The enzymes of this category can be long or short RSHs (Fig. 1.8) and can only synthesize (p)ppGpp with no hydrolase activity, for example, RelA of *E. coli*.
- Monofunctional (p)ppGpp hydrolases - These enzymes are homologous to the hydrolases but having mysterious phosphatase activity of NADPH. So far, these enzymes have only been reported in the metazoan group (*Mesh1*) [120, 121].
- Bifunctional (p)ppGpp synthetase and/or hydrolase enzymes. These enzymes are long RSHs which generally carry fully functional hydrolase and synthetase domains in addition to C-terminal regulatory domains for instance the Rel protein of *Mtb* (*MtRel*) (Fig. 1.8). Enzymes like *E. coli* SpoT are also considered bifunctional RSHs although they carry domains for fully functional hydrolysis and a synthetase exhibiting weak synthesis activity.

Figure 1.8: Schematic representation of the RelA/SpoT homologs (RSH) domain structure. RSH family proteins can be classified as long and short enzymes. Long RSHs can be monofunctional hydrolases/synthetases or bifunctional hydrolases and synthetases and possess C-terminal regulatory domains. The short RSHs can be either synthetases or hydrolases and lacks the C-terminal regulatory domains. The figure is created based on information provided in [121].

It is common among bacteria to have single bifunctional RSH protein abbreviated as Rel, whereas members of β and γ proteobacteria possess two copies of long RSHs, RelA and SpoT [121, 122]. This is generally understood as an event of gene duplication. In addition to these long RSHs, short RSHs are commonly present in gram-positive bacteria and restrictedly found in some gram-negative bacteria [121]. Small alarmone synthetases (SASs) and small alarmone hydrolases (SAHs) are among such

short RSHs. These SASs have been reported in many bacterial species such as RelV in *Vibrio cholera* [123] and RelP (Yjbm) and RelQ (Ywac) in *Bacillus subtilis* and *Streptococcus mutans* [124, 125]. The presence of SASs sometimes has other beneficial roles including fine-tuning the stress response by amplifying the signal or by sensing the environmental cues that are not recognized by long RSHs [124-126]. For example, the presence of RelV in *Vibrio* enables the bacteria to produce (p)ppGpp when *rel* and *spot* genes are inactivated, during fatty acid or glucose starvation [123, 127]. The distribution of short RSHs suggests the event of horizontal gene transfer leading to a scattered appearance in organisms [121]. For simplicity, the long RSHs are referred as RSHs and bifunctional long RSHs as Rel throughout this thesis.

1.9 Mechanism behind the accumulation of stringent factor

To identify the regulator of (p)ppGpp, Haseltine et al. were able to purify a ribosome-RelA complex exhibiting synthesis activity [128]. The protein was isolated by washing ribosomes with ammonium chloride. They also provided the first evidence of the accumulation of (p)ppGpp as a result of idling of the ribosomes undergoing the process of protein synthesis using *in vitro* studies [129]. The first RSH protein which was identified to be ribosome-associated, was *EcRelA* and was later used to propose the model of the RelA mechanism by the stalled ribosomes [77, 130]. Several models and hypotheses have been put forward to explain the mechanism of activation and inactivation of long RSH proteins. According to one model, the ribosomes get stalled when deacylated tRNAs sit at the A-site of the ribosome. This interaction of RelA with a stalled ribosome activates the synthesis of (p)ppGpp with the help of ATP and GTP (Fig. 1.9)

The mechanism was further extended to explain the rapid accumulation of millimolar concentration of (p)ppGpp in *E. coli* in comparison to the low intracellular concentration of RelA. The model was named as the ‘hoping’ model where it was suggested that RelA hops from one stalled ribosome to another in order to rapidly accumulate (p)ppGpp produced per enzyme. It was also shown that the RelA affinity decreases to the ribosome during the event of (p)ppGpp synthesis [131].

Figure 1.9: Mechanism of stringent response by *E. coli* RelA. During amino acid starvation, uncharged tRNAs accumulate in the cytosol and binds to the ribosomal A-site. The stalled state of ribosome recognized by RelA which binds to the large subunit of ribosome in order to adopt active conformation. RelA activation leads to (p)ppGpp synthesis and eventually RelA gets dissociated from ribosome and carry out several rounds of (p)ppGpp in dissociated state. Increase in (p)ppGpp reprogram cellular gene expression and directs the resources to amino acid biosynthesis. The restoration of amino acids further restores the acylated tRNA which replaces deacylated tRNA from the A-site with its competitive binding affinity (The figure was reconstructed from [132, 133]).

In another study, positive allosteric regulation of ribosome dependent synthetic activity by ppGpp was proposed in order to explain the rapid accumulation of (p)ppGpp [134]. Another mechanism stating RelA extensively delayed inactivation upon dissociation from the ribosomes was studied in the single-molecule experiments in living cells [135]. On the contrary, recently an interesting study done by using UV crosslinking and deep sequencing proposed that *Ec*RelA can interact with the deacylated tRNAs without being activated. The hypothesis is explained in Figure 1.10.

The loading of deacylated tRNA-RelA complex at the hungry A-site of ribosomes occurs upon amino acid starvation. The loading of the complex brings RelA in the proximity of the Sarcin-Ricin loop of 23S rRNA which activates the synthesis of (p)ppGpp by RelA [136].

Figure 1.10: Model of RelA activation. During exponential growth while normal translation is happening, RelA interacts with deacylated tRNA and stays in transient RelA-tRNA binary complexes. Upon amino acid starvation, these binary complexes load at the “hungry” A-site in the absence of amino-acylated tRNAs complementing the codon on the mRNA. Cognate codon- anticodon pairing brings the Hydrolase domain in contact with the Sarcin-Ricin loop (SRL) in 30S and thereby stabilizes the Synthetase ON configuration of RelA (The figure is reconstructed from [132, 137, 138]).

1.10 Understanding of stringent effector regulation by RelA and SpoT

E. coli RelA and SpoT are the most studied enzymes of the RSH family. As explained earlier, when the translational machinery is stalled, RelA gets activated as per the hypothesized models and in return produces (p)ppGpp. The transferase reaction required for the production of (p)ppGpp in *E. coli* was first studied by Justensen et al, where they identified the catalytic action of RelA in transferring 5'-β, γ pyrophosphate group from ATP to the GTP occurs in the presence of divalent Mg^{2+} [139]. The stringent response in *E. coli* has been understood as a broad transcriptional program regulating 80 genes to bring starvation adaptation in the organism. The stringent response which was recognized to be linked with amino acid starvation only, further investigated and the other stimuli involved in the regulation were identified. Recent studies on transcriptional control of the *relA* gene in *E. coli* showed that *relA* is under control of 4 promoters. In addition to promoter P1 and P2 which are σ^{70} dependent and regulated either constitutively (by P1) or upon induction during carbon, temperature and osmotic stresses (by P2), P3 and P4 promoters which are dependent on σ^{54} controls the

transcription of *relA* during nitrogen starvation [140-144]. Additional transcriptional regulation of *relA* and *spot* genes from *E. coli* is understood in the toxin-antitoxin system and 6S RNA leading to the differential expression [145].

Upon expression, the proteins RelA and SpoT utilize preferentially GDP or GTP respectively to produce ppGpp or pppGpp *in vitro*. The preference of substrate has been linked to the presence of motif EXDD or RXKD in substrate binding site of monofunctional and bifunctional long RSHs [146]. The EXDD-motif stabilizes GDP for catalytic activity whereas RXKD utilizes GTP for the same [146, 147]. In contrast to *E. faecalis* and *B. subtilis*, *E. coli* prefers ppGpp and considered as a strong regulator of stringent response leading to a generalization for the different substrate preferences in gram-negative and gram-positive bacteria [80, 148, 149]. Whereas, *in vivo*, pppGpp was reported as the initial product of stringent response because the physiological concentration of GTP depicted to be more than GDP [139, 150]. The pppGpp is readily converted to more stable ppGpp by the activity of GPPA. Another catalytic activity responsible for the regulation of (p)ppGpp in *E. coli* was reported with the discovery of *spot* gene where the studies provided the evidence of 5'pyrophosphohydrolase activity by the product of *spot* gene in the presence of divalent ion Mn^{2+} [151-153]. These studies also explained the presence of weak synthetase activity by SpoT. The hydrolysis activity of SpoT leads to the production of GDP from ppGpp which is readily converted to GTP by the activity of nucleoside 5'-diphosphate kinase (NdK) in *S. typhimurium* in the presence of ATP [154]. The figure below shows the schematic depicting the generalized pathway to regulate (p)ppGpp (Fig. 1.11)

The weak synthetic activity exhibited by SpoT is also linked to the limitation of phosphorous, iron, carbon sources and/or fatty acids [155-157]. Various studies have also shown that the SpoT activity to synthesize (p)ppGpp is upon the interaction of acyl carrier protein (ACP) in *E. coli* during fatty acid starvation by regulating the bifunctional activity of SpoT [156]. The conformational switch induced upon the interaction of ACP to the TGS domain of SpoT aids in switching the role of SpoT as hydrolase to synthetase [156]. In addition to fatty acid starvation, the response is also proposed to be linked to carbon starvation which affects glycolysis and downregulates the synthesis of fatty acids [158]. The SpoT regulation during fatty acid starvation is limited to bacteria in which both RelA and SpoT are present. Another study has provided evidence of SpoT interaction directly to the CgtA associated with the ribosome in order to initiate hydrolysis of (p)ppGpp. The study further supports the essentiality of CgtA absence

required during nutrient starvation to accumulate (p)ppGpp [159, 160]. The aforementioned studies highlight the importance of SpoT in identifying the physiological state of the *E. coli* cells so as to exhibit stringent response independent to RelA.

Figure 1.11: Generalized schematic showing biosynthetic pathway of ppGpp. The enzymes involved in the different steps for the metabolism of (p)ppGpp are represented on the arrows. The substrates and product of respective reactive are highlighted symbolically. The symbols are explained below the diagram. The figure is reconstructed based on Figure 1b in [133].

In addition to all the mechanisms described above for the regulation of stringent effector, ligand-induced positive regulation of *EcRel* has been proposed. This is a rare mechanism of positive feedback mechanism by the product and explains the rapid accumulation of ppGpp required to elicit stringent response [134]. Although the mechanism of action is unknown, the positive regulation by alarmone has also been reported in SAS proteins from *E. faecalis* and *B. subtilis* [161, 162].

Recent structural studies have shed some light on the structure and mechanism of *EcRelA* complexed with stalled translational machinery [138, 163-166]. These studies were able to identify the components of translational machinery interacting with RelA and map the regions of interactions in *E. coli* (Fig. 1.12).

Figure 1.12: The structure of the complete ribosome 70S from *E. coli*. The lateral view of the 70 S ribosome shows 30S subunit (including rRNA and the ribosomal proteins) in sand colored surface and the surface of 50S subunit is colored in gray (including rRNA and ribosomal proteins). The three tRNAs at A, P and E site respectively colored in purple, forest green and mustard. The RelA from *E. coli* is shown in cylindrical ribbons and loops. The respective color of each domain is shown as, Blue - Synthetase domain, Teal- Hydrolase domain, Purple - TGS domain and Orange - ACT domain. The figure is generated by using the PDB ID 5IQR [165].

1.11 Metabolism of (p)ppGpp by bifunctional Rel protein- *SeRel*

Despite all the work carried out on monofunctional RSH RelA, there is no structural evidence of how these domains from bifunctional enzymes interact with components of the translational machinery and/or with other regulatory factors. The only structural study explaining the mechanism of (p)ppGpp regulation by Rel protein is reported from *S. equisimilis* describing a cross-talk between the two NTDs [167]. The NTDs of *SeRel* (1-385 amino acid residues) has revealed two monomers locked in two different states; state 1 showing Hydrolase ON and Synthetase OFF whereas state 2 explains the contrary (Fig. 1.13) [167]. The study explained that the conformational changes in the protein upon binding of substrates for the antagonistic activities using structural and mutagenesis experiments. The detailed analysis of crystal structure proposed the mechanism of synthetase activity explained how 3'OH of GDP/GTP

attacked the β phosphorous of ATP to form pppGpp. The role of E323 in this mechanism was investigated and confirmed by the mutant generation [167]. The synthetase activity was shown to be dependent on Mg^{2+} -binding, which was explored with the involvement of D264. It guided the structural confirmation required for ATP to interact with Mg^{2+} and bring the complex in the proximity of GTP for pyrophosphate transfer reaction [167]. The charge reversal in the synthetase domain motif as compared to *EcRelA* also explained why bifunctional Rel in *S. equisimilis* inhibited by high Mg^{2+} concentration. The inhibition of synthetic activity upon the increased concentration of Mg^{2+} was observed when a loop to helix transition of $\alpha 13$ - $\beta 4$ loop happened to close the substrate-binding pocket of the synthetase domain [147].

Figure 1.13: Crystal structure of Rel from *S. equisimilis* (*SeRel*). The *SeRel* (1-385 amino acid residues) showing two different conformations. The two conformations depict that how ligand binding in hydrolase domain affects the conformation of synthetase domain shown to bound with GDP in both the conformers. The ligand found bound in hydrolase domain was non-hydrolysable form of (p)ppGpp, abbreviated as ppG2':3'p.

Hogg et al. also demonstrated the comparative analysis of the *SeRel* hydrolase domain to the catalytic domain of Phosphodiesterase 4 (PDE4) showing that irrespective of the different substrates available for hydrolysis (p)ppGpp and cNMP, respectively, both domains shared the same HD family fold. They also highlighted the essential residues involved in bringing the hexa-coordinated geometry required for Mn^{2+} -chelation upon binding of the substrate so as to catalyse hydrolysis reaction. The ligand was shown to be clutched at the binding pocket by N148 and T151 of the 3-helix bundle and R44 of the hydrolase domain. Thereby, the Mn^{2+} that was present in the penta-

coordinated geometry comes in contact with the D78 with the disruption of the salt bridge between D78 and R44. This overall mechanism stabilizes the ligand with Mn^{2+} in hexa-coordinated geometry and hydrolysis reaction occurs [167].

1.12 Significance to study stringent response in *Mtb*

(p)ppGpp appeared as magic spots around 50 years ago and recently with the new insights into the mechanism of RSH enzymes has renewed the attention in this field. As explained before, the first structure of a bifunctional Rel from *S. equisimilis* came nearly 3 decades after the discovery of stringent response. Further studies reinforced the facts behind targets of (p)ppGpp which includes components of DNA replication, transcription and translation. The molecular targets of (p)ppGpp that have been studied so far in *Mtb* are RNA polymerase, gunaylate kinase, and polyphosphatases. Despite common targets among various bacterial species, (p)ppGpp differentiates the way it targets these cellular processes in mycobacteria [168, 169]. In the studies comparing the binding site of (p)ppGpp in *E. coli* to that of *Mtb*, the two residues conserved in site-1 to that of *E. coli* RNAP were *Mt*RNAP β' D714 and *Mt*RNAP β' K437. Although, to prove that (p)ppGpp interacts with *Mt*RNAP, *in-vitro* transcription reaction was shown to be inhibited by the addition of pppGpp. This further blocked the RNAP open complex formation at the rRNA promoter (*rrnAP3*) and transcript formation at *gyrB* promoters [170-172]. It was also shown that an RNAP negative regulator, a Gre factor homolog binds near the secondary channel and alleviate the effect of ppGpp on RNAP open complex formation at the rRNA promoters even in the absence of DksA [170]. In addition to direct regulation, indirect regulation by (p)ppGpp is also observed in *Mtb*. In the absence of (p)ppGpp hydrolysis, the over-accumulation of (p)ppGpp was shown in *M. smegmatis* to decrease the GTP levels in the cell and thus affecting the purine metabolism. In mycobacteria, rRNA production was also shown to get downregulated and two promoters upstream to the rRNA locus were shown to be initiated by GTP. This highlights that any change in the nucleotide pool by (p)ppGpp interferes with transcription of stable RNAs [96, 173]. The studies on animal models have shown that full virulence by *Mtb* require activation of stringent response [114, 174]. Additionally, the direct targets of (p)ppGpp in DNA replication are not known from mycobacteria, however, morphogenetic changes have been studied in *M. smegmatis* and the defect in cell division was reported by the formation of long multinucleate and multiseptate cells [175]. Thus, (p)ppGpp regulation is essential for virulence, long term persistence,

antibiotic tolerance and resistance, biofilm formation as a result of the stress response and has significance beyond stress response in metabolism and cell-cycle progression [114, 176-179].

The research in the field has gained interest and RSHs have showed potential as the targets of antimicrobial therapeutic discovery [180-182]. With all that is known about importance of (p)ppGpp for stringent response in mycobacteria, it became even more significant to understand the mechanism and significance of its regulator in persisters. Some of the key studies highlighting essentiality of Rel in mycobacteria are described in the following section.

1.13 Discovery of *MtRel* and its significance in persistent bacteria

After gaining state of art knowledge from model organisms such as *E. coli*, which are extensively studied, discovery and significance of *MtRel* in persisting bacteria will help in understanding the rationale behind studying it as a target in *Mtb*.

Significance of *MtRel* and its essential role in the survival of mycobacteria has been studied *in-vivo* and *in-vitro*. The *MtRel* (Fig. 1.14) was first isolated in 1999 as a RelA/SpoT homologue and was identified as a bifunctional enzyme which can synthesize and hydrolyze (p)ppGpp *in-vitro* independent of ribosome and deacylated tRNA unlike Rel from *S. equisimilis* [183].

Figure 1.14: Functional characterization of the *MtRel*. The *MtRel* protein was truncated into multiple fragments and important residues and domains were identified by Avarbock et al.[183, 184]. The minimum residues required to have hydrolase activity (HD domain with doublet) was identified to be from 1-181 amino acid residues, however residues 1-203 were necessary for fully functional hydrolase domain. G241 and H344 were identified as probable nucleotide binding residues in synthetase domain (87-394 amino acid residues) and critical for the activity of synthetase domain. The multimer forming regions were identified in the region 449 to 676 amino acid residues, however proposed to be not enough for multimerization. The mutation of C607 and C633 showed no role in multimerization thus

advocating non-covalent interaction to be the cause of oligomerization. The figure has been reconstructed from Figure 1 of [184]

The role of *MtRel* was confirmed *in-vivo* by generating *E. coli relA* and *spoT* deletion strain. The complementation of *MtRel* in such strain suppressed the (p)ppGpp null mutant phenotype and produced more (p)ppGpp than *EcRelA*. Eventually, two strong evidences showed the role of *relA* in long term survival and persistence of *Mtb* [74, 185]. The Rel knockout in the mycobacteria impaired survival during long term starvation and showed loss of viability in hypoxic conditions over the span of 6 months when compared to wild type *in-vitro* [74]. Such conditions were then linked to the stresses that challenge mycobacteria inside the human host [186]. These results were further validated in mice models and guinea pigs and *MtRel* was reported to be critical for the successful establishment of latent infection by bringing alteration to the transcriptional apparatus, expression of virulence factors, enzymes involved in cell wall biosynthesis, and heat shock response etc [60, 174]. In the study conducted in mouse models, after 2-5 weeks post-infection, no difference in the replication of WT and *rel* knockout *Mtb* in both mouse lungs and spleen have been observed, whereas after 5-7 weeks of infection, the knockout started losing viability in comparison to WT. By the end of 38 weeks, the infection was 500-fold less in the knockout strain in comparison to WT. The difference was also observed in the disease-associated weight loss of the mouse infected with WT and knockout Rel strain. The size of granulomas was also found to be smaller in knockout Rel strain infected mice [60]. Other studies validated the role of Rel homolog from *M. smegmatis* in persisters and showed that overexpression of *relA* in *M. smegmatis* resulted in similar morphology as of persisters of *Mtb* [175, 185, 187]. The *rel* gene disruption by using transposon also exhibited similar results as of knockout strains in terms of *Mtb* proliferation and the size of granulomas in mice [188]. Overall, these results indicated the essentiality of Rel in *Mtb* to enter dormancy and maintaining it during the infection.

Interestingly, it has also been shown that *M. smegmatis* has another functional synthetase apart from *MsRel* i.e. a small alarmone synthetase (SAS) and the homolog of *M. smegmatis* SAS is encoded in *Mtb* by *rv1366*. As studied in the $\Delta rv1366$ strains, the bacteria exhibited WT rate of growth in liquid media, on plates and in biofilms and also, WT level of virulence in mouse models. This further indicated that *MtRel* is the only producer of (p)ppGpp in *Mtb*, unlike *M. smegmatis* and thus the only regulator of

stringent response [189]. The studies have also highlighted the importance of maintain balance of (p)ppGpp concentration as too little concentration cannot cause chronic infection in mouse models and too much concentration in the absence of inactivated hydrolase is deleterious for *Mtb* at all stages of infection [60, 189].

The above mentioned studies of Rel in the pathogenesis of *Mtb* intensified the interest to understand Rel protein from mycobacteria. It was also validated *in-vitro* and *in-vivo* that it is important to study these (p)ppGpp regulators in species-specific manner and unravel the structural intricacies which makes them so unique in their functioning. For instance, the maximal basal rate of the two catalytic activities in *MtRel* provided first evidence of having two separate catalytic centers which were not antagonistically inhibited as in *SeRel* [184, 190]. Moreover, the differential activity has been shown to be regulated by concentration changes of uncharged tRNA to charged tRNA and association to ribosome-uncharged tRNA-mRNA complex [190]. The thesis thus addresses the gap in the knowledge about the structural features of *MtRel* and how its regulation is different than other gram-positive or gram-negative bacteria.

1.14 Organization of N-terminal catalytic domains

The N-terminus of *MtRel*, harbors the hydrolase and the synthetase domains for catalytic activities [183, 184, 191]. A similar domain architecture was observed by the studies carried out by Mechold et al in *S. equisimilis* (*SeRel*) [192, 193]. The hydrolase domain of RSHs was first identified by Aravind and Koonin and characterized as a member of HD superfamily of phosphohydrolases exhibiting the features of metalloenzymes [194]. The HD superfamily (Pfam: PF01966) was given the name due to the presence of a histidine (H) and aspartate (D) conserved motif which was later studied to be essential for the catalytic activity of the hydrolase domain of RSH enzymes [167, 184, 195]. The hydrolase domain is required for the hydrolysis of (p)ppGpp in RSH enzymes so as to keep the production of the global alarmone in check. The hydrolysis of (p)ppGpp is a metal ion, Mn^{2+} dependent reaction and had been characterized extensively for RSH enzymes of *E. coli*, *S. equisimilis* and *Mtb* [183, 193]. The first structure of the N-terminus of an RSH protein was characterized 16 years ago in *S. equisimilis* [167]. The study provided the novel structural evidence about the conformational antagonism in the bifunctional *SeRel* catalytic domains in the N-terminus. It proposed that the regulation of hydrolase and synthetase activity of *SeRel* occur by crosstalk between these two domains through a three-helix bundle acting as an

overlapping linker [167]. The synthetase domain (Pfam: PF04607), which is responsible for catalysing the transferase reaction so as to produce (p)ppGpp in the presence of metal ions Mg^{2+}/Mn^{2+} , was structurally characterized [167]. The structural homology suggests that the domain belongs to the superfamily of nucleotidyltransferases. The synthetase domain was well studied biochemically in the Rel proteins and the residues crucial for the activity of the synthetase domain were identified [183, 184, 190-193].

For *MtRel*, it was shown that residues 1-394 were able to carry out both the catalytic activities whereas 87-394 amino acids were required for synthetase activity, and amino acids 1-181 could only perform hydrolysis of (p)ppGpp [184]. The residues crucial for synthesis and hydrolysis activity were reported, where mutations of G241E and H344Y resulted in loss of function of synthetase, and mutants H80A and D81A resulted in loss of hydrolysis activity [184]. Recently, the minimal range of amino acids required for synthetase activity in *MtRel* was reported to be 136-378 amino acids (corresponding to 188-430 amino acids as reported in the presence of 52 extra residues at the N-terminus) [196]. Similar studies were carried out in *M. smegmatis* Rel (*MsRel*) which was characterized for the specific amino-acid boundaries of the respective domains [197, 198].

1.15 Organization of C-terminal regulatory domains

As shown in Figure 1.5, the RSHs carry hydrolase and synthetase as the NTDs and in addition to that, C-terminal harbours TGS (ThrRS, GTPase, and SpoT/RelA) and ACT (Aspartate kinase, Chorismate mutase, and TyrA) domains with helical and conserved cysteine (CC) regions in between. The presence of these domain is solely predicted based on sequence homology in RSHs [121]. The domains TGS and ACT acronyms are derived from enzymes in which they have been reported first. The C-terminal domains (CTD), TGS and ACT, are usually conserved in RSHs and display regulatory role. The CTD region is reported to be involved in inter- and intramolecular interactions by which RSHs regulate their catalytic activity. These two identifies environmental signals and integrate the signal in order to regulate the (p)ppGpp accumulation. TGS domain (Pfam: PF02824) which is present next to the N-terminal synthetase domain is a small stretch of about 50 amino acids and is identified to have predominantly β sheets. The domain is generally conserved among proteins of GTPase and RSH superfamily and predicted to have role in regulatory function (Fig. 1.15).

Figure 1.15: Sequence alignment of TGS domain from mycobacterial and non-mycobacterial species. The sequence alignment of TGS domain of RelA/SpoT homologs from *Mtb* (*MtRel*), *M. bovis* (*MsRel*), *M. smegmatis* (*MsRel*), *M. leprae* (*MlRel*), *Porphyromonas gingivalis* (*PgRel*), *Clostridium leptum* (*ClRel*), and *E. coli* (*EcSpoT* and *EcRelA*) are shown. The sequence alignment shows the residues in shades of blue which are identical at the threshold of 30%. The sequence alignment was performed using Clustal Omega.

In addition to inter-domain and intramolecular interactions with the ribosome-activation-complex, studies have shown the regulation of catalytic activities of the RSH proteins also depends upon the other molecular interactions. For instance, *E. coli* SpoT has shown interactions with an acyl carrier protein (ACP) with the help of the TGS domain present in the CTD [156]. The interaction of ACP with SpoT regulates the accumulation of (p)ppGpp which was reported to be a hydrolase enzyme with very weak synthetase. ACP is a central cofactor of lipid metabolism and thereby provides stimulus to SpoT during fatty acid starvation and elicit stringent response[106].

In addition to the TGS domain, the ACT domain (Pfam: PF01842) is present at the far most C-terminus of RSH enzymes and is believed to regulate the activity of a wide range of metabolic enzymes, including RSHs which are regulated by amino acids/nucleotides concentration [199]. ACT domains are generally 70-80 residues long and exhibit a conserved $\beta\alpha\beta\beta\alpha\beta$ topology which is a characteristic of ferredoxin fold [199, 200]. ACT domains are recently further classified into RNA recognition motif (RRM) proteins as a class of these proteins interacts with RNA species as well to bring regulatory mechanism and show additional binding sites than present in ACT domains [165]. In a sequence alignment, the ACT domain from mycobacterial species show appreciable conservation among themselves however demonstrate poor conservation with the non-mycobacterial homologs (Fig. 1.16).

Figure 1.16: Sequence alignment of ACT domains from mycobacterial and non-mycobacterial species. The sequence alignment of ACT domain of RelA/SpoT homologs from *Mtb* (MtRel), *M. bovis* (MsRel), *M. smegmatis* (MsRel), *M. leprae* (MlRel), *Staphylococcus aureus* (SaRel), *Chlorobium tepidum* (CtRel), and *E. coli* (EcSpoT and EcRelA) are shown. The sequence alignment shows the residues in shades of blue which are identical at the threshold of 30% across species. The sequence alignment was performed using Clustal Omega.

Apart from the TGS and the ACT domain, there are two structural elements in between TGS and ACT domains of 150-200 amino-acid residues. The first structural element is present downstream to the TGS domain and is predominantly helical whereas the second structural element which is upstream to the ACT domain, is named as Zinc-finger domain (ZFD) or conserved cysteine (CC) domain. The two names are due to the presence of three conserved cysteine residues along with the ZFD-like fold.

Recent studies of Rel in *C. crescentus* (CcRel) have shown that the ACT domain is indispensable for the activity of the hydrolase domain and the binding of EIIA^{Ntr} ~P (an enzyme of nitrogen related Phosphotransferase (PTS) system) to the ACT inhibits the hydrolase activity [201]. The authors also reported loss of both the catalytic activities of CcRel upon truncation of the entire C-terminus harboring TGS and ACT. The truncated CcRel cannot accumulate (p)ppGpp upon nitrogen starvation [202, 203]. These studies further suggested the role of the TGS domain in the regulation of the synthetase domain. There are other studies that have also described the interactions between EIIA^{Ntr} proteins and Rel enzymes in *S. meliloti* and *Rhodobacter sphaeroides* which belong to α -proteobacteria and *Ralstonia eutropha* of β -proteobacteria [202, 204].

In addition to CTD mediated regulation of RSHs by ACP during fatty acid limitation and EIIA during nitrogen starvation as described in above sections, branched-chain amino acids (BCAA) have also been reported to interact with CTD of Rel proteins. In a recent study, RcRel (Rel protein of *Rhodobacter capsulatus*) have shown to interact

with branched-chain amino acids in order to stimulate the hydrolase activity [205]. This is the first report suggesting another feedback mechanism in order to control the (p)ppGpp accumulation upon BCAA restoration. However, there is no evidence on how such small molecules like BCAAs get differentiated among other amino acids and how Rel of different bacterial classes behave so differently in recognizing a particular BCAA. The authors have reported the interaction of BCAAs in 12 other RSHs and further validated the species-specific behavior in BCAA mediated regulation [205]. The essentiality of BCAAs has also been understood in *Mycobacterium* pathogenesis. For instance, the deletion of L-leucine synthesizing genes leads to the attenuation of virulence in *Mtb* [206]. Further, studies have shown that BCAAs are crucial for the growth and survival of *Mtb* [207].

Despite being conserved, the CTDs and their regulatory mechanism within RSH enzymes, particularly in mycobacterial species, are not explored much.

1.16 Role of oligomerization by CTDs in regulation of activity

The studies on *MtRel* and *MsRel* reported different states and the cause of oligomerization. According to Avarbock et al, *MtRel* forms a trimer by strong non-covalent interactions (Fig. 1.14), whereas *MsRel* was reported to form various multimers due to covalent interaction mediated by free thiol groups within the protein [184, 208]. In *MtRel* NTD, it was shown that substrates tend to monomerize the protein and makes it catalytically active. In *MsRel* it was shown that inter-protein covalent interactions by disulfide bonds were responsible to form higher oligomers and monomerize in the presence of reducing agent [208]. The study also reported the importance of the C-terminal residues 394-738 in interacting with the ribosome-tRNA-mRNA complex and the catalytic activities were reported to differ from maximal basal rate in *MtRel* in the absence of C-terminal domains [184]. The studies in *M. smegmatis* also reported that the *MsRel* tends to oligomerize *in vitro* whereas the monomer of the enzyme was reported to be synthetically active [208]. The mutational studies enabled to mark the residues in the CTD required for the activation of synthetase in response to tRNA and rendering the protein to be monomeric [208]. It was further investigated that in the presence of uncharged tRNA, the mutational effect of C692S was restored and thereby abolished the interaction of the CTD with the synthetase active site [198]. The studies performed in *MtRel* have also shown that deacylated tRNA increase synthetic

activity which further increases several folds in the presence of ribosome, mRNA and deacylated tRNA complex showing complete activation of synthetase [190].

Several other studies have identified the role of the CTD in the auto-inhibition of RSH synthetic activity. It was proposed that the C-terminal homo-oligomerization involving cysteines for RelA-RelA interaction leads to the deactivation of synthetic activity of RelA in *E. coli* [209]. In bifunctional enzymes, the best example of the autoinhibitory behavior of the CTD was reported from *S. equisimilis*. The report stated that in the absence of the CTD the synthetase activity of the *SeRel* was 12 fold higher and the hydrolase activity 150-times lower than that of the full length *SeRel* [193]. This intrinsic regulation explained the behavior of the CTD in controlling the accumulation of (p)ppGpp. A model proposed by Mechold et al. justified the involvement of extreme structural regions in the CTD in the regulation of the hydrolase domain of *MsRel* [193].

1.17 Goals of the thesis

The mechanism governing the regulation of catalytic activity of the *MtRel* remains structurally and mechanistically uncharacterized. However, in the previous studies biochemical data have been provided for the critical residues required for the catalytic activity by the NTDs of *MtRel* [184, 196]. Also, the binding sites in the C-terminal linker region responsible to provide feedback mechanism by sensing (p)ppGpp was characterized in *MsRel* [198]. In addition to that, the auto-regulation by the CTD and the role of CTD residues in oligomerization allowed researchers to propose a model of NTDs regulation independent to the ribosome-tRNA-mRNA complex based on conserved cysteine (in the linker region of TGS and ACT domain) mutations in *MsRel* [198, 208]. It has been studied that the *MtRel* can carry out both hydrolysis and synthetase activities at a maximal basal rate, unlike *SeRel* or *EcRelA*, although, the mechanism behind the regulation of activities independent of ribosome interactions has not been understood in *MtRel*. Moreover, the role of stringent response in virulence, long term persistence of the pathogen and biofilm formations further emphasises relevance as a drug target against persistent bacteria. The increase and decrease in (p)ppGpp beyond a certain concentration have also shown to be lethal for the pathogen. In order to manipulate its catalytic activity and thus making it potentially lethal by exploiting the (p)ppGpp concentration dependence for persistence and virulence, it is crucial to understand the mechanism that regulates these activities. Further, there are several intriguing questions that remain unanswered about bifunctional Rel proteins

including how the environment is sensed by the Rel protein in *Mtb* and the allosteric role of such environmental cues in regulating catalytic activities. Considering the former, I am interested in:

1. Describing the oligomerization state of *MtRel* in solution using biophysical approaches and proposing the first solution structure of *MtRel* using SAXS.
2. Determination of the first atomic-resolution structure of catalytic domains of *MtRel* by using x-ray crystallography and understanding the effect of substrates on the overall structure of *MtRel* NTD by using SAXS.
3. Producing the critical domain *MtRel* TGS to homogeneity to carry out novel atomic resolution structure studies while investigating its role in synthetase activity and tRNA binding.
4. Elucidating the unique *MtRel* ACT by generating and purifying the domain and studying with help of SAXS, NMR spectroscopy and mutagenesis experiments while underpinning the mechanism behind *MtRel* regulation by branched-chain amino acids.

The structural and mechanistic insights gained into the *MtRel* and the description of the unique features could contribute to:

- Understanding the species-specific features of *MtRel* and how it regulates in solution independent to the ribosome.
- Providing scope for vast chemical-space search by identifying residues in regulatory domains involved in feedback mechanism or environment sensing and not limiting to (p)ppGpp analogs for therapeutic development.
- Presenting a new layer of *MtRel* regulation by means of small molecules like branched-chain amino acids.

Finding epitopes in a species-specific manner will help in developing potent therapeutics against latent TB infection and providing direction to target dormant *Mtb* during stringent conditions such as amino acid starvation. The role of (p)ppGpp concentration balance in maintaining the nucleotide pool, development of chronic infection and initiating the global stress response and hence persistence, justifies that the *MtRel* is a suitable target for drug development.

2 Materials and Methods

2.1 Materials

2.1.1 Chemicals

All the chemicals used for these studies were of at least analytical grade. Chemicals were purchased from the following companies:

Buffers, Media and Salts	Sigma (St. Louis, MO, USA)
	USB (Sampscott, MA)
	Calbiochem (Darmstadt, Germany)
	Fluka (Sigma, Buchs Germany)
	Roth (Karlsruhe, Germany)
	Serva (Heidelberg, Germany)
Dyes, ATTO fluorescent	ATTO-TEC GmbH (Siegen, Germany)
DTT	Hoefer (San Francisco, CA, USA)
Ni ²⁺ -NTA	QIAGEN (Hilden, Germany)
Pefabloc ^{SC}	BIOMOL (Hamburg, Germany)
PMSF	Sigma (St. Louis, MO, USA)
LB Media	BD (Sparks, MD, USA)
N-Acetyl-L-cysteine	Sigma (St. Louis, MO, USA)
Antibiotics	Calbiochem, (Darmstadt, Germany), Sigma and Gibco, (St. Louis, MO, USA)
IPTG	Fermentas (Vilnius, Lithuania)
BSA	GERBU (Heidelberg, Germany)
ATP / ADP / AMPPNP, Imidazole, Thrombin	Sigma Aldrich (St. Louis, MO, USA)
Agarose, SDS, Glycine, APS	Bio-Rad (Hercules, CA, USA)

2.1.2 Molecular biology materials

2.1.2.1 Enzymes and molecular biology kits

plasmid DNA	
KAPA HiFi polymerase	KAPA Biosystems (UK)
Pfu DNA Polymerase	Fermentas (Glen Burnie, MD, USA)
<i>NcoI</i> , <i>SacI</i> , <i>BamHI</i> , <i>HindIII</i> , T4 DNA Ligase	Fermentas and New England Biolabs, (UK)
T4 Quick Ligase	Promega, (Wisconsin, USA)
Miniprep Plasmid Kit	Qiagen (Hilden, Germany)
Nucleobond AX mediprep Kit	MN & Co (Düren, Germany)
<i>Escherichia coli</i> strains	DH5 α , BL-21 (DE3), C41

<i>E. coli</i> (MRE 600) tRNA	Sigma (St. Louis, MO, USA)
QIAquick PCR purification and Gel extraction kits	Qiagen (Hilden, Germany)
pET-9d (+)	[210]
AMP-Glo™ Assay	Promega (Wisconsin, USA)

2.1.2.2 Primers, synthetic oligonucleotides and plasmids

Primers for generating the gene constructs of *MtRel* and RNA oligos were synthesized by IDT. All the *M. tuberculosis* Rel constructs were cloned into a modified pET-9d-1 vector [210]. The plasmid allows the expression of His-tagged protein at the N-terminus and also confers resistance towards kanamycin to facilitate the selection of positive clones.

2.1.2.3 *E. coli* strains

The electrocompetent DH5 α cell strain was used for all the cloning processes. The genotype of the strain is F⁻, ϕ 80*lacZ*ΔM15, Δ(*lacZYA-argF*), U169, *recA1*, *endA1*, *hsdR17*(r_K⁻, m_K⁺), *gal*⁻, *phoA*, *supE44*, λ *-thi* -1 *gyrA96*, *relA1*. The *recA1* mutation inactivates the recombinases and prevents homologous recombination, while the *endA1* mutation inactivates the endonucleases, thus preventing degradation of the transformed plasmid.

The electrocompetent BL21 (DE3) cell strain was used for the expression of *MtRel* constructs. The genotype of this strain was: F⁻, *ompT*, *hsdS_B* (r_B⁻, m_B⁻), *gal*, *dcm*, *lon* (DE3). These strains utilize the T7 RNA polymerase promoter to facilitate the overexpression of the desired gene product. These strains lack the Lon and OmpT proteases which would ensure that the recombinantly produced proteins are not degraded.

The electrocompetent C41 (DE3) strains were used for the expression of *MtRel* full length, *MtRel*₃₉₅₋₄₅₉, and *MtRel*₆₆₀₋₇₃₈ constructs. The strain C41(DE3) was derived from BL21 (DE3) and has genotype as: F⁻, *ompT*, *hsdS*(r_B⁻, m_B⁻), *dcm*⁺, *gal*, (DE3). This strain has atleast one uncharacterized mutation that reduces the level of T7 RNAP activity, thereby preventing cell death associated with overexpression of many recombinant toxic proteins. This strain is effective in expressing toxic and membrane proteins from all classes of organisms, including viruses, eubacteria, archaea, yeasts, plants, insects, and mammals.

The BL21 (DE3) strain with GroEL-ES containing pBAD vector was used for the expression of *MtRel* so as to facilitate the proper folding of protein in the presence of sufficient chaperones. To induce the promoter of pBAD in order to express sufficient GroEL-ES, 0.5% arabinose is used in addition to IPTG which is required to induce expression of gene of interest from pET-9d-1[210].

2.1.3 Chromatography materials

Immobilized affinity chromatography

Ni²⁺-NTA agarose

QIAGEN (Hilden, Germany)

Anioin Exchange chromatography

ResourceTMQ, 6ml

GE Healthcare (Uppsala, Sweden)

Gel filtration chromatography

Superdex 75 HR (10/30)

GE Healthcare (Uppsala, Sweden)

Superdex 200 HR (10/30)

GE Healthcare (Uppsala, Sweden)

Superdex 200 HR Increase (10/30)

GE Healthcare (Uppsala, Sweden)

Instruments

Akta FPLC, Akta Prime

GE Healthcare (Uppsala, Sweden)

Millex Filters (0.45 µM)

Millipore (Bradford, USA)

Syringe, needles and accessories

BD Biosciences, (San Jose, California, USA)

Protein concentration devices and estimation

Centriprep YM10

Millipore (Co-cork, Ireland)

Amicon ultra (100 kDa)

Millipore (Co-cork, Ireland)

BCA Assay

Kit Pierce (Rockford, IL, USA)

BioSpec-nano

Shimadzu Scientific, (Kyoto, Japan)

2.1.4 Other instrumentation

PCR Thermocycler:

Biometra T personal and gradient

Biometra, (Göttingen, Germany)

Sonoplus Sonicator

Bendelin, (Göttingen, Germany)

Micropulser Electroporator

Bio-Rad, (Göttingen, Germany)

Ultraspec 2100 Pro Spectrophotometer

Amersham Biosciences, (Little Chalfont, UK)

MetalJet-NANOSTAR

Bruker AXS (Karlsruhe, Germany)

CytationTM 5

BioTek instruments

2.1.5 Computer software

ATSAS 2.8.3	EMBL (Hamburg, Germany)
Adobe Photoshop cs6	Adobe Systems, San Jose, CA, USA
Coot	[211]
HKL2000 package	[212]
PROCHECK	[213]
Ramachandran plot	[214]
Refmac5 (CCP4 package)	[215]
MOLREP	[216]
PHASER	[217]
CNS	[218]
CCP4	[219]
PRIMUS	[220]
Gasbor	[221]
SUBCOMB	[222]
MEGA 6.0	[223]
Vector NTI 10.3.0	Invitrogen (California, United States)
PyMOL v1.5	DeLano Scientific LLC (USA) [224]
Quantity One	Bio-Rad (Hercules, CA, USA)
SPARKY	University of California (USA)
Topspin 1.3	Bruker Biospin (Karlsruhe, Germany)
Jalview 2.10.4b1	[225]
UCSF Chimera	[226]
Genome compiler	http://www.genomecompiler.com/about-genome-compiler/
Snapgene software	GSL Biotech; www.snapgene.com
MolProbity	[227]

2.2 Methods

2.2.1 Generation of the plasmid pET9d: *MtRel*, and pET9d: *MtRel*₁₋₃₉₄

The plasmids with a backbone pET9d-1 and the gene inserts coding for *MtRel*, and *MtRel*₁₋₃₉₄ respectively, were cloned between *NcoI*/*SacI* restriction sites and kindly provided by Dr. Wilson Nartey (laboratory of Prof. Dr. G. Grüber) [228]. The following forward (*NcoI* site in bold) and reverse (*SacI* site in bold) primers were used to carry out PCR:

Construct		Oligonucleotide sequence 5'-3'
<i>MtRel</i>	forward	TAG GCC ATG GGG ATG ACC GCC CAG CGC AGC ACC ACC
	reverse	GCT GAG CTC CTA CGC GGC CGA GGT CAC CCG GTA GAC
<i>MtRel</i> ₁₋₃₉₄	forward	TAG GCC ATG GGG ATG ACC GCC CAG CGC AGC ACC ACC
	reverse	CGC GGA GCT CCT AGT CGT AGC GCA ATG ATT CCA AG

Table 2.1: Primers for generation of *MtRel* and *MtRel*₁₋₃₉₄.

2.2.2 Generation of the plasmid pET-9d: *MtRel*₆₆₀₋₇₃₈

The plasmid with a backbone pET9d-His6 was manipulated and gene coding for *MtRel*₆₆₀₋₇₃₈ was introduced. The gene was PCR amplified from pET-9d-1 vector containing *MtRel* coding gene as a template. The following primers were used:

Construct		Oligonucleotide sequence 5'-3'
<i>MtRel</i> ₆₆₀₋₇₃₈	forward	CCG TCG TCC ATG GTT CTG GTG GCA ATC CAG GTC GAG GCA CTC
	reverse	GCT GAG CTC CTA CGC GGC CGA GGT CAC CCG GTA GAC

Table 2.2: Primers for generation of *MtRel*₆₆₀₋₇₃₈.

A total 50 µl of PCR reaction mixture was set up having the reagents listed in table 2.3 in appropriate concentrations as mentioned. The whole process was performed on ice. A quick centrifugation was done using the table-top centrifuge (Eppendorf Mini-centrifuge) to mix the reagents homogenously.

Reagents	Amount
Pfu Buffer (10x) + 20 mM Mg ²⁺	5 µl
DNTP's (10 mM)	1.5 µl
Primers (100 µM)	0.5 µl each
Template (plasmid DNA) (50 ng)	1 µl
Nuclease free water	40.5 µl
Pfu DNA polymerase	1 µl

Table 2.3: Reaction mix for PCR.

PCR thermocycler (Biometra T personal) was preheated to 99 °C before placing the reaction tubes in the PCR machine. The PCR thermocycler program mentioned in table 2.4 was used for amplification.

Cycle steps	Temperature	Time	
Lid	99 °C		
Initial denaturation	95°C	3 min	
Cycle denaturation	95°C	30 sec	} 25 cycles
Annealing	68°C	45 sec	
Extension	72 °C	1 min	
Final extension	72 °C	5 min	
End	4 °C	pause	

Table 2.4: PCR cycling program.

After amplification, 5 µl of the PCR product were applied onto an analytical agarose gel (1.2%) to check its quality and the rest was stored at -20 °C. The gel was subsequently stained in ethidium bromide solution for 20 min and observed under UV light. The size of the PCR product was confirmed with the appropriate DNA marker. After confirming the size of the PCR product, a preparative agarose gel (1.5%) was run with the remaining reaction mixture (45 µl). The gel portion having the correct PCR product band was cut out immediately and purified by gel extraction kit (QIAGEN) following the manufacturer's protocol. Finally, to elute the DNA, 30 µl of elution buffer (10 mM Tris-HCl buffer, pH 8.5) was used. The purified DNA was used for the following sequential steps to get the plasmid transformed in *E. coli* DH5α:

1. Double digestion:

To introduce the sticky ends to the amplified and purified DNA, double digestion was performed for 1-1.5 hours at 37 °C by using appropriate amount of *NcoI* and *SacI* enzymes.

2. PCR purification:

After fast digestion, the reaction mixture was purified by QIAquick PCR purification kit (Qiagen) so as to remove the salts and enzymes from the buffers used for double digestion.

3. Ligation:

Subsequently, ligation reaction between *NcoI* and *SacI* pre-digested vector and amplified PCR product were setup at room temperature for 1 hour. The reagents used were as follows:

Reagent	Amount
Ligase buffer (10x)	1 µl
Vector (V)	60 ng
Insert (I)	16.32 ng
T4 Ligase	1 µl
MilliQ water	5.58 µl
Total	10 µl

Table 2.5: Reaction mix for ligation of DNA fragments

4. Butanol DNA precipitation:

In order to isolate DNA for subsequent transformation into the electro-competent cells, 90 µl of Milli Q water was added to the ligation reaction mixture to make a final volume to 100 µl. 1 ml of butanol was added into the mixture and mixed thoroughly. The mixture was centrifuged at 13,000 x g (using Eppendorf Mini-centrifuge) for 10 min. and the supernatant was discarded gently. The pellet was resuspended in 200 µl of 70% ethanol and centrifuged again for 5 min. Supernatant was removed and the pellet in the tube was left to dry inside the sterile hood for 10 minutes. The residual ethanol evaporated eventually. Pellet was re-dissolved in 10 µl of MilliQ water on ice.

5. Transformation by electroporation:

E. coli DH5 α electro-competent cells were prepared as per the protocol from “Current Protocols in Molecular Biology” (Wiley InterScience) manual. A volume of 10 μ l of ligation mixture was added to 80 μ l of competent cells and incubated for 1-2 min on ice. The mixture was transferred to a pre-chilled electroporation cuvette. Electroporation was performed using constant voltage of 2.5 kV on Micropulser Electroporator, Bio-Rad. The cuvette was then immediately placed back on ice, following washing of cells with 910 μ l of LB media. This new culture of freshly transformed cells was transferred into an Eppendorf tube and incubated at 37 °C for 1 hour. Appropriate dilutions were prepared to plate on the antibiotic selection plates. The plates contained 30 μ g/ml of kanamycin and the culture spread was incubated overnight at 37 °C. Single colonies were picked from the plate on the next day and were incubated at 37°C in the Luria-Bertani (LB) media. The LB Media was prepared by mixing 10g/L NaCl, 5g/L yeast extract and 10g/L tryptone distilled water. The plasmid DNA was isolated using standard protocol (QIAGEN mini-prep kit), subsequently double digested with NcoI and SacI and applied onto 1.2% agarose gel, in order to confirm the presence of the insert. Size of the insert and the vector were compared with appropriate controls and markers. Plasmids were sent for sequencing and the verified ones were finally transformed into *E. coli* C41(DE3) in the similar way. The cells were cultivated on 30 μ g/ml kanamycin-containing LB agar plates. The colonies were picked and inoculated in LB media at 37 °C to prepare glycerol stocks.

2.2.3 Site-directed mutagenesis of *MtRel*₆₆₀₋₇₃₈

In order to incorporate desired mutations for this study, mutagenic primers were used to clone the entire plasmid. The mutants *MtRel*₆₆₀₋₇₃₈ H12A, *MtRel*₆₆₀₋₇₃₈ L15A, *MtRel*₆₆₀₋₇₃₈ A23S and *MtRel*₆₆₀₋₇₃₈ N28A were generated by using the primers mentioned in Table 2.6 and template pET9-d1-His6 plasmid carrying coding strand for *MtRel*₆₆₀₋₇₃₈.

Construct		Oligonucleotide sequence 5'-3'
<i>MtRel</i> ₆₆₀₋₇₃₈ H12A	forward	GAC CGG <u>GCT</u> CGG CTG CTG TCG GAT GTG
	reverse	CAG CCG <u>AGC</u> CCG GTC GAG TGC CTC GAC
<i>MtRel</i> ₆₆₀₋₇₃₈ L15A	forward	GG CTG <u>GCG</u> TCG GAT GTG ACG CGC G
	reverse	ATC CGA <u>CGC</u> CAG CCG GTG CCG GTC
<i>MtRel</i> ₆₆₀₋₇₃₈ A23S	forward	TG ACG CGC GCA CTG <u>AGT</u> GAC GAG

<i>MtRel</i> ₆₆₀₋₇₃₈ N28A	reverse	T CTC GTC <u>ACT</u> CAGT GCG CGC GT
	forward	AAG GTC <u>GCT</u> ATC CTG TCC GCG TCG
	reverse	GA CAG GAT <u>AGC</u> GAC CTT CTC GTC GGC

Table 2.6: Primers for generation of *MtRel*₆₆₀₋₇₃₈ mutants.

In the PCR, 10 ng of plasmid DNA encoding *MtRel*₆₆₀₋₇₃₈ were used as a template. Total reaction volume has been set to 50 µl and pipetting of the reaction mixture to PCR tubes was done on ice for all the steps. The reaction was set following the order and the volumes of the reagents in the Table 2.7. The mixture was homogenized by pipetting and centrifugation on a conventional table-top centrifuge (Eppendorf Mini-centrifuge).

Reagents	Amount
5X KAPA HiFi GC Buffer	10 µl
KAPA dNTP mix (10 mM)	1 µl
Primers (100 µM)	0.5 µl each
Template (plasmid DNA) (~10 ng)	1 µl
MilliQ water	36.5 µl
KAPA DNA polymerase	0.5 µl

Table 2.7: Reaction mix for site directed mutagenesis PCR.

PCR Thermocycler (Biometra T personal) was preheated to 99 °C before placing the PCR tubes on the heating block. The following program was used for the amplification:

Cycle steps	Temperature	Time	
Lid	99 °C		
Initial denaturation	95 °C	3 min	
Cycle denaturation	95 °C	20 sec	} 18 cycles
Annealing	64.4 °C	15 sec	
Extension	72 °C	2 min 30 sec	
Final extension	72 °C	5 min	
End	4 °C	pause	

Table 2.8: PCR cycling program for complete plasmid.

To check the quality, 5 μ l of each the PCR product was applied on an 1.2% agarose gel. The gel was subsequently stained with Ethidium bromide (EtBr) solution for 20 min and observed under UV light upon de-staining. The size of the products was confirmed by comparing it with the appropriate DNA marker. For each of the constructs, the remaining reaction mixture (45 μ l) was used to isolate the amplified plasmid DNA containing the mutation and implemented to following protocol:

Figure 2.1: Schematic diagram of site-directed mutagenesis. Two step purification for *MtRel* by using affinity chromatography and gel filtration chromatography.

1. *DpnI* digestion and inactivation:

To digest the methylated plasmid DNA which was used as the template in PCR, restriction with endonuclease *DpnI* was setup. The components listed in table 2.9 were pipetted into a tube and mixed till homogeneity. The reaction mixture was incubated for 1 h in water bath at 37 °C followed by 20 min of incubation on 80 °C to inactivate the *DpnI* endonuclease.

Reagents	Volume (μ l)
PCR mixture	30.0
CutSmart® buffer (10x)	35.9
<i>DpnI</i>	2 x 0.30
MiliQ water	1.50

Table 2.9: Reaction mix for *DpnI* digestion.

2. DNA precipitation:

In order to isolate DNA for subsequent transformation into the electro-competent cells, 40 µl of MiliQ water followed by 900 µl of *n*-butanol (HPLC grade) was added to the reaction mixture. After thorough mixing and homogenizing, DNA was precipitated out using the conventional table-top centrifuge (Eppendorf Mini-centrifuge) for 10 min on 13,000 x g. The supernatant was pipetted out and replaced with 500 µl of 70% ethanol (HPLC grade) to remove the residual butanol. Centrifugation was repeated and ethanol was pipetted out. The tube containing the pellet with DNA was inverted on a paper inside of the sterile hood in duration of 10 minutes to allow residual ethanol to evaporate. Dry DNA pellet was re-dissolved in 7 µl of MiliQ water on ice in duration of 30 min.

3. Transformation into electrocompetent cells:

The mutated DNA was transformed into electrocompetent *E. coli* DH5a and later in *E. coli* C41 (DE3) by following the protocol mentioned as subsection “Transformation by electroporation” under section 2.2.2.

2.2.4 Generation of the *MtRel*₃₉₅₋₇₃₈ construct

To generate this construct, same template as mentioned in section 2.2.2 was used and PCR reaction was carried out by using the following primers having *NcoI* and *SacI* restriction site respectively.

Construct		Oligonucleotide sequence 5'-3'
<i>MtRel</i> ₃₉₅₋₇₃₈	forward	CAT TGC <u>CCA TGG</u> ACC TTG CGG CGG TGC AA
	reverse	GCT <u>GAG CTC</u> CTA CGC GGC CGA GGT CAC CCG GTA
		GAC

Table 2.10: Primers used for the generation of *MtRel*₃₉₅₋₇₃₈.

The complete process was performed on ice and all the steps were performed as stated for *MtRel*₆₆₀₋₇₃₈ in section 2.2.2. The plasmid encoding *MtRel*₃₉₅₋₇₃₈ was transformed into *E. coli* BL21 (DE3) cells for protein expression and production.

2.2.5 Generation of the *MtRel*₃₉₄₋₄₅₉ construct

Construct *MtRel*₃₉₄₋₄₅₉ was PCR amplified by using the following primers having *NcoI* and *SacI* sites.

Construct		Oligonucleotide sequence 5'-3'
<i>MtRel</i> ₃₉₄₋₄₅₉	forward	TAG <u>GCC ATG GGG</u> ATG ACC GCC CAG CGC AGC ACC ACC
	reverse	TGC GGC <u>GAG CTC</u> CTA CGT GAA AAC CTC GAC AAC TTC TCC G

Table 2.11: Primers used for the generation of *MtRel*₃₉₄₋₄₅₉.

The PCR reaction was carried out as follows to a total volume of 50 µl:

Reagents	Amount
5X KAPA HiFi buffer with 2 mM MgSO ₄	10 µl
DNTP's KAPA (10 mM)	1.5 µl
Primers (100 µM)	0.25 µl each
Template (plasmid DNA) (~50 ng)	1 µl
MilliQ water	36 µl
KAPA DNA polymerase	1 µl

Table 2.12: Reaction mix used for *MtRel*₃₉₄₋₄₅₉.

PCR thermocycler (Biometra T personal) was preheated to 99 °C before placing the reaction tubes in the PCR machine. The following PCR thermocycler program was used for amplification:

Cycle steps	Temperature	Time
Lid	99 °C	
Initial denaturation	98°C	3 min
Cycle denaturation	98°C	20 sec
Annealing	66°C	45 sec
Extension	72 °C	1 min
Final extension	72 °C	5 min
End	4 °C	pause

Table 2.13: PCR cycling program to amplify *MtRel*₃₉₄₋₄₅₉.

After amplification, similar protocol as described in section 2.2.2 of *MtRel*₆₆₀₋₇₃₈ was followed.

2.2.6 Generation of the *MtRel*₁₋₄₅₉ construct using mutagenesis

The gene coding for pET9d-1: *MtRel* was manipulated by introduction of a stop codon after TGS domain coding sequence. The introduction of stop codon led to amplification of desired length of gene coding for 1 - 459 amino acid residues from *MtRel*.

Construct	Oligonucleotide sequence 5'-3'	
<i>MtRel</i> ₁₋₄₅₉	forward	TTC ACG <u>TAG</u> TCC AAG GCG CCG AAC GC
	reverse	GCC TTG <u>GAC</u> TAC GTG AAA ACC TGG AC

Table 2.14: Primers used for the generation of *MtRel*₁₋₄₅₉

The PCR reaction was carried out as follows to a total volume of 50 µl:

Reagents	Amount
5X Phusion HiFi buffer with MgSO ₄	10 µl
dNTP's mix (10 mM)	1 µl
Primers (100 µM)	0.25 µl each
Template (plasmid DNA) (~10 ng)	0.14 µl
MilliQ water	36.36 µl
DMSO	1.5 µl
Phusion DNA polymerase	0.5 µl

Table 2.12: Reaction mix used for *MtRel*₃₉₄₋₄₅₉

A preheated PCR thermocycler (Biometra T personal) at 99 °C was used to carry out PCR. The reaction tubes mixed were placed in the thermocycler for the amplification to start. The following PCR thermocycler program was used for amplification:

Cycle steps	Temperature	Time
Lid preheated	98 °C	
Initial denaturation	98°C	2 min
Cycle denaturation	98°C	10 sec
Annealing and extension	71.5°C	3 min 30 sec
Final extension	72 °C	2 min
End	4 °C	pause

Table 2.13: PCR cycling program to amplify *MtRel*₃₉₄₋₄₅₉

After amplification, similar protocol as described in section 2.2.2 of *MtRel*₆₆₀₋₇₃₈ was followed.

2.2.7 Recombinant expression of *MtRel* and truncated constructs

2.2.7.1 Expression of *MtRel*, *MtRel*₁₋₃₉₄, *MtRel*₆₆₀₋₇₃₈ and *MtRel*₃₉₅₋₇₃₈ in *E. coli* BL21 (DE3) strain

Verified plasmids carrying the gene for *MtRel*₆₆₀₋₇₃₈ and *MtRel*₃₉₅₋₇₃₈ respectively were finally transformed into *E. coli* BL21 (DE3) cells for protein production, and grown on 30 µg/ml kanamycin-containing LB agar plates. *E. coli* BL21 (DE3) expression strains were purchased from Novagene (Darmstadt, Germany). Electroporation was done for the transformation of plasmids into electrocompetent cells, which were prepared as per the protocol from Current Protocols in Molecular Biology (Wiley InterScience) manual. 1 – 50 ng of plasmid was added to 80 – 100 µl of competent cells in an electroporation cuvette and incubated on ice for 5 minutes. Electroporation was done at constant voltage of 2500 V (Micropulser Electroporator, Bio-Rad). The cuvette was then immediately placed back on ice for 1 minute, following addition of LB media to total volume of 1 ml. The mixture was incubated at 37 °C for 1 hour with shaking. Appropriate dilutions were prepared before plating the cells on kanamycin selection plates. The expression strains of *E. coli* BL21 (DE3) for *MtRel*, *MtRel*₁₋₁₈₁, *MtRel*₁₋₃₉₄ were kindly provided by Dr. Wilson Nartey as glycerol stocks.

2.2.7.2 Expression of *MtRel*, *MtRel*₃₉₄₋₄₅₉, *MtRel*₁₋₄₅₉ and *MtRel*₆₆₀₋₇₃₈ wild type and its mutants in *E. coli* C41 (DE3) strain

Verified plasmids of *MtRel*, *MtRel*₃₉₄₋₄₅₉ and *MtRel*₆₆₀₋₇₃₈ were transformed into *E. coli* C41 (DE3) cells for protein production and grown on 30 µg/ml kanamycin-containing LB agar plates. *E. coli* C41 (DE3) expression strain was provided by Dr. Arvind Kumar (prepared in Prof. Dr. Gerhard Grüber's lab). Electroporation was done for the transformation of plasmids into electrocompetent cells, which were prepared as per the protocol from Current Protocols in Molecular Biology (Wiley InterScience) manual. 1 – 50 ng of plasmid was added to 80 – 100 µl of competent cells in an electroporation cuvette and incubated on ice for 5 minutes. Electroporation was done at constant voltage of 250 V (Micropulser Electroporator, Bio-Rad). The cuvette was then immediately placed back on ice for 1 minute, following addition of LB media to total

volume of 1 ml. The mixture was incubated at 37 °C for 1 hour with shaking. Appropriate dilutions were prepared before plating the cells on kanamycin selection plates.

2.2.7.3 Expression of *MtRel* in *E. coli* BL21 GroEL-ES strain

To ensure proper folding of proteins, chaperones are required as they assist the nascent polypeptide to achieve the right folding pattern to form a stable and well folded protein. The GroEL-ES class of chaperone-chaperonin in *E. coli* aid folding of synthesized proteins on ATP consumption. The GroEL-ES chaperone-chaperonin system forms a cage encapsulating the substrate protein, which might be a partially folded intermediate[229].

Using ATP, the folding is achieved and the folded protein is released with the disassembly of the GroEL-ES complex and making available the cavity for binding another substrate protein. This system was employed for folding the overexpressed *MtRel*. The cell strain of *E. coli* BL21 GroEL-ES containing pBAD vector was used, which codes for GroEL-ES chaperone-chaperonin and over expresses this, to allow sufficient population of the chaperone to be present to facilitate the folding of the recombinant protein. The promoter present in this vector is induced using 0.5% arabinose. Heat shock competent cells were kindly provided and further prepared by Dr Arvind Kumar in Prof. Dr. Gerhard Grüber's lab as per the following protocol as given by Prof. Oliver Mueller-Cajar's lab (SBS, NTU).

A tube of frozen competent cells was equilibrated on ice for 20 mins with the plasmid (pET9d with *MtRel*). Heat shock was performed by placing this mixture on a water bath set to 42 °C for 90 secs. and immediately returning this to ice for 2 mins. This was re-suspended in 200 µl of LB and allowed to revive for 30 mins at 37 °C in a shaking incubator. The re-suspended cells were plated on Kanamycin-Chloramphenicol LB plates and placed at 37 °C overnight. Single colonies obtained the next day, were picked and checked for expression of the desired protein.

2.2.8 Isotopic labelling of proteins

2.2.8.1 Minimal Media (M9) preparation

Reagents used for 1 litre minimal media:

42 mM Na₂HPO₄

2 mM MgSO₄

22 mM KH ₂ PO ₄	30 uM FeCl ₃
8.5 mM NaCl	5 ng thiamine
1g/l ¹⁵ NH ₄ Cl (labelled)	10 g glucose (unlabelled) or 2g ¹³ C (labelled)
0.1 mM CaCl ₂	

¹⁵N-labeled *MtRel*₆₆₀₋₇₃₈ and its mutants, ¹³C-¹⁵N-labelled and selectively labelled (Alanine/Arginine/Valine) *MtRel*₆₆₀₋₇₃₈, were prepared from freshly transformed *E. coli* C41 (DE3) cells. Single colony was picked to inoculate 50 ml Kanamycin supplemented LB broth and kept overnight to grow at 37 °C. Subsequently, the culture was centrifuged at 6,000 x g for 10 mins RT at 4 °C to pellet out the cells. The pellet was re-suspended and washed in M9 media. An appropriate volume was used to inoculate 1 liter of M9 media so as to achieve an O.D. of 0.1. Cells were incubated at 37 °C till an OD₆₀₀ of 0.6 was reached. The culture was induced with 1 mM Isopropyl-β-D-thiogalactoside (IPTG) a chemical analogue of galactose, to aid in protein production. Cells were incubated overnight at 18 °C and pelleted to purify the recombinant protein of desired label.

2.2.9 Production of recombinant *MtRel* and truncated constructs

2.2.9.1 Induction test

Reagents for 4 x lysis buffer: 250 mM Tris/HCl, pH 6.8

9.2% SDS

40% Glycerol

0.2% Bromophenol Blue

1 mM DTT

At least 3-4 single colonies were randomly picked up and inoculated in LB media so as to reach OD₆₀₀ of 0.6 - 0.7 at 37 °C by shaking at 180 rpm (Infors HT Minitron shaker). The colonies were then induced with Isopropyl β-D-1-thiogalactopyranoside (IPTG) with a final concentration of 1 mM and a non-induced control was kept for each colony. The induced and non-induced culture were incubated at 37 °C with shaking for 4 hours. Cells were then pelleted down and resuspended with 60 µl of 2x lysis buffer in the presence of 1 mM DTT. Resuspended cell mass was heated at 95 °C for 5 min and 15 µl of each sample was loaded onto a 17% SDS polyacrylamide gel [230]. The expression was more in the induced cultures ((+) supplemented with IPTG) than un-induced cultures ((-) not supplemented with IPTG).

2.2.9.2 Solubility test

Following the successful induction of proteins with IPTG at the optimised temperature and shaking conditions, solubility of the produced proteins was tested in various buffers mentioned below:

50 mM Tris-HCl, pH 8.5, 200 mM NaCl

50 mM Tris-HCl, pH 7.5, 200 mM NaCl

50 mM Hepes, pH 7.5, 200 mM NaCl

50 mM MES, pH 7.0 200 mM NaCl

50 mM Phosphate, pH 6.8 200 mM NaCl

50 ml of culture were grown and induced with IPTG as described above. Cells were then pelleted down into five equal fractions by centrifuging at 10,000 x g for 10 minutes. Freshly prepared protease inhibitors Pefabloc^{SC} in water and/or phenylmethylsulfonyl fluoride (PMSF, dissolved in 100% isopropanol) were added to a final concentration of 2 mM (Pefabloc^{SC}) or 1 mM PMSF. Reducing agent such as dithiothreitol, DTT (1 mM) was used in the preparation. Resuspension was sonicated at 15% power with KE 76 tip of sonicator (Bandelin Sonoplus) three times for one minute each, with a cooling interval of two minutes between each cycle. Supernatant was then separated from the pellet by centrifugation at 10,000 x g and transferred to a fresh tube. Pellets were resuspended in the respective buffers. All the steps were carried on ice. 15 µl of pellet and supernatant were loaded onto a 17% SDS gel [230].

2.2.10 Large scale production of recombinant *MtRel* and truncated constructs

2.2.10.1 Production of *MtRel*, *MtRel*₁₋₃₉₄, *MtRel*₃₉₄₋₄₅₉, *MtRel*₁₋₄₅₉, *MtRel*₃₉₅₋₇₃₈ and *MtRel*₆₆₀₋₇₃₈ in *E. coli* BL21 DE3 or C41 DE3 cell strains

Overnight inoculum of the transformed plasmid having *MtRel* or *MtRel*₁₋₃₉₄ coding gene in *E. coli* BL21 DE3 was setup in Kanamycin enriched LB at 37 °C. The starter culture was used next day to scale up the large culture, to a starting OD600 of 0.1. To produce the construct, liquid cultures were incubated in Kanamycin (30 µg/ml) LB medium at 37 °C with shaking of 180 rpm, until an OD600 of 0.6 - 0.7 was achieved. Expression of was obtained by adding IPTG to a final concentration of 1 mM. Following incubation for another 16-18 hrs at 15°C, the cells were harvested at 7,000 x g for 15 mins, at 4 °C on Beckman Coulter Avanti J26X centrifuge. The cells were frozen in

liquid Nitrogen and stored in -80 °C until further use. Likewise, the *MtRel*, *MtRel*₃₉₄₋₄₅₉, *MtRel*₁₋₄₅₉, *MtRel*₃₉₄₋₇₃₈ and *MtRel*₆₆₀₋₇₃₈ were produced from *E. coli* C41 DE3 strain.

2.2.10.2 Production of *MtRel* in *E. coli* BL21 GroEL-ES cell strain

Overnight inoculum of the transformed plasmid in *E. coli* BL21 GroEL-ES was setup in Kanamycin and Chloramphenicol enriched LB at 37 °C. This was used to scale up the large culture next morning, to a starting OD600 of 0.1. To produce *MtRel*, liquid cultures were inoculated in >2L LB media enriched with Kanamycin (30 µg/ml) and Chloramphenicol at 37 °C with shaking of 180 rpm, until an OD600 of 0.6 - 0.7 was achieved. Expression of all the constructs was obtained by adding IPTG (1 mM). Following incubation for another 16-18 hrs at 24 °C, the cells were harvested at 7,000 x g for 15 mins, at 4 °C on Beckman Coulter Avanti J26X centrifuge. The cells were immediately frozen in liquid nitrogen and stored in -80 °C in aliquotes.

2.2.11 Purification of *MtRel* constructs

Cells producing desired proteins were lysed on ice by sonication with an ultrasonic homogenizer (Bandelin, KE76 tip) for 3 x 1 min in respective buffer (refer Table 2.14) with 2 mM PMSF, 2 mM Pefabloc^{SC} and 0.8 mM DTT.

Construct	Buffer	Superdex column for Gel filtration	Extinction coefficient (M ⁻¹ cm ⁻¹)
<i>MtRel</i>	50 mM Tris-HCl (pH 8.5), 750 mM NaCl, 10 % glycerol	Superdex 200 HR Increase (10/30)	79,300
<i>MtRel</i> ₁₋₃₉₄	50 mM Tris-HCl (pH 8.5), 350 mM NaCl, 5 % glycerol	Superdex 200 HR Increase (10/30)	55,350
<i>MtRel</i> ₆₆₀₋₇₃₈	50 mM Tris-HCl (pH 8.5), 350 mM NaCl, 5 % glycerol	Superdex 75 HR (10/30)	2,980
<i>MtRel</i> ₃₉₅₋₇₃₈	50 mM Tris-HCl (pH 8.5), 750 mM NaCl, 10 % glycerol	Superdex 200 HR Increase (10/30)	23,950
<i>MtRel</i> ₃₉₄₋₄₅₉	50 mM Tris-HCl (pH 8.5), 750 mM NaCl, 10 % glycerol	Superdex 75 HR (10/30)	1490
<i>MtRel</i> ₁₋₄₅₉	50 mM Tris-HCl (pH 8.5), 750 mM NaCl, 10 % glycerol	Superdex 200 HR Increase (10/30)	56840

Table 2.14: Buffers used for the purification of *MtRel* and constructs

The sonicated cell lysate was centrifuged at 13,500 x g for 35 min at 277 K. The resulting supernatant was filtered by using 0.45 µm (Millipore, Billerica, MA, USA) filter and supplemented with Ni²⁺-NTA resin (pre-equilibrated in the respective buffer). The purification was carried out in two steps by using the purification scheme shown in Fig. 2.2.

The His-tagged protein was allowed to bind to the matrix for 1.5 hours at 277 K by mixing on a sample rotator (Neolab) after which the Ni²⁺-NTA was thoroughly washed with 10 column volumes of buffer (Table 2.14). The protein was subsequently eluted with an imidazole gradient (0–450 mM). The fractions containing the recombinant protein were identified by SDS–PAGE [230], pooled together and concentrated. The proteins were then applied onto a gel filtration column, Superdex column (GE Healthcare) as described in Table 2.14. The buffer used for the second step was the same as of buffers used for affinity

Figure 2.2: Schematic diagram of protein purification protocol. Two step purification for *MtRel* and constructs by using affinity chromatography and gel filtration chromatography.

chromatography except for constructs *MtRel*₃₉₅₋₇₃₈, *MtRel*₃₉₄₋₄₅₉ and *MtRel*₁₋₄₅₉. Gel filtration chromatography of these constructs was done using buffer of composition 50 mM Tris-HCl (pH 8.5), 350 mM NaCl, 5 % glycerol. For obtaining the reduced protein, DTT was added to the buffer for gel filtration column to a final concentration of 1 mM. The purity and homogeneity of protein sample was analysed by loading onto a 17% SDS-PAGE gel and stained with Coomassie Brilliant Blue R250. Protein concentrations were determined by absorbance at 280 nm in a spectrophotometer (BioSpec-Nano, Shimadzu) using an extinction coefficient mentioned in the Table 2.14.

2.2.12 Analysis of synthetase activity of *MtRel* and constructs using luminescence assay

The synthetase activity of the RSH enzymes takes place in the presence of substrates ATP and GTP/GDP leading to production of (p)ppGpp and AMP. The pyrophosphate group from ATP is transferred to GDP/GTP with the help of divalent

cation Mg^{2+} . To identify catalytically active constructs in the presence of substrates, luminescence assay to detect the production of AMP as a result of transferase enzymatic activity was used. AMP-Glo™ Assay is a luminescence based and generates a light signal upon production of AMP as a result of a reaction.

The assay is designed in such a way that it can be used to quantitatively determine the AMP concentration formed as a product of enzymatic reaction in a high throughput manner. The assay has two reagents: first one terminates the enzymatic reaction and simultaneously removes the ATP and also converts the AMP to ADP as produced; whereas the second reagent converts the ADP to ATP and the ATP to light with the help of luciferin/luciferase reaction.

The following schematic shows the principle of the assay:

Figure 2.3: Schematic diagram showing luminescence based AMP glo Assay. The assay can be used for all the enzymes which produces AMP as a result of catalytic reaction. The substrate does not affect the light signal generated in enzyme coupled reaction.

The standards reaction was set up in a master mix reaction buffer of ATP (100 μ M), GTP (100 μ M), $MgCl_2$ (200 μ M), BSA (0.1 mg/ml) and DTT (1 mM) dissolved in 50 mM Tris at pH 8.5, 350 mM NaCl and 5% glycerol. For standards, pure AMP was used in the concentration range of 10 μ M to 0.625 μ M serial diluted by using the reaction buffer. The assay was repeated several times to get consistent values. The reaction with *MtRel*, *MtRel* NTD, *MtRel* NTD-TGS and mutant *MtRel*_H671A was set up in the same reaction buffer for 1 hour at room temperature. The serial diluted standards and the AMP generating reaction carried out with *MtRel* and constructs were incubated/stopped with AMP-Glo™ Reagent I for 1 hour at room temperature after mixing. The AMP-Glo™ Reagent II along with Kinase –Glo One solution was used to prepare AMP detection solution as instructed in the assay's protocol upon completion of first step of incubation. The AMP detection solution was added to the above reactions and mixed. As step 1, the reactions were incubated for 1 hour at room temperature after which luminescence was recorded by Cytation™ 5 instrument. The ATP, AMP, AMP-Glo™ Reagent I, AMP-

Glo™ Reagent II and Kinase-Glo® One Solution were used from the Promega's AMP-Glo assay Kit (1000 assays).

2.2.13 Activity analysis of the *MtRel*₁₋₃₉₄ using HPLC

To determine the activity of *MtRel*₁₋₃₉₄, Ion-Pair Reverse-Phase High-Performance liquid chromatography (IP-RP HPLC) was used. The method was adopted from the study of RelA protein of *Francisella tularensis* [231]. The samples (40 µl) were injected on a reverse phase column (Agilent Eclipse XDB-C18 4.6 x 150 mm, 5 micro meter) was used. The HPLC was performed at a flow rate of 1 ml/min with UV detection at 260 nm. Two phases i.e. mobile aqueous phase (95% water with 5% methanol) and an organic phase (20% water and 80% methanol) were prepared. N,N-

dimethylhexylamine (DMHA) was added to a concentration of 15 mM. Acetic acid was used to adjust the pH to 7.0. The nucleotides were eluted by using the following gradient: 0–5 min isocratic, 25% organic buffer; 5–27 min, gradient 25%-60%; 27-28 min, gradient 60%-100%; 28-33 min, isocratic, 100%; 33-34 min, gradient 100%-25%; 34-44 min, isocratic 25%. The assay buffer having 20 mM Tris/HCl (pH 8.0), 15 mM KCl, 15 mM MgCl₂ and 1 mM β-mercaptoethanol was used to prepare the reaction mixture (1 ml) containing *MtRel*₁₋₃₉₄ (10 µM). The nucleotides GTP or ATP (2 mM each) were added to saturate the reaction mixture. The reactions were incubated at 30 °C and aliquots (100 µl) were withdrawn at several time points. The reactions were quenched by keeping the aliquots at 80 °C for 2 min. The protein was then fully precipitated by keeping the aliquots on ice for 10 mins. The precipitated protein was separated from the supernatant by centrifuging the aliquots at 13,500 x g, 4 °C for 5 min. The supernatant was analysed by using time-course HPLC for the activity of the recombinant *MtRel*₁₋₃₉₄.

Figure 2.4: Flow chart for determining the activity of *MtRel*₁₋₃₉₄. The protocol was adapted from *Francisella tularensis* RelA study [231].

2.2.14 Crystallization of *MtRel*₁₋₃₉₄

The crystal drops were set up for *MtRel*₁₋₃₉₄ at concentration varying from 3 mg/ml to 16 mg/ml. To perform the screening of crystallization conditions, vapour diffusion methods were opted namely hanging drop and sitting drop at 18 °C. The solutions used to initiate nucleation of the protein were used from Hampton Research crystal screen HR-110 and HR-112, Emerald Biosystems Wizard screen I, II, III, IV, and Molecular Dimensions Morpheus Screens I and II. Optimization was done by using different drop sizes (1-4 µl) having varied ratio of protein to the buffer (1:1, 1:2, 2:1, 2:2). The optimal conditions selected and drops of size 2 µl were set up on siliconized coverslips per well in an 1:1 ratio of protein in buffer 50 mM Tris-HCl pH 8.5, 350 mM NaCl, 5% glycerol and 1 mM DTT at 291 K. Many hits were observed in different conditions from Hampton crystal screen 1, Wizard Screen 2 and Wizard screen 4. The 2D plate like crystals obtained in Hampton crystal screen 1 condition #7 (0.1 M Sodium Cacodylate Trihydrate pH 6.5, 1.4 M Sodium Acetate Trihydrate) were first confirmed to be protein crystals. The optimization was carried out by performing grid screen in varying concentrations of precipitant, salt, buffer pH and protein concentration (Fig. 2.5). Cryo-protectant solution was also optimized by varying percentages of glycerol. The final optimised crystals ere quickly dipped in cryo-protectant solution and were flash frozen in liquid nitrogen at 100 K. Optimised *MtRel*₁₋₃₉₄ crystals were tested in-house at 100 K on a Rigaku RAXIS IV image plate detector with a Rigaku RA-Micro 7 HFM rotating copper anode generator (Rigaku/MSC). The optimized crystals diffracted poorly. Hence, we proceeded with optimizing the crystals in Wizard Screen 4.

Figure 2.5: Schematic diagram of crystallization optimization strategies. Steps followed to optimize the crystallographic conditions to improve size and x-ray diffraction of the crystals.

Well-formed 3D crystals were obtained in Wizard Screen 4 #13 (0.8 M succinic acid pH 7.0). This condition was further optimized by systematically performing a grid screen by varying the concentrations of precipitant, salt, buffer pH and protein concentration, which were then further subjected to Additive Screen from Hampton Research, USA. The final crystals obtained at 7 mg/ml protein concentration showed reflections in the diffraction image using the in-house machine (Rigaku). Cryoprotectant solution was also optimised by testing different percentages of glycerol. The crystals were quickly dipped in cryoprotectant solution and were flash-frozen in liquid nitrogen at 100 K.

2.2.14.1 Data collection and structure determination of *MtRel*₁₋₃₉₄

Together with Dr. Asha M. Balakrishna, Single-wavelength datasets for *MtRel*₁₋₃₉₄ crystals were collected remotely at 140 K on beamline 13B1 at the National Synchrotron Radiation Research Centre (NSRRC, Hsinchu, Taiwan) using the ADSC Quantum 315 CCD detector. Staff at beamline 13B1 (NSRRC) kindly provided technical assistance. All the diffraction data were indexed, integrated and scaled using

the HKL2000 suite program [212]. All the crystals belong to hexagonal space group $P3_2$ and had similar unit cell parameters.

2.2.14.2 Crystal structure determination of *MtRel*₁₋₃₉₄

Initial phases for *MtRel*₁₋₃₉₄ structure determination from *Mtb* were obtained by molecular replacement method by using the *SeRel*₁₋₃₉₃ (PDB ID: 1vj7 [232]) as model with the help of PHASER [217]. Due to severe anisotropy in the diffraction data set, the data were subjected to Diffraction Anisotropy Server. Data were anisotropically truncated and scaled with the diffraction anisotropy server (DAS) (Strong, M. *et. al*, 2006). Twinning operator was added to improve the R factor and R free difference. Subsequent rigid body and restrain refinement was done with REFMAC5; the starting model was improved manually using Coot ([211]; Fig. 2.2 and 2.3). During refinement, jelly body restraints and Non crystallographic symmetry, NCS were used. In the final cycle of restrained refinement TLS (Translation, liberation, Screw) restraints were included. Refinement was done until convergence and the geometry of the final model was validated with MolProbity [227]. All the figures were drawn using PYMOL [224]. Structural comparison analysis was carried out by the SUPERPOSE program [233] as included in the CCP4 suite. Interface areas and solvation energies were calculated by the Protein interfaces, surfaces and assemblies service PISA [233, 234]. The schematic described in Figure 2.6 summarize the strategy used to solve crystal structure of *MtRel*.

Figure 2.6: Scheme used to solve the crystal structure of *MtRel*₁₋₃₉₄. The data collected in the form of X-ray diffraction was processed by using HKL 2000 suite and solved by using CCP4 suite (Phaser, Refmac5, Coot). The model quality was checked by Procheck and visualized using Pymol.

2.2.15 Small angle X-ray scattering for low resolution studies

Small angle X-ray scattering (SAXS) is a low resolution experimental technique which aids in determining the global structural parameters and shape information. The range of resolution varies from nanometres to a micrometre [235]. The isotropic scattering by the random position of the particles in solution is calculated as scattering intensity represented as $I(s)$ which is a function of momentum transfer $q=4\pi\sin\theta/\lambda$ where λ represents the wavelength of the X-rays [236].

2.2.15.1 SAXS data collection

MtRel SAXS data was collected at the SAXS beamline P12 at the EMBL, Hamburg, Germany but also in-house [237]. The monochromatic X-rays were sourced from the controllable-gap low divergence U29 undulator of PETRA III [237]. The data was collected by Dr. Mertens H.D.T. at synchrotron facility. The X-ray beam was focused by adaptive bimorph mirrors positioned in a Kirkpatrick–Baez geometry and combined with a three-pair slit collimation. This set up produces a close to one-to-one optics and prevents the high divergence of the source by delivering an intense highly focused small beam to the sample position. The energy used was 10 KeV ($\lambda=0.124$ nm) for the data collection. Quartz capillary of 1.7 mm was used with a sample volume of 20 μ l and normalization was done to transmitted intensity by beam stop counter. A photon counting Pilatus 2M pixel detector was used at a sample –detector distance of 3.1 m. This covered the range of momentum transfer $0.02 < q < 5.0$ nm⁻¹. The standard BSA was used at 3.4 mg/ml in 50 mM HEPES buffer at pH 7.5. The scattering from the buffer alone was measured before and after each sample measurement and the average of the scattering before and after each sample is used for background subtraction. In total, 20 frames were collected for each sample at an interval of 0.045 seconds. SAXS experiments were carried out at 15 °C for a concentration series ranging from 0.6 to 0.9 mg/ml of *MtRel*.

SAXS-data of *MtRel*₁₋₃₉₄ and *MtRel*₆₆₀₋₇₃₈ were measured with an in-house BRUKER NANOSTAR SAXS instrument equipped with a Metal-Jet X-ray source (Excillum, Germany) and VÅNTEC 2000-detector system. For the in-house SAXS machine, X-ray radiation was generated from a liquid gallium alloy with microfocus electron source ($K\alpha = 0.13414$ nm with a potential of 70 kV and a current of 2.857 mA). The X-ray beam was focused and monochromated using Montel mirrors and collimated

by a two pinhole system. The two pinhole collimation systems helps in removing the parasitic scattering. Liquid samples are placed in a sealed quartz capillary, while the solid glassy carbon that is used as a standard is fixed into the path of the beam after the sample. The sample chamber and beam path is evacuated at around 1 Pa and all the measurements were performed at 288 K. The sample to detector distance was set at 0.67 m and the sample chamber and X-ray paths were evacuated. The set covers a range of momentum transfer of $0.16 < q < 4 \text{ nm}^{-1}$ [238-240]. SAXS experiments were carried out at 15 °C for 3.1, 3.8 and 5.0 mg/ml of *MtRel*₁₋₃₉₄ and 0.8, 1.5 and 3.3 mg/ml of *MtRel*₆₆₀₋₇₃₈ in a sample volume of 80 µl in a vacuum tight quartz capillary. Corresponding to each protein sample, data were collected by Dr. Manimekalai M.S.S. for 30 minutes and for each measurement a total of six frames at 5 min intervals were recorded. The data were tested for possible radiation damage by comparing the six data frames, which indicated no change. The scattering of the corresponding buffer was subtracted and the difference curves normalized for the concentration as well as initial intensity.

SAXS data of *MtRel*₁₋₃₉₄ were also collected in the absence and/or presence of substrates GTP (Guanosine triphosphate), ApCpp (Adenosine-5'-[(α,β)-methyleno] triphosphate), GTP + ApCpp and respectively, at a protein concentration of 3.4 mg/ml.

SAXS measurement for a standard compact globular protein, lysozyme, at a concentration of 5 mg/ml in buffer sodium acetate titrated with acetic acid to pH 4.5 was also done employing the similar strategy as above.

2.2.15.2 SAXS data analysis

All the data processing steps were performed using PRIMUS package from ATSAS v2.8.2 [241]. The data processing was done under the kind guidance of Dr. Malathy Sony S. Manimekalai. The experimental data obtained for all protein samples were analyzed for aggregation using the Guinier region [242]. The forward scattering $I(0)$ and the radius of gyration, R_g , were computed using the Guinier approximation assuming that at very small angles ($q < 1.3/R_g$) the intensity is represented as $I(q) = I(0) \cdot \exp(-(qR_g)^2/3)$ [243]. These parameters were also computed from the extended scattering patterns using the indirect transform package GNOM [244], which provides the distance distribution function $P(r)$ hence the maximum particle dimension, D_{max} , as well as the radius of gyration, R_g . Qualitative particle motion was inferred by plotting

the scattering patterns in the normalized Kratky plot [245]. The hydrated volume V_p , used to estimate the molecular mass of the protein, was computed using the Porod invariant. The molecular mass (MM) of the protein can be estimated in several ways from the SAXS data and cross-validation between them is very important in order to determine the oligomeric state of the protein. The MM was determined from the forward scattering of the sample $I(0)$, by comparison with a reference protein, lysozyme [246]. This method is limited by the accurate measurement of protein concentration and partial volume and the assumption of a compact near-spherical shape, which leads to error in the MM determination by about 10-15% [246]. Alternatively, the MM was estimated from the scattering data based on the excluded (i.e. hydrated) particle volume, V_p , which are although independent of concentration, break down when applied to flexible molecules and therefore gives an inaccuracy of about 20% [236, 247]. The MM determination using the volume of correlation, V_c (defined as the ratio of $I(0)$, to its total scattered intensity) is contrast and concentration independent and shown to have 5% error [248].

Ab initio low-resolution models of the proteins were built by the program DAMMIF [249] considering low angle data ($q < 0.2 \text{ nm}^{-1}$). The algorithm constructs bead models yielding a scattering profile with the lowest possible discrepancy (χ^2) to the experimental data while keeping beads interconnected and the model compact. Twenty independent *ab initio* reconstructions were performed for each protein and then averaged using DAMAVER [250]. The average excluded volume V_{Ex} computed using DAMAVER, was further used to estimate the molecular mass of the protein [247]. Superimposition between *ab initio* reconstruction and atomic model was performed using the software SUPCOMB [222].

The theoretical scattering curves from atomic structures were generated and evaluated against experimental scattering curves using CRY SOL [251]. Utilizing the intensities from each component, the volume fraction of each component can be determined by OLIGOMER [252].

2.2.16 Introduction to nuclear magnetic resonance (NMR) spectroscopy and its use in obtaining protein structure

Solution NMR spectroscopy is one of the principal techniques used to study the structural and functional aspects of macromolecules such as proteins depending upon their size. The technique enables to determine the solution structure of proteins and to

study protein-ligand and protein-protein interactions. Any nuclei having a mass number, which exhibits net intrinsic magnetic momentum and angular momentum, can be defined as NMR active. As hydrogen is one of the most receptive and abundant NMR active nuclei besides other isotopes such as ^{15}N , ^{13}C , ^{31}P *etc.*, it can be observed in magnetic resonance spectrum.

Due to the limited resolving power of the technique NMR for proteins is carried out in multi-dimension in order to resolve the overlapped peaks in 1 dimension. The multidimensional experiments provides correlation plots between the nuclei. These correlations are based on whether the two nuclei are sharing the bond or share the space. 2D NMR can be homonuclear (^1H , ^1H correlation) or heteronuclear (^1H , ^{13}C or ^1H ^{15}N correlations). The 2D NMR correlations are displayed as contour plots showing a specific nucleus at each axis. 3D NMR further results in several combinations of correlations (^1H ^1H ^{13}C , ^1H ^1H ^{15}N , ^1H ^{13}C ^{15}N). The 3D spectra is displayed as the slices of 2D NMR data.

A 3D structure of the protein can be solved by collecting “through space” correlations NMR data including Nuclear Overhauser effect (NOE). The NOEs are recorded in NOESY NMR experiments. The intensity of NOE gives an idea about the distance between the two nuclei. For instance, the nuclei which are 3 Å apart in space have strong NOE than those which are 6 Å apart having weak intensity NOE. Before acquiring information about the arrangement of nuclei in space in a protein fold, it is important to assign the polypeptide chains of the protein. To do so, sequential assignment is performed on the NMR spectroscopy data obtained via “through bond” correlation experiments. The results of this assignment are overlaid with the data obtained through experiments that determine the side chain for each residue. Together, this strategy completes the assignment of the amino acid residues in a protein. All the NOE data is fed into the programs like DYANA along with the protein sequence to obtain the protein fold with the maximum satisfaction of NOE distance constraints.

2.2.16.1 Preparation of NMR samples

The production of isotopic labelled proteins as explained earlier in this chapter was carried out to purify ^{15}N labelled, ^{13}C ^{15}N labelled and selectively labelled proteins. However, these protein samples can't be used directly for NMR spectroscopy experiments. Sample preparation differs from one NMR experiment to other. In order to obtain 2D experiments like Heteronuclear Single Quantum Coherence (HSQC),

namely ^1H - ^{15}N HSQC and ^1H ^{13}C HSQC, the ^{15}N and ^{13}C labelled proteins were used respectively. The sample to record these experiments for *MtRel*₆₆₀₋₇₃₈ and *MtRel*₃₉₄₋₄₅₉ were prepared to a total volume of 500 μl and/or 300 μl . All the samples were prepared with proteins in 90% H_2O based buffer and 10% D_2O to the final volume. A similar sample preparation strategy was used for selectively labelled *MtRel*₆₆₀₋₇₃₈ (Alanine/Arginine/Valine) and *MtRel*₆₆₀₋₇₃₈ mutants. The buffer used for *MtRel*₆₆₀₋₇₃₈ and *MtRel*₃₉₄₋₄₅₉ was composed of 50 mM Tris, 350 mM NaCl and 5% glycerol (pH 8.5) and purified with 1mM DTT.

For the 3D experiments, ^{13}C ^{15}N labelled protein sample were prepared in the same way as for 2D experiments except for 3D HCCH-TOCSY where the protein buffer was exchanged with 100 % D_2O buffer to have better signal.

2.2.16.2 NMR data collection and processing

All NMR experiments were performed by Dr Shin Joon from Prof. Grüber's lab at 20 °C on a 700 MHz NMR spectrometer (Bruker Avance) with 4 RF channels equipped with 5 mm TX1 triple-resonance ($^1\text{H}/^{15}\text{N}/^{13}\text{C}$) single axis gradient probes and a cryoprobe. All the protein samples were provided in respective buffer and 10 % D_2O . D_2O was added for the locking of the sample. Appropriate pulse calibrations and other parameters including temperature, buffer and protein concentration were optimized before making final measurements. To assign the backbone and obtain the structure of *MtRel*₆₆₀₋₇₃₈ and *MtRel*₁₋₄₅₉, the resonance assignments were achieved from conventional 2D and 3D heteronuclear NMR experimental data, collected using 0.1 mM to 0.5 mM of uniformly $^{13}\text{C}/^{15}\text{N}$ -labeled, ^{13}C , ^{15}N -double labeled and selectively labelled samples. The experiments conducted for *MtRel*₆₆₀₋₇₃₈ includes ^1H - ^{15}N HSQC, HNCACB, CBCA(CO)NH, HNCA and HN(CO)CA. These experiments were performed in order to obtain the backbone resonance assignment. In all practical cases, 3D triple resonance spectra were recorded in non-uniform sampling (NUS) [253] of the indirect dimension as 20-25% sampling rates and reconstructed using compressed sensing (CS) scheme using MDDNMR software [254, 255]. All data were processed with NMR-Pipe and NMRDraw [256, 257] and spectra were analyzed using SPARKY [258]. Baseline corrections were applied wherever necessary. The proton chemical shift was referenced to the methyl signal of DSS (2, 2-dimethyl-2-silapentane-5-sulphonate (Cambridge Isotope Laboratories)) as an external reference to 0 ppm [259, 260].

2.2.16.3 Resonance assignment

The resonance assignment of *MtRel*₆₆₀₋₇₃₈ was done under the guidance of Dr. Joon, whereas *MtRel*₁₋₄₅₉ was fully assigned by Dr. Joon. In the initial stage, protein molecules must be assigned in order to investigate NMR data for analysis of any kind of interaction or dynamic studies. Scalar couplings, via dipolar coupling and via chemical shift are useful information for sequence specific assignment of protein backbone. Depending on the type of sample and its isotopic label [261], different strategies are employed to assign the resonance. In NMR, a 1D spectrum is useful to investigate dispersion of resonance line of protein and also it gives an idea about the folding of the protein. But by doing an 1D experiment, it is impossible to resolve all resonance lines in a spectrum of a protein due to the presence of heavy crowding/overlapping of signals. To improve the spectral quality, extra dimensions are added in the NMR experiments. Assignments can be done using the sequential assignment approach based on Total Correlation Spectroscopy (TOCSY) [262] and 2D Nuclear Overhauser Effect Spectroscopy (NOESY). The TOCSY experiment is basically used to identify the resonance position within each amino acid spin system and the NOESY experiment is used for the sequential assignment of the spin systems present in the amino acids. For a protein/peptide NMR spectrum, amide hydrogens resonances are generally the best dispersed and observed between 6-12 ppm. The side chain resonance assignments are generally best done by examining the peaks in the backbone amide regions. This process is known as the sequential assignment strategy [263].

For larger molecules (>10 kDa) the sequential assignment strategy by using the above assignment approach is not useful, because overlaps of the NMR signals make the assignment complicated. For that it requires the addition of a third dimension by the replacement of naturally abundant isotope ¹⁴N with ¹⁵N and ¹²C by ¹³C. Hence, assignment is based on triple resonance experiments allowing further decrease in the overlap of NMR signals. There are six triple resonance experiments commonly used in chemical shift assignments of the backbone atoms of the protein: HNCA, HN(CO)CA, HNCACB, CBCA(CO)HN, HNCO and HN(CA)CO. These experiments vary in the way magnetization is passed from ¹H to ¹⁵N to ¹³C α and vice versa in order to come back to the atom desired to be detected in order to generate a 3D spectrum.

3 Results

3.1 Purification and biophysical characterization of *MtRel*

3.1.1 Introduction

The *MtRel* protein is a 738 amino acid long protein which belongs to category of long bifunctional enzymes of RSH superfamily. Its unique feature of regulating both the opposing activities namely hydrolysis and synthesis of (p)ppGpp makes it really fascinating to study in comparison to its homologs from *E. coli* which are either strictly (RelA) or preferentially (SpoT) monofunctional. Even after 19 years of the first study defining cloning and purification of the *MtRel*, no structural or mechanistic insights have been gained except few models proposed on the basis of mutagenesis studies. With the objective of finding better perspective in its species specific features and how they affect structure and function of *MtRel*, the RSH homologue from *Mtb* was cloned, expressed and purified.

From the literature, it has been established that these opposing activities are crucial for regulating stringent response and the crucial residues has been characterized by site directed mutagenesis studies involved in catalytic activity and regulation [183, 184, 190]. Prior to studies presented in this thesis on *MtRel*, there were no atomic resolution or low resolution structure of *MtRel* or any of its functional domains from any mycobacterial species. In this chapter, the strategies used to purify and optimize the stability of the *MtRel* and studies to understand the behaviour of *MtRel* in solution are presented. The assay to determine the catalytic activity of *MtRel* was performed to confirm the presence of active synthetase domain independent to ribosome-mRNA-tRNA complex. Appreciating the species specific differences will aid in understanding organism specific mechanism and carry out target based drug discovery.

3.1.2 Expression and purification of *MtRel*

To understand structural intricacies of *MtRel*, plasmid carrying gene coding for *MtRel* (1-738 aa, full length, FL) was provided by Dr. Wilson Nartey (described in Materials and Methods). The plasmid was transformed and overexpressed into the BL21 (DE3) by using 1 mM IPTG for induction and solubility test showed that the protein was soluble in 50 mM Tris, pH 7.5, 350 mM NaCl buffer and get induced at 18 °C overnight.

3.1.2.1 Optimization of buffer to improve solubility using cells from *E. coli* BL21(DE3)

1. 50 mM Tris, pH 7.5 and 350 mM NaCl

2. 50 mM Tris, pH 8.5 and 350 mM NaCl
3. 50 mM Tris, pH 8.5 and 750 mM NaCl
4. 50 mM Tris, pH 8.5, 750 mM NaCl and 10 % glycerol

The buffers mentioned above were optimized in a step wise manner. As the recombinant protein was produced with N-terminal 6-histidine-tag it had affinity for Ni^{2+} ion. The first step of purification, affinity chromatography was performed with Ni-NTA beads. The protein was eluted by using batch elution method using varying concentration of imidazole (25 - 600 mM) using the column containing Ni-NTA beads. Eluted fractions were analysed on a 17% SDS-PAGE gel and stained with Coomassie Brilliant Blue R250. Clear bands were identified on the gel for *MtRel* in the fractions eluted with 100-550 mM imidazole in buffer 1 and 2 having low salt. In order to prevent the formation of higher oligomeric species, buffer 3 was used with 750 mM salt. Buffer 4 further improved the stability of the protein upon addition of 10% glycerol (Fig. 3.1). The fractions eluted at 75-250 mM imidazole were collected and further used for second step of purification.

The Ni-NTA eluted fractions were pooled together, concentrated and applied to gel filtration column (Superdex 200 increase 10/300 GL). The recombinant protein *MtRel* collected at the elution volume of 12.4 ml. To ensure the purity of the purified protein, the eluted protein was applied on a 17% SDS-PAGE gel and stained with Coomassie Brilliant Blue R250. A single sharp band confirmed the presence of pure recombinant protein (shown as inset in Fig 3.1 B).

By strategizing step by step optimization of the buffer composition, the peaks were well resolved and concentration of protein eluted from S200 column improved significantly in buffer 4.

Figure 3.1: Purification of *MtRel* from *E. coli* BL21 (DE3). (A) 17% SDS-PAGE showing fractions eluted with 75-150 mM imidazole concentration from *E. coli* BL21 (DE3) (B) Superdex 200i purification chromatograms run at 0.5 ml/min flow rate, greyed area demarks the pooled fractions of the protein. The inset shows a 17% SDS-PAGE of the final protein product after size exclusion chromatography. The protein marker labelled as Lane M in (A) and Lane 1 in inset of (B) was Thermofisher Scientific PageRuler™ plus prestained protein ladder of 10 to 250 kDa.

3.1.2.2 Optimization of the expression system to improve solubility

E. coli GroEL-ES and C41 (DE3) expression strains were used to improve the yield of the protein and providing conditions for proper folding. It was hypothesized that co-expressing chaperons along with recombinant *MtRel* should improve the solubility of the protein suffered due to the presence of intrinsically disordered region of about 200 amino acids. Solubility and induction test (using 1 mM of IPTG) were performed for *MtRel* with appreciable fidelity. The similar patterns were observed in expression from *E. coli* C41 (DE3) and GroEL-ES strains (Fig. 3.2). The protein purified from *E. coli* GroEL-ES didn't show any improvement, whereas the *E. coli* C41 (DE3) improved the yield of *MtRel* protein (Fig. 3.2D).

Figure 3.2: Purification of *MtRel* from *E. coli* GroEL-ES and C41 (DE3) strains. (A) and (C) 17% SDS-PAGE showing fractions eluted with 75-150 mM imidazole concentration from GroEL-ES of *MtRel* and 100-250 mM imidazole concentration from *E. coli* C41 (DE3), respectively. (B) and (D) Superdex 200 increase purification chromatograms of *MtRel* expressed in *E. coli* GroEL-ES and C41 cell strains shown respectively. Greyed area demarks the pooled fractions of the protein. The inset shows a 17% SDS-PAGE of the final protein product after size exclusion chromatography. The protein marker labelled as Lane M in (A) & (C) and Lane 1 in inset of (B) & (D) was Thermofisher Scientific PageRuler™ plus prestained protein ladder of 10 to 250 kDa.

However, the solubility improved when the protein was expressed in C41 (DE3) strain whilst the issue with concentrating the protein beyond a concentration of ~2 mg/ml persisted. NanoDrop machine (BioSpec-nano, Shimadzu Scientific Instruments) was used for the estimation of protein concentration. For further experiments, the protein was expressed from *E. coli* C41 (DE3) cells and used. The identity of the protein was confirmed by sequencing peptides isolated from a trypsin digested products of purified *MtRel*. The fragments isolated (YGVYQ SLHTT VVGPE GKPLE VQIR) and (GDVIT LPTGS TPVDF AYAVH TEVGHR) corresponds exactly to residues 306 to 329 and 407 to 432, respectively, of the amino acid sequence of *MtRel* (Appendix I, Supplementary Table S1).

3.1.3 Existence of *MtRel* in multi-oligomeric forms in solution

3.1.3.1 Solution structure of the *MtRel* using SAXS

To elucidate the overall behaviour of recombinant protein *MtRel* in solution, SAXS data were collected at the SAXS beamline P12, EMBL, Hamburg, Germany. SAXS patterns for *MtRel* at concentrations of 0.6, 0.8, 0.9 mg/ml were recorded and analysed (Fig. 3.3-A). The Guinier plots analysed at low angles for all the three concentrations appeared linear, indicating no protein aggregation. (Fig. 3.3-A, inset). Further data analysis of *MtRel* was performed with 0.8 mg/ml concentration data. The R_g was calculated to be 5.14 ± 0.08 nm from a slope of the Guinier plot. The derived R_g from the Guinier plot is in agreement with the computed R_g (5.38 ± 0.05 nm) from the distance distribution function (Table 3.1). The maximum particle dimension, D_{max} , of *MtRel* was estimated to be 18 ± 10 nm and showed a tail which indicates that the molecule is in elongated confirmation (Fig. 3.3B) (Table 3.1).

Figure 3.3: SAXS data analysis of *MtRel*. (A) SAXS pattern ($\circ$) of *MtRel* at a concentration of 0.6 mg/ml (black), 0.8 mg/ml (green) and 0.9 mg/ml (pink), respectively. The inset shows the Guinier plots of *MtRel* at three different concentrations, which display linearity at all concentrations used, indicating no aggregation. (B) Pair-distance distribution $P(r)$ function of *MtRel* at 0.8 mg/ml shows a single peak slightly shifted to the right with a small tail.

Normalized Kratky plot was generated to understand the flexibility characteristic of the protein in solution (Fig. 3.4). The Normalized Kratky plot of the examined particle is compared with the distinct bell curve of the globular lysozyme. As shown in the figure 3.4, the broad bell-shaped curve gave information about the

conformational behaviour of *MtRel* which is shifted to the right with respect to the standard bell shaped curve of lysozyme. It indicates the incidence of high flexibility of *MtRel*.

Figure 3.4: Normalized Kratky plot of *MtRel*. The bell-shaped curve of *MtRel* (●; Green) compared to the compact globular lysozyme (●; grey) with a peak (---; grey). The *MtRel* curve is extended towards right from the theoretical peak of a globular particle (assuming an ideal Guinier region of a globular particle) and indicates that *MtRel* is really flexible in solution.

The molecular weight of *MtRel* was estimated by using a protein calibration technique against the standard protein, BSA, which revealed the mixed population of oligomers of *MtRel*. The solution shape of the putative dimer generated by SASREF was superimposed to the dimer of the *MtRel* homology model with an NSD of 1.63. DAMMIN was used to *ab-initio* reconstruct a low-resolution shape of the *MtRel* as shown in Fig. 3.5A (a putative dimer model). These studies however does not guarantee that the dimer model generated by SAREF is the only possible dimer arrangement. It is possible to have a dimer which is formed by parallel interaction of the domains. To evaluate the agreement of the experimental data of the solution shape, a monomer model was generated based on the *E. coli* RelA in an open confirmation (PDB ID: 5IQR, [165]) and the dimer of *MtRel* was generated by using SASREF. Both the models were fitted with the experimental data by using CRY SOL and the χ^2 values were recorded. The monomer had a poor fit to the experimental data with a χ^2 -value of 3.66, on the other hand, the SASREF head to tail dimer had a good fit to the experimental data with a χ^2 -value of 1.21. OLIGOMER yielded a fit of χ^2 -value = 1.21 for a mixture of monomer, dimer and trimer models in the ratio of 21:73:6 (Fig 3.5 B). The NSD value between the shapes calculated from DAMMIN is 0.88 ± 0.04 .

Since both the SASREF dimer model and OLIGOMER fit have same χ^2 -value (1.2), data comparison was performed. The R_g -value of monomer, dimer and trimer of the *MtRel* homology model computed as 4.725, 5.092 and 6.991 nm, respectively, whereas the R_g calculated from guinier and gnom were 5.14 and 5.38 nm for the dimer formation in solution. However, the P values for the SASREF fit and OLIGOMER fit were 0.0 and 0.001991 respectively. The P-value in statistics signifies the probability of obtaining observed results and rejecting the null hypothesis. Here the null hypothesis is either getting a dimer fit as by SASREF or getting a fit for mixed population as per OLIGOMER. In general, if the P-value is smaller than 0.01, there is strong evidence against null hypothesis. Here, as both the values are below 0.01, the p-value of OLIGOMER fit can still be considered a better fit in comparison to SASREF fit. So, based on R_g -value and P-value of fitting, OLIGOMER fit might be considered reliable, even though the P-value is low.

Figure 3.5: *Ab-initio* envelope reconstruction of *MtRel*. (A) DAMMIN envelope reconstructed onto the cartoon of SASREF dimer model of *MtRel*. The dimer is shown in head to tail fashion. (B) Scattering curve of *MtRel* superimposed with the theoretical scattering pattern derived from monomer (SwissModel), dimer (SASREF) and OLIGOMER models.

Sample details	MtRel
Organism	<i>Mycobacterium tuberculosis</i>
GenBank	AL123456.3
Extinction coefficient (A_{280} , $M^{-1} \text{ cm}^{-1}$)	79300
MM from chemical composition (Da)	81.826
Protein concentration (mg ml^{-1})	0.8 mg/ml
Solvent	50 mM Tris/HCl, pH 8.5, 750 mM NaCl, 10% glycerol, 1 mM DTT
Data collection parameters	
Instrument (source & detector)	EMBL P12 (PETRA-III, DESY, Hamnurg) and Pilatus 2M pixel detector ($253 \times 288 \text{ mm}^2$)
Beam geometry	$0.12 \times 0.2 \text{ mm}^2$
Wavelength (nm)	0.124
Sample-to-detector distance (m)	3.1
q range (nm^{-1})	0.02-5.0
Normalization	To transmitted intensity by beam stop counter
Monitoring for radiation damage	Frame comparison
Exposure time (sec)	0.9 (20 frames $\times$ 0.045 s)
Sample configuration	Flow-through capillary (1.2 mm I.D.)
Sample temperature (K)	288
Software employed	
Primary data reduction	I(s) versus s using RADAVER (ATSAS 2.8.3; Petoukhov et al., 2012), solvent subtraction using PRIMUSqt (ATSAS 2.8.3; Petoukhov et al., 2012)
Data processing	PRIMUSqt from ATSAS 2.8.2
<i>Ab initio</i> analysis	DAMMIN from ATSAS 2.8.2
Validation and averaging	DAMAVR from ATSAS 2.8.2
Atomic structure modeling	CRY SOL and SASREF from ATSAS 2.8.2
3D graphics representations	PyMOL v1.5.0.4
Structural parameters	
Guinier analysis	
$I(0)$ (arbitrary unit)	0.041
R_g (nm)	5.14 ± 0.08
q_{min} (nm^{-1})	0.0257

qR_g max	1.3
$P(r)$ analysis	
$I(0)$ (arbitrary unit)	0.04
R_g (nm)	5.28 ± 0.05
D_{max} (nm)	18 ± 10
q range (nm^{-1})	0.0257-5.08
χ^2 (Total estimate from GNOM)	1.21
Porod volume (V_p) (nm^3)	188.727
MM from V_p ($V_p/1.6$) (kDa)	117.954
DAMMIN excluded volume (V_{ex}) (nm^3)	335.289
MM from excluded volume ($V_{ex}/2$) (kDa)	167.6445

Table 3.1: Data collection and scattering derived parameters for *MtRel*

3.1.4 Determination of enzymatically active *MtRel*

The catalytic activity of *MtRel* in solution independent to stalled translational machinery was confirmed with the help of an end point synthetase assay AMP-GloTM from Promega. The assay was used to detect the formation of AMP which formed as a by-product of pyrophosphate transferase reaction from ATP to GTP to form (p)ppGpp. Since, the assay can detect the presence of AMP produced by the reactions, it gives an idea about the (p)ppGpp synthetase activity of *MtRel* in various concentrations. The different concentrations of *MtRel* showed the varied production of AMP as a result of synthetase reaction at a saturating concentration of ATP, GTP and MgCl_2 . The standards calculated were performed by using varying concentration of AMP (0.625 – 10 μM) (Fig. 3.6A).

Figure 3.6: Luminescence based AMP detection assay to check the synthetase activity of *MtRel*. (A) Luminescence obtained for AMP standards concentration (0-10 μM) titrated with 100 μM ATP and 100 μM in the presence of 200 μM MgCl₂. (B) Luminescence observed for AMP produced by *MtRel* synthetase reaction carried out at three different concentration (0.5, 1 and 5 μM).

Due to the inability to concentrate the protein above 2 mg/ml and its property to exist in different homo-oligomers restricted us to carry out structural studies on *MtRel*. This motivated to pursue domain-wise structural and mechanistic characterization of *MtRel*.

3.2 Crystallographic and solution structure of *MtRel* NTD

3.2.1 Introduction

N-terminal domain (1-394 amino acids) which is reported as the catalytic domain of *MtRel* based on mutagenesis experiments, comprises of the hydrolase and synthetase domain [183, 184]. The hydrolase domain spans amino acids 1-181 whereas the synthetase domain is characterized to be ranging from amino acids 87-394. The activity of the hydrolase domain containing fragment 1-203 is proposed to be similar to that of *MtRel*, whereas 1-181 show a 130-fold decrease in hydrolysis activity [184]. These domains are required for the regulation of (p)ppGpp formation and breakdown during stringent response.

With the first solution shape of the *MtRel*, we were motivated to go forward and unravel the structural intricacies of *MtRel* at atomic resolution. But the absence of single population of *MtRel* oligomer in the solution failed us to achieve good results in negative stain electron microscopy successfully. As reported in literature earlier and experienced with our *MtRel* construct, concentrating *MtRel* without getting it precipitated at higher concentration is impossible. The reason lies in its highly flexible nature and the intrinsically disordered region present in its C-terminus. Nevertheless, to reach our goal, we decided to learn from structural domains of *MtRel* and find out the significant differences in comparison to RSHs from non-mycobacterial species.

In order to understand the catalytic domain of *MtRel* (*MtRel*₁₋₃₉₄) i.e. the C-terminal truncated protein was produced and purified. *MtRel*₁₋₃₉₄ is mentioned as *MtRel* NTD in the following chapters. In this chapter, the structural and biochemical insights of *MtRel* NTD are presented. The enzymatically active nature was confirmed by IP-RP HPLC and AMPglo luminescence assay. The enzymatically active protein was used for crystallization strategies so as to get a well ordered crystal which can diffract X-rays to high resolution. The structure determination was carried out using X-ray crystallography and the protein was further studied for its solution behaviour in the presence and absence of substrates using SAXS.

3.2.2 Production and purification of *MtRel* NTD

The purification for *MtRel* NTD was carried out in two steps using the purification scheme described in materials and methods. The N-terminal 6-histidine tagged recombinant protein bound to Ni-NTA beads upon incubation, was eluted by using imidazole concentration gradient of 0-450 mM. The fractions eluted by affinity chromatography were analyzed on a 17% SDS-PAGE gel after staining with

Commassie Brilliant Blue R250. *MtRel* NTD was identified as clear bands on the gel in the fractions eluted with 75-150 mM imidazole (Fig. 3.7A). The corresponding fractions were pooled together, concentrated and applied to a gel filtration column for step two purification (Superdex 200 increase 10/300 GL). The recombinant protein *MtRel* NTD eluted at 14.5 ml. To ensure the purity of the protein, the eluted protein was applied on a 17% SDS-PAGE gel and stained with Coomassie Brilliant Blue R250. A single sharp band confirmed the presence of pure recombinant protein (shown as inset in Fig 3.7B).

Figure 3.7: Purification of *MtRel* NTD. (A) 17% SDS-PAGE showing fractions eluted with varying imidazole concentration of *MtRel* NTD. (B) Superdex 200 purification chromatogram of *MtRel* NTD, greyed area demarks the pooled fractions of the protein. The inset shows a 17% SDS-PAGE of the final protein product after size exclusion chromatography. The protein marker labelled as Lane M in (A) and Lane 1 in inset of (B) was Thermofisher Scientific PageRuler™ plus prestained protein ladder of 10 to 250 kDa.

3.2.3 Confirming the presence of enzymatically active protein in solution

3.2.3.1 HPLC studies of catalytically active *MtRel* NTD

Functional characterization of the *MtRel* was explored about 15 years ago [184, 190]. The dissection of *MtRel* into fragments and their function was explored biochemically. As explained in introduction, this led to the characterization of catalytic domains, the synthetase- and hydrolase domain, present in the N-terminus of the bifunctional Rel protein from *Mycobacterium* species. Recent studies have also extended this characterization by confirming the fragments involved in hydrolysis/synthesis of (p)ppGpp [196].

Till this study, radio-labelling has been the main technique of choice to determine the activity of the RSH enzymes. In this study, a time-course ion pair reverse phase HPLC experiment has been adopted from biochemical studies on RelA of *F. tularensis* [231]. The method has been used to confirm if the purified *MtRel* NTD is

catalytically active. The ion-pairing agent DMHA was used to counteract the poor retention issues with the highly polar nucleotides like AMP produced as a by-product of the transferase reaction. In a pyrophosphate transferase reaction ATP and GTP/GDP are required as substrates. The gradual disappearance of substrates in a time dependent manner leads to appearance of (p)ppGpp and AMP as products. Different peaks corresponding to different nucleotides were observed in the chromatographs shown in figure 3.8A at 0, 10, 20, 40 and 60-min. Similar pattern as observed by Wilkinson *et al.* was observed for *MtRel* NTD IP-RP HPLC chromatogram. Around a retention time of 7 min, most of the buffer salts eluted. At time zero, two notable peaks eluted at about 18 min and 21 min. These peaks when compared to *F. tularensis* experimental data corresponded to GTP and ATP, respectively. As the reaction proceeded, two peaks were observed, one eluting after 5 min and a double peak, which eluted after 25 min (*shown as an insert*). The double peak is the product of the synthesis, pppGpp which can be degraded to ppGpp. The nucleotide with high charge takes longer retention time as (p)ppGpp and the less charged nucleotide such as AMP elutes early in around 5 min. The data obtained at various time points were compared by superimposing the chromatographs as presented in figure 3.8B. The data reveal that the peak intensity decrease of both GTP and ATP correlates well with the appearance of (p)ppGpp over the time course of 60 min. A zoomed in figure (Fig. 3.8C) shows the decrease of the substrates intensity as the reaction proceed. In addition, a corresponding gradual appearance of the AMP-peak could be observed, confirming the by-product formation of AMP in (p)ppGpp synthesis. These observations demonstrate that the recombinant *MtRel* NTD is in its catalytically active conformation and can synthesize (p)ppGpp. The purpose of this experiment was not to do enzyme kinetics study rather to infer whether the protein is active or not.

Figure 3.8: IP-RP HPLC used to confirm the presence of enzymatically active *MtRel* NTD. (A) Time-courses of ppGpp formation in *MtRel* NTD activity assays determined by HPLC analysis. (B) *MtRel* NTD activity assays with GTP as the pyrophosphate acceptor as a time course measurement. (C) A zoomed in view of the chromatogram for the substrates of figure 3.3B.

3.2.3.2 Luminescence based detection of synthetic activity of *MtRel* NTD

As described earlier in chapter 3.1, the *MtRel* being enzymatically activity was confirmed with AMP Glo assay from Promega. The synthetic activity of *MtRel* NTD was also confirmed in the similar manner and was used to compare the difference in activity of WT to NTD. The AMP produced as by-product of the reaction was computed as per the luminescence signal of *MtRel* NTD [500 nM]. The graph below shows the comparison between the *MtRel* and *MtRel* NTD at 500 nM concentration (Fig. 3.9). The data clearly shows that at the equimolar concentration, *MtRel* NTD shows more synthetic activity than *MtRel* which is in line with the enzyme kinetics conducted by

Avarbock *et al.* in *Mtb* and by Bag *et al.* in *M. smegmatis* showing that the C-terminal plays its role in regulating the NTDs catalytic activities. For the assay, the AMP standards were obtained as mentioned in chapter 3.1. The reaction buffer had ATP and GTP in equimolar concentration and MgCl_2 was provided in equimolar concentration of [ATP + GTP] in the presence of 1 mM DTT, 0.1 mg/ml BSA in the buffer having 50 mM Tris at pH 8.5, 350 mM NaCl, and 5% glycerol.

Figure 3.9: AMP generation as a result of (p)ppGpp synthetase activity by *MtRel* and *MtRel* NTD. The bars corresponding to *MtRel* (blue) and *MtRel* NTD (cyan) are representing the luminescence produced by the AMP generated in the transferase reaction with equimolar concentration of both the proteins in the presence of GTP and ATP. *MtRel* NTD shows almost 3 fold increase in production of AMP as the by-product of transferase reaction.

3.2.4 Crystallization and data collection of *MtRel* NTD crystals

To crystallize *MtRel* NTD, hanging drop method of vapour diffusion was used. Upon screening of sparse matrix screens from various vendors, initial crystals were observed in Hampton crystal screen 1 condition #7 (0.1 M Sodium cacodylate trihydrate pH 6.5, 1.4 M Sodium acetate trihydrate) (Fig. 3.10, 3.11A), Wizard Screen 2 condition #37 (1.0M K/Na Tartrate, Tris pH 7.0, 0.2M Li_2SO_4) (Fig. 3.11B) and Wizard Screen 4 condition #13 (0.8M succinic acid pH 7.0) (Fig. 3.11C). The respective crystals diffracted the X-rays to 8 Å (Fig. 3.10), 9 Å and 6 Å resolution, respectively. Initially, a series of optimizations were done to improve the crystals and diffraction resolution in Hampton crystal screen 1 condition #7.

Figure 3.10: Initial crystal of *MtRel* NTD and diffraction pattern. Initial *MtRel* NTD crystals, which were crystallized in 0.1 M Sodium Cacodylate Trihydrate pH 6.5, 1.4 M Sodium Acetate Trihydrate. The diffraction pattern of *MtRel* NTD crystals before optimization shown in right. The circle shows resolution shell for 8 Å.

Figure 3.11: Screening of *MtRel* NTD crystals. (A) Hampton crystal screen 1 condition #7 (0.1 M Sodium Cacodylate Trihydrate pH 6.5, 1.4 M Sodium Acetate Trihydrate) (B) Wizard Screen 2 condition #37 (1.0M K/Na Tartrate, Tris pH 7.0, 0.2M Li₂SO₄) (C) Wizard Screen 4 condition #13 (0.8M succinic acid pH 7.0) (10 units = 0.1 mm)

The optimization was done by varying the protein concentration from 5 mg/ml to 16 mg/ml, drop size from 1 μ l to 2 μ l and precipitant solution ratio as 1:1, 1:2, and 2:2. The crystals were tested in-house for any improvement in the diffraction resolution. After performing series of variations in the crystallization parameters, no further improvement was observed in crystals growth using Hampton Crystal Screen 1 (#7), therefore, Wizard Screen 4 #13 was selected for further optimization. To fasten the nucleation process 0.2-0.4 M TMAO (Trimethylamine N-oxide) was added along with 0.8 M succinic acid pH 7.0. This reduced the crystal growth time from 3 weeks to 1.5-2 weeks, shown in figure 3.12A. To further optimize the crystals, variety of salts and precipitants were screened, while maintaining the same crystallization conditions and TMAO. 0.2 M MgCl₂ was found to be promising among all (Fig. 3.12B). An additional screen using Hampton Research additive screen was used to improve the crystal quality (Fig. 3.12C). After each step of optimization, the crystals were harvested and tested for improvement in diffraction. The crystals were cryo-soaked in 25% (w/v) glycerol and flash-cooled in liquid nitrogen at 100 K. The crystals obtained in Wizard Screen 4 #13

with TMAO, MgCl_2 and additive 20% (w/v) Benzamidinium hydrochloride diffracted in-house to a resolution of 3.8 Å.

Figure 3.12: Optimization of *MtRel* NTD crystals obtained in Wizard Screen 4 #13. 0.8M succinic acid pH 7.0 with (A) 0.2 M TMAO (B) 0.2M TMAO and 0.2 M MgCl_2 (C) 0.2M TMAO and 0.2 M MgCl_2 and Hampton Research additive #69 20% (w/v) Benzamidinium hydrochloride. (10 units = 0.1 mm)

Single wavelength datasets for *MtRel* NTD were collected remotely at the National Synchrotron Radiation Research Center (NSRRC), Taiwan and indexed, integrated and scaled using HKL2000 suite of programs as explained in *Materials and Methods*. The crystals were diffracted at the resolution of 3.7 Å (Fig. 3.13).

Figure 3.13: Final optimized *MtRel* NTD crystal in Wizard Screen 4 #13. The figure shows the final optimized crystal of *MtRel* NTD diffracting (3.7 Å resolution).

The diffraction data were observed to be anisotropic. It is important to cut off the resolution of the data considering the anisotropic nature of the crystal. Diffraction anisotropy server (DAS) was used to get the truncated and scaled data set which was further used in the refinement steps. The $\text{CC}_{1/2}$ was calculated to be 98% for the highest resolution shell by using CCP4, thereby supporting the cut off computed by DAS. The Matthews coefficient confirmed the presence of four molecules in each asymmetric unit, with a corresponding solvent content of 61.79%. To bring down the R-factor which was stalled at 38% even after number of refinements, the data were expanded to lower symmetry space group $P3_2$. The difference between the observed and predicted structure factors was brought down by introducing a twinning operator in the refinement. NCS

and jelly body refinement further improved the statistics of the refined structure. TLS was used in the final run of refinement. The *MtRel* NTD crystals data processing and refinement statistics are briefed in Table 3.2.

Data Processing Statistics	
Temperature (K)	100
Wavelength (Å)	1.00
Crystal System	Trigonal
Space group	P3 ₂
Unit cell parameters (Å)	$a = b = 161.71, c = 75.56$
	$\alpha = \beta = 90^\circ, \gamma = 120^\circ$
Resolution range (Å)	30-3.7
Solvent content (%)	61.79
Total number of reflections	73,768
Number of unique reflection	23,372
Average I/ $\sigma(I)$	19.2(3.9)
Completeness (%)	98.7 (99.7)
R _{merge} ^a (%)	5.0 (24.3)
Redundancy	3.2 (3.2)
No. of molecules in asymmetric unit	4
CC _{1/2}	98.74 (96.2)
CC*	99.6 (99.0)
R factor ^b (%)	35
R free ^c (%)	36.7
Number of amino acid residues	1,277
Ramachandran statistics	
Most favored (%)	91.8
Additionally allowed (%)	8.0
Generously (%)	0.2
Disallowed (%)	0.0
Validation Statistics	
Clash score	4.47
Molprobity score	1.62

Values in parentheses refer to the corresponding values of the highest resolution shell (3.81–3.7 Å).

^aR_{merge} = $\sum_i |I_h - \langle I_h \rangle| / \sum_i I_h$, where I_h is the mean intensity for reflection h .

^bR-factor = $\sum ||F_o| - |F_c|| / \sum |F_o|$, where F_o and F_c are measured and calculated structure factors, respectively.

^cR-free = $\sum ||F_o| - |F_c|| / \sum |F_o|$, calculated from 5% of the reflections selected randomly and omitted during refinement.

Table 3.2: Data processing, refinement and validation statistics of *MtRel* NTD crystals at the synchrotron.

3.2.5 Crystal Structure of *MtRel* NTD

The crystal structure of *MtRel* NTD at 3.7 Å resolution contains four molecules (chains A-D) in each asymmetric unit (Fig. 3.14). Each *MtRel* NTD monomer consists of 343 residues that fit well in the electron density map. Residues from 41-53, 115-137 and 348-363 were not visible in the density map and therefore were not modelled. Specifically, the missing residues in each of the chains are; Chain A: 41-53, 115-136, 348-366; Chain B: 41-53, 115-137, 348-362; Chain C: 41-53, 113-137, 349-363; Chain D: 41-52, 115-137, 351-363.

Figure 3.14: Representation of one asymmetric unit of *MtRel* NTD crystals. The figure shows four molecules chain A-D in one crystallographic asymmetric unit. The molecules are coloured A; ruby, B; lime green, C; blue and D; sand, respectively [228].

The overall structure of *MtRel* NTD (Fig. 3.15) comprises the fully functional hydrolase- (1-200) and synthetase (87-394) domains. The hydrolase domain consists of 11 helices conserved structurally among the proteins of HD superfamily and involves α_1 to α_4 forming a helix bundle. This three-helix bundle makes the binding pocket for the hydrolysis activity. The residues present in the β -turn that connects helix α_4 with α_5 , H80 and D81 (HD motif), are conserved among the metal-dependent phosphohydrolases superfamily. A three-helix bundle comprising of α_8 , α_9/α_{10} and α_{11} (137-197) forms an overlapping link between the hydrolase and synthetase domain. α_9 and α_{10} run along a common axis and separated by a kink formed by P174. The hydrolase domain requires binding of a divalent ion to a conserved tetrad (H-Xn-HD-Xn-D) for the catalytic activity [194]. In *MtRel* NTD, Mg^{2+} from the crystallization condition has occupied this site in hydrolase domain (shown in red sphere).

Figure 3.15: A monomer of *MtRel* NTD. The figure shows a monomer of *MtRel* NTD in two orientations rotated 180° degrees to each other on the horizontal axis. Domains colour coded as follows; hydrolase domain - cyan green, synthetase domain - salmon, overlapping three helix bundle - yellow. All the helices are marked as helices $\alpha 1$ - $\alpha 18$ and β -strands marked as $\beta 1$ to $\beta 5$ [228].

The synthetase domain (87-394) consists of five beta strands ($\beta 1$ to $\beta 5$) forming a mixed β -sheet, which are sandwiched between five alpha helices ($\alpha 13$ to $\alpha 17$). This fold is conserved among nucleotidyltransferase superfamily proteins. The ligand binding residues G241, R242, D265, Y300, H313, E325, Q327 and H344 for the synthetase activity are interspersed in the β -sheet which forms a solvent accessible cavity along with $\alpha 14$ [184, 196].

3.2.6 Intermolecular analysis of *MtRel* NTD

PDBePISA predicted that the crystal structure of *MtRel* NTD forms a stable dimer [234] with the complex formation significance score (CSS) of 0.68 (Fig. 3.16). The dimer subunits form an extended head to head orientation having a total surface

area $2002 \text{ \AA}^2$ buried at the dimer interface. Helix α_{16} and β_3/β_4 loop are found to be involved at the dimer interface in the formation of hydrogen bonds and salt bridges. Residues R277 and D278 of chain A and D (from symmetric unit) respectively forms the salt bridge to stabilize the dimeric interaction and residue D299 of chain A subunit interacts with S288 of chain D and *vice versa* to form a hydrogen bond with a distance of $2.63 \text{ \AA}$. In addition to that many hydrophobic and van der Waals interactions are also present. Residues T229, A232, S233, K234, I235, K236, E275, I276, R277, D278, Y280, A281, V283, G284, V285, H287, S288, L289, N291, P292, M293, A294, G295, F297, I301 and A302 of the two subunits are present at the dimer interface.

Figure 3.16: Dimer formation by *MtRel* NTD. The two molecules A (ruby) and D (sand), forming stable dimers (PDBePISA server), are shown. Inset, residues R277, D278, S288 and D299, forming hydrogen bonds, are shown as sticks at the dimer interface [228].

3.2.7 Solution structure studies by SAXS

3.2.7.1 Dimeric solution structure of the *MtRel* NTD

In order to investigate the oligomeric state and flexibility of *MtRel* NTD in solution, a SAXS analysis was done using freshly purified *MtRel* NTD. SAXS experiments were performed at concentrations of 3.1, 3.8 and 5.0 mg/ml. The Guinier plots revealed that the data is linear at low angles and shows no protein aggregation (Fig. 3.17, inset). Three datasets were measured and merged to improve the data quality. The slope of the Guinier plot which is also termed as R_g value was calculated to be $36.29 \pm 0.57 \text{ \AA}$. The derived R_g from Guinier plot is in agreement with the computed R_g ($36.76 \pm 0.33 \text{ \AA}$) from the $P(r)$ function (Table 3.2). To get an idea about the overall size of the molecule the maximum particle dimension (D_{max}) was calculated from the

distance- distribution curve. The D_{max} 120.9 ± 12 Å of *MtRel* NTD with a tail indicates that the molecule is elongated in solution (Fig. 3.17B) (Table 3.3).

Figure 3.17: SAXS studies of *MtRel* NTD and Pair-distance distribution $P(r)$ function of *MtRel* NTD generated after merging the three datasets. (A) SAXS pattern ($\circ$) of *MtRel* NTD at a concentration of 3.1 mg/ml (black), 3.8 mg/ml (green) and 5.0 mg/ml (Blue), respectively. The inset shows the Guinier plots of *MtRel* NTD at three different concentrations, which display linearity at all concentrations used, indicating no aggregation. (B) The $P(r)$ function shows a single peak slightly shifted to the right with a small tail [228].

Sample details	<i>MtRel</i> NTD
Organism	<i>Mycobacterium tuberculosis</i>
GenBank	AL123456.3
Extinction coefficient (A_{280} , $M^{-1} \text{ cm}^{-1}$)	55,350
MM from chemical composition (kDa)	44.6
Protein concentration (mg ml^{-1})	3.1, 3.8 and 5.0 mg/ml
Solvent	50 mM Tris/HCl, pH 8.5, 350 mM NaCl, 5% glycerol, 1 mM DTT
Data collection parameters	
Instrument (source & detector)	Bruker NanoStar equipped with MetalJet eXcillum X-ray source and VÅNTEC-2000 detector
Beam geometry	100 μm slit
Wavelength (nm)	0.134
Sample-to-detector distance (m)	0.67
q range ($\text{\AA}^{-1}$)	0.016 – 0.4
Normalization	To transmitted intensity by glassy carbon
Monitoring for radiation damage	Data frame-by-frame comparison
Exposure time (min)	30 (6 frames x 5 min)
Sample configuration	Vacuum tight quartz capillary
Sample temperature (K)	288.15
Software employed	
Primary data reduction	BRUKER SAXS program suite
Data processing	PRIMUSqt from ATSAS 2.8.2
<i>Ab initio</i> analysis	DAMMIN from ATSAS 2.8.2
Validation and averaging	DAMAVER from ATSAS 2.8.2
Atomic structure modeling	CRY SOL from ATSAS 2.8.2
3D graphics representations	PyMOL v1.5.0.4
Structural parameters	
Guinier analysis	
$I(0)$ (arbitrary unit)	102.0 ± 1.15
R_g ($\text{\AA}$)	36.29 ± 0.57
q_{min} (nm^{-1})	0.0257
qR_g max	1.3
$P(r)$ analysis	

$I(0)$ (arbitrary unit)	102.1 ± 0.8
R_g (Å)	36.76 ± 0.3
D_{max} (Å)	120.9 ± 12
q range (nm^{-1})	0.0257-5.08
χ^2 (Total estimate from GNOM)	1.0
Porod volume (V_p) (Å ³)	146,730
MM from V_p ($V_p/1.6$) (kDa)	91.7 ± 9
DAMMIN excluded volume (V_{ex}) (Å ³)	182,470
MM from excluded volume ($V_{ex}/2$) (kDa)	91.2 ± 9

Table 3.3: Data collection and scattering derived parameters for *MtRel* NTD

A normalized Kratky plot was generated to understand the flexibility characteristic of the protein in solution (Fig. 3.18). It helps in understanding the conformational behaviour of the particle into consideration. The Normalized Kratky plot of the examined particle is compared with the distinct bell curve of the well folded globular protein such as lysozyme. As shown in the figure 3.18, the broad bell-shaped curve gave information about the conformational behaviour of *MtRel* NTD which is slightly shifted to right with respect to standard bell shaped curve of lysozyme. It indicates the incidence of slight motion and extended nature of *MtRel* NTD [245, 264].

Figure 3.18: Normalized Kratky plot of *MtRel* NTD. The bell-shaped curve of *MtRel* NTD (●; black) compared to the compact globular lysozyme (●; grey) with a peak (---; grey). The *MtRel* NTD curve is slightly shifted from the theoretical peak of a globular particle (assuming an ideal Guinier region of a globular particle) and indicates that *MtRel* NTD is extended and thereby less rigid in solution [228].

DAMMIN was used to *ab-initio* reconstruct a low-resolution shape of *MtRel* NTD as shown in figure 3.19. The molecular weight of *MtRel* NTD was estimated using

a protein calibration technique against the standard protein which revealed the dimeric nature of *MtRel* NTD. To evaluate the agreement of the experimental data of the solution shape to the experimental data obtained for the dimer of the crystal structure (PDB code: 5XNX) by curve fitting was performed by using CRY SOL. The theoretically extracted scattering pattern nicely fitted to the experimental data of *MtRel* NTD with a χ^2 -value of 1.0 (Fig. 3.19 B). The solution shape of the dimer was superimposed to the dimer of crystal structure with an NSD of 1.2 (Fig. 3.19 A). Moreover, the R_g -value and the diameter of the crystal dimer, 36.2 and 120 Å, respectively, was equivalent to the R_g and D_{max} calculated (36.29 ± 0.57 Å and 120.9 ± 12 Å) for the dimer formation in solution.

Figure 3.19: Solution shape of dimer *MtRel* NTD and SAXS data analysis of crystal dimer fitting to the solution shape of *MtRel* NTD. (A) SAXS studies performed on *MtRel* NTD to complete the *ab initio* DAMMIN envelope shown superimposed onto the cartoon representation of the crystal dimer (PDB: 5XNX) of *MtRel* NTD with subunits colored in green and cyan. (B) Fitting of SAXS scattering pattern (○; black) and calculated scattering profiles (—; red) from the dimeric *MtRel* NTD crystal structure (PDB: 5XNX) [228].

3.2.7.2 SAXS studies of *MtRel* NTD with substrates

To investigate the effects of substrate-binding to the flexibility of the protein and the dimeric state of *MtRel* NTD, freshly purified *MtRel* NTD was used in the presence of GTP, ApCpp (non-hydrolysable analog of ATP), GTP+ApCpp and ppGpp, for the SAXS experiment, respectively. The data for each sample measurement was collected at 3.5 mg/ml concentration of *MtRel* NTD. The protein and substrates GTP, ApCpp, GTP+ApCpp and ppGpp were used in the molar ratio of 1:10. The divalent ions Mg^{2+}/Mn^{2+} were used in the equimolar ratio as that of substrates. The merged scattering

profiles of *MtRel* NTD in the presence and absence of different substrates are shown in figure 3.20A. The distance distribution curves of substrate free and substrate bound enzyme are shown in figure 3.20B. The data showed no sign of protein aggregation. As shown in Table 3.4 and figure 3.20, there was no significant change in the R_g - or D_{max} values observed between the substrate-free *MtRel* NTD ($R_g = 36.11 \pm 0.63$ Å) compared to the *MtRel* NTD in the presence of GTP ($R_g = 36.11 \pm 0.76$ Å), ApCpp ($R_g = 36.77 \pm 0.57$ Å), GTP+ApCpp ($R_g = 35.99 \pm 0.71$ Å) and ppGpp ($R_g = 36.52 \pm 0.65$ Å) and a D_{max} value ranging between 120.8 – 121.7 Å. The data reflect that substrate-binding to *MtRel* NTD leads to no significant deviation in the overall dimensions and dimeric state of the enzyme.

Figure 3.20: SAXS studies of *MtRel* NTD in the absence and presence of substrates. (A) Comparison of SAXS scattering patterns ($\circ$) between substrate-free enzyme (black) and in the presence of GTP (red), ApCpp (blue), GTP and ApCpp both (magenta), and ppGpp (olive). The inset shows Guinier plots with linearity in all the five experiments performed, indicating no aggregation. (B) Pair-distance distribution function $P(r)$ of substrate-free *MtRel* NTD and *MtRel* NTD in the presence of substrates. Comparison of $P(r)$ function *MtRel* NTD in substrate free (black) and in the presence of nucleotide substrates GTP (red), ApCpp (blue), GTP and ApCpp both (magenta), and ppGpp (olive), respectively [228].

<i>MtRel</i> NTD	$I(0)$ (Guinier) (arbitrary units)	R_g (Guinier) (Å)	$I(0)$ ($P(r)$) (arbitrary units)	R_g ($P(r)$) (Å)	D_{max} (Å)
Substrate free	86.94 ± 1.2	36.11 ± 0.6	87.26 ± 0.7	36.67 ± 0.3	121.0
+ GTP	89.94 ± 1.3	36.11 ± 0.8	90.80 ± 0.8	36.99 ± 0.3	120.9
+ ApCpp	90.39 ± 1.0	36.77 ± 0.6	90.34 ± 0.7	37.14 ± 0.3	121.7
+GTP- ApCpp	89.58 ± 1.4	35.99 ± 0.7	90.17 ± 0.9	36.74 ± 0.4	121.2
+ppGpp	84.01 ± 1.2	36.52 ± 0.7	86.53 ± 0.8	36.55 ± 0.4	120.8

Table 3.4: SAXS data of *MtRel* NTD in the presence and absence of nucleotide substrates.

3.3 Deciphering the structural features of the tRNA binding regulatory domain, the *MtRel* TGS

3.3.1 Introduction

The essentiality of understanding the regulatory mechanism of *MtRel* is as crucial as understanding the species-specific differences of *MtRel* NTD for the therapeutic discovery. The regulatory domain of *MtRel* spans from 395-738 amino acids and comprises of domains required for auto-regulating the enzymatic activities of *MtRel* NTD [208]. The C-terminal of *MtRel* was identified on the basis of sequence similarity to homologous proteins characterized from mycobacterium [121]. The structured regulatory domains present in the CTD are the TGS and ACT domains (see chapter 1). In addition to that, Atkinson et al. have predicted other significant regions namely the helical domain and the CC (ZFD) domain on the basis of sequence similarity search [121]. The CTD have been identified to have a role in interacting with components of the stalled ribosomal complexes and have been postulated to have feedback regulatory mechanism to keep the concentration of (p)ppGpp in check [165, 190]. In order to witness these unique features of long bifunctional RSH proteins in *MtRel*, the CTD (394-738 amino acids) and *MtRel* TGS domain (394-459 amino acids) were cloned, expressed and purified. After gaining insights into the catalytic domain as described in previous chapter, it became interesting to understand the effect of the tRNA binding TGS domain on the (p)ppGpp synthesis activity. *MtRel* TGS was further studied at atomic resolution with the aid of NMR spectroscopy. A construct comprising of NTD and TGS domain, was cloned, expressed and purified. The protein *MtRel* NTD-TGS (1-459 amino acids) was further studied for the effect on synthetic activity.

3.3.2 Cloning, generation and purification of *MtRel* CTD, *MtRel* TGS and *MtRel* NTD-TGS

MtRel CTD and *MtRel* TGS were generated as described in section 2.2. The amplified product for both the gene constructs were obtained from PCR by using *MtRel* plasmid as template. A clear band corresponding to about 1032 bp and 216 bp as PCR amplified gene products for *MtRel* CTD and *MtRel* TGS, respectively, were observed on agarose gel as shown in Fig. 3.21A, B. The PCR product was gel extracted and double digestion was performed by using *NcoI* and *SacI* restriction enzymes to produce sticky ends. Ligation was set up by using *NcoI/SacI* digested pET9d-1. DH5 α was used to successfully transform the ligated product by using electroporation. The plasmid from successful clones for which the sequence had been confirmed was transformed into *E. coli* C41 (DE3) for *MtRel* CTD and *E. coli* C41 (DE3) cells for overproduction of *MtRel*

TGS. Solubility test and induction test (using 1 mM of IPTG) were performed for recombinant proteins with appreciable fidelity.

Figure 3.21: PCR amplification of *MtRel* CTD and *MtRel* TGS. PCR amplification of gene coding for (A) *MtRel* CTD and (B) *MtRel* TGS from *MtRel*-pET9d-1 plasmid ([210]) using PCR. DNA sample of 5 μ l was loaded on a 1.2 % agarose gel. The DNA marker labelled as Lane 1 in (A) and (B) were GeneRuler (1kbp) and (100bp) ladder, respectively.

On the other hand, to get the coding sequence of *MtRel* NTD-TGS, site directed mutagenesis (SDM) was performed and a stop codon was introduced after the TGS domain in *MtRel*:pET9d-1 plasmid. Dpn1 digestion was performed on the PCR product (Fig. 3.22A) obtained after SDM and 1 μ l of digested plasmid was used to transform *E. coli* DH5 α cells. The colonies were picked to perform colony PCR by using *MtRel* primers and successful clones were inoculated in LB supplemented with kanamycin to perform plasmid extraction of the mutated *MtRel* from liquid culture. The colonies were selected after confirming the mutation in the extracted plasmids by DNA sequencing. The selected plasmid of interest from DH5 α cells was transformed into *E. coli* C41 (DE3) cells for overproduction. The production of *MtRel* NTD-TGS was confirmed by performing an induction test (Fig. 3.22B shows induction test result for colony 1). A prominent band was observed at ~55 kDa on the SDS-PAGE confirming the expression of desired construct *MtRel* NTD-TGS.

Figure 3.22: PCR amplification and induction test of *MtRel* NTD-TGS. (A) PCR amplification of gene carrying stop codon after the TGS domain in *MtRel*-pET9d-1 plasmid using PCR. DNA sample of 5 μ l was loaded on an 1.2 % agarose gel. The DNA marker labelled as Lane 1 in (A) was GeneRuler (1 kbp) ladder. (B) SDS gel showing induced (+ve) and uninduced (-ve) samples. Induction was carried out with 1 mM IPTG at 37 °C for 3 hours. A prominent band is seen around 55 kDa, which corresponds to recombinant *MtRel* NTD-TGS. The protein marker labelled as Lane M in (B) was Thermofisher Scientific PageRuler™ plus prestained protein ladder of 10 to 250 kDa.

The purification strategy was carried as described in chapter 2. The proteins were purified from the cell lysate in the following steps: sonication, centrifugation and incubation with NiNTA beads. The recombinant protein was separated from other impurities using an imidazole gradient of 10-450 mM by affinity chromatography in buffer described earlier. The specific band at 38 kDa, 10 kDa and 55 kDa were observed on an SDS gel for *MtRel* CTD, TGS and NTD-TGS, respectively (Fig. 3.23A, B and 3.24A) upon staining with Coomassie Brilliant Blue G250. The imidazole fractions containing pure protein, *MtRel* CTD (75-150 mM), *MtRel* TGS (100 to 250 mM) and *MtRel* NTD-TGS (75-250 mM) were pooled and injected into a size exclusion column selected as per the size of the protein (chapter 2). *MtRel* CTD was fractionated at an elution volume of 14.7 ml on Superdex™ 200 increase 10/300 GL column, whereas *MtRel* TGS and *MtRel* NTD-TGS were eluted at 14.5 ml and 13.5 ml on Superdex™ 75, respectively. The peaks containing the recombinant protein of interest were identified using SDS-PAGE and the fractions were collected from the top 15% of the peak (highlighted in grey in Fig.3.23 C, D and 3.24B).

Figure 3.23: Purification of *MtRel* CTD and *MtRel* TGS. (A) and (B) shows a 17% SDS-PAGE revealing fractions eluted with varying imidazole concentration while purifying *MtRel* CTD and *MtRel* TGS, respectively. (C) and (D) Superdex 200 purification chromatograms of *MtRel* CTD and Superdex™ S75 purification profile of *MtRel* ACT, respectively. Greyed area demarks the pooled fractions of the respective protein. The inset shows a 17% SDS-PAGE of the final protein product after size exclusion chromatography. The protein marker labelled as Lane M in (A) & (B) and Lane 1 in inset of (C) & (D) was Thermofisher Scientific PageRuler™ plus prestained protein ladder of 10 to 250 kDa.

A similar purification protocol was followed to purify ^{15}N labeled *MtRel* TGS, except the buffer used for size exclusion chromatography of *MtRel* NTD-TGS and isotopically labelled *MtRel* TGS was 50 mM Tris, pH 8.5, 350 mM NaCl, 5% glycerol and 1 mM DTT.

Figure 3.24: Purification of the *MtRel* NTD-TGS from *E. coli* C41 (DE3). (A) Ni^{2+} -NTA purification performed with imidazole gradient showing the desired band corresponding to the molecular weight of *MtRel* NTD-TGS. Fractions from 100-350 mM imidazole were pooled and used for further purification. (B) SuperdexTM S200i purification profile of *MtRel* NTD-TGS eluting at 14 ml. Inset in Figure (B) shows the SDS gel after size exclusion chromatography purification. In Figure (B), lane 1 and 2 represent a protein marker and protein eluted from size exclusion chromatography. The protein marker labelled as *Lane M* in (A) and Lane 1 in inset of (B) was Thermofisher Scientific PageRulerTM plus prestained protein ladder of 10 to 250 kDa.

Since the *MtRel* CTD had poor yield and had similar problems as with *MtRel* in concentrating the protein, further studies were carried out with *MtRel* TGS only and *MtRel* NTD-TGS.

3.3.3 NMR solution structure determination for *MtRel* TGS

To understand the structural features unique to *MtRel* TGS, single labelled and double labelled protein was produced and purified. The single labelled protein was used to perform ^1H - ^{15}N HSQC 2D experiment and observed that the protein was well folded with nicely dispersed resonance peaks in the spectra.

The following data were collected and assigned by Dr. Shin to solve the structure of *MtRel* TGS. The sequential backbone assignment was performed by adopting Non-uniform sampling strategy (NUS). The strategy is 40 years old but has recently been adopted in NMR spectroscopy due to its advantages in measuring time and spectrum resolution [265]. It specially has become unavoidable in triple resonance experiments and thereby enables detection of weak peaks in high dynamic range spectra with modern high-field instruments [265]. NUS can benefit in recording crowded high dynamic-range spectra such as NOESY if the crowding is reduced by shortening the mixing times of NOESY or by reducing overlaps by isotopic labelling [265].

Out of 64 non-proline residues of 66 amino acid long *MtRel* TGS construct, 53 residues were assigned in the $^1\text{H}^{15}\text{N}$ HSQC spectrum. The 11 residues did not appear in the $^1\text{H}^{15}\text{N}$ HSQC spectra. With the help of HNCA, HN(CO)CA, HCCH-TOCSY and $^1\text{H}^{13}\text{C}/^1\text{H}^{15}\text{N}$ simultaneous NOESY-HSQC experiments, most of the backbone C α and C β were assigned. With the assignment of 65/66 C α and 57/60 C β resonances, ~96.8% completeness was achieved. In total, 94.1% of resonances were assigned by NMR experimental data which includes backbone and side chains. The assignment of the resolved backbone residues is shown in $^1\text{H}^{15}\text{N}$ HSQC spectra (Fig. 3.25).

Figure 3.25: $^1\text{H}^{15}\text{N}$ HSQC spectrum of *MtRel* TGS. The spectrum was recorded in 50 mM Tris pH 8.5, 350 mM NaCl and 5% glycerol buffer on a Bruker Avance 700 MHz machine at 280 K. The dispersion of the peaks shows a well folded protein. The resonance were labelled by assigning 3D NMR experiments.

Upon complete assignment, the chemical shifts were used to plot the secondary structure relative to the average random coil chemical shifts of amino acids. The difference of observed chemical shifts and the average random coil chemical shifts gave information about the helices (positive) and beta sheets (negative) (Fig. 3.26).

Figure 3.26: NMR spectroscopic secondary structure characterization of *MtRel* TGS. The secondary structure elements of the NMR structure of *MtRel* TGS are represented above chemical shift index plot. The scores were derived from $\Delta C\alpha$ based on backbone assignment of *MtRel* TGS. The positive and negative chemical shifts represent presence of α -helical and β -strands of the secondary structure, respectively. Color scheme: α -helices – red, β -strands- yellow.

The structural calculations on the chemical shifts along with the secondary structure prediction resulted in the structure of *MtRel* TGS at atomic resolution as a monomer. For this purpose, a total of 996 NMR-derived distance restraints and 91 dihedral angle restraints were determined. The ensemble of 18 low-energy structures of *MtRel* TGS and their restrained energy minimized (REM) average structure are shown in figure 3.27 revealing 5 β -strands (β -1; E6-T11, β -2; G14-P20, β -3; I41-V45, β -4; R48-V50, β -5; V62-F65) and 2 α -helices (α -1; P25-V32, α -2; T34-H38) in a monomeric structure.

Figure 3.27: NMR solution structure of the *MtRel* TGS. (A) Superposition of the backbone traces from the final ensembles of 18 solution structures of *MtRel* TGS domain which were determined by NMR spectroscopy. (B) A ribbon representation of Restrained Energy Minimized (REM) structure of *MtRel* TGS domain is displayed in two orientations. The monomeric *MtRel* TGS is shown in magenta having $\beta\beta\alpha\beta\beta$ arrangement.

3.3.4 Molecular interaction between *MtRel* TGS with deacylated tRNA

The interaction of *MtRel* TGS with tRNA was investigated by NMR titration experiment. The ^{15}N labeled *MtRel* TGS was titrated with/without *E. coli* tRNA at a protein:tRNA molar ratio of 1:2. ^1H - ^{15}N -TROSY-HSQC experiment performed by Dr. Shin for *MtRel* TGS titrated with deacylated tRNA showed clearly separated cross peaks of most resonances (Fig. 3.28).

Figure 3.28: ^1H - ^{15}N -TROSY-HSQC experimental analysis of *MtRel* TGS titration with deacylated tRNA. Overlaid ^1H - ^{15}N -TROSY-HSQC spectra of *MtRel* TGS with (blue) and without (red) tRNA at a 1:2 molar ratio.

The newly appeared and separated resonances peaks correspond to tRNA bound conformation of *MtRel* TGS, indicating tight binding to tRNA. The spectrum shows that the protein undergoes dramatic structural rearrangement after binding to tRNA and slowly exchange between bound and unbound conformation of *MtRel* TGS.

To identify the binding epitope of *MtRel* TGS domain with tRNA, the relative decrease of peak intensities was calculated based on the peak height of each assigned resonances of unbound conformation (Fig. 3.29A). Significant decreases in peak intensities were observed for the backbone resonances of amino acids V9, I17, T18, T24-H33, G42-R48, V50-E52, L56 and N58-T66. Based on the intensity change of each amino acid residues, a possible binding epitope for tRNA to *MtRel* TGS was deduced. Regions including T24-H33 ($\alpha 1$ and H33 on the HLH loop) and $\beta 5$ (N58-T66) might interact directly to tRNA, which is similar with the previously proposed tRNA-*EcRelA* complex structure (PDB ID: 5IQR) [165]. Regions including V9, I17-V18 ($\beta 2$), G42-R48 (including $\beta 3$), V50-E52 and L56 might be indirectly affected by binding of tRNA, indicating tRNA binding to *MtRel* TGS causes significant conformational rearrangement of the entire domain structure.

Figure 3.29: Molecular interactions of *MtRel* TGS with tRNA at a molar ratio of 1:2. (A) A plot showing change in intensity of resonance peaks (I_1/I_0) at a molar ratio of 1:2. The average value of the intensity ratio was shown as dotted line. The residues showing significant decrease of intensities after addition of tRNA than average value is labelled as one letter code. The regions affected upon interaction are highlighted in *grey*. (B) Mapping of the residues and regions affected upon tRNA binding based on the NMR data; the residues/regions revealing more than average intensity decrease in intensity ratio are represented in *green* on the ribbon representative structure of *MtRel* TGS shown in *pale pink*.

3.3.5 Understanding the effect of the TGS domain on synthetase activity

Appreciating the importance of the TGS domain in regulating the synthetic activity by *MtRel* NTD when complexed with the stalled translational machinery as shown in RelA structural studies from *E. coli* [165] and biochemical studies from *M. tuberculosis* [191] and *M. smegmatis* [208], a construct containing *MtRel* NTD-TGS was generated and purified as explained in Chapter 2. This segment of protein was checked by the luminescence assay as used for *MtRel* and *MtRel* NTD to detect the production of AMP as a result of transferase reaction by the synthetase domain. The results showed that *MtRel* NTD and *MtRel* NTD-TGS had similar activity in comparison to *MtRel* which shows 4-fold lower production of AMP upon synthesis of (p)ppGpp (Fig. 3.30). Thus, the TGS domain has no obvious role in regulating the synthetase activity off the stalled translational machinery.

Figure 3.30: AMP generation as a result of (p)ppGpp synthetase activity by *MtRel*, *MtRel* NTD and *MtRel* NTD-TGS. The bars corresponding to *MtRel* (blue), *MtRel* NTD (cyan) and *MtRel* NTD-TGS (pink) are representing the luminescence signal produced by the AMP, generated in the transferase reaction with equimolar concentration of the proteins in the presence of GTP and ATP.

3.4 Unravelling the allosteric feedback mechanism exhibited by the *Mt*Rel ACT

3.4.1 Introduction

The ACT domain is a well conserved among the proteins involved in metabolic reactions and regulates their function by sensing the amino acid concentration. The name of the domain as stated in chapter 1 derived from the three proteins it was first reported in [266]. This domain is typically located at the C-terminal regulatory region of the proteins and exhibits a ferredoxin like fold. The domains of C-terminal of *MtRel* were always believed to be crucial for identifying the signal upon amino acid starvation and transduction of the signal to catalytic domain so as to bring desired activity [184, 190]. There is not enough information available on how the signal is sensed to regulate such crucial mechanisms independent to stalled translational machinery. The importance of ACT domain in hydrolysis activity has recently been studied in SpoT of *C. crescentus* and Rel protein of *R. capsulatus* as described in in chapter 1. The effect of branched chain amino acids (BCAAs) on hydrolysis has revealed another layer of mechanism among RSHs as studied by Fang and Bauer [205]. So far, these findings have not been explored in structurally in any of the RSH family protein. Moreover, no finding has reported the biochemical nature of *MtRel* in sensing BCAAs with the help of ACT domain.

In order to investigate the aforementioned mechanisms by *MtRel*, here the first solution structure of the regulatory domain for hydrolysis activity, *MtRel* ACT is presented. To shed some more light on the interacting amino acids and to map the residues serving as the binding pocket of these allosteric regulators, site directed mutagenesis and NMR titration studies have been carried out. The information about the structural rearrangement in the presence of amino acids, will serve as the first evidence of allosteric regulatory role of ACT in mycobacterial bifunctional Rel protein.

3.4.2 Cloning, production and purification of *MtRel* ACT

The amplified product *MtRel* ACT was obtained from PCR using the *MtRel* plasmid as a template. A clear band was observed on the agarose gel corresponding to about 255 bp (237 bp construct with codon for 6 His tags) as shown in Fig 3.31A. The PCR product was gel extracted and double digested with *NcoI* and *SacI* restriction enzymes to produce sticky ends. The digested product was ligated with *NcoI/SacI* pre-digested pET9d-1. DH5 α cells were electroporated to successfully transform the ligated product. Transformed colonies were selected upon control digestion of the extracted plasmid and correct insertion of insert confirmed. The sequence of plasmids from

confirmed clones was checked by Sanger sequencing service by BioBasic [267]. The plasmid from positive colonies was transformed into *E. coli* C41 (DE3) cells for overproduction. The production of *MtRel* ACT was confirmed by performing an induction test (Fig. 3.31B). A prominent band was observed slightly above the 10 kDa marker protein on an 12% SDS-PAGE gel.

Figure 3.31: PCR amplification and induction test of *MtRel* ACT. (A) PCR amplification of *MtRel* ACT gene from *MtRel*-pET9d-1 plasmid using PCR. DNA sample of 5 μ l was loaded on an 1.2 % agarose gel. The DNA marker labelled as *Lane 1* in (A) was GeneRuler (1 kbp) ladder. (B) SDS gel showing induced (+ve) and uninduced (-ve) samples. Induction was carried out with 1 mM IPTG at 37 $^{\circ}$ C for 3 hours. A prominent band is seen around 10 kDa, which corresponds to recombinant *MtRel* ACT. The protein marker labelled as *Lane M* in (B) was Thermofisher Scientific PageRulerTM plus prestained protein ladder of 10 to 250 kDa.

A solubility test was performed with the positive colony identified in the induction test. The recombinant protein was found to be soluble when cells were grown at low temperature (18 $^{\circ}$ C). Based on the solubility test, 50 mM Tris at pH 8.5, 350 mM NaCl, 5 % glycerol was selected as appropriate buffer for purification of the protein. Like other constructs of *MtRel*, *MtRel*₆₆₀₋₇₃₈ was cloned with a 6-His-tag at its N-terminus. Thereby, Ni²⁺-affinity chromatography was used as first step of purification by making use of histidine affinity to Ni²⁺ ion followed by size exclusion chromatography using buffer as described in materials and methods.

The recombinant protein was separated from other impurities using an imidazole gradient of 25-450 mM in buffer described in materials and methods. The specific band at 10 kDa was observed on an SDS gel in 50-450 mM imidazole as identified in the

induction test as well. (Fig. 3.32A) which was stained with Coomassie Brilliant Blue G250. Since the pure protein was found at 150 to 350 mM imidazole concentrations, these fractions were pooled together and concentrated. The concentrated fractions were passed through a size exclusion Superdex 75 (10/300) GL column (GE Healthcare) (Fig. 3.32B). The protein eluted at 13 ml on SuperdexTM 75, corresponding to a dimeric species of around 18 kDa. The peaks containing recombinant protein of interest were identified using SDS-PAGE and the fractions were collected from the top 15% of the peak (highlighted in grey in Fig. 3.32B).

Figure 3.32: Purification of the *MtRel ACT* from *E. coli* C41 (DE3). (A) Ni²⁺-NTA purification showing the desired band corresponding to the molecular weight of *MtRel ACT*. Fractions from 150-350 mM imidazole were pooled and used for further purification. (B) SuperdexTM S75 purification profile of *MtRel ACT*. Insert in Figure B shows the SDS gel after size exclusion chromatography purification. In Figure B, lane 1 and 2 represent a protein marker and protein eluted from size exclusion chromatography.

Similarly, the mutants *MtRel ACT*_H12A, *MtRel ACT*_L15A, *MtRel ACT*_A23S, and *MtRel ACT*_N28A were cloned and purified as described in Materials and Methods. To determine the NMR solution structure and to investigate the interactions, single labeled, double labeled and selectively labeled (for amino acids Alanine, Valine and Arginine) *MtRel ACT* were expressed and purified in the similar fashion.

3.4.3 NMR Solution structure of dimeric *MtRel ACT*

An atomic resolution NMR solution structure of *MtRel ACT* was determined. A high quality 2D ¹H-¹⁵N HSQC spectrum of the recombinant ACT domain (79 amino acids) was observed and assigned by using 3D experiments. This was followed by

titration experiments with RNA and branched chain amino acids to map the residues involved in the interactions.

A 2D $^1\text{H}^{15}\text{N}$ HSQC experiment was performed showing well folded protein and peaks were well dispersed. Out of 79 residues of *MtRel* ACT 78 residues were found to be non-proline from the sequence for which the spectrum is shown below in the Fig 3.33

Figure 3.33: $^1\text{H}^{15}\text{N}$ HSQC spectrum of *MtRel* ACT. The spectrum was recorded in 50 mM Tris pH 8.5, 350 mM NaCl and 5% glycerol buffer on a Bruker Avance 700 MHz machine at 280 K.

The NUS strategy was adapted to perform the sequential backbone assignment. Firstly, ^1H and ^{15}N chemical shifts of 66 residues which appeared in the 3D HNCA or HNCACB spectra were assigned leaving 12 residues which did not appear in any of these two experiments. After the assignment based on HNCA and HNCACB, to further complete the assignment to 91.2% of completeness for 79 Ca resonances, a large data set was collected and analysed. The 3D experiments acquired for this purpose were namely, NUS_CBCA(CO)NH, NUS_HN(CO)CA, HCCH-TOCSY, ^1H - $^{13}\text{C}/^{15}\text{N}$ -simultaneous NOESY-HSQC. Moreover, selective labelling of the amino acids alanine (A), valine (V) and arginine (R) helped in assigning respective residues with confidence.

Finally, approximately 92.5% of resonances were assigned including backbone and side chains based on the NMR experimental data (Fig. 3.34).

Figure 3.34: Assigned $^1\text{H}/^{15}\text{N}$ HSQC- spectrum of *MtRel* ACT. The peaks corresponding to all 79 amino acids were assigned and labelled. The spectrum was recorded in 50 mM Tris, pH 8.5, 100 mM NaCl, 5% glycerol. Buffer, on a Bruker Avance 700 MHz machine at 298 K.

To predict the secondary structure of *MtRel* ACT after the assignment of all the resonances, the $\text{C}\alpha$ -chemical shifts of the random coil values for each residue subtracted from observed $\text{C}\alpha$ chemical shifts ($\Delta\text{C}\alpha$) shown in Fig. 3.35. This $\Delta\text{C}\alpha$ is known as secondary chemical shift and hints a string correlation between the chemical shift and the local structure of the protein. The plot represents that the secondary chemical shift of $\text{C}\alpha$ atoms in α -helices, tends to be positive whereas that of $\text{C}\alpha$ atoms in β -strands will tends to have negative secondary chemical shifts.

Figure 3.35: NMR spectroscopic characterization of *MtRel* ACT. The secondary structure elements of the NMR structure of *MtRel* are represented above chemical shift index plot. The scores were derived from $\Delta C\alpha$ based on backbone assignment of *MtRel* ACT. The positive and negative chemical shifts represents presence of α -helical and β -strands of the secondary structure, respectively. Color scheme: α -helices – red, β -strands – yellow.

Using 3D experiments as described in material and methods, the solution structure of *MtRel* ACT was solved as a homodimer. For this purpose, a total of 2,378 NMR-derived distance restraints and 132 dihedral angle restraints were determined. The ensemble of 20 low-energy structures of *MtRel* ACT and their restrained energy minimized (REM) average structure are shown in Fig. 3.36.

Figure 3.36: NMR solution structure of *MtRel* ACT. (A) Superposition of the backbone traces from the final ensembles of 20 solution structures of *MtRel* ACT domain homodimer which were determined by NMR spectroscopy. (B) A ribbon representation of Restrained Energy Minimized (REM) structure of *MtRel* ACT domain dimer is displayed. In the homodimer, the monomeric units are arranged in head to tail arrangement, Each monomer unit (A and B) is shown in orange and green, respectively. The secondary structural element in each monomer unit are also indicated.

Analysis of restrained energy minimized structure using PROCHECK-NMR showed that 88.9% and 11.1% of the residues lie in the most favoured and allowed regions, respectively. A superposition of the final 20 NMR structures over the energy minimized average structure shows that it consists of 4 β -strands (β -1; A3-A8, β -2; L30-T36, β -3; A42-F48, β -4; Y70-V75) and 2 α -helices (α -1; H12-E25, α -2; P53-V66) in each monomer unit.

3.4.4 Dimeric interface of *MtRel* ACT

To understand the interactions prevalent in the dimeric interface, a total of 60 inter-monomer NOEs were observed between the two monomeric units such as H12/V27, H12/N28, H12/V29, L15/T19, L15/V29, L15/A32, S16/T19, V29/V34, S31/V34, A32/V34, and A32/T35. Some of the inter-monomer NOEs 2D slices from the 3D NOESY data are displayed in Fig. 3.37.

Figure 3.37: Molecular contact between each monomer unit of *MtRel* ACT. Some examples of representative strips of the ^{13}C -edited three-dimensional (3D) NOESY spectrum of *MtRel* ACT showing NOE cross-peaks between residues of the dimeric interface located on each subunit of *MtRel* ACT domain.

The solution structure of *MtRel* ACT confirmed that *MtRel* ACT forms a stable dimer through a number of hydrophobic interactions in such a way that residues in α -1A and α -1B interact with residues in β -2A and β -2B to form a stable dimer as shown in Fig. 3.38.

Figure 3.38: A ribbon representation of the molecular contact between each monomeric unit of *MtRel* ACT based on NOE data. Each monomeric unit are highlighted in orange and green, respectively. Crucial residues showing inter-domain NOEs are shown as sticks, and labeled as one-letter code.

3.4.5 Molecular interaction between *MtRel* ACT and single stranded RNA

Knowing the importance of interactions in *EcRelA* ACT domain with 23S rRNA A-site finger [165], the RNA interactions were investigated with *MtRel* ACT domain. The commercially synthesized RNA having the sequence of *Mtb* rRNA homologous to *E. coli* 23s rRNA interacting site was synthesized and NMR titration experiment was performed. The RNA was commercially synthesized in single stranded form by IDT. To understand the interaction, ^{15}N labelled protein was titrated with RNA at an 1:1, 1:2 and 1:5 molar ratio, respectively in 50 mM Tris, 100 mM NaCl and 5% glycerol at pH 8.5. The titration at 1:5 ratio is shown in Fig. 3.39. The 2D ^1H - ^{15}N -HSQC spectra of the protein in the presence and absence of ssRNA showed that most of the resonances in the HSQC spectra didn't become affected upon titration.

Figure 3.39: *MtRel* titration with ssRNA: Overlaid HSQC spectra of *MtRel* ACT domain in absence (red) and presence (blue) of ssRNA at molar ratio 1:5 is shown.

3.4.6 *MtRel* ACT selectively interacts with branched amino acid

To investigate the regulation mechanism reported earlier with proteins carrying an ACT domain and specific interaction of RSH-ACT domains to BCAAs, NMR titration studies were carried out. *MtRel* ACT was titrated with BCAAs namely V, L and I at an initial protein to BCAAs concentration in the molar ratio of 1:10. A ^1H - ^{15}N -HSQC experiment was recorded with and without BCAAs. As shown in Fig. 3.40, for L and I no significant change was observed for the chemical shift change or peak separation for ^{15}N and HN resonances. On the other hand, significant changes were observed in ^{15}N and HN resonances after addition of V to *MtRel* ACT. The results indicated selective interaction of *MtRel* ACT with V leading to conformational transition from free to complex form of *MtRel* ACT after addition of V.

Figure 3.40: Molecular interaction of *MtRel* ACT and branched chain amino acids based on NMR titration experiments. Overlaid ^1H - ^{15}N HSQC spectra of *MtRel* ACT domain with (blue)/without (red) BCAAs (Leu, Ile and Val) at a 1:10 molar ratio.

The transition of free *MtRel* ACT to ACT-Val complex was further investigated by performing series of titration with V. ^1H - ^{15}N HSQC spectra of *MtRel* ACT with different molar ratios of Val (1:0, 1:5, 1:10 and 1:20) showed a gradual decrease in resonance peak intensities of free *MtRel* ACT with a simultaneous increase in intensities of newly appeared separated peaks. This indicated fast structural transition of free to complex form of *MtRel* ACT after addition of V (Fig. 3.41).

Figure 3.41: Successive titration of *MtRel* ACT with valine for identification of interacting residues. Superposition of the ^1H - ^{15}N HSQC of *MtRel* ACT in the absence (red) and presence of increasing amount of V (cyan,1:5), (green,1:10), (magenta,1:20) and (blue,1:50). ^1H - ^{15}N HSQC-spectra were obtained on a Bruker Avance II 700 MHz machine at 298K.

Since no line broadening effects were observed after addition of V to *MtRel* ACT, the relative decrease of intensities were calculated based on the peak height of each assigned resonances using NMR spectra at 1:10 molar ratio (Fig. 3.42). The residues changes in peak intensities were observed for the backbone resonances of amino acids D10, L14, L15, T20, K26, K26, V27, L30, A32-V34, S44, M50, R64, V69 and Y70.

Figure 3.42: A plot of changes of peak intensity based on the ratio (I_1/I_0) at a molar ratio of 1:10. The average value of the intensity ratio and standard deviation were shown as dotted line. The residues showing higher intensity ratio than standard deviation are labelled as one letter code.

In order to get detailed information for binding epitope and molecular interaction of *MtRel* ACT with V, 3D triple resonance experiments and assignment of Ca resonances of the *MtRel* ACT -Val complex at molar ratio 1:50 were performed. Based on the assignment of ^{15}N , HN and ^{13}Ca resonances of the *MtRel* ACT -Val complex, significant chemical shift perturbations (CSPs) for most ^1H and ^{15}N resonances. Residues showing dramatic CSPs than average CSP value ($\Delta\delta_{\text{ppm}} > 0.38$) are E7, R11, L15, V18-T20, L22, D24, V27-L30, A32, V34, M43-M50 and V69-D71, respectively. Interestingly, residues including L15, T20, V27-L30, A32, V34, I43-M50 and V69-D71 show more significant CSP values, which is corresponding to analysis of decrease in intensity based on the NMR titration data at 1:10 molar ratio (Fig. 3.43 A-B).

Figure 3.43: 2D ^1H - ^{15}N HSQC experimental analysis of *MtRel* ACT titration with valine at an 1:50 molar ratio. (A) Overlaid HSQC spectra of *MtRel* ACT with (blue) and without (red) V at an molar ratio of 1:50. Residues showing CSPs over than average value (0.38 ppm) are labeled and indicated as one letter codes and arrows, respectively. (B) Weighted CSPs for the ^{15}N and ^1H resonance of *MtRel* ACT after binding of V at molar ratio 1:50; residues showing CSP above average CSP (~ 0.38 ppm) are labeled as one-letter code. The weighted CSPs between without and with Val of *MtRel* ACT for backbone ^{15}N and ^1HN were calculated by the formula $\Delta\delta = [(\Delta\text{N})^2 + (\Delta\text{HN}/5)^2]^{0.5}$. The average weighted CSP values were displayed by dashed line.

The mapping of CSP between free and complex form of *MtRel* ACT based on the NMR assignment data onto the surface of *MtRel* ACT also revealed similar binding epitope and structural alteration of *MtRel* ACT after addition of V. The results from NMR titration analysis and backbone assignment data of the *MtRel* ACT-Val complex strongly support selective interaction of V to *MtRel* ACT and overall structural rearrangements after selective binding of V (Fig. 3.44A).

The resonance peaks showing prominent decrease in intensity were mapped on the *MtRel* ACT 3D structure to identify the binding epitope of V to the domain. The mapping revealed the regions including residues D10-L15 and K26-L30 forming a binding cleft for V. Regions harbouring residues A32-V34, S44, M50, R64, V69 and Y70 might get affected indirectly upon binding of V. The intensity-based data helped in concluding that interaction of V to *MtRel* ACT leads to entire conformational transition of the protein structure including the dimeric interface (Fig. 3.44B).

Figure 3.44: Mapping of the valine binding surface and conformational changes using NMR data. (A) Residues revealing more than average CSP values (top) or intensity changes (bottom) are represented in red and blue on each monomeric unit. (B) Comparison between CSPs (left) and intensity changes (right) of *MtRel* ACT after addition of V. The residues showing significant intensity changes or CSPs are displayed and labeled red and blue color on each monomeric unit.

3.4.7 Low-resolution solution structure of the *MtRel* ACT

SAXS experiments were performed to further investigate the oligomeric state of *MtRel* ACT. The recombinant protein was used at 0.8, 1.5 and 3.3 mg/ml concentration, respectively, to determine the SAXS patterns. Data quality and absence of aggregation was confirmed by observing linearity at low angles in Guinier plot (Fig. 3.45A, inset). The derived R_g from Guinier analysis for *MtRel* ACT was determined to be 19.03 ± 0.36 Å (Table 3.7). The $P(r)$ function showed a single peak with a small tail (Fig. 3.45B) with a D_{max} of 61 ± 6.1 Å, indicating that *MtRel* ACT acquires an elongated shape in solution. The real space R_g -value extracted from the $P(r)$ function was 19.36 ± 0.13 Å, which agreed with the derived R_g -values extracted from the Guinier region of the scattering curve. Potential flexibility of the protein was further analyzed using the SAXS data. The normalized Kratky plot provides the information about the conformational behavior of examined particle. The dimensionless Kratky plot of *MtRel* ACT gave a characteristic broad bell-shaped profile shifted towards the right to the standard lysozyme (5 mg/ml), indicating the extended nature of *MtRel* ACT (Fig. 3.45C) [245, 264].

The SAXS solution experimental data was fitting nicely with the *MtRel* ACT dimer atomic model, obtained by NMR with a χ^2 -value of 0.59 obtained from Crysol analysis (Fig. 5.17E). The superimposition of the atomic *MtRel* ACT dimer to the solution shape was achieved by SUPCOMB at NSD of 1.5524 (Fig. 3.45D). The R_g (17.40 ± 1.7 Å) and D_{max} (57.32 ± 5 Å) calculated for the NMR solution dimer is comparable to the R_g and D_{max} obtained from the SAXS data in solution.

Figure 3.45: SAXS studies of *MtRel ACT*. (A) SAXS pattern ($\circ$) of *MtRel ACT* at a concentration of 0.8 mg/ml (wine), 1.5 mg/ml (green) and 3.3 mg/ml (blue), respectively. (Inset) Guinier plots show linearity at all concentrations used, indicating no aggregation. (B) Pair-distance distribution function $P(r)$ of *MtRel ACT*. (C) Dimensionless Kratky plot of *MtRel ACT* ($\bullet$; blue) compared to the compact globular lysozyme ($\bullet$; grey) with a peak ($---$; grey), representing the theoretical peak and assuming an ideal Guinier region of a globular particle. (D) *Ab initio* DAMMIF envelope superimposed onto the cartoon representation of the NMR dimer of *MtRel ACT* with subunits colored in orange and green. (E) Fitting of SAXS pattern ($\circ$; black) and calculated scattering profiles ($—$; red) from the dimeric *MtRel ACT* NMR solution structure.

Sample details	MtRel ACT		
Organism	Mycobacterium tuberculosis		
GenBank	AL123456		
Extinction coefficient (A ₂₈₀ , M ⁻¹ cm ⁻¹)	2,980		
MM from chemical composition (Da)	9,472.68		
Protein concentration (mg ml ⁻¹)	0.8 mg/ml	1.5 mg/ml	3.3 mg/ml
Solvent	50 mM Tris/HCl, pH 8.5, 350 mM NaCl, 5% glycerol, 1 mM DTT		
Data collection parameters			
Instrument (source & detector)	Bruker NanoStar equipped with MetalJet eXcillum X-ray source and VANTEC-2000 detector		
Beam geometry	100 μm slit		
Wavelength (nm)	0.134		
Sample-to-detector distance (m)	0.67		
q range (nm ⁻¹)	0.16 – 4		
Normalization	To transmitted intensity by glassy carbon		
Monitoring for radiation damage	Data frame-by-frame comparison		
Exposure time (min)	30 (6 frames x 5 min)		
Sample configuration	Vacuum tight quartz capillary		
Sample temperature (K)	288.15		
Software employed			
Primary data reduction	BRUKER SAXS program suite		
Data processing	PRIMUSqt from ATSAS 2.8.2		
Ab initio analysis	DAMMIF from ATSAS 2.8.2		
3D graphics representations	PyMOL v1.5.0.4		
Structural parameters			
Guinier analysis			
I(0) (arbitrary unit)	37.29 ± 0.98	37.50 ± 0.55	37.54 ± 0.31
R _g (Å)	19.03 ± 0.79	19.00 ± 0.46	19.03 ± 0.36
q _{min} (Å ⁻¹)	0.0229	0.0245	0.0245
qR _g max	1.29	1.29	1.30
P(r) analysis			
I(0) (arbitrary unit)	36.82 ± 0.59	37.34 ± 0.47	37.45 ± 0.27
R _g (Å)	19.10 ± 0.29	19.23 ± 0.24	19.36 ± 0.13
D _{max} (Å)	60	61.8	61
q range (Å ⁻¹)	0.0229 - 0.3444	0.0245 - 0.3035	0.0245 - 0.3035
χ ² (Total estimate from GNOM)	0.5628	0.6065	0.5696
Porod volume (V _p) (Å ³)	27387.00	28778.00	28734.10
Calculated monomeric MM (kDa) [from sequence*]	9.472	9.472	9.472
MM from V _p (V _p /1.6) (kDa)	17.116 ± 1.7	17.986 ± 1.8	17.958 ± 1.8
MM ratio (V _p : monomer)	1.8	1.9	1.9
Shape determination and model-fitting results			

DAMMIF (default parameters, slow mode, 20 calculations)	
q range ($\text{\AA}^{-1}$)	0.023 – 0.304
Symmetry, anisotropy assumptions	P2, none
NSD, number of clusters	1.807 ± 0.229 , 6
χ^2 range	0.291-0.293
Constant subtraction of intensities	0.300
DAMMIF Excluded volume (V_{ex}) ($\text{\AA}^3$)	37,105
MM from Vex ($V_{ex}/2$) (kDa)	18.55 ± 1
Atomistic modeling	
NMR structure	PDB entry
Symmetry	P1
q range	0.0229 - 0.300
χ^2	0.59
Predicted R_g ($\text{\AA}$)	17.40 ± 1.7
Vol ($\text{\AA}^3$) , Ra ($\text{\AA}$) , Dro	19950, 1.8, 0.055

Table 3.5: Data collection and scattering derived parameters for *MtRel*

3.4.8 Characterization of the dimeric interface and amino acid binding cleft

After analysing the dimeric interface and the amino acid binding clefts in the NMR structure, mutants *MtRel* ACT_H12A, *MtRel* ACT_L15A, *MtRel* ACT_A23S, and *MtRel* ACT_N28A were cloned, produced and purified as described in Materials and Methods. An equal amount of cells (~0.5 g) used for *MtRel* ACT was used for the purification of the mutant proteins. Fractions eluted at 150 mM and 350 mM imidazole in affinity chromatography using NiNTA raisins and were injected without pooling into Superdex 75 10/300 column. The chromatograms were compared for each mutant to the WT protein. The chromatograms revealed differences in the absorbance spectrum (Fig. 3.46 (A) and (B)) significantly for the peak II indicating the production of dimeric ACT. *MtRel* ACT_H12A has shown higher yield of dimeric protein in comparison to WT *MtRel* ACT and significantly higher than the other three mutants. Moreover, by comparing the stoichiometric ratios between oligomeric peak to dimeric protein peak, in case of *MtRel* ACT_L15A it is close to 1, whereas the ratio is significantly low in other cases. In addition to that, a shift of 1 ml to the right is observed in *MtRel* ACT_L15A (eluted at 14.5 ml) in the absorbance spectrum, indicating the presence of monomeric population in the sample in comparison to WT *MtRel* ACT (eluted at 12.75 ml). In accordance to NMR structural information, the comparison of UV absorbance spectrum reveals the importance of L15 in maintaining the dimer of *MtRel* ACT. Whereas, mutant H12A showed that peak II eluted almost at the same volume as of WT *MtRel* ACT and mutants A23S and N28A might have presence of mixed population.

Figure 3.46: Purification of WT *MtRel* ACT and mutants eluting at (A) 150 mM and (B) 350 mM imidazole. (a) SDS-PAGE gel showing the fractions of L15A, N28A, A23S, H12A and WT eluted (b) Gel filtration chromatograms showing the purification of ACT WT (orange) protein fraction (c) Gel filtration chromatograms showing the purification of ACT mutants L15A (red, solid), N28A (blue, dash), A23S (green, dots) and H12A (purple, dot-dash) protein fractions. Abbreviations: M (marker), L15A (*MtRel* ACT_L15A), N28A (*MtRel* ACT_N28A), A23S (*MtRel* ACT_A23S), H12A (*MtRel* ACT_H12A) and WT (*MtRel* ACT). The protein marker labelled as Lane M in all the figures was Thermofisher Scientific PageRuler™ plus prestained protein ladder of 10 to 250 kDa.

The phenomenon of change in oligomer formation upon introduction of point mutation was further confirmed using dynamic light scattering for *MtRel* ACT_L15A mutant. WT *MtRel* ACT and the L15A mutant were both used at the concentration of 1.5 mg/ml like in the SAXS experiments in the previous chapters. The DLS intensity

distribution is presented in Figure 3.48, showing a left shift for the *MtRel* ACT L15A mutant. The shift confirms the reduction in the hydrodynamic radius upon introduction of L15A mutation thereby a change in the higher oligomeric state to lower oligomeric state.

Figure 3.48: Dynamic light scattering analysis of *MtRel* ACT and the mutant L15A. Overall size analysis of WT (blue) as well as L15A (black) mutant of the ACT domain at a protein concentration of 1.5 mg/ml. Intensity distribution showed a left shift due to introduction of L15A mutation in the ACT domain, indicating the residue is important to maintain the dimeric interface of the protein.

3.4.9 NMR analysis of the *MtRel* ACT mutants H12A, L15A, A23S and N28A

To further test the role of specific residues located in dimeric interface and binding pocket of Val, we generated a series of mutants for dimeric interface and binding pocket including H12A, L15A, A23S and N28A. ^1H - ^{15}N -HSQC experiments for each generated mutant were performed in the presence and absence of V at 1:50 molar ratio to test the effect of mutation for structural change and V binding.

HSQC spectra of mutants H12A and A23S without V show almost similar peak dispersion patterns compared to WT, indicating these mutations do not affect the overall structure of the ACT domain. On the other hand, in the presence of V, the HSQC spectrum of H12A does not show any change in the most of resonances, providing evidence that H12 to alanine mutation critically affects the binding cleft of V in the *MtRel* ACT structure (Fig. 3.49).

Separated ^{15}N and HN resonances and significant chemical shift changes were observed in the HSQC spectrum of mutant A23S upon addition of V at an 1:50 molar ratio. However, the binding affinity of V to the A23S mutant showed a weaker binding than WT protein as ~50% of resonances did not shift to the bound conformation compared to almost 100% chemical shift changes of WT *MtRel* ACT when titrated at an 1:50 molar ratio with V (Fig. 3.49).

The L15A and N28A mutants showed most significant changes of resonances as compared to the WT protein, indicating that the mutation of these residues induces dramatic structural rearrangement of *MtRel* ACT dimer (Fig. 3.49). Therefore, L15 and N28 might be related to stability of dimeric interface of *MtRel* ACT. The N28A mutant showed significant line broadening of most resonances in the NMR spectra, highlighting oligomerization of the N28A mutant (Fig. 3.49).

In addition to the role of L15 and N28 in dimerization and stability of *MtRel* ACT, no chemical shift perturbations were observed upon titration of with V at an 1:50 molar ratio. The results thereby confirmed the importance of L15 and N28 in binding to valine as inferred from NMR structure of *MtRel* ACT. The L15A mutant showed significant protein precipitation during NMR data acquisition indicating its critical role in maintain the dimer and thereby stability of *MtRel* ACT. In conclusion, a stable dimer of *MtRel* ACT might be essential for the formation and maintenance of the V binding pocket.

Figure 3.49: Overlaid ¹H ¹⁵N HSQC spectra of WT (red) and mutants (H12A, L15A, A23S and N28A). *MtRel* ACT in the presence (green) and absence (blue) of valine at molar ratio 1:50.

4 Discussion

4.1 *MtRel* prefers forming homo-dimer in solution

RSH proteins are the key players behind metabolic regulation of second nucleotide messenger, (p)ppGpp. The synthesis of (p)ppGpp is carried out solely by the RSH family of enzymes. Despite its role in the stringent control, very limited information is available about the mechanisms that regulate the bifunctional RSHs. *MtRel* being identified as a potential target of the latent TB infection, seeks more attention than before to eradicate TB. In this thesis, the structural and functional characteristics of *MtRel* and its catalytic and regulatory domains have been studied. This work presented the first study of mycobacterial RSH mechanism on the basis of its structure. The figure below shows the architecture of *MtRel* and the domains studied in this thesis (Fig. 4.1).

Figure 4.1: Domain architecture of *MtRel*. The architecture of *MtRel* shows the presence of the (cyan green) hydrolase domain, (yellow) 3-helix linker bundle, (salmon) synthetase domain, (magenta) TGS domain, and (orange) ACT domain. The numbering shows the boundaries of the domains used in the study.

The *MtRel* is a 738 amino acid residues protein with the four conserved, hydrolase, synthetase, TGS, and ACT domains. Till now, least was known about the *MtRel* oligomerization state in solution and how it undergoes autoregulation independent to the ribosome. The studies performed on *Mtb* and *M. smegmatis*, showed that the Rel protein from mycobacteria can synthesize and hydrolyze (p)ppGpp *in vitro* in a ribosome independent manner at maximal basal rate [183, 190, 208]. These studies also highlighted that the catalytic activities of the NTDs were controlled by regulatory domains in the CTD by performing assays in the presence and absence of the CTD [184, 208].

In *MsRel* it has been shown that WT tends to form multimers on an SDS gel in the absence of DTT with an unambiguous presence of dimers [208]. However, the protein altered into a monomer after addition of an excessive amount of the reducing agent DTT according a performed SDS-PAGE [208]. Two dimensional gel analysis revealed that the oligomerization was a result of inter-protein disulfide bond formation and not due to the non-covalent interactions [208]. The multimer state was further

confirmed by gel filtration chromatography profiles performed in the presence and absence of 10 mM DTT and the minimum concentration of DTT to completely monomerize the *MsRel* was identified to be 10 mM with very less dimer formation at 1 mM [208].

In comparison, *MtRel* was demonstrated to form trimers in the presence of 0.1 mM DTT with the help of non-covalent interactions and the C-terminal cysteines have no role in trimer formation as 3 mM DTT concentration didn't disrupt the multimer formation according to size exclusion chromatography studies [184]. Based on the detailed biochemical analysis, in the same study the authors were able to propose that the presence of substrates such as GDP/GTP, ATP, Mg^{2+} , may affect the oligomerization state and complete activation of the protein upon monomerization.

In both the studies of *MsRel* and *MtRel*, C-terminal and N-terminal His-tag was used, respectively. The other difference observed was in the buffer condition where sodium phosphate buffer at pH 8.0 was used in *MtRel* while *MsRel* was purified with Tris-HCl, pH 6.8 buffer.

All the above discrepancies in *MsRel* and *MtRel* intrigued to explore these differences in oligomerization state by means of biophysical techniques. To further confirm the oligomerization state of the purified reduced protein, *MtRel* was studied by SAXS showing that the protein tends to form a mixture of oligomers namely monomers, dimers and trimers which were not separated by the gel-filtration chromatography. The OLIGOMER program allowed to deduce that with 73% population of molecules existing as dimers, 21% of monomers and 6% of trimers, the solution fits nicely with the experimental data. Based on SAXS data, *MtRel* in the presence of 1 mM DTT was proposed to form stable dimers (Fig. 4.2). These findings further agrees with the oligomeric behavior of the *EcRelA* and the *MsRel* proteins [208, 268]. This is suggestive of formation of *MtRel* oligomers due to interaction of thiols present in the C-terminal (395-659) of the protein excluding the terminal ACT domain which doesn't harbor any cysteine. However, the tendency to form higher unspecific oligomers while concentrating the protein didn't allow to pursue high resolution structural studies like X-ray crystallography. Nevertheless, the low resolution shape calculation and superimposition of shape to the *MtRel* dimer generated by SASREF showed that the homo-dimer of *MtRel* in the head to tail manner shows better fitting than the monomer.

To gain a better understanding into the individual domains of *MtRel*, the domain by domain characterization was performed.

Figure 4.2: Oligomeric state of *MtRel* in solution. *MtRel* exists as a mixture of various oligomeric forms in solution in the ratio of monomer (grey): dimer (light blue):trimer (pale pink)::21:73:6 as determined through SAXS studies. The mixture shows preference to form dimers in solution. The higher oligomers tend to form monomers upon addition of substrates as shown in *MtRel* NTD [184] or by increasing DTT as shown in *MsRel* [208].

4.2 *MtRel* NTD shows species specific differences to the non-mycobacterial bifunctional long RSH

The crystal structure of *MtRel* NTD forms a stable dimer which was also revealed by PDBePISA sever [43], with the complex formation significance score (CSS) of 0.68 (Fig. 4.3B). Chains B and C within the asymmetric unit forms one dimer and chains A and D form another dimer with symmetry-related molecules. The two subunits in the dimer form an extended head to head orientation and upon dimer formation, a total surface area of 2,002 Å² becomes buried at the dimer interface.

The only crystal structure available prior to this study was of *SeRel* NTD (1-385 amino acid residues). The homolog *SeRel* had been reported as two monomers with each locked in a different conformation state. One conformer represents a nucleotide-free and so-called hydrolase-ON/synthetase-OFF form, whereby the second conformer shows the nucleotide-bound and hydrolase-OFF/synthetase- ON form [167]. *MtRel* NTD was compared to *SeRel* to highlight the differences between two structures. *MtRel* NTD was aligned to the two monomers of *SeRel* and analysed by PDBeFOLD [233]. The nucleotide free monomer of *SeRel* aligned with the *MtRel* NTD with a Z score of 12.4 and r.m.s.d 1.67 for 326 residues. On the other hand, the nucleotide bound monomer of

SeRel showed $Z = 14.4$ and $r.m.s.d. = 1.92$ for 310 residues. (Fig. 4.3). The sequence alignment of *MtRel* NTD was also performed to *SeRel* to highlight the residues.

(A)

(B)

Figure 4.3: Sequence alignment between *MtRel* NTD, *EcRelA* NTD, *EcSpoT* NTD and *SeRel* NTD and structural alignment between *SeRel* NTD and *MtRel* NTD [228]. (A) Sequence alignment of *MtRel* NTD, *EcRelA* NTD, *EcSpoT*, and *SeRel* NTD. All the residues conserved among RSH family members are highlighted in *blue*. The secondary structure elements of *MtRel* NTD and *SeRel* NTD are shown. The tetrad residues, which are conserved in the hydrolase domain and bind to single divalent (Mn^{2+}) metal ion, are shown in the *red* box, and D265 and E325, which are responsible to bind to single divalent (Mg^{2+}) ion required for the catalytic activity of the synthetase domain, are shown within the *brownish crimson* box. The residues which are present in the catalytic site of synthetase domain, and required for binding to nucleotides, are shown in a *black* box with *red asterisks*. (B) Structural alignment between the nucleotide-free *SeRel* NTD (PDB: 1VJ7) (hydrolase, overlapping 3-helix bundle and synthetase domain colored in *wheat*, *orange* and *cyan* respectively) and *MtRel* NTD (PDB: 5XNX) (hydrolase domain - *cyan green*, synthetase domain - *salmon*, overlapping three helix bundle - *yellow*.) reveals the significant differences in the loop regions.

Upon careful examination of the sequence alignment, significant difference was observed in the loop region between $\beta 3$ and $\beta 4$, including residues 293MAGRFDYIAQPRYGVYQSL312, which encloses two conserved nucleotide pockets with the KDY- and NGYQSL/IH-motif, respectively (Fig 4.3A). It can also be observed in the structural alignment of two homologs (Fig 4.4). In *MtRel* NTD, the residues NG of the NGYQSL/IH-motif are substituted by the amino acids G307 and V308, respectively. Interestingly, the counterpart of the mycobacterial G307 in *SeRel* NTD is residue N306, which forms a hydrogen bond with the guanine base of GDP (Fig. 4.4). Furthermore, amino acid K304 of the *SeRel* NTD is also involved in guanine base recognition, which is replaced by arginine in the *MtRel* NTD, whereas the phenolic ring of the highly conserved Y308 of *SeRel* stacks against the guanine base (Fig. 4.4). The conserved 298KDY300 is involved in binding with the phosphate moiety of GDP [16].

The comparison reflects that the nucleotide-binding mode might be different in the mycobacterial Rel protein.

Figure 4.4: A structural alignment between the nucleotide-bound *SeRel* NTD monomer (PDB: 1VJ7) and *MtRel* NTD (color coded as in Fig 4.3B) [228]. The inset shows the residues of the KDY- and KANGYQSL/IH-motif of $\beta 3/\beta 4$ loop in the synthetase domain, that are involved in the binding of GDP. All the residues interacting with GDP are presented as sticks (cyan; 1VJ7 and salmon; *MtRel* NTD) and are labeled. The red and purple spheres represent Mg^{2+} and Mn^{2+} , respectively. The red arrow is pointed toward the loop region between $\alpha 8$ and $\alpha 9$ in the 3-helical bundle helix, which is stabilized in the nucleotide-bound form of *SeRel* NTD.

Importantly, the novel nucleotide-free *MtRel* NTD provides insights into the loop of residues 152TMRFLPPE159 between $\alpha 8$ and $\alpha 9$, which are missing in the nucleotide-free *SeRel* NTD form. In comparison, this loop is stabilized in the nucleotide bound state of *SeRel* NTD, and sandwiched between two nucleotide-binding pockets of the hydrolase domain. In contrast, the loop including the RXSGD/EPYI nucleotide binding motif between helix $\alpha 2$ and $\alpha 3$ of the *SeRel* NTD structure is missing in the *MtRel* NTD counterpart.

The C-terminal helix $\alpha 16$ and the loop joining $\beta 3$ and $\beta 4$ are found to be involved in hydrogen bonding and forming salt bridges between the two chains at the dimer interface. Residues R277 and D278 of both subunits stabilize the dimer interface through salt bridge interaction. Moreover, residue D299 of one subunit and S288 of other subunit forms a hydrogen bond with a distance of 2.63 Å, in addition to many hydrophobic and van der Waals interactions. In the crystal lattice, these *MtRel* NTD dimers form a helical packing in three dimension. The other noticeable difference is in the dimer interface, which is present in the synthetase domain and involves residues 276IRD-CYAAVGVVHSL289 of helix $\alpha 16$ and the loop region residues 293MAGRFDYIAQPRYGVYQSL312, positioned in between $\beta 3$ and $\beta 4$ (Fig. 4.5).

Figure 4.5: Intermolecular interactions between two monomers of *MtRel* NTD [228]. The two molecules A; ruby and D; sand, forming stable dimers (PDBePISA server), are shown. Inset, residues R277, D278, S288, and D299, forming hydrogen bonds, are shown as sticks at the dimer interface.

The dimer inter- face is primarily stabilized by hydrogen bonds and salt bridges. The residues R277 and D278 located in helices $\alpha 15$ and $\alpha 16$ are involved in salt bridge and hydrogen bonding which stabilizes the dimer interface. Amino acid D278 is conserved in the RSH family and involved in nucleotide binding as shown in the Fig. 4.3D. In addition, residue C279, which is adjacent to the conserved nucleotide binding D278, is also conserved and is in very close proximity to C273 inside $\beta 2$, to form a possible intra-disulfide bond. However, both cysteine residues are reduced in the presented structure, since the protein was purified and crystallized under reduced conditions. Interestingly, the cysteine residue of the 279CY280-motif, which is conserved among the RSH protein family, is substituted in the related *SeRel* NTD by a valine. Fluorescence resonance energy transfer experiments reveal that the distance between the cysteines of *MsRel* equivalent to *MtRel* C279 of synthetase domain and CTD C633 from the intrinsically disordered region increases upon binding of deacylated tRNA, indicating that upon uncharged tRNA binding, a conformational change occurs which abolishes the interaction between C633 and the active site of synthetase, and helps in regulating the synthetase activity [198].

The NTDs of the two Rel homologs reveal striking differences prevailing in the C-terminal region of the synthetase domain, including residues 366DDMAWMRQLLDWQREAA382 as shown in Fig. 4.6. While the mycobacterial protein reveals an entire α -helix formed by the peptide 366DDMAWMRQLLDWQREAA382, a short helix formed by the amino acids

363NWIKELEVEL371 was observed for the *SeRel* NTD. The reason might be the presence of a C-terminal His-tag in the latter. This region connects the synthetase domain with the TGS-domain of the regulatory C-terminal of *MtRel*, described to be responsible for tRNA binding in *E. coli* RelA [165].

Figure 4.6: Overlapping structure of *MtRel* NTD and *SeRel* NTD showing the C-terminal region of the synthetase domain [228]. A different orientation of Fig. 4.1 at 90° along the Y-axis revealing the significant difference in the C-terminal region of the aligned structures of *MtRel* NTD and *SeRel* NTD. The 2FO-FC map contoured at 1 σ for the C-terminal helix of *MtRel* NTD 366DDMAWMRQLLDWQREAA382 is shown in the inset.

4.3 The significance of understanding regulatory domains of *MtRel*

It is well understood that stringent factors which belong to RSH family gets activated upon accumulation of deacylated tRNA which is responsible to stall translating ribosomes [190, 193]. In recent years, the cryo-EM structures have provided more insights into the interactions of *EcRelA* to the stalled translational machinery (ribosome-deacylated tRNA-mRNA) which leads to activation of *EcRelA* synthetase domain [138, 165, 269]. However, unlike *EcRelA*, *MtRel* and *MsRel* can synthesize and hydrolyze (p)ppGpp at maximum basal rate and the affinity to substrates of transferase reaction further increases upon addition of uncharged tRNA [184, 208]. The stalled ribosome-RelA model explains the interactions made by the C-terminal domains of RelA with ribosomal RNA, ribosomal proteins and deacylated tRNA [165]. However, these interactions are not sufficient to explain the role of these domains in the regulation of catalytic domains in the absence of stalled ribosomal machinery. The homologous

region of *MsRel* to the C-terminal region (394-738) of *MtRel* has been explored with the help of site directed mutagenesis and Förster resonance energy transfer method and a model has been proposed explaining interaction of the C-terminal with the N-terminal domains with the help of the flexible and mostly unstructured linker region to regulate the catalytic activities [196, 198, 208, 270]. The TGS domain is known to regulate synthetase activity while interacting with components of the translation machinery [165]. The ACT domain is shown to be found in proteins which are regulated by amino acid concentration [200, 271]. It is also known that there is an intrinsically disordered region of ~200 amino acids between the TGS- and ACT domains which ensures domain-domain interactions by bringing flexibility to the structure [270]. The studies highlighting effect of deletion of CTD on synthetase activity has been summarized below in Fig. 4.7 [184, 193, 198, 208].

In *MsRel*, it was shown that TGS and ACT had a role in inter-domain cross-talk with the NTD, and the (p)ppGpp binds to the region between the TGS and ACT domain, which provides additional positive feedback regulation for the synthetase activity [198, 208, 272]. All the studies mentioned above explain the regulation of the synthetase domain whereas the mechanism behind the control of hydrolase activity is poorly understood. Nevertheless, the role of the TGS and ACT remain unknown in regulating catalytic domains in a ribosome independent manner in mycobacteria.

Figure 4.7: Comparison of *MtRel* and *MsRel* constructs for the synthetase activity. A pictorial representation of the deletion constructs studied from *MtRel* and *MsRel* to understand the effect on catalytic activity.

4.4 Novel *MtRel* TGS exhibits conserved structure across RSH homologs and interacts with tRNA

The structure of the TGS domain from *E. coli* RelA bound to deacylated tRNA (dtRNA) was extracted from PDB 5IQR where the interactions can be observed to be specific for the TGS domain (Fig. 4.8). The residues P411, K412, H432 and R438 were shown to interact with dtRNA. In a recent study, it has been shown that H432 is crucial for the *EcRelA* to recognize dtRNA and in turn formation of the RelA-tRNA complex, which enables the RelA to bind the ribosome and activates the stringent response.

Figure 4.8: Structure of *EcRelA* TGS domain interacting with deacylated tRNA [165]. The residues involved in stabilizing the interaction (blue) with deacylated tRNA (green) are labelled in red on the *EcRelA* TGS domain structure shown in pale blue.

The structure of *MtRel* TGS was aligned to the *EcRelA* TGS domain (PDB ID 5IQR). Both the structures showed r.m.s.d. of 1.01 Å of the Cα atoms. When these residues were compared sequentially to that of *MtRel* TGS and other homologs, overall conservation of H33 was observed corresponding to H432 in TGS domain of *EcRelA* (Fig. 4.9A). Moreover, when these residues were mapped on the structure of *MtRel* TGS, the orientation of the side chains specifically that of H33 was noted highlighting the interaction of the conserved histidine of helix-turn-helix motif stacked within the two cytosines of 3' dtRNA (Fig. 4.9B). Brown et al. also showed the interaction of K412 with the terminal adenosine and proposed that β5 precluded the binding of acylated tRNA by packing against 3'OH of the terminal adenosine of dtRNA [165]. In *MtRel* TGS the titration with dtRNA revealed a better understanding of the residues involved in the interaction of the same. The residues mapped upon titration of *MtRel* TGS to tRNA showed the identical residues being involved in the interaction sites as observed in the *EcRelA* TGS along with the residues indirectly affected upon binding as shown

(A)

	10	20	30	40	
<i>MtRel_TGS</i>	DLAVQEI FVFT	PKGDVIT LPT	GSTPVD FAYAV	HTEVGHR	CI GARVNGRLV 50
<i>MbRel_TGS</i>	DLAVQEI FVFT	PKGDVIT LPT	GSTPVD FAYAV	HTEVGHR	CI GARVNGRLV 50
<i>MsRel_TGS</i>	DLATQEI FVFT	PKGDVIT LPT	AGSTPVD FAYAV	HTEVGHR	CI GARVNGRLV 50
<i>MIRel_TGS</i>	DLAVQEI FVFT	PKGDVIT LPT	GSTPVD FAYAV	HTEVGHR	CI GARVNGRLV 50
<i>SaRel_TGS</i>	DLQSDKVYAFT	PASDVIEL	LPYGAVPID FAYAI	HS EVGNKM	IGAKVNGKIV 50
<i>PgRel_TGS</i>	- - - MEVVF	TPKGEIKRLP	QATALD FAYS	LHS DLGDHC	IGAKVNHKL 50
<i>CIRel_TGS</i>	DLAPEEVFVFT	PKGDVIS LP	IGSTVID FAYAI	HS AVGNRM	IGAKVDGRIV 46
<i>EcSpoT_TGS</i>	- - - DEIYVFT	PEGRIVE LP	AGATPVD FAYAV	HTDIGHAC	VGARVDRQPY 46
<i>EcRelA_TGS</i>	- - - DRVYVFT	PKGDVVD LP	AGSTPLD FAYAI	HS DVGHR	CI GAKIGGRIV 46

	60	70	
<i>MtRel_TGS</i>	ALERKLENGE	VVEVFT - - - -	66
<i>MbRel_TGS</i>	ALERKLENGE	VVEVFT - - - -	66
<i>MsRel_TGS</i>	ALERKLENGE	VVEVFT - - - -	66
<i>MIRel_TGS</i>	ALERKLENGE	FVEIFT - - - -	66
<i>SaRel_TGS</i>	PIDYILQTGD	IVEIRT - - - -	66
<i>PgRel_TGS</i>	PLSYVLNSGD	QVEVLSKSL E	67
<i>CIRel_TGS</i>	PIDYKVKTGE	IIDVLTTK - - E	69
<i>EcSpoT_TGS</i>	PLSQPLTSG	QTVEIIT - - - -	62
<i>EcRelA_TGS</i>	PFTYOLMGDO	IEIIT - - - -	62

(B)

Figure 4.9: Sequence and structure alignment of *MtRel* TGS with its homologs [165]. (A) The sequence alignment of TGS domain of RelA/SpoT homologs from *Mtb* (*MtRel*), *M. bovis* (*MbRel*), *M. smegmatis* (*MsRel*), *M. leprae* (*MlRel*), *S. aureus* (*SaRel*), *Porphyromonas gingivalis* (*PgRel*), *Clostridium leptum* (*ClRel*), and *E. coli* (*EcSpt* and *EcRelA*) are shown. The sequence alignment shows the residues in shades of blue which are identical at the threshold of 30%. The red boxes represent the conserved residues shown to interact in *EcRelA* TGS domain from PDB-5IQR with deacylated tRNA. (B) The structure alignment of *MtRel* TGS and *EcRelA* TGS domains show the conserved nature of residues involved in interacting with TGS confirmed by NMR data as well. The residues involved in stabilizing the interaction (blue) with deacylated tRNA (green) are labelled in red on the *EcRelA* TGS domain structure shown in pale blue. These conserved residues in *MtRel* TGS (magenta) structure are shown in yellow and labelled in black.

4.5 *MtRel* ACT exists as dimer in solution

Interestingly, the ACT domain which is conserved among various enzymes including RSHs, binds to amino acids and directly/allosterically regulates the enzyme activity [199]. Mutagenesis data have shown that despite the conserved ACT-fold, the proteins like *EcSerA* and *PfLrpA* harbor different binding sites for the ligands [273, 274]. In addition to that, the sequence comparison of these ACT domain homologs give clear idea about their conserved features in mycobacterial species and how the ligand binding sites are different from non-mycobacterial species (Fig. 4.10). The comparison shows that irrespective of the conserved fold of the ACT domains as seen in structural alignment, the unique sequence of *MtRel* makes it a closer relative to the RAM domain family proteins with an invariant glycine in the $\beta 2$ - $\beta 3$ connecting loop (Fig. 4.10). However, the glycine residue conserved in the $\beta 1$ - $\alpha 1$ loop, an atypical feature of the ACT domain carrying proteins is replaced with arginine residue in mycobacterial species. This also highlights the presence of mycobacterial-specific cavities for small molecules in *MtRel* (Fig. 4.10).

Figure 4.10: Sequence alignment of the ACT domain of mycobacterial and non-mycobacterial RSH family proteins with *E. coli* 3-phosphoglycerate dehydrogenase (*EcSerA*) and *P. furiosus* Leucine receptor protein A (*PfLrpA*). The mycobacterial Rel proteins from *Mtb*, *M. bovis*, *M. smegmatis* and *M. leprae* aligned with Rel homologs from non-mycobacterial species, *S. aureus*, *C. tepidium* and *E. coli* in addition to *EcSerA* and RAM domain carrying *PfLrpA*. The sequence alignment shows the conserved glycyl residue in $\beta 1$ - $\alpha 1$ loop of ACT domain carrying proteins is replaced by arginine in mycobacterial Rel proteins and by glutamic acid in *PfLrpA*, a representative of RAM domain carrying proteins (magenta highlighted residues). The residues highlighted in green shows the conserved glycyl residues in $\beta 2$ - $\beta 3$ loop of *PfLrpA* and mycobacterial Rel protein is not conserved in atypical ACT domain carrying proteins. The magenta asterisk marked residues shows the residues of *MtRel* ACT mutated in this study as H12A, L15A, A23S and N28A.

These intriguing features led us to propose that the ACT domain of *MtRel* is unique sequentially and therefore, the interactions of *MtRel* ACT might be different at intramolecular level with other domains or intermolecular level while forming oligomers or interacting with stalled ribosomes. In this thesis, *MtRel* ACT domain has been studied for its structure with the help of NMR spectroscopy and oligomeric state by SAXS.

The novel structure of *MtRel* ACT was analyzed for the residues involved in dimeric interface. The dimeric interface involves hydrophobic interactions and shows the involvement of $\alpha 1$ and $\beta 2$ in keeping the monomers as stable homo-dimer. As shown in the Fig. 4.11, H12 and L15 in chain A were found to interact with N28 and I29 in chain B thereby forming a possible binding cleft. As *MtRel* is known to bind ligands such as amino acids and nucleotides in its homologs, this cleft was proposed to be the regulatory site of *MtRel* ACT domain.

Figure 4.11: Surface representation of dimeric interface based on the NMR data and determined solution structure. Residues involved dimeric interface and possible binding epitope for ligand are highlighted as blue and red color and labeled as one letter code.

4.6 The *MtRel* ACT dimer shows biological relevance understood on the basis of the *EcRelA* complex with stalled ribosomal complex

The significance of *MtRel* ACT dimer can be understood by the interactions made by *EcRelA* with rRNA and r-protein when bound to stalled translational machinery (ribosome-mRNA-uncharged tRNA at A-site). In recent structural study presented for *EcRel*, the ACT domain is shown to interact with deacyl-tRNA as well as 23S rRNA while also coming in close contact of helical domain structured downstream to TGS upon binding of tRNA (Fig 4.12) [138, 275]. The authors suggested that interaction between the linker helix and the ACT domain may serve as a possible dimerization surface of the ACT domain in free *RelA* which gets disrupted upon ordering of *RelA* upon tRNA binding [138].

Figure 4.12: Inter and intramolecular interactions of the ACT domain of RSHs. (A) The structure of *EcRelA* (greyish blue) bound to 50S subunit (grey) showing 23S rRNA (blue surface) of *E. coli* is shown. The *MtRel* homology model as described in Fig. 4.1 is overlapped to *EcRelA*. (B) and (C) show the zoomed in 23S rRNA and helical domain interaction with the ACT domain, respectively.

The figure 4.13 shows the structural alignment of *MtRel* ACT to the *EcRelA* ACT (from PDB 5IQR) with a low r.m.s.d. score of 1.39 Å for all the atoms. This shows that the ACT domain is majorly conserved structurally between the two proteins however the residues in the RNA recognition motif and the ACT-likelike helix dimer interface are not conserved. These observations further point out a different arrangement of *MtRel* ACT in the full-length protein while interacting with stalled ribosome.

Figure 4.13: Structural alignment of *MtRel* ACT and *EcRelA* ACT from the cryo EM structure PDB 5IQR. The *MtRel* ACT (orange) shown to be completely aligned to *EcRelA* ACT (light blue) with an r.m.s.d. of 1.39 Å. The residues which are observed interacting with rRNA and linker helix are shown in blue (*EcRelA* ACT) and the corresponding residues from *MtRel* ACT are in brown.

The above findings along with our results are suggestive of why the NMR titration study of *MtRel* ACT with ssRNA harboring the conserved motif didn't show any interaction. As shown in the structural alignment, the residues R700 and K704 of *EcRelA* which interact with A-site finger of 23S rRNA are replaced by V34 and G38 in *MtRel* ACT (Fig. 4.13). These findings suggest that bifunctional *MtRel* indeed show species specific features which completely changes the way how *MtRel* interacts with the stalled translational machinery.

4.7 *MtRel* ACT dimer interacts with branched chain amino acid

In another study, the role of ACT domain was analyzed in the *RcRel* [205]. The authors proposed that RSH proteins are able to sense branched chain amino acids and maintain proper cellular concentration of (p)ppGpp by fine tuning the hydrolase activity. This can also be understood as an additional mechanism required to maintain the balance

between opposing catalytic activity independent of ribosomal translational machinery. In order to understand the essentiality and specificity of *MtRel* ACT in interacting with BCAAs NMR titration experiments were performed. In comparison to *RcRel* and other RSH enzymes (including *EcRelA*) studied by the authors for the BCAA specificity, our studies on *MtRel* ACT showed specific binding to valine as seen in the case of *EcRelA*. The inability of *MtRel* ACT to bind leucine and isoleucine further intrigued to pursue complex of *MtRel* ACT and valine and map the residues making such a specific binding pocket. The careful examination of the chemical shift perturbations upon addition of valine showed that the binding of valine brings drastic structural rearrangement in the *MtRel* ACT structure. The structural rearrangement upon binding of BCAA hints that restoration of amino acids might change the stability of *MtRel* interactions with the stalled ribosomal machinery and thereby interfere with the synthetase activity. On the other hand, this structural rearrangement might be essential in allosterically or directly regulating the hydrolase activity independent to the ribosomal complex when amino acids are sufficient.

The mapped residues further enabled to identify the binding cleft of valine as mentioned in chapter 3.4 which represents the dimeric interface. When the binding cleft of serine of *EcSerA* ACT was compared to *MtRel* ACT, it was observed that the serine binds at the loop of $\alpha 1$ and $\beta 1$. The conserved glycine in ACT domains interacting with amino acids, is replaced by an arginine in *MtRel* ACT (Fig. 4.14).

Figure 4.14: Structural comparison of *MtRel* ACT to a serine bound *EcSerA* ACT dimer. The dimer of *MtRel* ACT is structurally aligned to the *EcSerA* ACT having serine bound in its dimeric surface. The zoomed in structure shows the highly conserved glycine (G349) in amino acid binding ACT domains being replaced by arginine (R13) in *MtRel* ACT. The 45° rotation of the overlapped structure shows the residues involved in serine binding in *EcSerA* ACT (magenta, label in magenta) and residues mapped for valine binding in *MtRel* ACT (grey, label in black).

Moreover, the binding mode observed on the basis of CSPs in the presence of valine and the mutagenesis study of *MtRel* ACT is significantly different. The binding

of valine with *MtRel*ACT requires the dimeric interface while the serine was found to be bound to monomer of *EcSerA* ACT. The presence of valine binding site at the dimeric interface might be required for allosteric regulation by bringing structural rearrangement in the protein, upstream to ACT.

The overall topology of *MtRel* ACT was compared to the previously reported structures of the ACT domain from *E. coli* namely from 3-phosphoglycerate dehydrogenase of *E. coli* (*EcSerA*, PDB ID code 1PSD) and *E. coli* RelA (PDB 5IQR). The root-mean square deviation upon comparing the monomer structures of *EcSerA* and *EcRelA* to *MtRel* was 1.74 Å and 1.71 Å, respectively (Fig. 4.15A). However, significant difference was observed in the electrostatic charge distribution and surface shape of *MtRel* ACT in comparison to *E. coli* counterparts indicating species specific differences that might affect the molecular interaction and regulation mechanism during stringent response (Fig. 4.15B, C, and D).

Figure 4.15: Structural comparison of *MtRel* ACT with the ACT domain of *E. coli*. (A). Ribbon representation of *MtRel* (red), *E. coli* SerA (blue) and *E. coli* RelA (green) ACT domains after superposition for Ca atoms from C-terminal view. (B)-(D). Electrostatic potential surface of ACT domains of *MtRel* (B), *E. coli* SerA (C) and *E. coli* RelA (D), respectively.

This work has structurally revealed that how the mutation of conserved N28 as well as L15 in *MtRel* ACT shows defect in valine sensing capability of *MtRel* ACT. In addition to that, L15 is shown to be crucial to maintain the dimeric interface. Valine concentration sensing capability of *MtRel* ACT, which brings drastic rearrangement in the domain structure upon binding, might help in understanding the allosteric role of ACT domain in regulating catalytic domains in response to environmental cues. Moreover, the evidence was provided that the *MtRel* synthetase activity auto-regulated

by direct interaction of N-terminal domains with the ACT domain independent to ribosome, uncharged tRNA and mRNA complex. Upon H671 (corresponding to ACT H12) substitution to alanine, even though no structural rearrangement was observed in stand-alone ACT domain, the mutation affected the synthetase activity of *MtRel*.

In conclusion, the branched chain amino acid concentration as reviewed recently by Kaiser and Heinrichs [276] plays a significant role in alteration of bacterial physiology and virulence. This can be linked to the nutritionally starved *Mtb* which when devoid of branched chain amino acid can't maintain synthesis of branched chain fatty acids, required to support its complex cell membrane formation and thereby enters the stage of dormancy where it stops growing. This low concentration of BCAAs is sensed by the ACT domain which then regulates the accumulation of (p)ppGpp. Once the valine concentration in particular restores, the structural rearrangement occurs in *MtRel* which leads to activation of the opposite activity i.e. hydrolysis of (p)ppGpp, signaling amino acid restoration and stringent response shut down (Fig. 4.16).

Figure 4.16: The schematic presenting the role of *MtRel* and its regulatory domain in nutrient starved condition. Under nutrient deprived condition the concentration of branched chain amino acids declines due to which the fatty acid required for the formation of mycobacterial cell membrane are not available. The absence of surplus amount of branched chain fatty acid signals the growth and division to shut down which leads to cells to enter dormant stage. This metabolic shut down is sensed by *MtRel* by sensing the low valine concentration and thereby starts synthesis of (p)ppGpp. Upon amino acid restoration as a result of stringent response, the valine concentration rises which is sensed by *MtRel* ACT domain. This leads to allosterically/directly regulate the synthetase activity and increase the hydrolysis of (p)ppGpp to restore growth and division.

The dependence of hydrolysis activity on the ACT domain has also recently been understood in *Caulobacter crescentus* SpoT which is sensitive to nitrogen and carbon starvation [202]. The species specific differences of *MtRel* in the domains other than the catalytic NTDs will attract attention to develop therapeutics against different epitopes that will not generate (p)ppGpp analogs like Relacin which suffered impermeability through mycobacterial cell envelop [98].

Based on our structural and titration findings together with the information from the literature about *EcRelA* interactions with the ribosomal components, a mechanism for *MtRel* is proposed in which the binding of amino acids to the ACT domain and helical domain pseudo-dimer interface, brings conformational rearrangement in the domain, which leads to destabilization of the interaction of ACT with 23S rRNA. This can lead to activation of the hydrolysis activity upon release of *MtRel* being off the ribosome. The rearrangement of the *MtRel* ACT will induce changes into the consecutive zinc finger domain which makes the majority of the interaction with the ribosome and can also lead to destabilization of the helical domain which keeps ACT close to rRNA and r-protein. As drastic rearrangement has been observed upon valine binding to the *MtRel* ACT dimer, it explains that upon amino acid restoration, *MtRel* interacts with amino acid and brings back the hydrolysis activity by causing conformational changes in the overall protein structure.

5 Conclusions

Ever since the discovery of *Mtb* as the pathogen of TB, it remains among the diseases which causes the highest mortality despite all the efforts taken for its eradication. It is crucial to understand the stages of TB infection and develop a better therapy which could work for patients with MDR and XDR TB or can potentially work as preventive therapy in the LTBI patients. With the developments made in understanding the physiology of the persistent bacteria and the factors responsible to bring latency, the *MtRel* emerged as an attractive target to eradicate the dormant bacteria. It has been observed in the previous studies that the critical balance of (p)ppGpp concentration is required for regulation of stringent response. Maintaining this crucial balance of (p)ppGpp concentration during chronic infection is critical for *Mtb* survival under hostile conditions such as nutrient limitation. In addition to that, studies have shown that blocking the entry of *Mtb* in quiescence and reducing the tolerance to Isoniazid by targeting *MtRel* in nutrient starved condition might be beneficial in developing shorter regimens. It will also be interesting to understand the efficacy of *MtRel* inhibitor in patients with LTBI in high TB burden settings including patients with HIV and women with HIV who are pregnant or postpartum. The monotherapy of isoniazid in such cases is too long (36 months) which decreases the compliance rate and the other drugs such as Rifamycins induce drug-drug interaction leading to failure of other ongoing treatment. Such regimens will need further clinical trials to improve their dosage and duration and adding a potent inhibitor of *MtRel* in such regimes may help in sterilizing the infection completely in shorter time [10]. Although, LTBI is asymptomatic, as per CDC and WHO guidelines, it has been recommended to do the interferon- gamma release assays such as ELISA and ELISPOT in addition to Tuberculin Skin Test (TST) for LTBI patients. These tests are also used in the clinical trials of isoniazid and rifamycins for LTBI patients. The positive blood test indicates that the white blood cells from the patients shows immune response when mixed with the *MtRel* antigens confirming the previous interaction with the pathogen even if active TB disease was not developed. These tests are generally positive for LTBI patients with a typical normal chest radiograph and negative for smear or culture test. It is important to perform all these tests in addition to sputum examination and analysis of medical history before starting an LTBI treatment to eliminate chances of failure due to inadequate treatment or development of drug resistance [10].

In addition to TB preventive therapy, it will be interesting to study the efficiency of regimes for MDR/RR-TB, upon addition of stringent response inhibitor. Blocking

the hydrolysis of (p)ppGpp can have deleterious effect on persisters by compromising the viability of the cell and thereby aid in treating chronic infection and can have synergistic effect to new TB drugs in the pipeline [189].

A strategy to target the epitopes of *MtRel* which are crucial for the bacteria to sense these environmental cues in addition to the ones inhibiting the synthetase activity has been presented. These findings will aid in *Mtb* targeted therapeutics development with better precision to fight persisters. In this work, the novelty of *MtRel* brought into the light by means of structural and mechanistic studies. A diverse set of experiments including biophysical methods like SAXS, x-ray crystallography and NMR spectroscopy aided in to look deeper and gain better understanding of the unique features of *MtRel*. The SAXS results allowed to confirm the predominant oligomeric state of *MtRel* in solution and the co-existence of a dynamic mixture of monomer, dimer and trimer. The existence of an oligomeric state in solution emphasized the role of auto-regulation by oligomerization and the presence of catalytic activity to produce AMP as the by-product confirmed the activation of *MtRel* in the presence of substrates. The mixture of different populations of *MtRel* in solution and the aggregation upon concentrating the protein posed the challenge to carry out high resolution structural studies. However, the low resolution studies on *MtRel* directed the curiosity to gain atomic resolution details about the *MtRel* in domain-wise manner.

The unique features of the *MtRel* NTD sequence and the simultaneous regulation of both the activities by the *MtRel* at the basal level in comparison to the non-mycobacterial homologs which show a conformational switch, intrigued to carry out structural studies. The SAXS experiments revealed the homo-dimer formation of the *MtRel* NTD in solution with no effect on the dimeric shape upon addition of substrates showing no drastic movements in the domains at the macroscopic level. Additionally, the crystal structure of the NTDs of the *MtRel* gave insights into its uniqueness in comparison to the homolog structure of the *SeRel* (PDB: 1VJ7). The structural comparison enabled in suggesting unique binding pocket in the *MtRel* which were not conserved in the *SeRel* NTD or the *EcRelA* NTD. The work also proposed additional residues that can be studied by mutagenesis to understand their critical role in catalytic activities of the *MtRel*. The activity comparison of the *MtRel* and the *MtRel* NTD re-confirmed the role of the CTD in auto-regulating the catalytic activities as the synthetase activity showed a 3-fold drop by the *MtRel* in comparison to the *MtRel* NTD.

Furthermore, attempts were made to study the *MtRel* CTD. Due to the similar behaviour of the CTD fragment as in the case of the *MtRel* which aggregates while concentrating, the construct was not pursued for structural studies. However, the *MtRel* TGS and the *MtRel* ACT domains which mark the boundaries of the CTD fragment, were separately generated and studied by NMR techniques. Irrespective of the conserved nature of TGS domains across mycobacterial and non-mycobacterial homologs, the first structure of *MtRel* TGS was explored for structural changes in the presence and absence of deacylated tRNA. The results shed light on the critical residues of the TGS domain which can interact with deacylated tRNA even in the absence of ribosome. So far, it has only been shown in the *EcRelA* that the TGS interacts with tRNA even in the absence of ribosome. The histidine, which is shown to be critical in *EcRelA* to interact with deacylated tRNA CCA tail, is revealed to be conserved structurally in the *MtRel* TGS domain in the helix-turn-helix motif.

The NMR structure and SAXS studies of the *MtRel* ACT revealed the dimeric nature of the domain in solution. The dimer formed in the head to tail fashion showed the binding pockets at the dimeric interface for the interaction of small molecules which was confirmed with the NMR titration experiments. From the results, it is validated that among all other branched chain amino acids, valine interacts with the *MtRel* ACT at the dimeric interface and causes structural rearrangements. The structural studies highlighted species-specific differences in this regulatory domain and also enabled to map the valine binding site for the first time in any of the RSH proteins. The effect on the structure of the *MtRel* ACT and hence the role of sensing environmental cues by means of this domain confirmed and presented structurally. The critical residues in the dimeric interface and the amino acid binding site were identified by mutagenesis experiments. Further, the effect of mutations on valine binding was confirmed by NMR titration experiments.

The insights gained in this thesis are imperative to the future work in understanding the complex structure and mechanism of the *MtRel* in the presence of ribosomes. It is important to investigate the entire *MtRel* protein in the holistic manner. As mentioned earlier in this thesis, it is difficult to concentrate the *MtRel* to the desired high concentration in a homogenous manner. Nevertheless, SAXS results have validated the coexistence of a monomer, dimer and trimer in solution, making it difficult for cryo-electron microscopy studies. I hereby propose that the *MtRel* can be studied by either reconstituting it with the ribosome in the presence of uncharged tRNA and mRNA or

by introducing substrates for synthetase domain i.e. GTP and ATP, to stabilize the monomers in the solution. Now, it is possible to solve the structures of proteins of molecular weight close to 60 kDa at high resolution as reported in the study of human hemoglobin conducted by Khoshouei et al 2016, in which the structure was solved at a resolution of 3.2 Å [277]. Moreover, the difficulties in concentrating the homogeneous *MtRel* (81 kDa) to higher concentration can be avoided by using cryo-electron microscopy that can be utilized to solve the structure of the protein at concentration as low as 1 mg/ml.

The studies with the *MtRel* NTD have provided novel insights which can be utilised for the structure aided drug discovery by exploiting the atomic details of *Mtb* specific Rel protein. The structure with its unique features will serve as a template to design effective inhibitors of the *MtRel*, either by targeting the interface of the hydrolase and synthetase or by inhibiting the catalytic activity of synthetase/hydrolase. The current results also give the direction of solving the structure in the presence of substrate in order to map the involvement of identified residues showing diverse functional group in the structurally conserved fold of synthetase domain and C terminal helix of synthetase domain. The differences highlighted in the synthetase domain active site for the binding of ATP/GTP can be utilized as epitopes for pharmacophore modelling.

In addition, the studies with the *MtRel* ACT in the presence of RNA/amino acids gave a new direction to *MtRel* regulation. It will be interesting to perform FRET studies with the cysteine residues in the presence and absence of valine in *MtRel* which might help in validating the model proposed by Bag et al where it has been highlighted that the C-terminal cysteine interacts with the synthetase domain [196]. As described in this thesis, valine binding lead to drastic structural rearrangement, it will not be incorrect to propose that allosteric effects of this binding might affect the interactions with the catalytic domains. On the other hand, the structure of the TGS domain will aid in finding new target epitopes for the drug discovery and thereby affecting the synthetase activity which is shown to increase in the presence of deacylated tRNA binding. Overall, identification of novel epitopes in the *MtRel* will aid in structure based drug designing and further studies targeting residues identified in the dimerization of NTD and ACT will help in understanding the allosteric mechanism of oligomerization and valine binding in *MtRel* regulation.

In summary, the pleotropic changes happen as a result of (p)ppGpp accumulation which are regulated by the Rel protein in *Mtb*. The long-term survival of

this pathogenic bacteria can be addressed by targeting the synthesis/hydrolysis of (p)ppGpp by interfering with the ability of protein to sense the environmental signals. The structural and mechanistic insights into the species specific features of the *MtRel* provides an opportunity to explore various other epitopes and mechanisms which have not been targeted so far. The figure 5.1 shows the novel insights gained into this thesis.

Figure 5.1: Overall insights into *MtRel*. The reduced *MtRel* protein forms putative head to tail dimer in solution as the favorable oligomeric state (refer to Fig. 3.4 and 3.5). Likewise, *MtRel* NTD and *MtRel* TGS domains form homo-dimers in solution. *MtRel* NTD dimer is formed by the synthetase domain whereas the ACT dimer is formed in head to tail fashion with the valine binding site at the dimeric interface. *MtRel* TGS form monomer in solution and is crucial for *MtRel* binding to tRNA. The atomic resolution structure of all the structured domains of *MtRel* provides primary information about the possible binding epitopes for the therapeutics development.

6 References

1. Koch, R., *Die atiology der tuberkulose*. Berliner Klinische Wochenschrift, 1882. **15**: p. 221-30.
2. (WHO), W.H.O., *Global tuberculosis report 2019*. 2019: Geneva: World Health Organization. p. 1-297.
3. Cohn, D.L., et al., *A 62-Dose, 6-Month Therapy for Pulmonary and Extrapulmonary Tuberculosis: A Twice-Weekly, Directly Observed, and Cost-Effective Regimen*. Annals of Internal Medicine, 1990. **112**(6): p. 407-415.
4. Wood, T.K., S.J. Knabel, and B.W. Kwan, *Bacterial persister cell formation and dormancy*. Appl Environ Microbiol, 2013. **79**(23): p. 7116-21.
5. Tiberi, S., A. Zumla, and G.B. Migliori, *Multidrug and Extensively Drug-resistant Tuberculosis: Epidemiology, Clinical Features, Management and Treatment*. Infect Dis Clin North Am, 2019. **33**(4): p. 1063-1085.
6. Günther, G., et al., *Treatment Outcomes in Multidrug-Resistant Tuberculosis*. N Engl J Med, 2016. **375**(11): p. 1103-5.
7. Kiazzyk, S. and T.B. Ball, *Latent tuberculosis infection: An overview*. Can Commun Dis Rep, 2017. **43**(3-4): p. 62-66.
8. Syal, K., *BCG vaccine fails to prevent the reactivation of latent tuberculosis*. Med Hypotheses, 2016. **88**: p. 82.
9. Gengenbacher, M. and S.H. Kaufmann, *Mycobacterium tuberculosis: success through dormancy*. FEMS Microbiol Rev, 2012. **36**(3): p. 514-32.
10. Organization, W.H., *WHO consolidated guidelines on tuberculosis: tuberculosis preventive treatment: module 1: prevention: tuberculosis preventive treatment*. 2020.
11. Rustomjee, R., et al., *A Phase II study of the sterilising activities of ofloxacin, gatifloxacin and moxifloxacin in pulmonary tuberculosis*. Int J Tuberc Lung Dis, 2008. **12**(2): p. 128-38.
12. Cholo, M.C., et al., *Clofazimine: current status and future prospects*. J Antimicrob Chemother, 2012. **67**(2): p. 290-8.
13. Singh, D.K., et al., *Blockade of the Kv1.3 K⁺ Channel Enhances BCG Vaccine Efficacy by Expanding Central Memory T Lymphocytes*. J Infect Dis, 2016. **214**(9): p. 1456-1464.
14. Recht, M.I., S. Douthwaite, and J.D. Puglisi, *Basis for prokaryotic specificity of action of aminoglycoside antibiotics*. EMBO J, 1999. **18**(11): p. 3133-8.
15. Dookie, N., et al., *Evolution of drug resistance in Mycobacterium tuberculosis: a review on the molecular determinants of resistance and implications for personalized care*. J Antimicrob Chemother, 2018. **73**(5): p. 1138-1151.
16. Ford, C.W., et al., *Oxazolidinones: new antibacterial agents*. Trends Microbiol, 1997. **5**(5): p. 196-200.
17. Organization, W.H., *Rapid communication: key changes to the treatment of drug-resistant tuberculosis*. World Health Organization: Geneva, Switzerland, 2019.
18. Ahmad, N., et al., *Treatment correlates of successful outcomes in pulmonary multidrug-resistant tuberculosis: an individual patient data meta-analysis*. Lancet, 2018. **392**(10150): p. 821-834.
19. Conradie, F., et al., *Bedaquiline, pretomanid and linezolid for treatment of extensively drug resistant, intolerant or non-responsive multidrug resistant pulmonary tuberculosis*. The New England Journal of Medicine. **382**(10): p. 893.

20. de Miranda Silva, C., et al., *Effect of Linezolid plus Bedaquiline against Mycobacterium tuberculosis in Log Phase, Acid Phase, and Nonreplicating-Persister Phase in an In Vitro Assay*. Antimicrobial Agents and Chemotherapy, 2018. **62**(8): p. e00856-18.
21. Schulte, B., et al., *Emergence of increasing linezolid-resistance in enterococci in a post-outbreak situation with vancomycin-resistant Enterococcus faecium*. Epidemiol Infect, 2008. **136**(8): p. 1131-3.
22. Nguyen, T.V.A., et al., *Bedaquiline resistance: its emergence, mechanism, and prevention*. Clinical Infectious Diseases, 2018. **66**(10): p. 1625-1630.
23. Ibrahim, M., et al., *Synergistic activity of R207910 combined with pyrazinamide against murine tuberculosis*. Antimicrob Agents Chemother, 2007. **51**(3): p. 1011-5.
24. Somoskovi, A., et al., *A mutation associated with clofazimine and bedaquiline cross-resistance in MDR-TB following bedaquiline treatment*. Eur Respir J, 2015. **45**(2): p. 554-7.
25. Hartkoorn, R.C., S. Uplekar, and S.T. Cole, *Cross-resistance between clofazimine and bedaquiline through upregulation of MmpL5 in Mycobacterium tuberculosis*. Antimicrob Agents Chemother, 2014. **58**(5): p. 2979-81.
26. Andries, K., et al., *Acquired resistance of Mycobacterium tuberculosis to bedaquiline*. PLoS One, 2014. **9**(7): p. e102135.
27. Gupta, A., et al., *Isoniazid Preventive Therapy in HIV-Infected Pregnant and Postpartum Women*. N Engl J Med, 2019. **381**(14): p. 1333-1346.
28. Chiaradia, L., et al., *Dissecting the mycobacterial cell envelope and defining the composition of the native mycomembrane*. Sci Rep, 2017. **7**(1): p. 12807.
29. Kumar, A., et al., *Redox homeostasis in mycobacteria: the key to tuberculosis control?* Expert reviews in molecular medicine, 2011. **13**.
30. Aston, C., et al., *Early inhibition of mycobacterial growth by human alveolar macrophages is not due to nitric oxide*. Am J Respir Crit Care Med, 1998. **157**(6 Pt 1): p. 1943-50.
31. Schlesinger, L.S., S.R. Hull, and T.M. Kaufman, *Binding of the terminal mannosyl units of lipoarabinomannan from a virulent strain of Mycobacterium tuberculosis to human macrophages*. J Immunol, 1994. **152**(8): p. 4070-9.
32. Schlesinger, L.S., et al., *Phagocytosis of Mycobacterium tuberculosis is mediated by human monocyte complement receptors and complement component C3*. The Journal of Immunology, 1990. **144**(7): p. 2771.
33. Czop, J.K. and J. Kay, *Isolation and characterization of beta-glucan receptors on human mononuclear phagocytes*. The Journal of experimental medicine, 1991. **173**(6): p. 1511-1520.
34. Gaynor, C.D., et al., *Pulmonary surfactant protein A mediates enhanced phagocytosis of Mycobacterium tuberculosis by a direct interaction with human macrophages*. The Journal of Immunology, 1995. **155**(11): p. 5343.
35. Nicholson, S., et al., *Inducible nitric oxide synthase in pulmonary alveolar macrophages from patients with tuberculosis*. The Journal of experimental medicine, 1996. **183**(5): p. 2293-2302.
36. Kuo, H., et al., *Increased production of hydrogen peroxide and expression of CD11b/CD18 on alveolar macrophages in patients with active pulmonary tuberculosis*. Tubercle and lung disease, 1996. **77**(5): p. 468-475.
37. Rohde, K., et al., *Mycobacterium tuberculosis and the environment within the phagosome*. Immunological reviews, 2007. **219**(1): p. 37-54.

38. Ma, J., et al., *Tuberculosis antigen-induced expression of IFN- α in tuberculosis patients inhibits production of IL-1 β* . The FASEB Journal, 2014. **28**(7): p. 3238-3248.
39. Queval, C.J., et al., *Mycobacterium tuberculosis Controls Phagosomal Acidification by Targeting CISH-Mediated Signaling*. Cell Rep, 2017. **20**(13): p. 3188-3198.
40. Chen, Z., et al., *Inhibition of Autophagy by MiR-30A Induced by Mycobacteria tuberculosis as a Possible Mechanism of Immune Escape in Human Macrophages*. Jpn J Infect Dis, 2015. **68**(5): p. 420-4.
41. MacGurn, J.A. and J.S. Cox, *A genetic screen for Mycobacterium tuberculosis mutants defective for phagosome maturation arrest identifies components of the ESX-1 secretion system*. Infection and immunity, 2007. **75**(6): p. 2668-2678.
42. Ramakrishnan, L., *Revisiting the role of the granuloma in tuberculosis*. Nature Reviews Immunology, 2012. **12**(5): p. 352.
43. Augenstreich, J., et al., *ESX-1 and phthiocerol dimycocerosates of Mycobacterium tuberculosis act in concert to cause phagosomal rupture and host cell apoptosis*. Cell Microbiol, 2017. **19**(7).
44. Feng, C.G., et al., *NK cell-derived IFN-gamma differentially regulates innate resistance and neutrophil response in T cell-deficient hosts infected with Mycobacterium tuberculosis*. J Immunol, 2006. **177**(10): p. 7086-93.
45. Lockhart, E., A.M. Green, and J.L. Flynn, *IL-17 production is dominated by gammadelta T cells rather than CD4 T cells during Mycobacterium tuberculosis infection*. J Immunol, 2006. **177**(7): p. 4662-9.
46. Chakravarty, S.D., et al., *Tumor necrosis factor blockade in chronic murine tuberculosis enhances granulomatous inflammation and disorganizes granulomas in the lungs*. Infect Immun, 2008. **76**(3): p. 916-26.
47. Cardona, P.J., et al., *Evolution of granulomas in lungs of mice infected aerogenically with Mycobacterium tuberculosis*. Scand J Immunol, 2000. **52**(2): p. 156-63.
48. Driver, E.R., et al., *Evaluation of a mouse model of necrotic granuloma formation using C3HeB/FeJ mice for testing of drugs against Mycobacterium tuberculosis*. Antimicrob Agents Chemother, 2012. **56**(6): p. 3181-95.
49. Pichugin, A.V., et al., *Dominant role of the sst1 locus in pathogenesis of necrotizing lung granulomas during chronic tuberculosis infection and reactivation in genetically resistant hosts*. Am J Pathol, 2009. **174**(6): p. 2190-201.
50. Gil, O., et al., *Granuloma encapsulation is a key factor for containing tuberculosis infection in minipigs*. PLoS One, 2010. **5**(4): p. e10030.
51. Capuano, S.V., et al., *Experimental Mycobacterium tuberculosis infection of cynomolgus macaques closely resembles the various manifestations of human M. tuberculosis infection*. Infect Immun, 2003. **71**(10): p. 5831-44.
52. Flynn, J.L., et al., *Non-human primates: a model for tuberculosis research*. Tuberculosis (Edinb), 2003. **83**(1-3): p. 116-8.
53. MCCUNE, R.M. and R. TOMPSETT, *Fate of Mycobacterium tuberculosis in mouse tissues as determined by the microbial enumeration technique. I. The persistence of drug-susceptible tubercle bacilli in the tissues despite prolonged antimicrobial therapy*. J Exp Med, 1956. **104**(5): p. 737-62.
54. MCCUNE, R.M., W. MCDERMOTT, and R. TOMPSETT, *The fate of Mycobacterium tuberculosis in mouse tissues as determined by the microbial*

- enumeration technique. II. The conversion of tuberculous infection to the latent state by the administration of pyrazinamide and a companion drug. *J Exp Med*, 1956. **104**(5): p. 763-802.
55. Rustad, T.R., et al., *The enduring hypoxic response of Mycobacterium tuberculosis*. *PloS one*, 2008. **3**(1): p. e1502-e1502.
56. Stallings, C.L. and M.S. Glickman, *Is Mycobacterium tuberculosis stressed out? A critical assessment of the genetic evidence*. *Microbes Infect*, 2010. **12**(14-15): p. 1091-101.
57. Vandal, O.H., et al., *A membrane protein preserves intrabacterial pH in intraphagosomal Mycobacterium tuberculosis*. *Nat Med*, 2008. **14**(8): p. 849-54.
58. Vandal, O.H., et al., *Acid-susceptible mutants of Mycobacterium tuberculosis share hypersusceptibility to cell wall and oxidative stress and to the host environment*. *J Bacteriol*, 2009. **191**(2): p. 625-31.
59. Ng, V.H., et al., *Role of KatG catalase-peroxidase in mycobacterial pathogenesis: countering the phagocyte oxidative burst*. *Mol Microbiol*, 2004. **52**(5): p. 1291-302.
60. Dahl, J.L., et al., *The role of RelMtb-mediated adaptation to stationary phase in long-term persistence of Mycobacterium tuberculosis in mice*. *Proc Natl Acad Sci U S A*, 2003. **100**(17): p. 10026-31.
61. Stewart, G.R., et al., *Dissection of the heat-shock response in Mycobacterium tuberculosis using mutants and microarrays*. *Microbiology*, 2002. **148**(Pt 10): p. 3129-3138.
62. Ehlers, S. and U.E. Schaible, *The granuloma in tuberculosis: dynamics of a host-pathogen collusion*. *Front Immunol*, 2012. **3**: p. 411.
63. Brauner, A., et al., *Distinguishing between resistance, tolerance and persistence to antibiotic treatment*. *Nat Rev Microbiol*, 2016. **14**(5): p. 320-30.
64. Eldholm, V. and F. Balloux, *Antimicrobial Resistance in Mycobacterium tuberculosis: The Odd One Out*. *Trends Microbiol*, 2016. **24**(8): p. 637-648.
65. Dutta, N.K. and P.C. Karakousis, *Latent tuberculosis infection: myths, models, and molecular mechanisms*. *Microbiol Mol Biol Rev*, 2014. **78**(3): p. 343-71.
66. Gagneux, S., et al., *Impact of bacterial genetics on the transmission of isoniazid-resistant Mycobacterium tuberculosis*. *PLoS Pathog*, 2006. **2**(6): p. e61.
67. Wang, F., et al., *Evolution and transmission patterns of extensively drug-resistant tuberculosis in China*. *Antimicrob Agents Chemother*, 2015. **59**(2): p. 818-25.
68. Hu, Y., et al., *Acquisition of second-line drug resistance and extensive drug resistance during recent transmission of Mycobacterium tuberculosis in rural China*. *Clin Microbiol Infect*, 2015. **21**(12): p. 1093.e9-1093.e18.
69. Gagneux, S., *Fitness cost of drug resistance in Mycobacterium tuberculosis*. *Clin Microbiol Infect*, 2009. **15 Suppl 1**: p. 66-8.
70. Gandhi, N.R., et al., *Extensively drug-resistant tuberculosis as a cause of death in patients co-infected with tuberculosis and HIV in a rural area of South Africa*. *Lancet*, 2006. **368**(9547): p. 1575-80.
71. Mitchison, D.A., *The action of antituberculosis drugs in short-course chemotherapy*. *Tubercle*, 1985. **66**(3): p. 219-25.
72. Comstock, G.W., V.T. Livesay, and S.F. Woolpert, *The prognosis of a positive tuberculin reaction in childhood and adolescence*. *Am J Epidemiol*, 1974. **99**(2): p. 131-8.

73. Prevention, C.f.D.C.a., *The difference between latent TB infection and TB disease*. 2012.
74. Primm, T.P., et al., *The stringent response of Mycobacterium tuberculosis is required for long-term survival*. J Bacteriol, 2000. **182**(17): p. 4889-98.
75. Cashel, M. and B. Kalbacher, *The control of ribonucleic acid synthesis in Escherichia coli. V. Characterization of a nucleotide associated with the stringent response*. J Biol Chem, 1970. **245**(9): p. 2309-18.
76. SANDS, M.K. and R.B. ROBERTS, *The effects of a tryptophan-histidine deficiency in a mutant of Escherichia coli*. J Bacteriol, 1952. **63**(4): p. 505-11.
77. Cashel, M. and J. Gallant, *Two compounds implicated in the function of the RC gene of Escherichia coli*. Nature, 1969. **221**(5183): p. 838-41.
78. El-Gebali, S., et al., *The Pfam protein families database in 2019*. Nucleic Acids Res, 2019. **47**(D1): p. D427-D432.
79. Choi, M.Y., et al., *The two PPX-GppA homologues from Mycobacterium tuberculosis have distinct biochemical activities*. PLoS One, 2012. **7**(8): p. e42561.
80. Mechold, U., et al., *Differential regulation by ppGpp versus pppGpp in Escherichia coli*. Nucleic Acids Res, 2013. **41**(12): p. 6175-89.
81. Fiil, N.P., K. von Meyenburg, and J.D. Friesen, *Accumulation and turnover of guanosine tetraphosphate in Escherichia coli*. J Mol Biol, 1972. **71**(3): p. 769-83.
82. Paul, B.J., M.B. Berkmen, and R.L. Gourse, *DksA potentiates direct activation of amino acid promoters by ppGpp*. Proc Natl Acad Sci U S A, 2005. **102**(22): p. 7823-8.
83. Liu, K., A.N. Bittner, and J.D. Wang, *Diversity in (p)ppGpp metabolism and effectors*. Curr Opin Microbiol, 2015. **24**: p. 72-9.
84. Barker, M.M., et al., *Mechanism of regulation of transcription initiation by ppGpp. I. Effects of ppGpp on transcription initiation in vivo and in vitro*. J Mol Biol, 2001. **305**(4): p. 673-88.
85. Haugen, S.P., et al., *rRNA promoter regulation by nonoptimal binding of sigma region 1.2: an additional recognition element for RNA polymerase*. Cell, 2006. **125**(6): p. 1069-82.
86. Magnusson, L.U., A. Farewell, and T. Nystrom, *ppGpp: a global regulator in Escherichia coli*. Trends Microbiol, 2005. **13**(5): p. 236-42.
87. Artsimovitch, I., et al., *Structural basis for transcription regulation by alarmone ppGpp*. Cell, 2004. **117**(3): p. 299-310.
88. Vrentas, C.E., et al., *Still looking for the magic spot: the crystallographically defined binding site for ppGpp on RNA polymerase is unlikely to be responsible for rRNA transcription regulation*. J Mol Biol, 2008. **377**(2): p. 551-64.
89. Vrentas, C.E., et al., *Response of RNA polymerase to ppGpp: requirement for the omega subunit and relief of this requirement by DksA*. Genes Dev, 2005. **19**(19): p. 2378-87.
90. Zuo, Y., Y. Wang, and T.A. Steitz, *The mechanism of E. coli RNA polymerase regulation by ppGpp is suggested by the structure of their complex*. Mol Cell, 2013. **50**(3): p. 430-6.
91. Ross, W., et al., *The magic spot: a ppGpp binding site on E. coli RNA polymerase responsible for regulation of transcription initiation*. Mol Cell, 2013. **50**(3): p. 420-9.

92. Chatterji, D., et al., *The role of the omega subunit of RNA polymerase in expression of the relA gene in Escherichia coli*. FEMS Microbiology Letters, 2007. **267**(1): p. 51-55.
93. Paul, B.J., et al., *DksA: a critical component of the transcription initiation machinery that potentiates the regulation of rRNA promoters by ppGpp and the initiating NTP*. Cell, 2004. **118**(3): p. 311-22.
94. Perederina, A., et al., *Regulation through the secondary channel--structural framework for ppGpp-DksA synergism during transcription*. Cell, 2004. **118**(3): p. 297-309.
95. Ross, W., et al., *ppGpp Binding to a Site at the RNAP-DksA Interface Accounts for Its Dramatic Effects on Transcription Initiation during the Stringent Response*. Mol Cell, 2016. **62**(6): p. 811-23.
96. Kriel, A., et al., *Direct regulation of GTP homeostasis by (p)ppGpp: a critical component of viability and stress resistance*. Mol Cell, 2012. **48**(2): p. 231-41.
97. Liu, K., et al., *Molecular mechanism and evolution of guanylate kinase regulation by (p)ppGpp*. Mol Cell, 2015. **57**(4): p. 735-49.
98. Prusa, J., D.X. Zhu, and C.L. Stallings, *The stringent response and Mycobacterium tuberculosis pathogenesis*. Pathogens and Disease, 2018. **76**(5).
99. Levine, A., et al., *The stringent response blocks DNA replication outside the ori region in Bacillus subtilis and at the origin in Escherichia coli*. J Mol Biol, 1991. **219**(4): p. 605-13.
100. Wegrzyn, G., *Replication of plasmids during bacterial response to amino acid starvation*. Plasmid, 1999. **41**(1): p. 1-16.
101. Wegrzyn, G. and A. Wegrzyn, *Stress responses and replication of plasmids in bacterial cells*. Microb Cell Fact, 2002. **1**(1): p. 2.
102. Zyskind, J.W. and D.W. Smith, *DNA replication, the bacterial cell cycle, and cell growth*. Cell, 1992. **69**(1): p. 5-8.
103. Wang, J.D., G.M. Sanders, and A.D. Grossman, *Nutritional control of elongation of DNA replication by (p)ppGpp*. Cell, 2007. **128**(5): p. 865-75.
104. Maciag, M., et al., *ppGpp inhibits the activity of Escherichia coli DnaG primase*. Plasmid, 2010. **63**(1): p. 61-7.
105. Cashel, M., et al., *The stringent response*. Escherichia coli and Salmonella typhimulium: cellular and molecular biology, 1996. **2**: p. 1458-1496.
106. Milon, P., et al., *The nucleotide-binding site of bacterial translation initiation factor 2 (IF2) as a metabolic sensor*. Proc Natl Acad Sci U S A, 2006. **103**(38): p. 13962-7.
107. Yoshida, M., A. Travers, and B.F. Clark, *Inhibition of translation initiation complex formation by MS1*. FEBS Lett, 1972. **23**(2): p. 163-6.
108. Arai, K., et al., *Interaction of guanosine 5'-diphosphate, 2'-(or 3'-) diphosphate(ppGpp) with elongation factors from E. coli*. Biochem Biophys Res Commun, 1972. **48**(1): p. 191-6.
109. Wagner, E.G., M. Ehrenberg, and C.G. Kurland, *Kinetic suppression of translational errors by (p)ppGpp*. Mol Gen Genet, 1982. **185**(2): p. 269-74.
110. Rojas, A.M., et al., *ppGpp inhibition of elongation factors Tu, G and Ts during polypeptide synthesis*. Mol Gen Genet, 1984. **197**(1): p. 36-45.
111. Pingoud, A., et al., *The elongation factor Tu from Escherichia coli, aminoacyl-tRNA, and guanosine tetraphosphate form a ternary complex which is bound by programmed ribosomes*. J Biol Chem, 1983. **258**(23): p. 14200-5.

112. Dix, D.B., et al., *Effect of ppGpp on the accuracy of protein biosynthesis*. Arch Biochem Biophys, 1983. **223**(1): p. 319-24.
113. Mitkevich, V.A., et al., *Thermodynamic characterization of ppGpp binding to EF-G or IF2 and of initiator tRNA binding to free IF2 in the presence of GDP, GTP, or ppGpp*. J Mol Biol, 2010. **402**(5): p. 838-46.
114. Dalebroux, Z.D., et al., *ppGpp conjures bacterial virulence*. Microbiol Mol Biol Rev, 2010. **74**(2): p. 171-99.
115. Lemos, J.A., T.A. Brown, and R.A. Burne, *Effects of RelA on key virulence properties of planktonic and biofilm populations of Streptococcus mutans*. Infect Immun, 2004. **72**(3): p. 1431-40.
116. Mellies, J.L., A.M. Barron, and A.M. Carmona, *Enteropathogenic and enterohemorrhagic Escherichia coli virulence gene regulation*. Infect Immun, 2007. **75**(9): p. 4199-210.
117. Nakanishi, N., et al., *ppGpp with DksA controls gene expression in the locus of enterocyte effacement (LEE) pathogenicity island of enterohaemorrhagic Escherichia coli through activation of two virulence regulatory genes*. Mol Microbiol, 2006. **61**(1): p. 194-205.
118. Taylor, C.M., et al., *Listeria monocytogenes relA and hpt mutants are impaired in surface-attached growth and virulence*. J Bacteriol, 2002. **184**(3): p. 621-8.
119. Aberg, A., et al., *Similar and divergent effects of ppGpp and DksA deficiencies on transcription in Escherichia coli*. J Bacteriol, 2009. **191**(10): p. 3226-36.
120. Sun, D., et al., *A metazoan ortholog of SpoT hydrolyzes ppGpp and functions in starvation responses*. Nat Struct Mol Biol, 2010. **17**(10): p. 1188-94.
121. Atkinson, G.C., T. Tenson, and V. Haurlyuk, *The RelA/SpoT homolog (RSH) superfamily: distribution and functional evolution of ppGpp synthetases and hydrolases across the tree of life*. PLoS One, 2011. **6**(8): p. e23479.
122. Mittenhuber, G., *Comparative genomics and evolution of genes encoding bacterial (p)ppGpp synthetases/hydrolases (the Rel, RelA and SpoT proteins)*. J Mol Microbiol Biotechnol, 2001. **3**(4): p. 585-600.
123. Das, B., et al., *Stringent response in Vibrio cholerae: genetic analysis of spoT gene function and identification of a novel (p)ppGpp synthetase gene*. Mol Microbiol, 2009. **72**(2): p. 380-98.
124. Nanamiya, H., et al., *Identification and functional analysis of novel (p)ppGpp synthetase genes in Bacillus subtilis*. Mol Microbiol, 2008. **67**(2): p. 291-304.
125. Lemos, J.A., et al., *Three gene products govern (p)ppGpp production by Streptococcus mutans*. Mol Microbiol, 2007. **65**(6): p. 1568-81.
126. Geiger, T., et al., *Two small (p)ppGpp synthases in Staphylococcus aureus mediate tolerance against cell envelope stress conditions*. J Bacteriol, 2014. **196**(4): p. 894-902.
127. Das, B. and R.K. Bhadra, *Molecular characterization of vibrio cholerae DeltarelA DeltaspoT double mutants*. Arch Microbiol, 2008. **189**(3): p. 227-38.
128. Haseltine, W.A. and R. Block, *Synthesis of Guanosine Tetra- and Pentaphosphate Requires the Presence of a Codon-Specific, Uncharged Transfer Ribonucleic Acid in the Acceptor Site of Ribosomes*. Proceedings of the National Academy of Sciences, 1973. **70**(5): p. 1564-1568.
129. Haseltine, W.A., et al., *MSI and MSII made on ribosome in idling step of protein synthesis*. Nature, 1972. **238**(5364): p. 381-4.

130. Gentry, D.R. and M. Cashel, *Mutational analysis of the Escherichia coli spoT gene identifies distinct but overlapping regions involved in ppGpp synthesis and degradation*. Mol Microbiol, 1996. **19**(6): p. 1373-84.
131. Wendrich, T.M., et al., *Dissection of the mechanism for the stringent factor RelA*. Mol Cell, 2002. **10**(4): p. 779-88.
132. Arenz, S., et al., *The stringent factor RelA adopts an open conformation on the ribosome to stimulate ppGpp synthesis*. Nucleic Acids Research, 2016. **44**(13): p. 6471-6481.
133. Hauryliuk, V., et al., *Recent functional insights into the role of (p)ppGpp in bacterial physiology*. Nat Rev Micro, 2015. **13**(5): p. 298-309.
134. Shyp, V., et al., *Positive allosteric feedback regulation of the stringent response enzyme RelA by its product*. EMBO Rep, 2012. **13**(9): p. 835-9.
135. English, B.P., et al., *Single-molecule investigations of the stringent response machinery in living bacterial cells*. Proc Natl Acad Sci U S A, 2011. **108**(31): p. E365-73.
136. Winther, K.S., M. Roghanian, and K. Gerdes, *Activation of the Stringent Response by Loading of RelA-tRNA Complexes at the Ribosomal A-Site*. Mol Cell, 2018. **70**(1): p. 95-105 e4.
137. Winther, K.S., M. Roghanian, and K. Gerdes, *Activation of the Stringent Response by Loading of RelA-tRNA Complexes at the Ribosomal A-Site*. Mol Cell, 2018. **70**(1): p. 95-105.e4.
138. Loveland, A.B., et al., *Ribosome*RelA structures reveal the mechanism of stringent response activation*. Elife, 2016. **5**.
139. Justesen, J., et al., *The physiology of stringent factor (ATP:GTP 3'-diphosphotransferase) in Escherichia coli*. Biochimie, 1986. **68**(5): p. 715-22.
140. Metzger, S., et al., *The nucleotide sequence and characterization of the relA gene of Escherichia coli*. J Biol Chem, 1988. **263**(30): p. 15699-704.
141. Nakagawa, A., T. Oshima, and H. Mori, *Identification and characterization of a second, inducible promoter of relA in Escherichia coli*. Genes Genet Syst, 2006. **81**(5): p. 299-310.
142. Brown, D.R., et al., *Nitrogen stress response and stringent response are coupled in Escherichia coli*. Nat Commun, 2014. **5**: p. 4115.
143. Reitzer, L., *Nitrogen assimilation and global regulation in Escherichia coli*. Annu Rev Microbiol, 2003. **57**: p. 155-76.
144. Metzger, S., et al., *Characterization of the relA1 mutation and a comparison of relA1 with new relA null alleles in Escherichia coli*. J Biol Chem, 1989. **264**(35): p. 21146-52.
145. Wassarman, K.M. and G. Storz, *6S RNA regulates E. coli RNA polymerase activity*. Cell, 2000. **101**(6): p. 613-23.
146. Sajish, M., et al., *The significance of EXDD and RXKD motif conservation in Rel proteins*. J Biol Chem, 2009. **284**(14): p. 9115-23.
147. Sajish, M., et al., *A charge reversal differentiates (p)ppGpp synthesis by monofunctional and bifunctional Rel proteins*. J Biol Chem, 2007. **282**(48): p. 34977-83.
148. Gaca, A.O., et al., *Global transcriptional analysis of the stringent response in Enterococcus faecalis*. Microbiology, 2012. **158**(Pt 8): p. 1994-2004.
149. Samarrai, W., et al., *Differential responses of Bacillus subtilis rRNA promoters to nutritional stress*. J Bacteriol, 2011. **193**(3): p. 723-33.

150. Pedersen, F.S. and N.O. Kjeldgaard, *Analysis of the relA gene product of Escherichia coli*. Eur J Biochem, 1977. **76**(1): p. 91-7.
151. Weyer, W.J., et al., *The sequence of ppGpp and pppGpp in the reaction scheme for magic spot synthesis*. Biochim Biophys Acta, 1976. **442**(1): p. 123-7.
152. Kari, C., I. Torok, and A. Travers, *ppGpp cycle in Escherichia coli*. Mol Gen Genet, 1977. **150**(3): p. 249-55.
153. Hernandez, V.J. and H. Bremer, *Escherichia coli ppGpp synthetase II activity requires spoT*. J Biol Chem, 1991. **266**(9): p. 5991-9.
154. Ginther, C.L. and J.L. Ingraham, *Nucleoside diphosphokinase of Salmonella typhimurium*. J Biol Chem, 1974. **249**(11): p. 3406-11.
155. Spira, B., N. Silberstein, and E. Yagil, *Guanosine 3',5'-bispyrophosphate (ppGpp) synthesis in cells of Escherichia coli starved for Pi*. J Bacteriol, 1995. **177**(14): p. 4053-8.
156. Battesti, A. and E. Bouveret, *Acyl carrier protein/SpoT interaction, the switch linking SpoT-dependent stress response to fatty acid metabolism*. Molecular Microbiology, 2006. **62**(4): p. 1048-1063.
157. Vinella, D., et al., *Iron limitation induces SpoT-dependent accumulation of ppGpp in Escherichia coli*. Molecular Microbiology, 2005. **56**(4): p. 958-970.
158. Battesti, A. and E. Bouveret, *Bacteria possessing two RelA/SpoT-like proteins have evolved a specific stringent response involving the acyl carrier protein-SpoT interaction*. J Bacteriol, 2009. **191**(2): p. 616-24.
159. Jiang, M., et al., *G-protein control of the ribosome-associated stress response protein SpoT*. J Bacteriol, 2007. **189**(17): p. 6140-7.
160. Raskin, D.M., N. Judson, and J.J. Mekalanos, *Regulation of the stringent response is the essential function of the conserved bacterial G protein CgtA in Vibrio cholerae*. Proc Natl Acad Sci U S A, 2007. **104**(11): p. 4636-41.
161. Gaca, A.O., et al., *From (p)ppGpp to (pp)pGpp: Characterization of Regulatory Effects of pGpp Synthesized by the Small Alarmones Synthetase of Enterococcus faecalis*. J Bacteriol, 2015. **197**(18): p. 2908-19.
162. Steinchen, W., et al., *Catalytic mechanism and allosteric regulation of an oligomeric (p)ppGpp synthetase by an alarmones*. Proc Natl Acad Sci U S A, 2015. **112**(43): p. 13348-53.
163. Agirrezabala, X., et al., *The ribosome triggers the stringent response by RelA via a highly distorted tRNA*. EMBO Rep, 2013. **14**(9): p. 811-6.
164. *The stringent factor RelA adopts an open conformation on the ribosome to stimulate ppGpp synthesis*.
165. Brown, A., et al., *Ribosome-dependent activation of stringent control*. Nature, 2016. **534**(7606): p. 277-80.
166. Kudrin, P., et al., *The ribosomal A-site finger is crucial for binding and activation of the stringent factor RelA*. Nucleic Acids Res, 2018. **46**(4): p. 1973-1983.
167. Hogg, T., et al., *Conformational antagonism between opposing active sites in a bifunctional RelA/SpoT homolog modulates (p)ppGpp metabolism during the stringent response [corrected]*. Cell, 2004. **117**(1): p. 57-68.
168. Hauryliuk, V., et al., *Recent functional insights into the role of (p)ppGpp in bacterial physiology*. Nat Rev Microbiol, 2015. **13**(5): p. 298-309.
169. Steinchen, W. and G. Bange, *The magic dance of the alarmones (p)ppGpp*. Mol Microbiol, 2016. **101**(4): p. 531-44.

170. China, A., et al., *Inhibition of Mycobacterium tuberculosis RNA polymerase by binding of a Gre factor homolog to the secondary channel*. J Bacteriol, 2012. **194**(5): p. 1009-17.
171. Tare, P., A. China, and V. Nagaraja, *Distinct and contrasting transcription initiation patterns at Mycobacterium tuberculosis promoters*. PLoS One, 2012. **7**(9): p. e43900.
172. Tare, P., B. Mallick, and V. Nagaraja, *Co-evolution of specific amino acid in sigma 1.2 region and nucleotide base in the discriminator to act as sensors of small molecule effectors of transcription initiation in mycobacteria*. Mol Microbiol, 2013. **90**(3): p. 569-83.
173. Gaca, A.O., et al., *Basal levels of (p)ppGpp in Enterococcus faecalis: the magic beyond the stringent response*. MBio, 2013. **4**(5): p. e00646-13.
174. Klinkenberg, L.G., et al., *The stringent response is required for full virulence of Mycobacterium tuberculosis in guinea pigs*. J Infect Dis, 2010. **202**(9): p. 1397-404.
175. Wu, M.L., C.L. Chan, and T. Dick, *Rel Is Required for Morphogenesis of Resting Cells in Mycobacterium smegmatis*. Frontiers in Microbiology, 2016. **7**.
176. Dozot, M., et al., *The stringent response mediator Rsh is required for Brucella melitensis and Brucella suis virulence, and for expression of the type IV secretion system virB*. Cell Microbiol, 2006. **8**(11): p. 1791-802.
177. Geiger, T., et al., *Role of the (p)ppGpp synthase RSH, a RelA/SpoT homolog, in stringent response and virulence of Staphylococcus aureus*. Infect Immun, 2010. **78**(5): p. 1873-83.
178. Poole, K., *Bacterial stress responses as determinants of antimicrobial resistance*. J Antimicrob Chemother, 2012. **67**(9): p. 2069-89.
179. Geiger, T., et al., *Two small (p)ppGpp synthases in Staphylococcus aureus mediate tolerance against cell envelope stress conditions*. J Bacteriol, 2014. **196**(4): p. 894-902.
180. Wexselblatt, E., et al., *ppGpp analogues inhibit synthetase activity of Rel proteins from Gram-negative and Gram-positive bacteria*. Bioorg Med Chem, 2010. **18**(12): p. 4485-97.
181. Syal, K., et al., *Synthetic (p)ppGpp analogue: Inhibitor of stringent response in mycobacteria*. Antimicrob Agents Chemother, 2017.
182. Syal, K., N. Bhardwaj, and D. Chatterji, *Vitamin C targets (p)ppGpp synthesis leading to stalling of long-term survival and biofilm formation in Mycobacterium smegmatis*. FEMS Microbiol Lett, 2017. **364**(1).
183. Avarbock, D., et al., *Cloning and characterization of a bifunctional RelA/SpoT homologue from Mycobacterium tuberculosis*. Gene, 1999. **233**(1-2): p. 261-9.
184. Avarbock, A., et al., *Functional regulation of the opposing (p)ppGpp synthetase/hydrolase activities of RelMtb from Mycobacterium tuberculosis*. Biochemistry, 2005. **44**(29): p. 9913-23.
185. Ojha, A.K., T.K. Mukherjee, and D. Chatterji, *High intracellular level of guanosine tetraphosphate in Mycobacterium smegmatis changes the morphology of the bacterium*. Infect Immun, 2000. **68**(7): p. 4084-91.
186. Gupta, S. and D. Chatterji, *Stress Responses in Mycobacteria*. IUBMB Life, 2005. **57**(3): p. 149-159.
187. Dahl, J.L., et al., *The relA homolog of Mycobacterium smegmatis affects cell appearance, viability, and gene expression*. J Bacteriol, 2005. **187**(7): p. 2439-47.

188. Lamichhane, G., S. Tyagi, and W.R. Bishai, *Designer arrays for defined mutant analysis to detect genes essential for survival of Mycobacterium tuberculosis in mouse lungs*. Infect Immun, 2005. **73**(4): p. 2533-40.
189. Weiss, L.A. and C.L. Stallings, *Essential roles for Mycobacterium tuberculosis Rel beyond the production of (p)ppGpp*. J Bacteriol, 2013. **195**(24): p. 5629-38.
190. Avarbock, D., A. Avarbock, and H. Rubin, *Differential regulation of opposing RelMtb activities by the aminoacylation state of a tRNA.ribosome.mRNA.RelMtb complex*. Biochemistry, 2000. **39**(38): p. 11640-8.
191. Avarbock, D.H., *Biochemical characterization of Rel(Mtb), a dual-function ATP: GTP 3'-pyrophosphoryltransferase and (p)ppGpp 3'-pyrophosphohydrolase: Implications for Mycobacterium tuberculosis dormancy*. 2000, University of Pennsylvania: Ann Arbor. p. 173.
192. Mechold, U., et al., *Functional analysis of a relA/spoT gene homolog from Streptococcus equisimilis*. J Bacteriol, 1996. **178**(5): p. 1401-11.
193. Mechold, U., et al., *Intramolecular regulation of the opposing (p)ppGpp catalytic activities of Rel(Seq), the Rel/Spo enzyme from Streptococcus equisimilis*. J Bacteriol, 2002. **184**(11): p. 2878-88.
194. Aravind, L. and E.V. Koonin, *The HD domain defines a new superfamily of metal-dependent phosphohydrolases*. Trends Biochem Sci, 1998. **23**(12): p. 469-72.
195. Boehm, A., et al., *Second messenger signalling governs Escherichia coli biofilm induction upon ribosomal stress*. Mol Microbiol, 2009. **72**(6): p. 1500-16.
196. Bag, S., et al., *Mutational analysis of the (p)ppGpp synthetase activity of the Rel enzyme of Mycobacterium tuberculosis*. Arch Microbiol, 2014. **196**(8): p. 575-88.
197. Jain, V., et al., *Identification and characterization of rel promoter element of Mycobacterium tuberculosis*. Gene, 2005. **351**: p. 149-57.
198. Jain, V., R. Saleem-Batcha, and D. Chatterji, *Synthesis and hydrolysis of pppGpp in mycobacteria: a ligand mediated conformational switch in Rel*. Biophys Chem, 2007. **127**(1-2): p. 41-50.
199. Grant, G.A., *The ACT domain: a small molecule binding domain and its role as a common regulatory element*. J Biol Chem, 2006. **281**(45): p. 33825-9.
200. Chipman, D.M. and B. Shaanan, *The ACT domain family*. Curr Opin Struct Biol, 2001. **11**(6): p. 694-700.
201. Ronneau, S., et al., *Regulation of (p)ppGpp hydrolysis by a conserved archetypal regulatory domain*. Nucleic Acids Res, 2019. **47**(2): p. 843-854.
202. Ronneau, S., et al., *Regulation of (p)ppGpp hydrolysis by a conserved archetypal regulatory domain*. Nucleic acids research, 2019. **47**(2): p. 843-854.
203. Boutte, C.C. and S. Crosson, *The complex logic of stringent response regulation in Caulobacter crescentus: starvation signalling in an oligotrophic environment*. Molecular microbiology, 2011. **80**(3): p. 695-714.
204. Karstens, K., et al., *Phosphotransferase protein EIINtr interacts with SpoT, a key enzyme of the stringent response, in Ralstonia eutropha H16*. Microbiology, 2014. **160**(Pt 4): p. 711-22.
205. Fang, M. and C.E. Bauer, *Regulation of stringent factor by branched-chain amino acids*. Proceedings of the National Academy of Sciences, 2018. **115**(25): p. 6446.

206. Hondalus, M.K., et al., *Attenuation of and protection induced by a leucine auxotroph of Mycobacterium tuberculosis*. Infect Immun, 2000. **68**(5): p. 2888-98.
207. Sassetti, C.M., D.H. Boyd, and E.J. Rubin, *Genes required for mycobacterial growth defined by high density mutagenesis*. Mol Microbiol, 2003. **48**(1): p. 77-84.
208. Jain, V., et al., *Molecular dissection of the mycobacterial stringent response protein Rel*. Protein Sci, 2006. **15**(6): p. 1449-64.
209. Gropp, M., et al., *Regulation of Escherichia coli RelA requires oligomerization of the C-terminal domain*. J Bacteriol, 2001. **183**(2): p. 570-9.
210. Grüber, G., et al., *Expression, purification, and characterization of subunit E, an essential subunit of the vacuolar ATPase*. Biochemical and biophysical research communications, 2002. **298**(3): p. 383-391.
211. Emsley, P. and K. Cowtan, *Coot: model-building tools for molecular graphics*. Acta Crystallographica Section D: Biological Crystallography, 2004. **60**(12): p. 2126-2132.
212. Zbyszek, O. and M. Wladek, *Processing of X-ray diffraction data collected in oscillation mode*. Meth. Enzymol., 1997. **276**: p. 307-326.
213. Laskowski, R.A., et al., *PROCHECK: a program to check the stereochemical quality of protein structures*. Journal of applied crystallography, 1993. **26**(2): p. 283-291.
214. Ramachandran, G.N., C. Ramakrishnan, and V. Sasisekharan, *Stereochemistry of polypeptide chain configurations*. J Mol Biol, 1963. **7**: p. 95-9.
215. Murshudov, G.N., A.A. Vagin, and E.J. Dodson, *Refinement of macromolecular structures by the maximum-likelihood method*. Acta Crystallographica Section D: Biological Crystallography, 1997. **53**(3): p. 240-255.
216. Vagin, A. and A. Teplyakov, *Molecular replacement with MOLREP*. Acta Crystallographica Section D: Biological Crystallography, 2010. **66**(1): p. 22-25.
217. McCoy, A.J., et al., *Phaser crystallographic software*. Journal of applied crystallography, 2007. **40**(4): p. 658-674.
218. Brünger, A.T., et al., *Crystallography & NMR system: a new software suite for macromolecular structure determination*. Acta Crystallographica Section D: Biological Crystallography, 1998. **54**(5): p. 905-921.
219. Winn, M.D., et al., *Overview of the CCP4 suite and current developments*. Acta Crystallogr D Biol Crystallogr, 2011. **67**(Pt 4): p. 235-42.
220. Svergun, D., *A direct indirect method of small-angle scattering data treatment*. Journal of applied crystallography, 1993. **26**(2): p. 258-267.
221. Svergun, D.I., M.V. Petoukhov, and M.H. Koch, *Determination of domain structure of proteins from X-ray solution scattering*. Biophys J, 2001. **80**(6): p. 2946-53.
222. Kozin, M.B. and D.I. Svergun, *Automated matching of high-and low-resolution structural models*. Journal of applied crystallography, 2001. **34**(1): p. 33-41.
223. Tamura, K., et al., *MEGA6: Molecular Evolutionary Genetics Analysis version 6.0*. Mol Biol Evol, 2013. **30**(12): p. 2725-9.
224. DeLano, W.L., *Pymol: An open-source molecular graphics tool*. CCP4 Newsletter On Protein Crystallography, 2002. **40**: p. 82-92.
225. Waterhouse, A.M., et al., *Jalview Version 2--a multiple sequence alignment editor and analysis workbench*. Bioinformatics, 2009. **25**(9): p. 1189-91.

226. Pettersen, E.F., et al., *UCSF Chimera—a visualization system for exploratory research and analysis*. Journal of computational chemistry, 2004. **25**(13): p. 1605-1612.
227. Davis, I.W., et al., *MOLPROBITY: structure validation and all-atom contact analysis for nucleic acids and their complexes*. Nucleic acids research, 2004. **32**(suppl 2): p. W615-W619.
228. Singal, B., et al., *Crystallographic and solution structure of the N-terminal domain of the Rel protein from Mycobacterium tuberculosis*. FEBS Lett, 2017. **591**(15): p. 2323-2337.
229. Goyal, M. and T.K. Chaudhuri, *GroEL-GroES assisted folding of multiple recombinant proteins simultaneously over-expressed in Escherichia coli*. Int J Biochem Cell Biol, 2015. **64**: p. 277-86.
230. Laemmli, U.K., *Cleavage of structural proteins during the assembly of the head of bacteriophage T4*. Nature, 1970. **227**(5259): p. 680-5.
231. Wilkinson, R.C., et al., *Biochemical studies on Francisella tularensis RelA in (p)ppGpp biosynthesis*. Biosci Rep, 2015. **35**(6).
232. Dip, P.V., et al., *Key roles of the Escherichia coli AhpC C-terminus in assembly and catalysis of alkylhydroperoxide reductase, an enzyme essential for the alleviation of oxidative stress*. Biochim Biophys Acta, 2014. **1837**(12): p. 1932-43.
233. Krissinel, E. and K. Henrick, *Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions*. Acta Crystallographica Section D: Biological Crystallography, 2004. **60**(12): p. 2256-2268.
234. Krissinel, E. and K. Henrick, *Inference of macromolecular assemblies from crystalline state*. J Mol Biol, 2007. **372**(3): p. 774-97.
235. Bernadó, P., et al., *Structural characterization of flexible proteins using small-angle X-ray scattering*. Journal of the American Chemical Society, 2007. **129**(17): p. 5656-5664.
236. Mertens, H.D. and D.I. Svergun, *Structural characterization of proteins and complexes using small-angle X-ray solution scattering*. Journal of structural biology, 2010. **172**(1): p. 128-141.
237. Blanchet, C.E., et al., *Versatile sample environments and automation for biological solution X-ray scattering experiments at the P12 beamline (PETRA III, DESY)*. J Appl Crystallogr, 2015. **48**(Pt 2): p. 431-443.
238. Balakrishna, A.M., et al., *Crystal structure of subunits D and F in complex gives insight into energy transmission of the eukaryotic V-ATPase from Saccharomyces cerevisiae*. J Biol Chem, 2015. **290**(6): p. 3183-96.
239. Dip, P.V., et al., *Structure, mechanism and ensemble formation of the alkylhydroperoxide reductase subunits AhpC and AhpF from Escherichia coli*. Acta Crystallogr D Biol Crystallogr, 2014. **70**(Pt 11): p. 2848-62.
240. Tay, M.Y., et al., *The C-terminal 50 amino acid residues of dengue NS3 protein are important for NS3-NS5 interaction and viral replication*. Journal of Biological Chemistry, 2015. **290**(4): p. 2379-2394.
241. Konarev, P.V., et al., *ATSAS 2.1, a program package for small-angle scattering data analysis*. Journal of applied crystallography, 2006. **39**(2): p. 277-286.
242. Guinier, A. and G. Fournet, *Small-angle Scattering of X-rays*. Wiley, New York, 1955.
243. Guinier, A., *Diffraction of x-rays of very small angles-application to the study of ultramicroscopic phenomenon*. Ann Phys, 1939. **12**: p. 161-237.

244. Svergun, D., *Determination of the regularization parameter in indirect-transform methods using perceptual criteria*. Journal of applied crystallography, 1992. **25**(4): p. 495-503.
245. Durand, D., et al., *NADPH oxidase activator p67(phox) behaves in solution as a multidomain protein with semi-flexible linkers*. J Struct Biol, 2010. **169**(1): p. 45-53.
246. Mylonas, E. and D.I. Svergun, *Accuracy of molecular mass determination of proteins in solution by small-angle X-ray scattering*. Applied Crystallography, 2007. **40**(s1): p. s245-s249.
247. Petoukhov, M.V., et al., *New developments in the ATSAS program package for small-angle scattering data analysis*. Journal of applied crystallography, 2012. **45**(2): p. 342-350.
248. Rambo, R.P. and J.A. Tainer, *Accurate assessment of mass, models and resolution by small-angle scattering*. Nature, 2013. **496**(7446): p. 477-481.
249. Franke, D. and D.I. Svergun, *DAMMIF, a program for rapid ab-initio shape determination in small-angle scattering*. Journal of applied crystallography, 2009. **42**(2): p. 342-346.
250. Volkov, V.V. and D.I. Svergun, *Uniqueness of ab initio shape determination in small-angle scattering*. Journal of applied crystallography, 2003. **36**(3): p. 860-864.
251. Svergun, D., C. Barberato, and M.H. Koch, *CRY SOL—a program to evaluate X-ray solution scattering of biological macromolecules from atomic coordinates*. Journal of applied crystallography, 1995. **28**(6): p. 768-773.
252. Konarev, P.V., et al., *PRIMUS: a Windows PC-based system for small-angle scattering data analysis*. Journal of applied crystallography, 2003. **36**(5): p. 1277-1282.
253. Crane, E.J., J. Vervoort, and A. Claiborne, *¹³C NMR analysis of the cysteine-sulfenic acid redox center of enterococcal NADH peroxidase*. Biochemistry, 1997. **36**(28): p. 8611-8.
254. Rovnyak, D., et al., *Accelerated acquisition of high resolution triple-resonance spectra using non-uniform sampling and maximum entropy reconstruction*. J Magn Reson, 2004. **170**(1): p. 15-21.
255. Orekhov, V.Y. and V.A. Jaravine, *Analysis of non-uniformly sampled spectra with multi-dimensional decomposition*. Prog Nucl Magn Reson Spectrosc, 2011. **59**(3): p. 271-92.
256. Kazimierczuk, K. and V.Y. Orekhov, *Accelerated NMR spectroscopy by using compressed sensing*. Angew Chem Int Ed Engl, 2011. **50**(24): p. 5556-9.
257. Delaglio, F., et al., *NMRPipe: a multidimensional spectral processing system based on UNIX pipes*. J Biomol NMR, 1995. **6**(3): p. 277-93.
258. Goddard, T.D. and D.G. Kneller, *SPARKY, version 3*. University of California, San Francisco, CA.
259. Muchmore, D.C., et al., *Expression and nitrogen-15 labeling of proteins for proton and nitrogen-15 nuclear magnetic resonance*. Methods Enzymol, 1989. **177**: p. 44-73.
260. Schleucher, J., et al., *A general enhancement scheme in heteronuclear multidimensional NMR employing pulsed field gradients*. J Biomol NMR, 1994. **4**(2): p. 301-6.
261. Kalbitzer, H., *Protein NMR Spectroscopy. Principles and Practice*. Zeitschrift für Physikalische Chemie, 1997. **202**(1-2): p. 297-297.

262. Braunschweiler, L. and R.R. Ernst, *Coherence transfer by isotropic mixing: Application to proton correlation spectroscopy*. Journal of Magnetic Resonance (1969), 1983. **53**(3): p. 521-528.
263. Wüthrich, K., *Protein structure determination in solution by NMR spectroscopy*. J Biol Chem, 1990. **265**(36): p. 22059-62.
264. Receveur-Bréchet, V. and D. Durand, *How random are intrinsically disordered proteins? A small angle scattering perspective*. Current Protein and Peptide Science, 2012. **13**(1): p. 55-75.
265. Robson, S., et al., *Nonuniform Sampling for NMR Spectroscopy*. Methods Enzymol, 2019. **614**: p. 263-291.
266. Aravind, L. and E.V. Koonin, *Gleaning non-trivial structural, functional and evolutionary information about proteins by iterative database searches*. J Mol Biol, 1999. **287**(5): p. 1023-40.
267. Sanger, F. and A.R. Coulson, *A rapid method for determining sequences in DNA by primed synthesis with DNA polymerase*. J Mol Biol, 1975. **94**(3): p. 441-8.
268. Yang, X. and E.E. Ishiguro, *Dimerization of the RelA protein of Escherichia coli*. Biochem Cell Biol, 2001. **79**(6): p. 729-36.
269. Arenz, S., et al., *The stringent factor RelA adopts an open conformation on the ribosome to stimulate ppGpp synthesis*. Nucleic Acids Res, 2016. **44**(13): p. 6471-81.
270. Ekal, L., et al., *Evidence of a conserved intrinsically disordered region in the C-terminus of the stringent response protein Rel from mycobacteria*. FEBS Lett, 2014. **588**(9): p. 1839-49.
271. Liberles, J.S., M. Thórólfsson, and A. Martínez, *Allosteric mechanisms in ACT domain containing enzymes involved in amino acid metabolism*. Amino Acids, 2005. **28**(1): p. 1-12.
272. Syal, K., et al., *Novel pppGpp binding site at the C-terminal region of the Rel enzyme from Mycobacterium smegmatis*. FEBS J, 2015. **282**(19): p. 3773-85.
273. Schuller, D.J., G.A. Grant, and L.J. Banaszak, *The allosteric ligand site in the Vmax-type cooperative enzyme phosphoglycerate dehydrogenase*. Nat Struct Biol, 1995. **2**(1): p. 69-76.
274. Leonard, P.M., et al., *Crystal structure of the Lrp-like transcriptional regulator from the archaeon Pyrococcus furiosus*. EMBO J, 2001. **20**(5): p. 990-7.
275. Arenz, S., et al., *The stringent factor RelA adopts an open conformation on the ribosome to stimulate ppGpp synthesis*. Nucleic Acids Res, 2016.
276. Kaiser, J.C. and D.E. Heinrichs, *Branching Out: Alterations in Bacterial Physiology and Virulence Due to Branched-Chain Amino Acid Deprivation*. mBio, 2018. **9**(5): p. e01188-18.
277. Khoshouei, M., et al., *Cryo-EM structure of haemoglobin at 3.2 Å determined with the Volta phase plate*. bioRxiv, 2016.

List of author's publications:

- Singal, B., Balakrishna, A.M., Nartey, W., Manimekalai, M.S.S., Jeyakanthan, J. and Grüber, G. (2017) Crystallographic and solution structure of the N-terminal domain of the Rel protein from *Mycobacterium tuberculosis*. *FEBS Lett*, 591, 2323-2337.

Conference attendance, speaker invitation and posters:

- Young SBS Scientist Seminar Series (Y4S), NTU, Singapore held on Sep. 25th, 2019. Title: "Structural and mechanistic insights into the Rel protein from *Mycobacterium tuberculosis*" (Oral presentation).
- Tuberculosis Drug Discovery and Development – Gordon Research Conference, Castelldefels, Spain held from Jul. 07th – 12th, 2019. Title: "Structural and mechanistic insights into the Rel protein from *Mycobacterium tuberculosis*" (Poster presentation).
- 1st Annual Symposium, *From crisis to cures: New approaches to TB drug discovery*- TOPNet, Singapore, held from Nov. 19th – 20th, 2018. Title: "Structural and mechanistic insights into the Rel protein from *Mycobacterium tuberculosis*" (Poster presentation).
- EMBO course on Structural and biophysical methods for biological macromolecules in solution, NTU, Singapore, held on Dec. 6th-14th, 2017. Title: "Crystallographic and solution structure of the N-terminal domain of the Rel protein from *Mycobacterium tuberculosis*" (Poster presentation).

Awards:

- Nanyang Technological University Research Scholarship (2016-2020)

7 Appendix

Supplementary Table S1: MS+MS/MS data of the *MtRel* band of lane 2 in Figure 3.2D inset.

Protein name				Accession No.				Protein MW		
Bifunctional (p)ppGpp synthase/hydrolase RelA OS=Mycobacterium tuberculosis (strain ATCC 25618 /H3)				sp P9WHG9 RELA_MYCTU				82117.7		
Calc. Mass	Obsrv. Mass	$\pm$ da	$\pm$ ppm	Start seq.	End seq.	Sequence	ION score	C. I. %	Modification	Rank Result Type
948.5413	948.535	-0.0063	-7	484	490	IRQWFAK				Mascot
958.5104	958.5128	0.0024	3	373	379	QLLDWQR				Mascot
1124.5524	1124.5529	0.0005	0	245	252	HYWSIQK				Mascot
1137.6051	1137.6033	-0.0018	-2	297	305	FKDYIAQPR				Mascot
1213.5847	1213.5831	-0.0016	-1	724	733	NVEGVYDVYR	29	95.67		Mascot
1213.5847	1213.5831	-0.0016	-1	724	733	NVEGVYDVYR				Mascot
1261.6324	1261.6255	-0.0069	-5	470	479	DWQQFVVSPR				Mascot
1261.6324	1261.6255	-0.0069	-5	470	479	DWQQFVVSPR				Mascot
1274.6277	1274.6234	-0.0043	-3	336	346	TAEYGIAAHR	6	0		Mascot
1274.6277	1274.6234	-0.0043	-3	336	346	TAEYGIAAHR				Mascot
1433.6907	1433.6854	-0.0053	-4	380	392	EAADPGEFLESLR				Mascot

1613.7992	1613.7893	-0.0099	-6	700	713	VAISRFTFEMGDPK			Oxidation (M)[10]	Mascot
1632.9067	1632.8962	-0.0105	-6	6	20	STTNPVLEPLVAVHR				Mascot
1669.8835	1669.8582	-0.0253	-15	393	406	YDLAVQEIVFVTPK				Mascot
1692.7606	1692.7869	0.0263	16	631	645	TDCTNAASLQQQAER			Carbamidomethyl (C)[3]	Mascot
1692.7606	1692.7869	0.0263	16	631	645	TDCTNAASLQQQAER			Carbamidomethyl (C)[3]	Mascot
1697.8898	1697.8748	-0.015	-9	186	199	WELEDLSFAILHPK				Mascot
1955.9919	1955.975	-0.0169	-9	557	574	LLAELGGIDQAEELAER	70	100		Mascot
1955.9919	1955.975	-0.0169	-9	557	574	LLAELGGIDQAEELAER				Mascot
2012.0096	2011.991	-0.0186	-9	462	479	APNAGPSRDWQQFVVSPR	28	94.383		Mascot
2012.0096	2011.991	-0.0186	-9	462	479	APNAGPSRDWQQFVVSPR				Mascot
2133.9956	2133.9902	-0.0054	-3	354	372	NGVLHPHAAAEIDDMAW MR				Mascot
2331.1616	2331.1555	-0.0061	-3	530	551	ELHYADVSAITYAIGEGH VSAK				Mascot
2384.2683	2384.2437	-0.0246	-10	179	199	LGMASVKWELEDLSFAIL HPK				Mascot

2670.425	2670.4146	-0.0104	-4	306	329	YGVYQSLHTTVVGPEGKP LEVQIR	45	99.902	Mascot
2670.425	2670.4146	-0.0104	-4	306	329	YGVYQSLHTTVVGPEGKP LEVQIR			Mascot
2739.3735	2739.3623	-0.0112	-4	407	432	GDVITLPTGSTPVDFAYA VHTEVGHR	53	99.985	Mascot
2739.3735	2739.3623	-0.0112	-4	407	432	GDVITLPTGSTPVDFAYA VHTEVGHR			Mascot
3084.5564	3084.5332	-0.0232	-8	380	406	EAADPGEFLESLRYDLAV QEIFVFTP			Mascot
3084.5564	3084.5332	-0.0232	-8	380	406	EAADPGEFLESLRYDLAV QEIFVFTP			Mascot