

**PROTEIN PHASE SEPARATION BY THE ABC TRANSPORTER Rv1747 OF
MYCOBACTERIUM TUBERCULOSIS**

by

Florian Heinkel

B.Sc., Ludwig-Maximilians-Universität München, 2012

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

in

THE FACULTY OF GRADUATE AND POSTDOCTORAL STUDIES
(Genome Science and Technology)

THE UNIVERSITY OF BRITISH COLUMBIA

(Vancouver)

April 2018

© Florian Heinkel, 2018

The following individuals certify that they have read, and recommend to the Faculty of Graduate and Postdoctoral Studies for acceptance, the dissertation entitled:

Protein Phase Separation by the ABC Transporter Rv1747 of Mycobacterium Tuberculosis

submitted by Florian Heinkel in partial fulfillment of the requirements for

the degree of Doctor of Philosophy

in Genome Science and Technology

Examining Committee:

Lawrence McIntosh
Co-supervisor

Jörg Gsponer
Co-supervisor

Natalie Strynadka
Supervisory Committee Member

Steven Plotkin
University Examiner

Charles Thompson
University Examiner

Additional Supervisory Committee Members:

Franck Duong
Supervisory Committee Member

Abstract

The *Mycobacterium tuberculosis* ABC transporter Rv1747 is a putative exporter of cell wall biosynthesis intermediates. Rv1747 has a cytoplasmic regulatory module consisting of two pThr-interacting Forkhead-associated (FHA) domains connected by a conformationally disordered linker with two phospho-acceptor threonines (pThr). In chapter 2, I report the structures of FHA-1 and FHA-2 determined by X-ray crystallography and NMR spectroscopy, respectively. Relative to the canonical 11-strand β -sandwich FHA domain fold of FHA-1, FHA-2 is circularly permuted and lacking one β -strand. Nevertheless, the two share a conserved pThr-binding cleft. FHA-2 is less stable and more dynamic than FHA-1, yet binds model pThr peptides with moderately higher affinity ($\sim 50 \mu\text{M}$ versus $500 \mu\text{M}$ equilibrium dissociation constants). Based on NMR relaxation and chemical shift perturbation measurements, when joined within a polypeptide chain, either FHA domain can bind either linker pThr to form intra- and intermolecular complexes.

Protein phase separation has been recently shown to be a fundamental mechanism underlying the clustering of some proteins at the eukaryotic cell membrane. In chapter 3, I show that upon multi-site phosphorylation of the linker by several *Mtb* serine/threonine protein kinases including PknF, the isolated regulatory module readily multimerizes and phase separates into dynamic liquid droplets with diagnostic properties similar to those exhibited by eukaryotic proteins. The process is reversed by the sole *Mtb* serine/threonine phosphatase PstP. In the absence of phosphorylation, the Rv1747 regulatory module can still undergo phase separation, albeit at higher protein concentrations and into droplets with more fluid properties. This points to a synergy between specific FHA-pThr binding and additional weak association of the ID linker and/or the FHA domains leading to the pre-requisite multivalent interactions for demixing. Droplet formation of the regulatory module was replicated in a pseudo-two-dimensional system on model supported lipid bilayers. Potential clusters of Rv1747 were also detected in *Mtb* using ultra-high resolution Direct Stochastic Reconstruction Microscopy (dSTORM). This is the first reported example of phase separation by both a bacterial protein and an ABC transporter, and suggests possible mechanisms for the regulation of Rv1747 within *Mtb*. I hypothesize that the tunable, phosphorylation-dependent multimerization described here regulates Rv1747 transporter activity.

Lay Summary

Tuberculosis is having a renaissance, especially in its drug-resistant form and in a co-infection with HIV. Hence, a better understanding of the physiology of *Mycobacterium tuberculosis* – the bacterium causing it – is urgently needed in order to combat the disease. In this thesis, I present my investigation of the regulatory module of Rv1747, a membrane-spanning transporter of cell wall components that the pathogen needs to infect human cells. Phosphorylation of the regulatory module causes self-association. Akin to the separation of oil and water, this leads to its separation into protein-rich liquid droplets. Such phase separation has recently been discovered as a fundamental mechanism for the organization of macromolecules in eukaryotic cells. My research uncovers the first bacterial protein identified to phase separate, and suggests mechanisms by which Rv1747 functions within the bacterium. This may also inspire methods to prevent the phase separation of Rv1747 as therapies against tuberculosis.

Preface

Chapter 2 of this thesis is based on a reformatted version of the paper: **Florian Heinkel**, Leo Shen, Melissa Richard-Greenblatt, Christine L. Gee, Laurie M. Gay, Mark Okon, Jennifer M. Bui, Tom Alber, Yossef Av-Gay, Jörg Gsponer and Lawrence P. McIntosh (2018) “Biophysical Characterization of the Tandem FHA Domain Regulatory Module from the *Mycobacterium tuberculosis* ABC Transporter Rv1747” **Structure** (in press). I carried out all the experiments and analysis presented in this chapter with two exceptions: the crystal structure of FHA-1 was solved by Christine L. Gee and Laurie M. Gay in the laboratory of Tom Alber at the University of California, Berkeley, and ITC interaction studies were performed with the help of an undergraduate student, Leo Shen. Dr. Mark Okon assisted with the acquisition of NMR data. I received plasmids for the expression of PknF kinase constructs from Melissa Richard-Greenblatt, a graduate student in the laboratory of Dr. Yossef Av-Gay at UBC. She also helped with *in vitro* phosphorylation tests. Jörg Gsponer, Lawrence P. McIntosh and I wrote and edited the published manuscript.

Chapter 3 of this thesis is based on a reformatted initial version of the paper: **Florian Heinkel**, Libin Abraham, Joseph Chao, Gagandeep Narula, Mary Ko, Michael Gold, Keng Chou, Mark Okon, Yossef Av-Gay, Lawrence P. McIntosh and Jörg Gsponer (2018) “Protein Phase Separation Leads to Clustering of an ABC Transporter in *Mycobacterium tuberculosis*” (*in preparation*). I carried out all the experiments presented in this chapter assessing Rv1747 phase separation *in vitro*. I received purified samples of *Mtb* STPKs from the laboratory of Dr. Av-Gay and Mary Ko performed the *in vitro* phosphorylation assays using radioactively labelled ATP. Joseph Chao and Gagandeep Narula in Dr. Av-Gay's laboratory grew cultures and fixed samples of the various *Mtb* strains for microscopy studies. Dr. Horacio Bach generated the scFv's against Rv1747. Dr. Libin Abraham in the laboratories of Drs. Drs. Keng Chou and Michael Gold at UBC carried out dSTORM experiments and helped with the data analysis. Jörg Gsponer, Lawrence P. McIntosh and I are writing the final manuscript for submission.

I also contributed to a variety of additional projects in my time as a graduate student at the University of British Columbia that are thematically beyond the scope of this thesis but are summarized below.

Florian Heinkel, Alexander Cumberworth and Jörg Gsponer (2014) “Bridging Experiments and Simulations – Structure Calculations with a Dynamical Touch”, **Computational Approaches to Protein Dynamics: From Quantum to Coarse-Grained Methods**. Ed. Monika Fuxreiter (CRC Press). In this textbook chapter, we reviewed the current state on using restrained molecular dynamics simulations along with NMR spectroscopic data for probing protein dynamics on multiple time-scales. I wrote the manuscript with the exception of the section on dissecting ps-ns dynamics from ^{15}N relaxation experiments that Alexander Cumberworth, a fellow graduate student, prepared. Jörg Gsponer directed and edited the manuscript.

Florian Heinkel and Jörg Gsponer (2016) “Determination of Protein Folding Intermediate Structures Consistent with Data from Oxidative Footprinting Mass Spectrometry” **Journal of Molecular Biology** 428: 365-371. In this research article we presented a method to incorporate data from oxidative footprinting mass spectrometry in restrained molecular dynamics simulations to calculate structural ensembles consistent with experiments. Specifically, I used side-chain solvent accessible surface areas, calculated from atomic coordinates, as a proxy for the levels of oxidative labelling identified by mass spectrometry. The work presented is particularly interesting as the experimental method can probe important intermediate states of folding pathways that are otherwise hard to access. I wrote and implemented the code for the new restraint module in CHARMM and performed all experiments and analysis. Jörg Gsponer and I wrote the manuscript.

Ju Hun Yeon, **Florian Heinkel**, Minhui Sung, Dokyun Na and Jörg Gsponer (2016) “Systems-wide Identification of cis-Regulatory Elements in Proteins” **Cell Systems** 2: 89-100. This paper describes a sequence-based predictor for cis-regulatory elements (CREs). These are often disordered regions in proteins that can inhibit or enhance the function of this protein

through intra-molecular binding. Binding can be relieved or enforced, and thereby the function of the protein tuned, through post-translational modification, proteolysis or interaction with another binding partner. Identifying these CREs is of particular importance as mutations therein are often linked to disease. I helped with selection and analysis of data used for developing and testing the classifier. I contributed to writing the manuscript along with Ju Hun Yeon, Minhui Sung, Dokyun Na and Jörg Gsponer.

Antonio N. Calabrese, Scott M. Jackson, Lynsey N. Jones, Oliver Beckstein, **Florian Heinkel**, Jörg Gsponer, David Sharples, Marta Sans, Maria Kokkinidou, Arwen R. Pearson, Sheena E. Radford, Alison E. Ashcroft and Peter J.F. Henderson (2017) “Topological Dissection of the Membrane Transport Protein Mhp1 Derived from Cysteine Accessibility and Mass Spectrometry” **Analytical Chemistry** 89: 8844-8852. In this work, we applied cysteine-labelling chemistry based on N-ethylmaleimide to the sodium-coupled hydantoin transport protein Mhp1 from *Microbacterium liquefaciens* and quantified levels of modification in different conformations with mass spectrometry. We showed by different labelling of reporter cysteines that Mhp1 adopts a predominantly inward-facing conformation without substrate but upon addition of sodium ions and 5-benzyl-l-hydantoin (L-BH) it undergoes a conformational change towards an outward-facing conformation. I determined solvent accessibilities of residues in the known structures of Mhp1 to guide the selection of target residues as reporters for different conformational states. Dr. Gsponer and I helped to prepare the manuscript.

Cecilia Perez Borrajero, **Florian Heinkel**, Mark Okon, Jörg Gsponer, Lawrence P. McIntosh (2018) “Investigation of Relationship of Pax5 DNA Binding and Dynamics and Stability of DNA Binding Subdomains”. *Manuscript in preparation*. In this work, we investigated the structural and dynamic changes in the two Pax5 subdomains upon DNA binding and how binding affinity and thermodynamics are related to protein flexibility and stability. I performed and analyzed molecular dynamics simulations of the unbound Pax5 subdomains and DNA-bound Pax5 and helped to prepare and edit the manuscript.

Table of Contents

Abstract	iii
Lay Summary	iv
Preface	v
Table of Contents	viii
List of Tables	xiii
List of Figures	xiv
List of Abbreviations	xvii
Acknowledgements	xx
Chapter 1: Introduction	1
1.1 <i>Mycobacterium tuberculosis</i> Virulence and Infection	1
1.1.1 Tuberculosis Epidemiology and History of Discovery.....	1
1.1.2 Infection of Host Cells by <i>Mycobacterium tuberculosis</i>	2
1.1.3 The Role of Granulomas in Tuberculosis.....	6
1.2 Signaling in <i>Mycobacterium tuberculosis</i>	7
1.2.1 Serine Threonine Phosphorylation in <i>Mycobacterium tuberculosis</i>	8
1.2.1.1 Structural Analysis of Mycobacterial Serine Threonine Protein Kinases.....	13
1.2.1.2 Activation Mechanisms of Mycobacterial STPKs.....	16
1.2.1.3 Specificity of Mycobacterial Serine Threonine Phosphorylation.....	19
1.2.1.4 Biological Roles of <i>Mycobacterium tuberculosis</i> STPKs.....	19
1.2.2 The Serine Threonine Phosphatase PstP.....	22
1.2.3 Tyrosine Phosphorylation in <i>Mycobacterium tuberculosis</i>	22
1.2.4 Forkhead Associated Domains and their interplay with <i>Mtb</i> STPKs.....	22
1.3 Phase Separation of Biological Macromolecules	27
1.3.1 Phase Separation in Biology.....	27
1.3.1.1 Examples of Membraneless Organelles.....	30
1.3.1.2 Phase Separation and Membrane Protein Clustering.....	31
1.3.2 Biophysical Characterization of Phase Separation.....	32
1.3.2.1 Basic Characteristics of Biological Condensates.....	32

1.3.2.2	Flory Huggins Theoretical Description of Phase Separation of a Polymer	34
1.3.2.3	Molecular Basis of Biological Phase Separation.....	39
1.3.2.3.1	Interactions of Multivalent Proteins of Modular Architecture	39
1.3.2.3.2	Interactions Based on Intrinsically Disordered Regions in Proteins.....	41
1.3.2.3.3	The Role of Nucleic Acids in Biological Condensates	43
1.3.3	Functional Implication of Biological Phase Separation	44
1.3.3.1	Micro-reactors.....	44
1.3.3.2	Storage and Molecular Filters.....	45
1.3.3.3	Spatial Organization.....	46
1.4	Rationale and Thesis Overview	47
Chapter 2: Biophysical Characterization of the Tandem FHA Domain Regulatory Module		
from the <i>Mycobacterium tuberculosis</i> ABC Transporter Rv1747.....		
2.1	Overview	49
2.2	Introduction	49
2.3	Results	53
2.3.1	Rv1747 Regulatory Module is Composed of Independent FHA-1 and FHA-2 Domains Joined by a Disordered Linker	53
2.3.2	FHA-1 Has a Canonical FHA Domain Topology Whereas FHA-2 is Circularly Permuted	59
2.3.3	FHA-2 Binds pThr Peptides with Higher Affinity Than FHA-1	66
2.3.4	FHA-2 Exhibits Larger NMR Spectral Perturbations than FHA-1 Upon Phosphopeptide Binding.....	73
2.3.5	FHA-2 is More Dynamic and Less Stable Than FHA-1	74
2.3.6	FHA-1 and FHA-2 Bind the Phosphorylated Linker to Form Inter- and Intramolecular Complexes.....	79
2.3.7	Phosphorylated Regulatory Module Forms Higher Order Complexes	87
2.4	Discussion.....	90
2.4.1	Beads-on-a-String Regulatory Module.....	90
2.4.2	The More Dynamic FHA-2 has Higher Affinity for Phosphopeptides	90

2.4.3	FHA Domains Bind Phosphorylated Linker Threonines via Inter- and Intramolecular Equilibria.....	91
2.4.4	Potential Mechanisms for Regulating Rv1747 Transporter Function.....	92
2.5	Materials and Methods.....	93
2.5.1	Protein Expression and Purification of Rv1747 Constructs for NMR Spectroscopy	93
2.5.2	Protein expression and purification of Rv1747 ³⁻¹¹⁶ for crystallography.....	94
2.5.3	Protein Expression and Purification of the PknF Kinase Domain	95
2.5.4	Synthetic pThr-containing Rv1747 Peptides	96
2.5.5	<i>In vitro</i> Phosphorylation of Rv1747 Constructs.....	96
2.5.6	NMR Spectroscopy	97
2.5.7	NMR-based Structure Calculations of FHA-2 (Rv1747 ²⁰⁶⁻³¹⁰).....	97
2.5.8	Backbone Amide ¹⁵ N Relaxation Experiments.....	98
2.5.9	Amide HX Measurements.....	98
2.5.10	NMR-monitored Titrations	99
2.5.11	Crystallization of Rv1747 ³⁻¹¹⁶	100
2.5.12	X-ray Crystallography Data Collection, Processing, and Structure Determination of FHA-1 (Rv1747 ³⁻¹¹⁶).....	100
2.5.13	Isothermal Titration Calorimetry.....	101
2.5.14	Circular Dichroism Spectroscopy	101
2.5.15	Molecular Dynamics Simulations	102
2.5.16	Structural Alignment	103
2.5.17	Data and Model Depositions.....	103
Chapter 3: Phosphorylation-enhanced Protein Phase Separation of the Regulatory Module of a <i>Mycobacterium tuberculosis</i> ABC transporter		
3.1	Overview	104
3.2	Introduction.....	104
3.3	Results.....	106
3.3.1	Rv1747 Regulatory Module Undergoes Phase Separation upon Phosphorylation	106

3.3.2	Rv1747 ¹⁻³¹⁰ Droplets Exhibit the Diagnostic Properties of a Protein-rich Liquid Phase.....	110
3.3.3	Rv1747 ¹⁻³¹⁰ Droplet Formation is Cooperative	114
3.3.4	At Higher Concentrations, Unmodified Rv1747 ¹⁻³¹⁰ also Forms Droplets.....	114
3.3.5	Droplet Formation is Reversed upon Dephosphorylation by PstP.....	119
3.3.6	PknF and PstP Co-localize to Rv1747 ¹⁻³¹⁰ Droplets.....	119
3.3.7	Several <i>Mtb</i> STPKs Phosphorylate Rv1747 ¹⁻³¹⁰ and Induce Phase Separation	123
3.3.8	Rv1747 ¹⁻³¹⁰ Phase Separates in the Context of Model Membranes	127
3.3.9	Preliminary Evidence for Rv1747 Clustering along the <i>Mtb</i> Cellular Membrane.	130
3.4	Discussion.....	136
3.4.1	Phosphorylation-Independent Phase Separation	138
3.4.2	Multiple Kinases Induce Phosphorylation-enhanced Phase Separation.....	141
3.4.3	Biological Roles of Phosphorylation-enhanced Rv1747 Phase Separation.....	141
3.5	Materials and Methods.....	143
3.5.1	Expression and Purification of Rv1747 ¹⁻³¹⁰ for Droplet Studies	143
3.5.2	Expression and Purification of Active Kinase Domain Constructs of <i>Mtb</i> STPKs and the Phosphatase Domain of PstP	144
3.5.3	Fluorescent Labeling of Rv1747 ¹⁻³¹⁰ , GST-PknF ¹⁻²⁹² and PstP ¹⁻²⁴⁰	146
3.5.4	Imaging of Rv1747 ¹⁻³¹⁰ Droplets.....	147
3.5.5	Phase Separation of Rv1747 by <i>Mtb</i> STPK Phosphorylation.....	147
3.5.6	Phase Separation of Unmodified Rv1747 ¹⁻³¹⁰	148
3.5.7	Fluorescence Recovery After Photobleaching.....	148
3.5.8	Dephosphorylation of Rv1747 ¹⁻³¹⁰ by PstP and Dissolution Assay of Droplets	149
3.5.9	Co-localization Studies of Rv1747 ¹⁻³¹⁰ Droplets with PknF and PstP	149
3.5.10	Radioactive γ - ³² P Kinase Assay of <i>Mtb</i> STPKs on Rv1747 ¹⁻³¹⁰	149
3.5.11	Preparation of Rv1747 ¹⁻³¹⁰ Anchored in Supported Lipid Bilayers and Imaging of Membrane Clustering	149
3.5.12	Selection of Single Chain Variable Fragment (scFv) specific for Rv1747 ¹⁻¹⁵⁶	150
3.5.13	Co-localization Studies of scFv's with Rv1747 ¹⁻³¹⁰ Droplets.....	151
3.5.14	Functionalization of Glass Coverslips for dSTORM	151

3.5.15	Growth, Preparation and Labeling of <i>Mtb</i> for dSTORM	151
3.5.16	dSTORM Setup	152
3.5.17	Single Molecule Localizaion Microscopy (SMLM) Image Acquisition and Reconstruction	152
3.5.18	Rv1747 SMLM Cluster Analysis	153
3.5.19	Electron Microscopy.....	153
Chapter 4: Concluding Remarks.....		155
4.1	Structural and Dynamical Characterization of the Rv1747 FHA Regulatory Module.....	155
4.1.1	Significance, Limitations and Outstanding Questions	156
4.2	Phase Separation of the Rv1747 FHA Regulatory Module.....	158
4.2.1	Significance, Limitations and Outstanding Questions	158
Bibliography.....		162
Appendices		188
Appendix A ¹⁵N-HSQC Monitored Titrations of Rv1747 FHA Domains with Phospho- peptides.....		188
A.1	FHA-1 (Rv1747 ¹⁻¹⁵⁶) with pT152.....	188
A.2	FHA-1 (Rv1747 ¹⁻¹⁵⁶) with pT210.....	189
A.3	FHA-2 (Rv1747 ²⁰⁶⁻³¹⁰) with pT210.....	189
A.4	FHA-2 (Rv1747 ²⁰⁶⁻³¹⁰) with pT210.....	190

List of Tables

Table 1.1: Summary of Biological Roles of <i>Mtb</i> STPKs.....	21
Table 2.1: X-ray crystallography data collection and refinement statistics for FHA-1 (Rv1747 ³⁻¹¹⁶).....	61
Table 2.2: NMR spectroscopy data collection and refinement statistics for FHA-2 (Rv1747 ²⁰⁶⁻³¹⁰).....	62
Table 2.3: NMR Spectroscopic and ITC Studies of FHA – Phosphopeptide Interactions.	66
Table 3.1: Expression Conditions of STPK Constructs.....	145

List of Figures

Figure 1.1: Intracellular Fates of <i>Mycobacterium tuberculosis</i> after Uptake by Host Macrophages.....	4
Figure 1.2: Sequence Alignment of the <i>Mycobacterium tuberculosis</i> STPK Kinase Domains.....	9
Figure 1.3: Structures of the Kinase Domains of two STPKs of <i>Mycobacterium tuberculosis</i>	10
Figure 1.4: Domain Architecture of the <i>Mycobacterium tuberculosis</i> STPKs.	12
Figure 1.5: Structural studies of the Active Site of STPK Kinase Domains of <i>Mycobacterium tuberculosis</i>	15
Figure 1.6: Model of the Activation Mechanism of PknB.....	18
Figure 1.7: The FHA Domain and its Occurrence in <i>Mtb</i> Proteins.....	24
Figure 1.8: Membraneless Organelles in Eukaryotic Cells.....	29
Figure 1.9: An Example of Phase-Separated Protein Droplets.....	33
Figure 1.10: Flory-Huggins Theory of Polymer Phase Separation.....	36
Figure 1.11: Examples of Interactions Found in Biomolecular Phase Separation.....	40
Figure 2.1: Cartoon of the ABC Transporter Rv1747, Assuming a Minimal Functionally Active Homodimer.....	51
Figure 2.2: The Regulatory Module of Rv1747 is Composed of Two Independent FHA Domains Joined by an ID Linker.....	55
Figure 2.3: Assigned ¹⁵ N-HSQC Spectra of Rv1747 FHA Domain Constructs.....	57
Figure 2.4: Amide ¹⁵ N Relaxation Data of Rv1747 FHA Domain Constructs.....	58
Figure 2.5: FHA-2 has a Circularly Permuted FHA Domain Fold.....	60
Figure 2.6: Structural and Sequence Alignments of the Rv1747 FHA Domains.....	64
Figure 2.7: FHA-1 and FHA-2 have Conserved pThr-Binding Interfaces.....	65
Figure 2.8: FHA-1 and FHA-2 Bind Phosphothreonine Peptides.....	68
Figure 2.9: NMR Binding Isotherms of FHA-Phosphopeptide Titrations.....	69
Figure 2.10: ITC Binding Isotherms of FHA-Phosphopeptide Titrations at pH 6.....	71
Figure 2.11: ITC Binding Isotherms of FHA-Phosphopeptide Titrations at pH 8.....	72
Figure 2.12: FHA-1 is More Stable than FHA-2.....	76

Figure 2.13: Molecular Dynamics Simulations of Rv1747 FHA Domains.	78
Figure 2.14: NMR Spectral Comparison of Phosphorylated versus Unmodified Rv1747 Constructs.	80
Figure 2.15: FHA-pThr Mediated Intermolecular (<i>trans</i>) Interactions and Oligomerization.	81
Figure 2.16: Oligomerization of Rv1747 FHA Domain Constructs upon Phosphorylation.	83
Figure 2.17: Intra- and Intermolecular Interactions of Rv1747 upon Phosphorylation.	86
Figure 2.18: Phosphorylation of the Rv1747 Regulatory Module Leads to Higher Order Association.	89
Figure 3.1: Rv1747 ¹⁻³¹⁰ Undergoes Phase Separation <i>in vitro</i> upon Phosphorylation.	107
Figure 3.2: Sub-stoichiometric Phosphorylation of Rv1747 ¹⁻³¹⁰ by PknF is Sufficient to Induce Phase Separation.	108
Figure 3.3: Coalescence of Liquid Rv1747 ¹⁻³¹⁰ Droplets.	110
Figure 3.4: Diffusive Exchange of Rv1747 ¹⁻³¹⁰ Droplets measured by Fluorescence Recovery after Photobleaching.	111
Figure 3.5: Diffusive Exchange Occurs Between the Protein-rich and Protein-depleted Phases of an Rv1747 ¹⁻³¹⁰ Sample.	113
Figure 3.6: Rv1747 ¹⁻³¹⁰ Droplet Formation is Cooperative and Enhanced by Phosphorylation.	116
Figure 3.7: Higher Concentrations of PknF Phosphorylate Rv1747 ¹⁻³¹⁰ at Additional Sites that also Contribute to Phase Separation.	117
Figure 3.8: Treatment with PstP Dissolves PknF-induced Rv1747 ¹⁻³¹⁰ Droplets.	120
Figure 3.9: PknF and PstP Co-localize to Rv1747 ¹⁻³¹⁰ Droplets.	122
Figure 3.10: Several <i>Mtb</i> STPKs can Phosphorylate Rv1747 ¹⁻³¹⁰ at Multiple Sites and Induce Phase Separation.	126
Figure 3.11: Rv1747 ¹⁻³¹⁰ Anchored in Supported Lipid Bilayers Shows Concentration-dependent Phase Separation.	128
Figure 3.12: Diffusive Exchange of Rv1747 ¹⁻³¹⁰ Clusters on Supported Lipid Bilayers by Fluorescence Recovery after Photobleaching.	129
Figure 3.13: Single Chain Variable Fragment (scFv) Selected Against FHA-1 Co-localizes to Rv1747 ¹⁻³¹⁰ Droplets.	132
Figure 3.14: The scFv Selected against Rv1747 Localizes to Permeabilized Cells of <i>Mtb</i>	133

Figure 3.15: Analysis of Rv1747 Clustering in fixed Mtb Cells by dSTORM.	134
Figure 3.16: Clustering of Rv1747 in <i>Mtb</i> Cells Shown with Electron Microscopy.....	135
Figure 3.17: Model for Phosphorylation-enhanced Clustering and Activation Mechanism of Rv1747.	137
Figure 3.18: Sequence Analysis of Rv1747 ¹⁻³¹⁰	139
Figure 3.19: Model for Phosphorylation-enhanced Phase Separation and Different Liquid Properties of Rv1747 ¹⁻³¹⁰ Droplets.....	140

List of Abbreviations

ABC	ATP binding cassette
AF647	alexa fluor 647
ALS	amyotrophic lateral sclerosis
ATP	adenosine 5'-triphosphate
BCG	bacillus Calmette–Guérin
CD	circular dichroism
CLR	C-type lectin receptor
CR	complement receptor
CRE	cis regulatory element
DC	dendritic cell
ddH ₂ O	doubly distilled water
DDX	DEAD-box
DIC	differential interference contrast
DNA	deoxyribonucleic acid
dSTORM	direct stochastic optical reconstruction microscopy
DTT	1,4-Dithiothreitol
ECF	extracytoplasmic function
<i>E.coli</i>	<i>Escherichia coli</i>
ESI-MS	electrospray ionization mass spectrometry
FHA	forkhead associated
FRAP	fluorescence recovery after photobleaching
GFP	green fluorescent protein
GST	glutathione S-transferase
HIV	human immunodeficiency virus
HSQC	heteronuclear single quantum correlation
HX	hydrogen exchange
ID	intrinsically disordered
IDR	intrinsically disordered region
IFN- γ	interferon- γ

ITC	isothermal titration calorimetry
KD	kinase domain
LAT	linker for activation of T-cells
LCST	lower critical solution temperature
MAP	microtubule-associated protein
MD	molecular dynamics
MGC	multinucleated giant cell
MMP	matrix metalloproteinase
mRNP	messenger ribonucleoprotein
<i>Mtb</i>	<i>Mycobacterium tuberculosis</i>
<i>mym</i>	mycobacterial monooxygenase
NA	numerical aperture
NBD	nucleotide binding
NEM	N-ethylmaleimide
NMR	nuclear magnetic resonance
OG	oregon green
PASTA	penicillin-binding protein and serine threonine kinase associated domain
PBS	phosphate buffered saline
PCR	polymerase chain reaction
PF	protection factor
PG	peptidoglycan
PIM	phosphatidyl- <i>myo</i> -inositol mannoside
PML	promyelocytic leukemia
PPM	metal-dependent protein phosphatase
PRM	proline rich motif
PTB	polypyrimidine tract-binding protein
RMSD	root mean square deviation
RNA	ribonucleic acid
RNAi	RNA interference
RNP	ribonucleoprotein

ROI	region of interest
RRM	RNA recognition motif
rRNA	ribosomal RNA
scFv	single chain variable fragment
SDS-PAGE	sodium dodecyl sulfate polyacrylamide gel electrophoresis
SH2/SH3	src homology 2/3
SIM	SUMO interacting motif
SLB	supported lipid bilayer
SMLM	single molecule localization microscopy
STPK	serine threonine protein kinase
SR	scavenger receptor
SUMO	small ubiquitin modifier
TAU	tubulin-associated unit
TCEP	tris(2-carboxyethyl)phosphine
TCR	T-cell receptor
TEV	tobacco etch virus
TIRF	total internal reflection fluorescence
TM	transmembrane
TMD	transmembrane domain
TNF	tumor necrosis factor
TPR	tetratricopeptide repeat
TSSS	type VII (seven) secretion system
UCST	upper critical solution temperature
UV	ultraviolet

Acknowledgements

First and foremost, I want to thank my supervisors Drs. Jörg Gsponer and Lawrence McIntosh for their mentorship and guidance throughout my time as a graduate student. The freedom I had to explore new ideas and how I was encouraged to try new things and collaborate with many great scientists at UBC was truly amazing. In all my time here, I always felt like we let science take us to places that we had not anticipated, instead of trying to force ourselves down a preconceived path. If not for this we would surely have missed the really interesting and fun things along the way. Without Dr. Mark Okon and his endless patience for teaching me the basic technical aspects, I would not have been able to run a single NMR experiment and so he is a major piece in the puzzle that is my PhD. A big thanks also goes to all members past and present of the McIntosh and Gsponer labs (Laura, Adrienne, Gerald, Soumya, Gen, Desmond, Kody, Miriam, Uli, Leo, Jacob, Cecilia, Ben, Karl, Chloe, Stacy, Alex, Geoff, Roy, Guillaume, Dokyun, Ashwani, Eric, Nawar, Andrew, Nivi): always eager to discuss science or just talk about the latest weekend adventures and very good at coming up with an excuse to go for a coffee. Having someone go through the same stages of the PhD program at the same time as I did was amazing and comforting. Therefore, a special thanks goes to Dr. Cecilia Perez-Borrajero for many great conversations and I hope our paths will cross in Germany again, soon. Another special mention is due for Dr. Soumya De. Thank you for sharing your wisdom with me, for showing me the ropes of structural biology and being a great role model to look up to as a young scientist. You have been truly missed ever since you left the lab and are a wonderful mentor. Your students are a lucky bunch. I also have had the pleasure to collaborate with many great people at UBC that I owe a great deal of thanks to. In particular, the STORM collaboration with Dr. Libin Abraham stands out for his knowledge across many fields, excitement about science, and eagerness to help. He is a true inspiration. Furthermore, I want to thank my committee members Drs. Natalie Strynadka and Franck Duong for their support and mentorship throughout my PhD.

Science can be fickle, and while the feeling of discovery and being the first to ever lay eyes on something is special, and drives my curiosity, there are just as many dull moments and moments of failure and self-doubt. Without a support system, these are hard to endure. Fortunately, when the going got tough, I was surrounded by the most special, uplifting and

encouraging people. Most importantly, thank you Monika for always being so supportive, honest and forgiving. Thank you for being all around wonderful and understanding, for giving me confidence and for showing me that there is a lot of fun to be had just a little outside my comfort zone. You are my rock, my best friend, and you mean the world to me. I cannot wait to find out what life has in store for us. I also thank you for letting me be part of your amazing and loving family and for bringing Charlie, the best dog in the world, into my life. Thank you Lucie and Fred for your real-world perspective on many things. I sometimes get stuck in my academic bubble and a helping hand to step out of it every once in a while has been invaluable. I also want to thank my family and friends in Germany: Mama, Papa, Yannic, Magali, Oma Gisela, Peter, Nana, Timo, Jochen and Philipp. You have always supported me and the path that I chose, even when it meant moving an ocean and a continent away, not hearing from or seeing me when I was immersed in science, having fun in the mountains or training to swim, bike and run for extended periods of time. I cannot put into words how much it means to me to have your caring and unconditional support.

There are many wonderful people in Vancouver that have touched my life in so many ways over the course of the last six years. You have made all the difference to my experience here, and I could not have made it without your support, through both good times and bad, and without you generally putting up with me. Thank you Yaseen for being my first and best friend in a country far away from home, someone to do all the touristy things with and still the first person I go to for advice. Thank you, Laura for having been the best roommate I have ever had and for always being there to reminisce about European food. Thank you, Jacob and Christel for teaching me to appreciate the other football that is not really played with the foot, which means there was always something to do on a lazy Sunday afternoon.

Having a balanced life and being healthy physically and mentally is instrumental to making it through graduate school and, more generally, life. Despite this, it is still often taken for granted. The UBC Triathlon Club and the friends I have made there, have taught and inspired me to live a healthier, more balanced and better life. With them I have achieved things both academically and athletically that I would not have thought were possible. Thank you, Monika, Winston, Melissa, Mike, Jen, Kately, Eric, Katrina, Kazune, Matt, Anna, Ali and so many more for welcoming me with open arms at a point where completing a triathlon seemed unattainable to

me. Thanks for pushing me to swim/bike/run just a little faster and a little further, and continuing to be my support system that I draw from for strength and security. I hope for many more fun hours in training and races with you. Or, at the very least to keep motivating one another through our Strava workouts. See you all in Kona and/or Boston one day!

Graduate school, all the people that helped me get to where I am right now, and the unbelievably beautiful place that I could call home for almost six years, have made me a better scientist and a better person, and I will forever be grateful for that. I have been blessed with experiencing more love, hope, tolerance and positivity than I could ever ask for. I am confident and ready to take on whatever comes next in life and science. Always curious, I hope to never stop learning and never tire of sharing what I have been fortunate enough to have been taught. I shall always strive to move forward, one foot in front of the other, always remembering to “Go With the Flo”.

Chapter 1: Introduction

1.1 *Mycobacterium tuberculosis* Virulence and Infection

1.1.1 Tuberculosis Epidemiology and History of Discovery

Mycobacterium tuberculosis (*Mtb*) is the etiological agent of tuberculosis, a disease that has been ailing humankind with a high mortality rate for millenia. References to illnesses accompanied by symptoms matching those of tuberculosis are even in the Old Testament (1) and lesions compatible with skeletal tuberculosis have been detected on a human skull found in Turkey that can be dated as far back as 500,000 years (2). PCR amplification and sequencing of genetic material (3), as well as the isolation of mycobacterial lipids (4), found on bone lesions of a 17,000-year-old bison skeleton in Wyoming, USA constitute the undisputed proof of the earliest existence of the *Mtb* complex. The oldest genetic evidence for human *Mtb* infections stems from the 9000-year-old remains of Neolithic humans in the Eastern Mediterranean (5).

Tuberculosis, known as “consumption”, was thought to be a hereditary disease until 1865 when Jean-Antoine Villemin showed it to be transmissible (6, 7). At that time, when mortality was high, the cause unknown and the scare of the disease at its height, people in New England even blamed its transmission on vampires (8). Indeed, some of the widespread practices of exhumation and burning of the corpses from tuberculosis victims may be in part credited as inspiration for Bram Stoker’s “Dracula” (9). *Mtb* was finally identified in 1882 as the causative agent of tuberculosis by Robert Koch (10). Through research that was eventually recognized by a Nobel Prize in 1905, Koch succeeded in staining this Gram-negative bacterium in samples of lesions from tuberculosis patients, and subsequently isolating it on solid cultures.

In 1910, tuberculin, a protein that could be extracted from liquid cultures of *Mtb*, was shown by Charles Mantoux (11) to cause an immune-response and intra-dermal reaction in infected individuals. This discovery led to the well-known tuberculin skin test for diagnosis of tuberculosis. In the search for a vaccine against *Mtb*, Leon Charles Albert Calmette and Camille Guérin passed a virulent strain of *Mycobacterium bovis* on media more than 200 times over the course of more than 10 years and obtained a non-virulent strain of mycobacteria (12, 13). This strain, termed bacillus Calmette–Guérin (BCG), was later shown to have large DNA deletions (14). BCG is still used as live vaccine, preventing meningitis in infants and military forms of

tuberculosis in adults (15). Unfortunately, it has proven ineffective in preventing pulmonary tuberculosis.

A century later, and even with the wide availability of effective antibiotics, *Mtb* is still a major bacterial killer (16). Currently, about one-third of the world's population is thought to be infected. In most of these humans, growth of the bacillus is kept at bay by the immune system and the only indication of infection is immuno-sensitivity to mycobacterial antigens. However, about 5-10 % of infected individuals develop active (and transmissible) tuberculosis that manifests itself with a chronic cough with blood-containing sputum, fever, weight loss, and if left untreated, death (17). Severity is increased by its interplay with human immunodeficiency virus (HIV), which puts infected individuals at a higher risk for developing the active disease, and by the emergence of drug resistant strains of *Mtb* due to improper use and administration of antibiotics (18). Some of these strains virtually do not respond to treatment with any available antibiotic. Although the need for new and more effective ways to treat tuberculosis is critical, the first drug for the treatment of highly multi-drug resistant versions of the disease approved by the FDA in more than 40 years was bedaquiline in 2012 (19).

1.1.2 Infection of Host Cells by *Mycobacterium tuberculosis*

Transmission of tuberculosis most commonly occurs via the inhalation of an aerosol of bacilli-containing droplets and deposition of the bacteria in distal alveoli (reviewed in (17)). *Mtb* are then recognized and ingested by alveolar macrophages and other phagocytic cells, including neutrophils, monocyte-derived macrophages and dendritic cells, that are recruited to a locus of infection. Of course, common strategies of most pathogens for survival include evasion of phagocytic cell recognition and uptake. In contrast, *Mtb* has evolved to intrude and survive in phagocytic cells and thereby not only circumvent, but exploit this host defense mechanism. A summary of the intracellular survival strategies of *Mtb* is shown in Figure 1.1.

Although several host cells play a role in the uptake of bacteria (20), macrophages are the best studied system for mycobacterial infections. *Mtb* cells are recognized by macrophage receptors that include C-type lectin receptors (CLRs), scavenger receptors (SRs) and complement receptors (CRs). Many receptors interacting with different *Mtb* specific antigens results in many routes of uptake into the phagocytic host cell. Although the exact mechanisms of

specificity for targets recognized on the *Mtb* surface and the interplay between receptors remain to be elucidated, it has been shown that the type of receptor-antigen interaction plays a role in determining the intracellular fate of the bacilli (21, 22).

Upon ingestion by macrophages, bacilli are typically engulfed in phagosomes. Thus, an important pathway for intracellular survival of *Mtb* involves the prevention of phagosomal acidification and thereby the arrest of phagosomal maturation into lysosomes (Figure 1.1A) (23). This less acidic, endosome-like environment can then be exploited for essential nutrients like iron from recycling endosomes (24). Several *Mtb* virulence factors have been shown to help ensure this phagosomal arrest (17, 25). Two central processes that are inhibited in phagosomes by multiple virulence factors are the conversion of inactive, GDP-bound Rab5 to active GTP-bound Rab7 (26, 27) and the production of phosphatidylinositol-3-phosphate (PI3P) in the phagosome membrane (28-30). Both GTP-Rab7 and high contents of PI3P are associated with recruitment of late endosomal markers, like the Endosomal Sorting Complex Required for Transport (ESCRT), that assures lysosomal delivery and acidification of the phagosome (25). While the mechanism of phagosomal arrest seems to be a major route to intracellular survival, there is evidence that *Mtb* can still survive and grow in a more acidic, more lysosome-like environment if early inhibition of phagosomal maturation fails (Figure 1.1B) (21, 31).

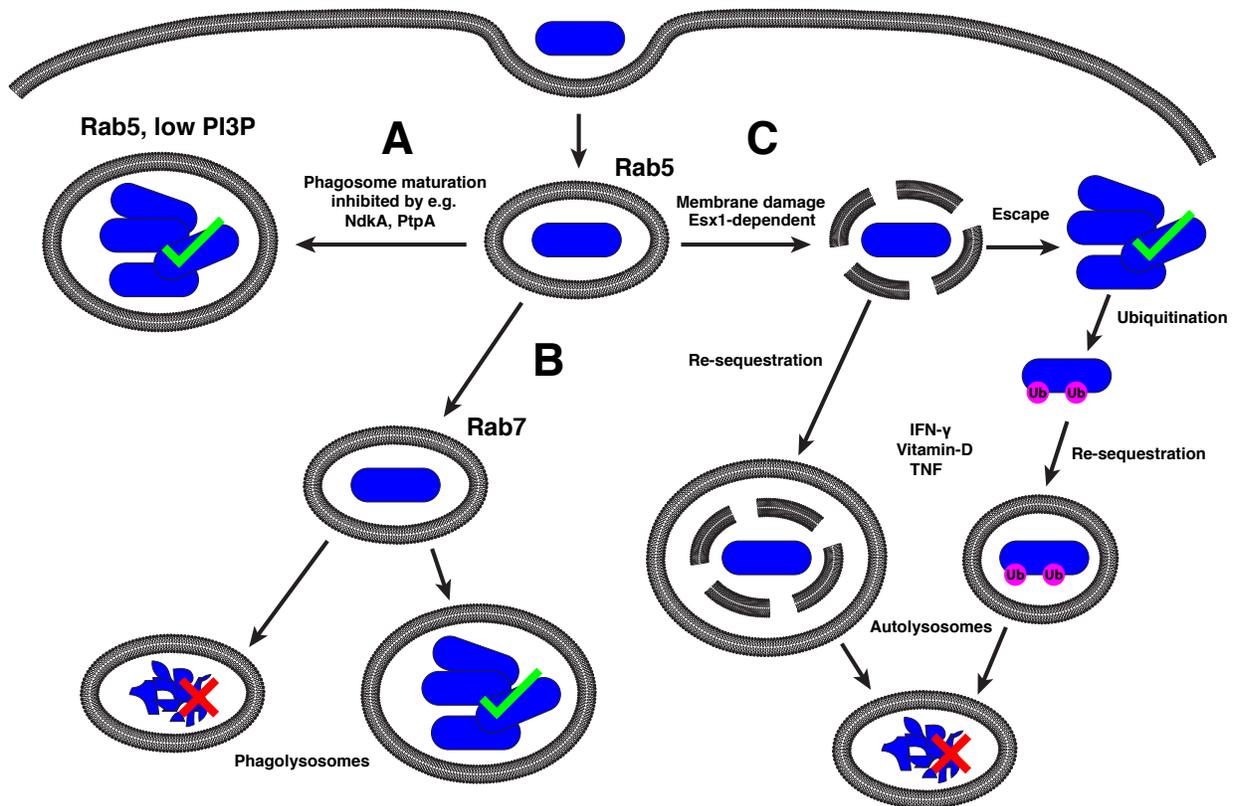


Figure 1.1: Intracellular Fates of *Mycobacterium tuberculosis* after Uptake by Host Macrophages.

Bacilli are shown in blue. A) Inhibition of Rab5-Rab7 conversion and low PI3P leads to arrest of phagosomal maturation and bacterial survival. B) If phagosomes mature into lysosomes normally, bacteria are cleared. However, there is evidence for survival even in acidic, later stage phagosomal organelles (21, 31). C) Mycobacteria can damage the phagosomal membrane through activity of Esx-1. Upon phagosomal escape, intracellular survival is possible. However an autophagy-like immune response (xenophagy) in activated macrophages (activated by IFN- γ , Vitamin-D, TNF) can occur. In this case, damaged phagosomes and intracellular bacilli are decorated with ubiquitin, re-sequestered and cleared in autolysosomes.

As an additional route of survival, mycobacteria are able to cause damage to the phagosomal membrane and thereby escape into the cytoplasm, where survival and replication are possible (Figure 1.1C). An important factor in this mechanism is Esx-1, one of the five type VII secretion systems (TSSS) that are capable of exporting endogenous virulent proteins across the complex *Mtb* envelope. The set of genes encoding the Esx-1 secretion system is missing in the BCG strain and has been identified to be largely responsible for its attenuation of virulence (32-34). Importantly, Esx-1 is also an essential factor for phagosomal membrane damage and escape (35-37). In addition to the components making up the secretory machinery, the genomic locus of Esx-1 contains genes encoding secreted factors of ~ 100 amino acids that contain a characteristic Trp-X-Gly (W-X-G) hairpin motif. Two of these so-called WXG100 proteins, Esx-A and Esx-B, are implicated in mediating escape of *Mtb* from the phagosome. The Esx-A/Esx-B heterodimer is thought to dissociate in the acidic environment of the phagosome, releasing Esx-A which displays membranolytic activity (33, 36, 37) (38). Additionally, Esx-A and Esx-B are important secreted antigens recognized by CD4⁺ and CD8⁺ T cells. In line with this finding, by mediating cytoplasmic release, Esx-1 is also responsible for activation of the cytoplasmic innate immune response to a *Mtb* infection (39-41). This points to a role for Esx-1 in the adaptive immune response and highlights the intricate balance between mechanisms of intracellular survival and the host immune response in a mycobacterial infection.

Another process called xenophagy that counteracts the mycobacterial infection after phagosomal damage and escape is similar to autophagy. In autophagy, harmful intracellular particles such as protein aggregates and damaged organelles are enclosed by and then degraded in a double-membrane organelle. This organelle is termed the autophagosome in the early stages and the autolysosome after fusion with lysosomes and acidification. In the case of xenophagy, the host engulfs and clears bacteria and viruses residing in its cytoplasm (42). The intracellular pathogens, including *Mtb*, are modified with ubiquitin and then enclosed by and eventually cleared in an autolysosome (43-47). Autophagy is therefore required to be negatively regulated to allow for cytoplasmic survival of *Mtb* (48) in resting macrophages. Conversely, in activated macrophages, host cellular factors that are associated with autophagy activation promote killing of intracellular mycobacterial bacilli and their clearance by autophagy. They include vitamin D (46) and interferon- γ (IFN- γ) (47). Additionally, tumor necrosis factor (TNF), in concert with

IFN- γ , is also able to stimulate macrophages to kill and clear invading mycobacteria (49), whereas an excess of TNF can lead to enhanced immunopathology (50, 51). This points to the tight balance of an adequate immune reaction and minimal tissue damage during a tuberculosis infection. *Mtb* actively exploits this and modulates TNF production to its advantage in macrophages by secretion of an adenylate cyclase which causes enhanced inflammation and favors virulence (52).

1.1.3 The Role of Granulomas in Tuberculosis

After successful survival and replication in the infected host cell, mycobacteria are released by apoptosis or necrosis of that cell and taken up by nearby phagocytic cells (17). Also, upon reaching a certain density inside one host cell, uninfected macrophages are recruited to the locus of infection by mycobacterial release of Esx-A and the successive induction of matrix metalloprotease 9 (MMP9) in nearby epithelial cells (53). These can then ingest the released bacilli after cell death of the initially encountered cell, effectively amplifying the number of infected cells (54). Eventually, the enrichment of macrophages in the area of infection leads to the formation of granulomas. These are organized aggregates of macrophages seen as the histopathological lesions in the lungs of patients with pulmonary tuberculosis. A variety of macrophages, including mature macrophages, epithelioid macrophages, foamy macrophages and multinucleated (or Langhans) giant cells (MGCs) (17, 55-57) are found in granulomas. Additionally, myeloid dendritic cells (58), T lymphocytes (59, 60) and B cells (59) are present and help mediate a proper immune response. By facilitating amplification and establishment of infection, the formation of granulomas benefits the bacilli yet also ensures an optimal host immune response. This leads to a stalemate between *Mtb* and the host. Inside granulomas, the pathogen evades complete eradication through various mechanisms mentioned above. At the same time, the host immune system can keep the bacterial burden at bay, which results in the chronic disease characteristics of tuberculosis.

1.2 Signaling in *Mycobacterium tuberculosis*

Signalling, the conversion of a stimulus into a cellular response, is an essential mechanism for cells to adapt to their environment. For bacterial pathogens, a rapid change in the extracellular environment is brought about upon infection of a host and signalling pathways that lead to quick adaptation are therefore important factors for virulence. They are especially important for an intracellular pathogen like *Mtb*, enduring the harsh phagosomal environment and having to interfere with and inhibit host intracellular clearance mechanisms upon infection.

A class of widespread signalling systems in bacteria are two-component systems (61). In their simplest form, a homodimeric transmembrane sensor histidine kinase is stimulated by an extracellular signal, which promotes auto-phosphorylation of a His residue. That phosphate is then transferred to an Asp residue in a response regulator, such as a transcription factor, rendering it active and thereby mediating the cellular response (62, 63). In contrast, simpler and evolutionarily older one-component systems, such as the lactose operon repressor, contain a sensor and an output domain in a single polypeptide chain. These proteins typically bind small intracellular signalling molecules in their sensor domain, eventually leading to transcriptional changes caused by the DNA binding output domain (64). A third mechanism of bacterial signalling is based on another group of transcriptional regulators, the extracytoplasmic function (ECF) σ factors. These factors are small interchangeable units of the RNA polymerase holoenzyme that convey specificity for distinct promoter sequences (65). Often, for a given ECF σ factor there also exists a membrane spanning anti- σ factor that serves as an inactivating binding partner. Upon reception of an external stimulus via the extracellular domain of the anti- σ factor, the σ factor is released and transcription is activated (66).

In eukaryotic cells, one central signalling mechanism is based on the phosphorylation of serine and threonine residues in proteins by serine threonine protein kinases (STPKs). The human genome encodes for more than 500 STPKs and an estimated two-thirds of human proteins are post-translationally modified by phosphorylation (67-69). Although typically rare, it is now well recognized that STPKs are present in prokaryotes. Indeed, *Mtb* is a particularly interesting example of a small subset of bacteria, which also include *Acidobacteria*, *Actinobacteria*, some *Cyanobacteria* and one order of the *Deltaproteobacteria*, that possess a relatively high number of eukaryotic-like STPKs and a relatively low number of two-component systems (70).

1.2.1 Serine Threonine Phosphorylation in *Mycobacterium tuberculosis*

The *Mtb* genome encodes 11 STPKs. Although differing significantly in sequence (< 30 % identity) (70), the kinase domains (KDs) of these STPKs show the same overall sequence features as their eukaryote counterparts (Figure 1.2). All of the characteristic subdomains of the eukaryotic STPKs (71) are found in the mycobacterial kinases. These enzymes are comprised of an N-terminal lobe, spanning subdomains I – IV and a part of V, and a C-terminal lobe, spanning subdomains V through XI (Figure 1.3). The N-terminal lobe contains mostly β -sheets and one long α -helix, whereas the C-terminal lobe only has α -helical secondary structure. The C-terminal lobe is responsible for substrate binding and phosphate transfer. ATP and Mg^{2+} are bound in a cleft between the two lobes.

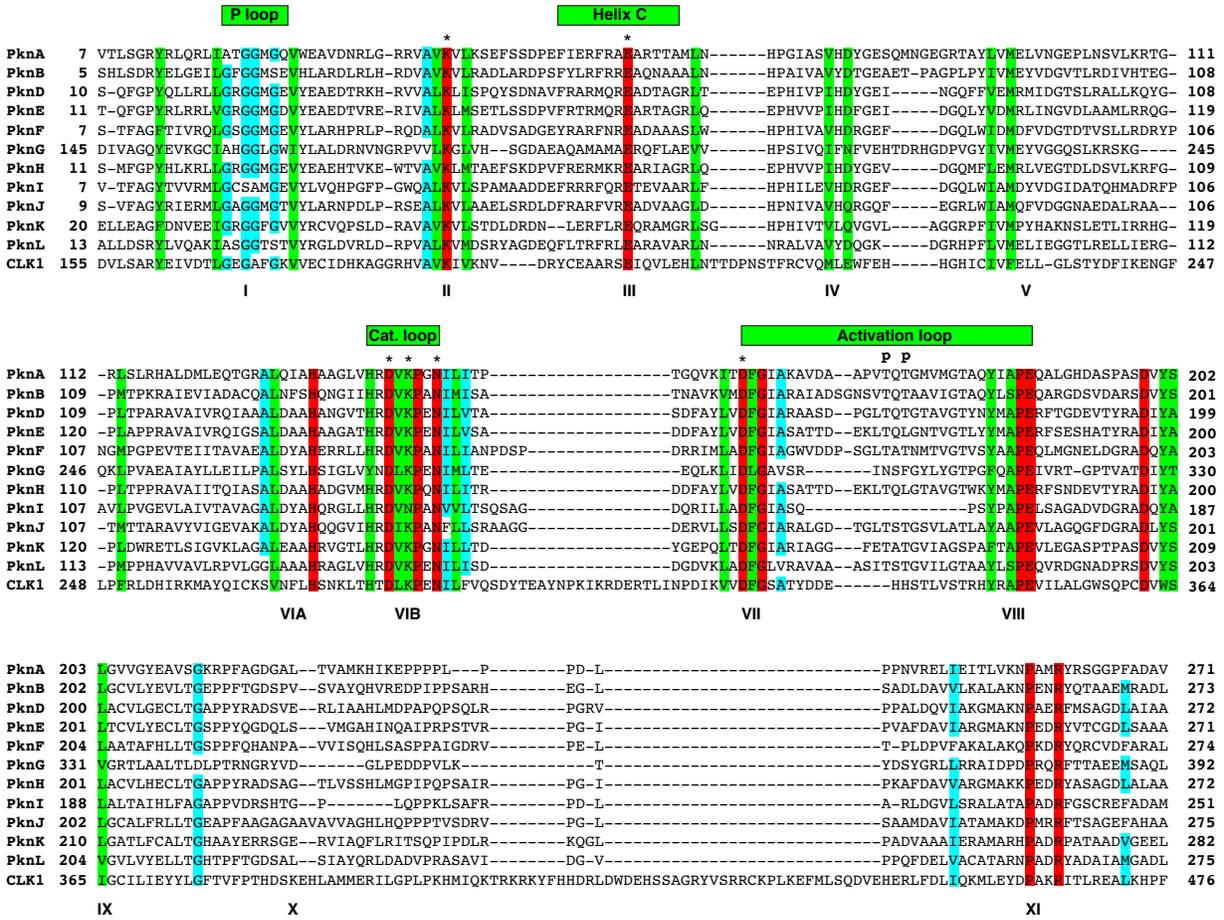


Figure 1.2: Sequence Alignment of the *Mycobacterium tuberculosis* STPK Kinase Domains. Sequence boundaries are indicated for the 11 *Mtb* STPKs (PknA through PknL) and the human kinase CLK1. The alignment was generated with Clustal Omega (72). Invariant residues are highlighted in red, highly conserved residues with similar properties in green, and conserved residues with some outliers in cyan. Common phospho-acceptors in the activation loop are indicate with a “P” and important, conserved residues discussed in the text are indicated with an asterisk. The locations of conserved motifs are shown in green at the top, and characteristic kinase subdomains are marked with Roman numerals (I through XI) at the bottom of the alignment.

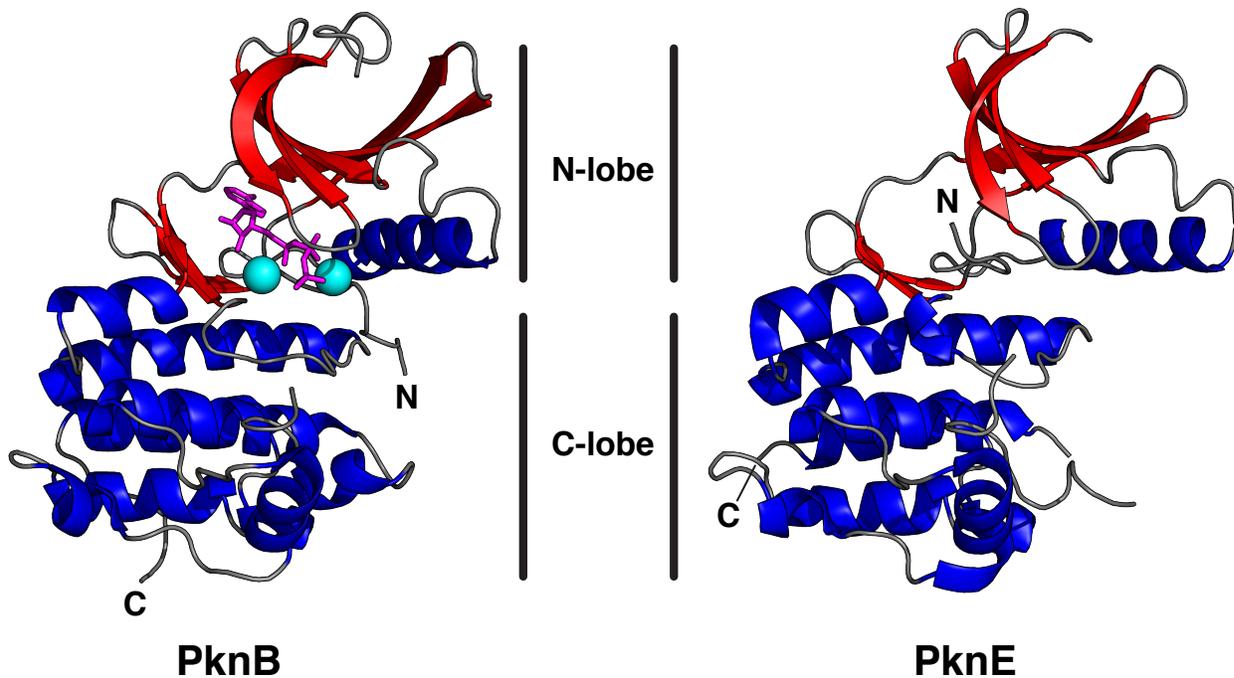


Figure 1.3: Structures of the Kinase Domains of two STPKs of *Mycobacterium tuberculosis*. Cartoon representations of PknB (left, PDB ID: 1MRU) and PknE (right, PDB ID: 2H34). β -strands are in red, α -helices in blue and the N- and C-terminal lobes are indicated. ATP- γ -S and Mg^{2+} bound to PknB are shown in pink and cyan, respectively.

In addition to the KDs, the mycobacterial STPKs have appended, divergent sequences that provide specificity for a wide variety of signalling inputs (Figure 1.4). Nine of the 11 STPKs in *Mtb* possess a predicted helical membrane spanning region and all, except PknL, have extracellular domains of various sizes, architectures and functions. These membrane-anchored STPKs can be subdivided into three clusters of three evolutionarily related sequences, suggesting that they stem from the duplication of ancestral genes (69). The two cytoplasmic STPKs, PknG and PknK, are more evolutionarily distant and were probably individually acquired (69). Of the STPKs found in *Mtb*, PknA, PknB, PknG, and PknL are present in all *Mycobacteria*, pointing to a role of these kinases in central aspects of mycobacterial physiology. Corroborating this notion, only PknA and PknB are essential for *Mtb*. Counteracting the activity of its 11 STPKs, *Mtb* possesses one serine threonine phosphatase (PstP; section 1.2.2).

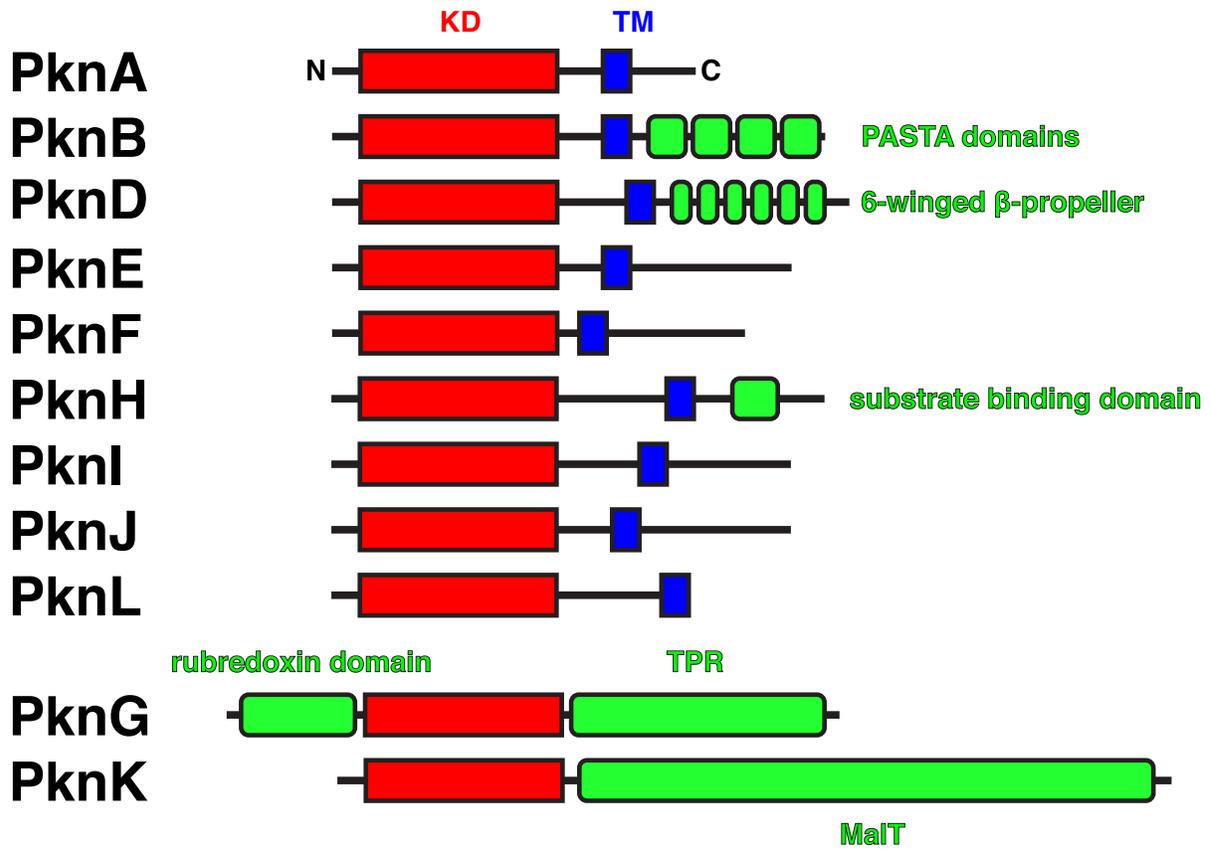


Figure 1.4: Domain Architecture of the *Mycobacterium tuberculosis* STPKs.

Schematic representation of the 11 *Mtb* STPKs, aligned at their KDs (red). Transmembrane domains (TM) are in blue, and accessory domains in green with their identities indicated. The two cytoplasmic kinases PknG and PknK are shown separately at the bottom.

1.2.1.1 Structural Analysis of Mycobacterial Serine Threonine Protein Kinases

The three-dimensional structures of the KDs of PknA (73, 74), PknB (75-78), PknE (79), PknG (80, 81) and PknI (82) have been determined experimentally and give valuable insights into the activation and catalytic mechanisms of mycobacterial STPKs. Broadly speaking, STPKs exist in "open" inactive and "closed" active conformations. Snapshots of these conformations are provided by the X-ray crystallographic structures of PknE and PknB, respectively (Figure 1.3 and 1.5A). In the closed state of PknB, helix C is shifted towards the active site, positioning Glu59 close to Lys40 of the nucleotide binding domain and thereby orienting its sidechain aminium group near the α and β phosphates of ATP (Figure 1.5B and C). In contrast, these critical interactions do not occur in the open conformation of apo PknE due to an alternative positioning of helix C. Furthermore, Asp138 and Asn143, two residues in the catalytic loop of PknB, are properly paired in the active, closed conformation to facilitate nucleophilic attack of the substrate serine/threonine hydroxyl group on the γ -phosphate of ATP (Figure 1.5B and C). Other conserved features in the mycobacterial STPKs are the glycine-rich P-loop in the N-terminal lobe that binds the ATP phosphates, the hydrophobic adenine binding pocket and the Asp-Phe-Gly (DFG) motif on the N-terminal end of the "activation loop." The aspartate of this motif plays an important role in positioning Mg^{2+} in the active site.

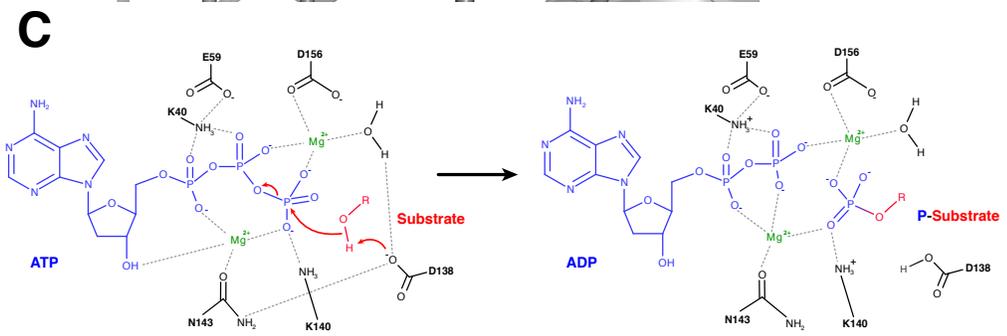
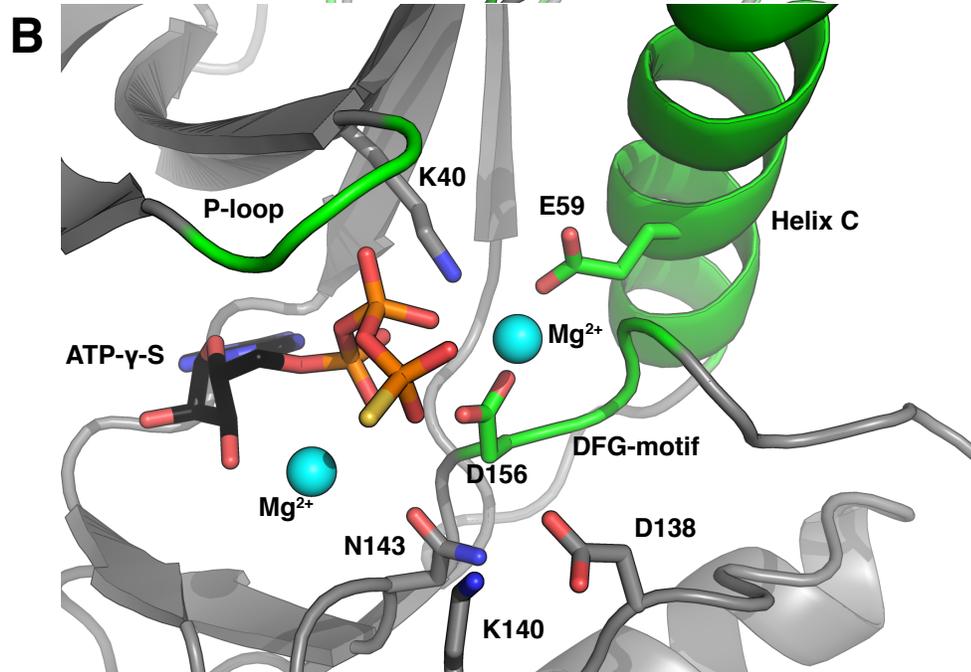
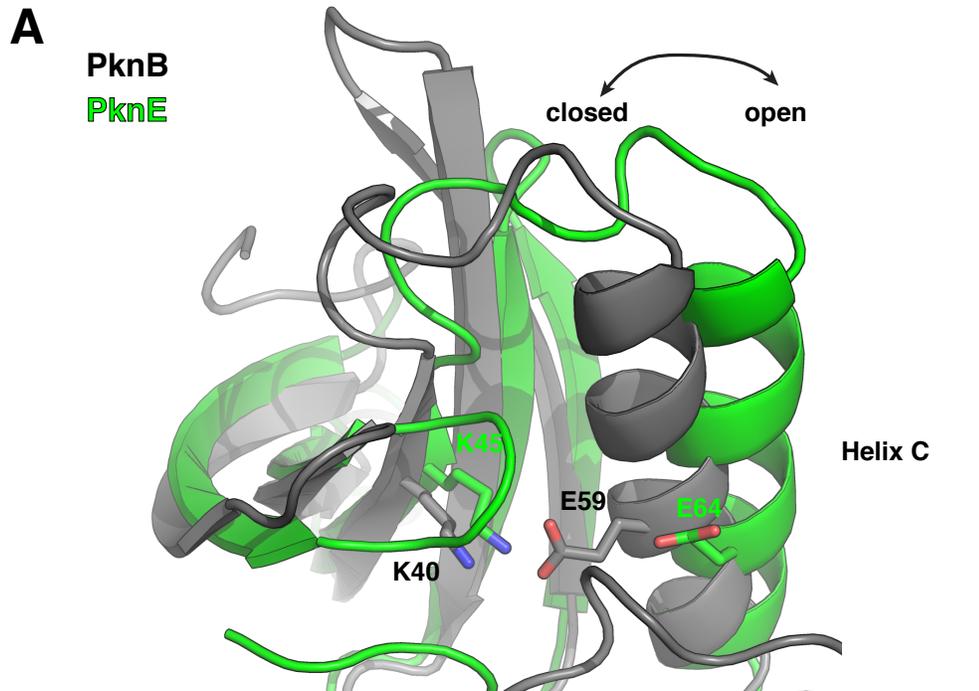


Figure 1.5: Structural studies of the Active Site of STPK Kinase Domains of *Mycobacterium tuberculosis*.

(A) Overlay of the closed conformation of PknB (grey, PDB ID: 1MRU) and PknE (green, PDB ID: 2H34). The shift of helix C between the open inactive and closed active state is indicated. The side chains of conserved lysine and glutamate residues are optimally positioned for to assist nucleophilic attack of the substrate on ATP in the closed conformation, and too far apart in the open conformation. (B) The catalytic site of PknB with conserved motifs (helix C, P loop and DFG motif of the activation loop) highlighted in green and side chains of important residues in stick representation. Bound ATP- γ -S and Mg²⁺ are also shown. (C) Proposed mechanism of substrate phosphorylation catalyzed by PknB. Asp138 serves as a general base to facilitate nucleophilic attack of the substrate hydroxyl on the terminal γ -phosphate of ATP. Important residues binding ATP, Mg²⁺ and substrate in the active site are also indicated.

1.2.1.2 Activation Mechanisms of Mycobacterial STPKs

Important insights into the mechanisms of activation of the mycobacterial STPKs have been gained through a battery of experimental studies. Generally, threonine/serine residues in the activation loop of a STPK are phosphorylated and this leads to enhanced kinase activity towards its substrate targets (83-87). Although termed "auto-phosphorylation", the post-translational modification generally occurs in an intermolecular manner ("trans-phosphorylation") and thus is linked with STPK dimerization or oligomerization. The interaction between KDs is generally of low affinity and thus association of accessory domains upon interaction with signalling molecules is likely required for activation. This is seen for PknG, where tetratricopeptide repeat domains (TPRs) C-terminal of the KD aid dimerization (80). Additionally, the N-terminal accessory rubredoxin domain senses the redox state of the environment via two Cys-X-X-Cys-Gly iron binding motifs and possibly relays that information to regulate kinase activity by interacting with the N- and C-terminal lobes of the PknG KD (80, 81, 88, 89). One notable outlier is PknG that possesses an unusually ordered activation loop that does not need to be phosphorylated for activation.

A "back-to-back" homodimerization of previously non-phosphorylated KD is thought to be the first step in the activation process for PknB (78, 90) (Figure 1.6), PknD (91) and PknE (79). This is also seen for eukaryotic kinases such as PKR (92). In this arrangement, the catalytic sites of the monomer subunits are well separated and unable to phosphorylate one another. Hence, this requires trans-phosphorylation of another homodimeric STPK. Once phosphorylated, the KDs are then fully active even in a monomeric state as shown for PknB (78, 90). A PknB mutant in complex with the Kt5720 inhibitor also shows a "front-to-front" dimer arrangement such that the activation loop of one monomer can be trans-phosphorylated by inserting into the active site of the partner KD (90).

Although details remain to be elucidated, both the back-to-back and front-to-front interfaces have to be intact for full activation of PknB. Whereas this constitutes a homotypic interaction pathway to activation, there also is evidence for cross-phosphorylation between different STPKs of *Mtb* which can directly lead to activation and links signalling pathways (93). Furthermore, several additional phospho-acceptor sites, other than in the activation loop, are

phosphorylated in the *Mtb* STPKs. These may serve as docking sites for substrates recruitment (94).

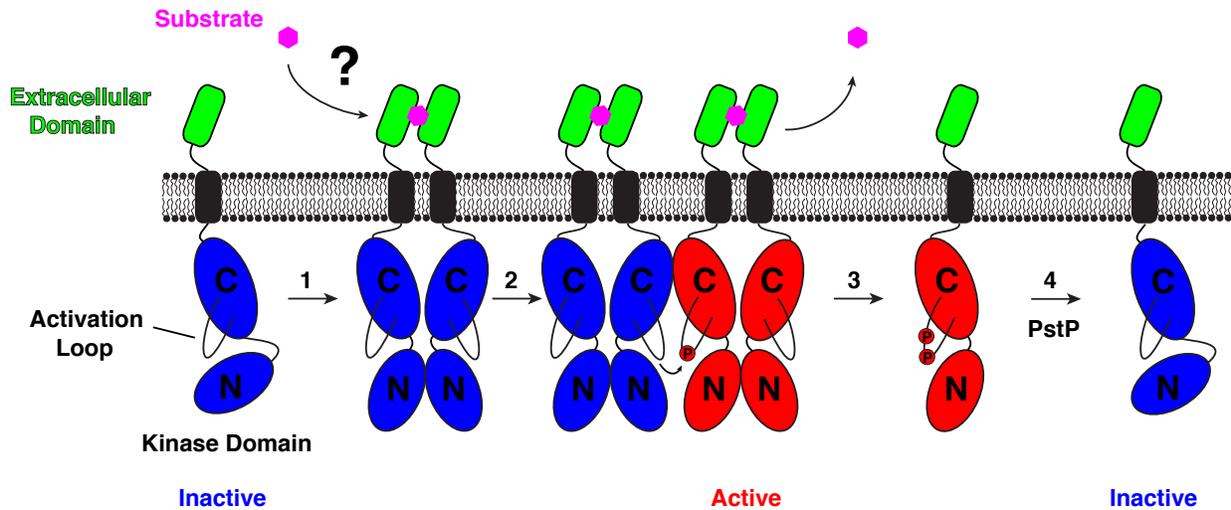


Figure 1.6: Model of the Activation Mechanism of PknB.

Cartoon representation of the STPK with the extracellular domain (green), the transmembrane domain (black) and the N- and C-terminal lobes of the kinase domain (blue: inactive, red: active) indicated. A hypothetical substrate is in magenta. 1) Possibly aided by binding of a substrate and dimerization of extracellular sensory domains, the intracellular kinase domains dimerize in a back-to-back fashion. 2) This dimerization allows trans-phosphorylation of the activation loop of another PknB homodimer. 3) Once phosphorylated, the kinase remains active even in a monomeric state. 4) Inactivation is achieved through dephosphorylation by the serine threonine phosphatase PstP.

1.2.1.3 Specificity of Mycobacterial Serine Threonine Phosphorylation

In *Mtb*, 500 phosphorylation sites in more than 300 proteins with functions in all regimes of its physiology have been identified via tandem mass spectrometry phosphoproteomics studies (87). Additional STPK target proteins were identified through *in vitro* studies with KD constructs (69). In humans, serine is more often phosphorylated than threonine (9:1), whereas phospho-threonine dominates in *Mtb* (1.5:1). Generally, the STPK substrate motif with threonine at the centre shows preference for N-terminal Asp/Glu residues at the -2 and -3 positions and C-terminal hydrophobic residues at the +3 and less so at the +5 site (69).

1.2.1.4 Biological Roles of *Mycobacterium tuberculosis* STPKs

A summary of the roles of the *Mtb* STPKs is given in Table 1. Of the 11 kinases, PknA and PknB are the only two essential for *Mtb* growth. Their encoding genes are found adjacent to each other in the same operon that includes other genes important for cell wall synthesis and controlling cell-shape (95). In accordance with their genomic location, their overexpression leads to abnormalities in cell shape such as branching, elongation and incomplete septation (86). Furthermore, several proteins directly involved in cell wall and peptidoglycan (PG) synthesis and linked to cell morphology are phosphorylated by these kinases (22, 96, 97). The four extracellular penicillin binding protein and serine/threonine kinase associated (PASTA) domains of PknB bind muropeptides, fragments of PG. The PASTA domains can thereby sense changes in the state of the cell wall and relay that information to regulate activity of the kinase (98).

PknD is a non-essential STPK in *Mtb* for which the encoding gene is found in a gene cluster associated with phosphate uptake (95, 99). Consistently, the kinase plays a role in phosphate transport (100). *Mtb* that is *pknD*-deficient cannot invade the central nervous system as seen by an abolished infectivity for brain endothelial cells *in vitro* (101). Based on *in vitro* phosphorylation studies of phosphoacceptor proteins, PknE has been implied to play a role in the type II fatty acid synthase system (102). Links to the regulation of apoptosis in macrophages has also been suggested (103). *PknF* is in the same operon as the gene for a characterized PknF phospho-target, the ABC transporter Rv1747. This transporter is implicated in exporting cell-wall components, such as phosphatidyl-*myo*-inositol mannosides (PIMs) (104, 105), and will be discussed in detail below. PknH can phosphorylate the transcriptional regulator EmbR (106), and

thereby regulate expression of the *embCAB* operon (107). This in turn controls the PG composition of the cell wall (107-110). Additionally, PknH controls the expression of genes of the dormancy operon under hypoxic and nitric oxide stresses (111). The structure of the extracellular domain of PknH forms a v-shaped lipoprotein-like cleft that is presumed to be a substrate binding site (112). However, the exact role of PknH in virulence remains unclear as the Δ *pknH* strain is more susceptible to peroxide and oxide stresses and shows increased virulence in macrophages while being less susceptible to nitric oxide and hypovirulent in mice (110). Relatively little is known about PknI, PknJ and PknL. While a PknI deficient *Mtb* strain has been shown to be hypervirulent in macrophages and mice (113), a transposon insertion in *pknJ* did not have any phenotypic effects. PknJ has been shown to dimerize (114) and both PknJ and PknL are reported to be autophosphorylated in their respective activation loops (114, 115), in line with the activation mechanism common for *Mtb* STPKs.

PknG and PknK are the two *Mtb* STPKs that do not possess a membrane spanning domain and their KDs differ significantly in sequence from the ones of the directly membrane bound kinases. PknG is a virulence factor that is secreted into the macrophage phagosome and plays a role in preventing phagosome-lysosome fusion (116, 117). It also functions intracellularly in glutamate metabolism (118). In addition to the KD, PknG possesses a C-terminal TPR that helps dimerization (80) and an N-terminal rubredoxin domain that is important for kinase activity (88, 89). PknG also shows notable differences to other *Mtb* kinases as activation loop phosphorylation is not required for activation (80). PknK has been implied to be important in virulence due to its involvement in the regulation of translation and the adaption to growth and environmental conditions (119, 120) as well as controlling the mycobacterial monooxygenase (*mym*) operon via phosphorylation of VirS (121). Its region C-terminal of the KD shows similarity to the ATP-, maltotriose- and DNA-dependent transcriptional regulator MalT (69) and a steric mechanism of PknK regulation by this C-terminal region has been proposed (119).

Table 1.1: Summary of Biological Roles of *Mtb* STPKs.

STPK	Importance for Virulence	Notable Genomic Location	Biological Role	Accessory Domains and their Function
PknA	Essential (69)	In operon adjacent to <i>pknB</i>	Cell wall and peptidoglycan synthesis (86), lipid synthesis, gene regulation (69)	Extracellular domain of unknown structure and function
PknB	Essential (69)	In operon adjacent to <i>pknA</i>	Cell wall and peptidoglycan synthesis (86, 96, 97), oxygen triggered revival of dormant <i>Mtb</i> (86, 122), lipid synthesis, gene regulation (69)	Four extracellular PASTA domains that bind peptidoglycan fragments and sense cell wall changes (98, 123)
PknD	$\Delta pknD$ is impaired in central nervous system invasion (101)	In gene cluster associated with phosphate uptake (95, 99)	Phosphate transport (100)	Extracellular six-winged beta propeller that restores CNS infectivity by itself, forms cup structure, binding partner unknown (124)
PknE	-	Separated from <i>pknF</i> by only 2 kb	Type II fatty acid synthase system (102), regulation of apoptosis in macrophages (103)	Extracellular domain of unknown structure and function
PknF	-	In same operon as <i>rv1747</i>	Transport of cell wall components like PIMs via Rv1747 (104, 105)	Extracellular domain of unknown structure and function
PknG	$\Delta pknG$ is hypovirulent (116, 117)	In operon with glutamine-binding lipoprotein <i>glnH</i>	Secreted in macrophage phagosome, roles in preventing phagosome-lysosome fusion (116, 117) and in regulation of glutamate metabolism via GarA (118)	C-terminal TPR that are important dimerization interface (80) N-terminal rubredoxin domain: iron binding important for PknG activity (88, 89)
PknH	$\Delta pknH$ is more susceptible to oxide and peroxide and hypovirulent in macrophages but less susceptible to NO and hypervirulent in mice (110)	Next to transcriptional regulator <i>embR</i>	Activates transcription of <i>embCAB</i> operon via phosphorylation of <i>embR</i> (106), thereby increasing lipoarabinomannan in cell envelope (107), regulation of dormancy operon via phosphorylation of DosR (111)	Structure with V-shaped lipoprotein-like cleft that is proposed substrate binding site (112)
PknI	$\Delta pknI$ is hypervirulent (113)	-	Unknown	Extracellular domain of unknown structure and function
PknJ	-	-	Unknown	Extracellular domain of unknown structure and function
PknK	Increased survival of $\Delta pknK$ during persistent infection in mice (120)	Close to <i>virS</i> and <i>mym</i> -operon	Regulation of translation and the adaption to growth and environmental conditions (119, 120), activation of VirS-dependent gene regulation of <i>mym</i> operon (121)	C-terminal domain shows similarity to the ATP-, maltotriose- and DNA-dependent transcriptional regulator MalT (69), proposed steric regulation of KD activity (119)
PknL	-	-	Unknown	-

1.2.2 The Serine Threonine Phosphatase PstP

The sole serine threonine phosphatase found in *Mtb* is PstP. Not surprisingly, this "eraser" enzyme for the 11 "writer" STPKs has been shown to be essential for cell survival and proper cell division (125). PstP is comprised of an intracellular metal-dependent protein phosphatase (PPM) family domain that belongs to the PP2C class and strictly requires Mn^{2+} ions for activity (126). However, in the structure of the PstP phosphatase domain determined by X-ray crystallography, three metal binding sites were identified instead of two found in most other PP2C structures (127, 128). PstP also has a transmembrane region and an extracellular domain. The gene encoding for PstP is found in an operon with genes for PknA and PknB and it is phosphorylated by both STPKs. This leads to an increased activity of PstP (129), possibly acting in a feedback mechanism for regulation of phosphorylation levels.

1.2.3 Tyrosine Phosphorylation in *Mycobacterium tuberculosis*

In addition to the STPKs, the *Mtb* genome also encodes for one tyrosine kinase (PtkA) and two tyrosine phosphatases (PtpA and PtpB). PtpA and PtkA are located in the same operon and have both been identified as factors important for *Mtb* pathogenesis (26, 130). Both are secreted (131) and PtpA interacts with the vacuolar ATPase complex in macrophages to block its transport to the mycobacterial phagosome and thereby inhibit its acidification and the lysosome-phagosome fusion (26, 132). PtkA has been shown to phosphorylate PtpA (133), thereby enhancing its activity (134) and contributing to virulence in a concerted way with the phosphatase.

1.2.4 Forkhead Associated Domains and their interplay with *Mtb* STPKs

Forkhead Associated (FHA) domains are protein domains, found across all kingdoms of life, that specifically bind phosphothreonine-containing polypeptides and thereby selectively mediate protein-protein interaction networks. The structures and functions of FHA domains have been extensively reviewed (135-138) along with a detailed discussion of the mechanisms underlying their specificity for phosphothreonine (139). The core FHA domain structure consists of 11 β -strands assembled into a two-sheet β -sandwich. The pThr interaction interface is a shallow cleft crossing one of the loop regions connecting the β -sheets (135) (Figure 1.7A). The

Mtb genome encodes six FHA domain-containing proteins (Figure 1.7B) that are closely linked with its regulatory STPKs and PstP.

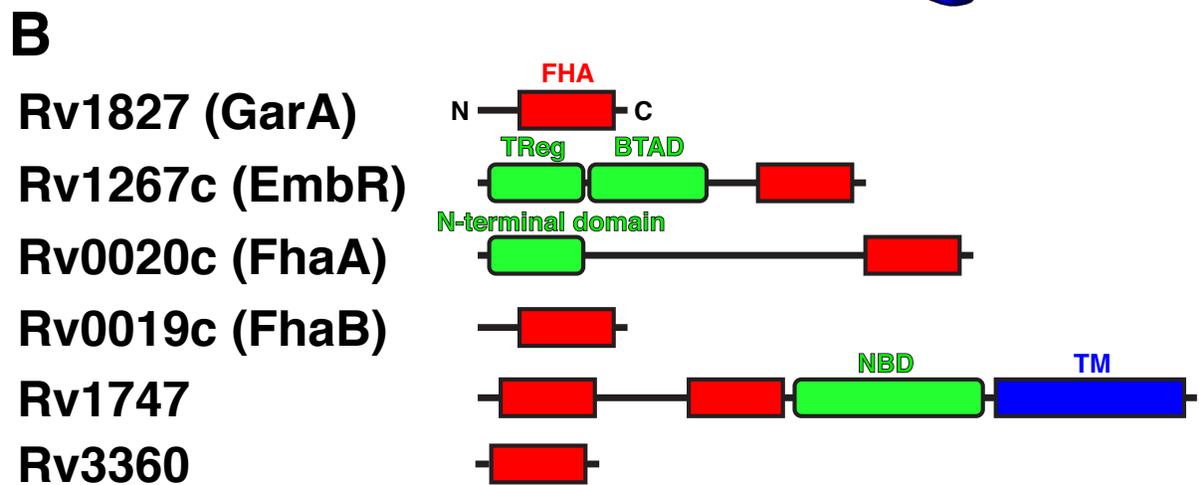
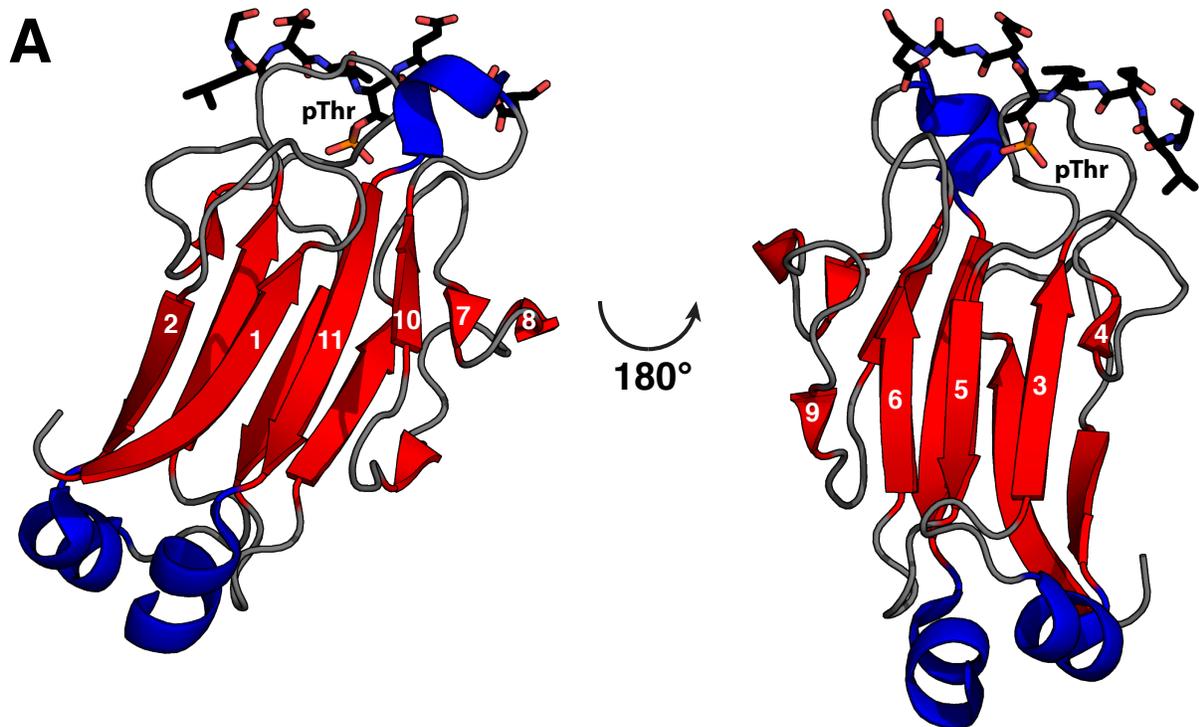


Figure 1.7: The FHA Domain and its Occurrence in *Mtb* Proteins.

(A) Canonical FHA domain structure as exemplified by Rad53 FHA-1 in complex with a phosphothreonine peptide (PDB ID: 1G6G). The β -strands of the conserved core FHA domain structure are shown in red and their conventional numbering is indicated. Non-conserved helices are in blue. (B) Schematic representation of the six *Mtb* proteins containing one or more FHA domains (red). Transmembrane domains (TM) are in blue and accessory domains in green with their identities indicated.

One well studied *Mtb* protein that contains an FHA domain is GarA (Rv1827). GarA also has an N-terminal intrinsically disordered region (IDR) with an identified phospho-acceptor threonine for PknG. In addition to pThr binding, the FHA domain of GarA can interact with and thereby regulate the activity of three enzymes found in central metabolic pathways (α -ketoglutarate dehydrogenase, NAD-dependent glutamate dehydrogenase, α -subunit of the glutamate synthase complex). Upon phosphorylation, the IDR of GarA is proposed to bind in an intramolecular fashion to its FHA domain and thereby occlude the binding interfaces for the enzymatic partners. This leads to an elegant switch-like auto-inhibitory mechanism by which GarA helps regulate a key aspect of *Mtb* metabolism (118).

The transcriptional regulator EmbR (Rv1267c) binds and controls expression of the *embCAB* operon that is important for arabinogalactan biosynthesis (106). In addition to an N-terminal winged-helix DNA-binding domain that forms a tight complex with a bacterial transcriptional activation domain, EmbR has a C-terminal FHA domain that binds to a phosphorylated threonine in the *Mtb* STPK PknH (140). This FHA domain thus directs the kinase to phosphorylate and activate the transcriptional regulatory domain of EmbR.

The genes encoding the two FHA-containing proteins FhaA (Rv0020c) and FhaB (Rv0019c) are found next to each other. Additionally, they are close to the genes for PknA, PknB, PstP, as well as proteins associated with cell wall biosynthesis. Both FHA proteins have been linked functionally to cell wall integrity and PG synthesis (141). FhaA has a C-terminal FHA domain connected to an N-terminal structured domain of unknown function by an ~ 300 residue IDR. The FHA domain of FhaA interacts with a phosphothreonine in the juxtamembrane domain of PknB, which in turn phosphorylates FhaA (142). This is another example of an FHA-pThr docking interaction between an FHA-containing substrate and the STPK to direct phosphorylation specificity. Less characterized, FhaB only has an FHA domain with a shorter N-terminal IDR. It also interacts with and is phosphorylated by PknA and PknB.

The only *Mtb* protein containing two FHA domains is the putative ABC transporter Rv1747. Although its exact functions remain to be defined, Rv1747 is proposed to play a role in the transport of cell wall biosynthesis intermediates (104). As will be discussed in detail below, Rv1747 has the core topology expected for a homodimeric ABC transporter with a Nucleotide Binding Domain (NBD), the location of ATP hydrolysis, followed by a transmembrane portion

consisting of six helices. Additionally, it contains a putative N-terminal regulatory module with two FHA domains (FHA-1 and FHA-2) connected by an ~ 100 amino acid IDR. Rv1747 is phosphorylated by PknF and to a lesser extent by PknE (*143* , *144* , *145*). For PknF, two phospho-acceptor threonines have been identified in the linker connecting the two FHA domains (*144*). Rv1747 is a virulence factor in macrophage and mice infections and phosphorylation at these mapped sites, as well as the integrity of the binding site of FHA domains in the regulatory module, is crucial for this activity (*144*). Recently, the universal stress protein Rv2623 was shown to interact via a phosphorylated threonine with FHA-1 of Rv1747 (*105*). Based on the hypervirulence of a $\Delta Rv2623$ strain, this interaction was proposed to have an inhibitory effect on the transporter activity of Rv1747.

The sixth *Mtb* protein predicted to contain an FHA domain is Rv3360. Its function and binding partners, however, remain to be elucidated.

1.3 Phase Separation of Biological Macromolecules

1.3.1 Phase Separation in Biology

Compartmentalization is a key feature of cellular life, allowing for the coexistence of different functionally specialized environments in a spatially separated form. Canonical cellular compartmentalization is achieved by the formation of membrane-enclosed vesicle-like organelles in which a phospholipid bilayer, inherently impermeable to larger polar molecules and macromolecules, constitutes a physical barrier between two distinct biological environments. To allow for exchange of molecules across this barrier, either passive transport along a concentration gradient or active transport against a concentration gradient typically occurs across the membrane via membrane-spanning transport proteins. Examples of classical membrane enclosed organelles are mitochondria and chloroplasts, both believed to have originated from an endosymbiotic origin. The Golgi apparatus, the endoplasmatic reticulum and, of course, the nucleus are additional examples.

Compartmentalization can not only be achieved by physical separation via a lipid bilayer, but also by the spontaneous formation of protein- and nucleic acid-rich "membraneless" organelles based on the phenomenon of phase separation. Here, the biological polymers phase separate above a threshold concentration into droplets that are highly polymer-enriched. Alternatively the phenomenon is referred to as liquid phase condensation. A summary of identified eukaryotic membraneless organelles is shown in Figure 1.8. Ribonucleoprotein (RNP) bodies/granules, including nucleoli, Cajal bodies, nuclear speckles, processing (P) bodies, P granules and stress granules are a prominent class of phase separated systems found in the nucleus and cytoplasm. Other RNA-independent systems include Dv1 puncta (146, 147), promyelocytic leukemia (PML) bodies (148, 149) and condensates of tau acting as nucleation sites for microtubule bundles (150, 151). T-cell signalling clusters (152, 153) are an interesting case of membrane-associated, pseudo two-dimensional organelles.

Many membraneless organelles have the same highly dynamic, liquid behaviour as seen for classical phase separating systems *in vitro*. Their spontaneous formation is based on a constantly re-shaped network of weak, multivalent interactions between a large number of molecules. The edge of membraneless organelles constitutes a sharp boundary separating the two distinct phases. In contrast to membrane-enclosed organelles, this boundary is not an

impermeable barrier and diffusive exchange of molecules between the surrounding phase and the body of the organelle can readily occur. Due to their liquid behaviour and high density, membraneless organelles may have to be stabilized against sedimentation inside cells. This is achieved in the nucleus by an actin scaffold (154).

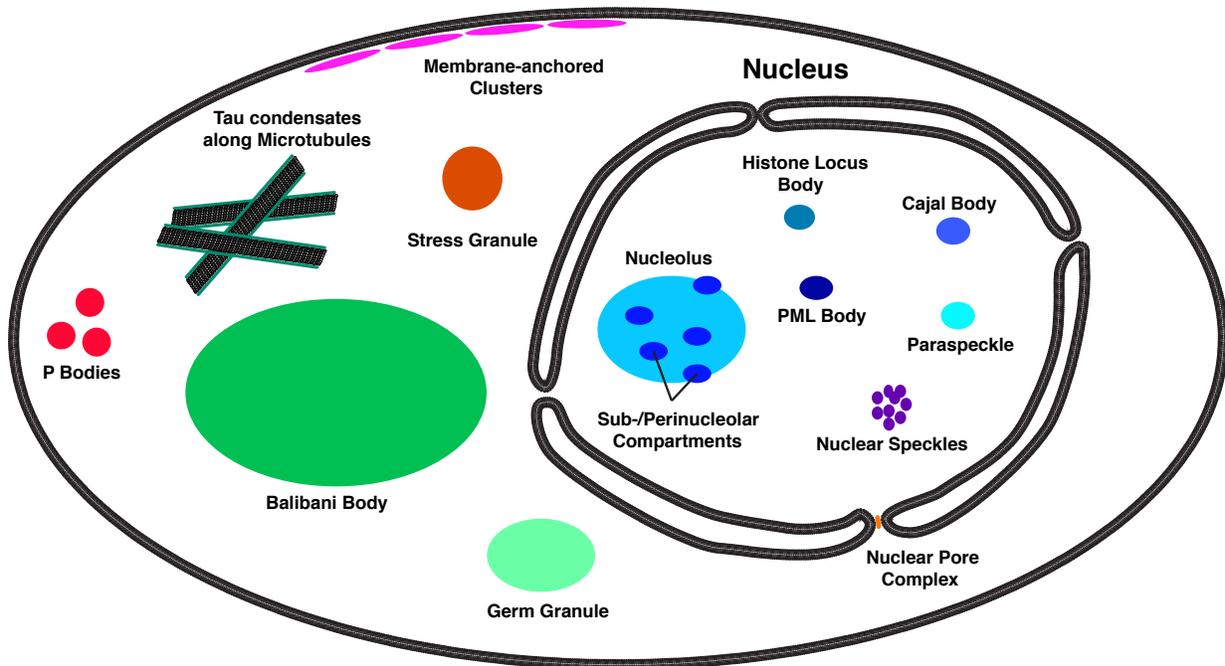


Figure 1.8: Membraneless Organelles in Eukaryotic Cells.

Schematic of representative membraneless organelles and their typical locations in eukaryotic cells. Some organelles are specific to certain cell types (e.g. germ granules in germ cells), but all are depicted for completeness. Sizes are not to scale.

1.3.1.1 Examples of Membraneless Organelles

Germ granules are a class of heterogeneous membraneless RNP organelles (155) that are found in the germ cells of sexually reproducing organisms and play an important role in cell polarization (156). In *Caenorhabditis elegans* one-cell embryos, they are termed P granules. In mammalian cells, they are referred to as nuage bodies and contain RNAi components that help prevent the activity of transposable elements in spermatocytes and spermatids (157). Germ granules exhibit dynamic, liquid-like behaviors, such as fusion, dripping and wetting (158). While driven by mRNA-enhanced phase separation, droplet formation also occurs *in vitro* by some of the major protein constituents of germ granules, including the RNA helicase LAF-1, PGL-3 and MEG-3 (156, 159, 160). DDX4, a close relative of LAF-1 (161), is central to phase separation of nuage bodies (162).

Another example of well-studied intracellular condensates that show liquid-like behaviour and form above a certain concentration threshold of a key component are stress granules (163). They were discovered as intracellular granules that form upon exposure of the cells to stresses such as heat (164). Recent data suggests that these granules act as a place of dynamic sorting for translating messenger ribonucleoproteins (mRNP) complexes stalled in the process of initiation during periods of cellular stresses (165, 166). The stalled mRNAs are either degraded, stored or further translated. When overexpressed, several key components of stress granules such as G3BP (167), CPEB1 (168) and TIA-1 (169) have been shown to be sufficient to induce stress granule formation even in the absence of a cellular stress. This indicates a concentration threshold for their formation, just as found for phase separating systems *in vitro*. Proteins such as FUS (170-172), hnRNPA1 (163) and TDP43 (172) that are found in stress granules and play an important role in amyotrophic lateral sclerosis (ALS), also undergo phase separation *in vitro*. For these systems, a metastability of the liquid-like state and the transition to less dynamic, more gel/glass-like states is important in their connection to pathology (163, 170, 172-176).

Another protein linked to neurodegenerative diseases that has recently been shown to form condensates under conditions of molecular crowding *in vitro* is the microtubule associated protein (MAP) tau (151). Tau droplets may serve as a non-centrosomal nucleation site of microtubule formation due to the increased concentration of tubulin, which associates with tau,

inside the condensed phase (150). Furthermore, condensed tau stays present along the microtubule bundles, possibly acting to provide stability and elasticity to microtubules in axons (150).

It is well known that intracellular membraneless bodies are often comprised of a heterogeneous set of different proteins and RNA (155, 177, 178). Strikingly, these bodies, including nucleoli (179), stress granules (177), P bodies (180), P granules (181, 182) paraspeckles (183) and PML bodies (149) also show substructures different from the homogeneous, spherical morphologies found for homogeneous phase-separating systems *in vitro*. The most comprehensively studied example is the nucleolus, which exhibits a core-shell structure that is important in sequential processing of ribosomal rRNA. The rRNA is thought to mature as it moves through different regions of different composition in the nucleolus (179). There is evidence that the self-assembly of this core-shell structure is driven by the immiscibility of the different phases involved and the different relative surface tensions between these phases (184, 185). The latter is a well-known phenomenon in mixtures of organic solvents (186).

1.3.1.2 Phase Separation and Membrane Protein Clustering

It is now well established that soluble proteins undergo phase separation to form droplet-like membraneless organelles in cells. There is also evidence that phase separation can be the underlying mechanism for clustering of membrane-associated components. This is seen for a wide variety of membrane proteins including signalling receptors (187-189), cadherins (190) and GPI anchored proteins (191). One well-studied cellular process that employs phase separation is the clustering of components of the T-cell receptor (TCR) signalling complex (153, 192). Here, LAT (linker for activation of T-cells), with its transmembrane spanning module, acts as an anchor in the membrane. The cytoplasmic domain of LAT is phosphorylated by the T-cell receptor (CD3 ζ)-activated kinase ZAP70. The phosphorylated tyrosine motif is in turn recognized by a SH2 domain in Grb2, which also possesses multiple SH3 domains that can interact with several proline rich motifs in Sos1. This provides multivalency necessary for phase separation and cluster formation. Furthermore, when anchored in a membrane *in vitro* via phosphorylated LAT (pLAT), these clusters possess liquid-like behaviour similar to droplets found in solution (153, 193). Intriguingly it has been shown that if the downstream signalling

proteins Nck, N-WASP, Arp2/3 and actin are added, actin filaments nucleate from the phase-separated clusters while the shape of the clusters is changed to align with the filaments. This points to an important role of LAT clustering as a means of promoting actin filament nucleation by an increased concentration inside the protein-rich cluster.

1.3.2 Biophysical Characterization of Phase Separation

1.3.2.1 Basic Characteristics of Biological Condensates

Although many of the biological phase-separating systems described are highly heterogeneous, *in vitro* one or two purified key factors are often sufficient to form droplets with similar properties as those formed *in vivo*. This has been instrumental for understanding the underlying principles governing biological phase separation.

Many protein/RNA condensates have a liquid character as is evident from a few key characteristics (194) that we intuitively associate with simple systems such as a well-shaken oil and vinegar salad dressing. As exemplified in Figure 1.9, they tend to adopt a more-or-less spherical shape, which is a hallmark of liquids due to surface tension. They often coalesce, referring to the fusion of two touching droplets and the following reversion to a larger spherical shape. Furthermore, shearing, dripping and wetting behaviour can be seen for the droplets, as well as internal rearrangement, exchange across the phase boundary, and the resulting growth by Ostwald ripening.

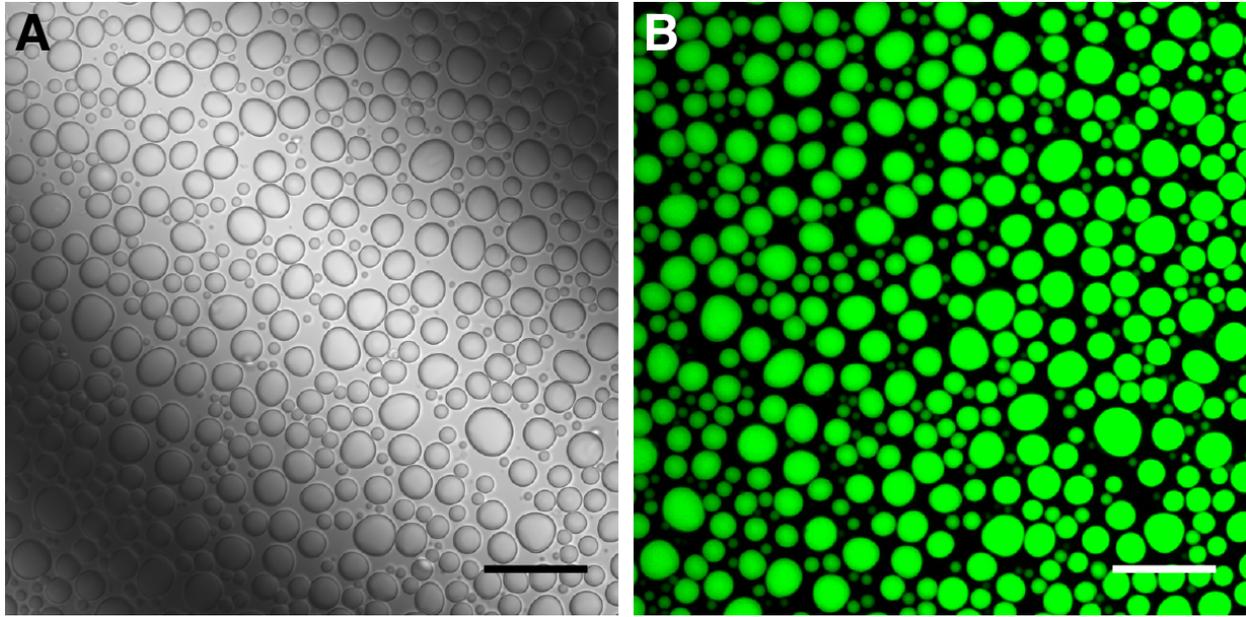


Figure 1.9: An Example of Phase-Separated Protein Droplets.

(A) Differential interference contrast (DIC) and (B) fluorescent microscopy images of the phase-separated regulatory module from the *Mtb* transporter Rv1747 (residues 1-310). The phosphorylated protein is covalently labeled at an N-terminal cysteine with an Oregon Green fluorophore (scale bar: 20 μm).

1.3.2.2 Flory Huggins Theoretical Description of Phase Separation of a Polymer

To understand the driving forces behind protein/RNA condensate formation, a basic review of the physics of polymer phase separation is necessary. This can be done with the help of a simple lattice model of a polymer-solvent mixture. Consider an infinite lattice with coordination number z , indicating the number of interacting neighbors. Each lattice point is filled with either a low molecular weight solvent (s) particle or a monomer subunit of a homopolymer (p) of length N (Figures 1.8A and B). The Flory-Huggins “mean-field” free energy G of mixing per lattice site is comprised of a mixing entropy component S and an enthalpic part H , taking account of the interactions between particles (195, 196).

$$G = H - TS$$

In a system with no interaction between particles in the lattice, $H = 0$ and therefore $G = -TS$. The volume fractions of solvent and polymer are denoted as ϕ_s and ϕ_p , respectively. Since $\phi_s + \phi_p = 1$, it follows that $\phi_s = 1 - \phi_p$ (Figure 1.10C). The mixing entropy per lattice site S can then be described as:

$$\frac{S}{k_B} = -\frac{\phi_p}{N} \ln \phi_p - \phi_s \ln \phi_s = -\frac{\phi_p}{N} \ln \phi_p - (1 - \phi_p) \ln (1 - \phi_p)$$

In the unmixed states for $\phi_p = 0$ and $\phi_p = 1$, the entropy vanishes to $S = 0$. In contrast, the formation of mixed states ($0 < \phi_p < 1$) is accompanied by an increase in entropy (Figure 1.10D). Favorable mixing thus lowers the free energy, yielding a concave function shown in black in Figure 1.10E.

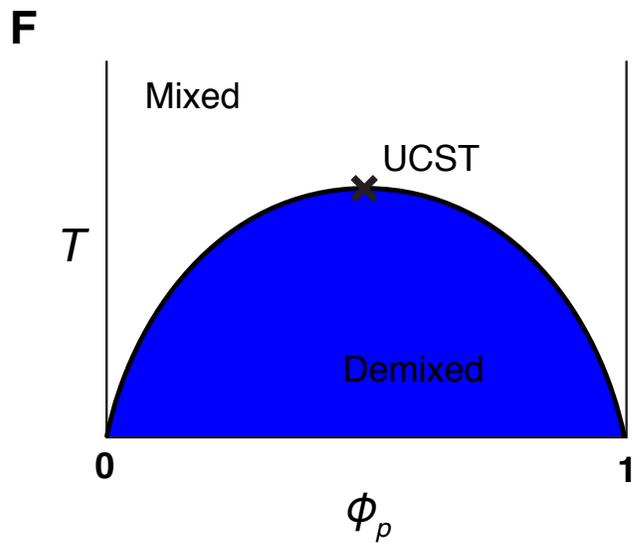
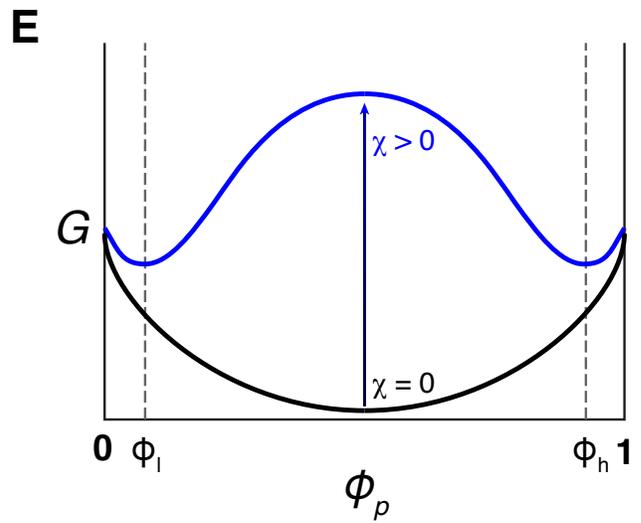
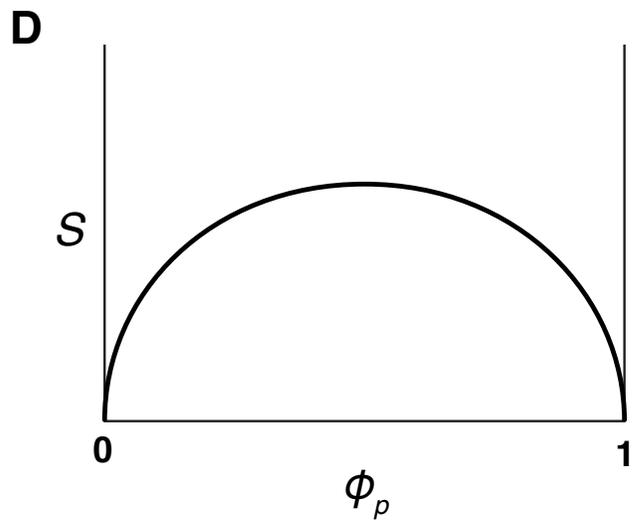
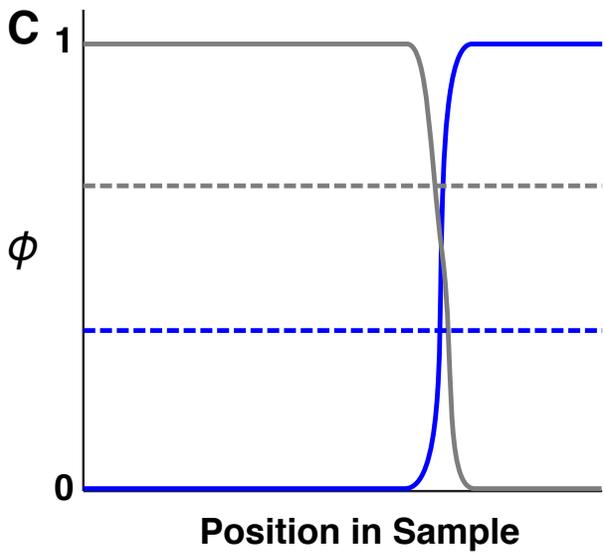
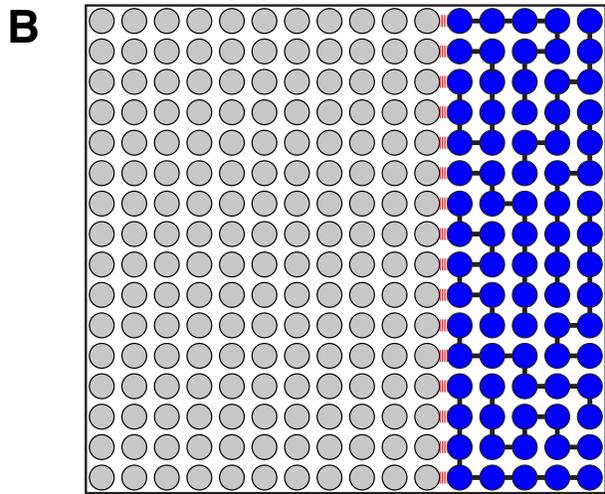
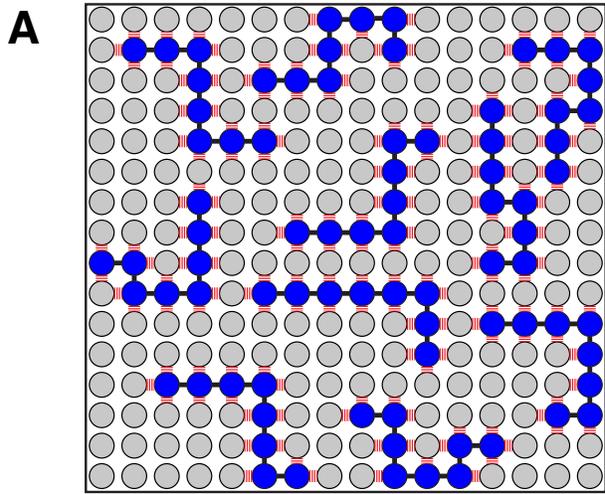


Figure 1.10: Flory-Huggins Theory of Polymer Phase Separation.

Lattice model of a polymer (blue particles) that is (A) mixed or (B) demixed with solvent (grey particles). Polymer-solvent interactions are depicted in red. (C) Volume fraction, ϕ of polymer (blue) and solvent (grey) across one dimension of a sample in mixed (dashed lines) and demixed (solid lines) states. (D) Dependence of mixing entropy, S , on volume fraction of polymer, ϕ_p . For simplicity the polymer length N is set to 1. (E) Dependence of free energy of mixing, G on volume fraction of polymer, ϕ_p . For simplicity the polymer length N is set to 1. Functions are shown for a Flory-Huggins parameter favouring mixing ($\chi = 0$, black) and favouring demixing ($\chi > 0$, blue). Volume fractions in polymer-depleted (ϕ_l) and polymer-rich phases (ϕ_h) are indicated. These are local minima for the free energy of mixing in a phase separating system. (F) Schematic phase diagram of phase-separating polymer. The demixed, two-phase region is shown in blue and the upper critical solution temperature (UCST) above which the system is always mixed is indicated.

To take interactions between particles into account, the mean interaction enthalpy term per lattice site H can be defined as

$$\frac{H}{k_B T} = \chi \phi_p (1 - \phi_p)$$

Here, χ is the Flory parameter that quantifies the balance of mean interaction energies per lattice site of solvent-solvent (u_{ss}), polymer-solvent (u_{ps}) and polymer-polymer (u_{pp}) interactions as

$$\chi = z \left[u_{ps} - \frac{1}{2} (u_{pp} + u_{ss}) \right]$$

The magnitude of the Flory parameter quantifies the energetic benefit/cost of solvation. A value of $\chi > 0$ indicates a poor solvent where polymer-solvent interactions are disfavoured compared to polymer-polymer and solvent-solvent interactions. Above a threshold value for χ , the energetic cost of solvation will overcome the entropic benefit of mixing and the free energy curve will have a region of negative curvature (blue function in Figure 1.10E). This leads to the formation of two distinct phases with high and low polymer volume fractions of ϕ_h and ϕ_l , respectively (Figure 1.10E and F).

The driving mechanism for a particle flux that is lowering the free energy in this model system is diffusive exchange. Here, the rate of transport J is proportional to the gradient in chemical potential, $J \propto \frac{d\mu}{dx}$. The chemical potential μ can be expressed as a function of the free energy G as

$$\mu = \frac{v}{V} \frac{dG}{d\phi} = \frac{v}{V} \left(\frac{dH}{d\phi} - T \frac{dS}{d\phi} \right)$$

and the chemical potential gradient $\frac{d\mu}{dx}$ as

$$\frac{d\mu}{dx} = \frac{v}{V} \left(\frac{d^2H}{d\phi^2} - T \frac{d^2S}{d\phi^2} \right) \frac{d\phi}{dx}$$

where v is the molecular volume of the particle and V is the total volume of the system. For the rate of transport, it follows that

$$J \propto -\frac{v}{V} \left(\frac{d^2H}{d\phi^2} - T \frac{d^2S}{d\phi^2} \right) \frac{d\phi}{dx}$$

Here, the sign of the rate of transport J , and therefore the direction of diffusion and whether or not a system is prone to phase separation, is again dependent on the balance between interaction enthalpic and entropic contributions. Because

$$\frac{d^2H}{d\phi^2} - T \frac{d^2S}{d\phi^2} = \frac{d^2G}{d\phi^2}$$

the direction of flux J is dependent on whether, at the volume fraction of interest ϕ , the free energy function has a convex or a concave curvature, promoting mixing and phase separation of solvent and polymer, respectively.

An important factor has not been taken into account here is the dampening of the mixing entropy contribution to the overall free energy of mixing. This due to the connectivity between the polymer beads, which effectively makes polymeric solutions more prone to phase separation (197, 198).

1.3.2.3 Molecular Basis of Biological Phase Separation

One way to increase favourable inter-polymer interactions and thereby promote phase separation is multivalency. The multivalent interactions that promote phase separation of biological systems have been shown to involve both folded domains and IDRs of proteins. In either case, intrinsic disorder is often a key factor in phase separating systems as IDRs also serve as flexible linkers between repeated domains in modular proteins. Furthermore, protein-nucleic acid, as well as solely nucleic acid- or protein-based interactions, can be the driving force behind phase separation of biological macromolecules.

1.3.2.3.1 Interactions of Multivalent Proteins of Modular Architecture

The degree of multivalency is a key factor for phase separation in model systems such as single chain repeats of folded SH3 domains interacting with repeats of proline-rich motifs (PRMs) (Figure 1.11A). In a ground-breaking study by the Rosen group (*152*), the propensity for phase separation was found to be highly dependent on the number of repeats of SH3 domains and PRMs in the protein constructs. A minimum of 3 repeats of each modules was required to observe droplet formation. A higher number of repeats correlated with a lower threshold concentration for demixing (*152*). Additional interactions between the IDRs linking the SH3 domains were subsequently found to also contribute to phase separation (*199*). Confirming this synergy of interactions of different nature, specific modular interactions have recently been identified to be complemented by non-specific IDR interactions to cause phase separation in stress granules (*200*).

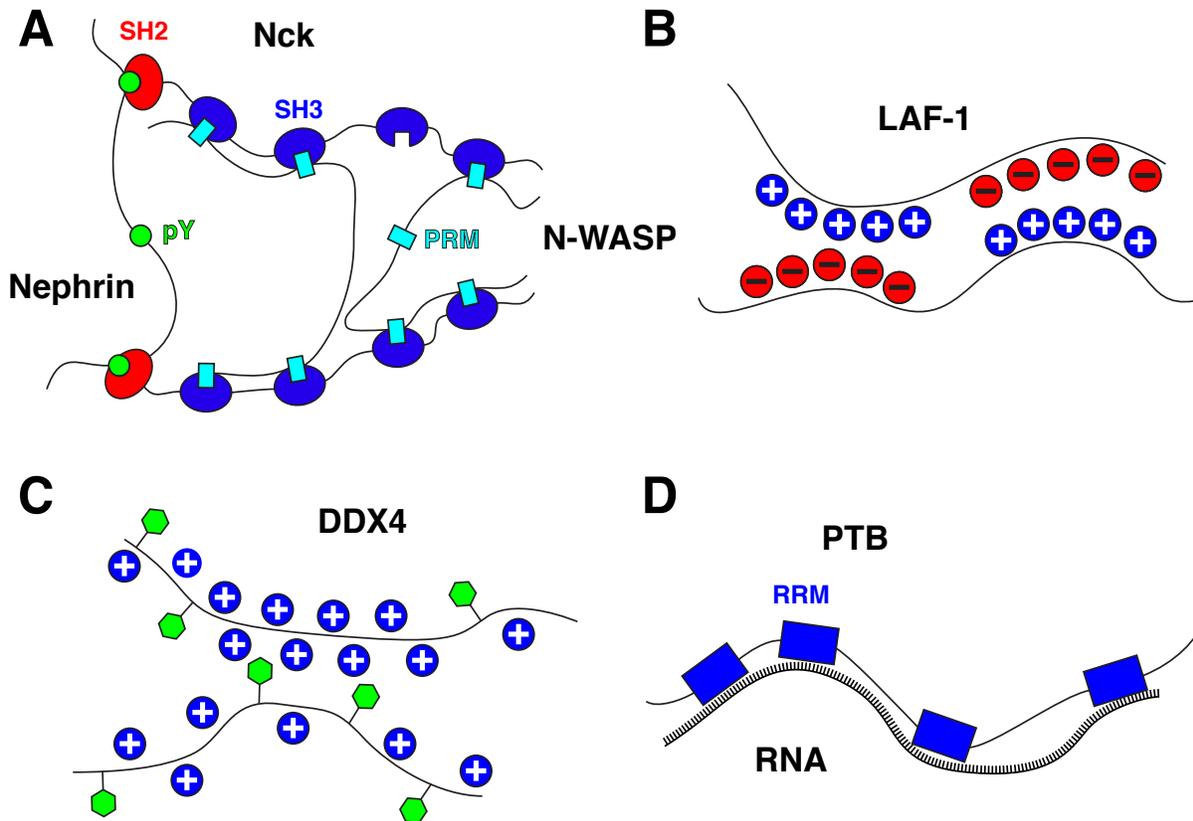


Figure 1.11: Examples of Interactions Found in Biomolecular Phase Separation.

(A) Multivalent modular interactions, based on folded domains such as SH2/SH3 and their binding partners pY/PRM, is seen for associations between Nephrin, Nck and N-WASP in T-cell receptor signalling. (B) Phase separation of the polyampholytic P granule protein LAF-1 is based on block charge-charge interactions. (C) π -cation interactions between aromatic residues, such as phenylalanine (green hexagons), and arginines play a role in the phase separation of the germ granule protein DDX4. (D) RNA also often serves as a source of multivalency and is therefore important in phase separation with the polypyrimidine tract-binding protein (PTB) that itself harbours four RNA recognition motifs (RRMs).

Another example of modular, multivalent interactions driving phase separation are repeats of SUMO (small ubiquitin-like modifier) interacting with repeating units of the SUMO interacting motif (SIM). These are both constituents of numerous cellular condensates including PML bodies (148). Unbound SUMO/SIM modules that are part of the repeating chains within the droplets can also recruit monomeric, non-phase separating constructs into the condensed phase. Which of the monomers is recruited depends on which of the modules is present in excess in the network forming the droplet. This is controlled by the relative concentration of SUMO/SIM repeat constructs. The process is reversible and a change of concentration in these droplet “scaffold” proteins leads to a change in which of the monomeric species is recruited (201).

1.3.2.3.2 Interactions Based on Intrinsically Disordered Regions in Proteins

Interactions between IDRs in proteins that drive droplet formation involve low-complexity sequences. Here, multivalency is often brought about by repetitive stretches enriched in amino acids with small and/or polar and charged side chains such as Gly, Glu, Gln, Asn, Ser, Pro, Lys and Arg (202, 203). Whereas hydrophobic side chains that promote protein folding are generally rare in IDRs, sequences driving phase separation often contain the aromatic residues Tyr and Phe. Based on the presence of these residues in the low-complexity IDRs, the weak, transient interactions driving phase separation can be of charge-charge, cation- π , dipole-dipole and π - π character (197). The salt dependence of the strength of the interaction decreases in this order. Charge-charge interactions provide a source of long range interactions. In contrast, dipole-dipole and interactions involving aromatic and non-aromatic side chains that contain π -orbitals are shorter-range and are thought to provide more specific points of contact. As the physicochemical basis of many biological phase separating systems is being investigated, efforts have been made to link phase separation behaviour of IDRs to features of their amino acid sequence (204). The eventual goal is to predict phase behaviour based on sequence. Advances in this regard have been made for sequences capable of π -interactions (205).

Systems where droplet formation is sensitive to salt concentration are exemplified by the RNA helicases DDX4 (206) and its close relative LAF-1 (159) (Figure 1.11B). In these cases, charge-charge interactions play an important role in phase separation. DDX4 is found in nuage bodies whereas LAF-1 plays a role in P granule formation. DDX4 and LAF-1 contain a DEAD-

box helicase domain (named after the amino acid sequence D-E-A-D of a conserved motif) flanked by N- and C-terminal IDRs. In both cases, the N-terminal IDRs, rich in positively charged RG sequence motifs and negatively charged amino acids, are drivers of phase separation (159, 206).

Charge distributions and patterning in IDRs have been shown to affect phase separation in general (206, 207), but the specific mechanisms remain to be unveiled. Charge patterns in IDRs affect the conformational space sampled by the chain and the compactness of the ensemble (208-210). This affects the balance of protein-solvent versus protein-protein interactions, and thus phase separation behaviour. In the case of LAF-1, a high conformational flexibility correlates with a permeable and low-density condensate (211). This feature is found for other droplet forming systems (154, 212). Furthermore, DDX4-based droplet formation is highly temperature dependent with lower temperatures promoting phase separation and higher temperatures dissolving droplets. It was also shown for DDX4 that arginine methylation impedes droplet formation, pointing to cation- π interactions also playing a contributing role in phase separation (Figure 1.11C). These cation- π have been confirmed to be important in a recent NMR spectroscopic study of DDX4 condensates (213). The study also revealed that the N-terminal IDR of DDX4 remains disordered and highly mobile inside the condensed phase.

Another system where charges play an important role in phase separation is the IDR of the nephrin intracellular domain (NICD). Here, the highly negatively charged (net charge of - 21) NICD forms complexes with positively "supercharged" green fluorescent proteins that go on to phase separate above a concentration threshold (207). This process, termed complex coacervation (214), has long been described for synthetic systems where two polymers or colloids of opposite charge undergo complex formation and phase separation at higher concentration.

In contrast to these highly charged IDRs, the N-terminal low complexity IDR driving phase-separation of the ALS-related stress granule protein FUS is mostly comprised of amino acids with neutral polar side chains (Glu, Gly, Ser and Thr) interspersed frequently with aromatic residues (Tyr). The physicochemical driving forces for phase separation of this IDR are therefore likely to involve polar dipole-dipole, as well as π - π -stacking, interactions. This is reflected in increasing phase separation propensity with increasing ionic strength (171). The importance of

aromatic interactions has further been confirmed by mutational studies (215). Similar studies showed aromatic residues to be paramount for phase separation of other systems such as other stress granule proteins like hnRNPA1 (175) and the mitotic spindle protein BuGZ (216).

BuGZ constitutes a phase separating system that relies predominantly on hydrophobic interactions between low complexity regions of IDRs. This is associated with a so-called lower critical solution temperature (LCST), with droplet formation favored at higher temperatures (216-218). This stands in contrast to the above mentioned systems with an upper critical solution temperature (UCST), where droplets form at lower and dissolve at higher temperatures. The latter is the classical case for many systems where hydrophobic interactions do not play a dominant role.

1.3.2.3.3 The Role of Nucleic Acids in Biological Condensates

Many intracellular membraneless organelles also contain RNA and are therefore referred to as ribonucleoprotein (RNP) bodies. In these systems, protein-RNA interactions enhance phase separation and lower the saturation concentration (corresponding to ϕ_l in Figure 1.10) for droplet formation. In addition to phase separating IDRs, some of the proteins involved contain multiple repeats of RNA recognition motif domains (RRMs) that provide multivalent interactions with long RNA polymers. For example, an N-terminal fragment of hnRNPA1 containing 2 RRM and no low complexity regions, does not phase separate on its own, but will form a condensate upon addition of RNA (163). Similarly, phase separation of the nucleolar component nucleophosmin (NPM1/B23), a pentamer of single RRM and IDR chains of alternating acidic and basic sequence blocks, is enhanced by the addition of RNA to the multivalent complex (219).

An example of RNA-dependent phase separation of a protein that contains very little intrinsic disorder is the polypyrimidine tract-binding protein (PTB). This protein harbours 4 RRM and undergoes phase separation upon addition of RNA (152) (Figure 1.11D). Other examples of RNA-enhanced condensate formation include hnRNPA1 (163, 176), Whi3 (220), PGL-3 (156) and another nucleolar component, FIB-1 (221). These findings are in line with RNA being identified as the trigger for the formation of various nuclear membraneless bodies such as nucleoli, paraspeckles, histone locus bodies, cajal bodies and nuclear stress bodies (221-223). As RNA-interacting motifs in proteins are predominately positively charged and the

interactions with the negatively charged phosphate backbone of RNA are of electrostatic nature, phosphorylation of these sequences in the protein by kinases is proposed to be a general mechanism for inhibiting the formation of RNP bodies (224).

Pathogenicity in neurological and neuromuscular repeat expansion disorders like Huntington's disease, spinocerebellar ataxia and myotonic dystrophy has been linked to sequence repeats in certain genes (225-231). It has recently been discovered that aberrant nuclear RNA foci that sequester RNA binding proteins are formed based on phase separation of repeat-containing RNAs alone. This happens above a threshold number of repeats (232). Such behaviour is consistent with a critical number of repeats, above which the symptoms of these diseases typically manifest.

1.3.3 Functional Implication of Biological Phase Separation

It is well established that phase separation of biologically relevant macromolecules is the physicochemical basis for the formation of many intracellular membraneless bodies and that it is involved in many important biological processes. However, the functional implications of this widespread phenomenon are only starting to be uncovered.

1.3.3.1 Micro-reactors

One important role proposed for phase separation is the concentration of molecules to enhance biological reactions that would otherwise be inefficient due to low binding affinities between partners. The specialized and selective reaction environment inside a condensate could exclude costly and potentially harmful side reactions and thereby serve as a molecular filter. Due to the dynamic nature of biological condensates and the rapid exchange across the phase boundary, reactants of biological reactions can quickly diffuse in and out of the condensed phase. This allows for a constant influx of substrates and counteracts an unwanted build-up of products. Examples for this mechanism at work include the TCR signalling complex in which actin polymerization factors are concentrated inside condensates, thereby inducing actin polymerization (152, 193). Similarly, phase separation of Alzheimer's Disease-associated tau has been shown to concentrate tubulin, thereby serving as a seed for microtubule growth (150). In addition to enhancing actin polymerization, condensation in TCR signalling also increases the

phosphorylation-dependent output of the mitogen-activated protein kinase (MAPK) signalling pathway (153). For RNP bodies such as nucleoli, it has been suggested that concentration of RNA and RNA processing enzymes increases efficiency of RNA processing reactions (184, 233). The prevalence of phase separation in RNP bodies has been linked to the idea of RNA playing the central role in early life. Droplets as “reaction crucibles” have been suggested to be a form of an ancient protocell (234, 235) and ongoing reactions inside these environments have been shown to be able to influence their size and even cause budding instead of fusion of droplets (236).

1.3.3.2 Storage and Molecular Filters

Another proposed function for biological condensates is the sequestration of molecules such as signalling particles to quickly dampen the output of a signal. Droplets can also act as protective storage sites for proteins otherwise prone to aggregation under stress conditions. An example of sequestration of signalling particles is the target of rapamycin complex 1 (TORC1) that localizes to stress granules and thereby downregulates TORC1 signalling (237, 238). The kinase DYRK3 can also be found in stress granules and in the inactive form promotes their formation and thereby the shutting down of the TORC1 signalling pathway. In the active form however, DYRK3 promotes the dissolution of stress granules, the release of TORC1 and thereby the activation of the pathway (238). Furthermore, stalled translational complexes are temporarily stored in stress granules during times of cellular stress. It has also been shown that the budding yeast prion protein and important translation termination factor Sup35 forms biomolecular condensates upon pH-stress by means of the interplay of a charged pH-sensory region and its prion domain. This condensation is essential to keep the GTPase domain of Sup35 functional and from forming detrimental aggregates during times of cellular stresses (239). This poses an interesting example of a prion domain playing an important functional role as opposed to being pathogenic. It functions here as a stress sensor and an important protein homeostasis safety measure.

1.3.3.3 Spatial Organization

Finally, biomolecular condensation is used as a strategy for the organization of cellular space. As stated above, it can serve as a nucleation site for actin filaments as well as concentrate tubulin for the formation of microtubules. Similarly, BuGZ from *Xenopus laevis* phase separates and is thought to concentrate tubulin for the formation of mitotic spindles (216). In endocytosis, the surface tension of membrane-associated phase separation potentially plays a role for membrane invagination (240). It also has long been known that the great variety of nuclear substructures is entirely based on non-membraneous compartmentalization. Nucleoli and their substructure are shown to originate from immiscible condensation of their components (184) and they have also been associated with chromatin organization (241). The nuclear sub-compartmentalization has been shown to be closely related to gene activity (242, 243). Heterochromatin formation and the associated gene repression based on heterochromatin protein 1a (HP1a) results in a biomolecular condensate with liquid properties (244, 245).

1.4 Rationale and Thesis Overview

The emergence of drug resistant *Mtb* strains highlights the need for further development of more effective and specific ways to treat tuberculosis. Interesting starting points for developments in that area are virulence factors that the bacteria needs for proper establishment of infection. Several such factors have been identified for *Mtb* (246). Some of these have been shown to be modified and regulated by the eukaryotic-like STPKs of the bacilli. This poses a fascinating and intertwined signalling system that remains to be fully dissected. Further understanding of these factors and how phosphorylation regulates their function on a molecular level is vital for exploiting this aspect of the biology of *Mtb* for therapeutic development.

A system that has been identified as a virulence factor and where phosphorylation is important for virulence is the ABC transporter Rv1747 (144). As briefly introduced above, Rv1747 is regulated by STPKs including PknF (143 , 144 , 145) and is proposed to transport cell envelope intermediates across the bacterial membrane (104). Phosphorylation at multiple sites in the IDR of the unique regulatory module appended to the regular ABC transporter architecture is important for the establishment of proper virulence (144). The IDR of this regulatory module is flanked by two FHA domains (Figure 1.7B), of which the pThr binding capability of at least one is also vital for virulence (144). Exactly how phosphorylation affects the regulatory module structurally and how it might in turn regulate the function of the transporter, however, was previously unknown.

Accordingly, the underlying theme of my thesis was to understand this fascinating system and the changes it undergoes upon phosphorylation. In chapter 2, I present a detailed characterization of the Rv1747 regulatory module, including structural and dynamic studies of the two FHA domains and an analysis of their binding to linker phosphothreonines. The key conclusion of this research is that both FHA domains bind the phosphorylated linker via intramolecular ("cis") and intermolecular ("trans") pathways, with the latter leading to higher order oligomerization. I propose that this multivalent binding equilibria with a range of possible microscopic affinities, combined with the activity of multiple STPKs and the sole phosphatase PstP, leads to tunable phosphorylation-dependent assembly or quaternary interactions of Rv1747 to regulate its transporter functions.

The regulatory module of Rv1747 presents an interesting case of multivalent association between two FHA domains and at least two phosphothreonines in the linker IDR. These are prerequisites for phase separating systems. Thus, in the second part of this thesis (chapter 3), I demonstrate and characterize the phase separation behaviour of the regulatory module. In describing the physicochemical and material properties of resulting droplets, I discovered that phase separation results from a combination of modular FHA-pThr binding and additional phosphorylation-independent interactions likely involving the linker IDR. Using a model bilayer system, I also investigated how condensation might lead to membrane clustering of the transporter. Finally, using ultra-high resolution microscopy, I carried out initial studies of the potential phase separation of Rv1747 within *Mtb*. At the time of writing my thesis, this stands as the first description of droplet formation in any bacterial system. My research also uncovers several potential mechanisms for the regulation of the Rv1747 ABC transporter vis SPTK signalling networks. This provides new insights into the mechanisms underlying *Mtb* virulence which someday may be exploited for therapeutic development.

Chapter 2: Biophysical Characterization of the Tandem FHA Domain Regulatory Module from the *Mycobacterium tuberculosis* ABC Transporter Rv1747

2.1 Overview

The *Mycobacterium tuberculosis* ABC transporter Rv1747 is a putative exporter of cell wall biosynthesis intermediates. Rv1747 has a cytoplasmic regulatory module consisting of two pThr-interacting Forkhead-associated (FHA) domains connected by a conformationally disordered linker with two phospho-acceptor threonines (pThr). In this chapter, the structures of FHA-1 and FHA-2 were determined by X-ray crystallography and NMR spectroscopy, respectively. Relative to the canonical 11-strand β -sandwich FHA domain fold of FHA-1, FHA-2 is circularly permuted and lacking one β -strand. Nevertheless, the two share a conserved pThr-binding cleft. FHA-2 is less stable and more dynamic than FHA-1, yet binds model pThr peptides with moderately higher affinity ($\sim 50 \mu\text{M}$ versus $500 \mu\text{M}$ equilibrium dissociation constants). Based on NMR relaxation and chemical shift perturbation measurements, when joined within a polypeptide chain, either FHA domain can bind either linker pThr to form intra- and intermolecular complexes. I hypothesize that this enables tunable phosphorylation-dependent multimerization to regulate Rv1747 transporter activity.

2.2 Introduction

Mycobacterium tuberculosis (*Mtb*), the etiological agent of tuberculosis (TB), primarily infects host alveolar macrophages, which represent the first line of cellular defense against microbial invasion. Normally, macrophages engulf and kill invading microorganisms using acidification, lysis and proteolytic digestion, but pathogenic species of *Mycobacterium* have developed strategies to circumvent this defense mechanisms (247, 248). These strategies rely in part on an intricate stimulus-response system that has a set of serine/threonine protein kinases (STPKs) as its central players.

Mtb encodes eleven STPKs, nine of which are predicted to be localized in the cell membrane due to the presence of a hydrophobic transmembrane helix linking the cytoplasmic

kinase domain and the extracellular sensor domain of each protein (249). Genomic and proteomic studies revealed that STPKs exhibit extensive crosstalk, thereby creating complex signaling networks (87, 93). Endogenous substrates have been identified for the STPKs PknA, PknB, PknD, PknF, PknH and PknG (250, 251). Many of these substrates ensure intraphagosomal survival of *Mtb* (117, 252). Intriguingly, several substrates of the *Mtb* STPKs contain a Forkhead-associated (FHA) domain (145). This is a widespread phosphothreonine (pThr)-recognition module, found in eubacteria and eukaryotes, that mediates phosphorylation-dependent protein-protein interactions (135-138).

One of the FHA domain-containing substrates of several *Mtb* STPKs is an ABC transporter encoded by the open reading frame *Rv1747* (Figure 2.1). *Rv1747* has the single polypeptide chain topology expected for a homodimeric exporter with an N-terminal cytoplasmic nucleotide binding domain (NBD), the location of ATP hydrolysis, followed by a helical transmembrane domain (TMD) through which substrate is transported (253). Uniquely, the NBD of *Rv1747* is preceded by a postulated regulatory module consisting of two FHA domains (FHA-1 and FHA-2) connected by an ~ 100 residue intrinsically disordered (ID; *vide infra*) linker. The FHA domains are required for specific interactions with *Rv1747*'s "cognate" STPK PknF (143) and likely with other *Mtb* STPKs (145). PknF is considered *Rv1747*'s cognate kinase because the two are present in the same operon. PknF phosphorylates the ID linker at least at two confirmed acceptor sites (T152 and T210), and short synthetic peptides corresponding to phosphorylated pT152 or pT210 can bind either of the isolated FHA domains with micromolar affinities (144).

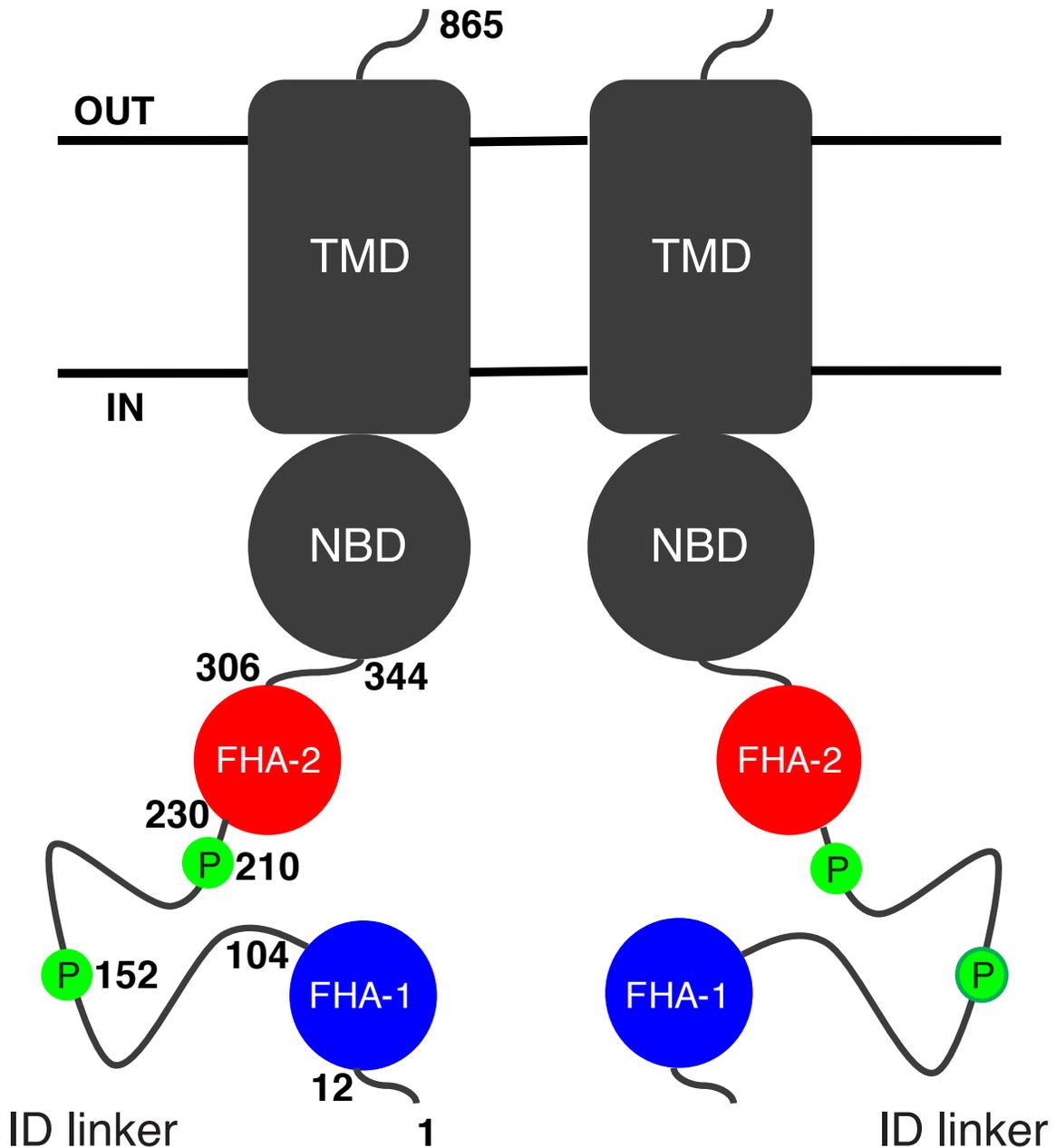


Figure 2.1: Cartoon of the ABC Transporter Rv1747, Assuming a Minimal Functionally Active Homodimer.

In addition to the transmembrane (TMD) and nucleotide binding domains (NBD), Rv1747 contains a postulated N-terminal cytoplasmic regulatory module composed of the FHA-1 (blue) and FHA-2 (red) domains joined by an intrinsically disordered (ID) linker with phospho-acceptor sites T152 and T210 (green).

The exact function of Rv1747 has not yet been elucidated. However, four of the remaining five *Mtb* proteins with FHA domains are involved in processes connected to cell wall synthesis or remodeling (96, 107, 109, 254, 255). Also, experiments on *Rv1747* knockout strains revealed altered *Mtb* sedimentation phenotypes, decreased levels of phosphatidyl-*myo*-inositol mannosides (PIMs) in lipid extracts (105), and increased expression of the efflux pump-related *iniBAC* operon (104). Collectively, these data link Rv1747 with the transport of cell wall biosynthesis intermediates. Most importantly, Rv1747 function is associated with *Mtb* infectivity. The $\Delta Rv1747$ mutant shows significantly attenuated growth in macrophages and mice (94, 105). Similarly, strains with the phosphor-ablative T152A/T210A mutations exhibit reduced growth in macrophages and mice, and those with a mutation in the canonical pThr-binding site of FHA-1 have impaired growth in macrophages (144). Thus the functions of Rv1747 are phosphorylation-dependent, but very little is currently known about the structural details underlying the roles played by the regulatory module.

In this chapter, I investigated the structural and functional features of the Rv1747 regulatory module using a combination of NMR spectroscopy, X-ray crystallography, isothermal titration calorimetry (ITC), and molecular dynamics (MD) simulations. The two FHA domains are structurally independent and joined by a flexible linker ("beads-on-a-string"). FHA-1 has the canonical 11-strand β -sandwich fold of a FHA domain. In contrast, FHA-2 possesses an unusual, circularly permuted topology lacking one β -strand. FHA-2 is also more dynamic and less stable than FHA-1, yet binds pThr-containing peptides with higher affinity. Constructs of Rv1747 containing either FHA domain and the ID linker with either or both phospho-acceptor threonines exhibit inter- and intramolecular modes of FHA-pThr binding with a range of net affinities. This enables a complex interaction network that leads to higher order oligomerization of the full regulatory module upon phosphorylation. I hypothesize that the FHA regulatory module association can be fine-tuned based on the level and sites of phosphorylation and that this in turn modulates the function of the transporter by aiding its assembly or activation.

2.3 Results

2.3.1 Rv1747 Regulatory Module is Composed of Independent FHA-1 and FHA-2 Domains Joined by a Disordered Linker

I initially focused on three non-phosphorylated constructs: residues 1-156 (Rv1747¹⁻¹⁵⁶) corresponding to the FHA-1 domain and part of the following interdomain linker; residues 206-310 (Rv1747²⁰⁶⁻³¹⁰) corresponding to the FHA-2 domain and a small part of the preceding linker; and residues 1-310 (Rv1747¹⁻³¹⁰) spanning the entire regulatory module (Figure 2.2A). The assigned ¹⁵N-HSQC spectra of the first two constructs showed dispersed cross-peaks attributed to the well-ordered regions of the FHA domains and more intense, sharper peaks with random coil ¹H^N chemical shifts of ~ 8 – 8.5 ppm originating from linker residues (Figures 2.2B and 2.3). Consistent with this interpretation, an analysis of the mainchain ¹H, ¹³C, and ¹⁵N chemical shifts with the MICS algorithm (256) indicated the presence of folded domains that are rich in β -strands (Figure 2.2C). The chemical shift-derived secondary structure of FHA-1 matches that of the canonical FHA fold. A minor exception is a 4-residue segment between strands 6 and 7 that has α -helical propensity in solution yet appears as a loop in the crystal structure (*vide infra*). A similar segment with helical propensity is also seen in the chemical shift-based secondary structure of FHA-2. Intriguingly, chemical shift analyses also indicated the absence of the canonical β -strands 1 and 2 in FHA-2, as well as the presence of an additional C-terminal β -strand. This conclusion is confirmed in the NMR-derived tertiary structure of FHA-2, presented below. In addition to lacking any predominant secondary structure, the linker residues exhibited MICS-calculated random coil index squared order parameters (RCI-S²) indicative of conformational disorder (257). The presence of ordered FHA domains and intrinsically disordered linker residues is corroborated by amide ¹⁵N relaxation (Figure 2.4) and amide hydrogen exchange experiments, as will be discussed below.

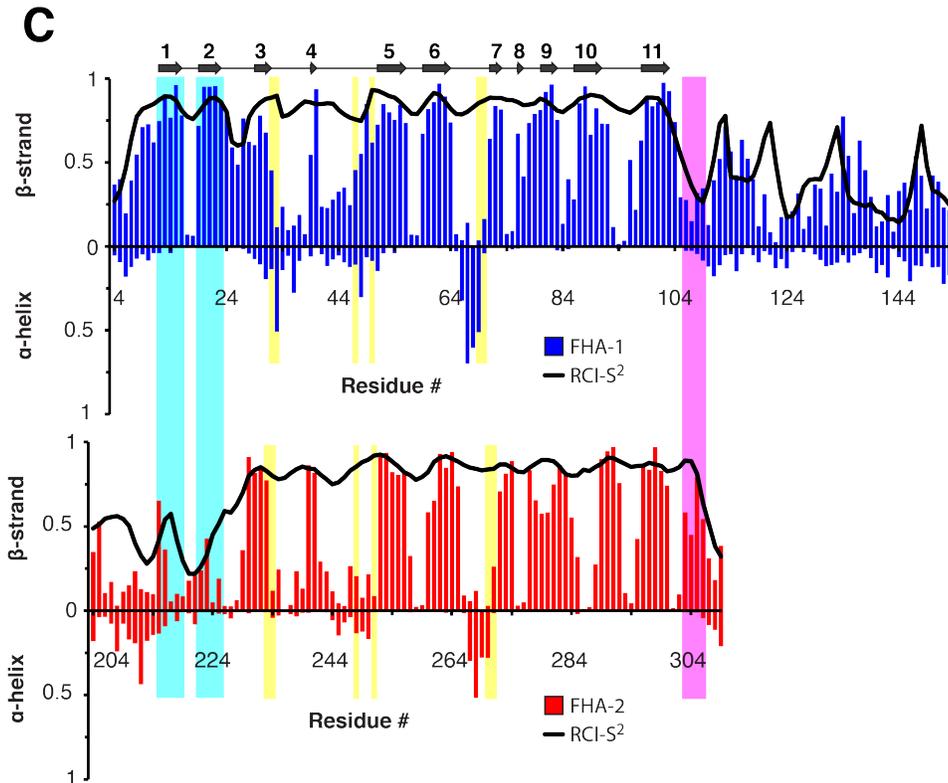
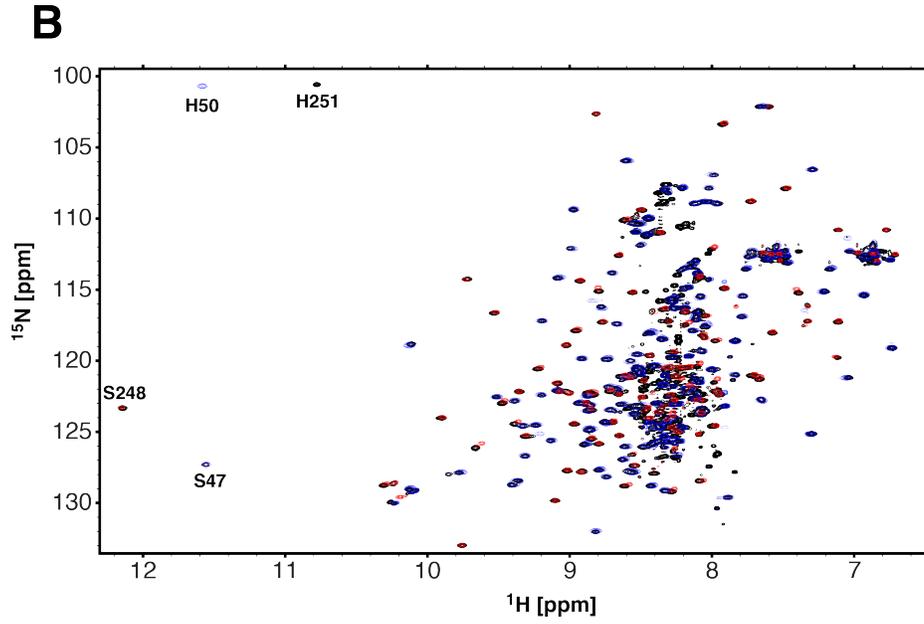
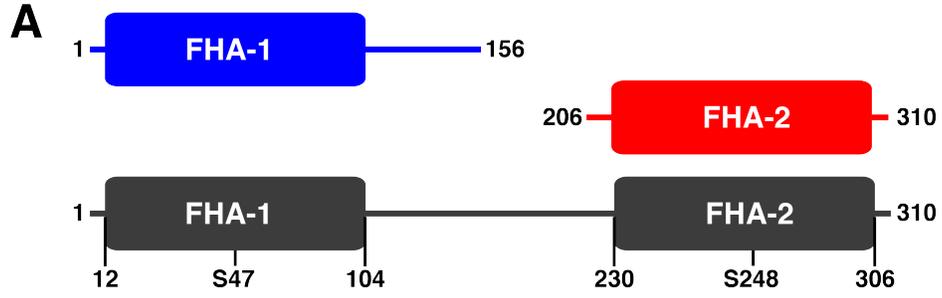


Figure 2.2: The Regulatory Module of Rv1747 is Composed of Two Independent FHA Domains Joined by an ID Linker.

(A) Cartoon representations of the FHA-1 (Rv1747¹⁻¹⁵⁶; blue), FHA-2 (Rv1747²⁰⁶⁻³¹⁰; red) and full FHA regulatory module (Rv1747¹⁻³¹⁰; black) constructs. (B) The ¹⁵N-HSQC spectra of the FHA-1 (blue) and FHA-2 (red) constructs overlaid on that of the full FHA regulatory module (black). The near identical ¹H^N-¹⁵N chemical shifts of corresponding residues indicates that the FHA domains are structurally independent and do not interact when joined to form the regulatory module. Signals from additional ID linker residues in the regulatory module have random coil ¹H^N chemical shifts near 8 – 8.5 ppm. The highly downfield shifted amide signals of S47 and S248, as well as the slowly exchanging imidazole ¹H^{ε2}-¹⁵N^{ε2} signals (aliased) of neutral H50 and H251, in the peptide-binding clefts of FHA-1 and FHA-2, respectively, are identified. Separate, assigned spectra are provided in Figure 2.3. (C) Normalized β-strand and α-helix propensities (histograms) and RCI-S² values (black lines; decreasing values from 1 to 0 indicate increasing flexibility) determined from mainchain chemical shifts with MICS (256, 257). See Figure 2.4 for complementary ¹⁵N relaxation measurements. Conserved residues important for binding are identified in yellow, β-strands 1 and 2 of the canonical FHA fold that are present in FHA-1, yet absent in FHA-2, are in cyan, and the additional C-terminal β-strand 1* of FHA-2 is in magenta. The secondary structure (β-strand: arrow) calculated with DSSP from the crystal structure of FHA-1 is also indicated.

Figure 2.3: Assigned ^{15}N -HSQC Spectra of Rv1747 FHA Domain Constructs.

Assigned ^{15}N -HSQC spectra of the (A) FHA-1 (Rv1747¹⁻¹⁵⁶) and (B) FHA-2 (Rv1747²⁰⁶⁻³¹⁰) constructs (25 °C, 100 mM NaCl, 20 mM sodium phosphate, pH 6.0). (C) Also shown is the spectrum of the full FHA regulatory domain (Rv1747¹⁻³⁰¹). These spectra are superimposed in Figure 2.2B.

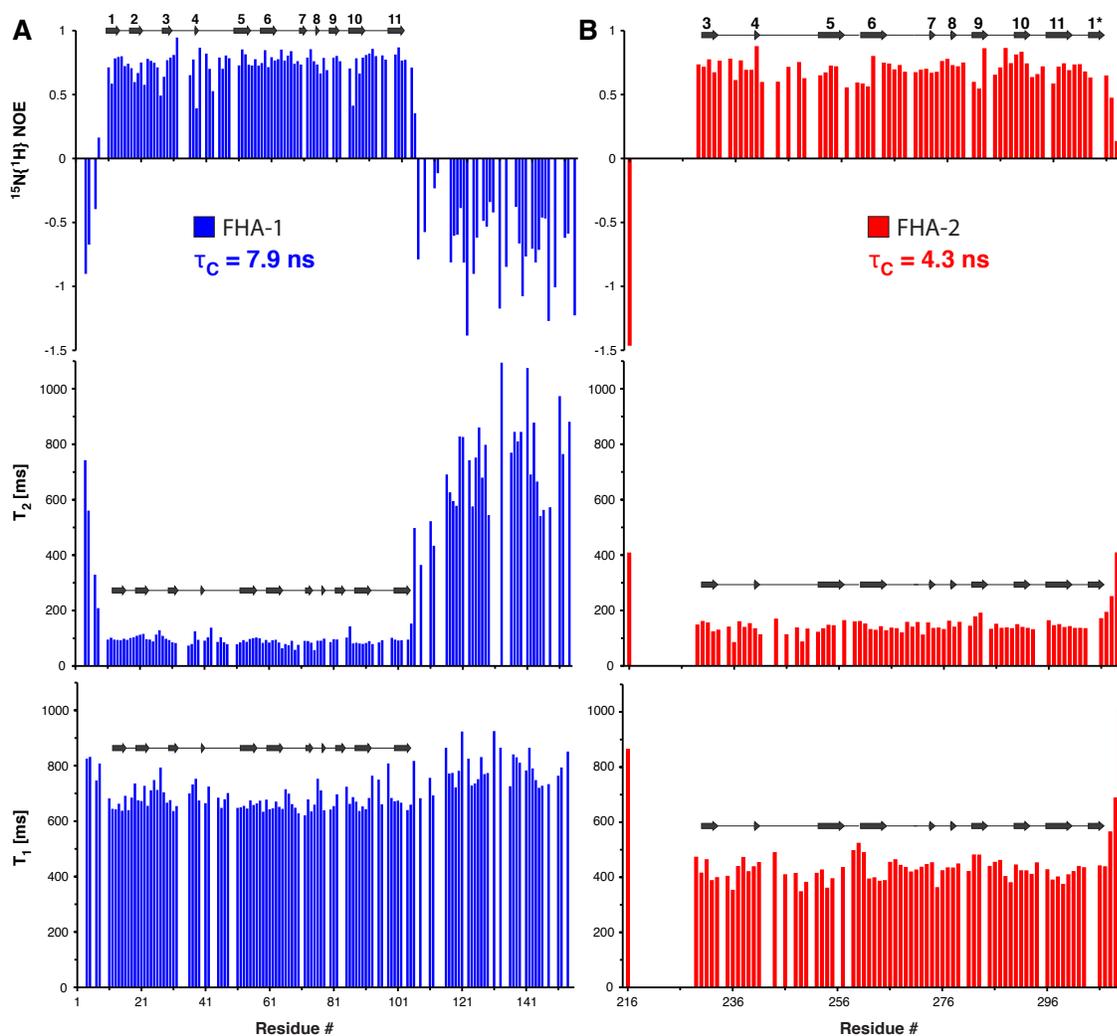


Figure 2.4: Amide ^{15}N Relaxation Data of Rv1747 FHA Domain Constructs.

Amide ^{15}N relaxation data (T_1 , T_2 , and heteronuclear NOE) for the FHA-1 (Rv1747¹⁻¹⁵⁶, blue) and FHA-2 (Rv1747²⁰⁶⁻³¹⁰, red) constructs, recorded with a 600 MHz NMR spectrometer (25 °C, 100 mM NaCl, 20 mM sodium phosphate, pH 6.0). The correlation times for isotropic global rotational diffusion (τ_c) were determined from the T_1 and T_2 values for ordered FHA domain residues using Tensor2 (258). These residues exhibit relatively uniform relaxation parameters with heteronuclear NOE values ~ 0.8 , indicative of restricted fast timescale motions. In contrast, the conformationally disordered terminal and ID linker residues have negative heteronuclear NOE values and long T_2 lifetimes due to sub-ns timescale motions. Secondary structures calculated with DSSP are indicated (arrow, β -strand). Missing data points correspond to prolines and amides with overlapping or unassigned signals.

The dispersed, assigned signals arising from the ordered domains in the ^{15}N -HSQC spectra of the FHA-1 and FHA-2 constructs overlap closely with those in spectrum of the full regulatory module (Figure 2.2C). Based on the lack of amide chemical shift perturbations, it can be inferred that, in the non-phosphorylated state, the two folded FHA domains are not interacting when joined into a continuous polypeptide. As expected, additional signals with random coil $^{15}\text{N}^{\text{H}}$ chemical shifts are also observed in the ^{15}N -HSQC spectrum of the longer construct due to the presence of the full linker. Collectively, these data demonstrate that the Rv1747 regulatory module has a "beads-on-a-string" organization with two structurally independent FHA domains joined by an ID linker.

2.3.2 FHA-1 Has a Canonical FHA Domain Topology Whereas FHA-2 is Circularly Permuted

The structure of FHA-1 (Rv1747³⁻¹¹⁶) was calculated in the group of Dr. Alber by X-ray crystallography (Figure 2.5A and Table 2.1) and I determined the tertiary structural ensemble of FHA-2 (Rv1747²⁰⁶⁻³¹⁰) using NMR spectroscopy (Figure 2.5B and Table 2.2). FHA-1 has the canonical FHA domain topology with 11 β -strands in a β -sandwich fold and a shallow pThr-binding cleft extending across its apical surface. FHA-2 also has the expected two sheet β -sandwich fold of a FHA domain. However, the canonical architecture exhibited by FHA-1 involves a 6-stranded anti-parallel β -sheet (strands 2, 1, 11, 10, 7, 8) and 5-stranded mixed β -sheet (4, 3, 5, 6, 9) (136). In contrast, FHA-2 is circularly permuted such that an additional C-terminal strand, denoted 1*, adopts the position of strand 1 in the first sheet. Also, the equivalent of strand 2 is missing, and thus the anti-parallel β -sheet only has 5 strands (1*, 11, 10, 7, 8). The permutation and the absence of strand 2 is particularly interesting as it places the T210 phospho-acceptor within the unstructured linker N-terminal to FHA-2. In a hypothetical homology model of this domain without the circular permutation, T210 would be part of strand 2 and thus likely inaccessible for modification by the *Mtb* kinases and for FHA domain binding.

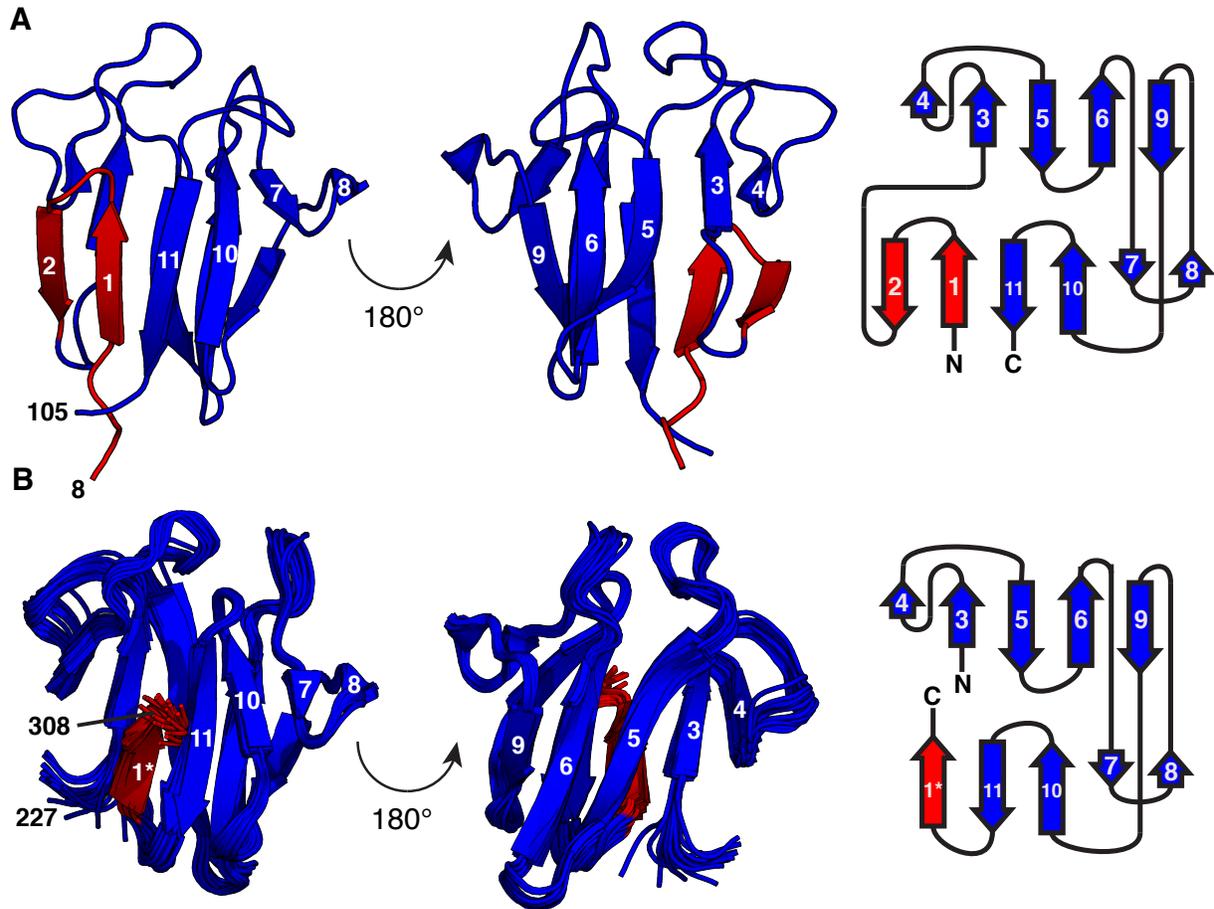


Figure 2.5: FHA-2 has a Circularly Permuted FHA Domain Fold.

(A) Structure and topology diagram of FHA-1 (Rv1747³⁻¹¹⁶) determined by X-ray crystallography. (B) Structural ensemble and topology diagram of FHA-2 (Rv1747²⁰⁶⁻³¹⁰) determined by NMR spectroscopy. Highlighted in red are the N-terminal β -strands 1 and 2 of FHA-1 and the corresponding C-terminal strand 1* of FHA-2. See Figure 2.6A for overlaid structures.

Table 2.1: X-ray crystallography data collection and refinement statistics for FHA-1 (Rv1747³⁻¹¹⁶).

Derivative	Native	NaBr		
Space Group	P3 ₂ 21	P3 ₂ 21		
Data collection				
Unit cell dimensions				
a = b, c (Å)	58.92, 58.92, 68.63	59.05, 59.05, 68.39		
α = β, γ (°)	90, 90, 120	90, 90, 120		
Wavelength (Å)	1.1159	0.920	0.9202	1.1159
Resolution (Å)	68.6 – 1.8 (1.84-1.8)	50 - 2.0	50 - 2.0	50 - 2.0
Observations	109053	289614	289834	164976
Unique reflections	11508	9723	9753	9697
Multiplicity †	9.5 (4.9)			
<I/σI> †	26.0 (1.9)	11 (2.5)	12.4 (1.8)	35.9 (8.2)
Data coverage (%)	87.0 (46.2)	100	100	100
R _{merge} *†	6.9 (82.2)	15.1 (52.0)	13.8 (67.0)	6.0 (26.0)
Mean figure of merit			0.72 (0.50)	
Refinement statistics				
Resolution range (Å)	51.03 - 1.80			
Reflections used	21394			
Mean B-factor (Å ²)	24.05			
R _{cryst} (%) ††	16.89			
R _{free} (%)	20.84			
RMS deviations #				
Bonds (Å)	0.007			
Angles (°)	1.13			
Ramachandran				
Most favored (%)	97.0			
Allowed (%)	3.0			

† Parentheses denote values for highest resolution shell

* $R_{merge} = \sum |I - \langle I \rangle| / \sum I$; I , intensity

†† $R_{cryst} = \sum |F_o - F_{calc}| / \sum F_o$; F_o observed structure-factor amplitude; F_{calc} calculated structure-factor amplitude

Root mean square deviations from ideal values

¥ Mean figure of merit (after density modification) = $\langle |\sum P(\alpha) e^{i\alpha} / \sum P(\alpha)| \rangle$; α , phase; $P(\alpha)$, phase probability distribution

Table 2.2: NMR spectroscopy data collection and refinement statistics for FHA-2 (Rv1747²⁰⁶⁻³¹⁰).

NMR distance and dihedral constraints	
Distance constraints	
Total NOE	1302
Intraresidue	347
Sequential ($ i-j = 1$)	360
Medium-range ($ i-j < 4$)	135
Long-range ($ i-j > 5$)	460
Hydrogen bonds (present in > 6 models)	63
Dihedral angle restraints	
φ	68
ψ	68
Structure statistics	
Violations (mean and std. dev.)	
Distance restraints (Å)	0.005 ± 0.002
Dihedral angle restraints (°)	4.1 ± 0.2
Max. distance restraint violation (Å)	0.54
Max. dihedral angle violation (°)	26.28
Ramachandran plot (%)	
Most favored	91.0
Additionally allowed	7.6
Generously allowed	0.3
Disallowed	1.0
Average pairwise RMS deviations* (Å)	
Backbone atoms	0.48 ± 0.10
Heavy atoms	0.96 ± 0.10

*Pairwise RMSD calculated among the 20 refined structures for residues 28 – 104

Despite significant sequence differences between FHA-1 and FHA-2 (27 % identity, 36 % similarity), as well as the circular permutation of FHA-2 and the absence of one β -strand, the two FHA domains have very similar overall tertiary structures (backbone RMSD = 1.4 Å; Figure 2.6A). Importantly, residues that have been shown to be involved in phospho-threonine peptide binding by canonical FHA domains are conserved in sequence and structure (R33, S47 and H50 in FHA-1; R234, S248, and H251 in FHA-2; Figures 2.7A,B and 2.6B). Furthermore, a hydrogen bond network supporting the binding site that is distinctive for FHA domains (137, 259) is present in both FHA-1 and FHA-2. This is readily seen by the characteristic downfield amide $^1\text{N}^{\text{H}}$ chemical shifts of S47/S248 (11.58 ppm and 12.15 ppm), which reflect the formation of a hydrogen bond to the deprotonated $\text{N}^{\delta 1}$ in the neutral H50/H251 sidechains of FHA-1/2, respectively. In turn, the nitrogen-bonded $\text{H}^{\epsilon 2}$ on these imidazole rings hydrogen bond to the backbone carbonyl of G71/G271. This protects the labile $^1\text{H}^{\epsilon 2}$ from rapid hydrogen exchange, yielding detectable signals at 11.61 ppm and 10.77 ppm, respectively, in the ^{15}N -HSQC spectra of the two FHA domains (Figure 2.2B). H251 shows characteristic N^{δ} and N^{ϵ} chemical shifts of 253.1 ppm and 169.5 ppm respectively that are consistent with a neutral protonation state at N^{ϵ} (260).

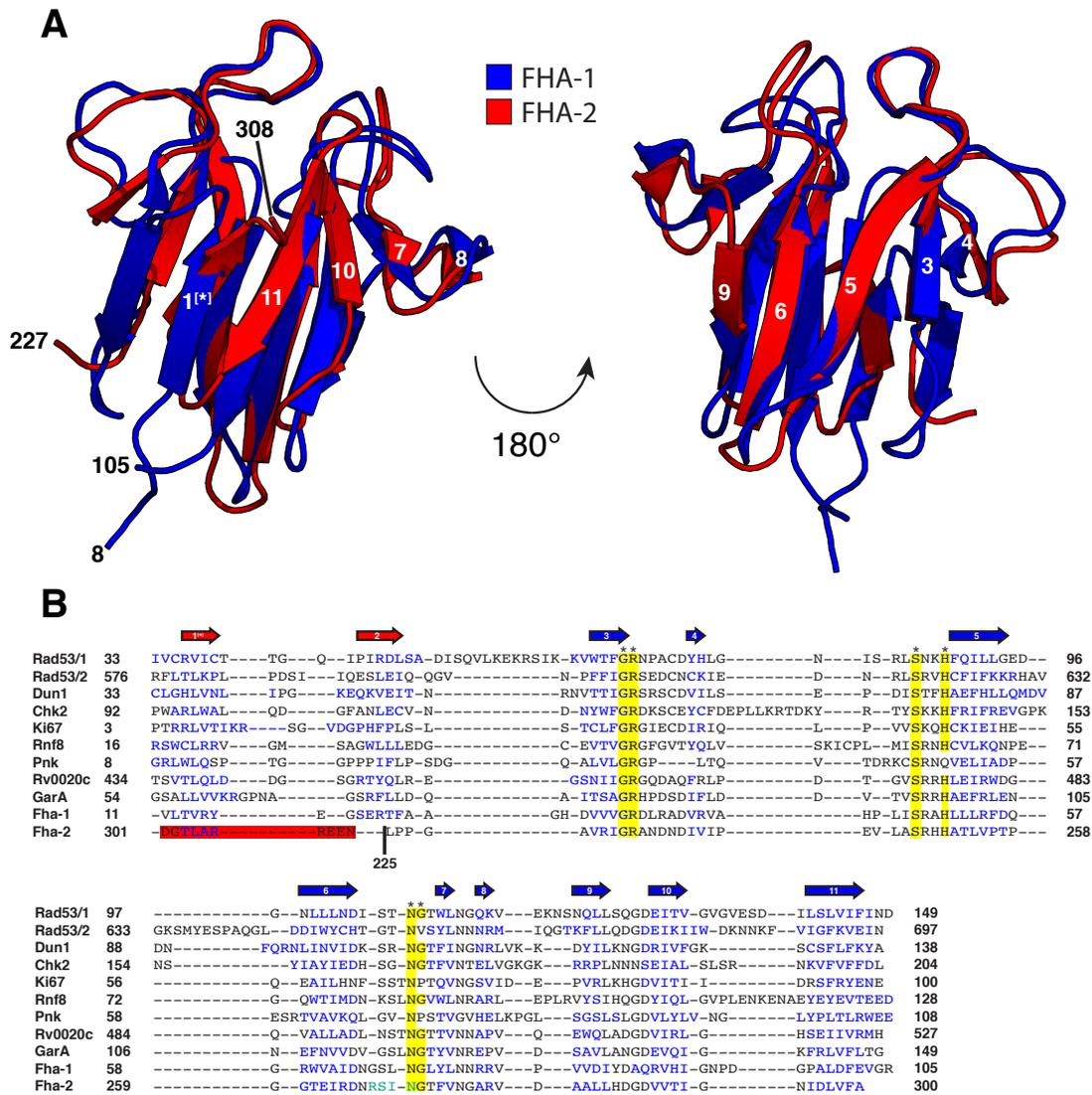


Figure 2.6: Structural and Sequence Alignments of the Rv1747 FHA Domains.

(A) Alignment of the X-ray crystallographic structure of FHA-1 (blue) and the lowest energy member of the NMR-derived structural ensemble of FHA-2 (red). The RMSD for all mainchain atoms, including those in loop regions (residues 8-105 for FHA-1 and 225-310 for FHA-2), is 1.4 Å. (B) Structure-based sequence alignment of FHA domains generated with the RCSB PDB jCE server (<https://www.rcsb.org/>). Conserved residues that interact with the bound peptide are highlighted in yellow and with an asterisk. Residues in β -strands are in blue font. The circularly permuted residues of Rv1747 FHA-2 are highlighted in red. (*S. cerevisiae* Rad51 FHA-1, PDB ID 1G6G.pdb; *S. cerevisiae* Rad51 FHA-2, 1K2N.pdb; *S. cerevisiae* Dun1, 2JQJ.pdb; *H. sapiens* Chk2, 1GXC.pdb; *H. sapiens* Ki67, 1R21.pdb; *H. sapiens* Rnf8, 2PIE.pdb; *M. musculus* Pnk, 1YJM.pdb; *M. tuberculosis* Rv0020c, 2LC1.pdb; *M. tuberculosis* GarA, 2KFU.pdb; *M. tuberculosis* FHA-1; *M. tuberculosis* FHA-2)

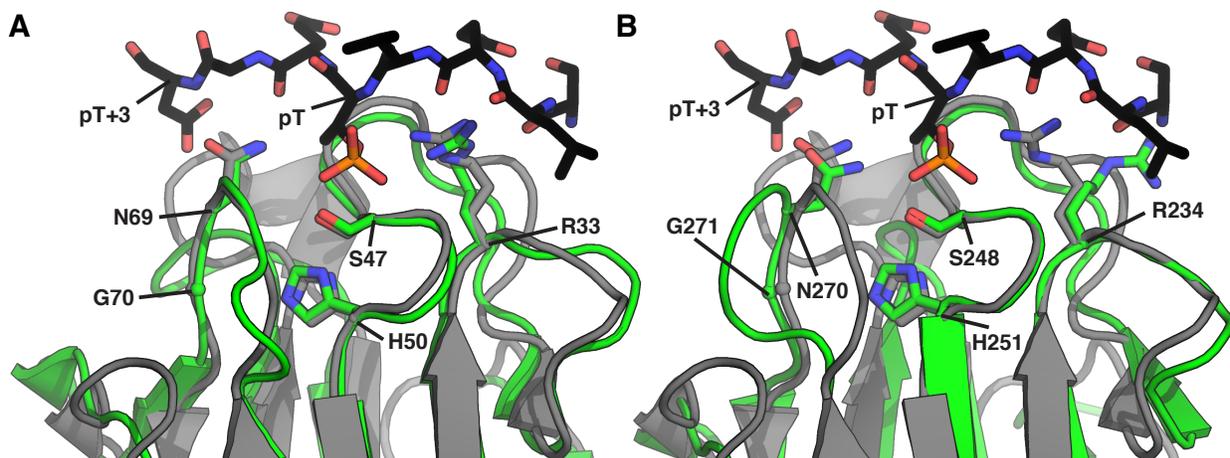


Figure 2.7: FHA-1 and FHA-2 have Conserved pThr-Binding Interfaces.

Overlaid cartoon representations of the X-ray crystallographic structure of pThr peptide-bound Rad53 FHA1 domain (1G6G.pdb, gray) with (A) the crystal structure of unbound FHA-1 (green) and (B) the lowest energy member of the NMR-derived structural ensemble of unbound FHA-2 (green). Important residues in the binding interfaces are shown as sticks (oxygen, red; nitrogen, blue). The Rad53-bound peptide has the phosphorous atom of pThr in orange and carbons in black. In all three structures, the central histidine is in its neutral $N^{\epsilon}H$ tautomeric form to accept a hydrogen bond from the amide of the adjacent serine and to donate a hydrogen bond to the carbonyl of a conserved glycine. This conclusion is based on structural requirements for the X-ray crystallographic models, combined with diagnostic NMR chemical shifts for the two Rv1747 constructs. See Figure 2.6B for sequence alignments.

2.3.3 FHA-2 Binds pThr Peptides with Higher Affinity Than FHA-1

The interactions of the Rv1747 FHA domains with short phosphopeptides corresponding to the reported PknF acceptor sites pT152 and pT210 in the ID linker (KKYAGQQLPPApT¹⁵²TRIPAA and KKYAGTEAGNLApT²¹⁰SMMK) were characterized by NMR spectroscopy and isothermal titration calorimetry (ITC). In ¹⁵N-HSQC monitored titrations, both FHA domains bound both phosphopeptides (Figures 2.8A and Appendix A.1-4). Furthermore, based on the fit K_d values from these titrations (Table 2.3 and Figure 2.9), neither FHA domain showed any significant discrimination between the two peptides. This is consistent with the observation that, beyond the invariant pThr, FHA domain specificity is often set by the pT+3 residue (Figure 2.7) (136, 137). In the case of the two peptides (pT¹⁵²TRI and pT²¹⁹SMM), these residues both have long, hydrophobic sidechains. In contrast, the peptides bound the two FHA domains with substantially different, albeit modest, affinities. The K_d values for either pT152 or pT210 phosphopeptide with FHA-1 were in the near mM range, whereas those with FHA-2 were ~ 20 – 30 fold lower.

Table 2.3: NMR Spectroscopic and ITC Studies of FHA – Phosphopeptide Interactions.

Protein construct	pH	phospho-peptide	K_d (NMR) (μ M)	K_d (ITC) (μ M)	ΔH (kcal/mol)	ΔS (cal/mol•K)
FHA-1 (Rv1747 ¹⁻¹⁵⁶)	6	pT152	680 ± 110	350 ± 190	N.D.	N.D.
	6	pT210	750 ± 170	560 ± 40	N.D.	N.D.
	8	pT152	N.D.	100 ± 60	N.D.	N.D.
FHA-2 (Rv1747 ²⁰⁶⁻³¹⁰)	6	pT152	38 ± 15	72 ± 12	- 31 ± 2	- 84 ± 7
	6	pT210	23 ± 5	23 ± 3	- 11.9 ± 0.5	- 19 ± 2
	8	pT152	N.D.	32 ± 7	-12 ± 2	- 20 ± 7

See Figures 2.9-11 for binding isotherms

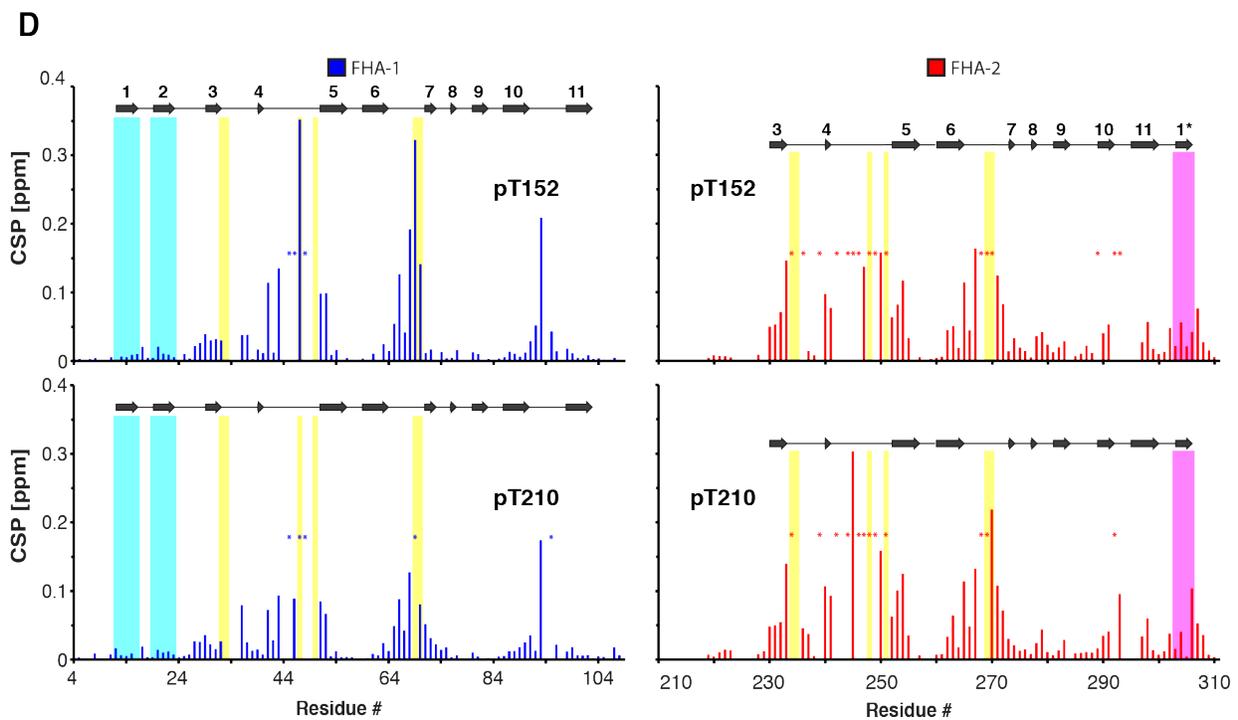
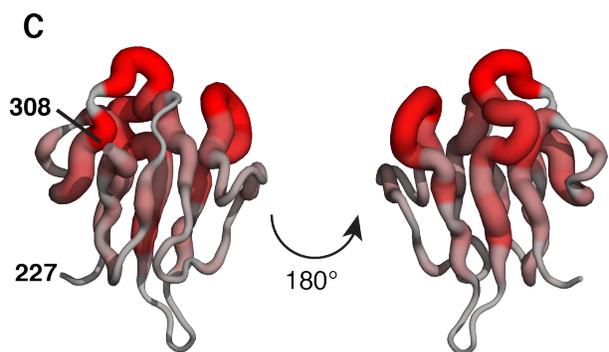
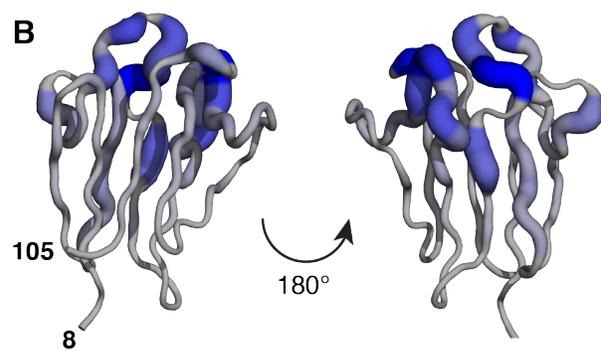
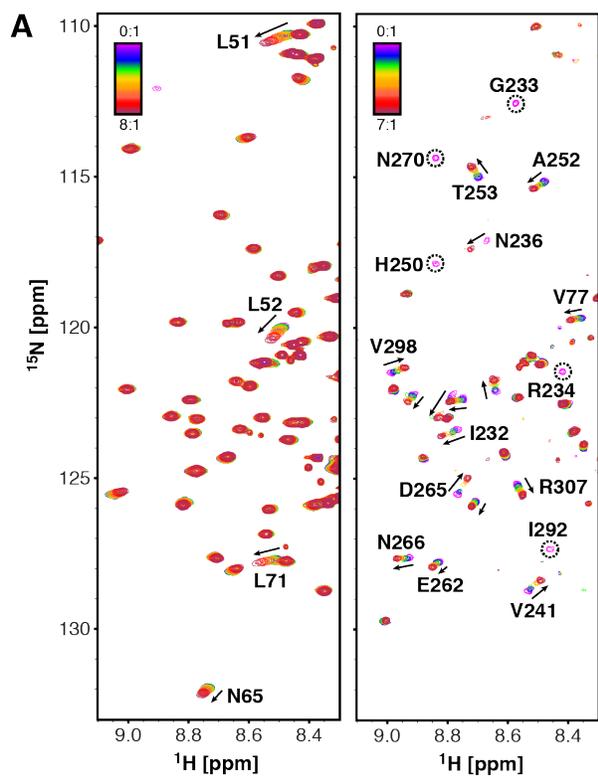


Figure 2.8: FHA-1 and FHA-2 Bind Phosphothreonine Peptides.

(A) Expanded regions of ^{15}N -HSQC spectra of uniformly ^{15}N -labeled FHA-1 (Rv1747¹⁻¹⁵⁶; left) and FHA-2 (Rv1747²⁰⁶⁻³¹⁰; right) titrated with unlabeled phosphopeptides corresponding to pT152 and pT210, respectively. Peptide:protein ratios in the overlaid spectra are indicated with rainbow coloring. The progressive shift changes of selected amides showing binding in the fast exchange limit are shown by arrows. Amides exhibiting severe intermediate exchange broadening are identified with dashed circles. Full spectra and fit binding isotherms are provided in Appendices A.1-4 and Figure 2.9, respectively. (B) Chemical shift perturbations upon binding the pT152 peptide mapped in a grey-blue color and tube thickness gradient onto a cartoon of the structure of FHA-1. Similar changes occur with the pT210 peptide (not shown). (C) Chemical shift perturbations upon binding the pT210 peptide mapped in a grey-red color gradient onto a cartoon of the structure of FHA-2. Similar changes occur with the pT152 peptide (not shown). (D) Amide chemical shift perturbations upon binding of pT152 (8:1 molar excess, 59 % and 94 % saturated for FHA-1 and FHA-2, respectively, based on K_d values in Table 2.3) and pT210 (7:1 molar excess, 52 % and 92 % saturated for FHA-1 and FHA-2, respectively) peptides by FHA-1 and FHA-2. Amides showing exchange broadening are indicated by an asterisk. Conserved residues important for binding are highlighted in yellow, β -strands 1 and 2 of FHA-1 in cyan, and the additional C-terminal strand 1* of FHA-2 in magenta. Missing data points correspond to prolines and amides with overlapping or unassigned signals.

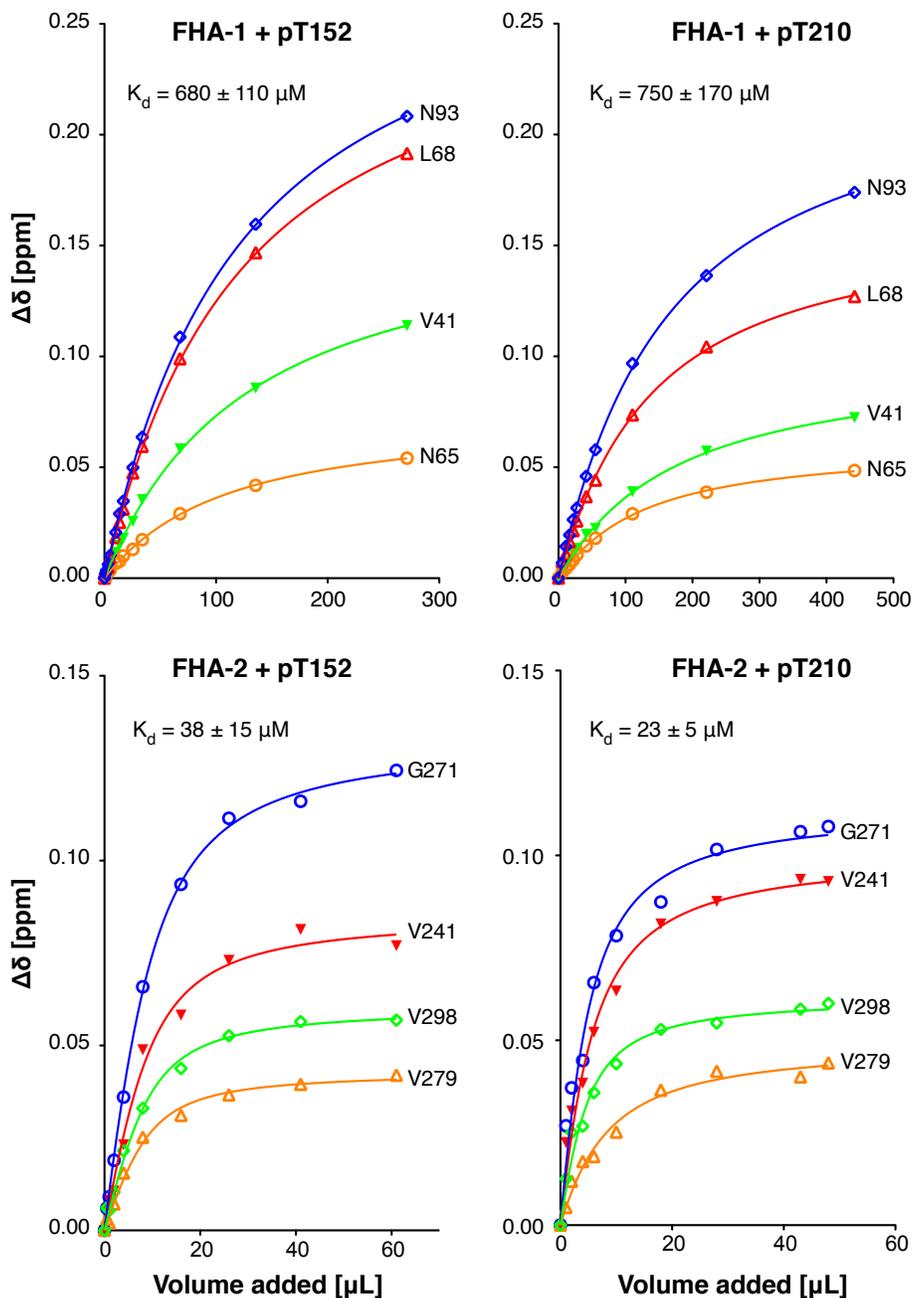


Figure 2.9: NMR Binding Isotherms of FHA-Phosphopeptide Titrations.

Representative binding isotherms for the NMR-monitored titrations of FHA-1 (Rv1747¹⁻¹⁵⁶) and FHA-2 (Rv1747²⁰⁶⁻³¹⁰) with peptides corresponding to the pT152 and pT210 phospho-acceptor sites (25 °C, 100 mM NaCl, 20 mM sodium phosphate, pH 6.0). The lines are best individual fits to a 1:1 binding isotherm, with the resulting K_d values averaged for Table 2.3. See Materials and Methods for experimental details.

The same series of titrations were also carried out using the complementary technique of ITC, and comparable K_d values were obtained (Table 2.3 and Figures 2.10,11). Due to the relatively weak interactions of FHA-1 with the phosphopeptides, reliable ΔH values could not be obtained. In the case of FHA-2, binding was accompanied by a favorable enthalpic change offset and an unfavorable net loss of entropy. Previously published ITC measurements of the same protein-phosphopeptide combinations that were carried out at a higher pH also revealed larger negative ΔH values with FHA-2 than FHA-1, yet relatively uniform K_d values of $\sim 2 \mu\text{M}$ (144). For a direct comparison with this previous study, I repeated the ITC experiments at pH 8 (Table 2.3 and Figure 2.11). Decreased K_d values were found for both FHA domains at pH 8 compared to pH 6, which can be explained by different net charges for the phosphate groups (unperturbed second $pK_a \sim 6.3$ for pThr (261)). Differences to the study by Spivey et. al are probably due to general differences in the experimental setup, and I decided to perform all other experiments here at pH 6 for internal consistency.

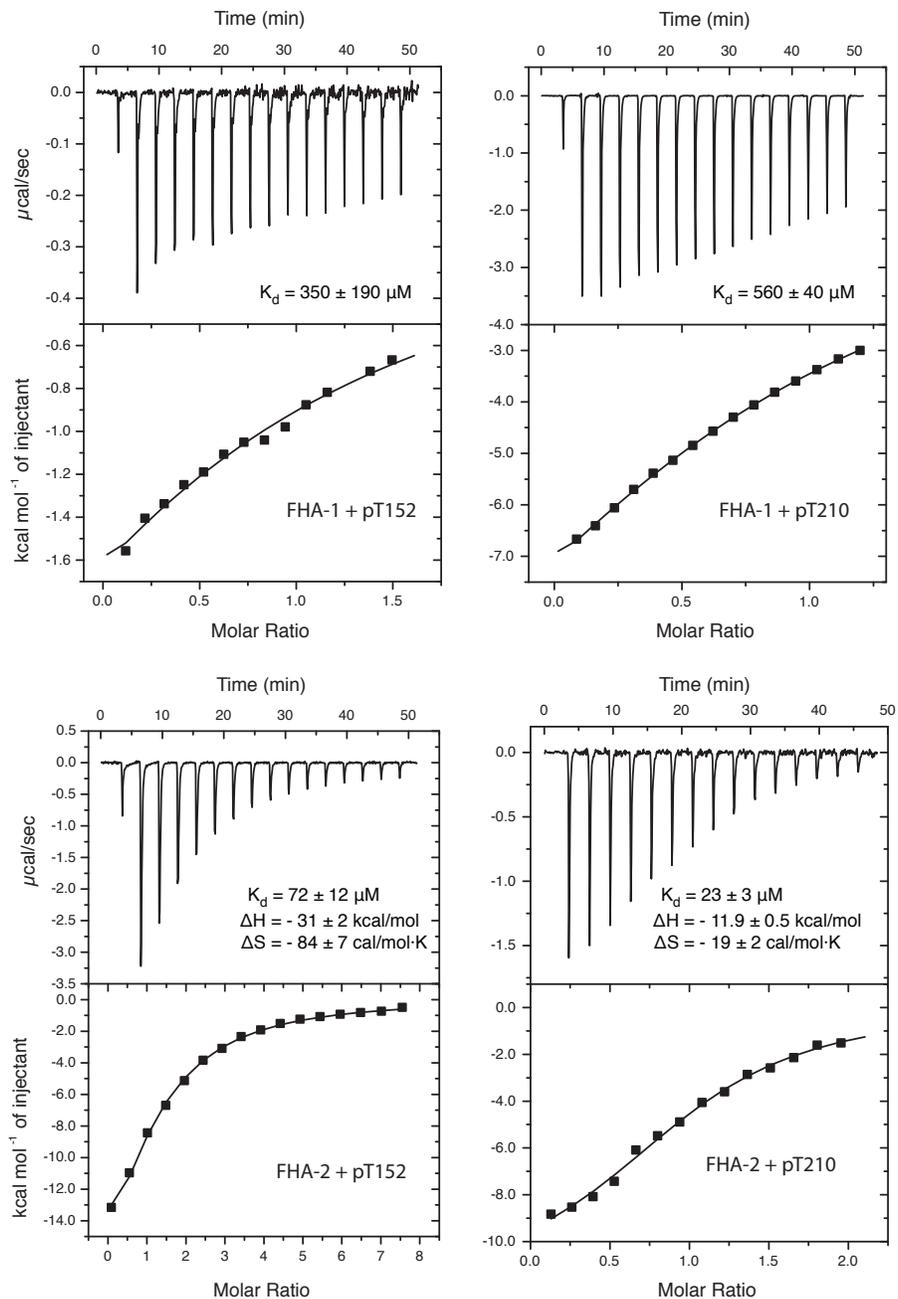


Figure 2.10: ITC Binding Isotherms of FHA-Phosphopeptide Titrations at pH 6.

Isothermal titration calorimetry data for titrations of FHA-1 (Rv1747¹⁻¹⁵⁶) and FHA-2 (Rv1747²⁰⁶⁻³¹⁰) with peptides corresponding to the pT152 and pT210 phospho-acceptor sites (25 °C , 100 mM NaCl, 20 mM sodium phosphate, pH 6.0). The baseline corrected data were fit to a 1:1 binding isotherm, with protein concentrations adjusted to obtain this stoichiometry. See Methods for experimental details.

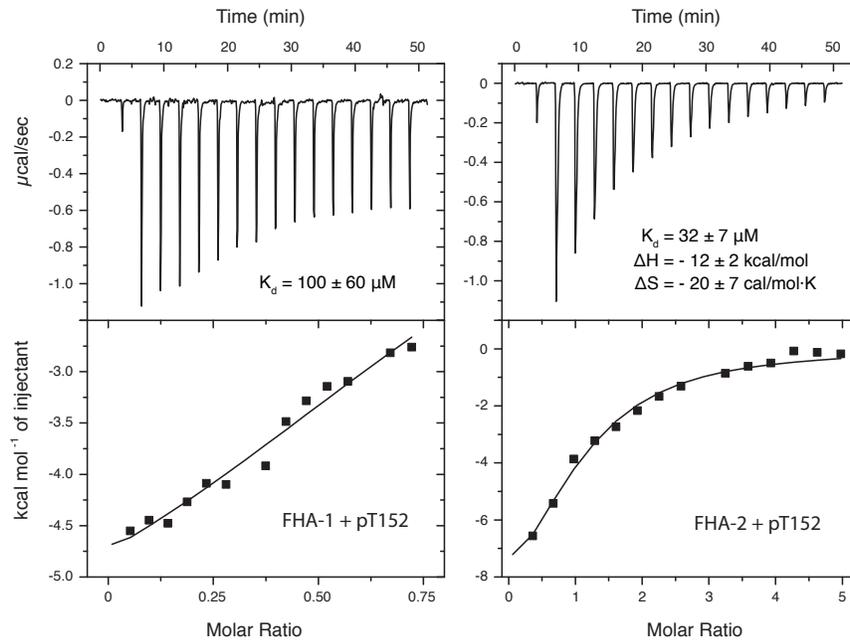


Figure 2.11: ITC Binding Isotherms of FHA-Phosphopeptide Titrations at pH 8.

Isothermal titration calorimetry data for titrations of FHA-1 (Rv1747¹⁻¹⁵⁶) and FHA-2 (Rv1747²⁰⁶⁻³¹⁰) with peptides corresponding to the pT152 phospho-acceptor site (25 °C , 100 mM NaCl, 20 mM sodium phosphate, pH 8.0). The baseline corrected data were fit to a 1:1 binding isotherm, with protein concentrations adjusted to obtain this stoichiometry. See Methods for experimental details.

2.3.4 FHA-2 Exhibits Larger NMR Spectral Perturbations than FHA-1 Upon Phosphopeptide Binding

The NMR-monitored titration experiments also provide insights into the mechanism of phosphopeptide binding. When mapped onto the structures of FHA-1 and FHA-2, residues showing pronounced amide chemical shift perturbations upon addition of either the pT152 or pT210 peptides and found them to cluster near the canonical FHA domain binding interface (Figures 2.8B-D). However, a visual comparison of the ^{15}N -HSQC titration spectra presented in Figure 2.8A shows clear differences in the effects of binding to the peptides on the two FHA domains. In the case of FHA-1, only a relatively small number of amides had perturbed NMR signals and of those, most exhibited progressive chemical shift changes upon titration with either phosphopeptide (Figure 2.8A and Appendices A.1,2). This is diagnostic of binding in the fast exchange limit, whereby the exchange rate constant $k_{\text{ex}} = k_{\text{on}}[\text{peptide}] + k_{\text{off}}$ is much greater than the resonance frequency difference $\Delta\omega$ between free and bound states (262). This is also consistent with the relatively weak $K_{\text{d}} = k_{\text{off}}/k_{\text{on}}$ values for binding (Table 2.3). A few amides, including S47 and N69, did exhibit intermediate exchange broadening ($k_{\text{ex}} \sim \Delta\omega$). This can be explained by their intimate involvement in the binding interface and hence a large $\Delta\omega$ accompanying phosphopeptide association. Since only 50 – 60 % saturation was achieved in the titration experiments with FHA-1, sharpening and reappearance of these signals from the fully bound state was not observed.

In contrast to FHA-1, numerous amides in FHA-2 showed chemical shift perturbations upon binding either phosphopeptide (Figure 2.8A and Appendices A.3,4). In accordance with K_{d} values in the 10 μM range, amides with larger $\Delta\omega$ values also exhibited linebroadening at intermediate titration points (fast-intermediate exchange regime). Furthermore, the signals from numerous amides disappeared in the presence of sub-stoichiometric amounts of peptide without the concomitant appearance of new signals from the bound state (90 – 95 % saturation) as would be expected for slow exchange binding. This is suggestive of conformational exchange broadening in the bound state. Most residues with complete peak attenuation in spectra of FHA-2 are located in or close to the canonical peptide binding interface, including R234, S248 and N270. It is also noteworthy that the $^1\text{N}^{\text{H}}\text{-}^{15}\text{N}$ signals of residues in the appended C-terminal strand 1*, as well as strand 11, in FHA-2 were more perturbed than those in the equivalent

strands 1 and 11 of FHA-1. Thus, the circularly permuted region of FHA-2 is also affected by phosphopeptide binding.

Overall, our NMR titrations reveal that pT152 and pT210 phosphopeptide binding leads to significantly more amides showing larger chemical shift perturbations and exchange broadening in FHA-2 compared to FHA-1. This correlates with their relative affinities, and points to a larger binding interface and/or greater conformational perturbations upon binding for FHA-2 than FHA-1. Another explanation for a greater/more global chemical shift change could be a change in the hydrogen bond status in FHA-2 upon binding. The bound state of FHA-2 also appears to undergo conformational exchange as indicated by the absence of many amide signals upon saturation. Furthermore, results from ITC experiments indicate that pThr peptide binding is associated with an entropic loss in FHA-2, which is offset by a large enthalpy change. Taken together, these results suggest that the conformational adaptability of FHA-2 allows for a more favorable bound state. The direct correlation of phosphopeptide affinity with flexibility for the FHA domains is bolstered by denaturation and amide hydrogen exchange (HX) studies.

2.3.5 FHA-2 is More Dynamic and Less Stable Than FHA-1

To help understand their different structural and phosphopeptide binding properties, I probed the stability and dynamics of FHA-1 and FHA-2 both experimentally and computationally. In chemical denaturation studies monitored with circular dichroism (CD) spectroscopy, the domains unfolded at distinctly different mid-point $[\text{GuHCl}]_{1/2}$ of 1 M (FHA-1) and 0.5 M (FHA-2) (Figures 2.12A-C). Non-linear regression of these data yielded values of 3.2 ± 0.3 and 2.3 ± 0.8 kcal/mol for $\Delta G_{u,H_2O}$, the extrapolated free energy change upon unfolding of FHA-1 and FHA-2, respectively, in the absence of denaturant. By both criteria, FHA-2 is substantially less stable than FHA-1.

This conclusion is supported by protection factors (PFs) obtained from amide HX measurements (Figures 2.12D,E). Under EX2 conditions, PFs provide a measure of the residue-specific free energy changes, $\Delta G_{HX} = 2.303RT \log(\text{PF})$, governing local or global conformational equilibria leading to exchange. The most slowly exchanging amides in FHA-1 have PFs ~ 100 -fold greater than those in FHA-2 ($\log(\text{PF})_{\text{max}} \sim 5.75$ versus 3.5). Assuming these amides exchange via global unfolding (263, 264), this corresponds to ΔG_{HX} values of 7.8 kcal/mol and

4.8 kcal/mol for the FHA domains, respectively. Differences between these values and those obtained via chemical denaturation experiments are attributed to the different experimental conditions used for the two approaches and the extrapolations made to extract free energy changes. Nevertheless, the key conclusion is that FHA-1, with a canonical 11-strand FHA domain fold, is more stable than the circularly permuted 10-strand FHA-2.

HX measurements also reveal that circular permutation has a clear effect on the dynamics of various parts of the FHA domain. Whereas the interacting N- and C-terminal β -strands 1 and 11 are highly protected in FHA-1, the terminal strand 3 and 1* of FHA-2 are not interacting and much less protected. Furthermore, amides in or close to the binding interface of FHA-2, including those of residues following strand 3 and in the loop connecting strands 6 and 7, have lower PFs than the corresponding amides in FHA-1. This finding indicates higher relative flexibility for these regions of FHA-2. Importantly, this aligns well with the segments of FHA-2 that display pronounced chemical shift changes upon phosphopeptide binding (Figure 2.8). The correlation of large peptide-dependent spectral perturbations with low PFs in the unbound state suggests that FHA-2 has a conformationally dynamic binding interface. In contrast, FHA-1 shows somewhat more uniform PFs, with strands 1 and 2 well protected from HX. The greater stability and lower flexibility of FHA-1 may explain the relatively small chemical shift perturbations upon phosphopeptide binding.

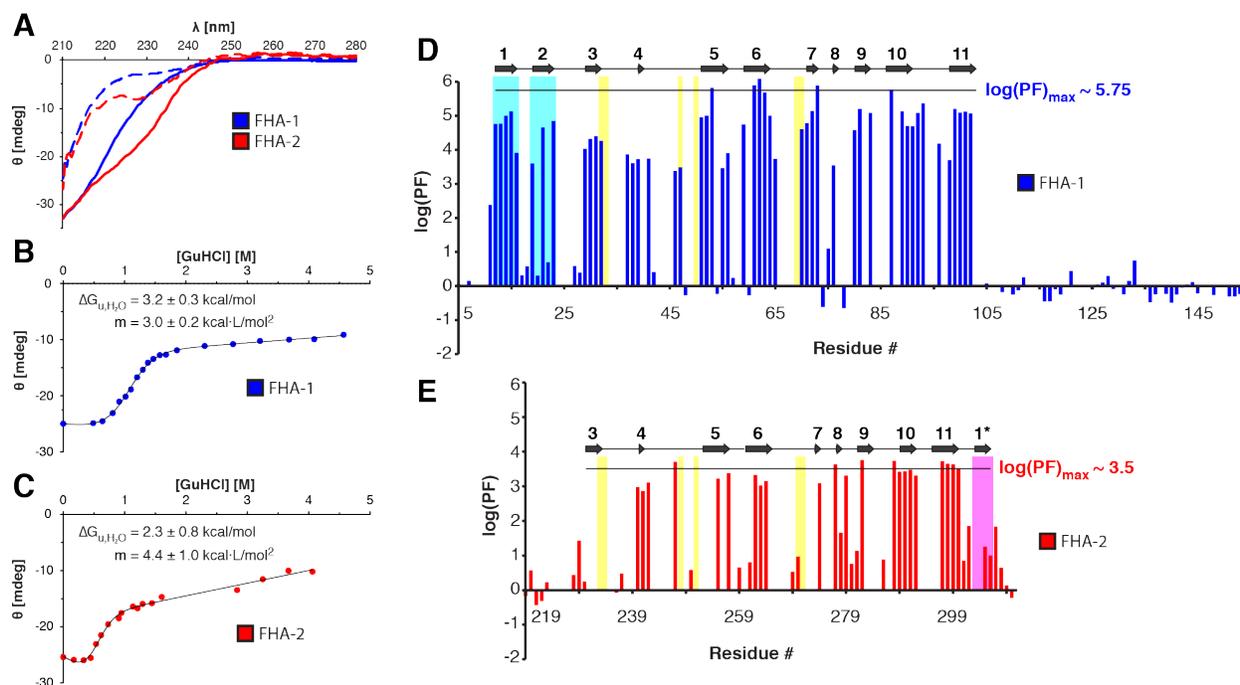


Figure 2.12: FHA-1 is More Stable than FHA-2.

(A) CD spectra of FHA-1 (Rv1747¹⁻¹⁵⁶; blue) and FHA-2 (Rv1747²⁰⁶⁻³¹⁰; red) in their folded (solid lines, 0 M GuHCl) and unfolded (dashed lines, 5 M GuHCl) states. CD signals for (B) FHA-1 (blue, 217 nm) and (C) FHA-2 (red, 217.5 nm) as a function of GuHCl concentration at 25 °C. The data were fit as outlined in the Methods to obtain extrapolated $\Delta G_{u,H_2O}$ values of 3.2 ± 0.3 and 2.3 ± 0.8 kcal/mol for the unfolding of FHA-1 and FHA-2, respectively, in the absence of denaturant. Histograms of combined amide $^1H/^2H$ (pH* 6) and $^1H/^1H$ (pH 6 – 8) exchange protection factors (PFs) measured for (D) FHA-1 and (E) FHA-2 at 25 °C. The horizontal lines indicate the largest PFs values for each domain, attributed to slowest exchange via global unfolding. In contrast, the linker residues have PFs ~ 1 , consistent with their conformational disorder. Conserved residues important for binding are highlighted in yellow. The β -strands 1 and 2 present in only FHA-1 and the additional C-terminal strand 1* in FHA-2 are highlighted in light cyan and magenta, respectively. Missing data points correspond to prolines, amides with overlapping or unassigned signals, or amides for which reliable slow ($^1H/^2H$) or fast ($^1H/^1H$) exchange rate constants could not be obtained at the conditions examined.

To further investigate the dynamics of both FHA domains, I conducted 900 ns all-atom MD simulations in explicit water using the AMBER force field. Over the course of the runs, FHA-1 and FHA-2 remained stable with average C α RMS fluctuations from the corresponding experimental structures of only 0.75 Å and 1.35 Å, respectively (Figures 2.13A,B). AMBER back-calculated B-factors, which report on the magnitude of structure dynamics in the nsec- μ sec timescale, showed that for both domains, the β -strands were rigid whereas the connecting hairpins and loops were mobile (Figures 2.13C-F). Most strikingly, for strand 1*, the fluctuations in FHA-2 were significantly larger than in the corresponding regions in FHA-1. This higher mobility in strand 1* might be caused by a lack of a stabilizing strand 2 and corroborates the low PFs measured for this region of FHA-2.

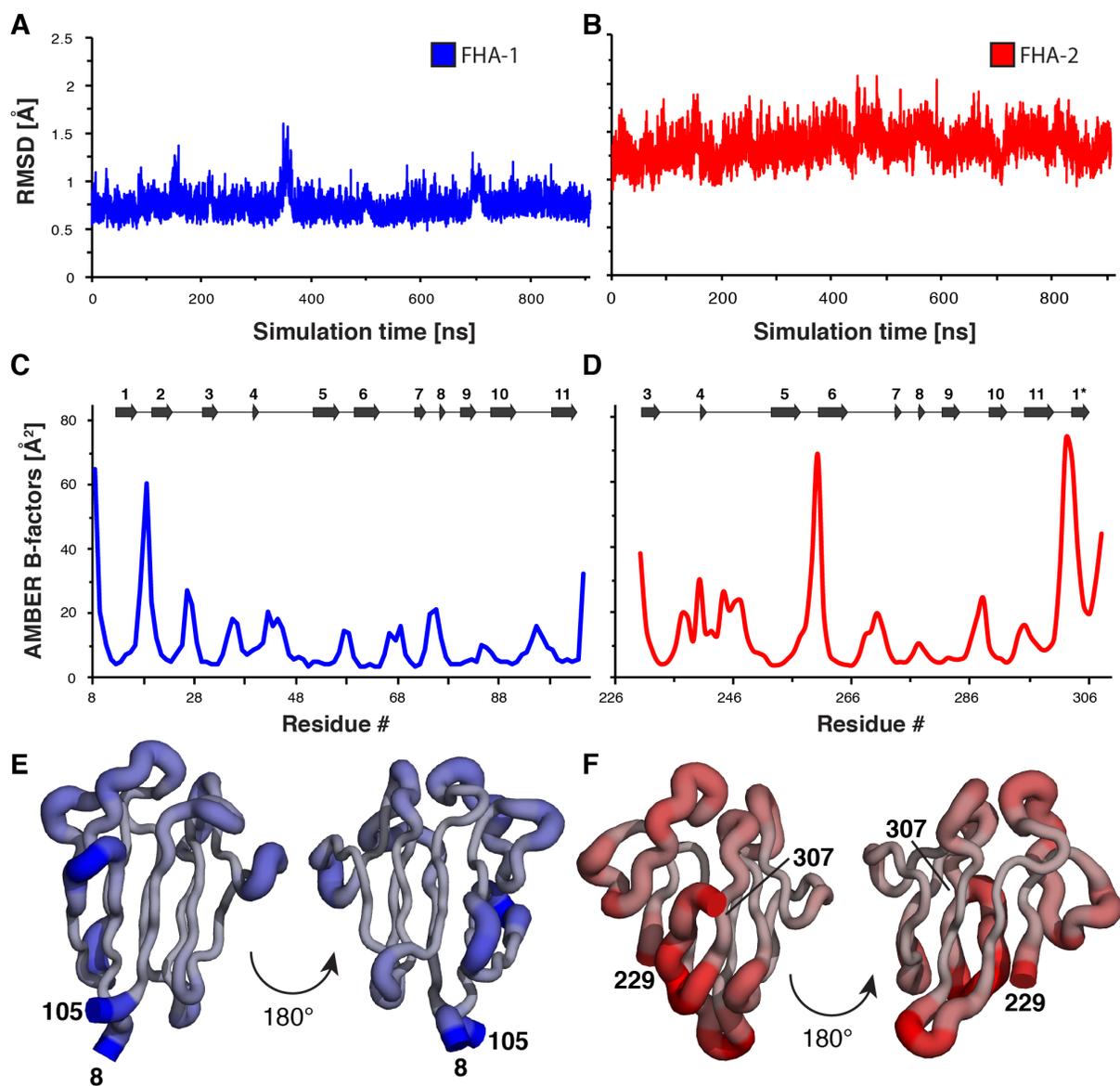


Figure 2.13: Molecular Dynamics Simulations of Rv1747 FHA Domains.

Backbone (N, C $^{\alpha}$, CO) RMS deviations of (A) FHA-1 (Rv1747⁸⁻¹⁰⁵) and (B) FHA-2 (Rv1747²²⁵⁻³¹⁰) over a 900 ns AMBER MD simulation with solvation in an explicit water model. Both domains remained stable with average RMSDs from the corresponding experimental structures of only 0.75 Å and 1.35 Å, respectively. Backbone AMBER B-factors, back-calculated from 900 ns MD simulations of (C, E) FHA-1 and (D, F) FHA-2. Higher values are indicated on the corresponding structures by a thicker backbone and by a grey-blue or grey-red gradient.

2.3.6 FHA-1 and FHA-2 Bind the Phosphorylated Linker to Form Inter- and Intramolecular Complexes

Both FHA-1 and FHA-2 bind short phosphopeptides corresponding to pT152 and pT210. However, within their native context, the FHA domains and phospho-acceptors lie within the same polypeptide chain. Therefore, I investigated the effects of phosphorylation on two constructs, Rv1747¹⁻²¹³ and Rv1747¹⁴⁸⁻³¹⁰, spanning FHA-1 or FHA-2 and the two reported phospho-acceptor threonines in the following or preceding linker region, respectively. The proteins were modified *in vitro* using the catalytic domain of PknF. The addition of two phosphates was confirmed by mass spectrometry. NMR spectral changes (Figures 2.14 and 2.15), as well as the effects of phospho-ablative alanine substitutions (Figures 2.17A,B), were consistent with T152 and T210 as the specific sites of modification.

First, I carried out ¹⁵N-HSQC-monitored titrations of ¹⁵N-labeled non-phosphorylated samples of Rv1747¹⁻²¹³ (Figures 2.15A,B) and Rv1747¹⁴⁸⁻³¹⁰ (Figures 2.15D,E) with doubly phosphorylated, unlabeled (NMR-silent) versions of the same construct. The chemical shift perturbations of backbone amides observed in these titrations paralleled those with either synthetic peptide. From this, I unambiguously conclude that intermolecular (*trans*) binding occurs with both constructs such that pT152 and/or pT210 on one molecule can bind the FHA domain on another molecule.

To investigate the possibility of intramolecular (*cis*) FHA-pThr binding, I compared the ¹⁵N-HSQC spectra of ¹⁵N-labeled Rv1747¹⁻²¹³ and Rv1747¹⁴⁸⁻³¹⁰ in their unmodified versus doubly phosphorylated forms (Figures 2.14 and 2.15C,F). *In vitro* phosphorylation of the linker threonines of these constructs (~ 300 μM) caused amide chemical shift perturbations in their FHA domains that approximately matched those observed at the titration endpoints with the phosphorylated peptides. However, these endpoints, with ~ 2.4 mM pT152 and ~ 2.1 mM pT210 phosphopeptide (~ 300 μM protein), corresponded to ~ 50 – 60 % saturation of the FHA-1 construct and 90 – 95 % of FHA-2 (Figures 2.15A,D). In contrast, the doubly phosphorylated proteins were ~ 300 μM, and thus seven- or eight-fold more dilute than the phosphopeptides.

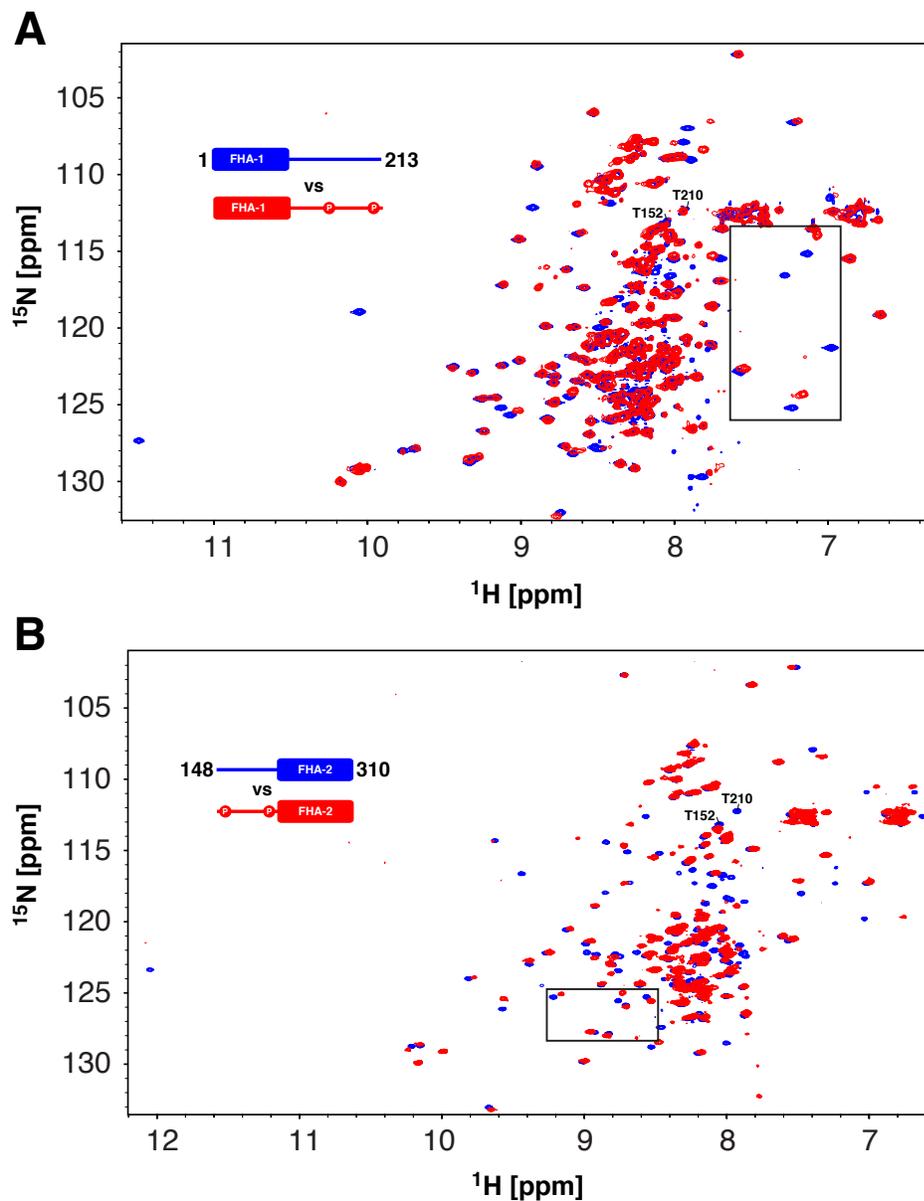


Figure 2.14: NMR Spectral Comparison of Phosphorylated versus Unmodified Rv1747 Constructs.

Overlaid ^{15}N -HSQC spectra of ^{15}N -labeled FHA-1 (Rv1747¹⁻²¹³; top) and FHA-2 (Rv1747¹⁴⁸⁻³¹⁰; bottom) with the linker residues T152 and T210 unmodified (blue) and phosphorylated by PknF (red). Although partially obscured by spectral overlap, the $^1\text{H}^{\text{N}}\text{-}^{15}\text{N}$ signals of T152 and T210 shift to unassigned chemical shifts upon phosphorylation. Additional changes in chemical shift and signal intensity (exchange broadening and increased molecular mass) result from inter- and intramolecular FHA-pThr interactions. The boxed regions are expanded in Figures 2.15C and F.

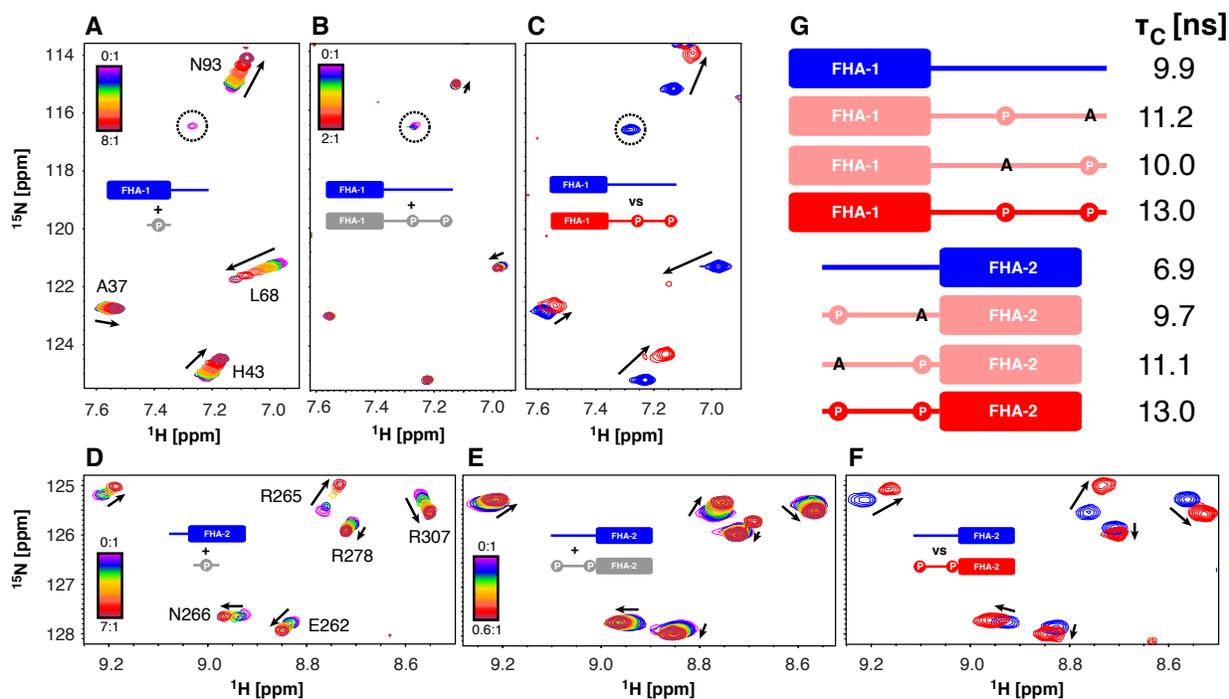


Figure 2.15: FHA-pThr Mediated Intermolecular (*trans*) Interactions and Oligomerization.

Overlaid ^{15}N -HSQC spectra showing the titration of (A) ^{15}N -labeled FHA-1 (Rv1747¹⁻¹⁵⁶) with unlabeled pT152 peptide (similar changes occur with the pT210 peptide; not shown) and (B) ^{15}N -labeled FHA-1 (Rv1747¹⁻²¹³) with unlabeled FHA-1 (Rv1747¹⁻²¹³) phosphorylated at T152 and T210. (C) Overlaid ^{15}N -HSQC spectra of ^{15}N -labeled FHA-1 (Rv1747¹⁻²¹³) in its unmodified (blue) and PknF-phosphorylated (red) forms. Overlaid ^{15}N -HSQC spectra showing the titration of (D) ^{15}N -labeled FHA-2 (Rv1747²⁰⁶⁻³¹⁰) with unlabeled pT210 peptide (similar changes occurred with the pT152 peptide; not shown) and (E) ^{15}N -labeled FHA-2 (Rv1747¹⁴⁸⁻³¹⁰) with unlabeled FHA-2 (Rv1747¹⁴⁸⁻³¹⁰) phosphorylated at T152 and T210. (F) Overlaid ^{15}N -HSQC spectra of ^{15}N -labeled FHA-2 (Rv1747¹⁴⁸⁻³¹⁰) in its unmodified (blue) and PknF-phosphorylated (red) forms. To enable comparison, panels A and D are reproduced from Figure 2.8A. For the titrations series, which detect intermolecular binding equilibria, rainbow coloring shows the molar ratios of phosphorylated unlabeled (NMR-silent) to unmodified labeled (NMR-observed) species, arrows highlight amide $^1\text{N}^{\text{H}}\text{-}^{15}\text{N}$ peaks that shift in fast exchange, and dashed circles denote exchange broadened signals. Note that the smaller chemical shift perturbations seen upon titration with the phosphorylated proteins versus the phosphopeptides reflect the lower final molar ratios for the former titrations, combined with unobserved intramolecular binding equilibria that compete with the observed intermolecular binding. This is particularly evident for FHA-1, which has higher K_d values for the phosphopeptides than FHA-2 (Table 2.3). In panels C and F, the arrows identify peaks that shift upon phosphorylation-mediated intra- and intermolecular association (see Figure 2.14 for full spectra). (G) Cartoon representations of the FHA-1 (Rv1747¹⁻²¹³) and FHA-2 (Rv1747¹⁴⁸⁻³¹⁰) constructs and their corresponding isotropic correlation times τ_c for rotational diffusion, measured via ^{15}N relaxation experiments (Figure 2.16). All

proteins were at $\sim 300 \mu\text{M}$. PknF-phosphorylated threonines are shown as a circled P, threonine to alanine phospho-ablative mutations are indicated by an A, and unlabeled species are in light grey.

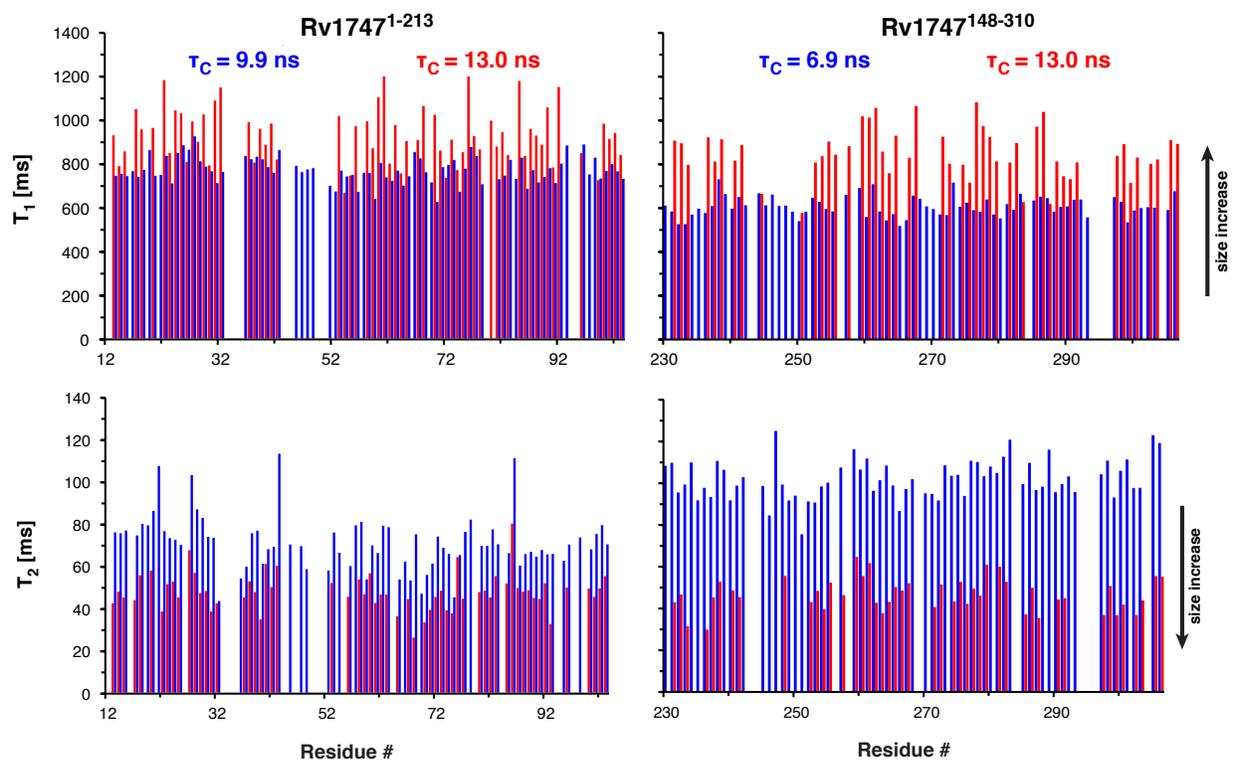


Figure 2.16: Oligomerization of Rv1747 FHA Domain Constructs upon Phosphorylation. Phosphorylation of the linker residues T152 and T210 in the FHA-1 (Rv1747¹⁻²¹³) and FHA-2 (Rv1747¹⁴⁸⁻³¹⁰) constructs leads to oligomerization. Shown are the amide ¹⁵N T₁ and T₂ lifetimes of the two constructs in their unmodified (blue) and dual phosphorylated (red) forms. The increased T₁ and decreased T₂ values of the phosphorylated species indicate slower rotational diffusion (longer isotropic correlation times τ_C for global tumbling) due to FHA-pThr mediated association. The two constructs differ in polypeptide chain length (213 versus 162 residues, respectively) and hence in relative relaxation lifetimes.

The apparent higher net affinity of the FHA domains for phosphorylated linkers (i.e. in the context of the Rv1747¹⁻²¹³ and Rv1747¹⁴⁸⁻³¹⁰ constructs) versus short phosphopeptides is attributed to two possible mechanisms that augment the demonstrated intermolecular interactions. The first of these is intramolecular (*cis*) binding in which one of the phosphorylated threonines binds the FHA domain within the same molecule. In the second mechanism, which can be considered as avidity or multivalent cooperativity (265), two or more molecules form a complex via intermolecular interactions and then FHA domains and phosphothreonines within the complex associate. In each case, the intramolecular or intracomplex interactions are entropically favored due to the increased local effective concentrations of the FHA domains and phosphorylated linkers. However, dissecting the contributions of the possible microscopic intra- and intermolecular binding equilibria to the overall macroscopic association of the Rv1747 constructs is not easily accomplished. Although in principle distinguishable through concentration dependent studies, this is limited in practice by the low sensitivity of NMR spectroscopy.

Further insights in the association states of the Rv1747 constructs were gained from their correlation times τ_C for global isotropic rotational diffusion as measured via amide ¹⁵N relaxation experiments (Figures 2.15G and 2.16). In their unmodified states, Rv1747¹⁻²¹³ (23 kDa) and Rv1747¹⁴⁸⁻³¹⁰ (17 kDa) had correlation times of 9.9 ns and 6.9 ns, respectively. These values are consistent with monomeric species composed of a folded FHA domain and an appended ID linker. Upon dual phosphorylation, these correlation times increased to 13 ns, confirming that the proteins form intermolecular complexes. Although roughly indicative of dimers, it is important to note that these correlation times will be averaged over all forms of the constructs present, including monomers with intramolecular FHA-pThr interactions, dimers with each FHA domain binding a partner phosphothreonine, and a variety of possible oligomers. Variants of Rv1747¹⁻²¹³ and Rv1747¹⁴⁸⁻³¹⁰ with alanine substitutions at positions 152 or 210 to restrict phosphorylation to a single site in the ID linker were also investigated. Although capable of dimerization, these species gave τ_C values intermediate to those of the unmodified and dual phosphorylated proteins. This further suggests that intramolecular interactions contribute to the FHA-pThr binding observed by NMR spectroscopy, and in the case of the singly phosphorylated proteins, compete with intermolecular binding. However, with two linker phosphothreonines, both binding modes

can occur to yield, on average, dimeric-like complexes. Indeed, it is interesting that Rv1747¹⁻²¹³ and Rv1747¹⁴⁸⁻³¹⁰ do not make extended head-to-tail polymers upon double phosphorylation. This also suggests that intramolecular or intracomplex interactions occur, thereby favoring dimeric or limited oligomeric species.

NMR spectral perturbations resulting from the single-site phosphorylation of the T152A and T210A variants provided evidence for differential interactions of the two FHA domains and the two linker phospho-acceptors (Figure 2.17). In the case of the Rv1747¹⁻²¹³ constructs, phosphorylation at pT152 yielded only small chemical shift changes of diagnostic amides relative to the unmodified protein, whereas phosphorylation at pT210 caused changes resembling those resulting from double phosphorylation of the wild-type protein. In the case of the Rv1747¹⁴⁸⁻³¹⁰ constructs, pT152 yielded intermediate chemical shift changes between the unmodified and fully phosphorylated end points, whereas pT210 also caused spectral perturbations resembling those resulting from double phosphorylation. These spectral changes are difficult to interpret in detail due to *cis* and *trans* binding modes and an order-of-magnitude difference in phosphothreonine peptide binding affinity exhibited by FHA-1 versus FHA-2 (Table 2.3). Nevertheless, the data point to differences in the accessibility, or local concentration, of the phosphothreonines in the ID linker with respect to the flanking FHA domains.

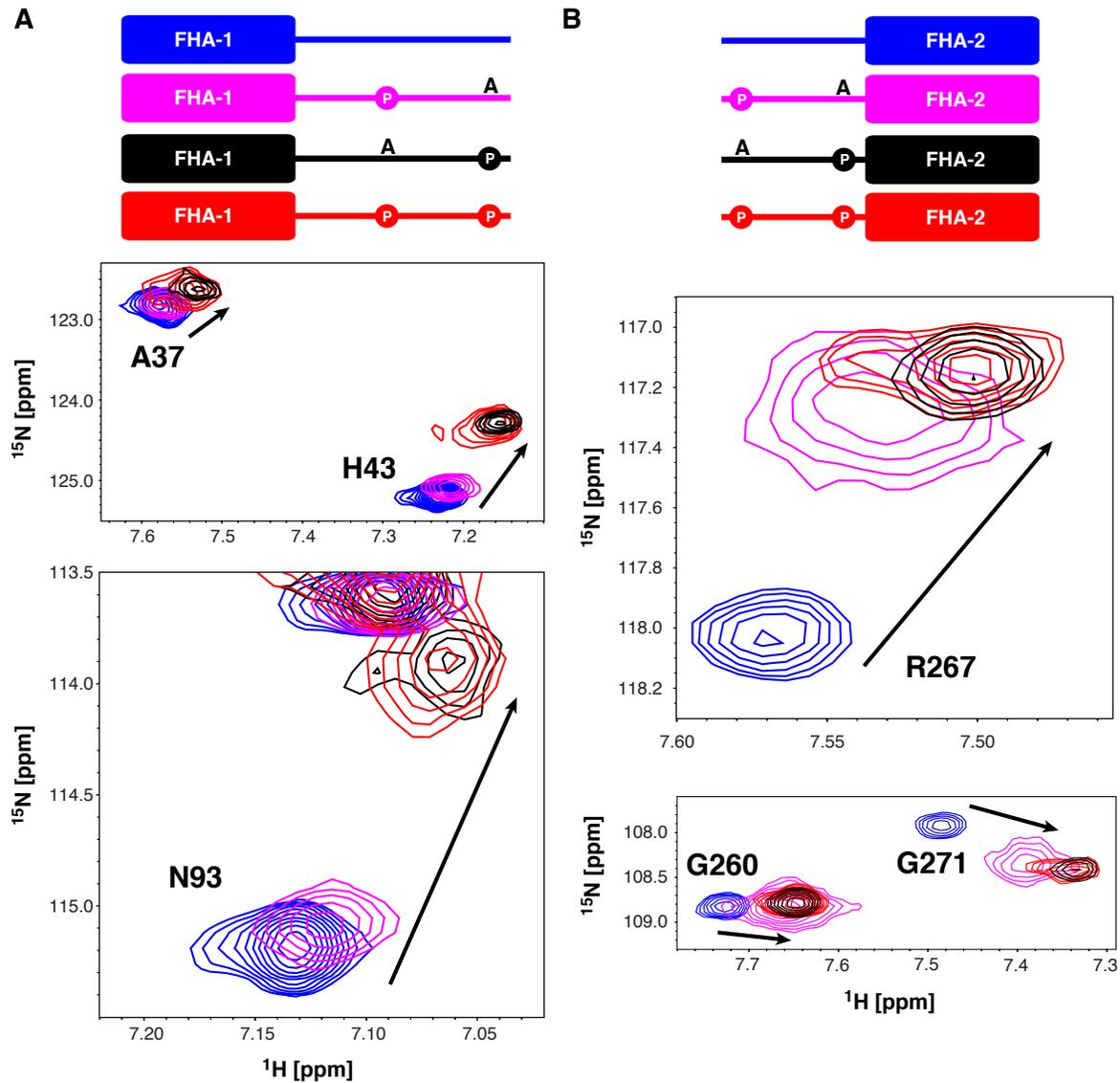


Figure 2.17: Intra- and Intermolecular Interactions of Rv1747 upon Phosphorylation

Overlaid regions of ^{15}N -HSQC spectra of uniformly ^{15}N -labeled (A) FHA-1 (Rv1747¹⁻²¹³) and (B) FHA-2 (Rv1747¹⁴⁸⁻³¹⁰) show differential effects of phosphorylation. Peaks from the unmodified wild-type proteins are in blue, from the T210A mutants with single phosphorylation at pT152 in pink, from the T152A mutants with single phosphorylation at pT210 in black, and from the wild-type proteins with dual phosphorylation at pT152 and pT210 in red. The directions of progressive chemical perturbation upon increased FHA-pThr binding are indicated with arrows.

2.3.7 Phosphorylated Regulatory Module Forms Higher Order Complexes

I also studied the effect of phosphorylation on the full FHA regulatory module (Rv1747¹⁻³¹⁰), which contains in one polypeptide chain both FHA-1 and FHA-2 as well as the ID linker with two phospho-acceptor threonines (Figure 2.18). Two hours after addition of PknF directly to the ¹⁵N-labeled protein in an NMR tube, I observed a significant decrease in the intensities of ¹H^N-¹⁵N signals from amides within the structured FHA domains, and to a lesser extent, from the ID linker. Also, NMR signals diagnostic of pThr-bound FHA domains did not appear. After longer incubation times, the sample became visibly turbid and only signals from the ID linker were detected (not shown). A simple interpretation of these results is that multivalent interactions between the two FHA domains and two phosphorylated threonines within the intervening linker leads to the formation of high molecular mass complexes. In these complexes, the FHA domains are effectively invisible due to rapid relaxation, whereas at least portions of the ID linker are sufficiently flexible to still be detectable in ¹⁵N-HSQC spectra.

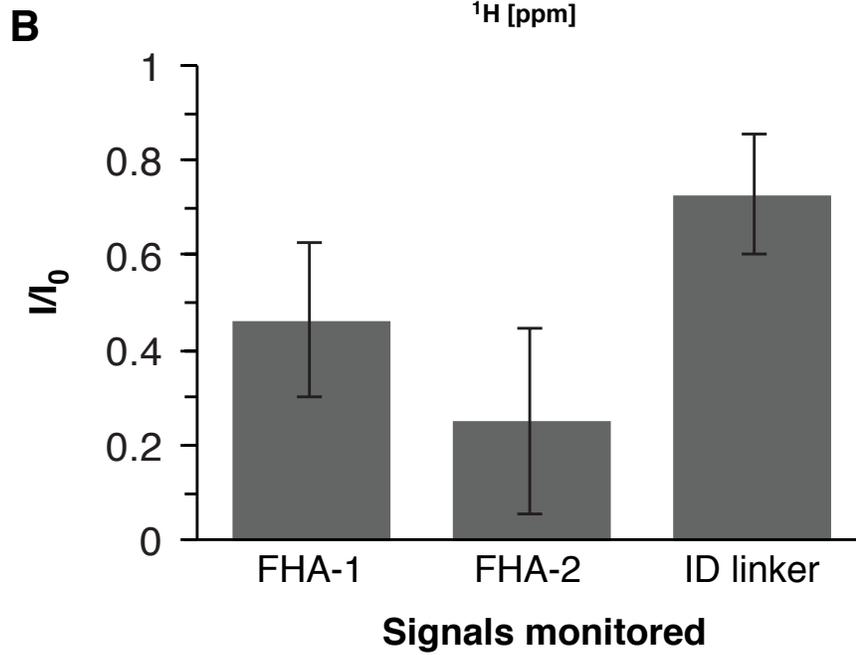
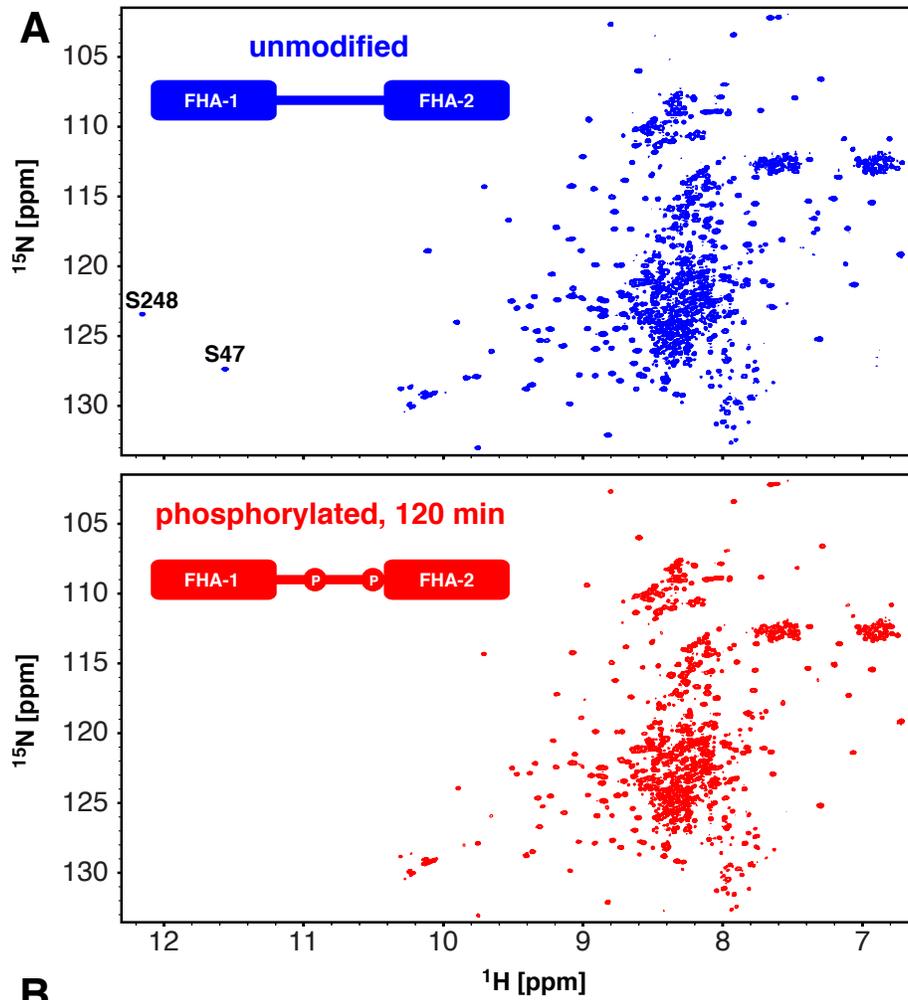


Figure 2.18: Phosphorylation of the Rv1747 Regulatory Module Leads to Higher Order Association.

(A) ^{15}N -HSQC spectra of the unmodified (top, blue) and PknF-phosphorylated regulatory module (bottom, red) spanning both FHA domains and the ID linker (Rv1747¹⁻³¹⁰). The unmodified protein was 200 μM in a buffer containing 20 mM sodium phosphate, pH 6.0, 100 mM NaCl, 5 mM MgCl_2 and 5 mM ATP. The spectrum of the phosphorylated protein was recorded starting 120 min. after addition of 2 μM PknF. Under these conditions, Rv1747¹⁻³¹⁰ becomes doubly phosphorylated as confirmed by ESI-MS. (B) Average intensity ratios and standard deviations of the corresponding $^1\text{H}^{\text{N}}\text{-}^{15}\text{N}$ signals originating from assigned residues in the structured cores of FHA-1 and FHA-2 before (I_0) and after PknF phosphorylation (I). These signals are attributed to the unbound domains or to residues that do not exhibit chemical shift perturbations upon phosphorylation. The overall loss of signal from these ordered domains indicates oligomerization to higher molecular mass species due to intermolecular FHA-pThr interactions. Unassigned amides from the ID linker show less intensity perturbations, likely due to a degree of mobility within these complexes.

2.4 Discussion

2.4.1 Beads-on-a-String Regulatory Module

I have characterized in detail the postulated regulatory module of the *Mtb* ABC transporter Rv1747. This module is composed of two independent FHA domain "beads" connected by an ID linker "string". In contrast to the well-ordered domains, the unmodified linker is conformationally disordered as judged by chemical shift, ^{15}N relaxation, and amide HX measurements. The structure of FHA-1, determined by X-ray crystallography, closely matches that of a canonical FHA domain with 6- and 5-stranded β -sheets forming a β -sandwich. Unexpectedly, the NMR-derived structural ensemble of FHA-2 revealed that this domain is circularly permuted such that a C-terminal strand 1* occupies the standard position of the first N-terminal strand 1. Also, the equivalent of strand 2 is absent. Although it is difficult to infer stability from structure, I speculate that the lower free energy of unfolding for FHA-2 relative to FHA-1, as measured globally by GuHCl denaturation ($\Delta G_{u,H_2O}$) and locally by HX studies (ΔG_{HX}), results from its altered secondary structure. Despite these differences, the two domains share similar global folds and peptide-binding interfaces, including a conserved hydrogen bond network involving a buried, neutral histidine. Thus the functional significance underlying the circular permutation of FHA-2 is unknown. A consequence of shifting the N-terminus of FHA-2 is that the phospho-acceptor T210 lies within the ID linker, and not in β -strand 2, and thus is accessible for modification and binding. However, from an evolutionary perspective, moving the position of the phospho-acceptor within a nonconserved ID linker seems simpler than permuting the conserved β -strand topology of an FHA domain.

2.4.2 The More Dynamic FHA-2 has Higher Affinity for Phosphopeptides

The affinities of the isolated Rv1747 FHA domains for short phosphopeptide models of the mapped acceptor sites, pT152 and pT210, were measured by NMR spectroscopy and ITC. Both FHA domains bound both peptides with little discrimination. This is consistent with the similarity of the peptides at the pT and pT+3 positions, which serve as key contact points to canonical FHA domains. However, FHA-2 bound the two peptides with 20 – 30 fold higher affinity than FHA-1. The reason for this modest difference (~ 2 kcal/mol on a free energy scale)

is not completely obvious since the two domains have similar residues within their binding interfaces.

More striking than the differences in K_d values are the differences observed in NMR-monitored titrations. In the case of FHA-1, a relatively small number of amides showed chemical shift perturbations in the fast exchange limit. These amides mapped closely to the expected peptide binding interface. Although the structure of a FHA-1/peptide complex was not determined, this spectroscopic behaviour suggests that FHA-1 has a well-defined interface, which undergoes little conformational change upon binding. In contrast, numerous amides mapping to a more extensive interface in FHA-2 exhibited chemical shift changes and/or exchange broadening upon addition of the phosphopeptides. Although attributable to a complex interplay of K_d , k_{on} , k_{off} , and $\Delta\omega$ values, this strongly suggests that peptide binding causes more extensive conformational changes in FHA-2 than in FHA-1. Furthermore, even at near saturation, signals from several FHA-2 amides did not appear, implying that msec- μ sec timescale conformational exchange still occurs within the bound state. Without high-resolution models of the bound complexes, the magnitudes of these changes are difficult to estimate, yet are likely on the Å scale given the similarities of the unbound FHA-1 and FHA-2 structures.

Combining these results, I hypothesize that the higher affinity of FHA-2 versus FHA-1 for phosphopeptides is linked to its lower stability and greater conformational dynamics. These dynamic properties are seen in particular through HX studies, with FHA-2 having low residue-specific PFs, as well as through MD simulations. The correlation of dynamics and ligand affinity suggests that the conformational plasticity of FHA-2 enables a more favorable bound state with phosphothreonines motifs in the Rv1747 linker. Although FHA domains have been considered to be relatively rigid pThr binding domains (137, 266, 267), there are cases in which dynamics play an important role in ligand recognition (137, 268, 269). Circularly permuted FHA-2 is a strikingly example of the latter.

2.4.3 FHA Domains Bind Phosphorylated Linker Threonines via Inter- and Intramolecular Equilibria

Using NMR spectroscopy, I demonstrated that fragments of Rv1747 spanning one or both FHA domains and the intervening linker exhibit intermolecular (*trans*) and intramolecular

(*cis*) modes of FHA-pThr association. Unfortunately, it was not possible to quantitatively dissect the contributions of a myriad of possible microscopic binding equilibria towards these processes. Qualitatively, phospho-ablative T152A and T210A variants with single phosphothreonines had correlation times τ_c for rotational diffusion intermediate to those of the monomeric unmodified proteins and the (on-average) dimer-like doubly phosphorylated species. Such diminished intermolecular association is expected because the overall concentration of phosphothreonines in the monophosphorylated proteins is effectively reduced relative to the doubly modified proteins. The position of the phosphothreonine (pT152 or pT210) in the linker relative to FHA-1 and FHA-2 also influenced the measured τ_c values and resulted in different patterns of diagnostic ^{15}N -HSQC spectral perturbations. This stands in contrast to the approximately equal affinity of peptide models of the pT152 and pT210 sites for each FHA domain. Collectively, these results indicate that, when connected within a polypeptide chain, the two phosphothreonines have different accessibilities or local concentrations for binding FHA-1 versus FHA-2 to form intra- and intermolecular complexes. This may enable tunable regulation Rv1747 through a gradation of phosphorylation-dependent self-association equilibria.

2.4.4 Potential Mechanisms for Regulating Rv1747 Transporter Function

The exact biochemical functions of Rv1747 remain to be established. However, growing evidence convincingly implicates this protein in the transport of *Mtb* cell wall precursors. A distinguishing feature of Rv1747 is its appended N-terminal regulatory module with phospho-acceptor sites for the *Mtb* STPK PknF. Evidence for the phosphorylation-dependent regulatory role of this module follows from reduced infectivity of *Mtb* strains harboring Rv1747 variants with phospho-ablative T152A/T210A mutations in the linker region or mutations in the binding cleft of FHA-1 (144). Complete phosphorylation of T152 and T210 is seen when the isolated regulatory module is treated with the catalytic domain of PknF *in vitro*. However, other *Mtb* STPKs have been reported to phosphorylate Rv1747 (145) and additional phosphorylation sites have been identified by proteomics studies (270). Thus *in vivo*, the phosphorylation levels of Rv1747 are likely dependent on the concentrations and activities of several kinases, which in turn are susceptible to environmental stimuli (78, 91), as well as on PstP, the sole *Mtb* serine/threonine protein phosphatase (126). Further layers of control could also arise, for

example via FHA domain recruitment of pThr-modified STPKs (142, 271) or phosphoproteins such as the recently identified stress factor Rv2623 (105). Collectively, this would lead to a sophisticated regulatory system for Rv1747 that could be tuned on multiple levels to enable the growth and survival of *Mtb* within host macrophages.

According to the canonical ABC transporter architecture, the expected active quaternary state of Rv1747 is that of at least a homodimer. I thus propose that intermolecular pThr-FHA binding between regulatory modules either favors the assembly of Rv1747 into functional dimers or higher order oligomers, or alters intracomplex interactions (e.g., of the nucleotide binding domains) to impact the activity of existing dimers/oligomers. With two FHA domains and at least two phospho-acceptor threonines showing differential levels of intra- and intermolecular binding, either general mechanism could lead to graded, rather than all-or-none, regulation of Rv1747. This is in line with previous studies showing that proteins containing both an FHA domain and phospho-acceptor threonines can self-associate upon phosphorylation (272, 273) or exhibit intramolecular binding to autoinhibit heterotypic partnerships (118, 274). Indeed, FHA domains were found to be enriched in proteins containing sequences predicted to be autoinhibitory (275). However, providing evidence for the existence of Rv1747 oligomers in the *Mtb* cell membrane requires further studies using electron or super resolution light microscopy.

2.5 Materials and Methods

2.5.1 Protein Expression and Purification of Rv1747 Constructs for NMR Spectroscopy

Using *M. tuberculosis* H37Rv genomic DNA as a template, genes encoding Rv1747¹⁻³¹⁰, Rv1747¹⁻¹⁵⁶, Rv1747²⁰⁶⁻³¹⁰, Rv1747¹⁻²¹³, Rv1747¹⁴⁸⁻³¹⁰ were PCR amplified and cloned with a TAA stop codon into the pET28MHL plasmid (Addgene, plasmid #26096) via *Nde*I and *Hind*III restriction sites. Phospho-ablative threonine to alanine mutants of the Rv1747 constructs, as well as an N-terminal Trp codon in Rv1747²⁰⁶⁻³¹⁰ for quantitation, were introduced via the QuikChange site-directed mutagenesis protocol (Stratagene). After cleavage of an N-terminal His₆-affinity tag with TEV protease, two non-native residues (Gly-His) preceded the expressed Rv1747 fragments.

Protein constructs were expressed in *Escherichia coli* Rosetta 2 (λ DE3) cells in media supplemented with 35 μ g/mL chloramphenicol and 35 μ g/mL kanamycin. Cells were grown in

LB media to produce unlabeled protein or in M9 minimal media supplemented with 3 g/L $^{13}\text{C}_6$ -glucose and/or 1 g/L $^{15}\text{NH}_4\text{Cl}$ as the sole carbon and nitrogen sources, respectively, to produce uniformly labeled samples. Cultures of 1 L were incubated shaking at 37 °C until $\text{OD}_{600} \sim 0.6$, induced with isopropyl- β -D-thiogalactopyranoside (IPTG; 1 mM final) and grown for another 16 h at 24 °C. After centrifugation for 15 min at 4000 g, the cell pellets were frozen at -80 °C, then later thawed, resuspended in lysis buffer (20 mM sodium phosphate, 1 M NaCl, 10 mM imidazole, pH 8.0) and sonicated. After centrifugation for 1 h at 30000 g, the supernatant was filtered (0.45 μm cutoff), applied to a Ni^{2+} -NTA HisTrap HP affinity column (GE Healthcare Life Sciences), washed with 10 column volumes of binding buffer (20 mM sodium phosphate, 500 mM NaCl, 30 mM imidazole, pH 8.0) and eluted in a 30 – 500 mM imidazole gradient. Fractions containing the desired protein were identified with SDS-PAGE, pooled and treated with His₆-tagged TEV protease while dialyzing against 4 L of cleavage buffer (20 mM sodium phosphate, 100 mM NaCl, 1 mM DTT, 0.5 mM EDTA, pH 7) for 16 h at 4 °C. Uncleaved protein, cleaved tag and TEV were removed using a second Ni^{2+} -NTA HisTrap HP affinity purification step, and the collected flow-through was concentrated and subjected to size-exclusion chromatography using a Superdex 75 (GE Healthcare Life Sciences) column with NMR sample buffer (100 mM NaCl, 20 mM sodium phosphate, pH 6.0). Final protein fractions were pooled, concentrated using a 10 kDa MWCO centrifugal filter (EMD Millipore), flash frozen and stored for further use at -80 °C. Samples were verified by SDS-PAGE and electrospray ionization mass spectrometry (ESI-MS), and their concentrations determined by UV-absorbance spectroscopy using predicted ϵ_{280} molar absorptivities of 24000, 15000, 5500, 24000 and 8500 $\text{M}^{-1} \text{cm}^{-1}$ for Rv1747¹⁻³¹⁰, Rv1747¹⁻¹⁵⁶, Rv1747²⁰⁶⁻³¹⁰, Rv1747¹⁻²¹³ and Rv1747¹⁴⁸⁻³¹⁰, respectively (276).

2.5.2 Protein expression and purification of Rv1747³⁻¹¹⁶ for crystallography

Residues 3-116 of Rv1747 (FHA-1) were amplified from genomic *M. tuberculosis* DNA (strain H37Rv) using primers designed to introduce *Bam*HI and *Eco*RI restriction sites outside the area of interest. These fragments were ligated into the plasmid pGEX-2T (GE Healthcare Life Sciences). The resulting plasmid contained amino acids for Rv1747³⁻¹¹³ (FHA-1) fused to the C-terminus of GST and linked by a thrombin protease cleavage site, with the entire protein

under the control of a T7 promoter. The construct was transformed and expressed in *E. coli* BL21 (λ DE3) Codon Plus cells (strain RP; Stratagene). Saturated 5 mL starter cultures were used to inoculate 500 mL cultures of auto-induction medium (277) at a dilution of 1:1000. These 500 mL cultures were incubated for 18-24 hours at 37° C with vigorous shaking (> 260 rpm). After centrifugation for 10 – 15 min at 5000 g, cell pellets were frozen at -20 °C and then thawed and resuspended in ~ 40 mL of cold PBS buffer (140 mM NaCl, 2.7 mM KCl, 10 mM sodium phosphate, 1.7 mM potassium phosphate, pH 7.3), containing 0.5 mM AEBSF, 5 μ M E-64 and 1 μ M pepstatin A, and 1 mM EDTA as protease inhibitors. Cells were lysed by sonication and the supernatant was separated by centrifugation at 17000 g. The decanted supernatant was then applied by gravity flow to an appropriate volume of Glutathione Sepharose 4 Fast Flow (GE Healthcare) equilibrated with cold PBS. After binding, the column was washed with > 10 volumes cold PBS. The GST-FHA-1 protein was eluted with 10 – 20 mM reduced glutathione (Sigma-Aldrich) in 50 mM Tris-HCl at pH 8. Fractions were collected and those containing the protein of interest were identified by SDS-PAGE, pooled, and supplemented with CaCl₂ to a final concentration of 1 mM. The GST-affinity-tag was then cleaved by treatment with thrombin (MP Biomedicals; 10 U protease per 1 mg GST- fusion protein) for 4 – 12 hours with gentle agitation at 4° C. The cleaved protein fragments were applied to GST-binding Sepharose equilibrated with cold PBS. The flow-through from this column was further purified by size-exclusion chromatography on a Superdex S75 column equilibrated with 50 mM NaCl, 25 mM HEPES, 0.5 mM TCEP at pH 7.8. The purified Rv1747³⁻¹¹⁶ was concentrated and used for crystal growth.

2.5.3 Protein Expression and Purification of the PknF Kinase Domain

The gene encoding Rv1746¹⁻²⁹² (the cytoplasmic kinase domain of PknF) was cloned from *Mtb* H37Rv genomic DNA into the pGEX-4T3 plasmid (GE Healthcare Life Sciences) with a TAA stop codon using *Bam*HI and *Xho*I restriction sites. The resulting construct, with an N-terminal glutathione S-transferase (GST) affinity tag, was expressed in *E. coli* Rosetta 2 (λ DE3) cells in LB media supplemented with 35 μ g/mL chloramphenicol and 100 μ g/mL ampicillin. Cultures of 1 L were incubated at 37 °C until OD₆₀₀ ~ 0.6, induced with IPTG (1 mM final) and grown for another 16 h at 24 °C. After centrifugation for 15 min at 4000 g, the cell

pellet was frozen at -80 °C, then later thawed, resuspended in buffer (10 mM sodium phosphate, 1.8 mM potassium phosphate, 140 mM NaCl, 2.7 mM KCl, pH 7.3) and sonicated. The supernatant was centrifuged for 1 h at 30000 g, filtered, applied to a GSTrap HP affinity column (GE Healthcare Life Sciences), washed with 10 column volumes of the same buffer, and eluted in GST elution buffer (50 mM Tris, 10 mM reduced glutathione, pH 7.4). After SDS-PAGE analysis, pooled fractions were subjected to size-exclusion chromatography using a Superdex 75 (GE Healthcare Life Sciences) column in storage buffer (100 mM NaCl, 20 mM sodium phosphate, pH 7.4). Final protein fractions were pooled, concentrated using a 10 kDa MWCO centrifugal filter (EMD Millipore), flash frozen and stored for further use at -80 °C. Samples were verified by SDS-PAGE and ESI-MS, and concentrations determined by UV-absorbance spectroscopy using a predicted $\epsilon_{280} = 77000 \text{ M}^{-1} \text{ cm}^{-1}$ (276).

2.5.4 Synthetic pThr-containing Rv1747 Peptides

Phosphorylated peptides corresponding to the reported Rv1747 modification sites at pT152 (KKYAGQQLPPApT¹⁵²TRIPAA) and pT210 (KKYAGTEAGNLApT²¹⁰SMMK), were purchased from Biomatik. The N-terminal acetylated and C-terminal amidated peptides were obtained in a lyophilized form after HPLC purification. These peptides, used in previously reported binding studies of the Rv1747 FHA domains (144), include a non-native N-terminal KKYAG tag for solubility and for quantification by UV-absorbance spectroscopy (predicted $\epsilon_{280} = 1500 \text{ M}^{-1} \text{ cm}^{-1}$) (276).

2.5.5 *In vitro* Phosphorylation of Rv1747 Constructs

Samples of Rv1747 constructs were diluted to ~ 5 μM in ~ 15 mL of buffer (20 mM sodium phosphate, 100 mM NaCl, 5 mM MgCl_2 , pH 7.4), to which a 1:100 molar ratio of PknF:substrate protein was added directly. The bacterially expressed kinase is active, and the phosphorylation reactions were started by the addition of 5 mM ATP and incubated at 22 °C with repeated inversion for 120 min. After concentration using a 10 kDa MWCO centrifugal filter (EMD Millipore), the phosphorylated samples were subjected to size-exclusion chromatography using a Superdex 75 (GE Healthcare Life Sciences) column for purification and buffer exchange (100 mM NaCl, 20 mM sodium phosphate, pH 6.0). Phosphorylation levels

were ascertained using ESI-MS and phosphorylation sites confirmed indirectly from diagnostic amide $^1\text{H}^{\text{N}}\text{-}^{15}\text{N}$ chemical shift changes.

2.5.6 NMR Spectroscopy

NMR spectroscopic data were recorded at 25 °C on cryoprobe-equipped 500, 600 and 850 MHz Bruker Avance III spectrometers. All ^{15}N - or $^{13}\text{C}/^{15}\text{N}$ -labeled samples (0.1 – 0.6 mM) were prepared in NMR buffer (100 mM NaCl, 20 mM sodium phosphate, pH 6.0) with 5 – 10 % lock D_2O , unless stated otherwise. Data were processed and analyzed using NMRpipe (278) and NMRFAM-Sparky (279), respectively. Signals from mainchain ^1H , ^{13}C , and ^{15}N nuclei of Rv1747¹⁻¹⁵⁶ and mainchain and sidechain ^1H , ^{13}C , and ^{15}N nuclei of Rv1747²⁰⁶⁻³¹⁰ were assigned by standard multidimensional heteronuclear scalar correlation experiments (280). The charge and tautomer states of the histidine sidechains were determined as described previously (281). Chemical shift-based secondary structure and RCI-S² predictions (257) were performed using MICS (256).

2.5.7 NMR-based Structure Calculations of FHA-2 (Rv1747²⁰⁶⁻³¹⁰)

The structural ensemble of Rv1747²⁰⁶⁻³¹⁰ was calculated with CYANA 3.0 (282) using chemical shift assignments, TALOS+ dihedral angle restraints (283), and an unassigned list of cross peaks from 3D $^1\text{H}\text{-}^{15}\text{N}/^{13}\text{C}\text{-}^1\text{H}$ NOESY–HSQC spectra ($\tau_{\text{mix}} = 100$ ms) (284). All X-Pro bonds were constrained to the *trans* conformation by chemical shift criteria (285). Based on diagnostic $^{15}\text{N}^{\delta 1}$ and $^{15}\text{N}^{\epsilon 2}$ chemical shifts, the histidine chains were set to the neutral $\text{N}^{\epsilon 2}\text{H}$ tautomer for H250 and H251 and to the fully protonated imidazolium cation for H285 (260). Seven iterations of automatic NOESY cross peak and distance restraint assignments, along with stereospecific chemical shift assignments, were carried out. At each step, an ensemble of 100 structures was calculated, followed by a final ensemble calculation from which the 20 lowest energy structures were further refined using the original NOE distance restraints, NOE-like inter-proton distance restraints from the initially calculated structures, and a knowledge based Statistical Torsion Angle Potential (STAP) in NMRe (286). Consensus secondary structure boundaries for the Rv1747²⁰⁶⁻³¹⁰ ensemble were determined using DSSP (287) and the figures rendered using PyMol (288). See Table 2.2 for data and refinement statistics.

2.5.8 Backbone Amide ^{15}N Relaxation Experiments

Amide ^{15}N relaxation data (T_1 , T_2 and heteronuclear NOE) were recorded for the Rv1747 protein constructs ($\sim 300 \mu\text{M}$) on a cryoprobe-equipped 600 MHz Bruker Avance III spectrometer (289). Relaxation rate constants were determined in NMRFAM-Sparky (279) by fitting $^1\text{H}^{\text{N}}\text{-}^{15}\text{N}$ peak heights to a single exponential decay. The heteronuclear NOE values were determined from the ratio of the peak heights versus a control reference spectrum without ^1H saturation. Correlation times for global rotational diffusion were calculated with Tensor2 using these relaxation rate constants and assuming isotropic tumbling (258).

2.5.9 Amide HX Measurements

Slow amide protium-deuterium HX rate constants, k_{ex} , were measured at 25 °C for uniformly ^{15}N -labelled samples of $\sim 0.3 \text{ mM}$ FHA-1 (Rv1747¹⁻¹⁵⁶) and FHA-2 (Rv1747²⁰⁶⁻³¹⁰). The proteins in H_2O buffer (100 mM NaCl, 20 mM sodium phosphate, pH 6.0) were lyophilized and subsequently resuspended in the same volume of 99.9 % D_2O . After an initial delay of ~ 6 min, successive ^{15}N -HSQC spectra were recorded with a 600 MHz spectrometer for Rv1747¹⁻¹⁵⁶ (2.5 min/spectrum) and Rv1747²⁰⁶⁻³¹⁰ (5 min/spectrum), until almost all backbone amide $^1\text{H}^{\text{N}}\text{-}^{15}\text{N}$ signals disappeared. The sample pH* values (uncorrected pH meter readings) were measured after completion of exchange, and peak heights versus time were fit to a single exponential decay with NMRFAM-Sparky (279) to obtain pseudo-first order k_{ex} rate constants. Rapid protium-protium HX was measured using the CLEANEX-PM (290) approach with uniformly ^{15}N -labelled samples of $\sim 0.3 \text{ mM}$ FHA-1 (Rv1747¹⁻¹⁵⁶) and FHA-2 (Rv1747²⁰⁶⁻³¹⁰) in H_2O buffer (20 mM sodium phosphate, 100 mM NaCl, 5 % v/v D_2O lock) at 25 °C and pH 6, 7 and 8. In each case, a series of 8 spectra with transfer periods ranging from 10 to 100 msec were recorded with a 600 MHz spectrometer using a recycle delay of 1.5 s. Corresponding reference spectra were recorded with a recycle delay of 12 s to ensure complete water relaxation. Pseudo-first-order rate constants for exchange were obtained by nonlinear least-squares fitting of peak intensities versus transfer time, as described previously (291). Protection factors (PF) were determined as the ratio of the intrinsic versus measured exchange rate constants. The former were predicted for a random coil polypeptide with the Rv1747 sequence and under the

corresponding solvent, pH and temperature conditions using the program Sphere (292). In cases where two or more reliable PF values were available from complementary measurements, these were averaged. This assumes that exchange occurs in the commonly observed EX2 limit and that the stability of the Rv1747 constructs do not change substantially over the pH range considered (264).

2.5.10 NMR-monitored Titrations

Phosphorylated peptides corresponding to pT152 or pThr210 were used in ^{15}N -HSQC-monitored titrations of FHA-1 (Rv1747¹⁻¹⁵⁶) and FHA-2 (Rv1747²⁰⁶⁻³¹⁰), recorded at 25 °C with a 600 MHz NMR spectrometer. Both protein and peptide samples were exhaustively dialyzed into the identical buffer (100 mM NaCl, 20 mM sodium phosphate, pH 6.0) using 3.5 – 5 kDa and 0.1 – 0.5 kDa MWCO dialysis tubing, respectively and the pH confirmed after dialysis. Aliquots of unlabeled phosphopeptide (initially ~ 4.3 and 2.8 mM for pT152 and pT210, respectively) were progressively added to uniformly ^{15}N -labelled protein (initially 500 μL of ~ 300 μM). Amide ^1H - ^{15}N chemical shift assignments of the bound proteins were extended from those of the assigned unbound proteins by tracking well resolved peaks exhibiting fast-exchange behavior over the course of the titration. The ^1H and ^{15}N chemical shift differences of backbone amides with respect to the unbound state, $\Delta\delta_{\text{H},i}$ and $\Delta\delta_{\text{N},i}$, respectively, were monitored and the overall chemical shift perturbations (CSP) $\Delta\delta_i$ at each titration point i were calculated as $\Delta\delta_i = [(0.14\Delta\delta_{\text{N},i})^2 + (\Delta\delta_{\text{H},i})^2]^{1/2}$. The resulting titration curves for individual amides were non-linear least squares fit in GraphPad Prism to the equation for a 1:1 binding isotherm (293),

$$\Delta\delta_i = \Delta\delta_{\text{sat}} \frac{([P]_{T,i} + [L]_{T,i} + K_d) - \sqrt{([P]_{T,i} + [L]_{T,i} + K_d)^2 - 4[P]_{T,i}[L]_{T,i}}}{2[P]_{T,i}}$$

where $[P]_{T,i}$ and $[L]_{T,i}$ are the dilution-corrected concentrations of protein and peptide, respectively at each titration point i , K_d is the fit equilibrium dissociation constant, and $\Delta\delta_{\text{sat}}$ is the extrapolated chemical shift perturbation at saturation. The reported K_d values and the

associated standard deviations were determined by averaging the fit values of the six most perturbed amides for each of the four FHA/peptide combinations.

2.5.11 Crystallization of Rv1747³⁻¹¹⁶

Preliminary crystals were obtained at 18° C via the microbatch technique using 10 mg/mL protein and the Wizard I screen, condition 08 (2 M ammonium phosphate, 100 mM citrate buffer, pH 5.5) with 10% silicon, 90% paraffin overlay. Optimal crystals were obtained via the same microbatch approach with an initial protein concentration of 18 mg/mL in 1.8 M ammonium sulfate, 100 mM citrate buffer, pH 5.5. These were single, rod-shaped crystals of approximately 0.4 mm length that diffracted to < 2 Å resolution when soaked in n-paratone oil and flash frozen in liquid nitrogen. They indexed to a primitive trigonal space group, and contained one molecule in the asymmetric unit (Table 2.1). Heavy atom-bound crystals were obtained by a short incubation (30 – 60 seconds) in cryoprotectant containing 0.5 M NaBr, 1 M ammonium sulfate, 100 mM citrate pH 5.5, and 25 % ethylene glycol (294) followed by flash-freezing in liquid nitrogen.

2.5.12 X-ray Crystallography Data Collection, Processing, and Structure Determination of FHA-1 (Rv1747³⁻¹¹⁶)

X-ray data were collected on beamline 8.3.1 at the Advanced Light Source at Lawrence Berkeley National Laboratory. Data were integrated and scaled using the program HKL2000 (295). Seven bromide ions were identified in a crystal frozen in cryoprotectant containing 0.5 M NaBr via multi-wavelength anomalous diffraction (MAD) analysis with SHELXD (296) of data collected at wavelengths 0.9200 Å, 0.9202 Å, and 1.1159 Å. Further analysis with SOLVE/RESOLVE (297, 298) generated two bromide sites. These proved sufficient to calculate accurate phases and produce an interpretable electron density map. ARP/wARP (299) was used to build a partial C^α backbone into this density, and the remainder of the backbone trace was completed by hand. This preliminary model was improved through iterative cycles of restrained refinement with REFMAC5 (300) and manual rebuilding using the visualization program O (301) until the R_{cryst} and R_{free} were 0.21 and 0.25, respectively. The model was used to calculate phases for the highest resolution data set (1.8 Å) collected a native crystal without bromide.

Molecular replacement solutions were generated using Phaser (302), refined with PHENIX (303) and rebuilt in COOT (304). See Table 2.1 for data and refinement statistics.

2.5.13 Isothermal Titration Calorimetry

Isothermal titration calorimetry (ITC) studies were carried out at 25 °C using a MicroCal ITC200 (GE Healthcare Life Sciences). Protein and peptide samples were exhaustively dialyzed into the identical buffer (100 mM NaCl, 20 mM sodium phosphate, pH 6.0) using 3.5 – 5 kDa and 0.1 – 0.5 kDa MWCO dialysis tubing, respectively and the pH confirmed after dialysis. Phosphorylated peptides corresponding to pT152 or pThr210 (initially ~ 4.3 and 2.8 mM, respectively, ~ 30 uL added) were titrated in a stepwise manner into solutions of FHA-1 (Rv1747¹⁻¹⁵⁶) and FHA-2 (Rv1747²⁰⁶⁻³¹⁰), initially at ~ 0.5 and 0.2 mM, respectively. Dilution heats were determined by repeating the same titrations of peptide into buffer and subsequently subtracted from the data to obtain the final corrected binding heats. These titration data were then fit to a 1:1 binding model using Origin (GE Healthcare Life Sciences). Protein concentrations were adjusted to yield a stoichiometry value $n = 1$. This accounts for loss in binding activity or errors in the concentrations of the proteins or phosphopeptides determined using predicted ϵ_{280} values. Thermodynamic parameters are only reported for FHA-2 as the uncertainties associated with fitting the FHA-1 isotherms are large due to weak binding and low enthalpy changes.

2.5.14 Circular Dichroism Spectroscopy

Circular dichroism (CD) spectra were recorded for ~ 10 μ M FHA-1 (Rv1747¹⁻¹⁵⁶) and FHA-2 (Rv1747²⁰⁶⁻³¹⁰) at 25 °C using a JASCO J-810 spectropolarimeter. Data were measured between 210 nm to 280 nm using a 100 nm/min scan rate, 100 mdeg sensitivity and 0.1 s response time. The proteins were in buffer (100 mM NaCl, 20 mM sodium phosphate, pH 6.0) with 0 – 5 M guanidinium hydrochloride (GuHCl; SIGMA \geq 99%, recrystallized (305)). The GuHCl concentrations were determined by refractive index (306). Ellipticity values, θ , at 217 nm and 217.5 nm for FHA-1 and FHA-2, respectively, versus [GuHCl] were fit for a two-state unfolding transition using GraphPad Prism according to the following equations,

$$\theta = \theta_{f_0} + [GuHCl]s_f + f_u(\theta_{u_0} + [GuHCl]s_u - \theta_{f_0} - [GuHCl]s_f)$$

$$f_u = \frac{K_u}{1 + K_u}$$

$$K_u = e^{-\Delta G_u/RT}$$

$$\Delta G_u = \Delta G_{u, H_2O} - m[GuHCl]$$

where s_f and s_u denote the slopes of the θ dependence on [GuHCl] for the fully folded and unfolded proteins, respectively and θ_{f_0} and θ_{u_0} are their extrapolated y-intercepts, f_u is the fraction of unfolded protein, K_u the equilibrium constant for the unfolding transition and ΔG_u the corresponding difference in free energy between unfolded and folded protein at a certain [GuHCl]. $\Delta G_{u, H_2O}$ is the free energy difference between unfolded and folded protein at [GuHCl] = 0 M and m is the slope of the assumed linear ΔG_u vs [GuHCl] dependence (306).

2.5.15 Molecular Dynamics Simulations

Molecular dynamics (MD) simulations were carried using X-ray crystallographic structure of FHA-1 and the lowest energy member of the NMR-derived structural ensemble of FHA-2 as starting points. Both were truncated to the core residues for FHA-1 (8 – 105) and FHA-2 (225 – 310). PROPKA3.1 (307) was used to predict protonation states at pH 7. Both proteins were solvated in a cuboid explicit TIP3P water box and 4 Na⁺ ions were added to FHA-2 to neutralize the system. The final number of atoms were 17189 and 18950 for FHA-1 and FHA-2, respectively, and the box dimensions were 59.0 Å · 62.0 Å · 59.6 Å and 65.3 Å · 58.5 Å · 63.1 Å. After 5000 steps of steepest descent energy minimization of the solvent with the protein coordinates fixed, and an additional 10000 steps for all atoms including the protein, the systems were heated to 300 K over 50 ps and 1 ns of equilibration was performed. The production run of a total of 900 ns of Langevin dynamics was performed with an integration step size of 2 fs in the modified AMBER ff14SB all-atom force field using the PREMD module in AMBER14 (308). The isobaric isothermal ensemble was used at 300 K and a pressure of 1 atm with periodic boundary conditions and the long range electrostatic interactions were accounted for using the

particle-mesh Ewald sum. A cutoff of 10 Å was used for long-range non-bonded interactions. The SHAKE algorithm was used to constrain bonds involving hydrogens. Backbone (N, C^α, CO) RMSD time courses were calculated from the trajectories aligned to the starting crystal and NMR structures as well as the solvated, equilibrated structures using CPPTRAJ, a module of AMBER14. This module was further used to calculate the per-residue backbone (N, C^α, CO) AMBER B-factors and they were mapped onto the structures and visualized using PyMol.

2.5.16 Structural Alignment

Structures of previously reported FHA domains (*S. cerevisiae* Rad51 FHA-1, PDB ID 1G6G.pdb; *S. cerevisiae* Rad51 FHA-2, 1K2N.pdb; *S. cerevisiae* Dun1, 2JQJ.pdb; *H. sapiens* Chk2, 1GXC.pdb; *H. sapiens* Ki67, 1R21.pdb; *H. sapiens* Rnf8, 2PIE.pdb; *M. musculus* Pnk, 1YJM.pdb; *M. tuberculosis* Rv0020c, 2LC1.pdb; *M. tuberculosis* GarA, 2KFU.pdb), as well as the lowest energy structure of the NMR determined ensemble of Rv1747 FHA-2, were aligned to Rv1747 FHA-1 using CE-CP, an incremental combinatorial extension approach that was optimized for comparison of circularly permuted protein structures (309).

2.5.17 Data and Model Depositions

The coordinates and X-ray crystallographic structure factors for Rv1747³⁻¹¹⁶ (FHA-1) have been deposited in the Protein Data Bank (<http://www.rcsb.org>) under ID code 6CCD.pdb. The coordinates and NMR-derived restraints for the Rv1747²⁰⁶⁻³¹⁰ (FHA-2) ensemble have been deposited as 6CAH.pdb. The chemical shift assignments for Rv1747¹⁻¹⁵⁶ (FHA-1) and Rv1747²⁰⁶⁻³¹⁰ (FHA-2) have been deposited in the Biological Magnetic Resonance Bank (<http://www.bmrb.wisc.edu/>) as 27394 and 30399, respectively.

Chapter 3: Phosphorylation-enhanced Protein Phase Separation of the Regulatory Module of a *Mycobacterium tuberculosis* ABC transporter

3.1 Overview

Protein phase separation has been recently shown to be a fundamental mechanism underlying the clustering of some proteins at eukaryotic cell membranes. The virulence-linked ABC transporter Rv1747 from *Mycobacterium tuberculosis* (*Mtb*) possesses a cytoplasmic regulatory module containing two phosphothreonine-binding FHA (Forkhead associated) domains connected by an intrinsically disordered (ID) linker. Upon multi-site phosphorylation of the ID linker by several *Mtb* serine/threonine protein kinases including PknF, the isolated regulatory module readily phase separates into dynamic liquid droplets with diagnostic properties similar to those exhibited by eukaryotic proteins. The process is reversed by the sole *Mtb* serine/threonine phosphatase PstP. In the absence of phosphorylation, the Rv1747 regulatory module can still undergo phase separation, albeit at higher protein concentrations and with more dynamic properties of the resulting droplets. This points to a synergy between specific FHA-pThr binding and additional weak association of the ID linker and/or the FHA domains leading to the pre-requisite multivalent interactions for phase separation. Droplet formation of the regulatory module was replicated in a pseudo-two-dimensional system on a model supported lipid bilayers. Potential clusters of Rv1747 were also detected in *Mtb* using ultra-high resolution Direct Stochastic Optical Reconstruction Microscopy (dSTORM). This is the first reported example of phase separation by both a bacterial protein and an ABC transporter, and suggests possible mechanisms for the regulation of Rv1747.

3.2 Introduction

Recent work has demonstrated that many proteins can phase separate into liquid droplets (condensates) *in vitro* and *in vivo* (310, 311). Weak multivalent inter-molecular interactions between proteins and/or nucleic acids drive this phase separation and define the dynamic behavior of the resulting droplets. Consistent with this view, proteins that undergo phase separation often contain multiple ligand binding domains (152, 201) and/or intrinsically disordered segments with repetitive short motifs that mediate weak inter-molecular contacts. The

latter include charge-charge, cation- π , or π - π interactions (158, 170, 205, 206). Droplet formation can be regulated through changes in ionic strength, pH, temperature, post-translational modifications and protein/nucleic acid concentration.

In eukaryotic cells, phase separation can give rise to membraneless organelles such as nucleoli (184) and RNA granules (e.g. stress granules or P-bodies (158)). These membraneless organelles can act as sites of molecular storage, as well as active biomolecular filters and reactors (194). Phase separation has also been shown to be the driving force behind clustering of downstream targets of the T-cell receptor pathway in the membrane of immune cells. In this case, the soluble region of an otherwise membrane-anchored protein undergoes liquid-liquid demixing and thereby promotes clustering of a signaling complex at the membrane (153, 193). In contrast to the growing recognition for the importance of protein phase separation in eukaryotic cells, there is a lack of evidence for this phenomenon in prokaryotes.

One bacterial protein capable of modular, multivalent interactions is the ABC transporter Rv1747 from *Mtb*. The protein has a core topology common to single chain homodimeric ABC transporters with a cytoplasmic nucleotide binding domain (NBD), the location of ATP hydrolysis, followed by a helical transmembrane (TM) portion through which substrates move (Figure 2.1). Appended to this core, Rv1747 contains a putative N-terminal regulatory module composed of two Forkhead associated (FHA) domains connected by an \sim 100 amino acid intrinsically disordered (ID) linker that was structurally investigated in Chapter 2. FHA domains interact primarily with phosphothreonine peptides, thereby mediating phospho-dependent protein-protein interactions. Rv1747 has been shown to be phosphorylated within the ID linker by the *Mtb* serine/threonine protein kinase (STPK) PknF, which is expressed in the same operon as the transporter, and to a lesser extent, PknE (143-145). The sole *Mtb* serine/threonine protein phosphatase acting on Rv1747 is PstP (125). Although the exact functions of Rv1747 have not yet been elucidated, it is implicated in the export of cell wall biosynthesis intermediates (104, 105). Importantly, Rv1747 is a *Mtb* virulence factor and its deletion results in significantly attenuated growth of the bacterium in macrophages and mice (144). The presence of the phospho-acceptor sites in the ID linker, as well as the phospho-peptide binding capability of one of the FHA domains, are vital for the Rv1747-dependent virulence of *Mtb*.

In Chapter 2, I demonstrated that *in vitro* the isolated regulatory module of Rv1747 is capable of a variety of intra- and inter-molecular FHA-pThr interactions. This leads to the formation of higher-order multimers that can be tuned by the phosphorylation state of the protein. Here, I show that this multimerization results in phosphorylation-enhanced phase separation of the regulatory module, both free in solution and when tethered to a model membrane. These observations prompt potential models for the regulation of Rv1747 involving membrane clustering driven by phase separation of its regulatory module. Preliminary support of this model is obtained through ultra-high resolution microscopy of the intact transporter in *Mtb* cells.

3.3 Results

3.3.1 Rv1747 Regulatory Module Undergoes Phase Separation upon Phosphorylation

In chapter 2, I showed that the isolated regulatory module (denoted as Rv1747¹⁻³¹⁰) self-associates upon phosphorylation by PknF. This was attributed to a myriad of possible intra- and intermolecular binding equilibria between the two FHA domains and phosphothreonine residues within the ID linker. To further investigate the consequences of this association, the protein samples were visualized by differential interference contrast (DIC) microscopy. Strikingly, Rv1747¹⁻³¹⁰ (50 μ M) existed as a clear solution when unmodified, yet spontaneously formed spherical droplets upon phosphorylation by PknF (Figure 3.1). We confirmed that Rv1747¹⁻³¹⁰ was the main constituent of these droplets through fluorescence microscopy of the protein selectively labeled with Oregon Green (OG) at an introduced N-terminal cysteine (Figure 3.1). Control samples without ATP or Mg²⁺ or performed on the phospho-ablative mutant Rv1747¹⁻³¹⁰ T152A/T210A failed to show the same phase separation upon addition of the kinase under these conditions (Figure 3.2). Thus PknF phosphorylation of the two previously reported modification sites (T152/T210) promotes of Rv1747¹⁻³¹⁰ droplet formation.

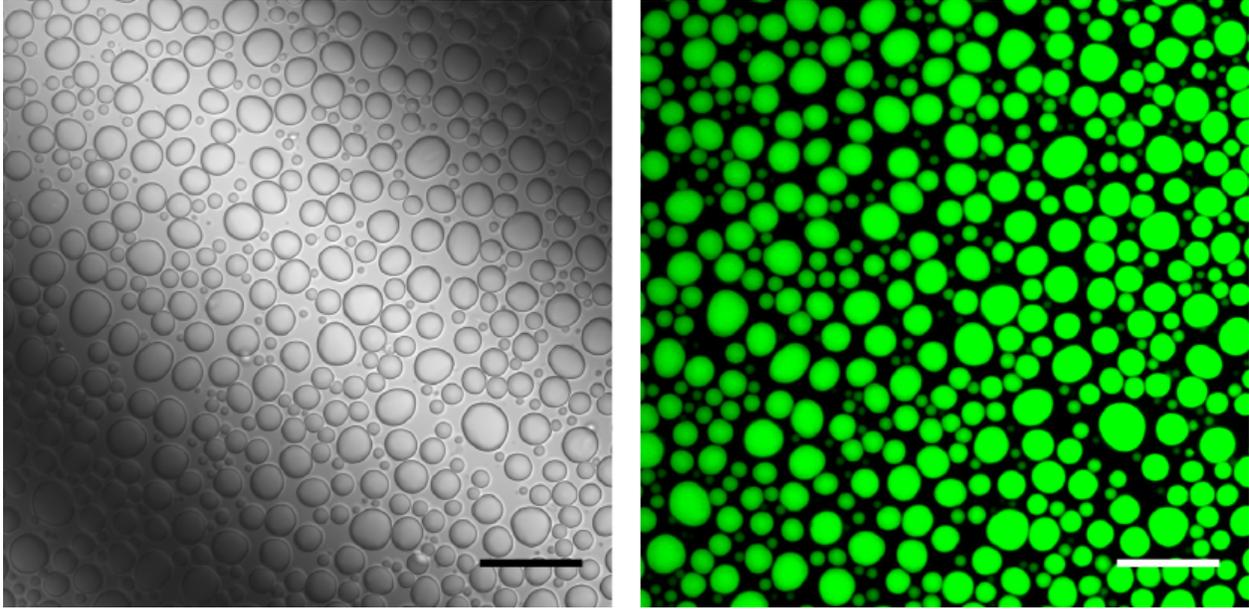


Figure 3.1: Rv1747¹⁻³¹⁰ Undergoes Phase Separation *in vitro* upon Phosphorylation.

Liquid droplets observed by DIC (left) and fluorescence microscopy (right) formed by 50 μM OG-labeled Rv1747¹⁻³¹⁰ upon phosphorylation by 5 μM PknF in buffers containing 5 mM ATP (scale bars: 20 μm).

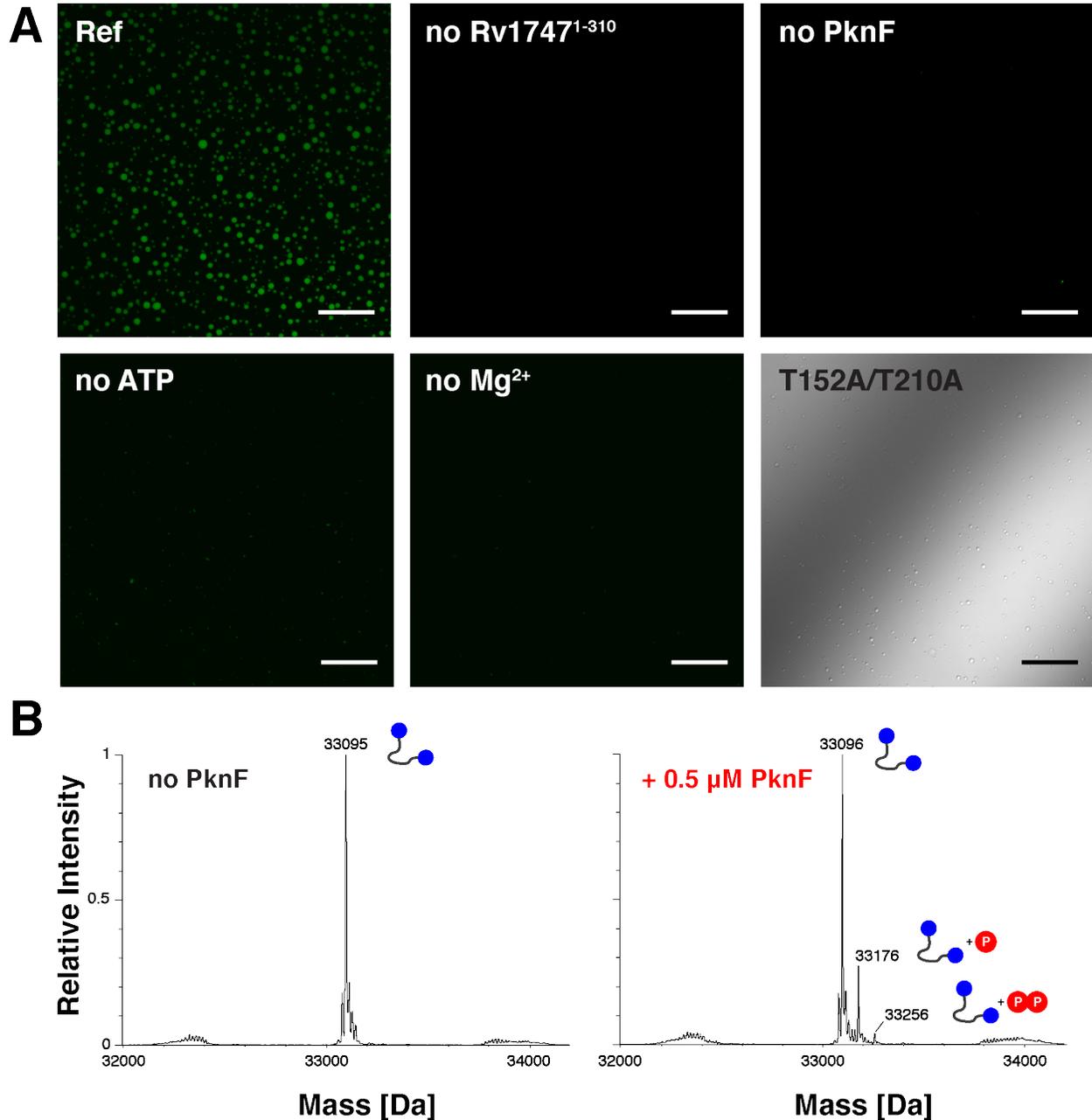


Figure 3.2: Sub-stoichiometric Phosphorylation of Rv1747¹⁻³¹⁰ by PknF is Sufficient to Induce Phase Separation.

(A) Reference fluorescent image of 50 μM OG-labeled Rv1747¹⁻³¹⁰ droplets, formed in buffer containing 5 mM Mg^{2+} , 100 μM ATP, and 0.5 μM PknF. Also shown are fluorescent images of controls where the indicated component was omitted. The DIC image (lower right) corresponds to unlabeled Rv1747¹⁻³¹⁰ T152A/T210A subjected to the same conditions. Droplets did not form to the same degree due to mutation of the two phospho-acceptor threonines. (scale bars: 20 μm)

(B) ESI MS spectra of Rv1747¹⁻³¹⁰ before phosphorylation (left) and after 120 min of treatment with PknF (right) under the conditions used in (A). Masses of major peaks are indicated and the

corresponding phosphorylation states are shown as cartoons with phosphates represented by red circles. Under these conditions, the PknF-treated protein was predominantly unmodified, with ~ 25 % single phosphorylated and a small amount of double phosphorylated forms also present.

3.3.2 Rv1747¹⁻³¹⁰ Droplets Exhibit the Diagnostic Properties of a Protein-rich Liquid Phase

The phosphorylated Rv1747¹⁻³¹⁰ droplets have the typical liquid-like, dynamic properties exhibited by other biological macromolecular condensates (152, 158). In particular, the droplets are spherical with sharp boundaries, they form rapidly upon kinase addition, they grow in size over time, they have dynamic boundaries showing wetting and dripping on surfaces (not shown), and they coalesce into larger droplets on a timescale of seconds (Figure 3.3). Also, diffusive exchange inside the droplets can be seen through Fluorescence Recovery After Photobleaching (FRAP) experiments (Figure 3.4, "phosphorylated" dataset). In a complementary approach, we see diffusion of OG-labeled Rv1747¹⁻³¹⁰ into droplets previously formed by non-fluorescent Rv1747¹⁻³¹⁰ (Figure 3.5). Thus molecules associated in the protein-rich droplets readily exchange with those in the surrounding protein-depleted phase.

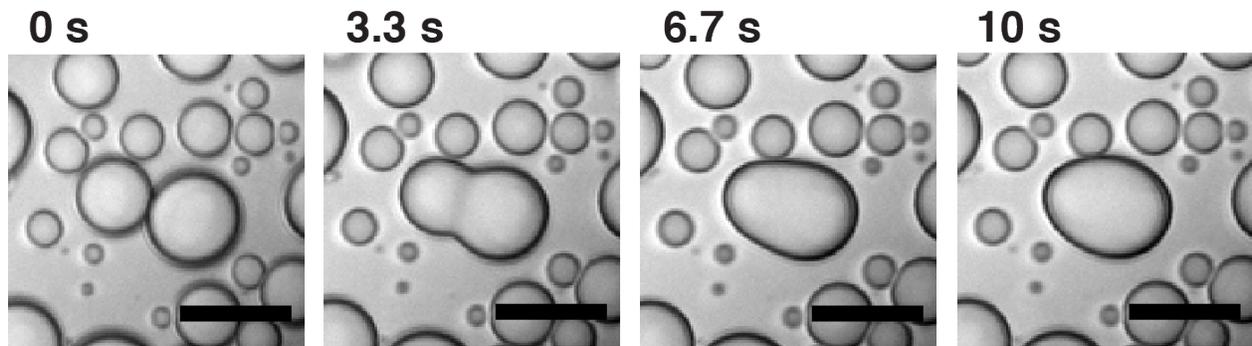


Figure 3.3: Coalescence of Liquid Rv1747¹⁻³¹⁰ Droplets.

DIC images showing the time course for coalescence of two Rv1747¹⁻³¹⁰ droplets. The droplets were formed by addition of 0.5 μM PknF to 50 μM Rv1747¹⁻³¹⁰ in buffers containing 100 μM ATP (scale bars: 5 μm).

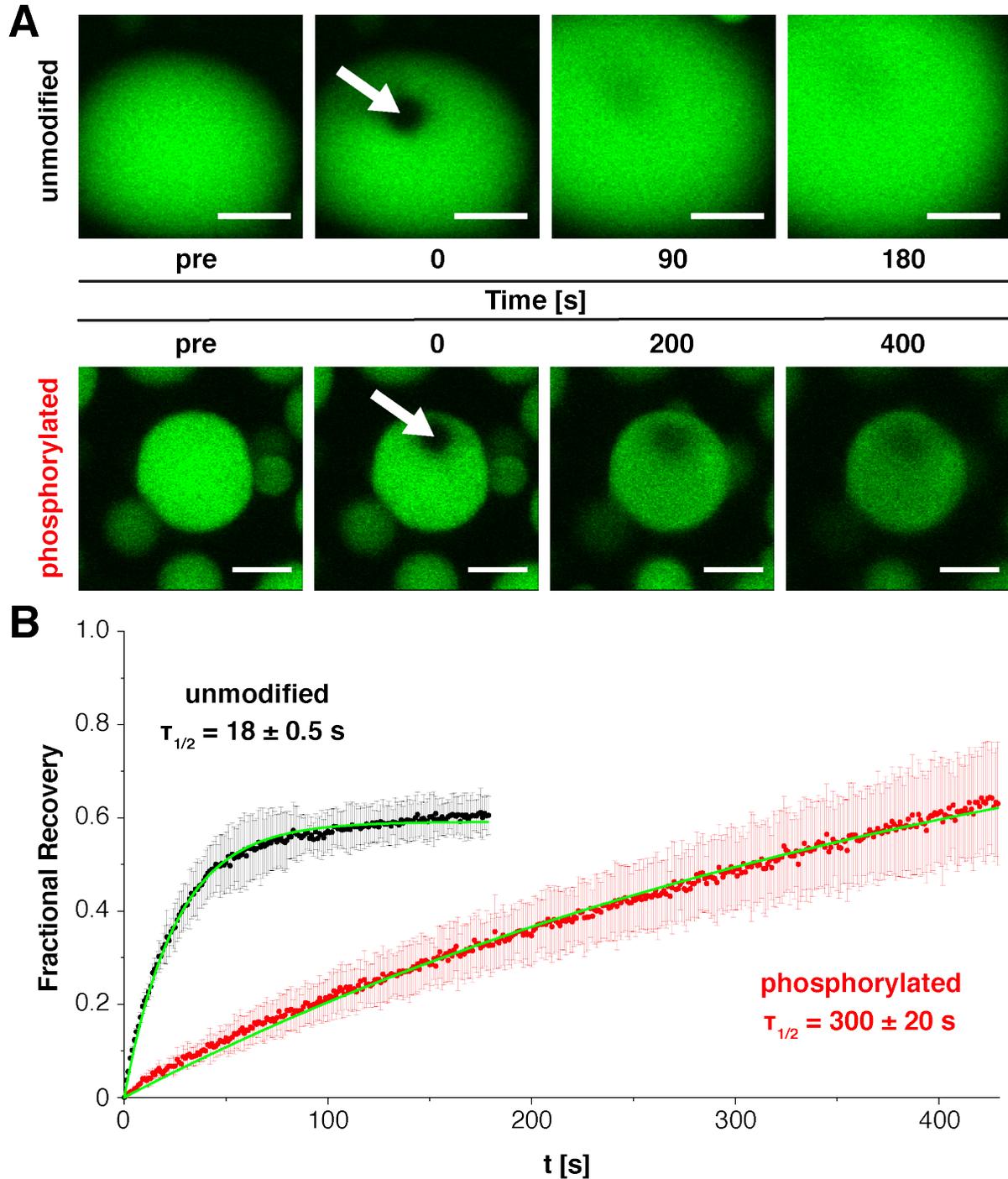


Figure 3.4: Diffusive Exchange of Rv1747¹⁻³¹⁰ Droplets measured by Fluorescence Recovery after Photobleaching.

(A) Diffusive exchange at 22 °C in condensates formed by unmodified Rv1747¹⁻³¹⁰ at > 250 μM (top) or by phosphorylated Rv1747¹⁻³¹⁰ (50 μM treated with 0.5 μM PknF and 100 μM ATP; bottom). The bleached droplet sector is indicated by an arrow in the $t = 0 \text{ s}$ time point image,

followed by representative images taken at two subsequent time points (scale bars: 5 μm). Growth by Ostwald ripening during the course of the FRAP experiment of the unmodified droplet can be seen. (B) Recovery was quantified by fitting the average normalized intensity (solid dots with bars showing standard deviations) of three different FRAP measurements on droplets of unmodified and phosphorylated Rv1747¹⁻³¹⁰ to a single exponential recovery function (green lines). The times to 50 % recovery ($\tau_{1/2}$) are shown.

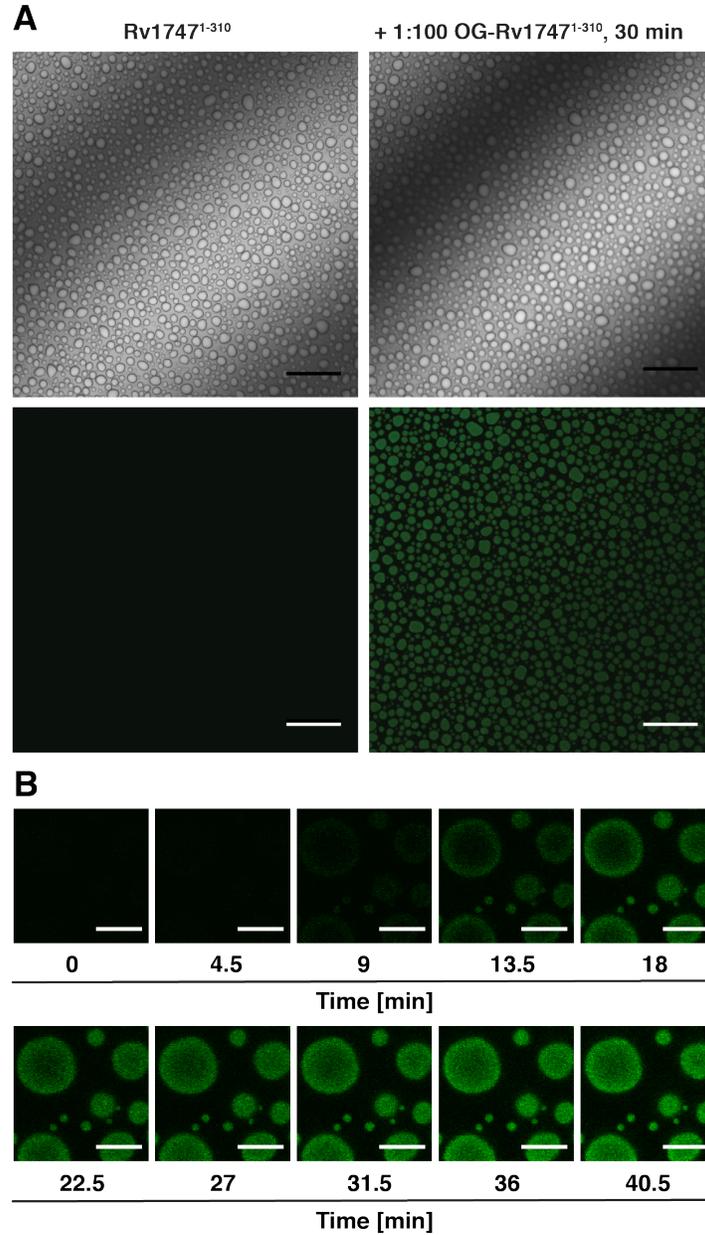


Figure 3.5: Diffusive Exchange Occurs Between the Protein-rich and Protein-depleted Phases of an Rv1747¹⁻³¹⁰ Sample.

(A) DIC (top) and fluorescent images (bottom) of non-fluorescent, phosphorylated Rv1747¹⁻³¹⁰ droplets formed at 50 μ M by addition of 0.5 μ M PknF (with 100 μ M ATP) are shown on the left. Similar images were recorded 30 min after the addition of 0.5 μ M OG-labeled, unmodified Rv1747¹⁻³¹⁰ (right). Incorporation of the fluorescently labeled protein into the previously non-fluorescent droplets can be seen. Under these conditions, the initially unmodified OG-Rv1747¹⁻³¹⁰ is too dilute to form new droplets by itself (scale bars: 20 μ m). (B) Representative fluorescent images taken at different time points after addition of OG-labeled Rv1747¹⁻³¹⁰ at 22 $^{\circ}$ C as in (A) (scale bars: 5 μ m).

3.3.3 Rv1747¹⁻³¹⁰ Droplet Formation is Cooperative

The concentration dependence of Rv1747¹⁻³¹⁰ droplet formation was investigated with fluorescence microscopy. A standardized protocol was followed, in which images were taken at 22 °C and 120 min after the addition of PknF to different concentrations of Rv1747¹⁻³¹⁰. For details, see Materials and Methods. The resulting data were analyzed using a simple criteria of measuring the fractional intensity of fluorescence arising from pixels classified as part of droplets versus that of the entire recorded image. As shown in Figure 3.6, the PknF-treated protein cooperatively formed droplets at a low threshold concentration of ~ 10 μM. We define the threshold as the concentration where the fractional intensity first reaches 0.5. This is the average concentration in the sample as it is measured by UV-absorbance spectroscopy of a well resuspended aliquot. The concentrations inside the droplets and in the surrounding, protein-depleted phase, ϕ_h and ϕ_l , respectively (Figure 1.10), were not determined.

3.3.4 At Higher Concentrations, Unmodified Rv1747¹⁻³¹⁰ also Forms Droplets

Measurements with ESI-MS on samples of well-resuspended droplets of 50 μM samples of Rv1747¹⁻³¹⁰ revealed that an average level of sub-stoichiometric phosphorylation was sufficient to induce phase separation (Figure 3.2). Although the phosphorylation states of protein within the droplets versus in solution were not determined, this result was somewhat unexpected if classical FHA-pThr association is assumed to be the only form of multivalent interaction leading to phase separation. Further investigation showed that completely unmodified Rv1747¹⁻³¹⁰ also formed droplets, albeit at significantly higher threshold concentrations (~ 250 μM) than the PknF-treated protein (Figure 3.6). This points to additional weak intermolecular interactions of Rv1747¹⁻³¹⁰ complementing FHA-pThr binding. Diffusive exchange as measured in FRAP experiments was found to be more rapid in droplets formed from unmodified ($\tau_{1/2} = 18 \pm 0.5$ s, $I_{\text{mobile}} = 0.60 \pm 0.01$) than from phosphorylated Rv1747¹⁻³¹⁰ ($\tau_{1/2} = 300 \pm 20$ s, $I_{\text{mobile}} = 0.99 \pm 0.04$) (Figure 3.4). Importantly, while showing much slower recovery, the mobile fraction I_{mobile} of phosphorylated droplets is higher than in the unmodified droplets. This could be due to the much lower average protein concentration that is required for phosphorylated phase separation compared to the unmodified case. In unmodified samples, some background aggregation caused by a high overall protein concentration might cause a reduced I_{mobile} . The shorter fluorescence

recovery times of unmodified droplets suggests that the dominant driving force for phase separation involve non-specific interactions that are weaker than specific FHA-pThr binding. We speculate that these non-specific interactions arise between residues within the ID linker.

Due to the rapid and cooperative formation of droplets by Rv1747¹⁻³¹⁰ upon treatment with PknF, we were unable to define a threshold concentration for phase separation by the protein with stoichiometric phosphorylation at T152 and T210. Furthermore, in the presence of 10-fold higher concentrations of PknF (5 μ M versus 0.5 μ M) and excess ATP, Rv1747¹⁻³¹⁰ was phosphorylated at additional, unmapped sites to yield a mixture with 1 to at least 3 modifications (Figure 3.7B). Under these conditions, PknF-treated Rv1747¹⁻³¹⁰ T152A/T210A also phase separated at protein concentrations below the threshold for the unmodified species (Figure 3.7A). This indicates that the FHA domains can interact with phosphothreonines at sites other than 152 and 210 to favor droplet formation.

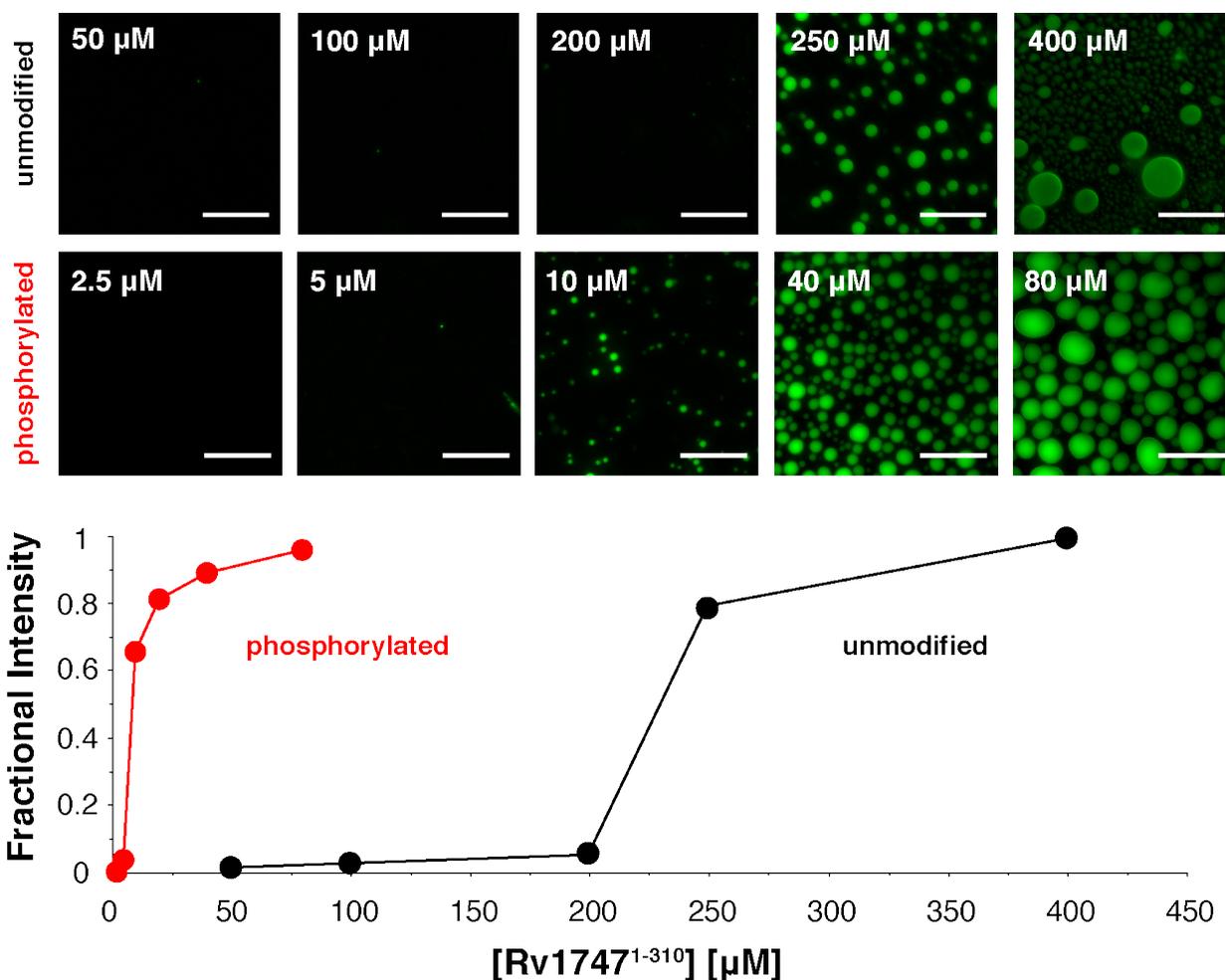


Figure 3.6: Rv1747¹⁻³¹⁰ Droplet Formation is Cooperative and Enhanced by Phosphorylation.

Unmodified and phosphorylated Rv1747¹⁻³¹⁰ phase separate at different critical threshold concentrations. Representative images of samples of unmodified (top) and phosphorylated (middle) OG-labeled Rv1747¹⁻³¹⁰ (scale bars: 40 μm). Images were taken at 22 °C and 120 min after the addition of PknF or the removal from the concentrator for phosphorylated and unmodified samples, respectively. For details, see Materials and Methods. The degree of phase separation is quantified by fractional fluorescence intensity (bottom).

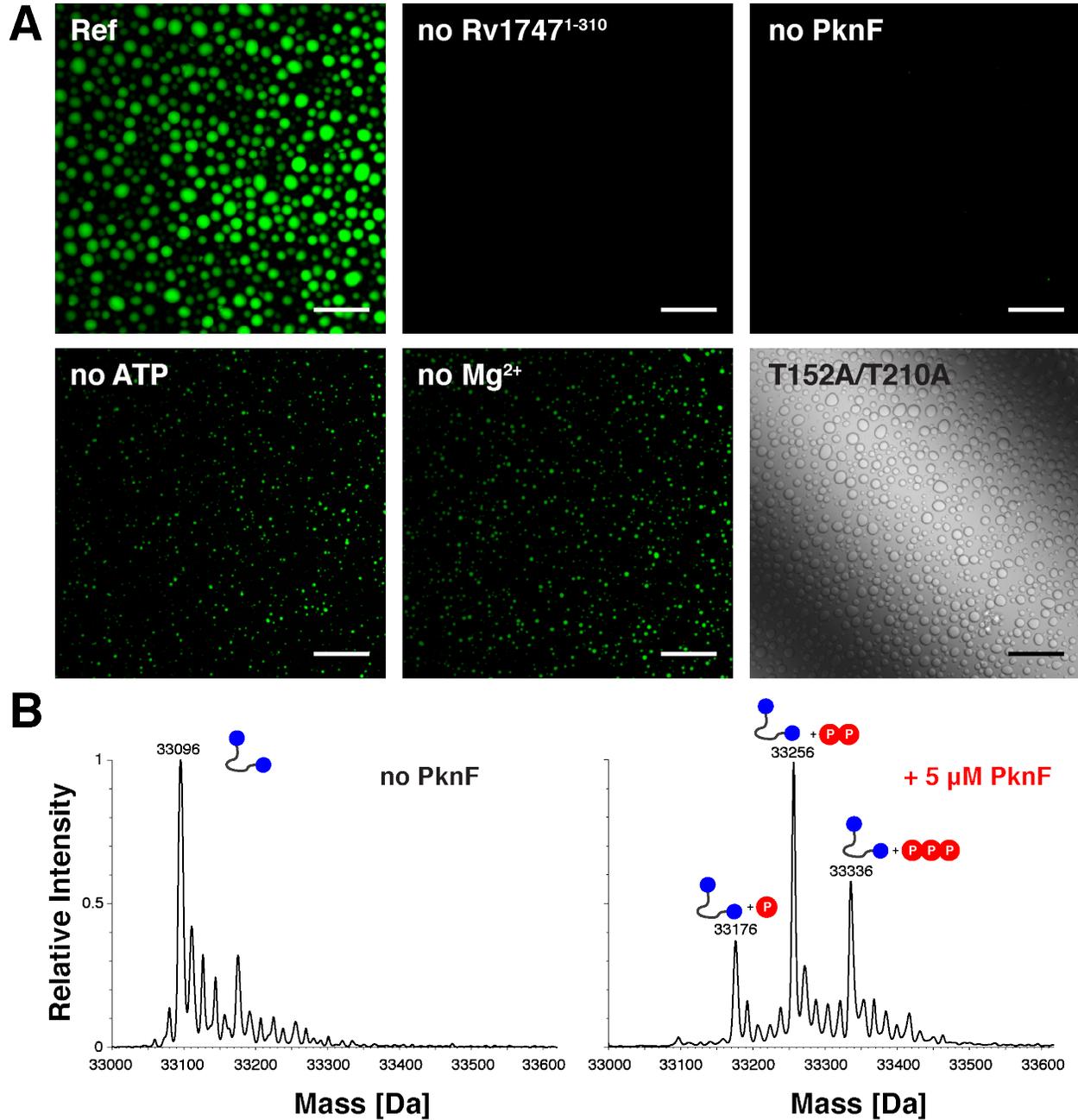


Figure 3.7: Higher Concentrations of PknF Phosphorylate Rv1747¹⁻³¹⁰ at Additional Sites that also Contribute to Phase Separation.

(A) Reference fluorescent image of 50 μM sample of OG-labeled Rv1747¹⁻³¹⁰ droplets, formed in buffer containing 5 mM Mg²⁺ and 5 mM ATP treated with 5 μM PknF. Also shown are fluorescent images of controls where the indicated component was omitted. The small droplets formed in the absence of added ATP or Mg²⁺ are attributed to sub-stoichiometric phosphorylation of Rv1747¹⁻³¹⁰ by the high concentration of added active kinase, heterologously expressed in the presence of Mg²⁺ and ATP in *E. coli*. The DIC image is of unlabeled Rv1747¹⁻

³¹⁰ T152A/T210A subjected to the same conditions (scale bars: 20 μm). Although lacking T152 and T210, droplets still formed due to phosphorylation at additional unmapped sites. Note that 50 μM is below the threshold concentration of $\sim 250 \mu\text{M}$ for phase separation by unmodified Rv1747¹⁻³¹⁰. (B) ESI-MS of Rv1747¹⁻³¹⁰ before phosphorylation and after 120 min of treatment with PknF under the conditions used in (A). Masses of major peaks are indicated and the corresponding phosphorylation states are shown as a cartoons with phosphates represented by red circles. The PknF-treated protein is a mixture of at least single, double and triple phosphorylated forms. This shows that phosphorylation can occur at sites other than T152 and T210.

3.3.5 Droplet Formation is Reversed upon Dephosphorylation by PstP

Treatment of PknF-induced Rv1747¹⁻³¹⁰ droplets with the phosphatase PstP resulted in their disappearance (Figure 3.8). Thus phase separation is reversible. Under the standard sub-stoichiometric phosphorylation conditions used above, droplet dissolution due to dephosphorylation was significantly slower than the induction of phase separation by phosphorylation. This likely reflects the fact that sub-stoichiometric levels of phosphorylation leads to demixing, and hence extensive dephosphorylation is required to return Rv1747¹⁻³¹⁰ to a homogenous state in solution, particularly when occurring against a background of continuing kinase activity. More interestingly, this observation also leads to the question of whether or not PknF and PstP were acting on protein in the solution and/or in the droplet phase.

3.3.6 PknF and PstP Co-localize to Rv1747¹⁻³¹⁰ Droplets

Using fluorescence microscopy, AF647-labeled constructs of both PstP and the catalytic domain of PknF were found to co-localize to droplets of phosphorylated OG-labeled Rv1747¹⁻³¹⁰ (Figure 3.9A). Control experiments confirmed that neither enzyme phase separated on its own (not shown), and thus each can join a heterogeneous protein-rich phase based on a Rv1747¹⁻³¹⁰ scaffold. Within this dynamic liquid-like phase, the enzymes are likely functional, phosphorylating or dephosphorylating the regulatory module to re-enforce or weaken droplet formation. This co-localization could certainly result from even weak active site binding of PknF and PstP to the highly concentrated Rv1747¹⁻³¹⁰. However, STPKs such as PknF are phosphorylated at multiple threonines (94). These could be also recognized by the FHA domains of Rv1747¹⁻³¹⁰ and serve as docking sites between kinase and FHA-containing substrate. It is also noteworthy that PknF appeared to be homogeneously distributed within the Rv1747¹⁻³¹⁰ droplets. In contrast, PstP tended to form clusters at the interface between the main Rv1747¹⁻³¹⁰-rich phase and the protein-depleted surrounding buffer (Figure 3.9B). Functional multiphase immiscibility has been described before (184) and potentially limits the access for PstP to pRv1747¹⁻³¹⁰. This could also explain the slow droplet dissolution by PstP.

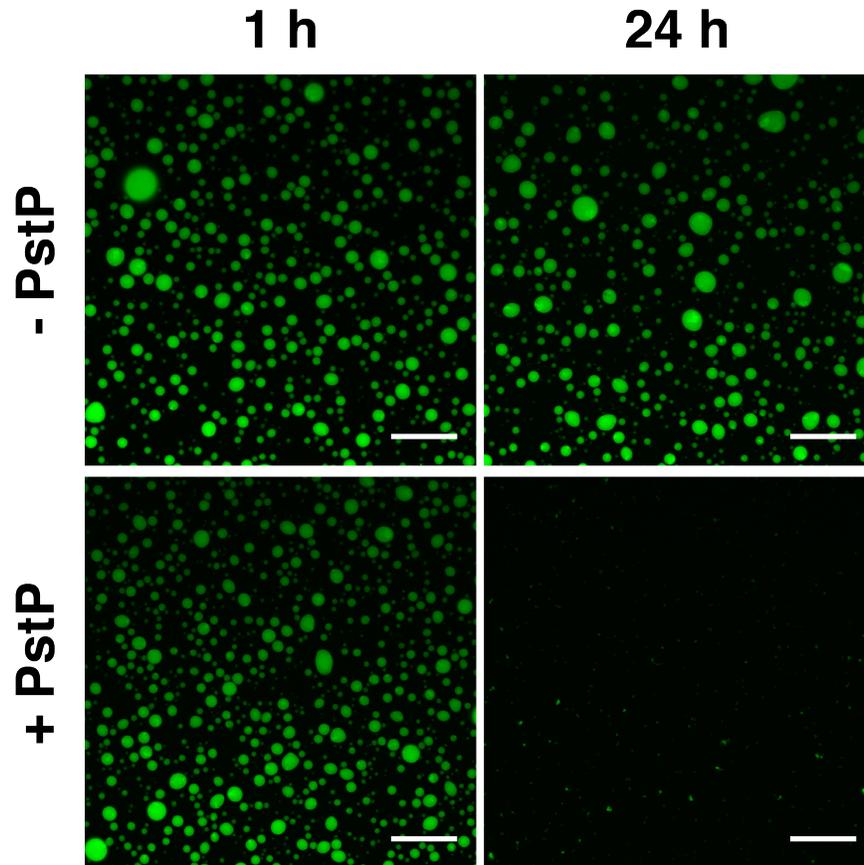


Figure 3.8: Treatment with PstP Dissolves PknF-induced Rv1747¹⁻³¹⁰ Droplets.

Condensates of 50 μM OG-Rv1747¹⁻³¹⁰ were formed in buffer containing 100 μM ATP, 5 mM Mg^{2+} and 5 mM Mn^{2+} and 0.5 μM PknF (22 $^{\circ}\text{C}$). After 120 min, PstP (5 μM) was added to one of two samples (bottom) and fluorescent images of representative time points after the addition of PstP are shown (scale bars: 20 μm). The concentration of Rv1747¹⁻³¹⁰ is below the threshold for phosphorylation-independent phase separation.

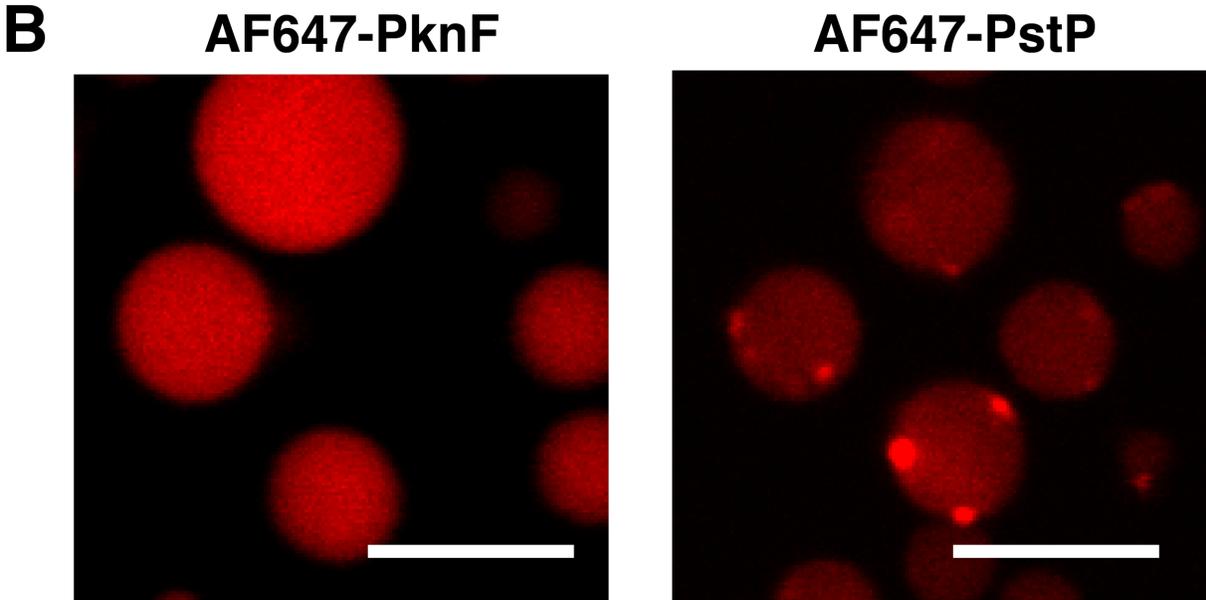
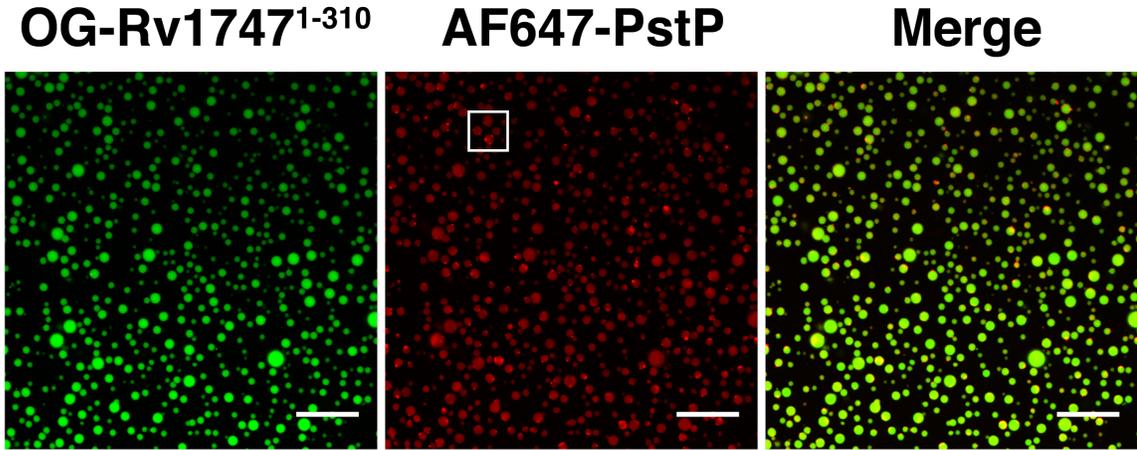
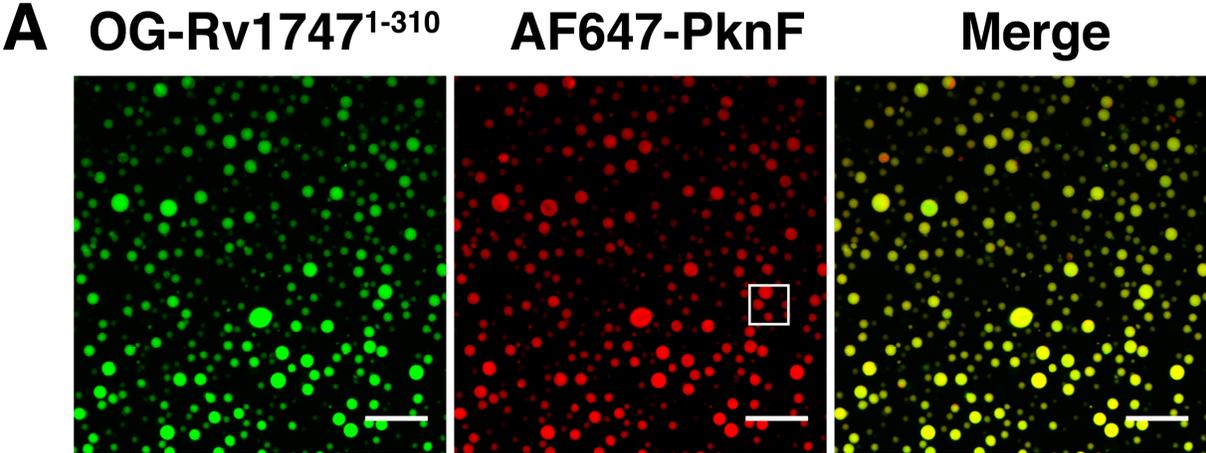


Figure 3.9: PknF and PstP Co-localize to Rv1747¹⁻³¹⁰ Droplets.

Both PknF and PstP co-localize to pRv1747¹⁻³¹⁰ droplets. (A) After formation with non-labeled PknF in buffer containing 100 μM ATP, 5 mM Mg^{2+} and 5 mM Mn^{2+} and 0.5 μM PknF, 0.05 μM AF647-labeled PknF (top) and AF647-labeled PstP (bottom) were added. Images of OG (green) and AF647 (red) fluorescence were taken and overlaid to monitor co-localization (scale bars: 20 μm). (B) Whereas PknF appears uniformly distributed in droplets, PstP shows heterogeneous distribution with foci of higher concentration predominately on the surface of droplets. Zoomed-in views of highlighted areas of images in (A) are shown (scale bars: 5 μm).

3.3.7 Several *Mtb* STPKs Phosphorylate Rv1747¹⁻³¹⁰ and Induce Phase Separation

PknF resides in the same operon as Rv1747 and therefore this "cognate" kinase was used for most *in vitro* studies of phase separation. However, the purified catalytic domains of at least eight out of nine additional *Mtb* STPKs were found to phosphorylate Rv1747¹⁻³¹⁰ *in vitro* (Figure 3.10A). This occurred with varying levels of efficiency and specificity for the mapped PknF phospho-acceptors T152 and T210. Furthermore, most of these STPKs were able to enhance phase separation of Rv1747¹⁻³¹⁰ (Figure 3.10B). This is consistent with both the promiscuity and overlapping specificity of the STPKs for their substrates (87), and the observation that even sub-stoichiometric phosphorylation of Rv1747¹⁻³¹⁰, as well as the T152A/T210A mutant, by PknF can induce droplet formation.

Stated in more detail, like PknF, the catalytic domains of PknA, PknB, PknG, PknK and PknL showed some specificity for T152/T210 yet also phosphorylated additional sites in the T152A/T210A mutant (Figure 3.10A). In contrast, PknD and PknE effectively phosphorylated sites other than these two linker threonines and showed little specificity towards T152/T210. It is important to note that only broad qualitative trends can be inferred from these data because the activity of the STPKs was not quantified in more detail. Regardless of phospho-acceptor residues, all of these STPKs were able to induce phase separation. One interesting exception is PknH, which phosphorylated Rv1747¹⁻³¹⁰ with some specificity for T152/T210 but did not enhance phase separation significantly. This could be due to a lower phosphorylation activity than reflected in the radioactive assay or due to PknH phosphorylation of sites not contributing to phase separation. The sample of PknJ used for these studies showed barely detectable phosphorylation of Rv1747¹⁻³¹⁰ and did not induce phase separation. However, its activity was not tested using known control substrates.

A closer inspection of the fluorescent images in Figure 3.10B reveals that the resulting droplets showed a range of different droplet morphologies. For example, phosphorylation of Rv1747¹⁻³¹⁰ by PknB and PknD resulted in phase separation to small spherical droplets, whereas those induced by PknE and PknL had non-spherical shapes. The latter may reflect a less dynamic liquid-like state than occurring upon PknF-induced phase separation of Rv1747¹⁻³¹⁰. That is, the differing morphologies and therefore potentially different coalescence timescales may be due to

different sets of pThr sites being available for FHA-pThr interactions. This could affect the dynamic behavior and material properties of the resulting Rv1747¹⁻³¹⁰ condensates.

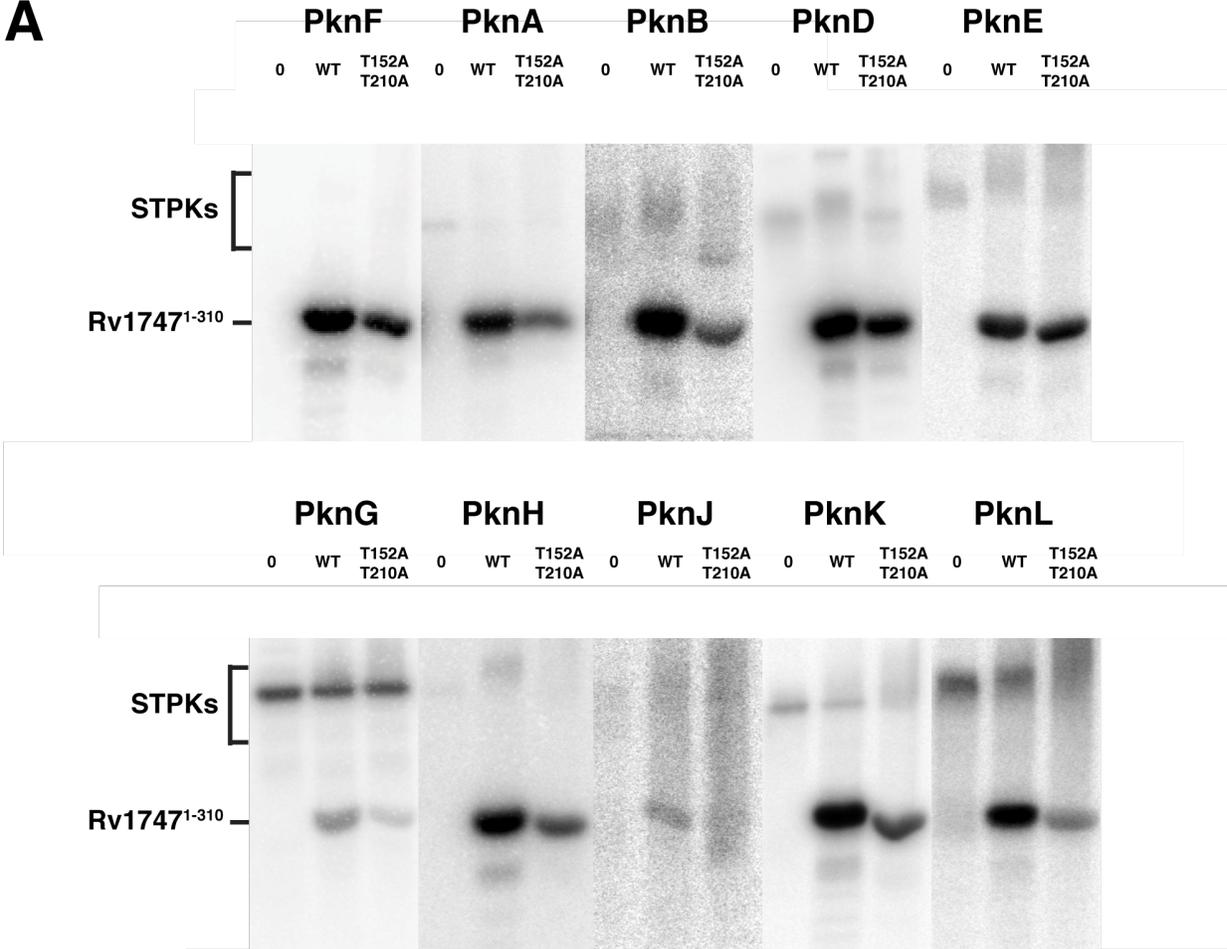
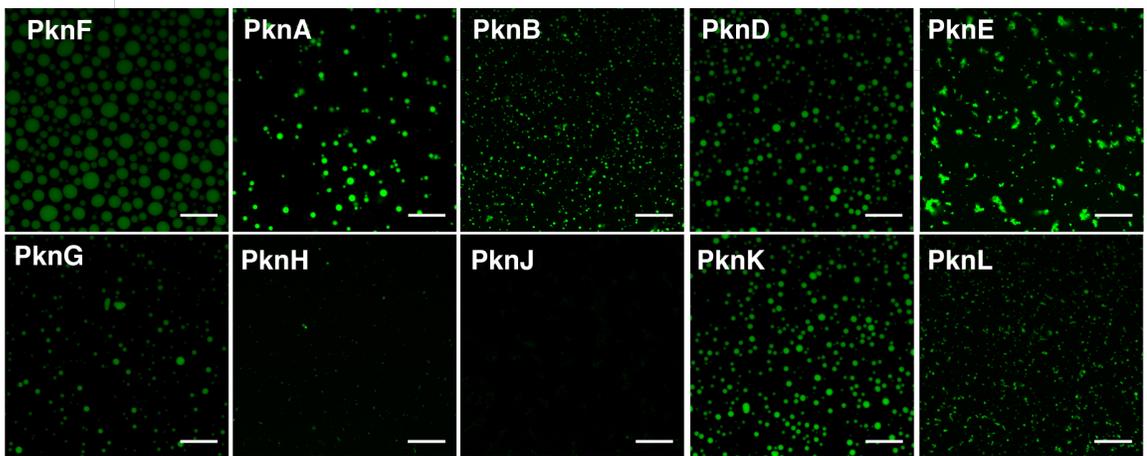
A**B**

Figure 3.10: Several *Mtb* STPKs can Phosphorylate Rv1747¹⁻³¹⁰ at Multiple Sites and Induce Phase Separation.

(A) Autoradiograms of SDS-PAGE gels of 10 μM Rv1747¹⁻³¹⁰ and Rv1747¹⁻³¹⁰ T152A/T210A phosphorylated by 0.5 - 2 μM of the kinase domains of a subset of *Mtb* STPKs (5 μCi of [γ -³²P]ATP, 5 mM Mg²⁺ and 5 mM Mn²⁺). Rv1747¹⁻³¹⁰ T152A/T210A is still phosphorylated at additional unmapped sites. Expected sizes of Rv1747¹⁻³¹⁰ and the autophosphorylated STPKs are indicated. (B) Fluorescent images of OG-labeled Rv1747¹⁻³¹⁰ droplets formed by these STPKs under the same conditions as in (A) except using only 100 μM of non-radioactive ATP (scale bars: 20 μm). At this concentration, unmodified Rv1747¹⁻³¹⁰ does not phase separate.

3.3.8 Rv1747¹⁻³¹⁰ Phase Separates in the Context of Model Membranes

Rv1747¹⁻³¹⁰ is the cytosolic regulatory module of a membrane-spanning ABC transporter. As a step towards determining whether the demixing induces clustering in a membrane context, Rv1747¹⁻³¹⁰ was attached to a Supported Lipid Bilayer (SLB). This is a simple bilayer of defined lipid composition on a glass coverslip. Membrane-association was achieved via an N-terminal His₆-tag on Rv1747¹⁻³¹⁰ that binds with high affinity to a Ni²⁺-containing DGS-NTA lipid within the SLB. Varying the amount of this lipid allowed the pseudo-2D-concentration of Rv1747¹⁻³¹⁰ at the membrane to be controlled. Using this system, even in the absence of phosphorylation, OG-labeled Rv1747¹⁻³¹⁰ readily separated into foci in a cooperative process with a sharp concentration threshold (Figure 3.11). The membrane-anchored clusters still possessed a dynamic, liquid character, as evidenced by diffusive exchange in FRAP experiments (Figure 3.12). This exchange however, was slower than seen for unmodified droplets formed in solution ($\tau_{1/2} = 32 \pm 2$ s, $I_{\text{mobile}} = 0.60 \pm 0.02$ vs $\tau_{1/2} = 18 \pm 0.5$ s, $I_{\text{mobile}} = 0.60 \pm 0.01$ for unmodified Rv1747¹⁻³¹⁰ droplets in solution), potentially caused by the membrane-anchoring of the condensate. Furthermore, at higher DGS-NTA (Ni²⁺) concentrations in the SLB, a characteristic spinodal decomposition of the membrane-anchored clusters into coarse extended networks was seen (193).

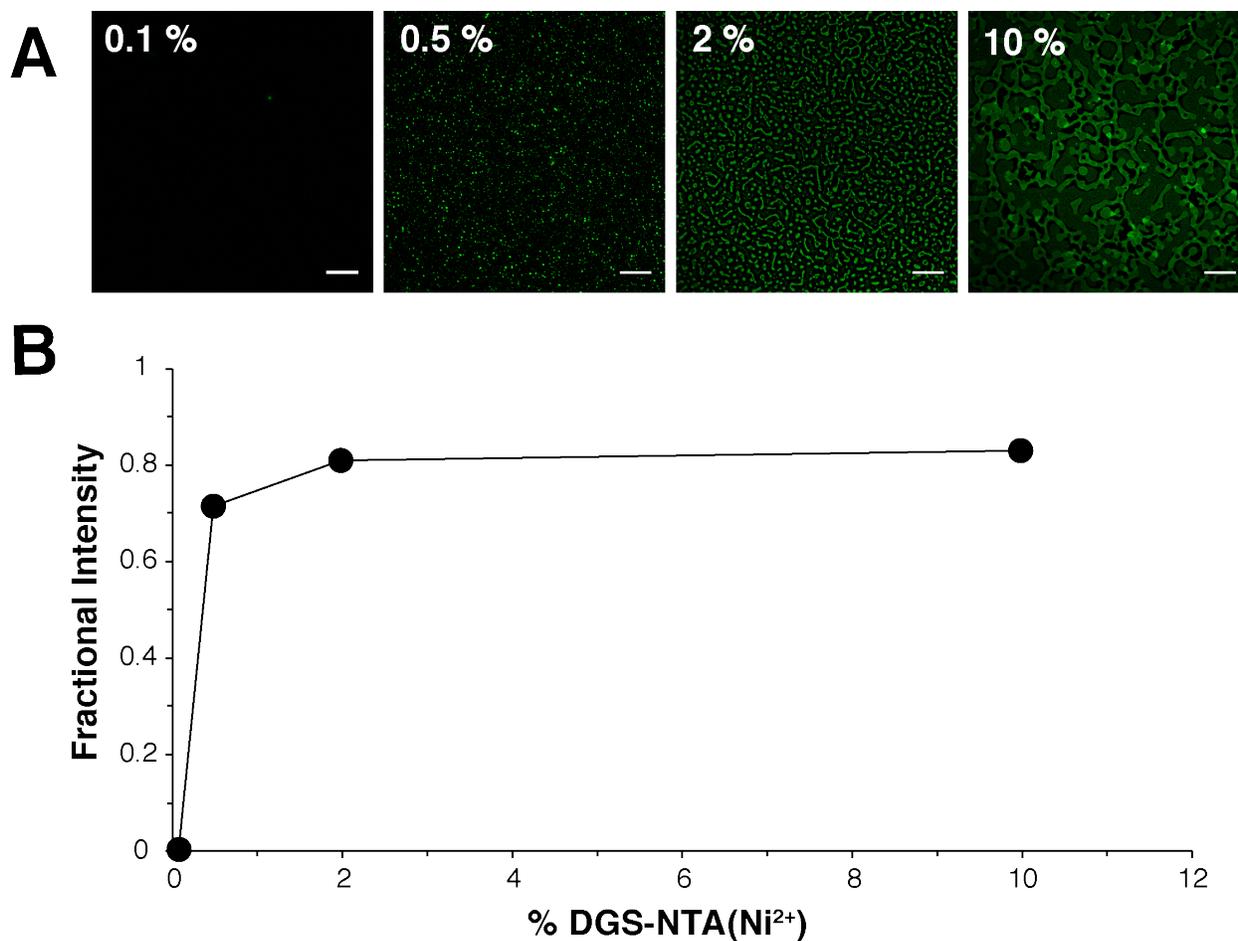


Figure 3.11: Rv1747¹⁻³¹⁰ Anchored in Supported Lipid Bilayers Shows Concentration-dependent Phase Separation.

(A) Fluorescence images of His₆-tagged OG-labeled Rv1747¹⁻³¹⁰ anchored to SLBs containing DGS-NTA(Ni²⁺) (scale bars: 20 μm). The weight percentage of the NTA(Ni²⁺) lipid sets the pseudo-2D-concentration of Rv1747¹⁻³¹⁰ bound to the SLB. (B) Even without phosphorylation, droplet formation as quantified by the fractional fluorescence intensity occurred at a relatively low threshold concentration when anchored in a membrane.

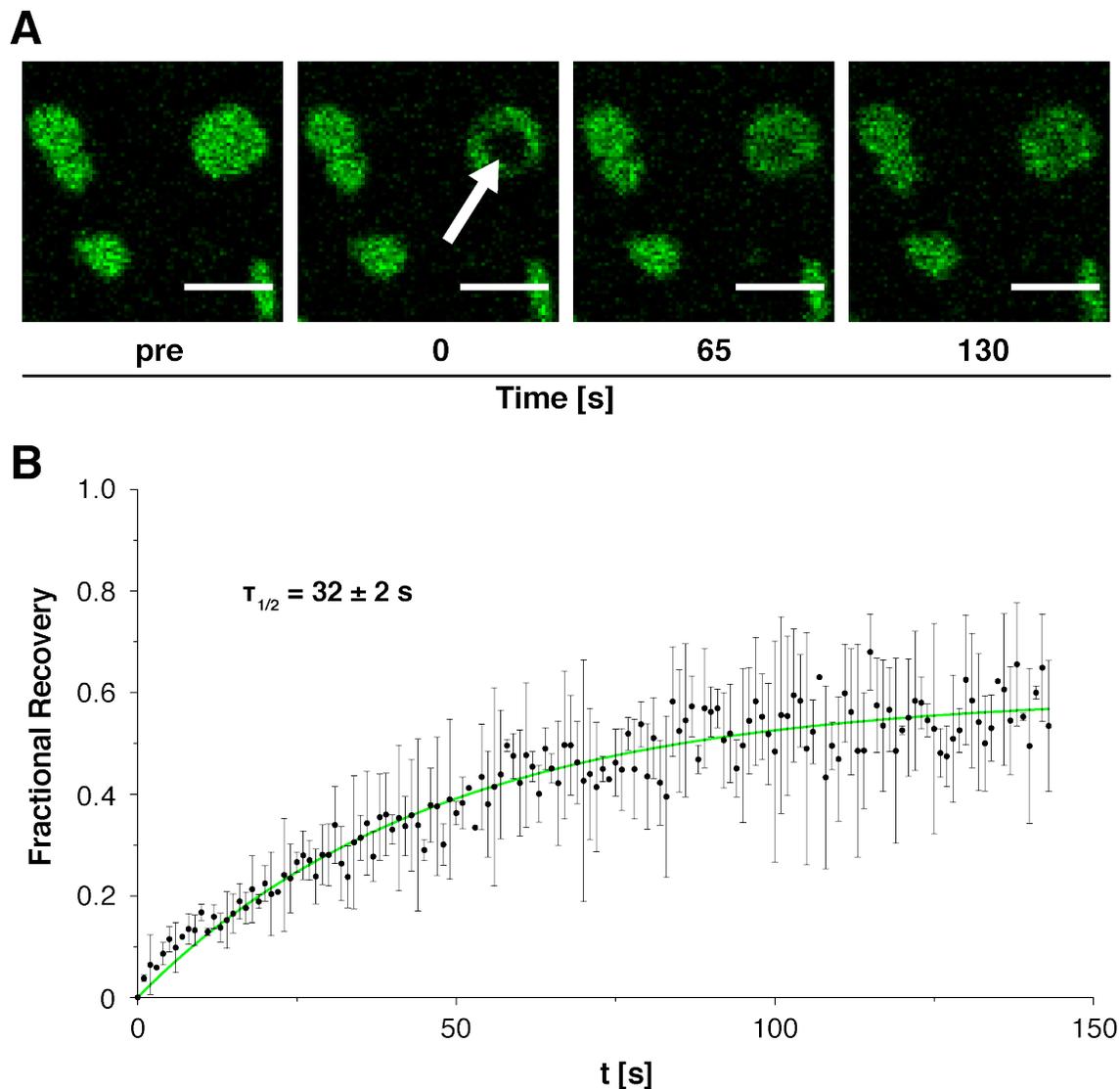


Figure 3.12: Diffusive Exchange of Rv1747¹⁻³¹⁰ Clusters on Supported Lipid Bilayers by Fluorescence Recovery after Photobleaching.

Diffusive exchange at 22 °C in condensates of Rv1747¹⁻³¹⁰ anchored in SLBs shown by FRAP. SLBs were formed as in (Figure 3.11) with 2 % DGS-NTA(Ni²⁺) in the lipid mixture. (A) The bleached sector is indicated by an arrow in the t = 0 s time point image, followed by representative images taken at two subsequent time points (scale bars: 3 μm). (B) Recovery was quantified by fitting the average normalized intensity (solid dots with bars showing standard deviations) of two different FRAP measurements on membrane-anchored, unmodified condensates of Rv1747¹⁻³¹⁰ to a single exponential recovery function (green lines). The time to 50 % recovery ($\tau_{1/2}$) is indicated.

3.3.9 Preliminary Evidence for Rv1747 Clustering along the *Mtb* Cellular Membrane

Having discovered that Rv1747¹⁻³¹⁰ undergoes phosphorylation-enhanced phase-separation *in vitro*, we asked if this translates into membrane clustering of native Rv1747 *in vivo*. This hypothesis was made based on previous studies showing clustering of membrane-anchored T-Cell receptor signaling components due to phase separation (153, 193). As a step towards addressing this critical question, a phage display protocol was used to generate a Single Chain Variable Fragment (scFv) that binds the N-terminal FHA-1 domain of Rv1747. In contrast to a control "non-specific" scFv from the same library (i.e., specific for an unrelated protein), the selected scFv co-localized to pre-formed phosphorylated Rv1747¹⁻³¹⁰ droplets (Figure 3.13). The anti-FHA-1 scFv also bound *Mtb* H37Rv cells that were fixed in para-formaldehyde and detergent permeabilized (Figure 3.14). However, the resolution of these conventional fluorescence microscopy experiments was insufficient to ascertain the cellular distribution of scFv.

Using super resolution Direct Stochastic Reconstruction Microscopy (dSTORM), the fluorescent signals from the AF764-labeled scFv were found in non-uniform clusters along the bacterial membrane (Figure 3.15). Cluster analysis of the Single Molecule Localization Microscopy (SMLM) data was performed using StormGraph (Scurll et al; Manuscript in preparation). This analysis showed that a number of clusters were significantly larger than what would be expected from the scFv binding to a simple homodimer of the transporter. In collaboration with Dr. Libin Abraham, we estimate the cutoff area indicative of higher order multimerization to be about 1800 nm² (see Materials and Methods for details). Although highly encouraging, it must be stressed that these are very preliminary observations and need corroboration by on-going experiments. Important controls include studies of *Mtb* lacking Rv1747 and with altered expression of STPKs such as PknF.

In parallel, electron microscopy was also used to examine the distribution of Rv1747 in *Mtb*. Thin sections of fixed cells were treated with the same scFv, along with 5 nm gold nanoparticles for detection. The scFv contained an N-terminal His₆-tag to which the functionalized gold nanoparticles could bind. Electron micrographs revealed isolated assemblages of gold particles in the cell envelope of *Mtb* (Figure 3.16). Again, although

indicative of scFv, and hence Rv1747, clustering, further experiments are required to justify this conclusion.

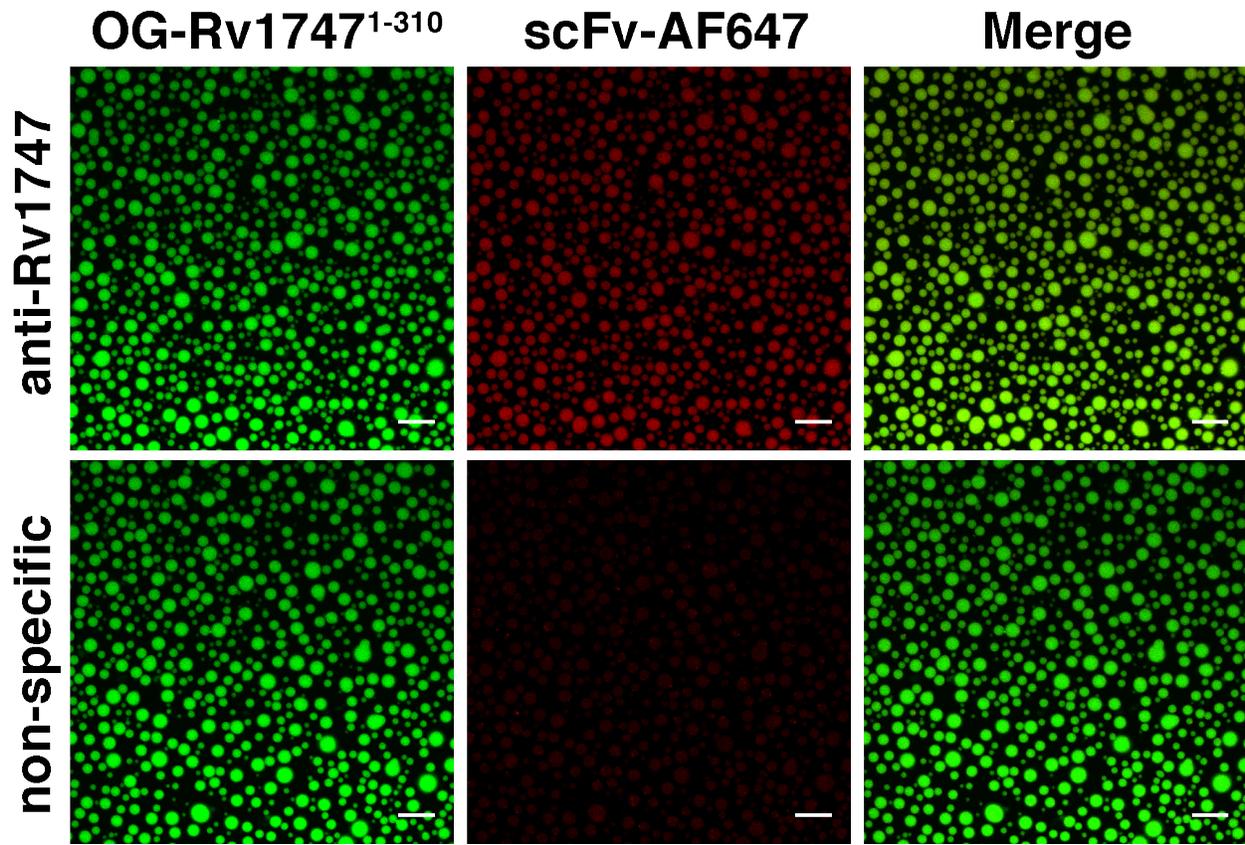


Figure 3.13: Single Chain Variable Fragment (scFv) Selected Against FHA-1 Co-localizes to Rv1747¹⁻³¹⁰ Droplets.

Condensates of 50 μM OG-labeled Rv1747¹⁻³¹⁰ were formed in buffer containing 5 mM Mg^{2+} , 100 μM ATP and 0.5 μM PknF. Subsequently, 0.05 μM of an AF647-labeled scFv that was selected in a phage display to bind FHA-1 was added (top). OG (green) and AF647 (red) fluorescence was measured and overlaid to show co-localization. As a control, the same experiment was repeated with a non-specific AF647-scFv that only differs from the specific one in the variable region. No co-localization was observed (bottom; scale bars; 20 μm).

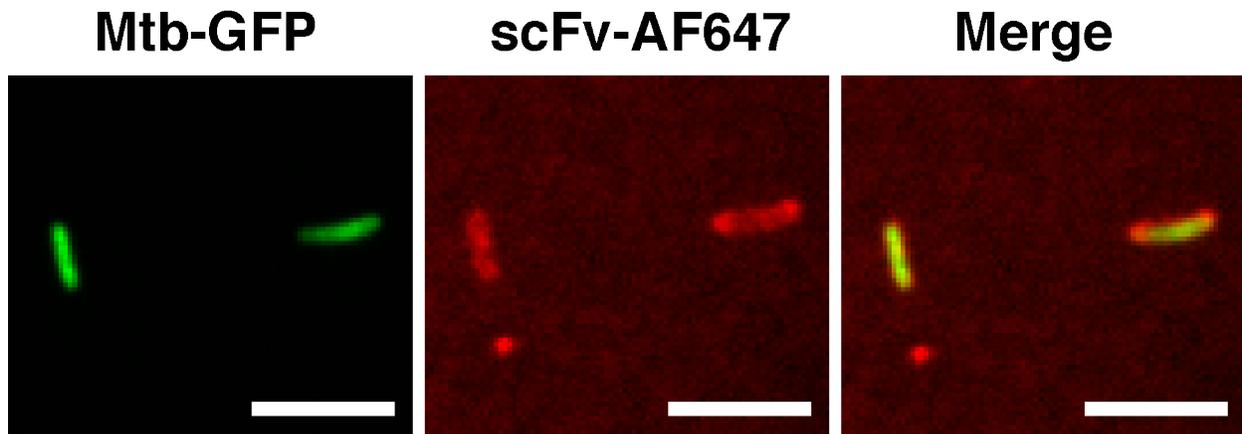


Figure 3.14: The scFv Selected against Rv1747 Localizes to Permeabilized Cells of *Mtb*. *Mtb* H37Rv cells expressing green fluorescent protein (GFP) were fixed with para-formaldehyde, permeabilized with lysozyme/tritonX100, and treated with the AF647-labeled anti-FHA-1 scFv. For details, see methods section. GFP (left) and AF647 (middle) Total Internal Fluorescence (TIRF) images of fixed cells were taken and overlaid (right) (scale bar: 2 μ m).

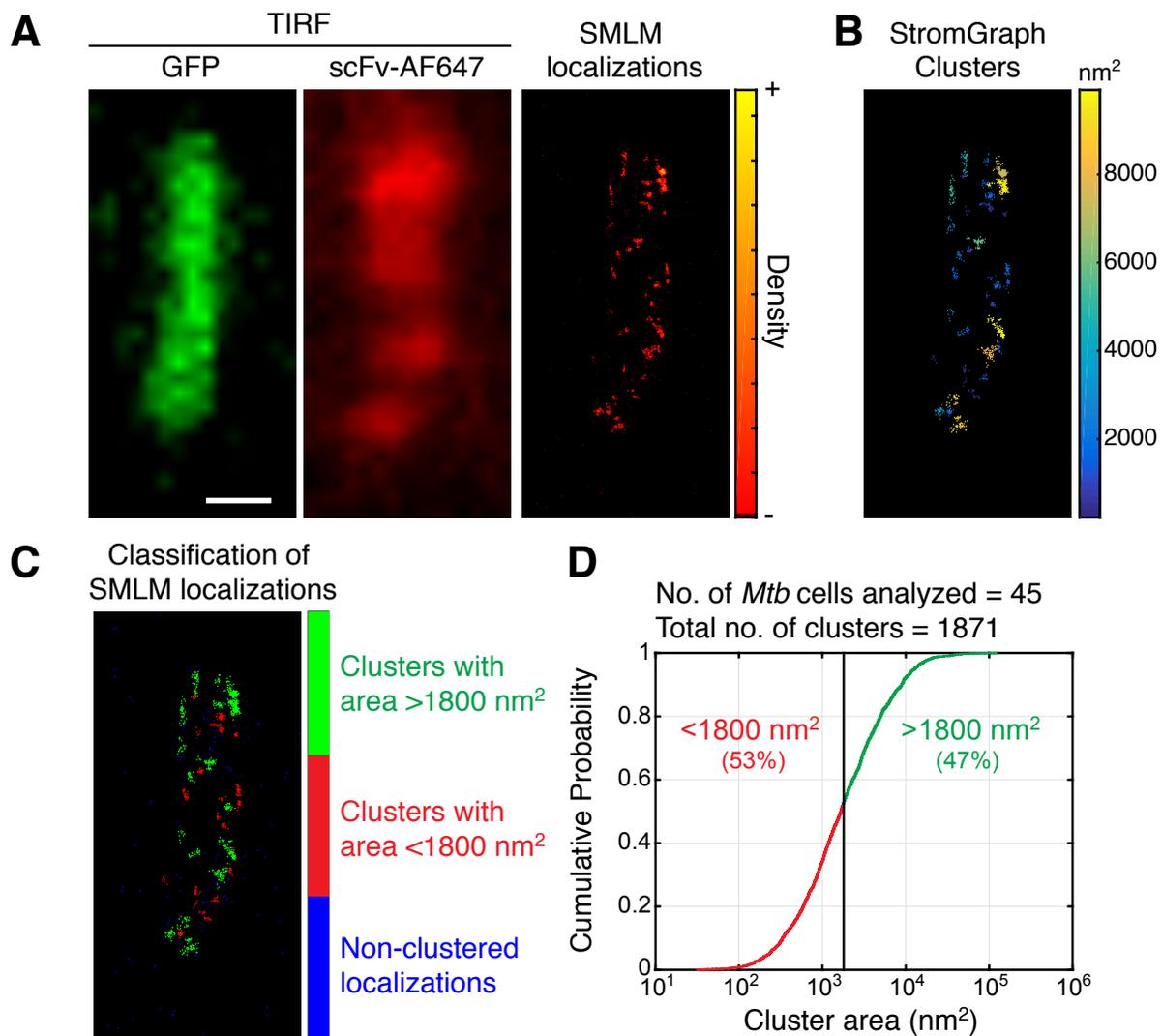


Figure 3.15: Analysis of Rv1747 Clustering in fixed *Mtb* Cells by dSTORM.

(A) GFP (left) and AF647 (middle) TIRF images of *Mtb* H37Rv cells expressing GFP, treated with AF647-labelled Rv1747-specific scFv. These images are reproduced from Figure 3.14 in a side-by-side comparison with the (right) SMLM locations from a dSTORM experiment for one sample *Mtb* cell (for experimental details, see Materials and Methods) (scale bar: 500 nm). The density of reconstructed locations is indicated in a red-yellow gradient. (B) Sizes of clusters identified by StormGraph mapped onto the SMLM locations from (A). Their distribution is indicated by a blue-green-yellow gradient. (C) Classification of clusters from (B) into unambiguous oligomers (green $> 1800 \text{ nm}^2$) and those possibly arising from a simple dimer of the transporter (red $< 1800 \text{ nm}^2$). Non-clustered locations are shown in blue. (D) Results from the size analysis of clusters from 45 *Mtb* cells. Shown is the cumulative probability of finding clusters of certain sizes. Nearly one-half (47 %) of clusters were found to be larger than 1800 nm^2 and classified as arising from oligomers of Rv1747.

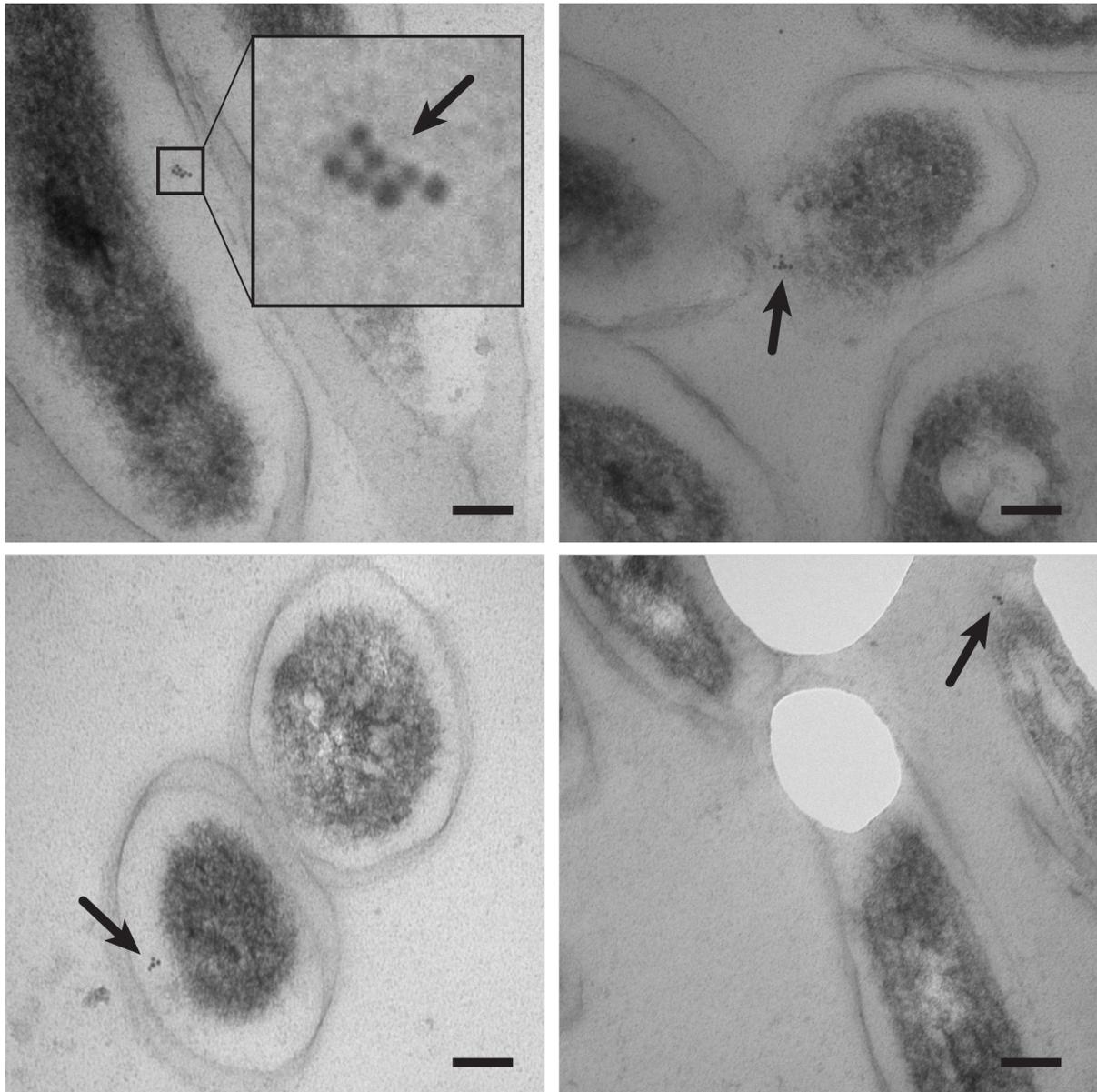


Figure 3.16: Clustering of Rv1747 in *Mtb* Cells Shown with Electron Microscopy.

Electron micrographs of ~ 70 nm microtome sections of para-formaldehyde fixed H37Rv *Mtb* cells treated with His₆-tagged Rv1747-specific scFv and Ni²⁺-functionalized 5 nm gold nanoparticles. The nanoparticles can bind the scFv, which in turn can bind Rv1747 in the bacteria. Potential clusters of gold nanoparticles are indicated with arrows in the lighter cell envelope region of fixed *Mtb* cells (scale bar: 100 nm). A magnified view of one of the clusters is shown.

3.4 Discussion

In summary, the work presented in this chapter provides the first evidence for a prokaryotic protein to undergo phase separation into liquid-like, protein-rich droplets, as well as initial data linking this phenomenon to the biological context of Rv1747. I show in an *in vitro* system that the regulatory module of the ABC transporter undergoes phosphorylation-enhanced clustering triggered by multiple *Mtb* STPKs, as well as the phosphatase-dependent dissolution of these droplets. Furthermore, I see clusters of the regulatory module anchored in a model membrane system. Very preliminary collaborative studies using electron microscopy and ultra-high resolution dSTORM microscopy hint at the clustering of wild type Rv1747 in *Mtb* cells. A proposed model for the phosphorylation-enhanced clustering of Rv1747 is shown in Figure 3.17. In this model, Rv1747 possesses an intrinsic ability to form phase-separation-based clusters at high concentrations. This phase separation is enhanced by phosphorylation and can then happen at much lower concentrations.

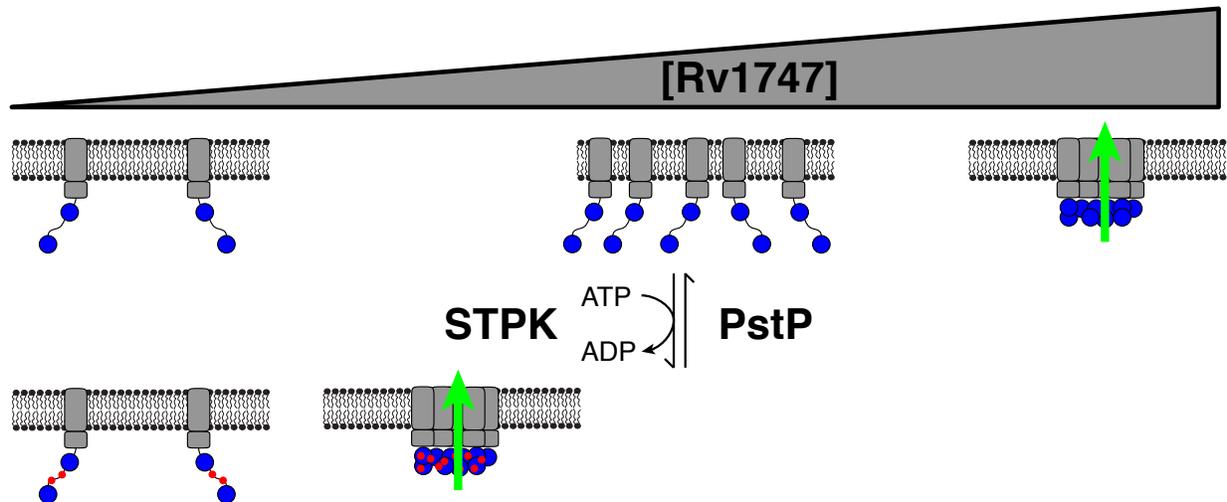


Figure 3.17: Model for Phosphorylation-enhanced Clustering and Activation Mechanism of Rv1747.

Rv1747 (core ABC structure: grey) forms clusters due to the phase separation of the FHA regulatory module (FHA domains: blue circles). This in turn increases the transporter activity of Rv1747, albeit by an undefined mechanism (green arrow). The critical concentration for clustering (upper triangle) is significantly decreased upon phosphorylation (red circles) by *Mtb* STPKs compared to the unmodified transporter. The phosphatase PstP can revert the phosphorylation-enhancement of clustering. Additional interactions with regulatory proteins or even transporter substrates could shift the phase diagram for droplet formation (not shown).

3.4.1 Phosphorylation-Independent Phase Separation

Initially the regulatory module of Rv1747 was found to undergo phase separation upon phosphorylation by PknF. However, analysis by mass spectrometry indicated that, on average, the protein was only modified at sub-stoichiometric levels. Subsequent studies demonstrated that the unmodified regulatory module also forms droplets, albeit at a higher threshold concentration ($\sim 250 \mu\text{M}$ versus $10 \mu\text{M}$) than the kinase-treated Rv1747¹⁻³¹⁰. This indicates that phase separation results from a combination of specific FHA-pThr interactions and additional interactions involving the unmodified protein.

The basis of this basal capability for demixing might involve interactions of the structured FHA domains via non-canonical interaction surfaces with each other. Indeed, dimerization of FHA domains has been reported (272, 273, 312). Alternatively, the FHA domains could be binding non-phosphorylated residues of the ID linker. However, weak, multivalent associations could also arise between linker residues. The linker is of low sequence complexity with a high glycine, proline and threonine content (Figure 3.18). Furthermore, as previously found in other phase-separating disordered proteins (206), its sequence shows features such as RG/GR motifs and sparsely distributed aromatic residues. Together, these are capable of π -cation interactions. It will be important to determine the exact molecular basis of the inherent tendency for phase separation of Rv1747 and how it is enhanced by phosphorylation. In a preliminary experiment, a mutant of Rv1747¹⁻³¹⁰ with all 22 linker prolines mutated to glycine shows enhanced phase separation (not shown). This is in line with RG/GR motifs playing a role as the mutant introduces several more of these by changing RP/PR to RG/GR. Unmodified droplets show a more liquid character in FRAP experiments than their phosphorylated counterparts. This is in line with less specific, weaker interactions forming the condensate scaffold. In our proposed phosphorylation-enhanced model for Rv1747 phase separation, phosphorylation introduces the more specific and tighter FHA-pThr interactions which explains the less liquid character based on a tighter network inside phosphorylated droplets overall (Figure 3.19).

10	20	30	40	50
MPMSQPAAPP	VLTVRYEGSE	RTFAAGHDVV	VGRDLRADVR	VAHPLISRAH
60	70	80	90	100
LLLRFDQGRW	VAIDNGSLNG	LYLNNRRVPV	VDIYDAQRVH	IGNPDGPALD
110	120	130	140	150
FEVGRHRGSA	GRPPQTSIR	LPNLSAGAWP	TDGPPQTGTL	GSGQLQQLPP
160	170	180	190	200
A TRIPAAAPP	SGPQPRYP TG	GQQLWPPSGP	QRAPQIYRPP	TAAPPPAGAR
210	220	230	240	250
GGTEAGNLAT	SMMKILRPGR	LTGELPPGAV	RIGRANDNDI	VIPEVLASRH
260	270	280	290	300
HATLVPTPGG	TEIRDNRSIN	GTFVNGARVD	AALLHDGDVV	TIGNIDLVFA
310				
DGTLARREEN				

Figure 3.18: Sequence Analysis of Rv1747¹⁻³¹⁰.

The sequence of the Rv1747 regulatory module is shown. Residues forming the two structured FHA domains are in bold. Within the ID linker, positively (red) and negatively (blue) charged residues are indicated, along with aromatic residues (green), glycines (cyan), and threonines (magenta). The two mapped phospho-acceptor threonines, T152 and T210, for PknF are highlighted in yellow.

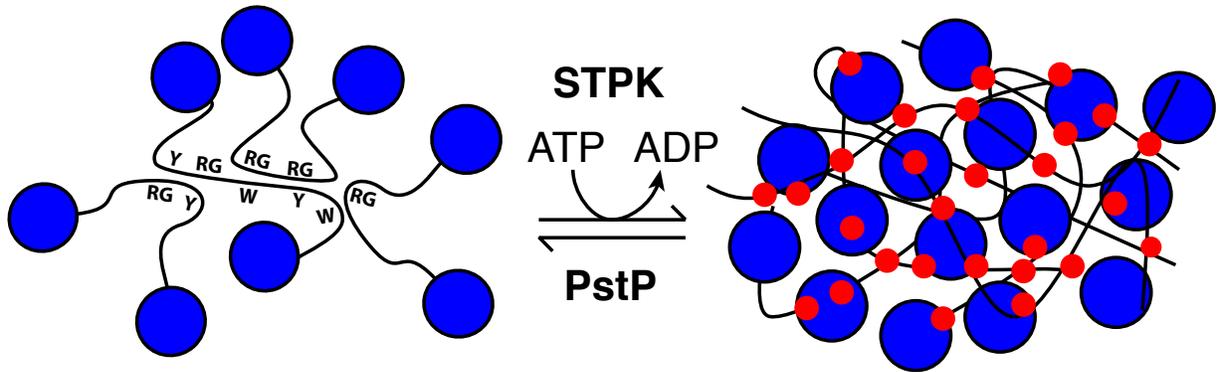


Figure 3.19: Model for Phosphorylation-enhanced Phase Separation and Different Liquid Properties of Rv1747¹⁻³¹⁰ Droplets.

Model of interactions making up unmodified (left) vs phosphorylated (right) Rv1747¹⁻³¹⁰ droplets (FHA domains: blue circles; phospho-threonine: red circles). Unmodified droplets are proposed to be formed based on non-specific interactions of the disordered linker. Linker residues forming possible cation- π interactions are schematically indicated with their one letter code (R, G, Y, W). These droplets are highly dynamic and have a more liquid character than phosphorylated droplets, indicated by a less tight packing in the condensate. Conversely, the added, more specific FHA-pThr interactions in droplets phosphorylated by *Mtb* STPKs cause a less liquid, more gel-like character, indicated by a tighter packing of protein molecules inside the droplets.

3.4.2 Multiple Kinases Induce Phosphorylation-enhanced Phase Separation

Although PknF is considered to be the cognate kinase for Rv1747, most *Mtb* STPKs tested phosphorylated Rv1747¹⁻³¹⁰ *in vitro* and induced phase separation. This is consistent with the observations that PknF is not essential for the activation of the transporter and for virulence (144). Furthermore, modifications also occurred at additional sites other than the two well-documented PknF phospho-acceptors (T152 and T210). This can be explained by differences in the partially overlapping specificities of the kinases (87) and the phenomenon of "hyper-activity" of their catalytic domains under non-physiological *in vitro* conditions with high kinase and ATP concentrations (313). I speculate that the low complexity of the ID linker with twelve threonines that are potentially available for modification may lead to a "multivalency-dial" (Figure 3.18). That is, increasing levels of Rv1747 phosphorylation could progressively promote clustering at the membrane and thereby alter transporter activity, albeit by currently undetermined mechanisms. This could also affect the physicochemical properties, and hence biochemical functions, of the resulting Rv1747 condensates as seen with the different morphologies of the droplets induced by different STPKs.

3.4.3 Biological Roles of Phosphorylation-enhanced Rv1747 Phase Separation

Rv1747 is implicated in transporting intermediates of the cell wall biosynthesis pathway across the *Mtb* cell membrane. The thick cell wall is unique to mycobacteria and especially important for host infection and to survive the harsh conditions after ingestion by macrophages. Hypotheses for the biological roles of Rv1747 phase separation include an increased transport efficiency due to higher local concentrations of transporter in regions of the cell where cell wall intermediates are produced, signal amplification by acting as a molecular filter for these intermediates, or by acting as a scaffold, bringing several players of cell wall biosynthesis pathways together in a heterogeneous membraneless cell-wall-biosynthesis-"organelle".

One potential component that could be found in a heterogeneous, FHA-pThr based Rv1747 cluster is the FHA-domain containing protein Rv0020c. This protein interacts with clusters of penicillin binding protein A (PbpA) and plays a role in cell wall biosynthesis (314). In this context it has been shown to interact with the pseudokinase MviN which is essential for growth and peptidoglycan synthesis at a site that gets phosphorylated by PknB (96). Through

that interaction, Rv0020c also localizes to clusters in poles and septa of cells, potentially reflecting the hypothesized cell-wall-biosynthesis-“organelle”. Furthermore, similar to that seen for Rv1747, there is evidence for a “spiraled” clustering of FtsZ, which is part of the large divisome complex (315). A component of this complex, Rv0019c, also contains an FHA domain and therefore could be a link, possibly interacting with the Rv1747 scaffold and compartmentalizing the cell wall biosynthesis and cell division components upon the right STPK signal input (254, 316). Lastly, structural proteins in bacteria also can also form dynamic clusters/patches (317).

Inhibiting Rv1747 phase separation could be a viable strategy in the quest for novel therapeutic strategies against the growing threat of drug-resistant *Mtb*. Our findings are also especially interesting as clustering of ABC transporters has been the source of controversy and phase separation could be the broader underlying mechanism. The most notable example is the Cystic Fibrosis Transmembrane Conductance Regulator (CFTR) for which there is evidence for a fully disordered regulatory domain and phosphorylation therein playing an important role in facilitating interactions (318).

3.5 Materials and Methods

3.5.1 Expression and Purification of Rv1747¹⁻³¹⁰ for Droplet Studies

Using *M. tuberculosis* H37Rv genomic DNA as a template, the gene encoding Rv1747¹⁻³¹⁰ was PCR amplified and cloned with a TAA stop codon into the pET28MHL plasmid (Addgene, plasmid #26096) via *Nde*I and *Hind*III restriction sites. An N-terminal single cysteine for fluorescent labeling, as well as phospho-ablative threonine to alanine mutants (T152A/T210A) of the Rv1747 construct, were introduced via the QuikChange site-directed mutagenesis protocol (Stratagene). After cleavage of an N-terminal His₆-affinity tag with TEV protease, two non-native residues (Gly-His) preceded the expressed Rv1747 fragments.

Protein constructs were expressed in *E. coli* Rosetta 2 (λ DE3) cells in LB media supplemented with 35 μ g/mL chloramphenicol and 35 μ g/mL kanamycin. Cultures of 1 L were incubated shaking at 37 °C until OD₆₀₀ ~ 0.6, induced with IPTG (1 mM final) and grown for another 16 h at 24 °C. After centrifugation for 15 min at 4000 g, the cell pellets were frozen at -80 °C, then later thawed, resuspended in lysis buffer (20 mM sodium phosphate, 1 M NaCl, 10 mM imidazole, pH 8.0) and sonicated. After centrifugation for 1 h at 30000 g, the supernatant was filtered (0.45 μ m cutoff), applied to a Ni²⁺-NTA HisTrap HP affinity column (GE Healthcare Life Sciences), washed with 10 column volumes of binding buffer (20 mM sodium phosphate, 500 mM NaCl, 30 mM imidazole, pH 8.0) and eluted in a 30 – 500 mM imidazole gradient. Fractions containing the desired protein were identified with SDS-PAGE, pooled and treated with His₆-tagged TEV protease while dialyzing against 4 L of cleavage buffer (20 mM sodium phosphate, 100 mM NaCl, 1 mM DTT, 0.5 mM EDTA, pH 7) for 16 h at 4 °C. Uncleaved protein, cleaved tag and TEV were removed using a second Ni²⁺-NTA HisTrap HP affinity purification step, and the collected flow-through was concentrated and subjected to size-exclusion chromatography using a Superdex 75 (GE Healthcare Life Sciences) column with sample buffer (100 mM NaCl, 20 mM sodium phosphate, pH 6.0). In the case of preparation of Rv1747¹⁻³¹⁰ samples for studies on SLBs, cleavage with TEV protease was omitted and the concentrated fractions from the initial Ni²⁺-NTA HisTrap purification were subjected to size-exclusion chromatography directly. Final protein fractions were pooled, concentrated using a 10 kDa MWCO centrifugal filter (EMD Millipore), flash frozen and stored at -80 °C until thawed further use. Samples were verified by SDS-PAGE and electrospray ionization mass spectrometry

(ESI-MS), and their concentrations determined by UV-absorbance spectroscopy using the predicted molar absorptivity ϵ_{280} of 24000 M⁻¹ cm⁻¹ (276).

3.5.2 Expression and Purification of Active Kinase Domain Constructs of *Mtb* STPKs and the Phosphatase Domain of PstP

Genes encoding PknA¹⁻³³⁷, PknD¹⁻³⁶⁹, PknE¹⁻³³⁶, PknF¹⁻²⁹², PknJ¹⁻³⁴³, PknK¹⁻²⁹⁹ and PknL¹⁻³⁶⁹ were cloned from *Mtb* H37Rv genomic DNA into the pGEX-4T3 plasmid (GE Healthcare Life Sciences) using *Bam*HI and *Xho*I restriction sites. The resulting constructs, with an N-terminal glutathione S-transferase (GST) affinity tag, were expressed in *E. coli* BL21 (λ DE3) cells in LB media supplemented with 100 μ g/mL ampicillin. Cultures of 1 L were incubated at 37 °C until OD₆₀₀ ~ 0.6, induced with IPTG and grown as summarized in Table 3.1. After centrifugation for 15 min at 4000 g, the cell pellet from each given preparation was frozen at -80 °C, then later thawed, resuspended in buffer (10 mM sodium phosphate, 1.8 mM potassium phosphate, 140 mM NaCl, 2.7 mM KCl, pH 7.3) and sonicated. The supernatant was centrifuged for 1 h at 30000 g, filtered, applied to a GSTrap HP affinity column (GE Healthcare Life Sciences), washed with 10 column volumes of the same buffer, and eluted with 50 mM Tris, 10 mM reduced glutathione at pH 7.4. After SDS-PAGE analysis, pooled fractions were subjected to size-exclusion chromatography using a Superdex 75 (GE Healthcare Life Sciences) column in storage buffer (100 mM NaCl, 20 mM sodium phosphate, pH 7.4). Final protein fractions were pooled, concentrated using a 10 kDa MWCO centrifugal filter (EMD Millipore), flash frozen and stored at -80 °C until thawed for further use. Samples were verified by SDS-PAGE and concentrations determined by UV-absorbance spectroscopy using predicted molar absorptivities ϵ_{280} (276) (Table 3.1).

Table 3.1: Expression Conditions of STPK Constructs.

Kinase	Amino Acids Cloned	Vector	Expression Temperature/Time	IPTG Concentration [mM]	Predicted ϵ_{280} [M⁻¹ cm⁻¹]
PknA	1-337	pGEX-4T3	22 °C, 16 hr	0.5 mM	59000
PknB	1-332	pET22b	22 °C, 16 hr	0.5 mM	15000
PknD	1-369	pGEX-4T3	16 °C, 16 hr	1 mM	71000
PknE	1-336	pGEX-4T3	22 °C, 16 hr	0.5 mM	76000
PknF	1-292	pGEX-4T3	22 °C, 16 hr	0.5 mM	70000
PknG	1-750	pET22b	22 °C, 16 hr	0.1 mM	74000
PknH	1-399	pET22b	37 °C, 4 hr	1 mM	46000
PknJ	1-343	pGEX-4T3	22 °C, 16 hr	0.5 mM	83000
PknK	1-299	pGEX-4T3	22 °C, 16 hr	0.5 mM	56000
PknL	1-369	pGEX-4T3	22 °C, 16 hr	0.5 mM	64000

Genes for PknB⁸⁻²⁹², PknG¹⁻⁷⁵⁰ and PknH¹⁻³⁹⁹ were cloned into and pET22b (via *BamHI/XhoI*). These plasmids added a C-terminal His₆-affinity tag for purification. Protein constructs were expressed in *E. coli* BL21 (λ DE3) cells in LB media supplemented with 100 μ g/mL ampicillin. Cultures of 1 L were incubated with shaking at 37 °C until OD₆₀₀ ~ 0.6, induced with IPTG and grown as shown in Table 3.1. After centrifugation for 15 min at 4000 g, each cell pellet was frozen at -80 °C, then later thawed, resuspended in lysis buffer (20 mM sodium phosphate, 1 M NaCl, 10 mM imidazole, pH 8.0) and sonicated. After centrifugation for 1 h at 30000 g, the supernatant was filtered (0.45 μ m cutoff), applied to a Ni²⁺-NTA HisTrap HP affinity column (GE Healthcare Life Sciences), washed with 10 column volumes of binding buffer (20 mM sodium phosphate, 500 mM NaCl, 30 mM imidazole, pH 8.0) and eluted in a 30 – 500 mM imidazole gradient. Fractions containing the desired protein were identified with SDS-PAGE, pooled, concentrated and subjected to size-exclusion chromatography using a Superdex 75 (GE Healthcare Life Sciences) column with sample buffer (100 mM NaCl, 20 mM sodium phosphate, pH 6.0). Final protein fractions were pooled, concentrated using a 10 kDa MWCO centrifugal filter (EMD Millipore), flash frozen and stored at -80 °C until thawed for further use. Samples were verified by SDS-PAGE and concentrations determined by UV-absorbance spectroscopy using predicted molar absorptivities ϵ_{280} (276) (Table 3.1).

The gene encoding PstP¹⁻²⁴⁰ was PCR amplified and cloned with a TAA stop codon into the pET28MHL plasmid (Addgene, plasmid #26096) via *NdeI* and *HindIII* restriction sites. An N-terminal single cysteine for fluorescent labeling were introduced via the QuikChange site-directed mutagenesis protocol (Stratagene). Protein constructs were expressed in *E. coli* Rosetta 2 (λ DE3) cells in LB media supplemented with 35 μ g/mL chloramphenicol and 35 μ g/mL kanamycin. Cultures of 1 L were incubated shaking at 37 °C until OD₆₀₀ ~ 0.6, induced with IPTG (1 mM final) and grown for another 16 h at 24 °C. After centrifugation for 15 min at 4000 g, the cell pellets were frozen at -80 °C, then later thawed, resuspended in lysis buffer (20 mM sodium phosphate, 1 M NaCl, 10 mM imidazole, pH 8.0) and sonicated. After centrifugation for 1 h at 30000 g, the supernatant was filtered (0.45 μ m cutoff), applied to a Ni²⁺-NTA HisTrap HP affinity column (GE Healthcare Life Sciences), washed with 10 column volumes of binding buffer (20 mM sodium phosphate, 500 mM NaCl, 30 mM imidazole, pH 8.0) and eluted in a 30 – 500 mM imidazole gradient. Fractions containing the desired protein were identified with SDS-PAGE, pooled, concentrated and subjected to size-exclusion chromatography using a Superdex 75 (GE Healthcare Life Sciences) column with sample buffer (100 mM NaCl, 20 mM sodium phosphate, pH 6.0). Final protein fractions were pooled, concentrated using a 10 kDa MWCO centrifugal filter (EMD Millipore), flash frozen and stored for further use at -80 °C. Samples were verified by SDS-PAGE and concentrations determined by UV-absorbance spectroscopy using $\epsilon_{280} = 6100 \text{ M}^{-1} \text{ cm}^{-1}$ (276). For simplicity, the constructs of tagged catalytic domains of the STPKs and PstP are denoted by the name of the wild type protein throughout Chapter 3.

3.5.3 Fluorescent Labeling of Rv1747¹⁻³¹⁰, GST-PknF¹⁻²⁹² and PstP¹⁻²⁴⁰

Rv1747¹⁻³¹⁰ with a non-native N-terminal cysteine was conjugated Oregon Green (OG, Invitrogen). Similarly, GST-PknF¹⁻²⁹² and PstP¹⁻²⁴⁰ were linked to AlexaFluor647 (AF647, Invitrogen). For the conjugation reactions, a 10-fold molar excess of maleimide-functionalized fluorescent dye was added to protein samples in conjugation buffer (100 mM NaCl, 20 mM sodium phosphate, 1 mM TCEP, pH 7.0) and incubated at 22 °C in the dark while constantly inverting for 120 min. The reactions were stopped by the addition of 1 mM of reduced glutathione. Samples were then subjected to size-exclusion chromatography using a Superdex 75 (GE Healthcare Life Sciences) column with sample buffer (100 mM NaCl, 20 mM sodium

phosphate, pH 6.0). Final protein fractions were pooled, concentrated using a 10 kDa MWCO centrifugal filter (EMD Millipore), flash frozen and stored at -80 °C. Concentrations of protein samples were determined by UV-absorbance as stated above. Labeling efficiencies with the fluorescent dyes of around 5 – 10 % were determined using UV-absorbance and the molar absorptivities at the maximum absorbance for OG and AF647 of $\epsilon_{496} = 70000 \text{ M}^{-1} \text{ cm}^{-1}$ and $\epsilon_{650} = 270000 \text{ M}^{-1} \text{ cm}^{-1}$, respectively.

3.5.4 Imaging of Rv1747¹⁻³¹⁰ Droplets

DIC and fluorescence microscopy was performed on a Olympus FV1000 inverted confocal microscope and an UplanSApo 60×/1.35 numerical aperture (NA) oil immersion objective. For each experimental setup, all images were acquired maintaining consistent lamp/laser power, exposure time, gain, and offset. For fluorescent images, excitation lasers generated 473 nm (OG) and 633 nm (AF647) wavelength light and the appropriate emission filters for maxima at 488 nm and 647 nm for OG and AF647 were used. All images were processed and analysed with ImageJ (National Institutes of Health, Bethesda, MD).

3.5.5 Phase Separation of Rv1747 by *Mtb* STPK Phosphorylation

Samples of OG-Rv1747¹⁻³¹⁰ at defined concentrations were incubated at 22 °C for 120 min with different molar ratios of PknF kinase domain in 20 mM sodium phosphate, 100 mM NaCl, and 5 mM MgCl₂ at pH 7.4. For sub-stoichiometric phosphorylation, PknF:OG-Rv1747¹⁻³¹⁰ molar ratio was 1:100 and the ATP was 100 μM. For stoichiometric phosphorylation, the ratio was 1:10 with 5 mM ATP. Reactions for microscopy were performed in CoverWell imaging chambers (Sigma-Aldrich) that were pre-treated with 30 mg/mL BSA for 30 mins, rinsed with double-distilled H₂O and dried before use. Phosphorylation levels after 120 mins were determined by ESI-MS on thoroughly re-suspended samples of droplets. After a background correction that was carried out in ImageJ as described before (319), the classification of pixels into droplets and non-droplets pixels was done using the triangle automated thresholding algorithm in ImageJ (193). For quantification of phase separation, the fractional intensity I_{frac} was calculated as the integrated intensity of pixels classified as part of droplets $I_{droplet}$ divided by the integrated intensity of all pixels I_{all} of the image:

$$I_{frac} = \frac{I_{droplet}}{I_{all}}$$

Thresholding using the Huang method in ImageJ gave very similar critical concentrations for phase separation and quantification is therefore not dependent on the thresholding algorithm used.

3.5.6 Phase Separation of Unmodified Rv1747¹⁻³¹⁰

Samples of OG-Rv1747¹⁻³¹⁰ were slowly concentrated at 22 °C and 2000 g in 10 kDa MWCO centrifugal filters (EMD Millipore). Frequently, centrifugation was stopped, the sample thoroughly resuspended and the concentration determined using UV-absorbance as described above. Part of the resuspended sample was also transferred to pre-treated CoverWell imaging chambers (Sigma-Aldrich) and incubated for 120 mins before fluorescent imaging and quantification of phase separation as described above.

3.5.7 Fluorescence Recovery After Photobleaching

FRAP studies of unmodified and phosphorylated Rv1747¹⁻³¹⁰ droplets, prepared in imaging chambers as described above, were carried out at room temperature (22 °C) using a 473 nm laser. Laser power was set to 5 % and the Tornado scanning tool (Olympus) was used. Fluorescence recovery was recorded at 1 frame/s for 180 to 500 s. The FRAP Profiler Plugin in ImageJ was used for intensity analysis and the normalization to the fluorescence intensity of a non-bleached area of the droplet. The recovery data of three independent experiments were averaged and analysed in GraphPad Prism using a non-linear least squares fit to a single exponential recovery:

$$I(t) = I_{mobile}(1 - e^{-t/\tau})$$

Here, I is the fractional intensity of the bleached region as a function of time, I_{mobile} is the limiting intensity of the mobile fraction and τ is the time constant for recovery. The latter was

convert to the halftime of recovery, $\tau_{1/2} = \tau/\ln 2$, the time it takes to reach half the mobile fraction intensity.

3.5.8 Dephosphorylation of Rv1747¹⁻³¹⁰ by PstP and Dissolution Assay of Droplets

Phase separation was induced by addition of 0.5 μM PknF to samples of 50 μM OG-Rv1747¹⁻³¹⁰ in 20 mM sodium phosphate, 100 mM NaCl, 100 μM ATP, 5 mM MgCl₂ and 5 mM MnCl₂ at pH 7.4. After 120 min, 5 μM PstP was added to one of two samples and droplets were monitored over 24 h at 22 °C by confocal fluorescence microscopy.

3.5.9 Co-localization Studies of Rv1747¹⁻³¹⁰ Droplets with PknF and PstP

Non-fluorescently labeled PknF at 0.5 μM was added to 50 μM samples of OG-Rv1747¹⁻³¹⁰ in 20 mM sodium phosphate, 100 mM NaCl, 100 μM ATP, 5 mM MgCl₂ and 5 mM MnCl₂ at pH 7.4 to induce droplet formation. After 120 min, 0.05 μM AF647-labeled PknF or PstP were added and images of OG (473 nm excitation) and AF647 (633 nm excitation) fluorescence taken and overlaid to monitor co-localization.

3.5.10 Radioactive γ -³²P Kinase Assay of *Mtb* STPKs on Rv1747¹⁻³¹⁰

Purified samples of 10 μM of Rv1747¹⁻³¹⁰ and Rv1747¹⁻³¹⁰ T152A/T210A were incubated with 5 μCi of [γ -³²P]ATP in a buffer containing 20 mM sodium phosphate, 100 mM NaCl, 1 mM DTT, 5 mM MgCl₂ and 5 mM MnCl₂ and 0.5 – 2 μM of STPK kinase domain at pH 7.4. The incubation was performed at room temperature (~ 22 °C) for 5 to 10 min. The reactions were stopped with SDS-sample loading buffer and heated at 95 °C for 10 min. Samples were resolved by a 12% SDS-PAGE. The gels were silver stained and dried. The ³²P-radioactively labeled bands of phosphorylated protein were detected using a PhosphorImager SI (Molecular Dynamics).

3.5.11 Preparation of Rv1747¹⁻³¹⁰ Anchored in Supported Lipid Bilayers and Imaging of Membrane Clustering

Synthetic 1,2-dioleoyl-sn-glycero-3-phosphoethanolamine-N-[methoxy(polyethyleneglycol) -5000] (ammonium salt, PEG5000-PE), 1,2-dioleoyl-sn-glycero-

3- [(N-(5-amino-1-carboxypentyl)iminodiacetic acid)succinyl] (nickel salt, DGS-NTA-Ni) and 1,2-dioleoyl-sn-glycero-3-phosphocholine (POPC) were purchased from Avanti Polar Lipids. Mixtures of 0.1-10 % DGS-NTA-Ni, 99.8-89.9% POPC and 0.1 % PEG5000-PE at 1 mg/mL total final lipid concentration were prepared in 1:1 v/v chloroform:methanol solvent. Glass coverslips (Marienfeld #1.5H, round, 18 mm; catalogue #0107032, Lauda-Königshofen, Germany) were rinsed with ethanol and ddH₂O and then treated with 1 % Hellmanex III solution in a sonication bath for 10 min. After another thorough rinse with ddH₂O, the cover slips were dried. The lipid mixture (40 µL) was added onto the clean cover slips, followed by spin-coating for 30 s at 2000 rpm using a Laurell WS-650 instrument. Any remaining solvent was evaporated for another 5 min. During that time, the cover slips were mounted into ChamSlide CMB 18 mm Magnetic Chambers (Live Cell Instruments) and with 500 µL of buffer (20 mM sodium phosphate, 100 mM NaCl, pH 7.4) immediately after the drying period. Finally, 2 µM of His₆-tagged OG-Rv1747¹⁻³¹⁰ was added to the imaging chambers and, after incubation for 10 mins at 22 °C, the membrane clusters were investigated using confocal fluorescence microscopy as stated above. FRAP studies of unmodified Rv1747¹⁻³¹⁰ condensates anchored in SLBs were carried out at room temperature (22 °C) using a 473 nm laser. Laser power was set to 0.5 % and the Tornado scanning tool (Olympus) was used. Fluorescence recovery in this region was recorded at 1 frame/s for 150 s. The FRAP data was analyzed as stated above.

3.5.12 Selection of Single Chain Variable Fragment (scFv) specific for Rv1747¹⁻¹⁵⁶

Selection, production and purification of the scFv was carried out by the Antibody Engineering Facility at the Immunity and Infection Research Centre (IIRC) at the University of British Columbia (UBC) as described before (320). Briefly, Rv1747¹⁻¹⁵⁶ (FHA-1) was used for biopanning against a M13 bacteriophage phagemid library presenting an optimized library of ~ 10⁹ different scFv clones fused to the major viral coat protein g3p. scFv's with affinity to Rv1747¹⁻¹⁵⁶ were eluted after washing away unspecific bound antibodies and enriched by repeating the biopanning process three times. The highest affinity scFv, as judged by an enzyme linked immunosorbent assay (ELISA), was used for subsequent studies presented herein. After production and purification, the antibody was labelled non-specifically at exposed amines with

an N-hydroxysuccinimide-functionalized AF647 fluorescent dye by the AbLab at the Biomedical Research Centre (BRC) at UBC.

3.5.13 Co-localization Studies of scFv's with Rv1747¹⁻³¹⁰ Droplets

PknF at 0.5 μ M was added to 50 μ M samples of OG-Rv1747¹⁻³¹⁰ in 20 mM sodium phosphate, 100 mM NaCl 100 μ M ATP, 5 mM MgCl₂ and 5 mM MnCl₂ at pH 7.4 to induce droplet formation. After 120 min, 0.05 μ M AF647-labeled Rv1747-specific or non-specific scFv were added. Images of OG and AF647 fluorescence taken and overlaid to monitor co-localization.

3.5.14 Functionalization of Glass Coverslips for dSTORM

Glass coverslips were cleaned and functionalized as previously described (321). Briefly, acid-cleaned glass coverslips (Marienfeld #1.5H, 18 \times 18 mm; catalogue #0107032, Lauda-Königshofen, Germany) were incubated with 0.01% poly-L-lysine (Sigma-Aldrich; catalogue #P4707) for at least 1 h at 37°C. The slides were then washed with PBS prior to the experiment.

3.5.15 Growth, Preparation and Labeling of *Mtb* for dSTORM

All protocols were approved by the University of British Columbia and experiments were carried out in accordance with institutional regulations. *M. tuberculosis* H37Rv harboring pTEC15 (GFP expressing) was grown in Middlebrook 7H9 broth (BD Diagnostic Systems) supplemented with 10% (v/v) OADC (oleic acid, albumin, dextrose and catalase solution) and 0.05% (v/v) Tween-80 (Sigma-Aldrich) at 37 °C in standing cultures. Cultures were grown to OD₆₀₀ ~ 0.5-0.6, washed two times with PBS and resuspended in PBS. Roughly 7.5 million *Mtb* cells were settled on poly-L-lysine coated coverslips for 2 hours at 37 °C and subsequently fixed with PBS containing 4% paraformaldehyde (PFA) for 20 minutes at 37 °C and 4 °C overnight. The cell-containing coverslips were washed thoroughly with PBS and processed for intracellular staining, as previously described (322). Briefly, *Mtb* samples were first treated with Lysozyme (Sigma Aldrich; Catalogue #10837059001; 2 mg/ml in H₂O) for 30 min at 37°C, followed by Triton X-100 (0.1% in H₂O) for 5 min at room temperature 22 °C. After PBS washes, the sample was blocked with 10% normal goat serum in PBS for 30 min at 22 °C, and stained with the

AF647-conjugated, Rv1747-specific scFv, for 1 hour at 22 °C. The coverslips were washed thoroughly with PBS and fluorescent beads for image stabilization (F8799, ThermoFisher Scientific) were added and incubated overnight at 4°C.

3.5.16 dSTORM Setup

Imaging was performed using a home-built microscope with a sample stabilization system as previously described (323, 324). Briefly, three lasers were used in the excitation path: a 639 nm laser (Genesis MX639, Coherent) for exciting AF647, a 488 nm laser (DHOM-100B, Fine Mechanics) for exciting GFP, and a 405 nm laser (LRD 0405, Laserglow Technologies) for reactivating AF647. All three lasers were coupled into an inverted microscope equipped with an apochromatic TIRF oil-immersion objective lens (60×; 1.49 NA; Nikon). The emission fluorescence was separated using appropriate dichroic mirrors and filters (Semrock), and detected by EM-CCD cameras (Ixon, Andor). A feedback loop was employed to lock the position of the sample during image acquisition (323, 324). Sample drift was controlled to be less than 1 nm laterally and 2.5 nm axially.

3.5.17 Single Molecule Localization Microscopy (SMLM) Image Acquisition and Reconstruction

Imaging was performed in an oxygen-scavenging GLOX-thiol buffer containing 50 mM Tris-HCl, 10 mM NaCl, 0.5 mg/ml glucose oxidase, 40 µg/ml catalase, 10% (w/v) glucose and 140 mM 2-mercaptoethanol at pH = 8.0) (325). The coverslip was mounted onto depression slides and sealed with Twinsil two-component silicone-glue (#13001000, Picodent). For SMLM imaging, cells were exposed to a laser power density of 1 kW/cm² for the 639 nm laser to activate the AF647. For each reconstructed SMLM image, 40,000 images were acquired at 50 Hz.

3.5.18 Rv1747 SMLM Cluster Analysis

SMLM data was analyzed and quantified using StormGraph, an automated graph-based algorithm developed by Josh Scurll (UBC Math). Briefly, StormGraph first converts the pointillist data into a graph (or network) in which the localization coordinates specify the nodes, and weighted edges between the nodes describe their similarity (how close they are), and then borrows from the extensive field of “community detection” on graphs to detect clusters of densely connected nodes. In addition, StormGraph can identify and quantify clusters in 3D and is able to factor localization uncertainties into cluster identification (Scurll et al; Manuscript in preparation). The SMLM data was analyzed in 3D with localization uncertainties factored in, and clusters were identified. The area of all clusters from 45 individual *Mtb* cells was plotted as a Cumulative Distribution Function (CDF) plot.

For estimation of the area cutoff of a higher-order multimerization, a simple consideration of the sum of sizes of components of the regulatory module of Rv1747 was done. This included the end-to-end distance of a flexible linker between the FHA domains (~ 8.3 nm), the diameters of FHA domains ($2 \times \sim 3.3$ nm) and scFv (~ 3.8 nm), and a localization error of ~ 5 nm. The end-to-end distance $\langle r^2 \rangle^{1/2}$ of the intervening linker was calculated using a predicted radius of gyration R_g of 3.4 nm according to (326) and assuming a Gaussian chain and therefore the relationship $R_g^2 = 1/6 \times \langle r^2 \rangle$ (327). The length of the whole regulatory module was thereby estimated (~ 23.7 nm) and considered to be the maximum distance of a fluorescent signal from the center of a simple homodimeric ABC transporter pore. This was then used to calculate the maximum circular area covered by this flexible “arm” (~ 1800 nm²).

3.5.19 Electron Microscopy

Samples of *Mtb* H37Rv were produced as described for dSTORM samples above. About 7.5 million cells were fixed with PBS containing 4% PFA for 20 minutes at 37 °C and 4 °C overnight. These fixed samples were then pelleted by centrifugation and resuspended in ~ 100 μ L PBS. Samples were then embedded in Lowicryl HM20 and sample sections of ~ 70 nm thickness were cut using a Leica UC7/FC7 Ultracut microtome and placed on nickel grids. Grids were prepared for immunostaining by blocking in buffer containing 1% BSA, 0.1% Tween-20 in PBS for 30 min using a Leica IGL automatic immunogold labeling system. The grids were then

immunolabeled with Rv1747 scFv (1:10) for 1h at room temperature, washed with buffer containing 0.1% Tween-20 in PBS and blocked with the same blocking solution that was used before. This was followed by incubation with 5 nm Ni-NTA-Nanogold (Nanoprobes) (1:10) for 1h at room temperature. In order to immobilize the scFv-Nanogold complexes, sections were fixed again in 2% PFA (EM grade) for 5 min. Images were acquired using a Tecnai G2 TEM electron microscope (FEI) at the UBC Bioimaging Facility (University of British Columbia, Vancouver, BC, Canada).

Chapter 4: Concluding Remarks

Tuberculosis, especially in its multi-drug-resistant form and in combined infections with HIV, is having a renaissance as a major health concern. It is therefore essential to better understand the physiology of *Mycobacterium tuberculosis*, the pathogen causing the disease. Its signaling system that is based on a set of STPKs is unusual for a prokaryotic organism and one important aspect of its biology. *Mtb* also harbors many interesting targets for interfering with its infectivity. An identified virulence factor with essential STPK phosphorylation is the ABC transporter Rv1747, a putative exporter of cell wall intermediates. The overreaching goal of my thesis was to understand the structural and dynamic changes brought about by phosphorylation of the regulatory module of this virulence-linked transporter and how it acts on a biophysical and cellular level.

4.1 Structural and Dynamical Characterization of the Rv1747 FHA Regulatory Module

Rv1747 is a unique ABC transporter in that it possesses a regulatory module appended to its canonical ATP-binding cassette and transmembrane domains. This module is comprised of two FHA domains that bind phosphothreonine-containing peptides, connected by a disordered linker. There are at least two STPK phospho-acceptor sites in this linker that bind to both of the FHA domains. The presence of these phosphorylation sites, as well as the integrity of the binding sites in at least one of the FHA domains, is essential for virulence of the pathogen (144).

In Chapter 2 of this thesis, I investigated this regulatory module from a structural and dynamic perspective and how phosphorylation by the cognate STPK PknF changes these properties. I determined the unbound structure of FHA-2 by NMR spectroscopy, whereas the X-ray crystallographic structure of FHA-1 was provided by our collaborators. FHA-1 has the expected 11-strand β -sandwich fold of a canonical FHA domain. In contrast, the structural ensemble of FHA-2 revealed the unexpected features of a circular permutation - the new positioning of N- and C-terminus - and the loss of one β -strand. The sequence-dependent origin of these structural differences likely causes the lower stability and greater dynamics of FHA-2 relative to FHA-1 as measured by CD-monitored denaturation and NMR-monitored HX studies. I also speculate that the higher affinity of FHA-2 for phosphopeptides results from its structural

plasticity. Most importantly, I show that upon phosphorylation of the regulatory module, a combination of multivalent intra- and inter-molecular interactions leads to self-association and the formation of higher order multimers and ultimately, phase separation.

4.1.1 Significance, Limitations and Outstanding Questions

FHA-1 shows the expected fold, with a relatively rigid core and pThr peptide binding interface, that has been well established for FHA domains (137, 266, 267). In contrast, FHA-2 is an interesting aberration. With its circular permutation and the lack of one β -strand, combined with its dynamic character, FHA-2 is an outlier amongst many proteins with this ubiquitous domain. My characterization of FHA-2 thus expands the repertoire of known FHA domains that mediate phosphorylation-dependent protein-protein interactions in proteins of many architectures and across all kingdoms of life.

It is important to understand these nuances and how they provide functional specificity in a biological context. In the case of FHA-2, tighter binding appears to correlate with a more flexible binding interface. This constitutes an example of a protein where evolution optimized the enthalpic advantage of an adaptable binding interface to outweigh the entropic penalty of the loss of flexibility in the peptide and the protein upon binding. However, dissecting the link between structural dynamics and the affinity and specificity of peptide binding certainly requires more extensive studies. For example, I only collected data giving insights into the structure and flexibility of the unbound state. To complete the picture, complementary experimental and computational measurements of bound states with a range of phosphopeptides would have to be undertaken.

In these studies, I aimed to determine contributions of FHA-pThr interactions upon phosphorylation of various Rv1747 constructs. Qualitatively, I was able to show higher order association resulting from both intra- and inter-molecular interactions, with relative contributions differing depending on the FHA domain and the modification site. However, it was not possible to tease apart all of the underlying microscopic binding equilibria. Indeed, this system highlights the difficulty of studying homotypic interactions in multi-valent proteins. Common methods for probing binding, such as NMR spectroscopy and ITC, will only show a macroscopic sum of generally indistinguishable intra- and inter-molecular events. Because the local concentrations

for intramolecular (or intracomplex) binding tends to be higher than can be achieved for proteins in solution, they often mask the weaker apparent inter-molecular binding. To separate these processes, the local or effective concentration for intra-molecular binding needs to be estimated. This can be done computationally by sampling possible conformations of the flexible linker in MD simulations and determining average distances between pThr-containing motif and the FHA domain binding interface (328). Experimentally, intramolecular and intermolecular binding are (at least in principle) distinguishable by their concentration independence or dependence, respectively. Thus binding studies as a function of protein concentration, along with varying the sequence separation between the FHA domain interface and pThr site, are required along with model fitting to deconvolute the contributions of intra- and inter-molecular binding.

The importance of these studies lies with understanding how the regulatory module contributes to the transporter function of Rv1747. I speculate that intermolecular FHA-pThr interactions drive the dimerization, or more so, oligomerization of Rv1747 to enable tunable activity depending on sites and levels of phosphorylation. This in turn depends on the balance of kinase and phosphatase activities within the STPK signaling network of *Mtb*, along with other potential regulatory factors. However, the biochemical and cellular functions of the transporter were not directly probed herein.

One path to this goal would involve the expression of wild type Rv1747, followed by reconstitution in liposomes (329) or nanodiscs (330). Although technically challenging, its ATPase and transport activities could be compared as a function of STPK-phosphorylation and oligomerization state. Currently, the exact substrate(s) for Rv1747 are not known, and thus assays with "generic" compounds such as Hoechst 33342 commonly used in the ABC transporter field could be initially pursued (331, 332). In the case of large liposomes or in preparations of SLBs from liposomes (153, 333) association and clustering in the membrane could be monitored at the same time. With these assays in place, it should also be possible to better define the native substrates of Rv1747. As it is implied to transport cell wall components (104, 105), different PIMs should be tested, perhaps along with fractionated *Mtb* extracts. The assay for native substrates could either be done by assaying the ATPase activity of Rv1747 or by detecting transport into inside-out vesicles (334).

4.2 Phase Separation of the Rv1747 FHA Regulatory Module

In Chapter 3, I demonstrated that multivalent self-association of the Rv1747 regulatory module leads to phase separation into liquid droplets. At high concentrations, unmodified Rv1747¹⁻³¹⁰ formed condensates. I speculate that this arises from weak interactions between ID linker residues, or possibly between the FHA domains or via a non-canonical binding of these domains with the linker. Phosphorylation enhanced phase separation due to specific multivalent, FHA-pThr interactions. As clearly seen in FRAP studies, the phosphorylated droplets were less dynamic for diffusion than those formed without phosphorylation. This is in line with the contribution of relatively tight FHA-pThr interactions versus weaker, and presumably non-specific, interactions occurring between the unmodified proteins. In the context of the STPK signaling system of *Mtb*, most kinases tested phosphorylated the regulatory module at T152/T210 and elsewhere, and induced droplet formation. PstP, the sole *Mtb* serine threonine phosphatase, dissolved droplets. Notably, condensates of differing appearance resulted from treatment with different kinases, suggesting a possible dependence on the sites of phosphorylation.

Rv1747¹⁻³¹⁰ also readily phase separated when attached to SLBs as a model membrane system. This indicated that wild-type Rv1747 might also undergo a similar pseudo-2D process at the *Mtb* cell membrane. In support of this notion, very preliminary super resolution microscopy (dSTORM) and electron microscopy studies revealed the apparent clustering of Rv1747 in fixed *Mtb* cells.

4.2.1 Significance, Limitations and Outstanding Questions

At the time of writing this thesis, Rv1747 represents the first bacterial protein that I am aware of to be shown to phase separate. However, physiologically important phase separation in prokaryotes is fully expected. This is a ubiquitous phenomenon exhibited by macromolecules such as proteins, for which we are only beginning to understand its broader significance. Nonetheless, the characterization of Rv1747 marks an important milestone in the field, potentially pointing to a greater evolutionary role of protein condensates. This is consistent with a recent hypothesis that spontaneous phase separation of biologically important molecules played an important role in the origin of life (236).

Rv1747 also represents an interesting case of phosphorylation-enhanced phase separation likely involving a synergy of specific FHA-pThr binding with a background of weaker non-specific interactions between elements of the unmodified protein. Phosphorylation has been shown to affect protein phase separation in different ways including promoting condensation via modular SH2-pThr interactions (153), as well as promoting or disrupting condensates of IDRs, including those with RNA (238, 335). Conversely, disordered linker residues have also been shown to enhance phase separation dominated by classical modular interactions, such as between SH3 domains and proline-rich motifs (152).

Although specific FHA-pThr binding provides a key mechanism for the phase separation of Rv1747, the effects of phosphorylation on the other interactions driving droplet formation, possibly between the disordered linker, need to be tested. Obviously, the isolated ID linker should be expressed for phase separation studies. The dependence of separation on solution conditions (pH, ionic strength, temperature), phosphorylation, and sequence alterations, including patterns of aromatic and charged residues, could then be established. Analysis of the resulting phase diagram according to models, such as that developed by Flory and Huggins for simple polymers (195, 196), would yield valuable insights into the nature of the interactions promoting droplet formation, and their synergy with the specific FHA-pThr binding.

The demonstration that Rv1747¹⁻³¹⁰ phase separates when attached to SLBs is an important step towards understanding the role of this phenomenon in transporter function. It is somewhat difficult to compare the "pseudo-2D-concentration" of the bound protein, set by the fraction of NTA(N⁺2)-lipid in the SLB, with the more conventional "3D-concentration" of the protein in solution. Nevertheless, the unmodified protein readily formed droplets at the lowest level of binding-lipid tested, whereas the threshold concentration for droplet formation by free Rv1747¹⁻³¹⁰ was ~ 250 μM. This suggests that confinement of the intact transporter to a small volume near the membrane surface of an *Mtb* cell would significantly facilitate phase separation.

My work on Rv1747 also marks the first ABC transporter where phase separation of at least a portion of the protein is observed. This has wide-spread significance to understanding the clustering/higher order oligomerization that has been suggested and a source of controversy for other proteins of ABC transporter architecture such as peroxisomal ABCD transporters (336) and the cystic fibrosis transmembrane conductance regulator, CFTR (337-339). In the case of

CFTR, phosphorylation-dependent and independent interactions of disordered appended regulatory sequence that might also lead to phase separation modulate the activity of the transporter (318, 340). Although the active quaternary structure of an ABC transporter minimally involves only two transmembrane domains (TMs) and two nucleotide binding domains (NBDs) (341), clustering of transporters could pose a higher level of regulation. For example, regulation could be achieved by increasing the local concentration of transporter subunits that need to dimerize in order to hydrolyze ATP and transport substrates. Alternatively, higher-order regulation could be achieved on a spatial level by clusters forming in regions of the cell where transport is required and/or substrates are ubiquitous. In the case of Rv1747, clustering could be important in regions of cell growth and division, where remodeling of the cell wall is required by the large protein complex of the divisome (315). Interestingly, other *Mtb* proteins containing FHA domains also localize to this complex (254, 316) which could recruit Rv1747. The observation that several STPKs other than the presumed canonical PknF can phosphorylate, and presumably regulate, Rv1747 is in line with the notion of an interwoven system and cross-talk within the eukaryotic-like serine threonine phosphorylation signaling system in *Mtb* (93).

The very preliminary dSTORM results hinting at clustering of Rv1747 are certainly tantalizing. While this lets me speculate that phase separation is a mechanism of clustering of Rv1747 *in vivo*, further studies are needed to confirm this. Critical control studies with a Δ Rv1747 *Mtb* strain are on-going. However, the pathogenic nature of *Mtb* and its notoriously slow growth in culture make this seemingly simple test unexpectedly challenging. If validated, important future experiments also involve probing the impact of phosphorylation on cluster size using *Mtb* strains lacking or over-expressing PknF and other STPKs. The effect of over-expression of Rv1747, both in its wild-type form and with mutations of the phospho-acceptor threonines and FHAn domains, will also be important to study in this regard. Physical properties of the clusters *in vivo* are hard to probe in the bacilli due to its small size and the requirement for fixing before a more complex imaging process due to safety constraints.

One more convenient system to study due to its non-pathogenicity and much faster growth is *Mycobacterium smegmatis*. This bacterium has close orthologs to Rv1747 and the SPTKs of *Mtb*. The regulatory module of this Rv1747 ortholog has two FHA domains that overlap closely in sequence with those of Rv1747 (> 98 % identity), connected by a slightly

longer and less conserved ID linker (~ 150 vs ~ 100 residues). After initial experiments probing the phosphorylation-dependent phase separation of this system, the *Mycobacterium smegmatis* transporter might be used as a proxy for Rv1747 and its ease of genetic manipulation might enable further in-vivo studies that will be required.

A thorough characterization of the regulatory module of the Rv1747 ABC transporter in *Mtb* might inspire future strategies for the therapeutic targeting of this virulence factor. For example, disrupting Rv1747 clustering may inhibit its function. Fortunately, phase separation of the regulatory module of Rv1747 is a macroscopic phenotype that can be easily accessed *in vitro* and assayed with low-magnification fluorescence microscopy combined with simple automated image processing. This would enable high-throughput screening to identify and optimize molecules that disrupt phase separation.

As phase separation of biological macromolecules is understood better, the phenomenon can be harnessed for its material properties and have potential industrial applications as e.g. micro-reactors in the *in vitro* assembly of complex biosynthetic pathways. Model systems of different properties will be required to achieve such goals. Rv1747 is an interesting candidate with phosphorylation and de-phosphorylation able to act as a switch between mixed and demixed states at low concentrations of the protein. Such behaviour would allow the quick assembly/disassembly of the condensates as needed for selected applications.

Bibliography

1. V. S. Daniel, T. M. Daniel, Old Testament biblical references to tuberculosis. *Clin Infect Dis* **29**, 1557-1558 (1999).
2. J. Kappelman *et al.*, First Homo erectus from Turkey and implications for migrations into temperate Eurasia. *Am J Phys Anthropol* **135**, 110-116 (2008).
3. B. M. Rothschild *et al.*, Mycobacterium tuberculosis complex DNA from an extinct bison dated 17,000 years before the present. *Clin Infect Dis* **33**, 305-311 (2001).
4. O. Y. Lee *et al.*, Mycobacterium tuberculosis complex lipid virulence factors preserved in the 17,000-year-old skeleton of an extinct bison, Bison antiquus. *PLoS One* **7**, e41923 (2012).
5. I. Hershkovitz *et al.*, Detection and molecular characterization of 9,000-year-old Mycobacterium tuberculosis from a Neolithic settlement in the Eastern Mediterranean. *PLoS One* **3**, e3426 (2008).
6. J. A. Villemin, *Études sur la tuberculose: preuves rationnelles et expérimentales de sa spécificité et de son inoculabilité.* (J.-B. Baillière et fils, 1868).
7. J. A. Villemin, Cause et nature de la tuberculose. *Bull Acad Med* **37**, 211-216 (1865).
8. P. S. Sledzik, N. Bellantoni, Brief communication: bioarcheological and biocultural evidence for the New England vampire folk belief. *Am J Phys Anthropol* **94**, 269-274 (1994).
9. A. Tucker, The Great New England Vampire Panic. *Smithsonian Magazine*, (2012).
10. R. Koch, Die Aetiologie der Tuberkulose. *Klinische Wochenschrift* **11**, 490-492 (1932).
11. C. Mantoux, L'intradermo-reaction à la tuberculine. *La Presse médicale* **2**, 10-13 (1910).
12. L. C. A. Calmette, J.-M. C. Guérin, B. Weill-Hallé, Essai d'immunisation contre l'infection tuberculeuse. *Bull Acad Med* **91**, 787-796 (1924).
13. L. C. A. Calmette, J.-M. C. Guérin, Nouvelles recherches expérimentales sur la vaccination des bovidés contre la tuberculose. *Ann Inst Pasteur (Paris)* **34**, 553-560 (1920).
14. R. Brosch *et al.*, A new evolutionary scenario for the Mycobacterium tuberculosis complex. *Proc Natl Acad Sci U S A* **99**, 3684-3689 (2002).

15. J. D. Aronson, C. F. Aronson, H. C. Taylor, A twenty-year appraisal of BCG vaccination in the control of tuberculosis. *AMA Arch Intern Med* **101**, 881-893 (1958).
16. Who, "Global tuberculosis report 2017," (2017).
17. J. A. Philips, J. D. Ernst, Tuberculosis pathogenesis and immunity. *Annu Rev Pathol* **7**, 353-384 (2012).
18. L. Phillips, Infectious disease: TB's revenge. *Nature* **493**, 14-16 (2013).
19. A. H. Diacon *et al.*, The diarylquinoline TMC207 for multidrug-resistant tuberculosis. *N Engl J Med* **360**, 2397-2405 (2009).
20. A. J. Wolf *et al.*, Mycobacterium tuberculosis infects dendritic cells with high frequency and impairs their function in vivo. *Journal of Immunology* **179**, 2509-2519 (2007).
21. J. A. Armstrong, P. D. Hart, Phagosome-lysosome interactions in cultured macrophages infected with virulent tubercle bacilli. Reversal of the usual nonfusion pattern and observations on bacterial survival. *J Exp Med* **142**, 1-16 (1975).
22. P. B. Kang *et al.*, The human macrophage mannose receptor directs Mycobacterium tuberculosis lipoarabinomannan-mediated phagosome biogenesis. *J Exp Med* **202**, 987-999 (2005).
23. J. A. Armstrong, P. D. Hart, Response of cultured macrophages to Mycobacterium tuberculosis, with observations on fusion of lysosomes with phagosomes. *J Exp Med* **134**, 713-740 (1971).
24. D. G. Russell, Mycobacterium tuberculosis: here today, and here tomorrow. *Nat Rev Mol Cell Biol* **2**, 569-577 (2001).
25. J. A. Philips, Mycobacterial manipulation of vacuolar sorting. *Cell Microbiol* **10**, 2408-2415 (2008).
26. H. Bach, K. G. Papavinasundaram, D. Wong, Z. Hmama, Y. Av-Gay, Mycobacterium tuberculosis virulence is mediated by PtpA dephosphorylation of human vacuolar protein sorting 33B. *Cell Host Microbe* **3**, 316-322 (2008).
27. J. Sun *et al.*, Mycobacterial nucleoside diphosphate kinase blocks phagosome maturation in murine RAW 264.7 macrophages. *PLoS One* **5**, e8769 (2010).
28. O. V. Vieira *et al.*, Acquisition of Hrs, an essential component of phagosomal maturation, is impaired by mycobacteria. *Molecular and Cellular Biology* **24**, 4593-4604 (2004).

29. G. E. Purdy, R. M. Owens, L. Bennett, D. G. Russell, B. A. Butcher, Kinetics of phosphatidylinositol-3-phosphate acquisition differ between IgG bead-containing phagosomes and Mycobacterium tuberculosis-containing phagosomes. *Cell Microbiol* **7**, 1627-1634 (2005).
30. I. Vergne *et al.*, Mechanism of phagolysosome biogenesis block by viable Mycobacterium tuberculosis. *Proc Natl Acad Sci U S A* **102**, 4033-4038 (2005).
31. O. H. Vandal, C. F. Nathan, S. Ehrt, Acid resistance in Mycobacterium tuberculosis. *J Bacteriol* **191**, 4714-4721 (2009).
32. A. S. Pym, P. Brodin, R. Brosch, M. Huerre, S. T. Cole, Loss of RD1 contributed to the attenuation of the live tuberculosis vaccines Mycobacterium bovis BCG and Mycobacterium microti. *Mol Microbiol* **46**, 709-717 (2002).
33. T. Hsu *et al.*, The primary mechanism of attenuation of bacillus Calmette-Guerin is a loss of secreted lytic function required for invasion of lung interstitial tissue. *Proc Natl Acad Sci U S A* **100**, 12420-12425 (2003).
34. K. M. Guinn *et al.*, Individual RD1-region genes are required for export of ESAT-6/CFP-10 and for virulence of Mycobacterium tuberculosis. *Mol Microbiol* **51**, 359-370 (2004).
35. N. van der Wel *et al.*, M. tuberculosis and M. leprae translocate from the phagolysosome to the cytosol in myeloid cells. *Cell* **129**, 1287-1298 (2007).
36. J. Smith *et al.*, Evidence for pore formation in host cell membranes by ESX-1-secreted ESAT-6 and its role in Mycobacterium marinum escape from the vacuole. *Infect Immun* **76**, 5478-5487 (2008).
37. L. Y. Gao *et al.*, A mycobacterial virulence gene cluster extending RD1 is required for cytolysis, bacterial spreading and ESAT-6 secretion. *Mol Microbiol* **53**, 1677-1693 (2004).
38. M. I. de Jonge *et al.*, ESAT-6 from Mycobacterium tuberculosis dissociates from its putative chaperone CFP-10 under acidic conditions and exhibits membrane-lysing activity. *J Bacteriol* **189**, 6028-6034 (2007).
39. B. B. Mishra *et al.*, Mycobacterium tuberculosis protein ESAT-6 is a potent activator of the NLRP3/ASC inflammasome. *Cell Microbiol* **12**, 1046-1063 (2010).
40. I. C. Koo *et al.*, ESX-1-dependent cytolysis in lysosome secretion and inflammasome activation during mycobacterial infection. *Cell Microbiol* **10**, 1866-1878 (2008).

41. S. A. Stanley, J. E. Johndrow, P. Manzanillo, J. S. Cox, The Type I IFN response to infection with *Mycobacterium tuberculosis* requires ESX-1-mediated secretion and contributes to pathogenesis. *J Immunol* **178**, 3143-3152 (2007).
42. B. Levine, N. Mizushima, H. W. Virgin, Autophagy in immunity and inflammation. *Nature* **469**, 323-335 (2011).
43. M. A. Delgado, R. A. Elmaoued, A. S. Davis, G. Kyei, V. Deretic, Toll-like receptors control autophagy. *EMBO J* **27**, 1110-1121 (2008).
44. Y. Xu *et al.*, Toll-like receptor 4 is a sensor for autophagy associated with innate immunity. *Immunity* **27**, 135-144 (2007).
45. J. Harris *et al.*, T helper 2 cytokines inhibit autophagic control of intracellular *Mycobacterium tuberculosis*. *Immunity* **27**, 505-517 (2007).
46. J. M. Yuk *et al.*, Vitamin D3 induces autophagy in human monocytes/macrophages via cathelicidin. *Cell Host Microbe* **6**, 231-243 (2009).
47. M. G. Gutierrez *et al.*, Autophagy is a defense mechanism inhibiting BCG and *Mycobacterium tuberculosis* survival in infected macrophages. *Cell* **119**, 753-766 (2004).
48. D. Kumar *et al.*, Genome-wide analysis of the host intracellular network that regulates survival of *Mycobacterium tuberculosis*. *Cell* **140**, 731-743 (2010).
49. H. Clay, H. E. Volkman, L. Ramakrishnan, Tumor necrosis factor signaling mediates resistance to mycobacteria by inhibiting bacterial growth and macrophage death. *Immunity* **29**, 283-294 (2008).
50. D. M. Tobin *et al.*, The *lta4h* locus modulates susceptibility to mycobacterial infection in zebrafish and humans. *Cell* **140**, 717-730 (2010).
51. L. G. Bekker *et al.*, Immunopathologic effects of tumor necrosis factor alpha in murine mycobacterial infection are dose dependent. *Infect Immun* **68**, 6954-6961 (2000).
52. N. Agarwal, G. Lamichhane, R. Gupta, S. Nolan, W. R. Bishai, Cyclic AMP intoxication of macrophages by a *Mycobacterium tuberculosis* adenylate cyclase. *Nature* **460**, 98-102 (2009).
53. H. E. Volkman *et al.*, Tuberculous granuloma induction via interaction of a bacterial secreted protein with host epithelium. *Science* **327**, 466-469 (2010).
54. J. M. Davis, L. Ramakrishnan, The role of the granuloma in expansion and dissemination of early tuberculous infection. *Cell* **136**, 37-49 (2009).

55. D. O. Adams, The granulomatous inflammatory response. A review. *Am J Pathol* **84**, 164-192 (1976).
56. M. P. Puissegur *et al.*, An in vitro dual model of mycobacterial granulomas to investigate the molecular interactions between mycobacteria and human host cells. *Cell Microbiol* **6**, 423-433 (2004).
57. J. M. Davis *et al.*, Real-time visualization of mycobacterium-macrophage interactions leading to initiation of granuloma formation in zebrafish embryos. *Immunity* **17**, 693-702 (2002).
58. L. Tailleux *et al.*, DC-SIGN is the major Mycobacterium tuberculosis receptor on human dendritic cells. *J Exp Med* **197**, 121-127 (2003).
59. M. C. Tsai *et al.*, Characterization of the tuberculous granuloma in murine and human lungs: cellular composition and relative tissue oxygen tension. *Cell Microbiol* **8**, 218-232 (2006).
60. P. S. Randhawa, Lymphocyte subsets in granulomas of human tuberculosis: an in situ immunofluorescence study using monoclonal antibodies. *Pathology* **22**, 153-155 (1990).
61. E. J. Capra, M. T. Laub, Evolution of two-component signal transduction systems. *Annu Rev Microbiol* **66**, 325-347 (2012).
62. A. H. West, A. M. Stock, Histidine kinases and response regulator proteins in two-component signaling systems. *Trends Biochem Sci* **26**, 369-376 (2001).
63. T. Parish, Two-Component Regulatory Systems of Mycobacteria. *Microbiol Spectr* **2**, MGM2-0010-2013 (2014).
64. L. E. Ulrich, E. V. Koonin, I. B. Zhulin, One-component systems dominate signal transduction in prokaryotes. *Trends Microbiol* **13**, 52-56 (2005).
65. R. Manganelli, Sigma Factors: Key Molecules in Mycobacterium tuberculosis Physiology and Virulence. *Microbiol Spectr* **2**, MGM2-0007-2013 (2014).
66. A. Staron *et al.*, The third pillar of bacterial signal transduction: classification of the extracytoplasmic function (ECF) sigma factor protein family. *Mol Microbiol* **74**, 557-581 (2009).
67. G. Manning, D. B. Whyte, R. Martinez, T. Hunter, S. Sudarsanam, The protein kinase complement of the human genome. *Science* **298**, 1912-1934 (2002).

68. S. Braconi Quintaje, S. Orchard, The annotation of both human and mouse kinomes in UniProtKB/Swiss-Prot: one small step in manual annotation, one giant leap for full comprehension of genomes. *Mol Cell Proteomics* **7**, 1409-1419 (2008).
69. S. Prisic, R. N. Husson, Mycobacterium tuberculosis Serine/Threonine Protein Kinases. *Microbiol Spectr* **2**, 1-26 (2014).
70. A. Wehenkel *et al.*, Mycobacterial Ser/Thr protein kinases and phosphatases: physiological roles and therapeutic potential. *Biochim Biophys Acta* **1784**, 193-202 (2008).
71. S. K. Hanks, T. Hunter, Protein kinases 6. The eukaryotic protein kinase superfamily: kinase (catalytic) domain structure and classification. *FASEB J* **9**, 576-596 (1995).
72. F. Sievers *et al.*, Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol Syst Biol* **7**, 539 (2011).
73. T. Wagner *et al.*, The crystal structure of the catalytic domain of the ser/thr kinase PknA from *M. tuberculosis* shows an Src-like autoinhibited conformation. *Proteins* **83**, 982-988 (2015).
74. S. K. Ravala *et al.*, Evidence that phosphorylation of threonine in the GT motif triggers activation of PknA, a eukaryotic-type serine/threonine kinase from *Mycobacterium tuberculosis*. *FEBS J* **282**, 1419-1431 (2015).
75. A. Wehenkel *et al.*, The structure of PknB in complex with mitoxantrone, an ATP-competitive inhibitor, suggests a mode of protein kinase regulation in mycobacteria. *FEBS Lett* **580**, 3018-3022 (2006).
76. M. Ortiz-Lombardia, F. Pompeo, B. Boitel, P. M. Alzari, Crystal structure of the catalytic domain of the PknB serine/threonine kinase from *Mycobacterium tuberculosis*. *J Biol Chem* **278**, 13094-13100 (2003).
77. T. A. Young, B. Delagoutte, J. A. Endrizzi, A. M. Falick, T. Alber, Structure of *Mycobacterium tuberculosis* PknB supports a universal activation mechanism for Ser/Thr protein kinases. *Nat Struct Biol* **10**, 168-174 (2003).
78. T. N. Lombana *et al.*, Allosteric activation mechanism of the *Mycobacterium tuberculosis* receptor Ser/Thr protein kinase, PknB. *Structure* **18**, 1667-1677 (2010).
79. L. M. Gay, H. L. Ng, T. Alber, A conserved dimer and global conformational changes in the structure of apo-PknE Ser/Thr protein kinase from *Mycobacterium tuberculosis*. *J Mol Biol* **360**, 409-420 (2006).

80. N. Scherr *et al.*, Structural basis for the specific inhibition of protein kinase G, a virulence factor of *Mycobacterium tuberculosis*. *Proc Natl Acad Sci U S A* **104**, 12151-12156 (2007).
81. M. N. Lisa *et al.*, Molecular Basis of the Activity and the Regulation of the Eukaryotic-like S/T Protein Kinase PknG from *Mycobacterium tuberculosis*. *Structure* **23**, 1039-1048 (2015).
82. M. N. Lisa *et al.*, The crystal structure of PknI from *Mycobacterium tuberculosis* shows an inactive, pseudokinase-like conformation. *FEBS J* **284**, 602-614 (2017).
83. R. Duran *et al.*, Conserved autophosphorylation pattern in activation loops and juxtamembrane regions of *Mycobacterium tuberculosis* Ser/Thr protein kinases. *Biochem Biophys Res Commun* **333**, 858-867 (2005).
84. H. M. O'Hare *et al.*, Regulation of glutamate metabolism by protein kinases in mycobacteria. *Mol Microbiol* **70**, 1408-1423 (2008).
85. V. Molle *et al.*, Characterization of the phosphorylation sites of *Mycobacterium tuberculosis* serine/threonine protein kinases, PknA, PknD, PknE, and PknH by mass spectrometry. *Proteomics* **6**, 3754-3766 (2006).
86. C. M. Kang *et al.*, The *Mycobacterium tuberculosis* serine/threonine kinases PknA and PknB: substrate identification and regulation of cell shape. *Genes Dev* **19**, 1692-1704 (2005).
87. S. Pristic *et al.*, Extensive phosphorylation with overlapping specificity by *Mycobacterium tuberculosis* serine/threonine protein kinases. *Proc Natl Acad Sci U S A* **107**, 7521-7526 (2010).
88. M. Gil *et al.*, Inhibition of *Mycobacterium tuberculosis* PknG by non-catalytic rubredoxin domain specific modification: reaction of an electrophilic nitro-fatty acid with the Fe-S center. *Free Radic Biol Med* **65**, 150-161 (2013).
89. D. Tiwari *et al.*, Key residues in *Mycobacterium tuberculosis* protein kinase G play a role in regulating kinase activity and survival in the host. *J Biol Chem* **284**, 27467-27479 (2009).
90. C. Mieczkowski, A. T. Iavarone, T. Alber, Auto-activation mechanism of the *Mycobacterium tuberculosis* PknB receptor Ser/Thr kinase. *EMBO J* **27**, 3186-3197 (2008).
91. A. E. Greenstein, N. Echols, T. N. Lombana, D. S. King, T. Alber, Allosteric activation by dimerization of the PknD receptor Ser/Thr protein kinase from *Mycobacterium tuberculosis*. *J Biol Chem* **282**, 11427-11435 (2007).

92. M. Dey *et al.*, Mechanistic link between PKR dimerization, autophosphorylation, and eIF2alpha substrate recognition. *Cell* **122**, 901-913 (2005).
93. C. E. Baer, A. T. Iavarone, T. Alber, C. M. Sassetti, Biochemical and spatial coincidence in the provisional Ser/Thr protein kinase interaction network of *Mycobacterium tuberculosis*. *J Biol Chem* **289**, 20422-20433 (2014).
94. J. M. Curry *et al.*, An ABC transporter containing a forkhead-associated domain interacts with a serine-threonine protein kinase and is required for growth of *Mycobacterium tuberculosis* in mice. *Infect Immun* **73**, 4471-4477 (2005).
95. C. M. Sassetti, D. H. Boyd, E. J. Rubin, Genes required for mycobacterial growth defined by high density mutagenesis. *Mol Microbiol* **48**, 77-84 (2003).
96. C. L. Gee *et al.*, A phosphorylated pseudokinase complex controls cell wall synthesis in mycobacteria. *Sci Signal* **5**, ra7 (2012).
97. A. Dasgupta, P. Datta, M. Kundu, J. Basu, The serine/threonine kinase PknB of *Mycobacterium tuberculosis* phosphorylates PBPA, a penicillin-binding protein required for cell division. *Microbiology* **152**, 493-504 (2006).
98. M. Mir *et al.*, The extracytoplasmic domain of the *Mycobacterium tuberculosis* Ser/Thr kinase PknB binds specific muropeptides and is required for PknB localization. *PLoS Pathog* **7**, e1002182 (2011).
99. J. C. Camus, M. J. Pryor, C. Medigue, S. T. Cole, Re-annotation of the genome sequence of *Mycobacterium tuberculosis* H37Rv. *Microbiology* **148**, 2967-2973 (2002).
100. F. Vanzembergh *et al.*, Effect of PstS sub-units or PknD deficiency on the survival of *Mycobacterium tuberculosis*. *Tuberculosis (Edinb)* **90**, 338-345 (2010).
101. N. A. Be, W. R. Bishai, S. K. Jain, Role of *Mycobacterium tuberculosis* pknD in the pathogenesis of central nervous system tuberculosis. *BMC Microbiol* **12**, 7 (2012).
102. V. Molle, A. K. Brown, G. S. Besra, A. J. Cozzone, L. Kremer, The condensing activities of the *Mycobacterium tuberculosis* type II fatty acid synthase are differentially regulated by phosphorylation. *J Biol Chem* **281**, 30094-30103 (2006).
103. D. Kumar, S. Narayanan, pknE, a serine/threonine kinase of *Mycobacterium tuberculosis* modulates multiple apoptotic paradigms. *Infect Genet Evol* **12**, 737-747 (2012).
104. V. L. Spivey, R. H. Whalan, E. M. Hirst, S. J. Smerdon, R. S. Buxton, An attenuated mutant of the Rv1747 ATP-binding cassette transporter of *Mycobacterium tuberculosis* and a mutant of its cognate kinase, PknF, show increased expression of the efflux pump-related iniBAC operon. *FEMS Microbiol Lett* **347**, 107-115 (2013).

105. L. N. Glass *et al.*, Mycobacterium tuberculosis universal stress protein Rv2623 interacts with the putative ATP binding cassette (ABC) transporter Rv1747 to regulate mycobacterial growth. *PLoS Pathog* **13**, e1006515 (2017).
106. V. Molle *et al.*, An FHA phosphoprotein recognition domain mediates protein EmbR phosphorylation by PknH, a Ser/Thr protein kinase from Mycobacterium tuberculosis. *Biochemistry* **42**, 15300-15309 (2003).
107. K. Sharma *et al.*, Transcriptional control of the mycobacterial embCAB operon by PknH through a regulatory protein, EmbR, in vivo. *J Bacteriol* **188**, 2936-2944 (2006).
108. V. E. Escuyer *et al.*, The role of the embA and embB gene products in the biosynthesis of the terminal hexaarabinofuranosyl motif of Mycobacterium smegmatis arabinogalactan. *J Biol Chem* **276**, 48854-48862 (2001).
109. N. Zhang *et al.*, The Emb proteins of mycobacteria direct arabinosylation of lipoarabinomannan and arabinogalactan via an N-terminal recognition region and a C-terminal synthetic region. *Mol Microbiol* **50**, 69-76 (2003).
110. K. G. Papavinasasundaram *et al.*, Deletion of the Mycobacterium tuberculosis pknH gene confers a higher bacillary load during the chronic phase of infection in BALB/c mice. *J Bacteriol* **187**, 5751-5760 (2005).
111. J. D. Chao *et al.*, Convergence of Ser/Thr and two-component signaling to coordinate expression of the dormancy regulon in Mycobacterium tuberculosis. *J Biol Chem* **285**, 29239-29246 (2010).
112. A. Cavazos, D. M. Prigozhin, T. Alber, Structure of the sensor domain of Mycobacterium tuberculosis PknH receptor kinase reveals a conserved binding cleft. *J Mol Biol* **422**, 488-494 (2012).
113. R. Gopaldaswamy, S. Narayanan, B. Chen, W. R. Jacobs, Y. Av-Gay, The serine/threonine protein kinase PknI controls the growth of Mycobacterium tuberculosis upon infection. *FEMS Microbiol Lett* **295**, 23-29 (2009).
114. J. Jang *et al.*, Functional characterization of the Mycobacterium tuberculosis serine/threonine kinase PknJ. *Microbiology* **156**, 1619-1631 (2010).
115. M. J. Canova *et al.*, The Mycobacterium tuberculosis serine/threonine kinase PknL phosphorylates Rv2175c: mass spectrometric profiling of the activation loop phosphorylation sites and their role in the recruitment of Rv2175c. *Proteomics* **8**, 521-533 (2008).

116. S. Cowley *et al.*, The Mycobacterium tuberculosis protein serine/threonine kinase PknG is linked to cellular glutamate/glutamine levels and is important for growth in vivo. *Mol Microbiol* **52**, 1691-1702 (2004).
117. A. Walburger *et al.*, Protein kinase G from pathogenic mycobacteria promotes survival within macrophages. *Science* **304**, 1800-1804 (2004).
118. T. J. Nott *et al.*, An intramolecular switch regulates phosphoindependent FHA domain interactions in Mycobacterium tuberculosis. *Sci Signal* **2**, ra12 (2009).
119. V. Malhotra, B. P. Okon, J. E. Clark-Curtiss, Mycobacterium tuberculosis protein kinase K enables growth adaptation through translation control. *J Bacteriol* **194**, 4184-4196 (2012).
120. V. Malhotra, L. T. Arteaga-Cortes, G. Clay, J. E. Clark-Curtiss, Mycobacterium tuberculosis protein kinase K confers survival advantage during early infection in mice and regulates growth in culture and during persistent infection: implications for immune modulation. *Microbiology* **156**, 2829-2841 (2010).
121. P. Kumar *et al.*, The Mycobacterium tuberculosis protein kinase K modulates activation of transcription from the promoter of mycobacterial monooxygenase operon through phosphorylation of the transcriptional regulator VirS. *J Biol Chem* **284**, 11090-11099 (2009).
122. C. Ortega *et al.*, Mycobacterium tuberculosis Ser/Thr protein kinase B mediates an oxygen-dependent replication switch. *PLoS Biol* **12**, e1001746 (2014).
123. P. Barthe, G. V. Mukamolova, C. Roumestand, M. Cohen-Gonsaud, The structure of PknB extracellular PASTA domain from mycobacterium tuberculosis suggests a ligand-dependent kinase activation. *Structure* **18**, 606-615 (2010).
124. M. C. Good, A. E. Greenstein, T. A. Young, H. L. Ng, T. Alber, Sensor domain of the Mycobacterium tuberculosis receptor Ser/Thr protein kinase, PknD, forms a highly symmetric beta propeller. *J Mol Biol* **339**, 459-469 (2004).
125. A. K. Sharma *et al.*, Serine/Threonine Protein Phosphatase PstP of Mycobacterium tuberculosis Is Necessary for Accurate Cell Division and Survival of Pathogen. *J Biol Chem* **291**, 24215-24230 (2016).
126. P. Chopra *et al.*, Phosphoprotein phosphatase of Mycobacterium tuberculosis dephosphorylates serine-threonine kinases PknA and PknB. *Biochem Biophys Res Commun* **311**, 112-120 (2003).
127. K. E. Pullen *et al.*, An alternate conformation and a third metal in PstP/Ppp, the M. tuberculosis PP2C-Family Ser/Thr protein phosphatase. *Structure* **12**, 1947-1954 (2004).

128. A. Wehenkel, M. Bellinzoni, F. Schaeffer, A. Villarino, P. M. Alzari, Structural and binding studies of the three-metal center in two mycobacterial PPM Ser/Thr protein phosphatases. *J Mol Biol* **374**, 890-898 (2007).
129. A. Sajid *et al.*, Phosphorylation of Mycobacterium tuberculosis Ser/Thr phosphatase by PknA and PknB. *PLoS One* **6**, e17871 (2011).
130. D. Wong *et al.*, Protein tyrosine kinase, PtkA, is required for Mycobacterium tuberculosis growth in macrophages. *Sci Rep* **8**, 155 (2018).
131. A. Koul *et al.*, Cloning and characterization of secretory tyrosine phosphatases of Mycobacterium tuberculosis. *J Bacteriol* **182**, 5425-5432 (2000).
132. D. Wong, H. Bach, J. Sun, Z. Hmama, Y. Av-Gay, Mycobacterium tuberculosis protein tyrosine phosphatase (PtpA) excludes host vacuolar-H⁺-ATPase to inhibit phagosome acidification. *Proc Natl Acad Sci U S A* **108**, 19371-19376 (2011).
133. H. Bach, D. Wong, Y. Av-Gay, Mycobacterium tuberculosis PtkA is a novel protein tyrosine kinase whose substrate is PtpA. *Biochem J* **420**, 155-160 (2009).
134. P. Zhou, W. Li, D. Wong, J. Xie, Y. Av-Gay, Phosphorylation control of protein tyrosine phosphatase A activity in Mycobacterium tuberculosis. *FEBS Lett* **589**, 326-331 (2015).
135. A. W. Almawi, L. A. Matthews, A. Guarne, FHA domains: Phosphopeptide binding and beyond. *Prog Biophys Mol Biol* **127**, 105-110 (2017).
136. X. Liang, S. R. Van Doren, Mechanistic insights into phosphoprotein-binding FHA domains. *Acc Chem Res* **41**, 991-999 (2008).
137. A. Mahajan *et al.*, Structure and function of the phosphothreonine-specific FHA domain. *Sci Signal* **1**, re12 (2008).
138. M. Pallen, R. Chaudhuri, A. Khan, Bacterial FHA domains: neglected players in the phospho-threonine signalling game? *Trends Microbiol* **10**, 556-563 (2002).
139. S. Pennell *et al.*, Structural and functional analysis of phosphothreonine-dependent FHA domain interactions. *Structure* **18**, 1587-1595 (2010).
140. L. J. Alderwick *et al.*, Molecular structure of EmbR, a response element of Ser/Thr kinase signaling in Mycobacterium tuberculosis. *Proc Natl Acad Sci U S A* **103**, 2558-2563 (2006).
141. P. Fernandez *et al.*, The Ser/Thr protein kinase PknB is essential for sustaining mycobacterial growth. *J Bacteriol* **188**, 7778-7784 (2006).

142. C. Roumestand *et al.*, Structural insight into the Mycobacterium tuberculosis Rv0020c protein and its interaction with the PknB kinase. *Structure* **19**, 1525-1534 (2011).
143. V. Molle *et al.*, Two FHA domains on an ABC transporter, Rv1747, mediate its phosphorylation by PknF, a Ser/Thr protein kinase from Mycobacterium tuberculosis. *FEMS Microbiol Lett* **234**, 215-223 (2004).
144. V. L. Spivey *et al.*, Forkhead-associated (FHA) domain containing ABC transporter Rv1747 is positively regulated by Ser/Thr phosphorylation in Mycobacterium tuberculosis. *J Biol Chem* **286**, 26198-26209 (2011).
145. C. Grundner, L. M. Gay, T. Alber, Mycobacterium tuberculosis serine/threonine kinases PknB, PknD, PknE, and PknF phosphorylate multiple FHA domains. *Protein Sci* **14**, 1918-1921 (2005).
146. T. Schwarz-Romond, C. Merrifield, B. J. Nichols, M. Bienz, The Wnt signalling effector Dishevelled forms dynamic protein assemblies rather than stable associations with cytoplasmic vesicles. *J Cell Sci* **118**, 5269-5277 (2005).
147. J. R. Miller *et al.*, Establishment of the dorsal-ventral axis in Xenopus embryos coincides with the dorsal enrichment of dishevelled that is dependent on cortical rotation. *J Cell Biol* **146**, 427-437 (1999).
148. R. Bernardi, P. P. Pandolfi, Structure, dynamics and functions of promyelocytic leukaemia nuclear bodies. *Nat Rev Mol Cell Biol* **8**, 1006-1016 (2007).
149. M. Lang *et al.*, Three-dimensional organization of promyelocytic leukemia nuclear bodies. *J Cell Sci* **123**, 392-400 (2010).
150. A. Hernandez-Vega *et al.*, Local Nucleation of Microtubule Bundles through Tubulin Concentration into a Condensed Tau Phase. *Cell Rep* **20**, 2304-2312 (2017).
151. S. Ambadipudi, J. Biernat, D. Riedel, E. Mandelkow, M. Zweckstetter, Liquid-liquid phase separation of the microtubule-binding repeats of the Alzheimer-related protein Tau. *Nat Commun* **8**, 275 (2017).
152. P. Li *et al.*, Phase transitions in the assembly of multivalent signalling proteins. *Nature* **483**, 336-340 (2012).
153. X. Su *et al.*, Phase separation of signaling molecules promotes T cell receptor signal transduction. *Science* **352**, 595-599 (2016).
154. M. Feric, C. P. Brangwynne, A nuclear F-actin scaffold stabilizes ribonucleoprotein droplets against gravity in large cells. *Nat Cell Biol* **15**, 1253-1259 (2013).

155. D. Updike, S. Strome, P granule assembly and function in *Caenorhabditis elegans* germ cells. *J Androl* **31**, 53-60 (2010).
156. S. Saha *et al.*, Polar Positioning of Phase-Separated Liquid Compartments in Cells Regulated by an mRNA Competition Mechanism. *Cell* **166**, 1572-1584 e1516 (2016).
157. N. Kotaja, P. Sassone-Corsi, The chromatoid body: a germ-cell-specific RNA-processing centre. *Nat Rev Mol Cell Biol* **8**, 85-90 (2007).
158. C. P. Brangwynne *et al.*, Germline P granules are liquid droplets that localize by controlled dissolution/condensation. *Science* **324**, 1729-1732 (2009).
159. S. Elbaum-Garfinkle *et al.*, The disordered P granule protein LAF-1 drives phase separation into droplets with tunable viscosity and dynamics. *Proc Natl Acad Sci U S A* **112**, 7189-7194 (2015).
160. J. Smith *et al.*, Spatial patterning of P granules by RNA-induced phase separation of the intrinsically-disordered protein MEG-3. *Elife* **5**, 1-18 (2016).
161. L. Liang, W. Diehljones, P. Lasko, Localization of Vasa Protein to the Drosophila Pole Plasm Is Independent of Its Rna-Binding and Helicase Activities. *Development* **120**, 1201-1211 (1994).
162. O. Meikar, M. Da Ros, H. Korhonen, N. Kotaja, Chromatoid body and small RNAs in male germ cells. *Reproduction* **142**, 195-209 (2011).
163. A. Molliex *et al.*, Phase separation by low complexity domains promotes stress granule assembly and drives pathological fibrillization. *Cell* **163**, 123-133 (2015).
164. L. Nover, K. D. Scharf, D. Neumann, Cytoplasmic heat shock granules are formed from precursor particles and are associated with a specific set of mRNAs. *Mol Cell Biol* **9**, 1298-1308 (1989).
165. P. Anderson, N. Kedersha, Stress granules. *Curr Biol* **19**, R397-398 (2009).
166. J. R. Buchan, R. Parker, Eukaryotic stress granules: the ins and outs of translation. *Mol Cell* **36**, 932-941 (2009).
167. H. Tourriere *et al.*, The RasGAP-associated endoribonuclease G3BP assembles stress granules. *J Cell Biol* **160**, 823-831 (2003).
168. A. Wilczynska, C. Aigueperse, M. Kress, F. Dautry, D. Weil, The translational regulator CPEB1 provides a link between dcp1 bodies and stress granules. *J Cell Sci* **118**, 981-992 (2005).

169. N. Gilks *et al.*, Stress granule assembly is mediated by prion-like aggregation of TIA-1. *Mol Biol Cell* **15**, 5383-5398 (2004).
170. A. Patel *et al.*, A Liquid-to-Solid Phase Transition of the ALS Protein FUS Accelerated by Disease Mutation. *Cell* **162**, 1066-1077 (2015).
171. K. A. Burke, A. M. Janke, C. L. Rhine, N. L. Fawzi, Residue-by-Residue View of In Vitro FUS Granules that Bind the C-Terminal Domain of RNA Polymerase II. *Mol Cell* **60**, 231-241 (2015).
172. T. Murakami *et al.*, ALS/FTD Mutation-Induced Phase Transition of FUS Liquid Droplets and Reversible Hydrogels into Irreversible Hydrogels Impairs RNP Granule Function. *Neuron* **88**, 678-690 (2015).
173. M. Kato *et al.*, Cell-free formation of RNA granules: low complexity sequence domains form dynamic fibers within hydrogels. *Cell* **149**, 753-767 (2012).
174. S. C. Weber, C. P. Brangwynne, Getting RNA and protein in phase. *Cell* **149**, 1188-1191 (2012).
175. S. Xiang *et al.*, The LC Domain of hnRNPA2 Adopts Similar Conformations in Hydrogel Polymers, Liquid-like Droplets, and Nuclei. *Cell* **163**, 829-839 (2015).
176. Y. Lin, D. S. Protter, M. K. Rosen, R. Parker, Formation and Maturation of Phase-Separated Liquid Droplets by RNA-Binding Proteins. *Mol Cell* **60**, 208-219 (2015).
177. S. Jain *et al.*, ATPase-Modulated Stress Granules Contain a Diverse Proteome and Substructure. *Cell* **164**, 487-498 (2016).
178. J. S. Andersen *et al.*, Nucleolar proteome dynamics. *Nature* **433**, 77-83 (2005).
179. F. M. Boisvert, S. van Koningsbruggen, J. Navascues, A. I. Lamond, The multifunctional nucleolus. *Nat Rev Mol Cell Biol* **8**, 574-585 (2007).
180. A. Hubstenberger, S. L. Noble, C. Cameron, T. C. Evans, Translation repressors, an RNA helicase, and developmental cues control RNP phase transitions during early development. *Dev Cell* **27**, 161-173 (2013).
181. U. Sheth, J. Pitt, S. Dennis, J. R. Priess, Perinuclear P granules are the principal sites of mRNA export in adult *C. elegans* germ cells. *Development* **137**, 1305-1314 (2010).
182. J. T. Wang *et al.*, Regulation of RNA granule dynamics by phosphorylation of serine-rich, intrinsically disordered proteins in *C. elegans*. *Elife* **3**, e04591 (2014).

183. J. A. West *et al.*, Structural, super-resolution microscopy analysis of paraspeckle nuclear body organization. *J Cell Biol* **214**, 817-830 (2016).
184. M. Feric *et al.*, Coexisting Liquid Phases Underlie Nucleolar Subcompartments. *Cell* **165**, 1686-1697 (2016).
185. L. D. Zarzar *et al.*, Dynamically reconfigurable complex emulsions via tunable interfacial tensions. *Nature* **518**, 520-524 (2015).
186. J. Eckelmann, U. Lüning, Mixing Liquids—Mission Impossible? A Colorful Demonstration on Immiscible Systems. *Journal of Chemical Education* **90**, 224-227 (2012).
187. E. Seiradake *et al.*, Structurally encoded intraclass differences in EphA clusters drive distinct cell responses. *Nat Struct Mol Biol* **20**, 958-964 (2013).
188. F. Henkler *et al.*, The extracellular domains of FasL and Fas are sufficient for the formation of supramolecular FasL-Fas clusters of high stability. *J Cell Biol* **168**, 1087-1098 (2005).
189. A. B. Schreiber, T. A. Libermann, I. Lax, Y. Yarden, J. Schlessinger, Biological role of epidermal growth factor-receptor clustering. Investigation with monoclonal anti-receptor antibodies. *J Biol Chem* **258**, 846-853 (1983).
190. A. S. Yap, W. M. Briehner, M. Pruschy, B. M. Gumbiner, Lateral clustering of the adhesive ectodomain: a fundamental determinant of cadherin function. *Curr Biol* **7**, 308-315 (1997).
191. R. Varma, S. Mayor, GPI-anchored proteins are organized in submicron domains at the cell surface. *Nature* **394**, 798-801 (1998).
192. A. Grakoui *et al.*, The immunological synapse: a molecular machine controlling T cell activation. *Science* **285**, 221-227 (1999).
193. S. Banjade, M. K. Rosen, Phase transitions of multivalent proteins can promote clustering of membrane receptors. *Elife* **3**, e04123-e04123 (2014).
194. Y. Shin, C. P. Brangwynne, Liquid phase condensation in cell physiology and disease. *Science* **357**, (2017).
195. P. J. Flory, Thermodynamics of high polymer solutions. *Journal of Chemical Physics* **9**, 660-661 (1941).
196. M. L. Huggins, Some properties of solutions of long-chain compounds. *Journal of Physical Chemistry* **46**, 151-158 (1942).

197. C. P. Brangwynne, P. Tompa, R. V. Pappu, Polymer physics of intracellular phase transitions. *Nature Physics* **11**, 899-904 (2015).
198. F. Tanaka, *Polymer Physics: Applications to Molecular Association and Thermoreversible Gelation*. (Cambridge University Press, Cambridge, 2011).
199. S. Banjade *et al.*, Conserved interdomain linker promotes phase separation of the multivalent adaptor protein Nck. *Proc Natl Acad Sci U S A* **112**, E6426-6435 (2015).
200. D. S. W. Protter *et al.*, Intrinsically Disordered Regions Can Contribute Promiscuous Interactions to RNP Granule Assembly. *Cell Rep* **22**, 1401-1412 (2018).
201. S. F. Banani *et al.*, Compositional Control of Phase-Separated Cellular Bodies. *Cell* **166**, 651-663 (2016).
202. S. Vucetic, C. J. Brown, A. K. Dunker, Z. Obradovic, Flavors of protein disorder. *Proteins* **52**, 573-584 (2003).
203. P. Romero *et al.*, Sequence complexity of disordered protein. *Proteins* **42**, 38-48 (2001).
204. F. G. Quiroz, A. Chilkoti, Sequence heuristics to encode phase behaviour in intrinsically disordered protein polymers. *Nat Mater* **14**, 1164-1171 (2015).
205. R. M. Vernon *et al.*, Pi-Pi contacts are an overlooked protein feature relevant to phase separation. *Elife* **7**, (2018).
206. T. J. Nott *et al.*, Phase transition of a disordered nuage protein generates environmentally responsive membraneless organelles. *Mol Cell* **57**, 936-947 (2015).
207. C. W. Pak *et al.*, Sequence Determinants of Intracellular Phase Separation by Complex Coacervation of a Disordered Protein. *Mol Cell* **63**, 72-85 (2016).
208. R. K. Das, K. M. Ruff, R. V. Pappu, Relating sequence encoded information to form and function of intrinsically disordered proteins. *Curr Opin Struct Biol* **32**, 102-112 (2015).
209. A. S. Holehouse, R. K. Das, J. N. Ahad, M. O. Richardson, R. V. Pappu, CIDER: Resources to Analyze Sequence-Ensemble Relationships of Intrinsically Disordered Proteins. *Biophys J* **112**, 16-21 (2017).
210. R. K. Das, R. V. Pappu, Conformations of intrinsically disordered proteins are influenced by linear sequence distributions of oppositely charged residues. *Proc Natl Acad Sci U S A* **110**, 13392-13397 (2013).
211. M. T. Wei *et al.*, Phase behaviour of disordered proteins underlying low density and high permeability of liquid organelles. *Nat Chem* **9**, 1118-1125 (2017).

212. K. E. Handwerger, J. A. Cordero, J. G. Gall, Cajal bodies, nucleoli, and speckles in the *Xenopus* oocyte nucleus have a low-density, sponge-like structure. *Mol Biol Cell* **16**, 202-211 (2005).
213. J. P. Brady *et al.*, Structural and hydrodynamic properties of an intrinsically disordered region of a germ cell-specific protein on phase separation. *Proc Natl Acad Sci U S A* **114**, E8194-E8203 (2017).
214. C. G. de Kruif, F. Weinbreck, R. de Vries, Complex coacervation of proteins and anionic polysaccharides. *Curr Opin Colloid In* **9**, 340-349 (2004).
215. I. Kwon *et al.*, Phosphorylation-Regulated Binding of RNA Polymerase II to Fibrous Polymers of Low-Complexity Domains (vol 155, pg 1049, 2013). *Cell* **156**, 374-374 (2014).
216. H. Jiang *et al.*, Phase transition of spindle-associated protein regulate spindle apparatus assembly. *Cell* **163**, 108-122 (2015).
217. S. E. Reichheld, L. D. Muiznieks, F. W. Keeley, S. Sharpe, Direct observation of structure and dynamics during phase separation of an elastomeric protein. *Proc Natl Acad Sci U S A* **114**, E4408-E4415 (2017).
218. D. W. Urry *et al.*, Hydrophobicity scale for proteins based on inverse temperature transitions. *Biopolymers* **32**, 1243-1250 (1992).
219. D. M. Mitrea *et al.*, Nucleophosmin integrates within the nucleolus via multi-modal interactions with proteins displaying R-rich linear motifs and rRNA. *Elife* **5**, (2016).
220. H. Zhang *et al.*, RNA Controls PolyQ Protein Phase Transitions. *Mol Cell* **60**, 220-230 (2015).
221. J. Berry, S. C. Weber, N. Vaidya, M. Haataja, C. P. Brangwynne, RNA transcription modulates phase transition-driven nuclear body assembly. *Proc Natl Acad Sci U S A* **112**, E5237-5245 (2015).
222. S. P. Shevtsov, M. Dundr, Nucleation of nuclear bodies by RNA. *Nat Cell Biol* **13**, 167-173 (2011).
223. Y. S. Mao, H. Sunwoo, B. Zhang, D. L. Spector, Direct visualization of the co-transcriptional assembly of a nuclear body by noncoding RNAs. *Nat Cell Biol* **13**, 95-101 (2011).
224. W. M. Aumiller, Jr., C. D. Keating, Phosphorylation-mediated RNA/peptide complex coacervation as a model for intracellular liquid organelles. *Nat Chem* **8**, 129-137 (2016).

225. W. J. Krzyzosiak *et al.*, Triplet repeat RNA structure and its role as pathogenic agent and therapeutic target. *Nucleic Acids Res* **40**, 11-26 (2012).
226. A. R. La Spada, J. P. Taylor, Repeat expansion disease: progress and puzzles in disease pathogenesis. *Nat Rev Genet* **11**, 247-258 (2010).
227. V. V. Belzil, T. F. Gendron, L. Petrucelli, RNA-mediated toxicity in neurodegenerative disease. *Mol Cell Neurosci* **56**, 406-419 (2013).
228. P. K. Todd, H. L. Paulson, RNA-mediated neurodegeneration in repeat expansion disorders. *Ann Neurol* **67**, 291-300 (2010).
229. N. A. Timchenko *et al.*, RNA CUG repeats sequester CUGBP1 and alter protein levels and activity of CUGBP1. *J Biol Chem* **276**, 7820-7826 (2001).
230. E. G. Conlon *et al.*, The C9ORF72 GGGGCC expansion forms RNA G-quadruplex inclusions and sequesters hnRNP H to disrupt splicing in ALS brains. *Elife* **5**, (2016).
231. J. W. Miller *et al.*, Recruitment of human muscleblind proteins to (CUG)(n) expansions associated with myotonic dystrophy. *EMBO J* **19**, 4439-4448 (2000).
232. A. Jain, R. D. Vale, RNA phase transitions in repeat expansion disorders. *Nature* **546**, 243-247 (2017).
233. A. K. Henras *et al.*, The post-transcriptional steps of eukaryotic ribosome biogenesis. *Cell Mol Life Sci* **65**, 2334-2359 (2008).
234. S. Koga, D. S. Williams, A. W. Perriman, S. Mann, Peptide-nucleotide microdroplets as a step towards a membrane-free protocell model. *Nat Chem* **3**, 720-724 (2011).
235. T. Hyman, C. Brangwynne, In Retrospect: The Origin of Life. *Nature* **491**, 524-525 (2012).
236. D. Zwicker, R. Seyboldt, C. A. Weber, A. A. Hyman, F. Jülicher, Growth and division of active droplets provides a model for protocells. *Nature Physics* **13**, 408-413 (2016).
237. T. Takahara, T. Maeda, Transient sequestration of TORC1 into stress granules during heat stress. *Mol Cell* **47**, 242-252 (2012).
238. F. Wippich *et al.*, Dual specificity kinase DYRK3 couples stress granule condensation/dissolution to mTORC1 signaling. *Cell* **152**, 791-805 (2013).
239. T. M. Franzmann *et al.*, Phase separation of a yeast prion protein promotes cellular fitness. *Science* **359**, (2018).

240. L.-P. Bergeron-Sandoval *et al.*, Endocytosis caused by liquid-liquid phase separation of proteins. *bioRxiv*, (2017).
241. A. Nemeth, G. Langst, Genome organization in and around the nucleolus. *Trends Genet* **27**, 149-156 (2011).
242. T. Fukaya, B. Lim, M. Levine, Enhancer Control of Transcriptional Bursting. *Cell* **166**, 358-368 (2016).
243. A. Pombo, N. Dillon, Three-dimensional genome architecture: players and mechanisms. *Nat Rev Mol Cell Biol* **16**, 245-257 (2015).
244. A. G. Larson *et al.*, Liquid droplet formation by HP1alpha suggests a role for phase separation in heterochromatin. *Nature* **547**, 236-240 (2017).
245. A. R. Strom *et al.*, Phase separation drives heterochromatin domain formation. *Nature* **547**, 241-245 (2017).
246. M. A. Forrellad *et al.*, Virulence factors of the Mycobacterium tuberculosis complex. *Virulence* **4**, 3-66 (2013).
247. A. L. Hestvik, Z. Hmama, Y. Av-Gay, Mycobacterial manipulation of the host cell. *FEMS Microbiol Rev* **29**, 1041-1050 (2005).
248. J. L. Flynn, J. Chan, Immune evasion by Mycobacterium tuberculosis: living with the enemy. *Curr Opin Immunol* **15**, 450-455 (2003).
249. Y. Av-Gay, M. Everett, The eukaryotic-like Ser/Thr protein kinases of Mycobacterium tuberculosis. *Trends Microbiol* **8**, 238-244 (2000).
250. J. Chao *et al.*, Protein kinase and phosphatase signaling in Mycobacterium tuberculosis physiology and pathogenesis. *Biochim Biophys Acta* **1804**, 620-627 (2010).
251. F. L. Wu *et al.*, The Ser/Thr Protein Kinase Protein-Protein Interaction Map of M. tuberculosis. *Mol Cell Proteomics* **16**, 1491-1506 (2017).
252. R. M. Corrales *et al.*, Phosphorylation of mycobacterial PcaA inhibits mycolic acid cyclopropanation: consequences for intracellular survival and for phagosome maturation block. *J Biol Chem* **287**, 26187-26199 (2012).
253. M. Braibant, P. Gilot, J. Content, The ATP binding cassette (ABC) transport systems of Mycobacterium tuberculosis. *FEMS Microbiol Rev* **24**, 449-467 (2000).
254. K. Sureka *et al.*, Novel role of phosphorylation-dependent interaction between FtsZ and FipA in mycobacterial cell division. *PLoS One* **5**, e8590 (2010).

255. V. Molle *et al.*, Embr2, a structural homologue of EmbR, inhibits the Mycobacterium tuberculosis kinase/substrate pair PknH/EmbR. *Biochem J* **410**, 309-317 (2008).
256. Y. Shen, A. Bax, Identification of helix capping and β -turn motifs from NMR chemical shifts. *J Biomol NMR* **52**, 211-232 (2012).
257. M. V. Berjanskii, D. S. Wishart, A simple method to predict protein flexibility using secondary chemical shifts. *J Am Chem Soc* **127**, 14970-14971 (2005).
258. P. Dosset, J. C. Hus, M. Blackledge, D. Marion, Efficient analysis of macromolecular rotational diffusion from heteronuclear relaxation data. *J Biomol NMR* **16**, 23-28 (2000).
259. H. Liao, I. J. Byeon, M. D. Tsai, Structure and function of a new phosphopeptide-binding domain containing the FHA2 of Rad53. *J Mol Biol* **294**, 1041-1049 (1999).
260. G. Platzer, M. Okon, L. P. McIntosh, pH-dependent random coil (^1H) , (^{13}C) , and (^{15}N) chemical shifts of the ionizable amino acids: a guide for protein pK_a measurements. *J Biomol NMR* **60**, 109-129 (2014).
261. E. A. Bienkiewicz, K. J. Lumb, Random-coil chemical shifts of phosphorylated amino acids. *J Biomol NMR* **15**, 203-206 (1999).
262. I. R. Kleckner, M. P. Foster, An introduction to NMR-based approaches for measuring protein dynamics. *Biochim Biophys Acta* **1814**, 942-968 (2011).
263. R. Li, C. Woodward, The hydrogen exchange core and protein folding. *Protein Sci* **8**, 1571-1590 (1999).
264. M. M. Krishna, L. Hoang, Y. Lin, S. W. Englander, Hydrogen exchange methods to study protein folding. *Methods* **34**, 51-64 (2004).
265. M. Mammen, S. K. Choi, G. M. Whitesides, Polyvalent interactions in biological systems: Implications for design and use of multivalent ligands and inhibitors. *Angew Chem Int Edit* **37**, 2755-2794 (1998).
266. P. Wang *et al.*, II. Structure and specificity of the interaction between the FHA2 domain of Rad53 and phosphotyrosyl peptides. *J Mol Biol* **302**, 927-940 (2000).
267. G. I. Lee, Z. Ding, J. C. Walker, S. R. Van Doren, NMR structure of the forkhead-associated domain from the Arabidopsis receptor kinase-associated protein phosphatase. *Proc Natl Acad Sci U S A* **100**, 11261-11266 (2003).
268. I. J. Byeon, H. Li, H. Song, A. M. Gronenborn, M. D. Tsai, Sequential phosphorylation and multisite interactions characterize specific target recognition by the FHA domain of Ki67. *Nat Struct Mol Biol* **12**, 987-993 (2005).

269. Z. Ding *et al.*, PhosphoThr peptide binding globally rigidifies much of the FHA domain from Arabidopsis receptor kinase-associated protein phosphatase. *Biochemistry* **44**, 10119-10134 (2005).
270. S. Fortuin *et al.*, Phosphoproteomics analysis of a clinical Mycobacterium tuberculosis Beijing isolate: expanding the mycobacterial phosphoproteome catalog. *Front Microbiol* **6**, 6 (2015).
271. A. Villarino *et al.*, Proteomic identification of M. tuberculosis protein kinase substrates: PknB recruits GarA, a FHA domain-containing protein, through activation loop-mediated interactions. *J Mol Biol* **350**, 953-963 (2005).
272. J. H. Weng *et al.*, Uncovering the Mechanism of Forkhead-Associated Domain-Mediated TIFA Oligomerization That Plays a Central Role in Immune Responses. *Biochemistry* **54**, 6219-6229 (2015).
273. J. Y. Ahn, X. Li, H. L. Davis, C. E. Canman, Phosphorylation of threonine 68 promotes oligomerization and autophosphorylation of the Chk2 protein kinase via the forkhead-associated domain. *J Biol Chem* **277**, 19389-19395 (2002).
274. P. Barthe *et al.*, Dynamic and structural characterization of a bacterial FHA protein reveals a new autoinhibition mechanism. *Structure* **17**, 568-578 (2009).
275. J. H. Yeon, F. Heinkel, M. Sung, D. Na, J. Gsponer, Systems-wide Identification of cis-Regulatory Elements in Proteins. *Cell Syst* **2**, 89-100 (2016).
276. M. R. Wilkins *et al.*, Protein identification and analysis tools in the ExPASy server. *Methods Mol Biol* **112**, 531-552 (1999).
277. F. W. Studier, Protein production by auto-induction in high density shaking cultures. *Protein Expr Purif* **41**, 207-234 (2005).
278. F. Delaglio *et al.*, NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J Biomol NMR* **6**, 277-293 (1995).
279. W. Lee, M. Tonelli, J. L. Markley, NMRFAM-SPARKY: enhanced software for biomolecular NMR spectroscopy. *Bioinformatics* **31**, 1325-1327 (2015).
280. M. Sattler, J. Schleucher, C. Griesinger, Heteronuclear multidimensional NMR experiments for the structure determination of proteins in solution employing pulsed field gradients. *Prog Nucl Mag Res Sp* **34**, 93-158 (1999).
281. M. L. Ludwiczek *et al.*, Strategies for modulating the pH-dependent activity of a family 11 glycoside hydrolase. *Biochemistry* **52**, 3138-3156 (2013).

282. P. Guntert, L. Buchner, Combined automated NOE assignment and structure calculation with CYANA. *J Biomol NMR* **62**, 453-471 (2015).
283. Y. Shen, F. Delaglio, G. Cornilescu, A. Bax, TALOS+: a hybrid method for predicting protein backbone torsion angles from NMR chemical shifts. *J Biomol NMR* **44**, 213-223 (2009).
284. S. M. Pascal, D. R. Muhandiram, T. Yamazaki, J. D. Formankay, L. E. Kay, Simultaneous Acquisition of N-15-Edited and C-13-Edited Noe Spectra of Proteins Dissolved in H₂O. *J Magn Reson Ser B* **103**, 197-201 (1994).
285. M. Schubert, D. Labudde, H. Oschkinat, P. Schmieder, A software tool for the prediction of Xaa-Pro peptide bond conformations in proteins based on ¹³C chemical shift statistics. *J Biomol NMR* **24**, 149-154 (2002).
286. H. Ryu, G. Lim, B. H. Sung, J. Lee, NMRe: a web server for NMR protein structure refinement with high-quality structure validation scores. *Bioinformatics* **32**, 611-613 (2016).
287. W. Kabsch, C. Sander, Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers* **22**, 2577-2637 (1983).
288. Schrodinger, LLC. (2015).
289. N. A. Farrow *et al.*, Backbone Dynamics of a Free and a Phosphopeptide-Complexed Src Homology-2 Domain Studied by N-15 Nmr Relaxation. *Biochemistry* **33**, 5984-6003 (1994).
290. T. L. Hwang, P. C. van Zijl, S. Mori, Accurate quantitation of water-amide proton exchange rates using the phase-modulated CLEAN chemical EXchange (CLEANEX-PM) approach with a Fast-HSQC (FHSQC) detection scheme. *J Biomol NMR* **11**, 221-226 (1998).
291. C. Perez-Borrajero, M. Okon, L. P. McIntosh, Structural and Dynamics Studies of Pax5 Reveal Asymmetry in Stability and DNA Binding by the Paired Domain. *J Mol Biol* **428**, 2372-2391 (2016).
292. Y. Z. Zhang, *Protein and Peptide Structure and Interactions Studied by Hydrogen Exchange and NMR*. (University of Pennsylvania, 1995).
293. J. Granot, Determination of Dissociation-Constants of 1-1 Complexes from Nmr Data - Optimization of the Experimental Setup by Statistical-Analysis of Simulated Experiments. *Journal of Magnetic Resonance* **55**, 216-224 (1983).

294. A. Perrakis, R. Morris, V. S. Lamzin, Automated protein model building combined with iterative structure refinement. *Nat Struct Biol* **6**, 458-463 (1999).
295. Z. Otwinowski, W. Minor, Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol* **276**, 307-326 (1997).
296. T. R. Schneider, G. M. Sheldrick, Substructure solution with SHELXD. *Acta Crystallogr D Biol Crystallogr* **58**, 1772-1779 (2002).
297. T. C. Terwilliger, Maximum-likelihood density modification using pattern recognition of structural motifs. *Acta Crystallogr D Biol Crystallogr* **57**, 1755-1762 (2001).
298. T. C. Terwilliger, J. Berendzen, Automated MAD and MIR structure solution. *Acta Crystallogr D Biol Crystallogr* **55**, 849-861 (1999).
299. G. Langer, S. X. Cohen, V. S. Lamzin, A. Perrakis, Automated macromolecular model building for X-ray crystallography using ARP/wARP version 7. *Nat Protoc* **3**, 1171-1179 (2008).
300. P. Skubak, G. N. Murshudov, N. S. Pannu, Direct incorporation of experimental phase information in model refinement. *Acta Crystallogr D Biol Crystallogr* **60**, 2196-2201 (2004).
301. T. A. Jones, J. Y. Zou, S. W. Cowan, M. Kjeldgaard, Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Crystallogr A* **47 (Pt 2)**, 110-119 (1991).
302. A. J. McCoy *et al.*, Phaser crystallographic software. *J Appl Crystallogr* **40**, 658-674 (2007).
303. P. D. Adams *et al.*, PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr* **66**, 213-221 (2010).
304. P. Emsley, B. Lohkamp, W. G. Scott, K. Cowtan, Features and development of Coot. *Acta Crystallogr D Biol Crystallogr* **66**, 486-501 (2010).
305. Y. Nozaki, The preparation of guanidine hydrochloride. *Methods Enzymol* **26**, 43-50 (1972).
306. N. C. Pace, B. A. Shirley, J. A. Thomson, Measuring the conformational stability of a protein. *Protein Structure: A Practical Approach (Ed. T.E. Creighton) Oxford University Press*, 311-330 (1989).

307. M. H. Olsson, C. R. Sondergaard, M. Rostkowski, J. H. Jensen, PROPKA3: Consistent Treatment of Internal and Surface Residues in Empirical pKa Predictions. *J Chem Theory Comput* **7**, 525-537 (2011).
308. V. B. D.A. Case, J.T. Berryman, R.M. Betz, Q. Cai, D.S. Cerutti, T.E. Cheatham, III, T.A. Darden, R.E. Duke, H. Gohlke, A.W. Goetz, S. Gusarov, N. Homeyer, P. Janowski, J. Kaus, I. Kolossváry, A. Kovalenko, T.S. Lee, S. LeGrand, T. Luchko, R. Luo, B. Madej, K.M. Merz, F. Paesani, D.R. Roe, A. Roitberg, C. Sagui, R. Salomon-Ferrer, G. Seabra, C.L. Simmerling, W. Smith, J. Swails, R.C. Walker, J. Wang, R.M. Wolf, X. Wu and P.A. Kollman. (University of California, San Francisco, University of California, San Francisco, 2014).
309. I. N. Shindyalov, P. E. Bourne, Protein structure alignment by incremental combinatorial extension (CE) of the optimal path. *Protein Eng* **11**, 739-747 (1998).
310. A. A. Hyman, C. A. Weber, F. Julicher, Liquid-liquid phase separation in biology. *Annu Rev Cell Dev Biol* **30**, 39-58 (2014).
311. P. A. Chong, J. D. Forman-Kay, Liquid-liquid phase separation in cellular signaling systems. *Curr Opin Struct Biol* **41**, 180-186 (2016).
312. S. Luo, X. Xin, L. L. Du, K. Ye, Y. Wei, Dimerization Mediated by a Divergent Forkhead-associated Domain Is Essential for the DNA Damage and Spindle Functions of Fission Yeast Mdb1. *J Biol Chem* **290**, 21054-21066 (2015).
313. B. D. Manning, L. C. Cantley, Hitting the target: emerging technologies in the search for kinase substrates. *Sci STKE* **2002**, pe49 (2002).
314. G. Viswanathan, S. Yadav, S. V. Joshi, T. R. Raghunand, J. Jass, Insights into the function of FhaA, a cell division-associated protein in mycobacteria. *FEMS Microbiol Lett* **364**, (2017).
315. S. Ben-Yehuda, R. Losick, Asymmetric cell division in *B. subtilis* involves a spiral-like intermediate of the cytokinetic protein FtsZ. *Cell* **109**, 257-266 (2002).
316. K. J. Kieser, E. J. Rubin, How sisters grow apart: mycobacterial growth and division. *Nat Rev Microbiol* **12**, 550-562 (2014).
317. J. Dominguez-Escobar *et al.*, Processive movement of MreB-associated cell wall biosynthetic complexes in bacteria. *Science* **333**, 225-228 (2011).
318. Z. Bozoky *et al.*, Regulatory R region of the CFTR chloride channel is a dynamic integrator of phospho-dependent intra- and intermolecular interactions. *Proc Natl Acad Sci U S A* **110**, E4427-4436 (2013).

319. J. Q. Wu, C. D. McCormick, T. D. Pollard, Chapter 9: Counting proteins in living cells by quantitative fluorescence microscopy with internal standards. *Methods Cell Biol* **89**, 253-273 (2008).
320. H. Bach *et al.*, Escherichia coli maltose-binding protein as a molecular chaperone for recombinant intracellular cytoplasmic single-chain antibodies. *J Mol Biol* **312**, 79-93 (2001).
321. L. Abraham *et al.*, Limitations of Qdot labelling compared to directly-conjugated probes for single particle tracking of B cell receptor mobility. *Sci Rep* **7**, 11379 (2017).
322. M. Cimino, L. Alamo, L. Salazar, Permeabilization of the mycobacterial envelope for protein cytolocalization studies by immunofluorescence microscopy. *BMC Microbiol* **6**, 35 (2006).
323. R. Tafteh *et al.*, Real-time 3D stabilization of a super-resolution microscope using an electrically tunable lens. *Opt Express* **24**, 22959-22970 (2016).
324. R. Tafteh, D. R. Scriven, E. D. Moore, K. C. Chou, Single molecule localization deep within thick cells; a novel super-resolution microscope. *J Biophotonics* **9**, 155-160 (2016).
325. G. T. Dempsey, J. C. Vaughan, K. H. Chen, M. Bates, X. Zhuang, Evaluation of fluorophores for optimal performance in localization-based super-resolution imaging. *Nat Methods* **8**, 1027-1036 (2011).
326. J. A. Marsh, J. D. Forman-Kay, Sequence determinants of compaction in intrinsically disordered proteins. *Biophys J* **98**, 2383-2390 (2010).
327. A. Moglich, K. Joder, T. Kiefhaber, End-to-end distance distributions and intrachain diffusion constants in unfolded polypeptide chains indicate intramolecular hydrogen bond formation. *Proc Natl Acad Sci U S A* **103**, 12394-12399 (2006).
328. D. P. Goldenberg, Computational simulation of the statistical properties of unfolded proteins. *J Mol Biol* **326**, 1615-1633 (2003).
329. J. L. Rigaud, D. Levy, Reconstitution of membrane proteins into liposomes. *Methods Enzymol* **372**, 65-86 (2003).
330. T. K. Ritchie *et al.*, Chapter 11 - Reconstitution of membrane proteins in phospholipid bilayer nanodiscs. *Methods Enzymol* **464**, 211-231 (2009).
331. L. Balakrishnan, H. Venter, R. A. Shilling, H. W. van Veen, Reversible transport by the ATP-binding cassette multidrug export pump LmrA: ATP synthesis at the expense of downhill ethidium uptake. *J Biol Chem* **279**, 11273-11280 (2004).

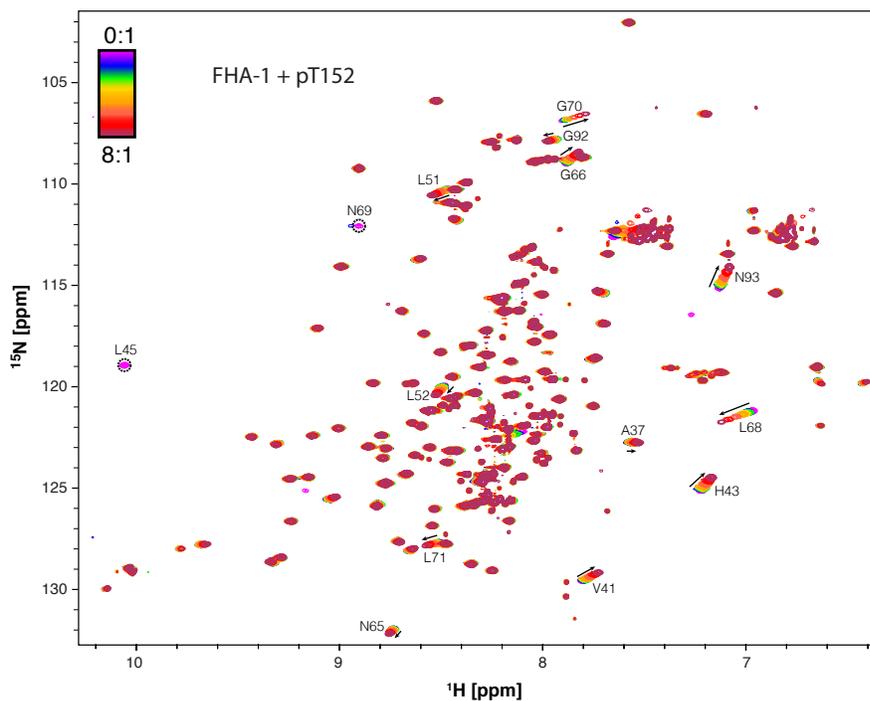
332. H. Venter, S. Velamakanni, L. Balakrishnan, H. W. van Veen, On the energy-dependence of Hoechst 33342 transport by the ABC transporter LmrA. *Biochem Pharmacol* **75**, 866-874 (2008).
333. S. Vardhana, M. Dustin, Supported Planar Bilayers for the Formation of Study of Immunological Synapses and Kinapse. *J Vis Exp*, (2008).
334. H. Glavinas *et al.*, Utilization of membrane vesicle preparations to study drug-ABC transporter interactions. *Expert Opin Drug Metab Toxicol* **4**, 721-732 (2008).
335. J. Guillen-Boixet, V. Buzon, X. Salvatella, R. Mendez, CPEB4 is regulated during cell cycle by ERK2/Cdk1-mediated phosphorylation and its assembly into liquid-like droplets. *Elife* **5**, (2016).
336. F. Geillon *et al.*, Peroxisomal ATP-binding cassette transporters form mainly tetramers. *J Biol Chem* **292**, 6965-6977 (2017).
337. M. Ramjeesingh *et al.*, A monomer is the minimum functional unit required for channel and ATPase activity of the cystic fibrosis transmembrane conductance regulator. *Biochemistry* **40**, 10700-10706 (2001).
338. M. Ramjeesingh, J. F. Kidd, L. J. Huan, Y. Wang, C. E. Bear, Dimeric cystic fibrosis transmembrane conductance regulator exists in the plasma membrane. *Biochem J* **374**, 793-797 (2003).
339. C. Li, K. Roy, K. Dandridge, A. P. Naren, Molecular assembly of cystic fibrosis transmembrane conductance regulator in plasma membrane. *J Biol Chem* **279**, 24673-24684 (2004).
340. Z. Bozoky *et al.*, Synergy of cAMP and calcium signaling pathways in CFTR regulation. *Proc Natl Acad Sci U S A* **114**, E2086-E2095 (2017).
341. D. C. Rees, E. Johnson, O. Lewinson, ABC transporters: the power to change. *Nat Rev Mol Cell Biol* **10**, 218-227 (2009).

Appendices

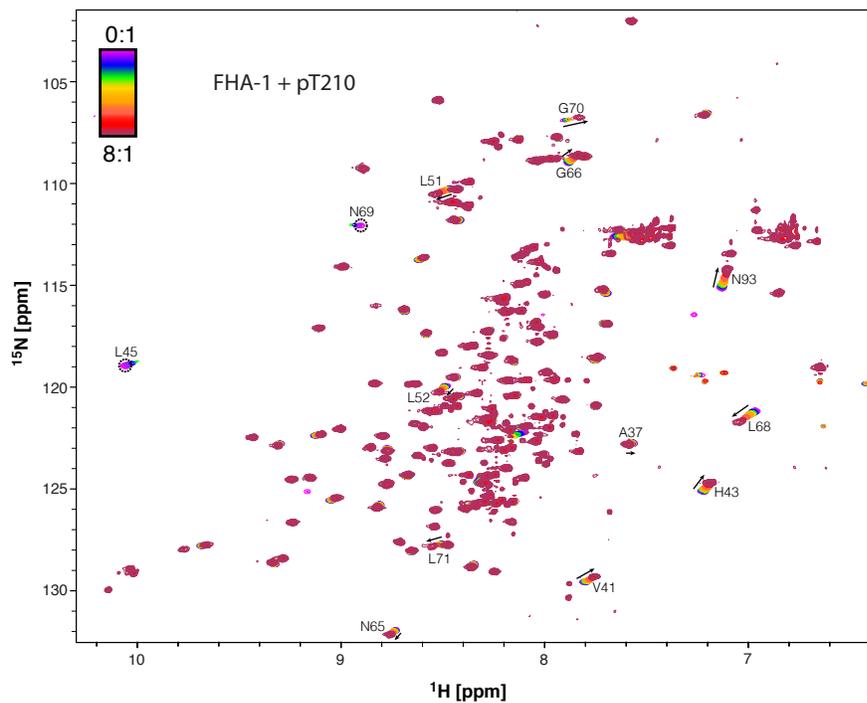
Appendix A ^{15}N -HSQC Monitored Titrations of Rv1747 FHA Domains with Phosphopeptides.

Overlaid spectra from ^{15}N -HSQC monitored titrations of ^{15}N -labeled (A) FHA-1 (Rv1747¹⁻¹⁵⁶) and (B) FHA-2 (Rv1747²⁰⁶⁻³¹⁰) with unlabeled synthetic phosphorylated peptide models of the pT152 and pT210 phospho-acceptor sites. The peptide:protein ratios are indicated with a rainbow color gradient. The progressive shift changes of selected amides showing binding in the fast exchange limit are shown by arrows. Amides exhibiting severe intermediate exchange broadening are identified with dashed circles.

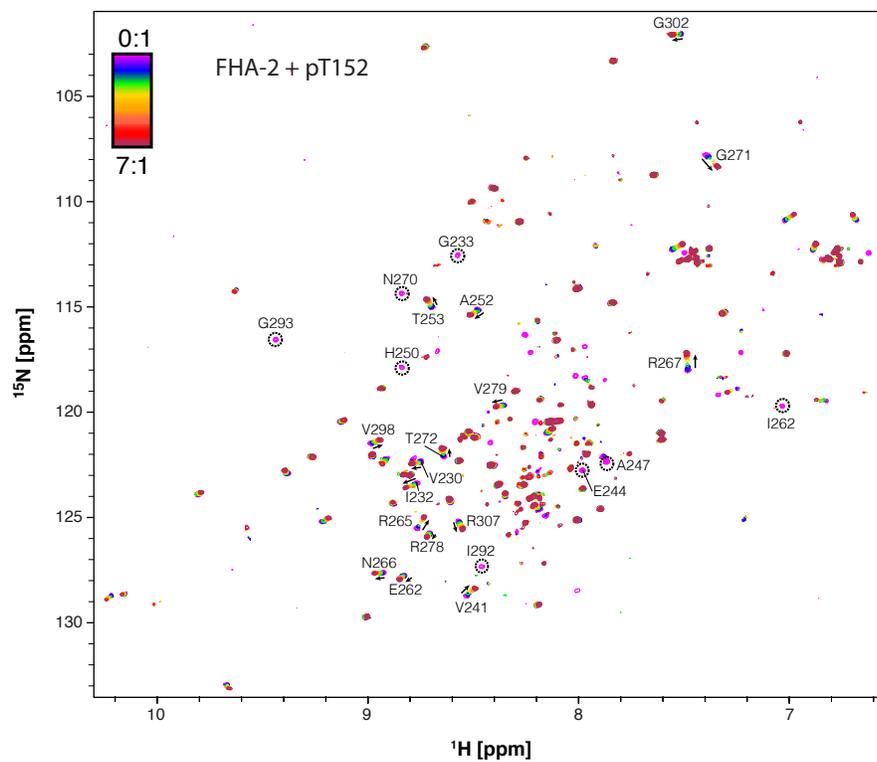
A.1 FHA-1 (Rv1747¹⁻¹⁵⁶) with pT152.



A.2 FHA-1 (Rv1747¹⁻¹⁵⁶) with pT210.



A.3 FHA-2 (Rv1747²⁰⁶⁻³¹⁰) with pT152.



A.4 FHA-2 (Rv1747²⁰⁶⁻³¹⁰) with pT210.

