Studying the mechanism of *Salmonella* effector protein, SteC

Timesh Dendy Pillay

Imperial College London
Faculty of Medicine, Department of Infectious Diseases
And
The Francis Crick Institute

PhD Supervisors: Katrin Rittinger and Teresa Thurston

A thesis submitted for the degree of
Doctor of Philosophy
Imperial College London
September 2024

Declaration of Originality

I, Timesh Dendy Pillay, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been appropriately referenced.

Copyright Declaration

The copyright of this thesis rests with the author. Unless otherwise indicated, its contents are licensed under a Creative Commons Attribution-Non-Commercial 4.0 International Licence (CC BY-NC).

Under this licence, you may copy and redistribute the material in any medium or format. You may also create and distribute modified versions of the work. This is on the condition that: you credit the author and do not use it, or any derivative works, for a commercial purpose.

When reusing or sharing this work, ensure you make the licence terms clear to others by naming the licence and linking to the licence text. Where a work has been adapted, you should indicate that the work has been changed and describe those changes. Please seek permission from the copyright holder for uses of this work that are not included in this licence or permitted under UK Copyright Law.

Abstract

SteC, a *Salmonella* effector kinase translocated across the *Salmonella*-containing vacuole (SCV) into the host cytosol, causes actin polymerisation in host cells. However, its mechanism of activity is poorly understood. The goal of this thesis was to structurally and biophysically characterise SteC using bioinformatic and experimental techniques; determine its interaction with substrates and molecular requirements for its kinase activity; and interrogate the role of kinase activity in host cells using ectopic expression and infection of mammalian cells with *Salmonella*. SteC has a C-terminal kinase domain (KD) that shares some sequence and predicted structural features with eukaryotic protein kinases, whilst lacking others. The purified kinase domain shows conformational flexibility and is monomeric in solution. One proposed substrate, FMNL1, was analysed using mass-spectrometry to determine SteC-dependent phosphorylation sites. SteC KD bound a peptide of FMNL1 spanning 3 phosphorylation sites. SteC recombinantly expressed in Sf9 cells was phosphorylated at S379. *In vitro* kinase assays showed that SteC KD was active and this activity was dependent on S379 phosphorylation. S379 phosphorylation was efficiently achieved *in vitro* by the mammalian kinase STK38, a reported binding partner of SteC. Phosphorylation on S379 was required for ATP binding. In *Salmonella* infection of 3T3 cells, S379 phosphorylation was required for dense focal actin polymerisation. In addition, a second capsular F-actin pattern was observed, which was SteC-dependent and kinase activity independent. SteC localised to one end of the SCV in a cap pattern, and *in vitro* lipid binding assays showed that the N-terminal domain bound phosphatidic acid.

In summary this thesis determined that SteC is structurally flexible and lacks some essential eukaryotic protein kinase features. Furthermore, the activating mechanism of the kinase domain was uncovered and its importance in *Salmonella*-infected mammalian cells was confirmed. Finally, a second kinase-independent actin polymerisation activity of SteC was described.

Acknowledgement

Firstly, I would like to thank Katrin and Teresa for giving me the opportunity to work on such a challenging and interesting project. You provided me with all the independence, resources and guidance I could have hoped for. Max Gutierrez, Mike Blackman and Alex McCarthy gave important independent academic guidance as thesis committee members and Philippa Matthews gave me the 'clinician scientist' support and reassurance I needed in moments when I was looking into the science abyss and couldn't see a way forward, backwards or sideways.

Mum, Dad and Kam have consistently listened to me moaning about the insignificance of everything, thank you. And I'd like to also recognise my friends. Letty has observed and supported me through the highs and lows of the whole three years. Billie has provided a constant and loving ear for my frustrations at the world of academia and real challenges inherent to attempting to do good biochemistry with a flexible and aggregation-prone protein. Allon has asked me difficult questions at important times, and in the process of trying to answer I learnt more. Ben, Sahra, Milly, Leah, Alma, Joel, Xalimo, Rita, Hannah, Lou, Hannah, Cecil, Simon, Marika and Leon have been so encouraging, even if I couldn't articulate to them (or myself) what I was *doing* most of the time.

The Crick's Science and Technology Platforms (STPs) have been instrumental to my learning, and the breadth of the science I could do in three years. Specific mentions are needed for the following scientists: Laura Masino from the Structural Biology STP, who has dedicated a huge amount of time to supporting me and teaching me the ways of biophysical experiments; Steve Howell, Sarah Maslen and Tania Auchynnikava from the Proteomics STP who have been responsive in the seemingly endless back-and-forth of doing good proteomics experiments; Dhira Joshi who synthesised peptides for my experiments; and Donald Bell and Matt Renshaw from the Light Microscopy STP who showed me how to capture and process nice super-resolution images.

Along the way, I have met a number of experts in their fields who have generously offered insights into relevant aspects of science or given technical and safety training and supervision for specialist techniques. These include Rajvee Shah Punatar (thanks for all the radioactivity stuff), Aleksandra (Ola) Michrowska (thanks for the actin), Ilaria Dalla Rosa, Michael Way, Mark Lemmon, Neil McDonald, Mathias Cobbaut, Acely Garza-Garcia, Danny Lang, Roger George, Svend Kjaer, Roksana Ogrodowicz and David Holden. Yan Li and Jay Hinton, collaborators from the University of Liverpool, were also instrumental. Thank you all. Thanks also to the academic training team at the Crick and the laboratory support teams, especially

4SW team at the Crick: Rob, Diaz, Cam and Luis, who have provided a friendly and committed service, for which I feel lucky.

The bedrock of my scientific learning has been formed by the ever-generous, patient and knowledgeable members of my two laboratories, who seemed surprisingly unphased by my relative lack of training when I started. Specifically, Jessica Huber, taught me how to be a protein biologist. Diego Esposito held my hand through many complex structural techniques and their subsequent analyses, which he often performed while I learnt. Xiujun Yu taught me how to engineer *Salmonella* strains. Ines Diaz del Olmo taught me how to culture mammalian cells. Ioanna Panagi gave me glimpses into how to perform the perfect western blot. Paul O'Sullivan entertained forays into flow cytometry, LifeAct-expressing cell lines and live cell imaging (all experimentally doomed, but I learnt a lot in the process!). Mei Liu helped me perform repeats of infection assays at the compressed end of my PhD time. Janina, Jane, Cassandra, Karen, Niek, Ben, Billy, Kat, George, Aini, Yizhou, Alison, Magda, Ada, Ioanna, Simeon and Francesca have all contributed acts of generosity over the years, big and small. And finally, Briac Lemetais and Patricia Diaz Belamendia, two fantastic MRes students based at Imperial who I supervised, contributed to the thesis directly and indirectly through reading, questioning and all the time in the lab. Thank you **all**.

Much of the duration of this work coincided with the prolonged assault on Gaza and the Palestinian people. I'd like to thank those at work with whom I could discuss, debate and share in activities of solidarity. This thesis is dedicated to the Palestinian people and those who have spoken up in their support.

Table of Contents

1. Introduction 21

1.1 Salmonella 21

1.1.1 Salmonella serovars 21

1.1.2 Disease caused by Salmonella 22

1.1.3 Type 3 Secretion System (T3SS) 24

1.1.4 Mechanisms of bacterial effector proteins 24

1.2 The actin cytoskeleton in *Salmonella* infection 33

1.2.1 Actin regulation in host cells 33

1.2.2 Manipulation of the actin cytoskeleton by intracellular bacteria 34

1.3 Kinase structure and function 37

1.3.1 Phosphorylation and its role in cells 37

1.3.2 Conserved eukaryotic protein kinase features 37

1.3.3 Kinase regulation 41

1.3.4 Mechanisms of kinase specificity 42

1.3.5 Kinase inhibitors 43

1.3.6 Prokaryotic kinases 43

1.3.7 Bacterial effector kinases 44

1.4 *Salmonella* effector kinase, SteC 46

1.4.1 Identification of an effector kinase 46

1.4.2 Reported substrates of SteC 46

1.4.3 Role of SteC in infectivity 49

1.5 Objectives 50

2. Materials and Methods 51

2.1 Bioinformatics 51

2.1.1 Protein sequence analysis 51

2.1.2 Protein structure prediction 51

2.1.3 Other bioinformatic tools 52

2.2 Plasmids, cloning and *Salmonella* transformation 53

2.2.1 Method of acquiring genetic material 53
2.2.2 Cloning and purifying plasmids 53
2.2.3 Salmonella strains 56
2.3 Recombinant protein expression and purification 57
2.3.1 Bacterial protein expression and purification....57
2.3.2 Size exclusion chromatography (SEC) and protein storage....57
2.3.3 Anion exchange chromatography (AIX) 58
2.3.4 SDS-PAGE 58
2.3.5 Insect cell (Sf9) expression and purification....58
2.3.6 Rabbit muscle actin purification....59
2.4 Structural and biophysical techniques....60
2.4.1 CD spectroscopy 60
2.4.2 1D NMR....60
2.4.3 SEC-MALLS 60
2.4.4 SAXS and SteC structure modelling 61
2.4.5 Biolayer Interferometry (BLI) 61
2.4.6 Hydrogen-deuterium exchange mass spectrometry (HDX-MS) 62
2.4.7 Protein crystallisation 62
2.5 Biochemical techniques 64
2.5.1 Kinase assay with Phos-stain....64
2.5.2 Radioactive kinase assay....64
2.5.3 PhosphoMS of recombinant protein 64
2.5.4 Intact mass spectrometry 65
2.5.5 Mant-AMPPNP binding 65
2.5.6 PIP strip analysis....65
2.6 Cell culture 66
2.6.1 Cell lines....66
2.6.2 Murine cell lysate....66
2.6.3 Transfection of HEK 293ET cells....66

2.6.4 Salmonella infection 66
2.6.5 Immunofluorescence microscopy 67
2.6.6 SDS-PAGE and immunoblotting 68
2.6.7 In cellulo effector translocation assay 69
2.6.8 Co-Immunoprecipitation mass spectrometry 69
2.6.9 Statistical analysis 70
3. Structural and biophysical characteristics of SteC 71
Introduction 71
3.1 Analysis of the sequence and predicted structure of SteC 71
3.1.1 Sequence analysis 71
3.1.2 Analysis of the AlphaFold2-predicted structure 74
3.1.3 Conserved sequence motifs across serovars 80
3.2 Biophysical and structural features of SteC 81
3.2.1 SteC expression and purification 81
3.2.2 SteC fragment design and purification 83
3.2.3 Analysis of the secondary structure of SteC 87
3.2.4 Analysis of the oligomeric state of the kinase domain using SEC-MALLS 90
3.2.5 Assessing the solution shape of kinase domain fragments using SAXS 92
3.3 SteC interaction with host substrate FMNL1 95
3.3.1 FMNL1 expression and purification 95
3.3.2 Establishing the kinase activity of SteC towards FMNL1 96
3.3.3 Identification of phosphorylation sites on FMNL1 97
3.3.4 Assessing substrate binding using biolayer interferometry 98
3.3.5 HDX-MS 100
3.3.6 Crystallisation attempts 102
Conclusion 105
4. Activity and regulation of SteC 106
Introduction 106

4.1 Characterisation of SteC phosphorylation and its effect on SteC enzymatic activity 107
4.1.1 SteC phospho-mass spectrometry identifies two phosphosites 107
4.1.2 Protein constructs produced for this chapter 109
4.1.3 Role of S379 in SteC activity 110
4.1.4 Impact of K256 and S379 mutation on protein folding 112
4.1.5 Impact of mutations on binding to substrate peptide 113
4.1.6 Impact of mutations on SteC binding to ATP 114
4.2 Studying the origin of SteC S379 phosphorylation 116
4.2.1 Intact MS of recombinantly expressed SteC 116
4.2.2 PhosphoMS of SteC WT and K256H from E. coli and Sf9 117
4.2.3 Autophosphorylation of SteC fragments 119
4.2.4 Using murine cell lysate to phosphorylate bacterially expressed SteC 120
4.2.5 STK38 as a candidate host kinase phosphorylating SteC S379 121
4.2.6 Relative activity of bacterial- and insect cell-expressed SteC 122
4.3 Role of C-lobe elements in kinase activity 124
4.3.1 Comparing the catalytic activity of different kinase domain fragments 124
4.3.2 Comparing SteC activity with and without the C-tail, and with N- or C-tags. 125
4.3.3 Comparing kinase activity of SteC constructs and tags with FMNL1 1-385. 126
4.3.4 S379 phosphorylation in kinase domain fragments 210-457 and 210-429 .. 128
4.3.5 HDX-MS to explore intramolecular interactions of the C-tail 130
4.4 Assessing kinase activity against a novel substrate, MYL12A 133
4.4.1 Using murine cell lysate to detect substrates of SteC 133
4.4.2 Expression of MYL12A 135
4.4.3 Kinase assay with MYL12A 136
4.4.4 Impact of SteC S379 on MYL12A phosphorylation 137
4.4.5 Assessing kinase activity with Michaelis-Menten kinetics 138
4.4.6 Kinase assay with calmodulin 140
Conclusion 141

5. Function of SteC in mammalian cells....143
Introduction....143
5.1 Screening for key catalytic residues by transfection....143
5.1.1 SteC-mediated actin polymerisation upon ectopic expression of SteC in HEK 293ET cells is kinase-dependent....143
5.1.2 Mutagenesis to identify functionally relevant residues in SteC....146
5.1.3 Novel SteC mutants R280A, H343A, D344A, F359A-F360A, D362A, D364A and S379A lost the ability to polymerise actin when ectopically expressed in HEK 293ET cells 146
5.2 The role of SteC phosphorylation at S379 and the C-tail in *Salmonella* infection ..150
5.2.1 Designing plasmids to complement a Salmonella knock-out strain....150
5.2.2 SteC is translocated into the host cytosol if its N-terminus is intact....150
5.2.3 Actin polymerisation in infected 3T3 cells is SteC-dependent....152
5.2.4 SteC produces two actin phenotypes in infection, one of which is kinase activity-independent....153
5.2.5 Quantifying actin phenotypes in infection....155
5.3 Understanding the kinase-independent activity of SteC in infection....159
5.3.1 The cellular localisation of SteC in infection....159
5.3.2 SteC does not bind to or phosphorylate actin....162
5.3.3 SteC 1-193 binds to phospholipids....164
5.3.4 Co-immunoprecipitation and mass spectrometry to identify stable binding partners of SteC in infected cells....167
5.4 MYL12A phosphorylation in infection....170
5.4.1 Assessing phospho-MYL12A antibodies....170
5.4.2 No SteC-dependent MYL12A phosphorylation was detected during infection 171
Conclusion....174
6. Discussion....176
6.1 A model of SteC activation through phosphorylation....177
6.1.1 What are the key catalytic residues for SteC kinase activity?....177

6.1.2 How is S379 phosphorylated? 179
6.1.3 How does S379 phosphorylation promote SteC-ATP binding? 181
6.1.4 Modelling S379 phosphorylation, ATP and magnesium binding using AlphaFold3 182
6.1.5 Summary 184
6.2 What is the role of the C-tail? 185
6.2.1 The role of the C-tail in substrate binding and kinase activity 185
6.2.2 The role of the C-tail in mammalian cells 187
6.2.3 Challenges with studying recombinantly expressed and purified SteC 187
6.2.4 A high-resolution structure of SteC 188
6.3 A model for the roles of SteC in actin polymerisation 189
6.3.1 SteC has a kinase-independent ability to polymerise actin 189
6.3.2 Evidence for substrates of SteC that mediate actin polymerisation 191
6.4 Summary 192
Conclusion 193
7. Appendices 194
References 200

List of Figures

Figure 1. 1: *Salmonella* phylogenetic nomenclature 22
Figure 1. 2: Schematic of *Salmonella* effector protein biochemistries 25
Figure 1. 3: Summary of the mechanisms of actin polymerisation, stabilisation and depolymerisation with associated regulatory factors 33
Figure 1. 4: *Salmonella* effector proteins manipulate the host actin cytoskeleton in all phases of the intracellular lifecycle 36
Figure 1. 5: Key kinase features using PKA, crystallised in the active conformation.....39
Figure 1. 6: Kinase recognition motifs.....43
Figure 1. 7: Domain map of SteC 46

Figure 2. 1: Gene sequence used for pWSK29 plasmid into *Salmonella* 54

Figure 3. 1: Reference SteC amino acid sequence 71
Figure 3. 2: Distribution of SteC across *Salmonella* serovars.....72
Figure 3. 3: Sequence alignment of SteC and homologs in non-*Salmonella* bacterial species 73
Figure 3. 4: Sequence alignment of SteC and RAF1.....74
Figure 3. 5: AlphaFold2 predicted structure of SteC.....75
Figure 3. 6: Secondary structural map of canonical kinase features 77
Figure 3. 7: Comparing PKA crystal and predicted SteC kinase domain structures.....78
Figure 3. 8: Sequence logos for hypothetical motifs of SteC across *Salmonella* serovars....80
Figure 3. 9: Full-length SteC protein purification attempts.....82
Figure 3. 10: Predicted structure-guided construct design 83
Figure 3. 11: Construct expression and purification.....85
Figure 3. 12: CD analysis of SteC, two domains and kinase domain fragments 87
Figure 3. 13: 1D ^{1}H-NMR spectra of SteC constructs.....89
Figure 3. 14: SEC-MALLS of SteC kinase domain fragments 91
Figure 3. 15: Scattering profiles and Kratky plots for SteC kinase domain fragments.....93
Figure 3. 16: SAXS low-resolution model of the shape the kinase domain fragments 94
Figure 3. 17: FMNL1 domain map and structure of the N-terminus.....95
Figure 3. 18: Purification of FMNL1 1-458 recombinantly expressed in Sf9 cells.....96
Figure 3. 19: Phos-stain kinase assay with SteC and FMNL1.....97
Figure 3. 20: Identifying SteC-induced FMNL1 phosphosites.....98
Figure 3. 21: Binding curves with estimated K_d values for SteC 210-457 and 210-429 100

Figure 3. 22: Hydrogen deuterium exchange of SteC 210-457 K256H and FMNL1 peptide 101
Figure 3. 23: Promising crystallisation conditions 104

Figure 4. 1: S379 and S75 are phosphosites of SteC.....108
Figure 4. 2: Symbols representing the recombinant protein boundaries, tags, tag positions and expression systems in subsequent figures 110
Figure 4. 3: Residue S379 of SteC is required for phosphorylation of FMNL1 peptide SteC 111
Figure 4. 4: CD spectra of kinase domain WT and mutants are comparable 113
Figure 4. 5: BLI to assess binding of SteC mutants and FMNL1 peptide 114
Figure 4. 6: Binding of SteC kinase domain to Mant-AMPPNP 115
Figure 4. 7: Separating SteC species by AIX.....116
Figure 4. 8: Relative quantification of S379 phosphorylation by phosphoMS.....118
Figure 4. 9: Radioactive kinase assay to distinguish trans- and cis-autophosphorylation...119
Figure 4. 10: SteC 210-457 K256H phosphorylated in the presence of murine cell lysate .121
Figure 4. 11: Radioactive kinase assay with SteC and mammalian kinase STK38 122
Figure 4. 12: Sf9-expressed SteC 210-457 is more active than the same construct expressed in *E. coli*.....123
Figure 4. 13: SteC kinase domain fragments 210-457 and 210-429 are active, whereas SteC 200-377 and 210-375 are not 124
Figure 4. 14: Tag position affects SteC 210-457 and 210-429 autophosphorylation and substrate peptide phosphorylation.....126
Figure 4. 15: Only N-tagged SteC 210-457 phosphorylates FMNL1 1-185.....127
Figure 4. 16: Relative quantification of S379 phosphorylation in SteC 210-457 and 210-429 by phosphoMS 128
Figure 4. 17: Radioactive kinase assay comparing SteC 210-429 WT and mutants with FMNL1 peptide alone and in the presence of murine cell lysate 129
Figure 4. 18: HDX-MS comparing SteC 210-457 and 210-429 131
Figure 4. 19: SteC phosphorylates a protein in murine cell lysate.....133
Figure 4. 20: Streptactin resin purification of MYL12A double mutant and WT 136
Figure 4. 21: Radioactive kinase assay comparing SteC constructs with MYL12A substrate 137
Figure 4. 22: Radioactive kinase assay of SteC 210-457 WT and mutants with MYL12A ..138
Figure 4. 23: Radioactive kinase assay with fixed SteC concentration and increasing MYL12A concentration.....139
Figure 4. 24: Radioactive kinase assay of SteC and calmodulin 140

Figure 5. 1: HEK 293ET cells ectopically expressing GFP-SteC WT, K256H, S379A or S379D 144
Figure 5. 2: Quantification of actin polymerisation in HEK 293ET ectopically expressing GFP-SteC WT and mutants 148
Figure 5. 3: SteC mutants E272A and D279A retain the ability to phosphorylate an FMNL1 peptide 149
Figure 5. 4: Translocation assay in 3T3 cells by western immunoblot 151
Figure 5. 5: Immunofluorescence of 3T3 cells infected with *Salmonella* WT, *ΔsteC* and *ΔsteC psteC* strains 153
Figure 5. 6: Immunofluorescence of 3T3 cells infected with *Salmonella* WT, *ΔsteC psteC*, K256H, S379A and S379D mutant strains 154
Figure 5. 7: Demonstration of the three actin phenotypes used for quantification 156
Figure 5. 8: proportion of infected cells with actin polymerisation 157
Figure 5. 9: Immunofluorescence of 3T3 cells infected with *Salmonella* strains shows SteC-2HA localised to filamentous structures 160
Figure 5. 10: Immunofluorescence of 3T3 cells infected with *Salmonella* strains shows SteC-2HA localised to the tip of solitary bacteria in the host cell 161
Figure 5. 11: SEC-MALLS of SteC 210-457 and Actin LatB 162
Figure 5. 12: SteC 210-457 does not demonstrate kinase activity towards Actin LatB 163
Figure 5. 13: Predicting SteC-membrane association using the AF prediction and MODA 164
Figure 5. 14: SteC 1-193 binds phosphatidic acid 165
Figure 5. 15: Skeletal formulae of the phospholipids that SteC 1-193 binds to 166
Figure 5. 16: Significantly enriched SteC-2HA binding partners from infected 3T3 cells 167
Figure 5. 17: Peptide coverage of SteC-2HA and SteC K256H-2HA from co-immunoprecipitation mass spectrometry during infection 168
Figure 5. 18: Testing anti-phosphoMYL12A antibodies with recombinantly expressed phosphorylated and unphosphorylated protein 171
Figure 5. 19: Detecting phosphorylated MYL12A in mammalian cells infected with *Salmonella* mutants 172
Figure 5. 20: Immunoblot analysis of infected RAW cells demonstrates no phosphorylation of MYL12A 173

Figure 6. 1: Highlighting SteC's experimentally determined key residues compared to PKA 178
Figure 6. 2: Sequence logo comparing the S379 region of SteC with 4 homologs 181
Figure 6. 3: Predicting the effect of S379 phosphorylation using AlphaFold3 183

Figure 6. 4: AlphaFold multimer predicts the site of FMNL1 peptide binding in SteC kinase domain 186
Figure 6. 5: Sequence motif analysis of SteC's substrates.....187
Figure 6. 6: Published crystal structure of SteC 202-375 188
Figure 6. 7: Schematic of SteC's regulation and roles in the host cell.....192

Figure S 1: Binding curves for SteC 210-457 with FMNL1 peptides.....194

List of Tables

Table 1. 1: List of *Salmonella* effectors 30
Table 1. 2: List of bacterial effector kinases and their kinase features.....45

Table 2. 1: Expression plasmids cloned, and summary of expression, purification and results where relevant 55
Table 2. 2: *Salmonella* strains prepared for infection assays.....56
Table 2. 3: Immunofluorescence staining reagents.....68
Table 2. 4: Antibodies used for western immunoblot and PIP strip 70

Table 3. 1: Top 10 hits searching SteC 1-193 AF predicted structure against human AF proteome with FoldSeek 79
Table 3. 2: Summary of constructs cloned, expressed and purified for this chapter.....86
Table 3. 3: SEC-SAXS parameters 92
Table 3. 4: Phosphorylated peptides of FMNL1 after incubation with SteC and ATP 98
Table 3. 5: Peptides covering FMNL1 phosphosites synthesised for binding studies 99
Table 3. 6: Summary of attempted crystallisation conditions 103

Table 4. 1: Methods used to assess kinase activity with recombinant expressed protein. ..106
Table 4. 2: Proteins expressed and purified for this chapter 109
Table 4. 3: Intact MS to determine phospho-species separated by AIX 117
Table 4. 4: Mass spectrometry identification of band from murine cell lysate showing SteC phosphorylation signal 135

Table 5. 1: List of 16 SteC kinase domain mutations used to screen for key catalytic residues by transfection, with rationale 145

Table S 1: Phosphopeptides identified by co-immunoprecipitation mass spectrometry 199

Abbreviations

1D	One-dimensional
αα	Amino acid
ADF	Actin depolymerising factor
ADP	Adenosine diphosphate
AEBSF	4-(2-aminoethyl)benzenesulfonyl fluoride hydrochloride
AF	AlphaFold2
AIX	Anion exchange chromatography
AL	Activation loop
AMPPNP	Adenylyl imidodiphosphate
ANOVA	Ordinary one-way analysis of variance
ARP2/3	Actin related protein 2/3 complex
ATP	Adenosine triphosphate
AS/DF	Actin severing or depolymerising factor
BLAST	Basic Local Alignment Search Tool (NCBI)
BLASTn	BLAST for nucleotide sequences
BLASTp	BLAST for protein sequences
BLI	Biolayer interferometry
BSA	Bovine serum albumin
CD	Circular dichroism
CDC	Cell division cycle
CL	Catalytic loop
CSA	Core *Salmonella* antigen
DAPI	4’6-diamidino-2-phenylindole
DC	Dendritic cell
DNA	Deoxyribonucleic acid
DMEM	Dulbecco’s Modified Eagle’s Medium
EDTA	Ethylenediaminetetraacetic acid
emPAI	Exponentially modified protein abundance index
ERK	Extracellular signal-regulated kinases
FCS	Foetal bovine serum
FERM	4.1 ezrin radixin moesin domain
FMNL	Formin-like protein
FT	Flow through
GAP	GTPase activating protein
GEF	Guanine nucleotide exchange factor

GFP	Green fluorescent protein
GlcNAc	N-acetylglucosamine
GSK	Glycogen synthase kinase
GST	Glutathione S-transferase
GTP	Guanosine triphosphate
GTPase	Guanosine triphosphatase
HA	Human influenza haemagluttinin
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HIV	Human Immunodeficiency Virus
HPLC	High-performance liquid chromatography
HRP	Horseradish peroxidase
HSP	Heat shock protein
ID	Identification
IgG	Immunoglobulin G
iNTS	Invasive non-Typhoidal *Salmonella* disease
IPTG	Isopropyl-β-D-thiogalactoside
JNK	c-Jun N-terminal kinases
LAMP	Lysosomal-associated membrane protein
LatB	Latrunculin B
LB	Luria Broth media
LCV	*Legionella*-containing vacuole
LIMK	Lim kinase
MAPK	Mitogen-activated protein kinases
MBP	Myelin basic protein
MHC	Major histocompatibility complex
MKK	Mitogen-activated protein kinase kinase
MLC	Myosin light chain
MLCK	Myosin light chain kinase
MLCP	Myosin light chain phosphatase
MOB	Monopolar spindle-one-binder
MS	Mass spectrometry
MW	Molecular weight
MYL	Myosin light chain
N-WASP	Neuronal Wiskott-Aldrich syndrome protein
NA	Not applicable
NCBI	National Centre for Biotechnology Information

ND	Not discovered
NF	Nuclear factor
NiNTA	Nickel nitrilotriacetic acid
NMR	Nuclear magnetic resonance
NOD	Nucleotide-binding oligomerisation domain
NTS	Non-Typhoidal *Salmonella* species
OD	Optical density
ORF	Open reading frame
PAE	Predicted alignment error
PAGE	Polyacrylamide gel electrophoresis
PBS	Phosphate buffered saline
PDB	Protein DataBase
PDZ	Post synaptic density protein
PEG	Polyethylene glycol
PFA	Paraformaldehyde
PH	Pleckstrin homology domain
PIP	Phosphoinositol
PKA/B	Protein kinase A/B
pLDDT	Predicted local distance difference test
PMN	Polymorphonuclear neutrophils
PNS	Post-nuclear supernatant
PtdIns	Phosphatidylinositol
PTM	Post-translational modifications
PVDF	Polyvinylidene difluoride
rMLST	Ribosomal multilocus sequence typing
RNA	Ribonucleic acid
ROCK	Rho kinase
SAXS	Small-angle X-ray scattering
SCV	*Salmonella*-containing vacuole
SDS	Sodium dodecyl sulphate
SDS-PAGE	Sodium dodecyl sulphate-polyacrylamide gel electrophoresis
SEC	Size exclusion chromatography
SEC-MALLS	Size exclusion chromatography coupled to multi-angle laser light scattering
SERBP1	SERPINE1 mRNA-binding protein 1
SERPINE1	Serine protease inhibitor E1
Sf9	*Spodoptera Frugiperda* 9

Sif	*Salmonella*-induced filament
SISTR1	*Salmonella* In Silico Typing Resource 1
SIT	*Salmonella*-induced tubule
siRNA	Small interfering ribonucleic acid
SPI	*Salmonella* pathogenicity island
STAT	Signal transducer and activator transcription
STP	Science technology platform (Francis Crick Institute)
T3SS	Type 3 Secretion System
TAB	Tandem actin binding
TABDN	Tandem actin binding domain nucleators
TBS-T	Tris-buffered saline with 0.1% Tween 20
TCEP	Tris(2-carboxyethyl)phosphine hydrochloride
TEV	Tobacco etch virus
TNF	Tumour necrosis factor
TOF	Time-of-flight
TRIM	Tripartide motif-containing protein
TS	TwinStrep tag
UDP	Uridine diphosphate
UK	United Kingdom
USA	United States of America
UV	Ultraviolet
WAVE	Wiskott-Aldrich syndrome protein family verprolin-homologous protein
WT	Wild-type

1. Introduction

Host-bacterium interactions represent the culmination of an evolutionary interplay lasting millennia. Intracellular bacteria today range from the essential energy factory for eukaryotic cells, mitochondria, to the deadliest human pathogen on earth, *Mycobacterium tuberculosis*. This diversity has manifested through the evolution of a plethora of molecular mechanisms to survive in the external environment, enter the host cell, escape host immunity (both extracellular and intracellular), create a safe replicative niche within the host cell and transmit from one host cell or host organism to another. *Salmonella*, one such intracellular bacterium, has evolved for between 120 and 160 million years with its various invertebrate and vertebrate hosts (Winfield & Groisman, 2004). In the modern day, a subset of *Salmonella* species causes an estimated 113 million cases of human illness and 860,000 deaths per year worldwide (Ao et al., 2015; Global Burden of Disease Collaborators, 2018; Stanaway et al., 2019).

1.1 Salmonella

1.1.1 Salmonella *serovars*

Salmonellae are Gram negative bacteria of the Entereobacteriaceae family. The genus is comprised of two species, *Salmonella enterica* and *Salmonella bongori*. The phylogenetic nomenclature is determined by laboratory assessment of components of cell wall (O) and flagellae (H1:H2) as well as other biochemical tests (Chattaway et al., 2021). *Salmonella enterica* is divided into 6 subspecies (Figure 1.1). *Salmonella enterica* subspecies *enterica* is divided into two clades. Clade A, containing over 2,600 species across two subclades, accounts for most of the *Salmonella* disease in humans (Worley et al., 2018) (Figure 1.1). The division of current subspecies and clades is under discussion (Chattaway et al., 2021; Worley et al., 2018). For consistency, *Salmonella enterica* subspecies *enterica* will be referred to as *S. enterica* and other serovars will be referred to with the common abbreviation: e.g. *Salmonella enterica enterica* Typhimurium will be named *S.* Typhimurium.

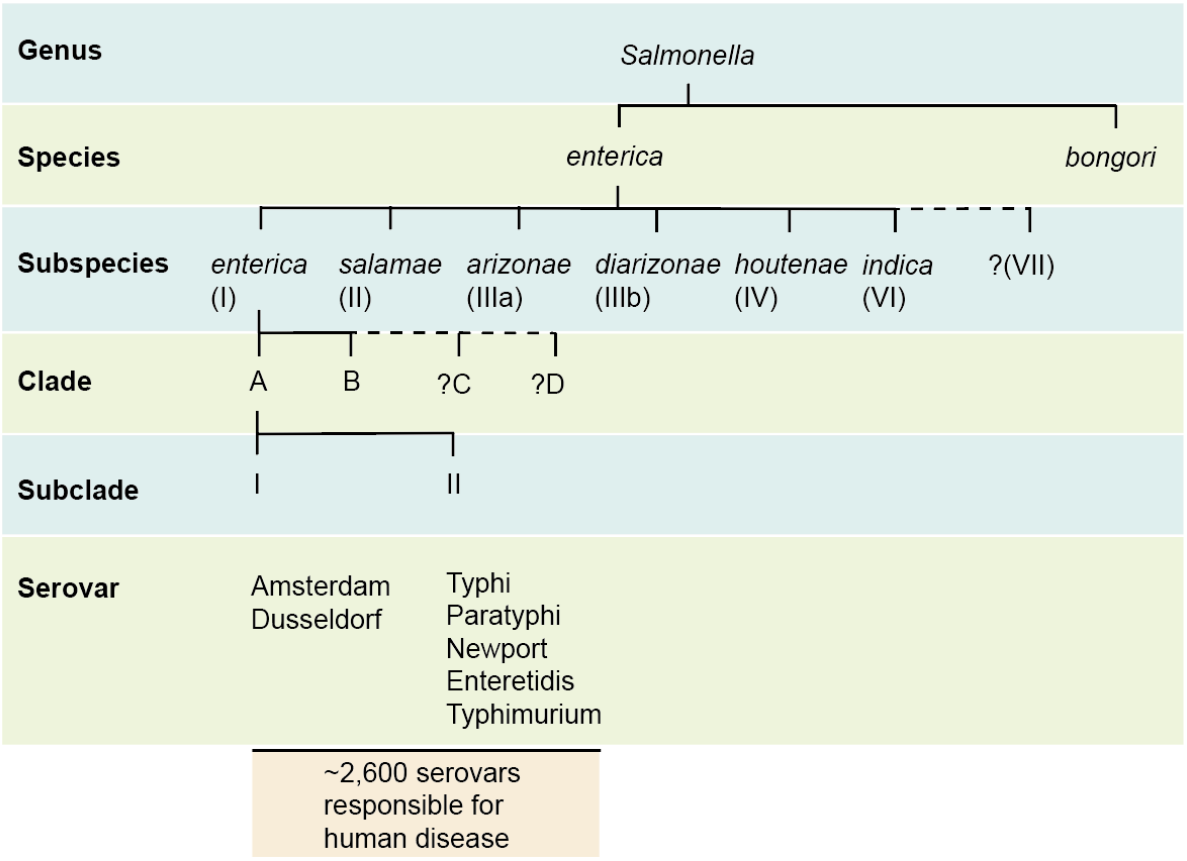


Figure 1. 1: *Salmonella* phylogenetic nomenclature

Salmonella species, subspecies, clades, subclades and serovars are determined by laboratory interrogation of cell wall type, and phase I and II flagella (O:H1:H2), alongside other biochemical characteristics such as the ability to use dextrorotatory (L+) d-tartrate as a source for fumarate respiration (Chattaway et al., 2021). Most *Salmonella* disease in humans is caused by ~2,600 serovars in *Salmonella enterica* Clade A. Phylogenetic definitions are under ongoing discussion and adjustment, for example a seventh subspecies and further two clades (represented with dotted lines) (Chattaway et al., 2021; Worley et al., 2018).

1.1.2 Disease caused by Salmonella

Salmonella spp. are transmitted by the faecal-oral route. Their disease manifestations in humans can be divided into three. The first, gastroenteritis caused by non-typhoidal serovars of *Salmonella*, is a disease characterised by bloody diarrhoea and fever, affecting 95.1 million people a year, and causing 50,771 deaths globally (Global Burden of Disease Collaborators, 2018).

The second, invasive non-typhoidal *Salmonella* disease (iNTS) caused by the gastroenteritis-causing serovars, is a more severe disease manifestation, characterised by *Salmonella* breaking through the epithelial barrier of the gut and entering the bloodstream (bacteriaemia),

where it can seed in various tissues including the meninges, spleen, bone and lung (Feasey et al., 2012). Those at high risk include children under 5 years, those with comorbidities such as sickle cell disease, recent malaria or HIV infection and those on immunosuppressive medication (Marchello et al., 2022). iNTS causes 3.4 million cases of disease per year, leading to 680,000 deaths (Ao et al., 2015), with the highest risk of death in African countries (Marchello et al., 2022).

The third, typhoid fever is a distinct disease caused by a distinct set of serovars: *S.* Typhi and *S.* Paratyphi. Its features include bacteraemia, prolonged high fever, headache and malaise - without diarrhoea. This can progress to altered mental state, gastrointestinal bleeding, septic shock and death if untreated. Despite four available vaccines that provide good protection in clinical trials (Batool et al., 2024), typhoid fever causes 14.3 million cases per year, and 135,900 deaths (Stanaway et al., 2019). This is due to partial vaccine coverage in endemically affected populations. Interestingly, while *S.* Typhi and *S.* Paratyphi are host-restricted to humans and primates, the manifestation of NTS infection in mice resembles the human typhoid, with systemic infection and no diarrhoea (Loeffler, 1892; Ørskov & Moltke, 1929), making it an attractive animal model for typhoid fever.

Complicating this picture, and presenting an ominous future, is the rise of antibiotic resistance. This poses a specific challenge in iNTS and Typhoid fever, which affect populations least served by microbiological drug susceptibility testing. What reports are available suggest that iNTS-causing serovars from clinical isolates can be multidrug resistant in 50-75% of cases. The better monitored typhoid fever-causing serovars are drug-resistant in a third of cases. Resistance to fluoroquinolones, the first line antibiotic class for iNTS and typhoid fever in many endemic regions, is now common (Kariuki et al., 2015). Poor vaccine coverage against typhoid fever and antibiotic resistance in *Salmonella* infection therefore present sad examples of the inverse care law, where the populations most in need of medical intervention have least access to it.

Understanding the biological mechanisms of the *Salmonella* lifecycle in its mammalian host has never been more important. The development of novel vaccines (for non-typhoidal strains); non-vaccine preventive measures; rapid and effective diagnostics; and *Salmonella*-targeting antibiotics are global health priorities as demonstrated by the prominent position of *Salmonella* in the World Health Organization's priority pathogen list (World Health Organization, 2024).

1.1.3 Type 3 Secretion System (T3SS)

As well as being of global health importance, *Salmonella*, and more specifically *S.* Typhimurium, is a popular organism for research in bacteriology, host cell biology and host-pathogen interaction due to its host-adapted intracellular lifestyle, and the relative ease with which it can be grown and genomically manipulated. Central to the coordination of an intracellular niche is a secretion system, which provides a route of communication between the bacterial and host cytosols. *Salmonella* uses the type 3 secretion system (T3SS), which is a multi-megadalton complex comprised of over 20 proteins: ring structures in the bacterial inner and outer membranes, and a translocation pore in the host cell membrane (Deng et al., 2017). Evolved from the flagellum that provides bacterial motility, the T3SS contains a hollow conduit like a needle. Its main role is to translocate bacterial effector proteins (effectors) into the host cytosol. Effectors are translocated in two phases, dictated by the *Salmonella* pathogenicity islands (SPI)-1 and -2. The first raft of effectors is translocated into the host cytosol prior to cell entry to trigger phagocytosis via the SPI-1 encoded T3SS. Once inside the cell in the *Salmonella*-containing vacuole (SCV), SPI-2 effectors are translocated through a second SPI-2 encoded T3SS (Jennings et al., 2017).

1.1.4 Mechanisms of bacterial effector proteins

Bacterial effector proteins are virulence factors that provide the main mechanism by which intracellular bacteria manipulate the host cell environment to create a niche suitable for survival and replication. Therefore, bacterial effectors and their host binding partners represent a key molecular site of host-pathogen interaction. The study of effectors has revealed examples of convergent evolution and molecular mimicry (Pillay et al., 2023), manifestations of the distant evolutionary relationship between bacteria and mammals, and the countervailing evolutionary pressures of the host-parasite interaction. Horizontal gene transfer has also shaped this process (Bakkeren et al., 2022). This context makes bacterial effector biology ripe for the discovery of novel biochemistries. Perhaps the best studied repertoire of bacterial effector proteins is that of *Salmonella*, summarised in Figure 1.2.

In this set of effectors alone, various eukaryotic-like biochemistries can be found including a kinase (see Section 1.4), phosphatase, E3 ligases, a deubiquitinase, guanine nucleotide exchange factors (GEFs) and GTPase activating proteins (GAPs). Table 1.1 lists the *Salmonella* effectors with structural, biochemical and functional features, and colours denote whether the effector protein has enzymatic activity (green), an adaptor function (orange) or are poorly defined / alternative function (blue).

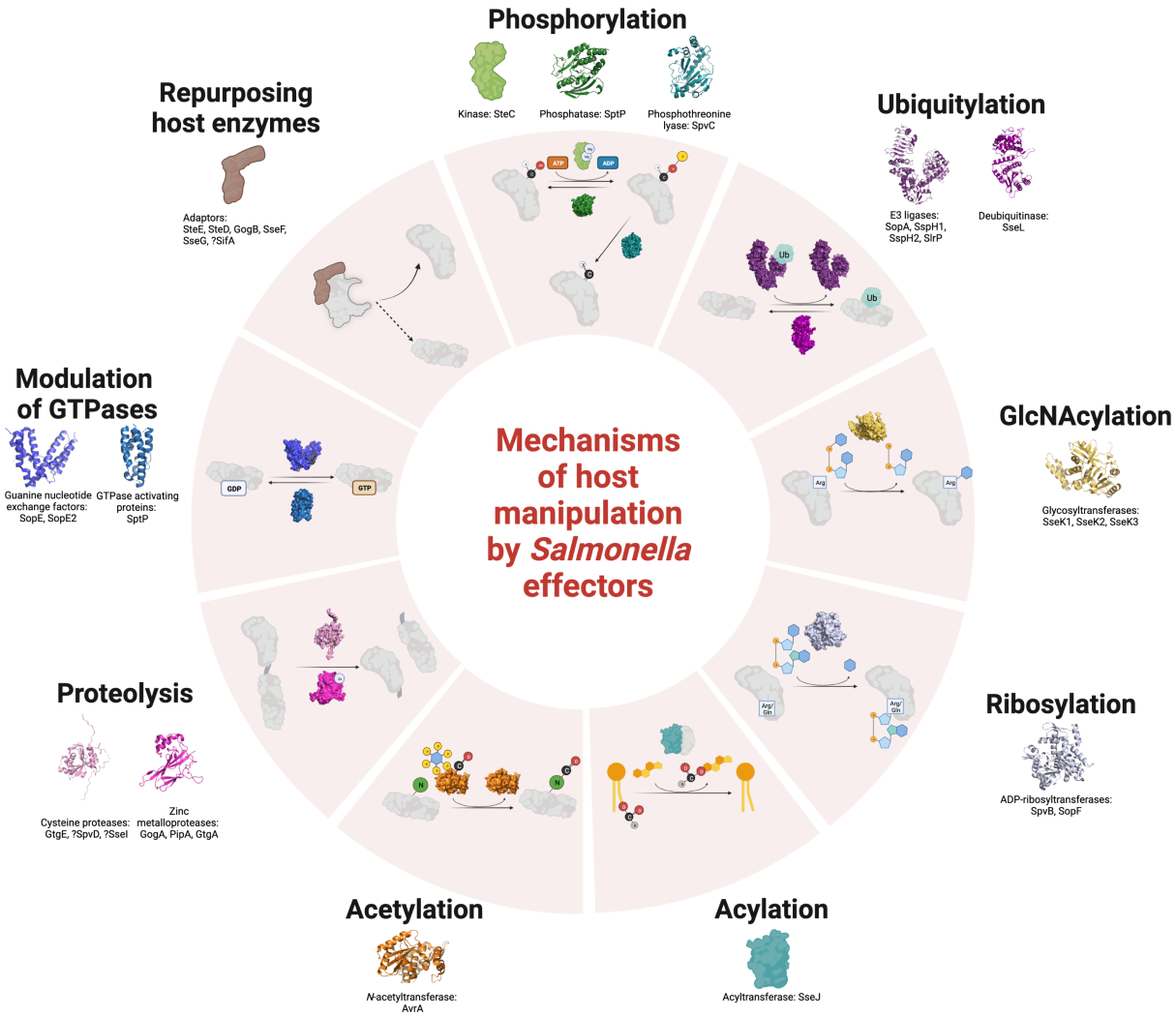


Figure 1. 2: Schematic of *Salmonella* effector protein biochemistries

Effectors are grouped by their enzymatic activities, which give them the catalytic capacity to modify host substrate molecules. Proteins displayed in cartoon form (outside) or surface form (inside), when available from their Protein DataBase (PDB) structures. Figure made with BioRender. Published as part of review of effector protein biology (Pillay et al., 2023).

Salmonella effectors can mimic host enzymatic activities, demonstrate novel catalytic activities to usurp established host signalling pathways, interact non-enzymatically to reprogram host enzymes (adaptors), and can even be activated themselves through post-translational modifications (PTM) by host enzymes.

Effector	Full name	SPI-1	SPI-2	PDB entry†	Biochemical activity	Interaction Partners	Functions	Related Effectors‡	References
AvrA	Avirulence gene A	Y	Y	6BE0	Acetyltransferase	ERK2, MKK4*/7*, P53	Inhibit inflammation and apoptosis	YopJ	(Labriola et al., 2018)
CigR		ND	ND	4EW5 (C-term. domain) (unpublished)	Poorly defined		Anti-virulence effector inhibiting replication and SCV development		(Kidwai et al., 2013)
GogB	Gifsy-one-gene B	Y	Y	-	Adaptor protein	SKP1, FBXO22	Inhibits NF-κB signalling	SspH1/2, SlrP	(Pilar et al., 2012)
GtgA (hom. PipA, GogA)	Gifsy-two-gene A	Y	Y	6GGR	Zinc metalloprotease	Class II NF-κBs (p65, RelB and cRel)*	Inhibits NF-κB signalling	NleC	(Jennings et al., 2018)
GtgE	Gifsy-two-gene E	Y	Y	5KDG 5OED (C45A mut.)	Cysteine protease	RAB32*/29*/38*	Prevents Rab accumulation on SCV and SITs		(Wachtel et al., 2018; C. Xu et al., 2016)
PipB	Pathogenicity island-encoded proteins A-B2	N	Y	-	Poorly defined	PDZD8		PipB2	(Walch et al., 2021)
PipB2		Y	Y	2LEZ (NMR) (unpublished)	Poorly defined	Kinesin-1, KIF5B, annexin A2	Recruits kinesin-1 to the SCV to reorganise late endosome/lysosomes to promote bacterial survival	PipB	(Knuff-Janzen et al., 2020)
SifA	*Salmonella* induced filament proteins A & B	ND	Y	3CXB (w. SKIP)	Adaptor protein	PLEKHM1/2, GDP-RhoA, RAB7, caspase-3	Induces SIT formation, detoxifies lysosomes, recruits late endosomes and lysosomes to SCV, maintain vacuolar membrane stability	SifB, SopE	(M. B. Ohlson et al., 2008)
SifB		N	Y	-	Poorly defined			SifA	(Brink et al., 2018)
SipA (SspA)	*Salmonella* invasion proteins A-D	Y	N	2FM9	Poorly defined	Caspase-3, F-actin, T-plastin, syntaxin8	Actin assembly, disruption of tight junctions, SCV positioning		(Lilic et al., 2006)
SipB		Y	N	3TUL (N-term)	Translocation pore	Caspase-1	Pyroptosis		(Barta et al., 2012)

SipC		Y		-	Translocation pore	Cytokeratin 8/18, EXO70, F-actin, syntaxin 6	Actin nucleation, SCV maturation		(Gong et al., 2010)
SipD		N	Y	3NZZ, 3O00	Poorly defined		Bind bile salts, possible bacterial sensor		(Chatterjee et al., 2011)
SlrP	*Salmonella* leucine-rich repeat protein	Y	Y	4PUF (w. thioredoxin)	E3 ubiquitin ligase	Thioredoxin*, SNRPD2*, ERdj3, UbcH5b	Inhibits release of IL-1beta for apoptosis, antigen presentation of dendritic cells and inflammasome activation	SspH1/2, IpaH family	(Zouhir et al., 2014)
SopA	*Salmonella* outer proteins A-F	Y	N	2QYU	E3 ubiquitin ligase	TRIM56/65, UbcH5a/5c/7, HsRMA1, Caspase-3	Invasion, escape from SCV, PMN migration	NleL	(Diao et al., 2008)
SopB (SigD)		Y	N	4DID (w. Cdc42)	Phosphoinositide phosphatase	CDC42	Invasion, nuclear responses, SCV maturation, fluid secretion	IpgD	(Burkinshaw et al., 2012)
SopD		Y	Y	5CPC	GAP & GEF	RAB8*/10*	Invasion, inflammation, fluid secretion		(D'Costa et al., 2015)
SopD2		ND	Y	5CQ9	GAP	RAB7*/32*, ANXA2	Disrupt host-driven regulation of microtubule motors, impairs trafficking of endocytic cargo to lysosomes for degradation		(D'Costa et al., 2015)
SopE		Y	N	1GZS	GEF	CDC42*, RAC1*, RAB5*	Actin remodelling, phagosome-early endosome fusion, inflammation		(Buchwald et al., 2002)
SopE2		Y	N	1R6E, 1R9K (NMR)	GEF	CDC42*, RAC1	Actin remodelling, inflammation		(Williams et al., 2004)
SopF		Y	ND	7DN8 (w. ARF1), 7DN9 (w. NAD&ARF1)	ADP ribosyltransferase	ATP6V0C*, ARF1	Preventing antibacterial autophagy (xenophagy)		(Y. Xu et al., 2022)
SptP	*Salmonella* protein tyrosine phosphatase	Y	N	1G4W (w. Cdc42) 1JYO (GAP domain w. chaperone)	GAP & tyrosine phosphatase	Cdc42*, RAC1*, VCP*, vimentin*, cSRC*, NSF*, Syk*	Reversion of actin reorganisation, inhibition of ERK activation	YopE/H, ExoS	(Stebbins & Galán, 2000, 2001)
SpvB		ND	Y	2GWJ-M	ADP-ribosyl transferase	G-actin*	Inhibits F-actin polymerisation, modifies actin to cause cytoskeletal disruption and		(Margarit et al., 2006)

	Salmonella plasmid virulence B-D						apoptosis, promotes macrophage apoptosis and P-body disassembly		
SpvC		Y	Y	2P1W (w. peptide), 2Q8Y (w. MAPK7), 2Z8M-P 4H43 (H106N mut.), 4HAH	Phosphothreonine lyase	ERK*, p38*, JNK*	Inhibits MAPK signalling, suppresses inflammation in spleen and liver to facilitate bacterial growth	OspF, VirA, HopAI	(Chen et al., 2008; Zhu et al., 2007)
SpvD		Y	Y	5LQ6 (R161 var.), 5LQ7 (G161 var.) (29)	Cysteine hydrolase	Exportin-2*	Inhibits NF-kappaB signalling and interferes with host immune signalling to promote virulence	OspI, AvrPphB	(Grabe et al., 2016)
SrfJ	SsrB regulated factor J	Y	Y	2WNW	Putative glycoside hydrolase				(Yeon-Gil et al., 2009)
SseB	*Salmonella* secreted effectors B-L	N	Y	-	Undefined				
SseC		N	Y	-	Undefined				
SseD		N	Y	-	Undefined				
SseF		ND	Y	-	Transmembrane protein / adaptor	SseG, ACBD3, RAB1A, plakoglobin, desmoplakin, TIP60	Localises, tethers and anchors SCV to golgi network to facilitate bacterial replication; inhibits RAB1-mediated autophagy	SseG	(Feng et al., 2018)
SseG		ND	Y	-	Transmembrane protein / adaptor	SseF, ACBD3, RAB1A, plakoglobin, desmoplakin, Caprin-1	Localises, tethers and anchors SCV to golgi network to facilitate bacterial replication; inhibits RAB1-mediated autophagy	SseF	(Feng et al., 2018)
SseI (SrfH)		N	Y	4G2B (catalytic domain), 4G29 (w. peptide)	Cysteine deamidase	GAI2*, IQGAP1	Inhibits directional migration of macrophages and DCs, promoting systemic infection	SpvD, OspI, AvrPphB	(Bhaskaran & Stebbins, 2012)

SseJ		N	Y	-	Acyltransferase	GTP-RhoA, Cholesterol*	Cholesterol esterification to modify SCV membrane, reduces cellular cholesterol to inhibit SifA-cholesterol interaction and LAMP-1 vesicles		(B. Ohlson et al., 2005)
SseK1		Y	Y	5H60 (w, UDP+Mn)	Glycosyltransferase	FADD*, TRADD*, RAB1*/5*/11*, SseK1*	Inhibits TNF-alpha-stimulated NF-kappaB signalling and necroptosis	SseK2/3, NleB effectors	(J. B. Park et al., 2018)
SseK2		ND	Y	5H61, 5H62 (w. UDP), 5H63 (w. UDP-GlcNAc)	Putative glycosyltransferase		Inhibits TNF-alpha-stimulated NF-kappaB signalling and necroptosis	SseK1/3, NleB effectors	(Esposito et al., 2018)
SseK3		ND	Y	6EYR, 6EYT (w. UDP-GlcNAc+Mn), 6CGI (w. UDP)	Glycosyltransferase	TNFR1*, TRAILR*, SseK3*, TRIM32	Inhibits TNF-alpha-stimulated NF-kappaB signalling and necroptosis	SseK1/2, NleB effectors	(Esposito et al., 2018)
SseL		ND	Y	5HAF, 5UBW (catalytic domain)	Deubiquitinase	K63 ubiquitin chains*, RPS3*, IkBα*, OSBP	Prevents accumulation of lipid droplets, inhibits autophagic clearance of cytosolic aggregates, induces late macrophage cell death	ChlaDub1/2, ElaD	(Pruneda et al., 2016)
SspH1	*Salmonella* secreted proteins H1-2	Y	Y	4NKH (LRR domain), 4NKG (LRR:PKN1 HR1b domain)	E3 ubiquitin ligase	PKN1*, Ube2D	Ubiquitinates host kinase PKN1 for degradation, suppresses NF-kappaB activation, inhibits androgen steroid receptor and macrophage activation	SspH2, IpaH effectors	(Keszei et al., 2014)
SspH2		N	Y	3G06	E3 ubiquitin ligase	NOD1*, SGT1, UbcH5-Ubiquitin	Activates NOD1 signalling	SspH1, IpaH effectors	(Quezada et al., n.d.)
SteA	*Salmonella* translocated effectors A-E	Y	Y	-	Poorly defined	PtdIns(4)P, Cullin-1	Regulates and partitions SCV vacuoles for bacterial growth		(Domingues et al., 2016)
SteB		Y	Y	-	Poorly defined				

SteC		N	Y	8JBI (N-lobe of C-domain, C276S mut., w. AMP+Mg)	Kinase	MEK1*, HSP27*, FMNL1/2*, MYL12A*	Induces formation of F-actin meshwork around SCV		(Dai et al., 2024)
SteD		N	Y	-	Adaptor protein	mMHCII, TMEM127	Inhibits antigen presentation and T cell activation, suppress adaptive immune responses, appears to act as an adaptor		(Alix et al., 2020)
SteE		Y	Y	-	Adaptor protein	GSK3α/β, STAT3	Transcriptional reprogramming towards anti-inflammatory phenotype		(Panagi et al., 2020)

Table 1. 1: List of *Salmonella* effectors

Summary of SPI-1 and SPI-2 effector protein functions with reference to key interaction partners and biochemical activity. Interaction partners identified in high-throughput screens that have not been verified by a second method have been excluded.

Rows are coloured as follows: green, enzymatic activity; orange, adaptor function; blue, poorly defined

† PDB files refer to structures obtained by crystallography unless otherwise stated

‡ Related effectors include those from *Salmonella* and other intracellular bacteria, listed here: AvrPphB from *Pseudomonas syringae*; *Chla*Dub1 and *Chla*Dub2 from *Chlamydia trachomatis*; ElaD from *E. coli*; ExoS is from *Pseudomonas aeruginosa*; HopAI from *Pseudomonas syringae*; IpaH effectors from *Shigella spp.* and enteroinvasive *E. coli* (EIEC); IpdG is from *Shigella spp.*; NleB effectors from Enteropathogenic *Escherichia coli* (EPEC); NleC is from *Escherichia coli* and *Citrobacter rodentium*; NleL is from *E. coli*; OspF from *Shigella flexneri*; OspI from *Shigella flexneri*; VirA from *Chromobacterium violaceum*; YopE is from *Yersinia spp.*; YopH is from *Yersinia spp.*; YopJ family is from *Yersinia enterocolitica*

* refers to interaction partners that are confirmed substrates

Abbreviations: ADP, adenosine diphosphate; DC, dendritic cell; GAP, GTPase activating protein; GEF, guanine nucleotide exchange factor; GTPase, guanosine triphosphatase; ND, not discovered; SIT, *Salmonella*-induced tubule; TNF, tumour necrosis factor.

For example, SopA is a *Salmonella* effector with host-like ubiquitin E3 ligase activity. Despite low amino acid sequence homology to eukaryotic proteins, it demonstrates a bilobal architecture, an E2 binding site close to the substrate binding site, and conformational flexibility between domains, all classic of the eukaryotic HECT E3 ligases (Diao et al., 2008). Its E3 ligase activity, dependent on a catalytic cystine C753, is required for its pro-inflammatory effect in the gut (Y. Zhang et al., 2006). This occurs through its ubiquitination of TRIM56 and TRIM65 (Fiskin et al., 2017; Kamanova et al., 2016).

Glycosyltransferase effectors, of which the *Salmonella* SseK family are three, catalyse transfer of N-acetylglucosamine (GlcNAc) from uridine diphosphate (UDP)-GlcNAc onto substrates. Rather than targeting serine and threonine residues by O-glycosylation as performed by eukaryotic glycosyltransferases, these effectors target arginine residues forming an N-glycosidic linkage. This novel PTM, only described in bacterial effector proteins, cannot be reversed by host cell machinery (S. Li et al., 2013; Pearson et al., 2013; Scott et al., 2017). SseK family members share high sequence similarity but have distinct cellular localisations in infection as well as substrate specificity (Gunster et al., 2017). The distinctive substrates of SseK1 and SseK3 are determined by a single residue in the catalytic core, while definitive substrate identification for SseK2 remains elusive (García-García et al., 2021). This heterogeneity may have arisen through the evolutionary process of duplication and divergence.

Another enzyme class only described in pathogenic bacteria is the phosphothreonine lyase, of which SpvC, a *Salmonella* effector, is one. SpvC and other class members can irreversibly dephosphorylate their host substrates by severing the Cβ-Oɣ bond of the threonine residue, rather than the O-P bond targeted by conventional phosphatases, meaning the threonine residue can no longer be phosphorylated (H. Li et al., 2007). SpvC contains a substrate-binding domain with a D motif that mediates its interaction with MAP kinases, and its recognition motif is pT-x-pY present on ERK, p38 and JNK. This phosphothreonine lyase activity arrests downstream immune signalling, leading to reduced inflammatory responses and promoting systemic infection (Zhu et al., 2007).

Other effector proteins bind to host enzymes to reprogram them. SteE, a small protein of 157 amino acids, binds to the host serine-threonine kinase GSK3, driving a change in its amino acid specificity to tyrosine, allowing it to phosphorylate STAT3 on Y705, its canonical site of activation (Panagi et al., 2020).

SifA is an effector conserved across all *Salmonella* serovars, suggesting its critical importance. It stabilises the SCV to prevent rupture and the release of *Salmonella* into the host cytosol. Remarkably, SifA undergoes cleavage by the host caspase 3 to become active (Patel et al., 2019). It has no confirmed enzymatic activity, though demonstrates structural homology to SopE, a *Salmonella* GEF, which facilitates its interaction with RAB7 and RhoA GTPases, possibly having an adaptor mechanism. For more details of SifA activity, see Section 1.2.2.

Alongside the diversity of *Salmonella* effector functions described, the effector repertoires of *Salmonella* strains vary, with combinations of effectors being associated with distinct life cycles and disease phenotypes. For example, SseI, a cysteine deamidase, prevents dendritic cell chemotaxis (McLaughlin et al., 2009, 2014). When it is lost through pseudogenisation, for example in *S.* Typhimurium isolate D23580, *Salmonella* can spread from the gut to mesenteric lymph nodes, manifesting in a systemic bacteraemic disease phenotype in the host (Carden et al., 2017).

1.2 The actin cytoskeleton in *Salmonella* infection

1.2.1 Actin regulation in host cells

In eukaryotic cells, the cytoskeleton both maintains cell morphology and drives processes of cell movement and phagocytosis. It is made up of actin filaments, microtubules and intermediate filaments. Actin filaments are dynamic, undergoing constant polymerisation at one end, the 'barbed end', and depolymerisation at the other, the 'pointed end', a process known as treadmilling (Shekhar et al., 2016; Svitkina, 2018). The formation of polymerised *filamentous* actin (F-actin) from soluble monomeric *globular* actin (G-actin) is adenosine triphosphate (ATP)-dependent and requires a nucleation factor as it is not energetically favourable in the cytosolic environment (Pollard, 2016; Rottner et al., 2017). Moreover, the factors that stabilise, sever and depolymerise actin are themselves regulated by several factors, including calcium concentration, calmodulin, Rho GTPases, kinases and phosphatases (Figure 1.3).

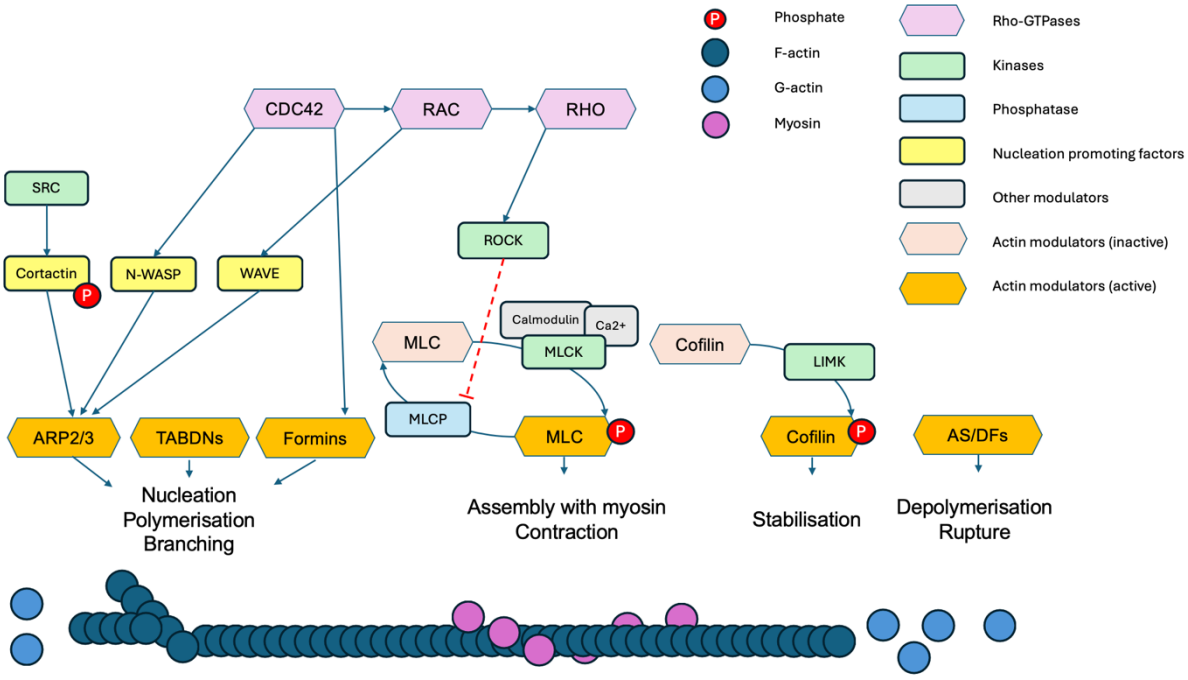


Figure 1. 3: Summary of the mechanisms of actin polymerisation, stabilisation and depolymerisation with associated regulatory factors

Actin polymerisation is promoted by ARP2/3, TABDNs and formins (left); regulated by MLC and Cofilin (middle) and depolymerised by AS/DFs (right). AS/DFs, actin severing or depolymerising factors; CDC42, cell division control protein 42; LIMK, Lim kinase; MLC, myosin light-chain; MLCK, myosin light-chain kinas; MLCP, myosin light-chain phosphatase; N-WASP, neuronal Wiskott-Aldrich syndrome protein; ROCK, Rho-associated protein kinase; TABDNs, tandem actin binding domain nucleators; WAVE, Wiskott-Aldrich syndrome protein-family verprolin homologous protein 1.

Three nucleation factor families have been identified. ARP2/3, a large 7-component protein complex, forms branched actin networks at the interface between the membrane and cytoplasm. It has roles in vesicular trafficking and cell motility (Papalazarou & Machesky, 2021), and requires the nucleation promoting factors N-WASP, WAVE and cortactin, regulated by the Rho-GTPase system and Src kinase, to promote its activity (Rotty et al., 2013). Tandem actin binding (TAB) domain nucleators (TABDNs) are a large family of proteins that initiate F-actin formation from G-actin via their TAB domain (Rottner et al., 2017). Formins dimerise head-to-tail to form a ring, trapping actin dimers and trimers, and binding F-actin's barbed end thereby modulating polymerisation rate. They can also bind along the length of actin filaments or microtubules, promoting cytoskeletal network bundling, coordination and disassembly. They too are regulated by Rho GTPases (Courtemanche, 2018).

Myosin light chains (MLC) provide another level of actin cytoskeletal regulation. They have many roles as calmodulin-family members of the myosin holoenzyme, but one is the stabilisation of F-actin-rich structures (Heissler & Sellers, 2014). They are activated by myosin light chain kinase (MLCK) and inhibited by the respective phosphatase (MLCP). MLCK is regulated by calcium and calmodulin, while MLCP is downregulated by Rho kinase (ROCK) (Amano et al., 2010), which itself is regulated by RhoA GTPase (Figure 1.3).

The actin depolymerising factor (ADF)/cofilin family are actin-binding proteins that can have the effect of severing, stabilising or nucleating actin branches depending on their concentration and that of G-actin and other factors. Lim kinases (LIMK), activated by ROCK, can phospho-activate cofilins (Bamburg et al., 2021). Other actin severing or depolymerising factors (AS/DFs) include gelsolin, villin and many others, each with their own regulatory system (Ono, 2007). Altogether, actin polymerisation is a dynamic process regulated through multi-layered signalling molecules and polymerisers, stabilisers and depolymerisers (Figure 1.3).

1.2.2 Manipulation of the actin cytoskeleton by intracellular bacteria

Intracellular bacteria can enter the host cell by being taken up by phagocytosing cells, e.g. macrophages (passive invasion) or by inducing their own uptake into non-phagocytosing cells (active invasion). To achieve the latter, they either engage host cell receptors via their own membrane proteins or translocate effector proteins into the host cytosol through their secretion system apparatus. Either process requires reorganisation of the host actin cytoskeleton (Colonne et al., 2016; de Souza Santos & Orth, 2015). In *Salmonella,* this process is primarily orchestrated through SPI-1 effectors (Figure 1.4, part 1). SipC is an effector protein with 3 putative domains, N- and C-terminal protein binding domains and a central hydrophobic region. It is an essential part of the T3SS translocation pore, which is the interface between

the tip of the T3SS and the inner surface of the host membrane (Myeni & Zhou, 2010). In this position it not only facilitates the translocation of effectors but also binds G-actin directly via its C-terminal tail (Myeni & Zhou, 2010), and nucleates it, a process requiring its self-association in dimers and trimers via its C-terminal domain (Chang et al., 2007).

SipA is a SPI-1 and SPI-2 effector that binds actin, again via its C-terminal domain. It displaces host actin binding proteins gelsolin and ADF/cofilin to prevent F-actin severing and depolymerisation, respectively (McGhie et al., 2004). Its N-terminal domain, which like SifA is released from the C-terminus by host caspase-3 cleavage, interacts with Syntaxin8 to promote fusion of the forming SCV with early endosomes. In so doing, SipA enhances the efficiency of *Salmonella* entry (Kaniga et al., 1995). SopE and SopE2 are 69% identical SPI-1 effector GEFs that activate Rho GTPases, leading to actin remodelling (Friebel & Hardt, 2000). Their divergent host Rho GTPase specificity, SopE for CDC42 and RAC1 and SopE2 only for CDC42, may explain their conservation despite their similarity (Friebel et al., 2001). To facilitate cell recovery after entry, *Salmonella* translocates SptP. SptP is a dual-specificity effector protein with a N-terminal GAP domain and C-terminal tyrosine phosphatase domain. Its GAP domain, containing a conserved Arg finger mimicking host GAPs, inhibits RAC and CDC42, leading to the downregulation of MAP kinases and actin polymerisation and recovery of the actin cytoskeleton (Fu & Galán, 1999). Though SopE and SptP are translocated in similar amounts and at the same time, SopE is barely detectable after 30 minutes of infection, whereas SptP is maintained in the host cytosol hours later. This has been proposed as a mechanism of timing, where SopE antagonises SptP activity until entry is complete, at which point SptP's actin depolymerising effects can proceed unchecked (Kubori & Galán, 2003).

Once inside the cytosol, actin manipulation remains important for intracellular bacteria to survive (Figure 1.4, part 2). In *Salmonella*-infected cells, latrunculin B, an actin depolymerising drug, leads to the rupture of the SCV, demonstrating the importance of F-actin in SCV stability (Méresse et al., 2001a). *Salmonella* effectors SifA and SipA, both previously introduced and with no confirmed enzymatic activities, are essential for the stabilisation of the SCV, preventing rupture and release of bacteria into the cytosol, and localising the SCV to the perinuclear region of the cell (Beuzón et al., 2000; Brawn et al., 2007). However, actin polymerisation is seen associated with the SCV in a dense focal or capsular pattern that is not dependent on either SifA or SipA, but rather SteC, a SPI-2 effector kinase (Poh et al., 2008). The role of SteC is discussed further in Section 1.4.

Intracellular bacteria can leave the cell through cell-to-cell spread, cell lysis or extrusion of bacteria-containing compartments. Again, all these processes require the manipulation of

actin. In the case of *Salmonella*, SPI-2 effector SpvB, with a C-terminal actin-targeting ADP-ribosyltransferase domain, inhibits F-actin polymerisation by preventing ATP hydrolysis in the actin ATP-binding pocket. This disrupts the cytoskeleton and promotes apoptosis of the macrophage from which it is propagating (Lesnick et al., 2001) (Figure 1.4, part 3). Interestingly, its N-terminal domain independently binds clathrin, the host molecular scaffold for vesicular uptake of cargo at the plasma membrane. Clathrin is not a substrate for its ADP-ribosyltransferase domain, and the interaction does not lead to actin rearrangement (Yuan et al., 2023). In summary, bacteria have evolved a number of mechanisms to manipulate the actin cytoskeleton to create a replicative niche.

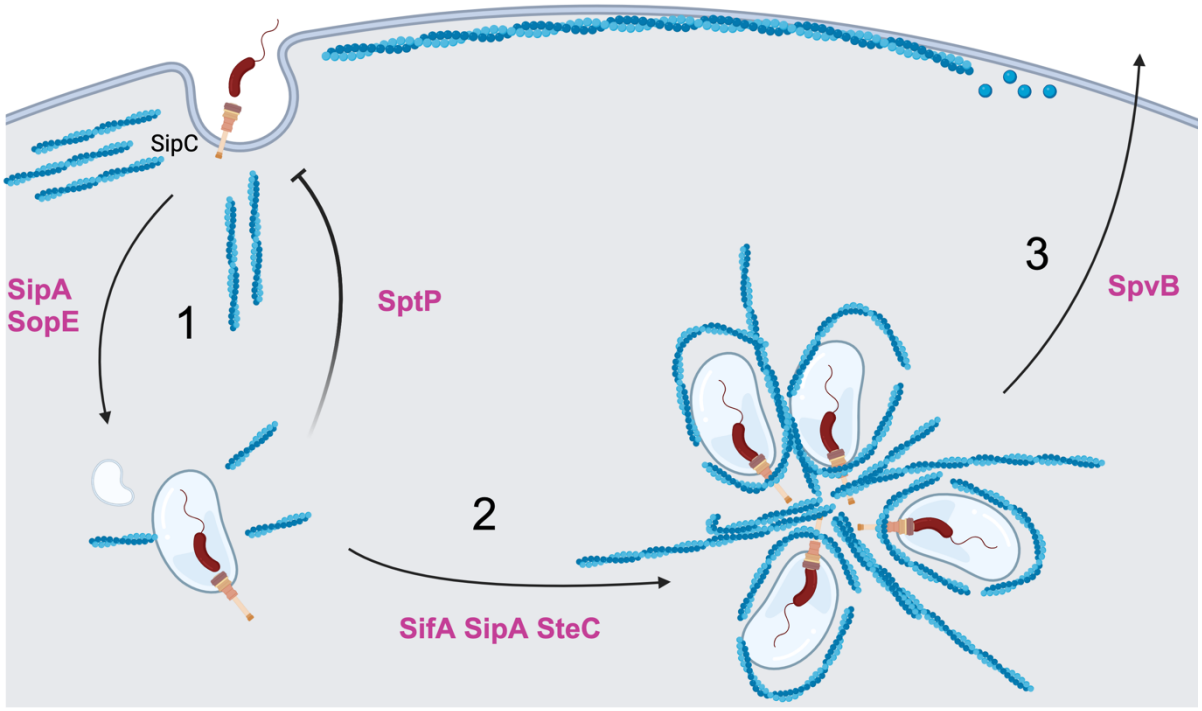


Figure 1. 4: *Salmonella* effector proteins manipulate the host actin cytoskeleton in all phases of the intracellular lifecycle

Salmonella bacteria (mauve) manipulate the actin cytoskeleton (blue) of the host cell (grey) to 1) enter host cells; 2) survive within the *Salmonella* containing vacuole; and 3) exit the cell. This is achieved through translocating effector proteins (named in pink bold) into the host cytosol using the type 3 secretion system (orange). Figure made with BioRender.

1.3 Kinase structure and function

Kinases are key regulators of the actin cytoskeleton (Figure 1.3). It is apt, therefore, that *Salmonella* recruits its only effector kinase, SteC, to manipulate the host actin cytoskeleton for its own ends. To understand how SteC functions, the principles of kinase enzymatic activity and regulation require exploration.

1.3.1 Phosphorylation and its role in cells

Phosphorylation is the archetypal PTM, and an essential language of cell signalling controlling nearly every cellular process in eukaryotic cells (Roskoski, 2015). Protein kinases catalyse phosphorylation by binding a molecule of adenosine triphosphate (ATP), two divalent cations (usually Mg^{2+}) and a substrate, bringing about the transfer of the ɣ-phosphate of ATP onto the target residue of the substrate (Hanks & Hunter, 1995). In chemical terms, phosphoryl-transfer is a dissociative elimination-addition reaction: dissociative because nucleophilic attack by the hydroxyl group of the substrate on the phosphorous atom of the ATP ɣ-phosphate leads to a transition state; and elimination-addition because the ɣ-phosphate is eliminated from ATP and added to the target amino acid of the substrate. The addition of a phosphate group onto a protein substrate has downstream effects by either creating a recognition motif for other domains or adding charge which can for example have allosteric effects intramolecularly (Ardito et al., 2017). Phosphatases reverse this process, by removing phosphate groups from substrates. The simplicity, energy efficiency and reversibility of this biochemical mechanism may provide its evolutionary advantage in the eukaryotic context, where rapid signal transduction in response to extracellular cues allows for coordination of multicellular processes (Ardito et al., 2017).

1.3.2 Conserved eukaryotic protein kinase features

Since the first description of the enzymatic activity of kinases (Burnett & Kennedy, 1954) a huge body of work has explored the key catalytic motifs, structural features, regulatory mechanisms, substrate specificities and signalling pathways of kinases. The kinase family is one of the largest enzyme families in eukaryotes, with at least 518 protein kinases identified in humans (Manning et al., 2002). Eukaryotic protein kinases, which I will refer to as kinases, can be broadly grouped by their target residue: tyrosine kinases, serine-threonine kinases or dual-specificity kinases, all sharing core catalytic features (Hanks & Hunter, 1995). The proportion of phosphosites in the human phosphoproteome by tyrosine, threonine and serine residue is 1.8%, 11.8% and 86.4%, respectively (Olsen et al., 2006). Prokaryotic kinases will be introduced in Section 1.3.6. Several kinase features are conserved across the eukaryotic kinome, each with roles in catalysing the phosphoryl-transfer reaction (Figure 1.5).

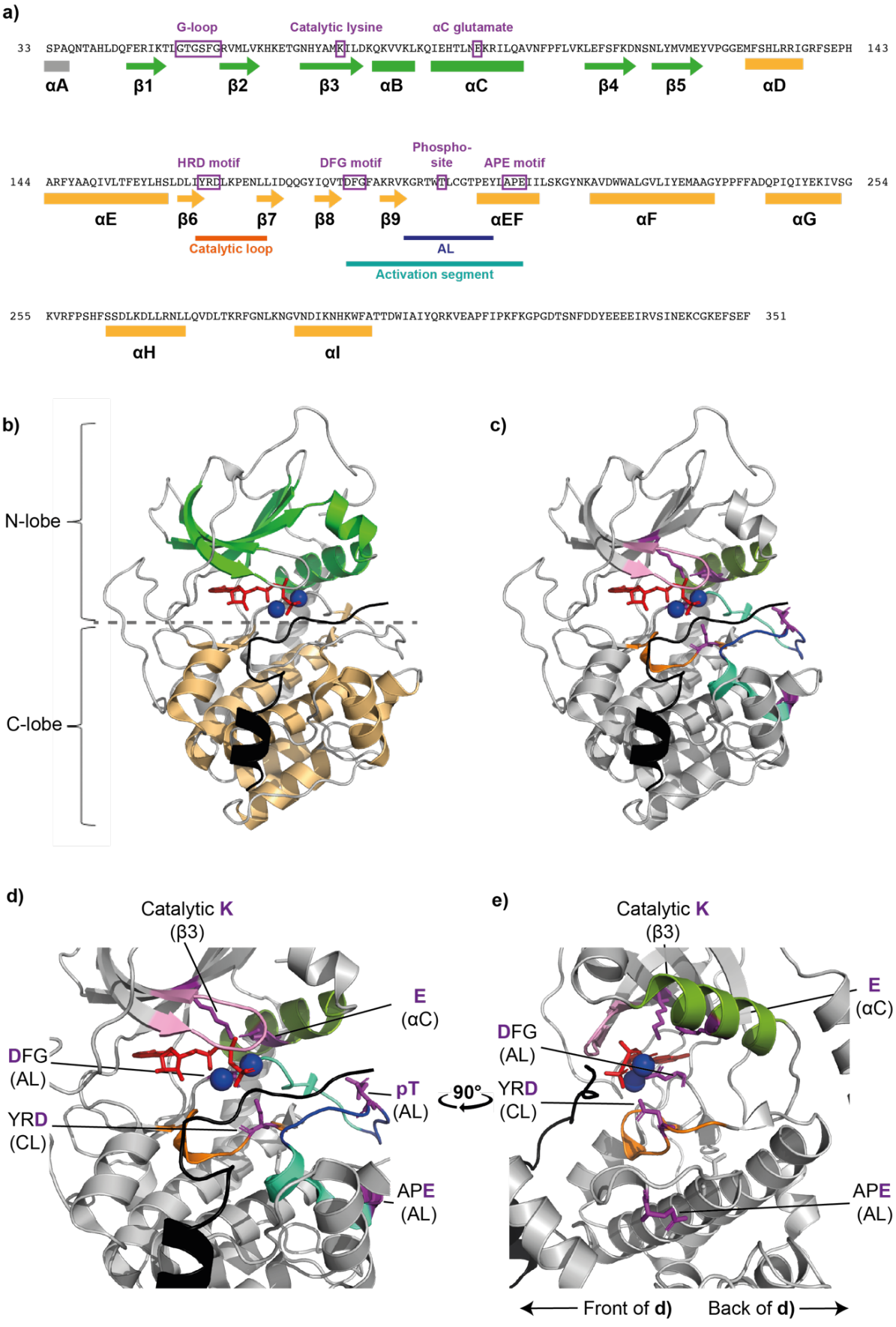

a)
G-loop
Catalytic lysine
αC glutamate
33 SPAQNTAHLDQFERIKTLGTGSFGRVMLVKHKETGNHYAMKILDKQKVVKLKQIEHTLNEKRILQAVNFPFLVKLEFSFKDNSNLYMVMEYVPGGEMFSHLRRIGRFSEPH 143
αA β1 β2 β3 αB αC β4 β5 αD
HRD motif
DFG motif
Phospho-site
APE motif
144 ARFYAAQIVLTFEYLHSLDLIYRDLKPENLLIDQQGYIQVTDFGFAKRVKGRTWTLCGTPEYIAPEIILSKGYNKAVDWWALGVLIYEMAAGYPPFFADQPIQIYEKIVSG 254
αE β6 β7 β8 β9 αEF αF αG
Catalytic loop
AL
Activation segment
255 KVRFPSHFSSDLKDLLRNLLQVDLTKRFGNLKNGVNDIKNHKWFATTDWIAIYQRKVEAPFIPKFKGPGDTSNFDDYEEEEIRVSINEKCGKEFSEF 351
αH αI
b)
N-lobe
C-lobe
c)
d)
Catalytic K (β3)
E (αC)
DFG (AL)
pT (AL)
YRD (CL)
APE (AL)
e)
Catalytic K (β3)
E (αC)
DFG (AL)
YRD (CL)
APE (AL)
90°
Front of d)
Back of d)

Figure 1. 5: Key kinase features using PKA, crystallised in the active conformation

PKA amino acid sequence, structural features and motifs are demonstrated based on the crystal structure (PDB: 1ATP) and kinase literature.

a) PKA amino acid sequence (residues 33-351) is demonstrated annotated with secondary structural features based on the crystal structure from αA-I and β1-8 (N-lobe in green; C-lobe in light orange); key catalytic motifs (purple); catalytic loop (red orange); activation segment (turquoise) and activation loop (blue). αA is partially covered by this map. Adapted from Sheetz & Lemmon (2022).
b) Ribbon demonstration of the PKA structure, with secondary structural elements in the N-lobe (green), larger C-lobe (light orange), loops (grey), catalytic cleft between (dotted line), ATP (red), Mg^{2+} cations (blue) and substrate peptide (black).
c) Key functional motifs are highlighted on PKA, including key residues in stick form (purple), glycine-rich loop (pink), αC helix (green), CL (orange), activation segment (turquoise), AL (blue), ATP (red), Mg^{2+} cations (blue), substrate peptide (black) and other residues (grey).
d) Zoom into the PKA catalytic site from b), with the same colours. Note, catalytic lysine (β3), glutamic acid (αC) and aspartic acid (YRD motif in CL) are oriented towards the nucleotide and Mg^{2+}.
e) 90° rotation in y-axis compared to c), with the activation segment from F185-P207 removed to visualise the catalytic site. D184 in DFG motif in the AL is now visible behind the divalent cations.

Abbreviations: AL, activation loop; CL, catalytic loop; pT, phosphothreonine.

Kinase domains range from 250-300 amino acids in length and are divided into **N-** and **C-lobes** (Figure 1.5a & b). The **N-lobe** tends to be smaller and contains the ATP binding site. The **C-lobe**, usually larger and predominantly α-helical, binds the substrate, orienting it towards the catalytic cleft, between the two lobes. By convention, the α-helices of a kinase domain are denoted with letters and the β-strands with numbers, i.e. αA-I and β1-8 (Figure 1.5a) (Hanks & Hunter, 1995).

Kinase motifs, loops and segments have been annotated on the amino acid sequence (Figure 1.5a) and the crystal structure of PKA (Figure 1.5b-e). The **glycine-rich loop**, in the turn between β1 and β2, anchors ATP (Lauber et al., 2016). The **catalytic lysine** on β3 in the ATP-binding pocket is essential due to its role in orientating ATP and transferring the ɣ-phosphate onto the substrate residue (Carrera et al., 1993). In the active conformation, a conserved **glutamate** on the **αC helix** forms a salt bridge with the **catalytic lysine**, and unstructured parts of this region gain helical structure (Gógl et al., 2019).

The **HxD motif**, including the **catalytic aspartic acid**, makes up the first residues of the **catalytic loop** (Figure 1.5c-e). It plays a role in polarisation of the substrate hydroxyl group for nucleophilic attack (Reinhardt & Leonard, 2023), and transfer of the ɣ-phosphate of ATP onto the substrate (Hanks & Hunter, 1995). In the case of some kinases, including PKA, the

His is replaced with a Tyr. Furthermore, the middle amino acid is Arg, making PKA an 'RD' kinase (see section 1.3.3 for more details).

The activation segment starts at the **DFG motif**, contains the β9 and **activation loop** and ends with the **APE motif** (Figure 1.5a). The **DFG motif** (Figure 1.5d) coordinates divalent cations, usually Mg^{2+}, through its aspartic acid oriented towards the ATP binding pocket (Hanks & Hunter, 1995; Knighton et al., 1991). The DFG motif can be in two positions, DFG-in or DFG-out. In the DFG-out position, the hydrophobic spines, key to the structural integrity of the active kinase domain and explained below, are disrupted. Whereas, in the DFG-in conformation, an active kinase conformation of the spines can arise (Hari et al., 2013; Taylor & Kornev, 2011).

The **activation loop**, a central part of the activation segment, spans the N- and C-lobes (Hanks & Hunter, 1995) (Figure 1.5c). As its name suggests, it is an important site of allosteric kinase regulation including phosphorylation, as is the case for PKA at T197 (Figure 1.5d) (see Section 1.3.3 for more details). The **APE motif** (Figure 1.5e), which anchors the activation loop to the structure of the F-helix in the C-lobe, has roles in substrate binding and kinase homodimer formation (Gógl et al., 2019).

To create kinase-inactive mutants different approaches have been used: mutation of catalytic Lys residue to Met to abrogate ATP binding; mutation of the HxD aspartic acid to Asn to prevent polarisation of the substrate hydroxyl and block catalysis; or mutation of the activation loop DFG Asp to Ala to abrogate magnesium and ATP binding. The second option (mutation of the HxD motif) is deemed the least structurally disruptive (Reinhardt & Leonard, 2023).

Finally, two conformational features common to active kinases require mentioning. The R and C spines comprise hydrophobic residues spanning the N- and C-lobes. The R spine is comprised of residues from β4, αC and HRD motif. The C spine is comprised of 8 residues from the β7, αD and αF, and completed by the adenine ring of ATP. These spines are coordinated by αF in the C-lobe, and their alignment is thought to be a key factor stabilising the active conformation of the kinase domain (J. Kim et al., 2017; Kornev et al., 2008).

1.3.3 Kinase regulation

Generally, kinases are expressed in an inactive form in the cell, such that the need for subsequent activation prevents aberrant downstream signalling and allows for a fine-tuned coordination of kinase signalling pathways (Pinna, 2003). The active conformation is relatively conserved across kinases, yet there is substantial variation in inactive structural conformations (Huse & Kuriyan, 2002). This may be due to the evolutionary demand for optimal catalytic efficiency when active, along with requirements for a diversity of regulatory mechanisms to allow for a signalling network exquisitely responsive to a range of cues.

The activation loop, which has the highest sequence divergence in the eukaryotic kinome (Modi & Dunbrack, 2019), can exist in a stabilised inhibitory position (that prevents substrate binding), or an open position. Phosphorylation of the activation loop is, perhaps, the best described mechanism of activation, sometimes referred to as the 'phosphoswitch' (Philpott et al., 2020). Phosphorylation can be by the kinase itself (autophosphorylation) or in trans (by another kinase). Autophosphorylation can be in cis (intramolecular autophosphorylation) or trans (another protomer of the same kinase) (Cobbaut et al., 2018). On phosphorylation, a basic pocket is formed by the Arg from HRD motif in the catalytic loop and a Lys in the activation loop close to the DFG motif. This houses the phosphorylated residue of the activation loop, moving the activation loop from a position inhibiting substrate binding to one that allows it (L. Johnson & Lewis, 2001). Substrate binding affinity is not the only determinant of enzyme efficiency, given the need to also dissociate from the modified substrate prior to binding the next unmodified substrate molecule, and the other essential catalytic steps required. Indeed, substrate binding affinity changes on phosphorylation for some phospho-activated kinases, but not others (Adams, 2003). Alternatively, ATP binding may be enhanced, though it is unclear whether possible improvements in affinity on activation loop phosphorylation would have a meaningful impact on enzymatic activity in an environment where ATP is present at millimolar concentration. A third model involves phosphorylation of the activation loop leading to αC helix reorientation with improvement in Glu position (Adams, 2003).

Phosphorylation can also inactivate kinases. For example, CDK1 is inhibited by WEE1 kinase phosphorylation on its glycine-rich loop, preventing substrate access to the catalytic cleft (Russell & Nurse, 1987), a process that can be reversed by phosphatase CDC25 (Frazer & Young, 2012). Another example is GSK3, which is constitutively active and can be inactivated by PKA and PKB by phosphorylation of serine 21 or 9 of GSK-3α and GSK-3β, respectively (Fang et al., 2000).

Inactive kinases can also be activated through protein-protein interactions. This can be interaction with another domain of the same protein – as in cyclin B (Petri et al., 2007) – or, another protomer of the same kinase (i.e. dimerisation) (Lavoie et al., 2014). Alternatively it can be another protein binding partner, for example Aurora B activation by INCENP (Sessa et al., 2005). These binding events can allosterically align key catalytic residues; bring the C- and R-spines into alignment; or expel an inhibitory pseudosubstrate allowing the true substrate to bind (Lavoie et al., 2014a). One of the many mechanisms by which pseudokinases – kinase-like proteins that have lost their catalytic ability – function is through dimerisation with their enzyme-active paralogous counterparts (Lavoie et al., 2014a).

1.3.4 Mechanisms of kinase specificity

Once activated, kinases must target specific protein sites for phosphorylation if they are to act as signalling molecules. This can be achieved by either target residue specificity, or specificity to the substrate per se. Residue specificity is achieved through the depth and local characteristics of the catalytic cleft, determining if a kinase can accommodate a tyrosine residue, serine-threonine residues or both. More broadly, local interactions between the kinase C-lobe and the amino acids either side of the target residue, (-5 to +4), especially in the -5 to -1 positions determine what substrate motif specificity (Figure 1.6), possibly because the target sequences are often in solvent-accessible flexible loops on the substrate, meaning there aren't other nearby residues. Kinases can thus be classified according to their substrate motif. Interestingly, acidophilic kinases require substrate priming by another kinase prior to recognition (J. Johnson et al., 2023) (Figure 1.6), the equivalent of an AND gate in signalling terminology. Furthermore, kinase motifs overlap with specific phosphate-binding domains of non-kinases, coupling the kinase and its downstream signalling molecule, via its substrate (Yaffe & Smerdon, 2001).

Other ways of achieving substrate specificity include docking sites distant from the substrate motif that can facilitate or block protein substrates from access to the cleft; subcellular localisation of the kinase; formation of complexes with other proteins; competition for the substrate motif by other proteins; and expression which may only occur in certain cell types or in response to certain stimuli (Ubersax & Ferrell, 2007).

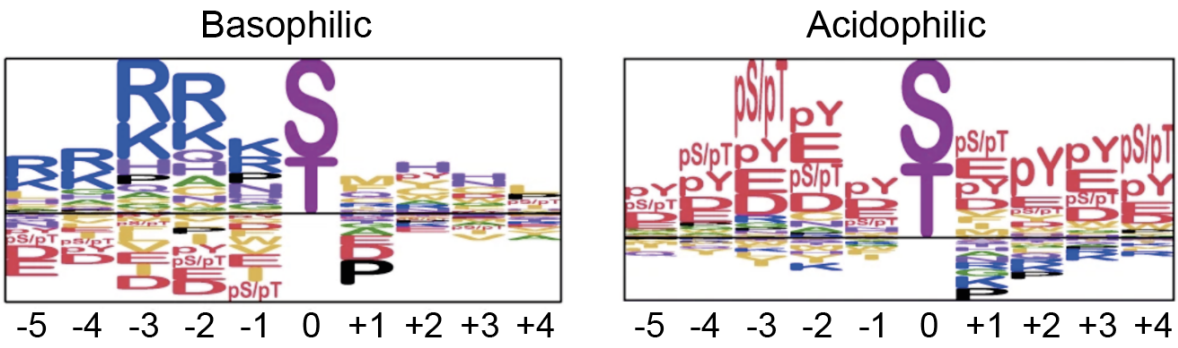


Figure 1. 6: Kinase recognition motifs

Basophilic kinases have a preference towards phosphorylating serine or threonine residues with basic residues in the -5 to -1 positions, as demonstrated by the sequence logo of the known substrate motifs of basophilic kinases. Meanwhile acidophilic kinases phosphorylate serine or threonine residues if surrounded by previously phosphorylated serine, threonine or tyrosine residues or other negatively charged residues such as aspartic acid or glutamic acid in the -5 to +4 positions. Adapted from J. Johnson et al. (2023) using WebLogo (https://weblogo.berkeley.edu/logo.cgi).

1.3.5 Kinase inhibitors

The centrality of kinases to eukaryotic cell signalling processes and the depth of their structural and biochemical study have made them a focus of drug discovery recently. This pursuit was boosted by the success of the first kinase inhibitor, imatinib which targets Abelson (Abl) tyrosine kinase, approved in 2001 for the treatment of chronic myeloid leukaemia (Cohen et al., 2021). This paved the way for over 70 approved kinase inhibitors to date. Most inhibitors commonly work by outcompeting ATP in the kinase ATP-binding pocket in the kinase-active DFG-in conformation (Type I inhibitors, the most common). These inhibitors have problems with specificity due to the similarity in the active conformation ATP-binding pocket. Alternative approaches have been deemed more specific by targeting more variable conformations and sites: either competing with ATP by binding to an inactive and therefore more structurally variable ATP binding pocket (Type II inhibitors) or binding distant to the ATP binding pocket and triggering an allosteric change that inactivates the kinase (Type III and IV) (J. Li et al., 2024). Bacterial kinases have become targets for antimicrobial drug discovery, though no bacterial kinase inhibitors are currently licensed (King & Blackledge, 2021).

1.3.6 Prokaryotic kinases

Kinases are essential for eukaryotic cell function, and their prevalence and complexity are higher in eukaryotes, including plants (Shankar et al., 2015) and fungi (Albataineh & Kadosh, 2016), than prokaryotes (Houles et al., 2024). In prokaryotes, membrane-bound histidine kinases have been described as part of the two-component system, that sense external stimuli

leading to activation and phosphorylation of intracellular response regulators (Mascher et al., 2006). However, many kinases have now been identified in bacteria, including serine/threonine, tyrosine, histidine and aspartate kinases (Kannan et al., 2007). In an analysis of kinase sequences identified in the Global Ocean Sampling database, the prokaryotic kinase families identified generally followed the same sequence motif rules as eukaryotes. However, for each of the 'canonical' motifs described in section 1.3.2, there was at least one example of a prokaryotic kinase family that had alternative residues in that position. In these cases, compensatory changes could often be seen in other residues in the catalytic site (Kannan et al., 2007).

One important difference is the DFG motif, which appears to have become essential for eukaryotic protein kinases but is not seen in many prokaryotic families (Kannan et al., 2007). Interestingly, Kannan *et al.* hypothesise that eukaryotic kinases have evolved more conformational flexibility than prokaryotic kinases and this may be due to the selection advantage of more tuneable regulatory mechanisms within the eukaryotic cell.

1.3.7 Bacterial effector kinases

A subset of bacterial effector proteins displays kinase activity. They commonly manipulate the cytoskeleton of the host or inhibit cell-intrinsic immune mechanisms (St Louis et al., 2023). Bacterial effector kinases are usually discovered by sequence or structural homology. Recently, with the development of accurate structural prediction tools, such as AlphaFold2 (referred to as AF), novel bacterial effector kinases have been discovered that lack sequence homology but share structural characteristics with known kinases (Louis et al., 2023). A common feature in effector kinases is a depleted C-lobe compared to the canonical kinase fold. Some bacterial effector kinases with depleted C-lobes require host binding partners to become stable and active. For example, OspG, the *Shigella* effector kinase, requires binding to a ubiquitin-E2 conjugate to become active (Pruneda et al., 2014). A list of reported bacterial effector kinases and categorisation of their putative C-lobes is presented in Table 1.2.

Name of effector	Organism	Stabilising host partner needed?	Stable binding partner	PDB of kinase domain	Depleted C-lobe?	Host Substrates	Targeted Host Pathways	References
SteC	*Salmonella enterica*	ND	ND	8JBI	Yes	MEK, FMNL1, HSP27, MYL12A, SERBP1	Cytoskeleton Rearrangement	(Birk et al., 2024; Dai et al., 2024; Odendall et al., 2012; Poh et al., 2008; Walch et al., 2021)
OspG	*Shigella flexineri*	Yes	Ubiquitin- UbcH5/7 complex	4BVU	Yes	ND	NFκB	(D. W. Kim et al., 2005; Pruneda et al., 2016)
YpkA (YopO)	*Yersinia pseudotuberculosis*	Yes	Actin	2H7V	No	Gαq, Otubain-1, gelsolin (+others)	Cytoskeleton Rearrangement	(Juris et al., 2000, 2006; Navarro et al., 2007)
HopBF1	*Pseudomonas syringae*	Yes	Host chaperone HSP90	-	Yes	ND	NFκB, Immune sensing	(Lopez et al., 2019)
LegK7	*Legionella pneumophila*	Yes	MOB1	6MCQ	No	MOB1	Hippo pathway	(Lee & Machner, 2018)
Lpg2603 (Lem28)	*Legionella pneumophila*	Yes	IP6	6VCC	No	ND	LCV	(Sreelatha et al., 2020)
SidJ	*Legionella pneumophila*	Yes	Calmodulin	7PPO	ND	ND	ND	(Sulpizio et al, 2019) (Gan et al, 2019)
PknG	*Mycobacterium tuberculosis*	Yes	RAB14	2PZI	No	TBC1D4	Lysosome trafficking; Autophagy	(P. Ge et al., 2022)
NleH1	*Escherichia coli*	ND	RPS3	4LRJ	Yes	CRKL, EPS8L2	NFκB, Micropillus formation	(Pham et al., 2013; Pollock et al., 2022)
NleH2	*Escherichia coli*	ND	RPS3	4LRK	Yes	CRKL, EPS8L2	NFκB, Micropillus formation	(Pham et al., 2013; Pollock et al., 2022)
LegK1	*Legionella pneumophila*	ND	ND	-	Yes	IκB	NFκB	(J. Ge et al., 2009)
LegK2	*Legionella pneumophila*	ND	ND	-	Yes	ARPC1B, ARP3	Cytoskeleton Rearrangement	(Céline et al., 2015)
LegK4	*Legionella pneumophila*	ND	ND	5CKW	Yes	HSP70	Host translation	(Moss et al., 2019)
SboH	*Salmonella bongori*	ND	ND	-	Yes	ND	ND	(Fookes et al., 2011)
YspK	*Yersinia enterocolitica*	ND	ND	-	No	E2 ubiquitin-conjugates	NFκB	(Matsumoto & Young, 2009)
XopC2	*Xanthomonas oryzae pv. oryzicola*	ND	ND	-	No	OSK1	Stomatal Closure	(S. Wang et al., 2021)
VopG	*Vibrio parahaemolyticus*	ND	ND	-	ND	ND	NFκB	(Plaza et al., 2021)
Stk1	*Staphylococcus aureus*	ND	ND	3M9G	Yes	ND	ND	(Michel et al., 2009)

Table 1. 2: List of bacterial effector kinases and their kinase features

Table content compiled together with Patricia Diaz Belamendia, MRes student, Thurston lab. Bacterial effector kinases identified in the literature are listed, and relevant literature summarised. Their PDB structure or AF prediction were used to class their C-lobe as either depleted (two α-helices or less) or not. LCV, *Legionella*-containing vacuole; ND, not determined.

1.4 *Salmonella* effector kinase, SteC

1.4.1 Identification of an effector kinase

The only *Salmonella enterica* effector kinase, SteC, was first discovered through insertions made by electroporation of transposon/transposase complexes into the *S.* Typhimurium strain 14028s genome (Geddes et al., 2005). SteC has an N-terminal regulatory domain and a C-terminal predicted kinase domain (Heggie et al., 2021) (Figure 1.7). It is a SPI-2 effector protein, meaning it is translocated into the host cytosol after *Salmonella* has entered the cell and resides in the SCV.

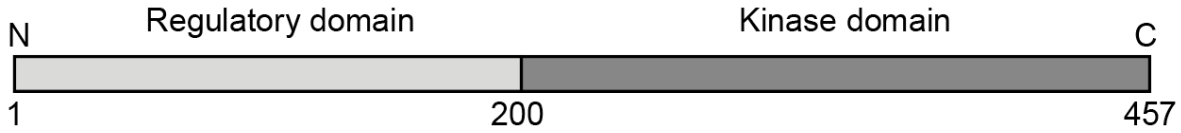


Figure 1. 7: Domain map of SteC

SteC is expressed in its intact form in all gastrointestinal serovars of *Salmonella* as well as Choleraesuis, Dublin and Typhi. It is disrupted in Gallinarum and Paratyphi (Jennings et al., 2017). Bacterial homologs outside of the *Salmonella* genus include proteins from *Yokenella* and *Cedecea* species, to which SteC has 41% amino acid sequence identity. Both these species are opportunistic human pathogens and the former is a cause of an enteric-like fever (Heggie et al., 2021). SteC's amino acid sequence has 28% homology to human RAF1 kinase (Poh et al., 2008). Sequence homology will be discussed in more detail in Chapter 3.

According to SteC-RAF1 sequence alignment, SteC K256 was hypothesised to be the catalytic lysine. Indeed, when mutated to His SteC's *in vitro* kinase activity against the model substrate myelin basic protein (MBP) is abolished (Poh et al., 2008). In Swiss 3T3 mouse embryonic fibroblast cells either transfected with tagged SteC or infected with *Salmonella* containing tagged SteC, SteC triggers F-actin polymerisation in a K256- dependent manner, suggesting that the actin polymerisation associated with intracellular *Salmonella* microcolonies at 8 hours of infection, is dependent on the kinase activity of SteC (Poh et al., 2008).

1.4.2 Reported substrates of SteC

Five substrates of SteC have been reported: MEK1, HSP27, FMNL1, MYL12A and SERBP1 (Birk et al., 2024; Dai et al., 2024; Imami et al., 2013; Odendall et al., 2012; Walch et al., 2021). As RAF1 activates MEK1, a human kinase in the MAPK pathway through phosphorylation at S218 and S222, Odendall et al. (2012) hypothesised that MEK1 might be a substrate of SteC.

Interestingly, *in vitro* kinase assays and mass spectrometry experiments revealed that SteC phosphorylates MEK1 at S200, a non-canonical site, which triggers MEK1 autophosphorylation at canonical sites S218 and S222, MEK1 activation and downstream ERK phosphorylation in a RAF1-independent manner (Odendall et al., 2012). Similarly, host kinase PAK triggers MEK1 auto-phosphorylation at S218 and S222 through phosphorylation of S298 (E. R. Park et al., 2007), suggesting a degree of plasticity in MEK1 autophosphorylation that SteC is exploiting. MEK1 and a downstream signalling pathway containing ERK, MLCK and myosin II were partially implicated in SteC-dependent F-actin polymerisation, but some activity remained even when this pathway was silenced with scrambled small interfering (si)RNA (Odendall et al., 2012), suggesting an alternative mechanism to F-actin polymerisation exists. At this point, MEK1 has not been validated as a substrate of SteC in cells, and whether phosphorylation of MEK1 by SteC at S200 activates the kinase in cells remains unclear.

In another study, parallel mass spectrometry experiments in HeLa cells transfected with HA-tagged SteC, one enriching the digested lysate for phosphopeptides and another identifying binding partners by immunoprecipitation mass spectrometry (MS), were performed. This identified HSP27 as a binding partner and substrate of SteC. HSP27 is a protein that, when active, binds G-actin monomers directly, preventing actin polymerisation (Doshi et al., 2009). This was further validated by detection of phosphorylated HSP27 in infection in an SteC-dependent manner, and an *in vitro* kinase assay of recombinantly expressed protein confirmed SteC phosphorylates HSP27 at 6 sites. Finally, actin polymerisation in HeLa cells stably expressing SteC was assessed with siRNA treatment targeting HSP27 and MEK1 inhibitors, suggesting that SteC-mediated actin polymerisation is partially dependent on HSP27, but unaffected by MEK1 inhibition (Imami et al., 2013).

In immunoprecipitation MS experiments from *Salmonella*-infected human cells with tagged effector proteins, formin-like proteins (FMNLs) were among the proteins identified as binding to SteC. FMNL1, 2 and 3 are members of the formin protein family (Higgs & Peterson, 2004), one of the three actin nucleators, as introduced in Section 1.2.1. This makes them an attractive alternative mediator of SteC-dependent actin polymerisation, though a canonical eukaryotic mechanism of formin phospho-activation has not been described. FMNL1 and 2 are phosphorylated when incubated with recombinantly expressed and purified SteC WT but not K256H, on several serine and threonine residues (Walch et al., 2021). Interestingly, in cells, data from transfection of SteC in FMNL2/3 knock-out Swiss 3T3s retained some actin polymerisation. The authors subsequently detected expression of FMNL1 in this cell line,

which they hypothesised to be the cause of the residual SteC-dependent effect (Walch et al., 2021).

In a recent study of *Salmonella* infection in macrophages, SteC was implicated in macrophage migration out of the epithelial layer. The MLC pathway, as introduced in Section 1.2.1, was explored as a potential mediator of this SteC-dependent effect using various drug inhibitors, which detected that the SteC-mediated macrophage migration was dependent on MLC phosphorylation not any of the upstream regulators, such as MEK, ERK, MLCK or ROCK. A yeast two-hybrid screen identified MYL12A, a myosin light chain protein component of the myosin holoenzyme. *In vitro*, SteC directly phosphorylates MYL12A. MYL12A phosphorylation is also seen in infection of macrophages. This substrate will be discussed in more detail in Chapter 4.

In summary, SteC-mediated phosphorylation of MEK1, HSP27, FMNL1 and MYL12A have all been implicated in the reorganisation of F-actin (Imami et al., 2013; Odendall et al., 2012; Walch et al., 2021), though caution must be applied as their relative contributions *in vivo* remain unclear.

Very recently, a study was published where thermal protein profiling was used to identify host proteins that change in thermal stability on *Salmonella* infection (Birk et al., 2024). Paired with publicly available *Salmonella*-host protein-protein interaction data (Walch et al., 2021), they identified that SERBP1, an SteC binding partner, changed in thermal stability on *Salmonella* infection of macrophages compared to macrophages exposed to a heat-killed *Salmonella* control. SERBP1 was also phosphorylated by SteC *in vitro* (Birk et al., 2024).

Using yeast as a model system to identify SteC binding partners, Cdc24 was found to bind to the N-terminal domain (amino acids 1-201) of SteC. The equivalent human protein of Cdc24 is VAV1, which has been shown to interact with SteC *in vitro*. Furthermore, the N-terminal domain of SteC inhibits the activation of CDC42 by its upstream GEF (Fernandez-Piñar et al., 2012). In a transfection experiment, the removal of the N-terminal domain of SteC does not affect actin polymerisation, but it does alter the subcellular localisation of the actin meshwork from membrane-localised to cytosolic (Poh et al., 2008). Intriguingly, VAV1 is a GEF for CDC42, which in its active form induces formins to polymerise actin. This raises the possibility of a SteC-VAV1-CDC42-FMNL1 interactome which would unify the activity of SteC enzymatic and non-enzymatic domains.

1.4.3 Role of SteC in infectivity

Attempts have been made to study the contribution of SteC to survival and replication of *S.* Typhimurium, with conflicting results. Competitive index experiments performed by intraperitoneal injection of mice demonstrate an advantage to bacteria that lack SteC over wild-type (Geddes et al., 2005; Poh et al., 2008). In mice inoculated by oral lavage, deemed to be a biologically relevant method of assessing replication, a similar effect was seen in one study (Odendall et al., 2012) while no significant effect was seen in another (Poh et al., 2008). In epithelial cells and macrophages there are also conflicting effects of SteC knock-out (Odendall et al., 2012; Buckner et al., 2011). Furthermore, in a signature-tagged mutagenesis screen, SteC (then known by its gene name *STM1698*) was identified as important for colonisation of the chick intestine (Morgan et al., 2004). Ultimately, SteC is likely to play an important role *in vivo* given the high survival cost of secreting an effector protein and its evident retention in a relatively broad range of serovars (Jennings et al., 2017). Studies to date have not explored SteC's contribution to morbidity and mortality in NTS diarrhoeal and non-diarrhoeal illness.

1.5 Objectives

With the clear need to better understand and treat pathogenic bacterial infections, there is an increasing interest in deciphering the molecular and cellular mechanisms of bacterial effectors. The conflicting reports regarding the effects of the *Salmonella* kinase SteC have highlighted a need for deeper understanding of its structure, enzymatic function and role in the host cell. To this end, this thesis focuses on four objectives.

First, I aimed to develop a model describing the molecular mechanism of SteC's kinase activity based on the amino acid sequence and available bioinformatic tools (Chapter 3, Section 3.1). I then recombinantly expressed and purified SteC with the aim of assessing its biophysical and structural features, and its interaction with a substrate (Chapter 3, Section 3.2-3.3). Thirdly, I aimed to biochemically determine the key aspects of SteC's enzymatic activity and activation using *in vitro* kinase assays (Chapter 4). Fourthly, to validate the described *in vitro* mechanisms, the effect of SteC on actin polymerisation was studied with the use of ectopic expression and *Salmonella* infection in mammalian cells (Chapter 5). Finally, in the discussion, I propose a model of SteC's kinase activation and how this relates to its activity in cells (Chapter 6).

2. Materials and Methods

2.1 Bioinformatics

2.1.1 Protein sequence analysis

The *steC* gene sequence with UniProt reference D0ZIB5 from *Salmonella* Typhimurium strain 14028s was used for all bioinformatic analyses and protein expression and purification (Section 2.5). The NCBI Basic Local Alignment Search Tool (BLAST) was used to identify homologues of the provided protein sequences (https://blast.ncbi.nlm.nih.gov/Blast.cgi).

To compare the different SteC protein sequence types among *Salmonella* serovars, 879 complete *Salmonella* genomes were downloaded from Enterobase by searching “Complete Genome” in the “Status” field, which represents the highest assembly quality with circular chromosomes and plasmids (https://enterobase.warwick.ac.uk/, Accessed on 2023/06/30). The SISTR1 results from Enterobase were used to identify the subspecies and serovars of the genomes. The *steC* nucleotide sequence from *Salmonella* Typhimurium LT2 (RefSeq: GCF_000006945.2) was used as a reference. A BLAST database was constructed from the *steC* sequence. Each of the 879 *Salmonella* genomes was queried against the *steC* database using BLASTn v2.14.0+ (Camacho et al., 2009). The aligned DNA sequences were then extracted and translated into protein sequences using Seqkit v2.4.0 (Shen et al., 2016). The unique SteC protein sequences were summarised and aligned using Clustalo v1.2.4 (Sievers et al., 2011). To visualise the SteC types in the *Salmonella* subspecies and serovars, an MStree of the 879 complete *Salmonella* genomes was generated on Enterobase using the rMLST scheme with the MSTree2 algorithm (Achtman et al., 2022). The tree was visualised with GrapeTree (Zhou et al., 2018). The sequence logo figure for the regions of interest was generated using the ggseqlogo package (https://github.com/omarwagih/ggseqlogo) in R (Wagih, 2017).

2.1.2 Protein structure prediction

Experimentally derived protein structures were retrieved from the Protein DataBank (PDB) (https://www.rcsb.org/). Published AlphaFold2 (AF) predicted structures were used where available from the AlphaFold Protein Structure Database (https://alphafold.ebi.ac.uk/). For SteC, the same UniProt reference was used as above (https://www.alphafold.ebi.ac.uk/entry/D0ZIB5). The FoldSeek online server (https://search.foldseek.com/search) was used to perform structural homology searches. The AF SteC prediction .pdb file was submitted and searched against either the complete or the human AF proteome. To search against individual regions of the SteC structure, the full AF

prediction was used as a template in Pymol and regions of the protein copied to create new .pdb files. PDB and predicted structures were visualised using Pymol (https://www.pymol.org/).

2.1.3 Other bioinformatic tools

Sequences were aligned with Jalview (https://www.jalview.org/), using the Tcoffee package. CLC Sequence Viewer 7 (https://clc-sequence-viewer.software.informer.com/7.0/) was used to design primers and analyse sequencing data. UniProt was used to determine protein domain boundaries. Expasy ProtParam tool (https://web.expasy.org/protparam/) was used to determine biophysical parameters of proteins including molecular weights (MW), isoelectric points (pI) and extinction coefficients (ε).

2.2 Plasmids, cloning and *Salmonella* transformation

2.2.1 Method of acquiring genetic material

The *steC* gene (strain 14028s) was available from the Thurston laboratory. The *steC* gene sequence with promoter region and C-terminal 2HA tag (Figure 2.1), mutants K256H, S379A, S379D and WT sequences removing amino acids 26-209, 430-457 or both (1-25+210-457, 1-429 and 1-24+210-429 respectively) were ordered as G-blocks from Integrated DNA Technologies. The FMNL1 gene was synthesised by ThermoFischer (https://www.thermofisher.com/uk/en/home.html). MYL12A (Q6ZWQ9) and MYL12A T19AS20A double mutant genes were ordered from Integrated DNA Technologies. Oligonucleotide primers were obtained from Merck.

2.2.2 Cloning and purifying plasmids

Expression plasmids were produced using Gibson assembly (Gibson et al., 2009), and all plasmids are listed in Table 2.1. Vector backbones used were pET49b (N-terminal GST-His6-3C protease cleavage site, kanamycin resistant, for *E. coli* expression), pET21 (N-terminal His6-3C protease cleavage site, carbenicillin resistant, for *E. coli* expression), pET28 (either with C-terminal uncleavable His6 or N-terminal Twin-Strep tag with TEV protease cleavage site, kanamycin resistant, for *E. coli* expression), pACEBac (either N-terminal His6-3c cleavage site or N-terminal Twin-Strep TEV protease cleavage site, gentamicin resistant, for insect cell expression), pTCMV (N-terminal GFP tag, for ectopic expression in mammalian cells) and pWSK29 (C-terminal 2HA tag, carbenicillin resistant, for transformation of *Salmonella*). Mutagenesis of these plasmids was performed using the Quickchange protocol. Plasmids were transformed into competent NEB DH5α *E. coli* cells by heatshock at 42 °C for 30 seconds, leaving in super optimal broth with catabolite repression (SOC) on ice for 5 min then incubation at 37 °C for a further 45 minutes and plating on agar with kanamycin, carbenicillin or gentamicin overnight. Single colonies were selected and grown overnight for DNA extraction using Qiagen miniprep kits. Sequencing reactions were performed by Azenta Genewiz (https://www.genewiz.com/en-GB) with in-house primers.

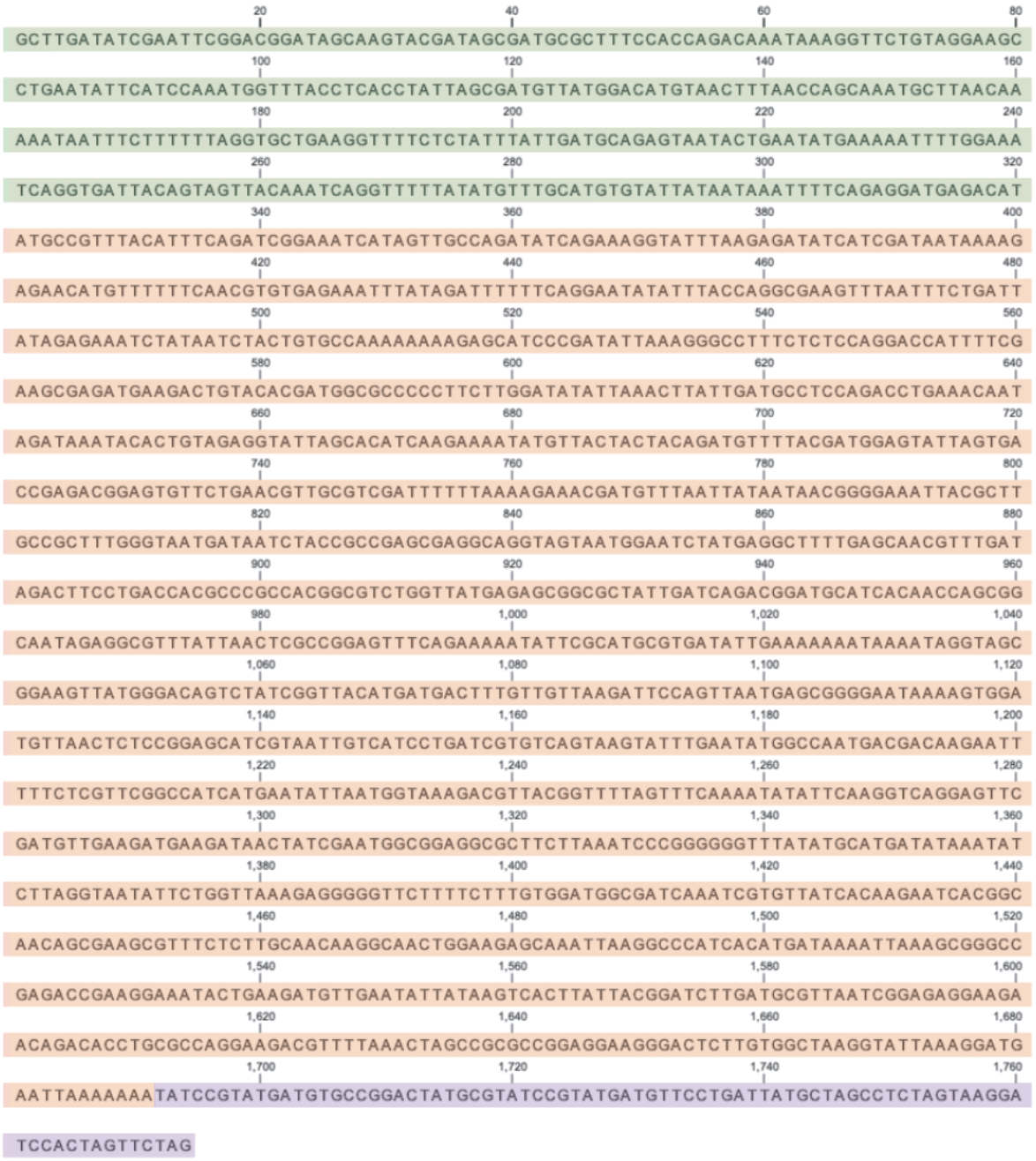


Figure 2. 1: Gene sequence used for pWSK29 plasmid into *Salmonella*

This gene sequence formed the template for all genes purchased from Integrated DNA Technologies. It contains *steC*'s endogenous promoter sequence (green), the *steC* open reading frame (orange) and C-terminal 2HA tag (purple). Mutants K256H, S379A and S379D, and shortened sequences 1-429, 1-25+210-457 and 1-25+210-429 were also ordered. All genes were cloned by Gibson Assembly into pWSK29 plasmid (available from Thurston Laboratory) then transformed into *Salmonella* 14028s *ΔsteC* strain (available from Thurston Laboratory) (see Section 2.2.3).

Protein (UniProt ID	Expression system	αα range	Mutants	Vector backbone	Tag	Antibiotic resistance	Purification methods	Quantity	Purity	Protein stability
SteC (D0ZIB5)	BL21 *E. coli*	1-457	WT	pET49b	N-GST-His-3C	Kanamycin	NiNTA, AIX	+	+/-	+/-
			WT	pET28	C-His	Kanamycin	NiNTA	+	-	NA
			WT, K256H, S379A/D	pET28	N-TS-TEV	Kanamycin	NiNTA	++	+	+/-
		210-457	WT, K256H, S379A/D, K256H S379D	pET28	C-His	Kanamycin	NiNTA, SEC	+++	++	+/-
				pET28	N-TS-TEV	Kanamycin	Streptactin, SEC	+++	++	+/-
		210-429	WT, K256H, S379A/D, K256H S379D	pET28	C-His	Kanamycin	NiNTA, SEC	+++	+++	+
				pET28	N-TS-TEV	Kanamycin	Streptactin, SEC	+++	+++	+
		200-377	WT	pET28	C-His	Kanamycin	NiNTA, SEC	+++	+++	++
		210-375	WT, K256H	pET28	C-His	Kanamycin	NiNTA, SEC	+++	+++	++
		1-193	WT	pET28	C-His	Kanamycin	NiNTA, SEC	+++	+++	+++
	Sf9 baculovirus	1-457	WT, K256H	pACEBac	N-TS-TEV	Gentamicin	Streptactin	+	+	+/-
		210-457	WT, K256H, S379A/D	pACEBac	N-TS-TEV	Gentamicin	Streptactin, SEC	++	++	+/-
		210-429	WT, K256H, S379A/D	pACEBac	N-TS-TEV	Gentamicin	Streptactin, SEC	++	++	+
	Mammalian cells	Table 5.1	Table 5.1	pTCMV	N-GFP	Kanamycin	NA	NA	NA	NA
	Salmonella	Table 5.2	Table 5.2	pWSK29	C-2HA	Carbenicillin	NA	NA	NA	NA
FMNL1 (O95466)	BL21 *E. coli*	1-385	WT	pET49b	N-GST-His-3C	Kanamycin	NiNTA, 3C cleavage, SEC	++	+	++
	Sf9	1-458	WT	pACEBac	N-His-3C	Gentamicin	NiNTA, SEC	++	+	++
MYL12A (Q6ZWQ9)	BL21 *E. coli*	1-172	WT, T19A S20A	pET28	N-TS-TEV	Kanamycin	NiNTA, SEC	+++	+++	+++

Table 2. 1: Expression plasmids cloned, and summary of expression, purification and results where relevant

S379A/D refers to two separate S379 mutants, K256H S379D and T19A S20A refer to double mutants; N- and C- refer to N-terminal and C-terminal tag positions. Disambiguation: 2HA, tag derived from human influenza haemagglutinin protein with sequence YPYDVPDYAYPYDVPDYASL; 3C, refers to 3C protease cleavage site with sequence LEVLFQGPG (Tag column) or 3C protease (Purification Methods column); αα, amino acid; GST, glutathione S-transferase; His, 6 x histidine tag; NA, not applicable; NiNTA, Nickel nitrilotriacetic acid; SEC, size exclusion chromatography; TEV, tobacco etch virus protease cleavage site with sequence ENLYFQSG; TS, twin strep tag with sequence MGWSHPQFEKGGGSGGGSGGSAWSHPQFEK; WT, wild-type.

2.2.3 Salmonella *strains*

Table 2.2 lists the *Salmonella* strains prepared for infection assays. *Salmonella* WT 14028s and the same strain in which either genomic *steC*, *invG* or *ssaV* had been replaced with a kanamycin cassettes were available from the Thurston laboratory. For an explanation of the approach to designing these plasmids, see Section 5.3.1. Plasmids were transformed into *E. coli* and incubated on agar in the presence of carbenicillin. Single colonies were selected for plasmid purification and sequencing. *Salmonella ΔsteC* strain was transformed with the relevant pWSK29 plasmid by electroporation. This entailed growing *Salmonella ΔsteC* overnight in 5 ml Luria Broth media (LB) then subculturing these bacteria for 2 hr 30 min at 1:200 in 5 ml LB per strain to be transformed at an OD_{600} of 0.4. The culture was cooled on ice for 30 min prior to washing in a sterile environment. Washing entailed pelleting bacteria at 2,700 x g for 6 min at 4 °C, discarding the supernatant, and resuspending the pellet in autoclaved MilliQ-purified H_2O (3 times) then autoclaved MilliQ H_2O with 10% glycerol (once). The pellet was resuspended in 200 µl 10% glycerol and electroporation was performed with 50 µl of cell suspension and 100 µl DNA in a 0.2 mm cuvette for 5 milliseconds, at 2.5 kV, 200 ohms and 25 µF with a Pulse Controller Plus (BioRad). The product was incubated for 1 hr at 37 °C in 500 µl SOC, plated on LB agar plates with kanamycin and carbenicillin along with non-transformed bacteria in the same antibiotic conditions as a control. Single colonies were selected and stored in 33% glycerol at -80 °C.

***Salmonella* strain**	**Genomic steC**	**pWSK29 plasmid (Carbenicillin resistance)**	**Antibiotic resistance**	**MW of protein (kDa)**
WT	Present	-	-	52.3
ΔsteC	Replaced with kanamycin resistance cassette	-	Kan	-
ΔsteC pE		Empty	Kan + Carb	-
ΔsteC psteC		*steC* 1-457		54.7
ΔsteC psteC K256H		*steC* 1-457 K256H		54.7
ΔsteC psteC S379A		*steC* 1-457 S379A		54.7
ΔsteC psteC S379D		*steC* 1-457 S379D		54.7
ΔsteC psteC 1-429		*steC* 1-429		51.6
ΔsteC psteC 1-25/210-457		*steC* 1-25/210-457		33.6
ΔsteC psteC 1-25/210-429		*steC* 1-25/210-429		30.5
ΔinvG	Present	NA	NA	NA
ΔssaV	Present	NA	NA	NA

Table 2. 2: *Salmonella* strains prepared for infection assays

Genomic knock-out strains were available from the Thurston laboratory. pWSK29 plasmids were cloned (Section 2.2.2) and transformed into *Salmonella ΔsteC* strain (Section 2.2.3) and stored in glycerol stocks at -80 °C. Carb, carbenicillin; Kan, kanamycin.

2.3 Recombinant protein expression and purification

2.3.1 Bacterial protein expression and purification

Competent BL21 (DE3) *E. coli* cells were transformed with relevant plasmids, grown to mid-log phase at 37°C in LB, measured by OD_{600} of 0.7-1, and expression was induced with 0.5 mM Isopropyl-β-D-thiogalactoside (IPTG) at 20 °C. After 16-24 hours of expression, cells were pelleted and either frozen for future use or lysed directly. To lyse, the fresh or frozen cells were resuspended in a buffer of 50 mM HEPES (4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid) pH 7.5, 300 mM NaCl, 10% glycerol (v/v), 10 mM $MgCl_2$, 20 mM imidazole (for His-tagged proteins only), 0.5 mM tris(2-carboxyethyl)phosphine hydrochloride (TCEP), 1 x protease inhibitor cocktail (Roche) and 2 µg/ml DNAse (Merck). Lysis was performed by sonication. The lysate was cleared by ultracentrifugation at 48,000 x g for 1 hour at 4 °C. Lysate supernatant was passed over affinity resin, either using nickel nitrilotriacetic acid (NiNTA) or Streptactin resin affinity chromatography, in a gravity column at 4 °C. Resin washing was performed with Buffer A (50 mM HEPES pH 7.5, 300 mM NaCl, 10% glycerol plus 20 mM imidazole for His-tagged proteins) and high salt (50 mM HEPES pH 7.5, 1 M NaCl) prior to elution with buffer B (50 mM HEPES pH 7.5, 300 mM NaCl, 10 mM $MgCl_2$, 10% glycerol, 0.5 mM TCEP and either 300 mM imidazole or 2.5 mM desthiobiotin). Tags were generally not removed by protease cleavage. In cases where cleavage was performed, the eluted protein was dialysed into Buffer A overnight at 4 °C with the addition of 100 µg of the relevant protease. Glutathione S-transferase (GST)-tagged 3C and His6-tagged tobacco etch virus (TEV) proteases were expressed and purified by standard procedures by members of the laboratory and stored at -80 °C in 20 µl aliquots at 5 mg/ml. After overnight cleavage, cleaved tags and proteases were removed by reverse affinity purification over the same resin combined with a second resin matching the protease tag.

2.3.2 Size exclusion chromatography (SEC) and protein storage

Where possible, size exclusion chromatography (SEC) was the final purification step. SEC buffer comprised 50 mM HEPES pH 7.5, 150 mM NaCl, 10 mM $MgCl_2$, 0.5 mM TCEP and 10% glycerol (v/v), unless otherwise stated. For large scale protein purifications, HiLoad 16/600 Superdex (200 or 75) columns were used on an AKTA prime system at 4 °C with a flow rate of 1 ml/min with 2 ml fractionation. For smaller scale purifications, Superdex 200 or 75 Increase 10/300 GL SEC columns were used, depending on whether the protein was above or below 50 kDa, respectively and with a flow rate of 0.5 ml/min with 500 µL fractionation. Where SEC was not successful due to protein aggregation, protein from previous purification steps was dialysed into SEC buffer overnight at 4°C. Protein was concentrated using Amicon Ultra centrifugal filters, of 50, 30 or 10 kDa filter sizes pre-equilibrated with buffer either to 3-4

ml for large scale purifications or 500 µl for smaller scale purifications. Protein concentration was measured using absorbance at 280 nm on a Nanodrop instrument, applying the Beer-Lambert Law, and the ultraviolet (UV) trace was checked for scatter suggestive of protein aggregation. Aliquots of protein were flash frozen in liquid nitrogen and stored at -80 °C. For biophysical techniques requiring highly pure protein in solution without aggregation, protein aliquots were thawed, potential aggregation cleared by ultracentrifugation at 17,000 x g at 4 °C for 10 min, protein dialysed where necessary, and the concentration re-estimated using a Jasco V-760 Spectrophotometer, applying the Beer-Lambert law and correcting for scatter.

2.3.3 Anion exchange chromatography (AIX)

When highly pure protein was required, anion exchange chromatography using a sodium chloride gradient elution was performed. AIX is a technique used to separate proteins based on their charge. Using AKTA prime at 4 °C, protein was loaded onto a 5 ml QHP HiTrap column (Cytiva) and eluted at 1 ml/min over a concentration gradient from 50 mM – 1 M NaCl over 40 min. Samples along relevant peaks were analysed by sodium dodecyl sulphate-polyacrylamide gel (SDS-PAGE). For acquiring the resolution necessary to separate phosphospecies, an AKTA pure with a Resource-Q column (Sigma-Aldrich) was used with a flow rate of 0.5 ml/min and a gradient of 50 – 500 mM over 30 min.

2.3.4 SDS-PAGE

To estimate the quantity and purity of protein at each purification step, SDS-PAGE analysis was performed. SDS solution was 2X NuPAGE LDS sample buffer (Thermo Fischer Scientific) with 200 mM DTT added 1:1 to protein samples. Samples were run on NuPAGE 4-12% Bis-Tris gels (Thermo Fisher Scientific) at 200 V for 35 min and stained with Quick Coomassie Stain (Protein Ark) before washing overnight with MilliQ water. All stained gels were imaged with a Chemidoc MP imaging system.

2.3.5 Insect cell (Sf9) expression and purification

This expression system involves producing a baculovirus with a genome containing the gene of interest and infecting cells of *Spodoptera frugiperda* (Sf9), the fall armyworm moth. The virus-infected cells then express the protein of interest. First, EmBacY cells were transformed with the pACEBac plasmid by electroporation pulsed at 2 kV with a Gene Pulser II (BioRad) and incubated at 37 °C for 6 hr to stimulate transposition. 100 µL of these cells were incubated for 72 hr on plates with antibiotics selecting for the bacmid (viral genome containing GFP gene to identify infected cells), the transposed plasmid and a helper plasmid. DNA was extracted from positive colonies and transfected into Sf9 cells. Cells expressing GFP denoted the

production of viable virus. Virus was amplified to reach the potency required for adequate protein expression. The P2 or P3 virus was deemed potent when > 80 % of cells were infected and the cell viability was 85-92 %. Sf9 at 2 x 10^6 cells/ml were infected with potent virus at a ratio of 500:1 and incubated for 72 hours at 28 °C with shaking. Protein derived from insect cells was purified as described for *E. coli* expressed protein with the following small adjustments: Lysis buffer included 0.1 mM 4-(2-aminoethyl)benzenesulfonyl fluoride hydrochloride (AEBSF) serine protease inhibitor. Cell lysis was performed with 0.1% Triton added to the lysis buffer with stirring at 4 °C for 45 min. For resin affinity purification, batch washing by resin centrifugation was performed prior to gravity column washing due to the viscosity of the lysate. Finally, elution of His-tagged proteins expressed in Sf9 required serial elution with buffers containing 300 mM and 500 mM imidazole.

2.3.6 Rabbit muscle actin purification

Actin was purified from rabbit skeletal muscle by Aleksandra Michrowska (Triesman laboratory) and given as a gift. Briefly, purification involved grinding rabbit skeletal muscle, before pelleting by centrifugation. Actin was polymerised for 30 min at 4 °C by adjusting the ATP, $MgCl_2$ and KCl concentrations to separate it from impurities before pelleting again and serial dialysis to reduce ATP concentration and depolymerise the actin. Latrunculin B was added at 5-fold excess with ON incubation at 4 °C with stirring. This drug binds the ATP binding pocket, to prevent polymerisation. It was stored at -80 °C in a buffer of 150 mM NaCl, 20 mM Tris pH 8, 3 mM $MgCl_2$, 0.5 mM TCEP and 0.2 mM ATP.

2.4 Structural and biophysical techniques

2.4.1 CD spectroscopy

Circular Dichroism (CD) spectroscopy is a method that uses the absorption of circularly polarised light by chiral molecules in an asymmetric environment to determine the presence of secondary structure in proteins in solution. Data collection and analysis were performed with the help of Laura Masino (Structural biology STP, The Francis Crick Institute). Proteins were first dialysed into a buffer of 15 mM phosphate buffer (pH 7.5), 250 mM NaF and 0.5 mM TCEP. A Jasco J-815 CD Spectrometer was used to collect an absorption signal in millidegrees across a wavelength range of 190-560 nm at 25 °C. The final spectra were averages of 20 separate scans. For thermal stability analysis, the signal at 220nm was collected across a temperature gradient from 2 °C – 95 °C and melting curves plotted. The delta epsilon molar unit was used in order to compare the signal between different sized fragments of the same protein.

2.4.2 1D NMR

One-dimensional ^{1}H nuclear magnetic resonance (1D NMR) data collection was performed by Diego Esposito (Rittinger lab, The Francis Crick Institute). All protein spectra were recorded at 25 °C on Bruker AVANCE spectrometers operating at 800 MHz in NMR buffer (50 mM HEPES pH 7.5, 150 mM NaCl, 10 mM MgCl2, 0.5 mM TCEP and 5% D_2O). Data were acquired and processed with Topspin (Bruker).

2.4.3 SEC-MALLS

Size exclusion chromatography coupled to multi-angle laser light scattering (SEC-MALLS) is a method used to obtain an absolute molecular mass of a molecule during analytical gel filtration. It is independent of shape therefore well-suited for determining the oligomeric state of proteins. It was performed with the help of Ian Taylor (Taylor lab, The Francis Crick Institute). In SEC-MALLS experiments 100 µL protein samples at concentrations of 0.25, 0.5, 1, 2, 4 and 8 mg/ml were first applied to a Superdex 200 10/300 INCREASE GL column equilibrated in 50 mM HEPES pH 7.5, 150 mM NaCl, 10 mM $MgCl_2$, 0.5 mM TCEP and 3 mM NaN_3 to separate species by hydrodynamic volume. Scattered light intensity was measured using a DAWN HELEOS II laser photometer, and dRI was measured using an OPTILAV-TrEX differential refractometer. The weight-averaged molecular mass of proteins was determined using the ASTRA software version 7.0 (Wyatt Technology Corp., Santa Barbara, CA) assuming Dn/dc to be 0.186 mL/g.

2.4.4 SAXS and SteC structure modelling

Small-angle X-ray scattering (SAXS) is a method of using X-ray scattering of a protein sample in solution to build a low-resolution map of a protein's flexibility, elongation, shape and oligomeric state. SAXS data collection and analysis were performed by Diego Esposito (Rittinger Lab, The Francis Crick Institute). SAXS data were collected at the SWING beamline at SOLEIL (GIF-sur-YVETTE CEDEX, France). 60 µL of proteins at 8 mg/ml were injected onto a 5 ml BioSEC-2.7 300 Å Agilent column and eluted at a flow rate of 0.2 ml/min at 15°C. Frames were collected continuously during the fractionation of proteins. Frames collected before the void volume (< 1.5 ml) were averaged and subtracted from the signal of the elution profile to account for background scattering. Data reduction, subtraction and averaging were performed using the software FOXTROT (SOLEIL). The scattering curves were analysed using the package ATSAS and reported as a function of the angular momentum transfer $q = 4\pi/\lambda \sin\theta$ where 2θ is the diffraction angle and λ the wavelength of the incident beam. Values of the cross-sectional radius of gyration were obtained with SCATTER (Rambo & Tainer, 2011). Low-resolution three-dimensional *ab initio* models for the SteC kinase domain fragments envelopes were generated by the program DAMMIF (Franke et al., 2017). The SAXS-derived dummy atom models were rendered and aligned to the AlphaFold2 predicted structures manually with PyMOL.

2.4.5 Biolayer Interferometry (BLI)

Biolayer interferometry is a method of assessing protein-ligand binding. A concentration range of protein is assessed against a ligand-loaded sensor. In this case, the N-terminally biotinylated FMNL1 peptides were synthesised by Dhira Joshi in the Chemical Biology STP (Francis Crick Institute). Peptides were solubilised in, and proteins were dialysed into, a buffer of 50 mM HEPES pH 7.5, 150 mM NaCl, 10 mM $MgCl_2$ and 0.5 mM TCEP. 0.03% Tween-20 was added to these samples. Protein-peptide interaction was analysed by BLI using an Octet BLI system. Briefly, sensors were equilibrated in buffer. Biotinylated peptides, synthesised by Dhira Joshi (Peptide Chemistry STP, The Francis Crick Institute) were loaded onto Streptavidin sensors at a concentration of 2 µg/ml prior to a buffer wash step. Protein binding was assessed at protein concentrations ranging from 7.5 nM to 200 µM. Data can be analysed either kinetically or at steady state. For the steady state estimate performed here, a point on the curve prior to the plateau phase was chosen and signal was plotted against protein concentration to estimate K_d, after correction for non-specific binding to the sensor. K_d was obtained from non-linear least squares fits, using the following equation:

$$\text{Response} = R_{max} * \text{Concentration} / (K_d + \text{Concentration})$$

where R_{max} is the response when the peptide on the sensors is fully saturated. The reported error was the mathematical error of the fit.

2.4.6 Hydrogen-deuterium exchange mass spectrometry (HDX-MS)

Hydrogen-deuterium exchange mass spectrometry looks at the exchange of hydrogen molecules in the protein backbone. It can be used to assess the relative exposure of protein residues to solution and hence to infer structural information about a protein in solution, as well as changes in this exposure in the presence of a ligand. HDX-MS was performed by Sarah Maslen (Proteomics STP, Francis Crick Institute). 20 µL of 5 µM proteins (individually or in combination) were incubated with 40 µL of D_2O buffer at room temperature for 3, 30, 300 and 3000 seconds in triplicate. The labelling reaction was quenched by adding chilled 2.4% v/v formic acid in 2 M guanidinium hydrochloride and immediately frozen in liquid nitrogen. Samples were stored at -80 °C prior to analysis.

The quenched protein samples were rapidly thawed and subjected to proteolytic cleavage by pepsin followed by reversed phase high-performance liquid chromatography (HPLC) separation. Briefly, the protein was passed through an Enzymate ethylene-bridged hybrid (BEH) immobilised pepsin column (Waters, UK) at 200 µL/min for 2 min and the peptides trapped and desalted on a 2.1 x 5 mm C18 trap column (Acquity BEH C18 Van-guard pre-column, 1.7 µm, Waters, UK) then eluted. Peptides were separated on a reverse phase column (Acquity UPLC BEH C18 column 1.7 µm, 100 mm x 1 mm (Waters, UK). Peptides were detected on a Cyclic mass spectrometer (Waters, UK). Peptide identification was performed by MS^e (Silva et al., 2005). The resulting MS^e data were analysed using Protein Lynx Global Server software (Waters, UK). Mass analysis of the peptide centroids was performed using DynamX software (Waters, UK). All time points in this study were prepared at the same time and individual time points were acquired on the mass spectrometer on the same day.

2.4.7 Protein crystallisation

The sitting drop vapour diffusion method was used for all crystallisation trials. 2-drop MRC crystal plates were loaded with a range of commercially available crystallisation screens including sparse matrix screens and polyethylene glycol (PEG) screens at 90 µl per well (see Table 3.6 and Figure 3.23 for details). For optimisation screens, buffer component concentrations ranges were created with the Formulatrix (https://formulatrix.com/) Rockmaker software and a Formulator liquid handling device was used to dispense the conditions. Once a plate was loaded with buffers, 100 nL protein was mixed with 100 nL reservoir solution and

loaded into drops using NT8 drop setter (Formulatrix). Plates were stored and imaged in a Formulatrix plate reader at 18 °C or 4 °C. The non-biotinylated 24 amino acid peptide of FMNL1 from K190 (pep_K190) was synthesised by Dhira Joshi (Chemical Biology STP, The Francis Crick Institute). Details of the protein constructs and concentrations are listed in Table 3.6.

2.5 Biochemical techniques

2.5.1 Kinase assay with Phos-stain

Kinase assays were carried out in standard SEC buffer. Each 15 µL reaction mixture contained 10 µM kinase (enough to visualise auto-phosphorylation) and substrate +/- 100 µM ATP. For reactions with insect cell-expressed protein, a phosphatase inhibitor cocktail (Roche) was also added. Murine lysate was produced from Swiss 3T3 mouse fibroblasts (See Section 2.6.2). Reactions were incubated at 30 °C for 30 min prior to SDS-PAGE analysis followed by fixing, staining and destaining according to the Pro-Q Diamond protocol (available here). Gels were imaged with chemiluminescence on the Chemidoc MP imaging system.

2.5.2 Radioactive kinase assay

Reactions were performed in standard SEC buffer. Each 15 µL reaction mixture contained 50 nM - 10 µM kinase (details stated in Results in each case), 100 µM ATP, 20 kBq [γ-32P]ATP (Hartmann Analytic, 9.25 MBq, 25 µL pot, 5000 Ci/mmol). Given the aggregation-prone nature of SteC 210-457, for concentrations below that detectable by Coomassie staining, protein aliquots were thawed, thoroughly mixed and concentration recalculated using Nanodrop measurement, prior to serial dilution to the concentration required. One substrate used was FMNL1 K190 24mer peptide (pep_K190) as used in crystallisation trials, at 100 µM. Other substrates were added at concentrations ranging between 0.5 and 10 µM as indicated in the relevant Results. Reactions were incubated at 30 °C for 30 min prior to SDS-PAGE analysis, Coomassie staining, gel drying and exposure to radiography film in the dark for between 30 min and 18 hours before developing. 1 µl Swiss 3T3 mouse fibroblast lysate (see Section 2.6.2 for details of the preparation of this lysate) was used to spike relevant kinase assay samples. STK38 from Thermo Fisher Scientific and calmodulin from Merck were purchased in buffer.

2.5.3 PhosphoMS of recombinant protein

Phospho-mass spectrometry (PhosphoMS) of recombinant protein was performed by Steve Howell (Proteomics STP, The Francis Crick Institute). Proteins were reduced and alkylated in-gel (10 mM TCEP, 40 mM chloroacetamide for 20 min at 70°C) prior to trypsin digestion (modified sequencing grade, Promega) in 10 mM NH_4HCO_3 overnight at 37°C. Acidified supernatant was separated by high-performance liquid chromatography and loaded into a Lumos Tribrid Orbitrap mass spectrometer (all Thermo Scientific). Raw files were searched using Maxquant (maxquant.org) against FASTA sequences of relevant recombinant constructs, recent downloads of UniProt baculovirus related databases and a common contaminants database and modified residues were assigned. Visualisation was in Perseus (maxquant.net/perseus) and Skyline (skyline.ms).

2.5.4 Intact mass spectrometry

Intact mass spectrometry was performed by Sarah Maslen (Proteomics STP, Francis Crick Institute). SDS-denatured proteins were injected onto a C4 BEH 1.7 µm, 1.0 x 100 mm UPLC column using an Acquity I class LC (Waters, UK). Proteins were eluted with a 15 min gradient of acetonitrile (2%v/v to 80%v/v) in 0.1%v/v formic acid using a flow rate of 50 µl/min. The analytical column outlet was directly interfaced via an electrospray ionisation source, with a time-of-flight (TOF) mass spectrometer (BioAccord, Waters, UK). Data were acquired over a *m/z* range of 300–8000, in positive ion mode with a cone voltage of 40 V. Scans were summed together manually and deconvoluted using MaxEnt1 (Masslynx, Waters, UK).

For the lysate experiment to identify an SteC substrate, MS analysis was performed by Lois Kent (Proteomics STP, The Francis Crick Institute). Acrylamide gel band was dissolved and digested with trypsin. ESI-TRAIP instrument was used. UniProt Mouse database was used to search for hits.

2.5.5 Mant-AMPPNP binding

This experiment was performed with the help of Laura Masino (Structural biology STP, The Francis Crick Institute). Proteins were analysed in SEC buffer. Mant-Adenylyl imidodiphosphate (AMPPNP) was acquired from the structural biology STP. A Jasco FP-8500 Fluorometer was used to create an excitation wavelength of 355 nm and collect an emission spectrum at 390-550 nm. Protein titrations were created by spiking Mant-AMPPNP at 500 nM with protein. Signal at 442 nm was recorded and corrected for differences in baseline signal and dilutional effect of spiking protein on Mant-AMPPNP concentration, prior to plotting against the calculated protein concentrations.

2.5.6 PIP strip analysis

All steps were performed at 4 °C. Phosphoinositide (PIP) strips membranes (ThermoFischer, UK) were first blocked in PBS with 0.1% Tween 20 (PBS-T) with 3% fat-free BSA for 1 hr. The strips were incubated with the purified recombinant protein of interest at the indicated concentrations in the same buffer for 4 hr at 4 °C. Then strips were washed in the same buffer for 10 min 3 times. Strips were then incubated with Anti-His antibody (Table 2.4) prior to repeating the wash step. The strips were imaged by chemiluminescence using ECL™ detection reagents and Pierce™ ECL Plus Western Blotting Substrate (Thermofisher Scientific, USA) and BioRad ChemiDoc MP imager.

2.6 Cell culture

2.6.1 Cell lines

Human Embryonic Kidney 293 cells (HEK 293ET cells) and RAW 264.7 macrophages (RAW cells) were obtained from the Thurston laboratory. Swiss 3T3 mouse fibroblasts (3T3 cells) were obtained from cell services at the Francis Crick Institute. Cells were grown in Dulbecco's Modified Eagle's Medium (DMEM) (Sigma, USA) supplemented with 10% foetal bovine serum (FCS) (Gibco Life Sciences, UK) at 37 °C in 5% CO_2. 3T3 cells were detached with 1 ml trypsin in EDTA per confluent 10 cm dish and incubated for 1-3 min at 37 °C for each passaging.

2.6.2 Murine cell lysate

2 confluent 10 cm plates of 3T3 cells were washed in phosphate buffered saline (PBS) once then lysed in the plate with 1 ml Lumier ++ (Tris pH 7.4, 150 mM NaCl, 0.1% Triton, 1% EDTA, protease inhibitor cOmplete (Roche) and phosphatase inhibitor phoSTOP (Roche)) with 0.3% Triton for 15 min on ice, before pelleting at 17,700 x g for 10 min at 4°C. Aliquots of 20 µl were stored at 20 °C.

2.6.3 Transfection of HEK 293ET cells

HEK 293ET cells were seeded onto sterile coverslips primed with Poly-L-lysine (Sigma, USA) to aid attachment in 24 well plates, 24 hours prior to transfection. At 70% confluency, the cells were transfected with Lipofectamine 2000 (Invitrogen, USA) hosting the pTCMV GFP-SteC mutant vectors (Table 2.1) as cargo, using 3 µL lipofectamine stock and 600 ng of vector in 250 µL Opti-MEM Reduced Serum Medium (Gibco, UK). The cells were incubated at 37 °C for 24 hours.

2.6.4 Salmonella *infection*

All infection experiments were performed in containment level 2 laboratory with appropriate infection control procedures in the Thurston Laboratory. For immunofluorescence microscopy, 3T3 cells were seeded on glass coverslips in 24 well plates at a density of 2 x 10^4 cells per well 24 hours before infection. For immunoblotting, 3T3 and RAW cells were seeded into 6-well plates. Overnight cultures of *S.* Typhimurium strains (Table 2.2) were diluted 1:33 in fresh LB and either grown for 3.5 hrs at 37 °C with shaking in an orbital shaking incubator to obtain logarithmic phase bacteria (to induce SPI-1 conditions to aid infection), or for one experiment grown to OD_{600} of 0.5 as per Dai et al. (2024) (see Results for details). Cultures were added directly to 3T3 and RAW cells at a multiplicity of infection of 100:1 and 10:1, respectively. Bacterial invasion was allowed to proceed for 20 min at 37 °C in 5% CO_2. To follow a synchronised population of intracellular bacteria, 20 min after inoculation 100 µg / ml

gentamicin (bacteriocidal concentration) was added to the media to kill extracellular bacteria. This was left for 1 hr, before reducing to 20 µg / ml gentamicin (bacteriostatic concentration) for the remaining experiment.

2.6.5 Immunofluorescence microscopy

All coverslip wash steps were in PBS twice. Coverslips with infected or transfected cells (see above) were washed; fixed using 3 % paraformaldehyde (PFA, Sigma-Aldrich) in PBS at room temperature for 30 min; washed; quenched with 100 mM NH_4Cl in PBS at room temperature for 30 minutes or at 4 °C overnight to quench PFA autofluorescence; washed; then permeabilised and blocked using 0.1 % Triton with 10 % horse serum (Sigma) or 0.1% Saponin with 10 % horse serum in PBS at room temperature for 1 hour. Staining of primary and secondary antibodies was performed in the permeabilisation and blocking solution with antibodies or staining reagents added at the concentration indicated in Table 2.3. For transfection experiments, actin was stained with phalloidin and DNA with4'6-diamidino-2-phenylindole (DAPI). For infection, two experiments were performed. For visualising and counting different actin phenotypes, cover slips were stained with DAPI, phalloidin Alexa Fluor 555 and anti-core *Salmonella* antigen (CSA)-1, an antibody specific for a *Salmonella* surface marker. A secondary antibody conjugated with Alexa Fluor 488 was used for CSA-1. For the second experiment, to visualise the localisation of SteC in the infected cell, coverslips were stained with phalloidin, CSA-1 as above, as well as anti-HA antibody with a secondary antibody conjugated with Alexa Fluor 647. Coverslips were mounted onto slides using Aqua-Poly/Mount (Polysciences, Inc.) and dried overnight at room temperature in the dark.

For transfection experiments, images were captured using a confocal laser scanning microscope (LSM710) (Zeiss GmbH) with a 63x objective. Counting was performed manually, by assessing the actin polymerisation of 100 transfected cells for each cover slip, with three coverslips per condition.

To capture super-resolution images of infected cells, an Olympus CSU-W1 SoRa Spinning Disk was used with 60x lens, under the supervision of Donald Bell and Matt Renshaw (Light Microscopy STP, The Francis Crick Institute). Images were collected with SoRa Disk Changer in Z-stack. Post-image analysis involved running OSR low Macro then deconvolution with SoRa software. All images were processed in Fiji (Image J) and are displayed in colour-blind friendly combinations.

To quantify actin phenotypes in infection, Axio Imager Upright Microscope (Zeiss) was used. 3-5 coverslips of each condition were code-labelled by Yizhou Huang (Thurston Laboratory, Imperial College London) to facilitate counting blinded to *Salmonella* strain.

Antibody	**Assay**	**Dilution / concentration**	**Source (catalogue number)**
DAPI	Transfection Infection	200 mM	Thermofischer (D3571)
Phalloidin Alexa Fluor 647	Transfection	1:200	Thermofischer (22287)
Phalloidin Alexa Fluor 555	Infection	1:150	Invitrogen A34055
Goat anti-CSA-1 (primary)	Infection	1:200	KPL 01-91-99
Mouse anti-HA 11 (primary)	Infection	1:200	BioLegend 901502
Rat anti-HA 3F10	Infection	1:200	Roche 1867423
Donkey anti-Goat Alexa Fluor 488	Infection	1:400	Life Technologies
Donkey anti-Mouse Alexa Fluor 647	Infection	1:400	Life Technologies
Donkey anti-Rat Alexa Fluor 647	Infection	1:400	Life Technologies

Table 2. 3: Immunofluorescence staining reagents

2.6.6 SDS-PAGE and immunoblotting

Each well of confluent transfected or infected cells from a 6-well plate (see sections 2.6.3 and 2.6.4) was washed twice in PBS, incubated in trypsin for 1-3 min at 37 °C, then resuspended in PBS and centrifuged at 10,000 x g for 2 min. After a second PBS wash, samples were lysed with Lumier ++ (Tris pH 7.4, 150 mM NaCl, 0.1% Triton, 1% EDTA, protease inhibitor cOmplete (Roche) and phosphatase inhibitor phoSTOP (Roche)) for 10 min on ice. Whole cell lysates were centrifuged for 20 min, at 18,000 x g at 4 °C. Samples were denatured with SDS loading buffer (125mM Tris-Cl pH 6.8, 4 % SDS, 10% glycerol, bromophenol blue and β-mercaptoethanol) and the resulting solution heated to 95 °C for 5 min before.

Denatured samples were run on 10-14% polyacrylamide gels by electrophoresis. A constant voltage of 110 V was applied for 90 min. Proteins were transferred onto a 0.2 µm polyvinylidene fluoride or polyvinylidene difluoride (PVDF) transfer membrane (Millipore) using a Trans-Blot® Turbo™ Transfer System (BioRad). Membranes were blocked in 5 ml 5% bovine serum albumin (BSA) in Tris-buffered saline with 100 mM Tris pH 7.4, 150 mM NaCl, 0.1% Tween 20 (TBS-T) for 30 min at room temperature on a roller. The membranes were incubated

with primary antibodies in 5% BSA overnight at 4 °C. Membranes were washed 3 times in TBS-T before being exposed to the appropriate secondary antibody in 5% BSA for 1 hr. The membranes were washed again in TBS-T then ECL™ detection reagents and Pierce™ ECL Plus Western Blotting Substrate (Thermofisher Scientific, USA) were used to visualise the substrate with an iBright FL1500 Imaging System (Invitrogen, USA). See antibody list for details (Table 2.4).

2.6.7 In cellulo effector translocation assay

This assay was performed with the help of Mei Liu and Ioanna Panagi from the Thurston laboratory. 3T3 cells seeded on 6-well plates were infected with *Salmonella* WT, *ΔsteC* and *ΔsteC psteC* strains (Table 2.2) as previously described. At 8 hours, cells were washed with 1 ml PBS, detached with trypsin, pelleted by centrifugation at 3,500 x g for 1 min at 4 °C, resuspension in PBS and re-centrifugation. The cell pellet was lysed with 50 µl Lumier ++ (Tris pH 7.4, 150 mM NaCl, 0.1% Triton, 1% EDTA, protease inhibitor cOmplete (Roche) and phosphatase inhibitor phoSTOP (Roche)) for 10 min on ice and centrifuged at 17,700 x g at 4 °C. After centrifugation two layers were generated: 1) a top soluble layer representing the post-nuclear supernatant (PNS) and 2) a viscous pellet. The PNS (~100 µl) which contains mammalian cytosol and translocated *Salmonella* effectors, was denatured by addition of 20 µl of 5x SDS buffer (125mM Tris-Cl pH 6.8, 4 % SDS, 10% glycerol, bromophenol blue and 5% β-mercaptoethanol). The pellet fraction, containing bacteria and mammalian cell nuclei, was denatured with addition of 100 µl of 2x SDS loading buffer (as above). Denatured samples were then boiled at 95°C for 7 min. Pellet samples were sonicated briefly to reduce viscosity prior to PAGE and immunoblot analysis, as previously described, followed by blotting for HA, DnaK and tubulin (Table 2.4).

2.6.8 Co-Immunoprecipitation mass spectrometry

10 cm dishes of 3T3 cells were infected with *Salmonella ΔsteC* strains expressing SteC-2HA, SteC K256H-2HA or *Salmonella ΔsteE* expressing SteE-2HA with the same method as described above, or left uninfected. After 8 hr, cells were washed in PBS, resuspended and centrifuged for 10 min at 300 x g at 25 °C. PBS was aspirated. Cells were lysed with 1 mL of mild lysis buffer (Tris pH 7.5, 50 mM KCl, 2 mM $MgCl_2$, 0.1% Triton, 1 x protease inhibitor cocktail (Roche), phosphatase inhibitor phoSTOP (Roche)) on ice for 10 min with regular vortexing. Cell lysates were centrifuged at 17,000 x g at 4 °C for 10 min. 900 µL of supernatant was added to 20 µL of cross-linked anti-HA IgG magnetic beads (Pierce, #88836). Samples were incubated on a wheel for 30 min at 25 °C then briefly centrifuged and washed in mild lysis buffer 3 times with a magnetic rack. Samples were eluted in 50 µL of 100 mM

triethylammonium bicarbonate buffer (Sigma, USA) with 10% added SDS at 50 °C for 10 min with shaking. 10 µL were taken from the eluted solution and frozen at -80 °C. Proteomic analysis of the samples and raw data processing were performed by Tania Auchynnikava by standard methods (Meier et al., 2020). Briefly, samples were digested with trypsin and analysed with a timsTOF MS machine by data-independent acquisition modes with parallel accumulation-serial fragmentation (diaPASEF). Spectronaut and DIA-NN were used to identify peptides present in the samples matching *Mus musculus* and *Salmonella* Typhimurium 14028s genomes and generate statistical significance values. Proteins identified were deemed significant if they had a Q value <0.413.

Antibody/reagent	**Experiment**	**Dilution**	**Source (catalogue number)**
Rat anti-GFP	Transfection	1:2000	ChromoTek (3H9-100)
Mouse anti-Tubulin	Transfection Translocation	1:2500	DSHB (E7)
HRP-conjugated Goat anti-Rat	Transfection	1:2500	Cell Signalling (7077)
HRP-conjugated Goat anti-Mouse	Transfection Translocation	1:2500	Agilent (Dako) (P0447)
DnaK	Translocation	1:10,000	Enzo ADI-SPA-880-F
Mouse anti-HA (primary)	Translocation	1:200	BioLegend 901502
Donkey anti-Mouse Alexa Fluorophore 647	Translocation	1:500	Life Technologies
Rabbit anti-His antibody (HRP-conjugated)	PIP strip	1:2000	Abcam (Ab1187)
Anti-MYL12A	MYL12A phosphorylation at T19/S20	1:1000	Proteintech (16287-1-AP)
Anti-phospho-MYL12A	MYL12A phosphorylation at T19/S20	1:1000	Abcam (Ab2480)
Anti-phospho- MYL12A	MYL12A phosphorylation at T19/S20	1:1000	Abcam (AB126739)

Table 2. 4: Antibodies used for western immunoblot and PIP strip

HRP, horseradish peroxidase; PIP, phosphoinositides

2.6.9 Statistical analysis

Statistical significances were calculated using an ordinary one-way analysis of variance (ANOVA) complemented with a *post-hoc* test for multiple comparison's corrections, as described in figure legends. All analyses were completed on GraphPad Prism (Version 10.1.1).

3. Structural and biophysical characteristics of SteC

Introduction

In the first part of this chapter, I use bioinformatic tools to identify sequence homologs of SteC in prokaryotes and eukaryotes, predict the structure of SteC and identify structural homologs to hypothesise molecular mechanisms behind the activities of SteC's kinase and regulatory domains. In the second part, I describe the recombinant expression and purification of SteC and its domains and apply biophysical and structural techniques to interrogate the validity of the predicted structure of SteC. In the final part, I introduce a recently identified substrate of SteC, FMNL1 (Walch et al., 2021), assess its phosphorylation sites, assess SteC-FMNL1 binding and describe crystallisation trials of the kinase domain with ligands.

3.1 Analysis of the sequence and predicted structure of SteC

The SteC amino acid sequence from *Salmonella* laboratory reference strain 14028s was used for all analyses in this chapter and experimental work described throughout this thesis (Figure 3.1).

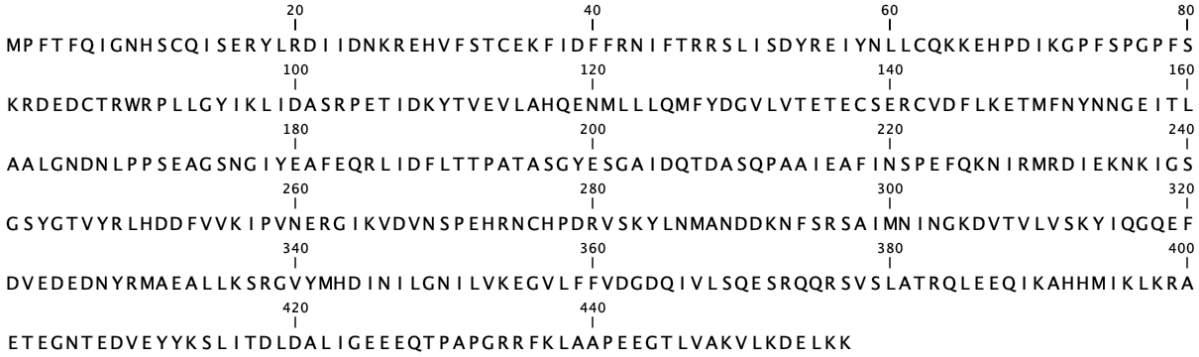


Figure 3. 1: Reference SteC amino acid sequence

This sequence is from reference strain *Salmonella* Typhimurium 14028s, UniProt D0ZIB5.

3.1.1 Sequence analysis

Analysis of a library of clinical isolates of *Salmonella* from multiple sites globally was performed by a collaborator Yan Li in the Hinton laboratory, University of Liverpool. The *steC* nucleotide sequence from the same serovar, 14028s, was used as the reference sequence and translated into the amino acid sequence. This analysis showed that SteC is present in the majority of serovars (Figure 3.2). SteC has 98% amino acid sequence identity across serovars compared

to the reference, except for SteC in *S.* Typhi which has 91% amino acid sequence identity. SteC was absent from *S. bongori*; absent or pseudogenised in subspecies IIIa; pseudogenised in subspecies IIIb and *S.* Paratyphi; and demonstrated evidence of pseudogenisation in a minority of isolates across the majority of serovars.

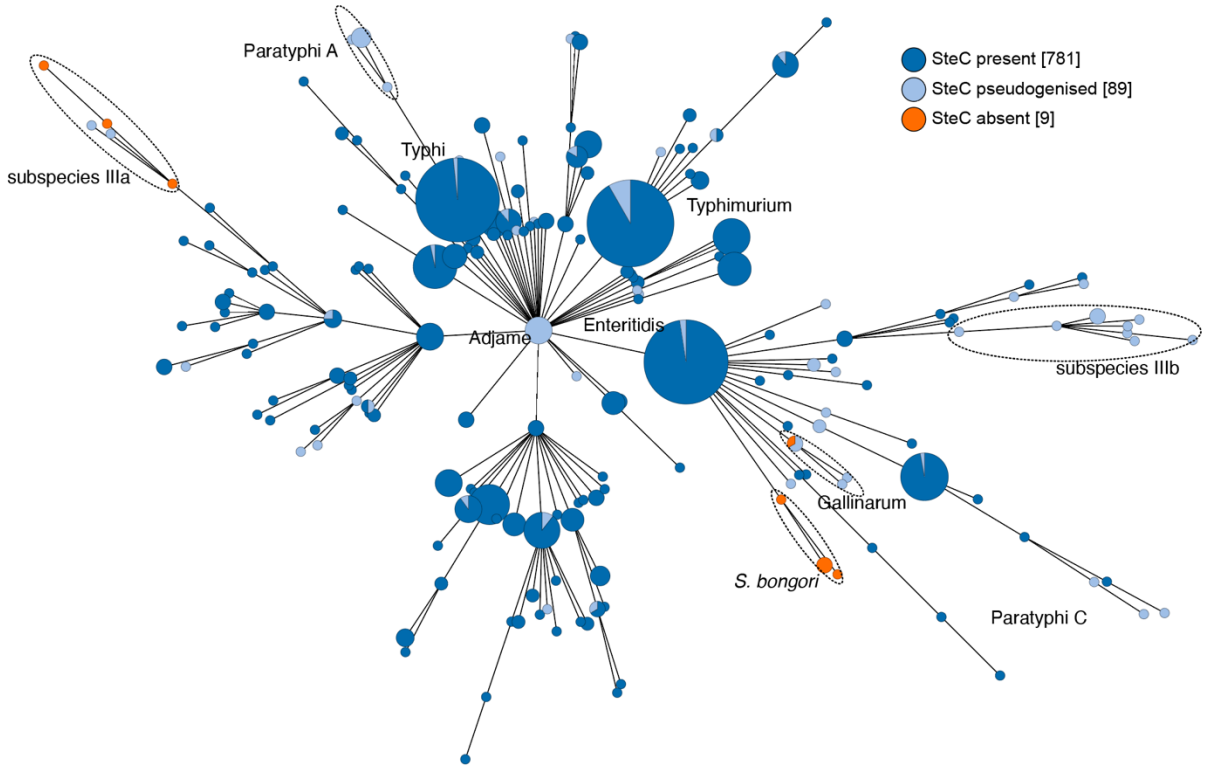


Figure 3. 2: Distribution of SteC across *Salmonella* serovars

Analysis of protein sequences using GrapeTree (Zhou et al., 2018) demonstrating the presence of SteC in 879 genomes from *Salmonella* clinical isolates and reference sequences. The size of each bubble is proportional to the number of genomes it represents. Analysis of sequences is not rooted meaning evolutionary relationships cannot be inferred. Figure generated by Yan Li (Hinton laboratory, University of Liverpool).

Using standard thresholds, a BLASTp query of the SteC amino acid sequence (14028s) against the NCBI database identified several SteC variants within *Salmonella* species, as well as similar sequences in *Yokenella regensburgei* (98% coverage, 41% amino acid identity), *Cedecea neteri* (92%, 42%), *Sodalis praecaptivus* (86%, 34%) and *Erwinia mallotivora* (42%, 49%), which are all bacterial species reported to be pathogenic in human. *C. neteri* and *E. mallotivora* have type 1 and 3 secretion systems, respectively (Tamizi et al., 2020; Thompson & Sharkady, 2021). When a multi-sequence alignment was performed of SteC and these four homologs, certain motifs appear conserved, largely in the region of the alignment from amino acid 260 to 420, which corresponds to SteC 238-390 (Figure 3.3).

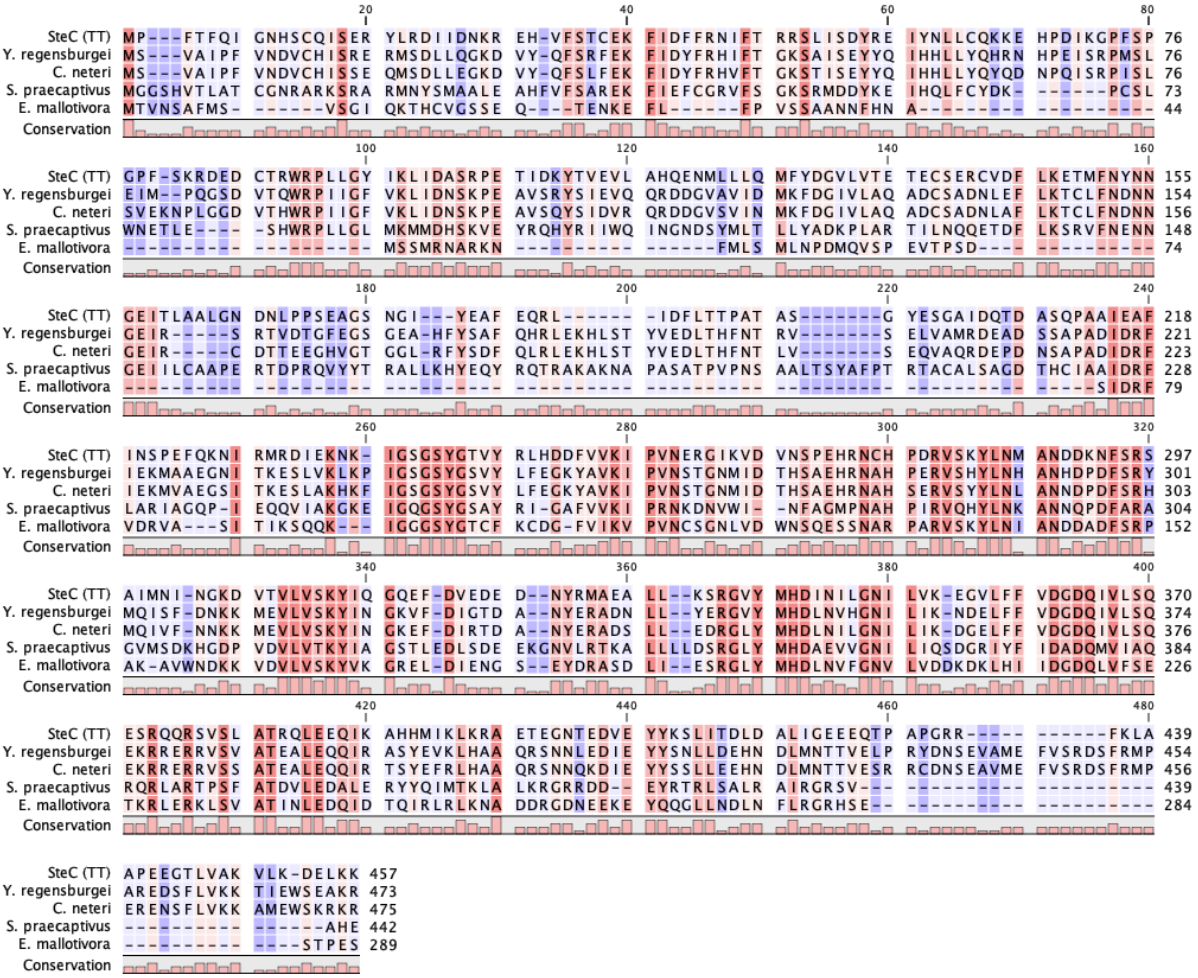


Figure 3. 3: Sequence alignment of SteC and homologs in non-*Salmonella* bacterial species

A multiple sequence alignment was performed of the amino acid sequences of SteC and its top 4 bacterial homologs identified by BLASTp search, using T-Coffee in CLC Sequence Viewer 7. The background colour scale represents a spectrum from full homology (red) to complete lack of homology (light blue). The conservation graph below demonstrates the number of amino acid sequences sharing any given residue.

Given effector proteins are translocated into the host cytosol and perform eukaryote-like biochemical activities, I hypothesised that SteC sequence homology to human kinases may reveal biologically meaningful insights. A standard BLASTp query constrained to the human proteome returned no significant results. However, increasing the expect threshold from 0.05 to 1 yielded one result: a 26% sequence identity between SteC 213-357 and human kinase RAF1 329-480, as reported previously (Poh et al., 2008). Alignment of SteC and RAF1 (Figure 3.4) demonstrated some conserved canonical kinase motifs as described in Section 1.3.2. Most notably, IGSGSxGTVY representing the glycine-rich loop, VKI containing the catalytic lysine which is catalytically essential in SteC (Poh et al., 2008), and PE containing a glutamic

acid were identified, three motifs that may be important for anchoring, orienting and stabilising ATP in the kinase N-lobe.



Figure 3. 4: Sequence alignment of SteC and RAF1

Sequence alignment using T-Coffee in Jalview. Conserved residues highlighted in red. The first three canonical kinase motifs, the glycine-rich loop, catalytic lysine and αC glutamate, are highlighted in red boxes.

3.1.2 Analysis of the AlphaFold2-predicted structure

The availability of AlphaFold2 (from now on referred to as AF) provides an opportunity to explore the structure of a protein of interest *in silico* and create hypotheses that can be tested experimentally. AF structural predictions significantly outperform previously available structural prediction software (Jumper et al., 2021). The AF algorithm reports two metrics of prediction confidence: predicted local distance difference test (pLDDT) that gives a local per residue confidence score related to its neighbouring atoms; and a predicted alignment error (PAE) which gives a confidence in the relative position of any two residues in the structure. pLDDT can be used to assess which residues have been predicted with higher confidence. A prediction of no secondary structure with a low pLDDT correlates with intrinsic disorder (Guo et al., 2022) though AF is not itself able to distinguish between a low pLDDT score due to insufficient training data for the sequence, versus intrinsic features of the protein. PAE provides a prediction of domains, because amino acids that are predicted with high confidence in relation to each other are likely to be within one globular protein region. The relative accuracy of residues across different domains relates to how confidently the prediction has associated domains in relative orientation.

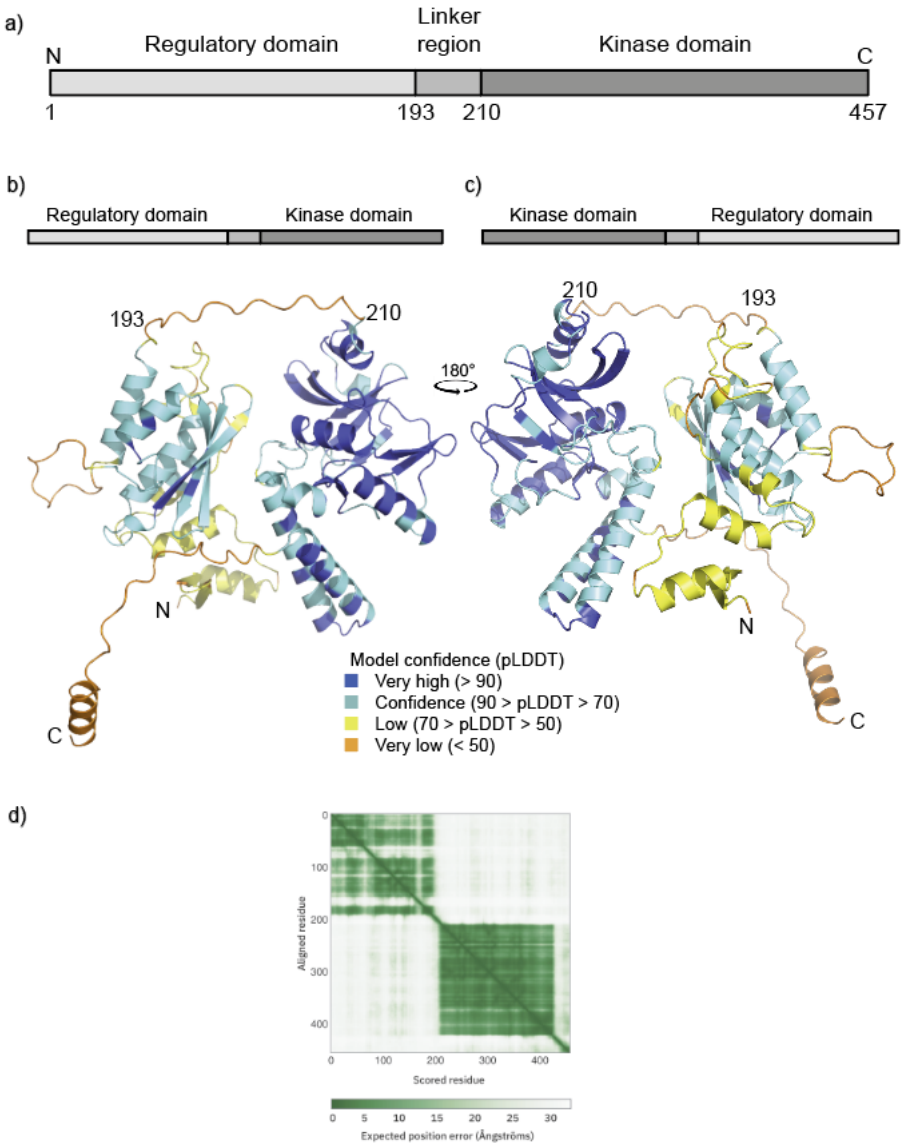


Figure 3. 5: AlphaFold2 predicted structure of SteC

a) Domain map of SteC, demonstrating two domains with linker region
b) AlphaFold2 prediction of SteC, coloured with AlphaFold2-generated predicted local distance difference test (pLDDT) score as per key. Domain map reproduced above.
c) 180° rotation of AlphaFold2 prediction around y axis. Domain map reproduced above.
d) Predicted alignment error (PAE) map demonstrating prediction confidence of the relative position of any two residue pairs.

The predicted structure of SteC contains two domains encompassing amino acids 1-193 and 210-457, connected by a linker region spanning 194-209 (Figure 3.5a-c). Overall, the prediction of the N-domain is of lower confidence, containing a β-sheet and six small α-helices. The N-terminal 40 amino acids, likely to contain the effector translocation signal, were predicted with the least confidence within this domain. The C-domain, the putative kinase

domain, is predicted with higher confidence from amino acids 210-429. However, the C-tail, 430-457, is predicted to largely lack secondary structure apart from one small alpha helix; not to associate with other parts of the structure; and with low confidence. This may suggest it is unstructured or intrinsically disordered. The PAE table suggests that the two domains are better predicted as stand-alone domains, compared to their relative orientation, which is poorly predicted (Figure 3.5d).

Eukaryotic kinase domains are comprised of N- and a C-lobes with conserved secondary structural features: α-helices A-I and β-sheets 1-8. Using the PKA amino acid sequence, secondary structural elements and kinase motifs as a template (Figure 3.6) as previously presented (Figure 1.5a), SteC's sequence, overlaid predicted secondary structural features and hypothetical motifs were compared (Figure 3.6b). This identified differences in the SteC predicted structure. Firstly, SteC demonstrates certain secondary structural elements, namely αA, β1-3, αC, β4-5, αD-E, β6-9 and αEF, F and G, but lacks α-helices B, D, H and I. This can be seen in the minimal nature of the SteC predicted C-lobe compared to PKA (Figure 3.7a). SteC is also predicted to contain a C-terminal tail lacking secondary structure except for a small α-helix seemingly not associated with the rest of the kinase domain (Figure 3.7a). This feature, from amino acids 430-457, will be referred to as the C-tail.

Alignment with PKA combined with AF-based structure prediction also presents the opportunity to interrogate the position of key catalytic motifs. SteC contains a glycine-rich loop between β1 and β2, and a catalytic lysine in β3, as expected (Figure 3.6b). Other motifs resemble those expected, however with some differences: instead of an HxD motif after β6, there is an 'HD' motif. Instead of a DFG motif after β8 at the start of the activation segment, there is an 'FFVDGD' motif. Other motifs seem lacking from SteC: there is no glutamic acid in the αC helix, the closest glutamic acid seems to be in a loop (E272). There also no APE motif (Figure 3.6b). The position of the catalytic loop can be predicted to be located between β6 and β7, following the HD motif. However, the position of the activation segment is less clear. The activation segment usually starts at the DFG motif, in this case FFVDGD, and ends at the APE motif. However, the APE motif is absent. Furthermore, αEF starts immediately after β9. Taken together, this implies that the activation loop was comprised of one serine residue (S369) at the junction of β9 and αEF. Alternatively, the activation loop may extend through αEF, from S369: SQESRQQRSVS. This would mean a further three serine residues in the vicinity (S372, S377 and S379), all of which could be the site of phospho-activation (Figure 3.6b). Alternatively, SteC may not require activation by phosphorylation, as suggested by its lack of an arginine in the HxD ('HD') motif, as explained in Section 1.3.3. The canonical sequence

motifs of PKA and the hypothetical motifs of SteC are highlighted on the structures in Figure 3.7b.

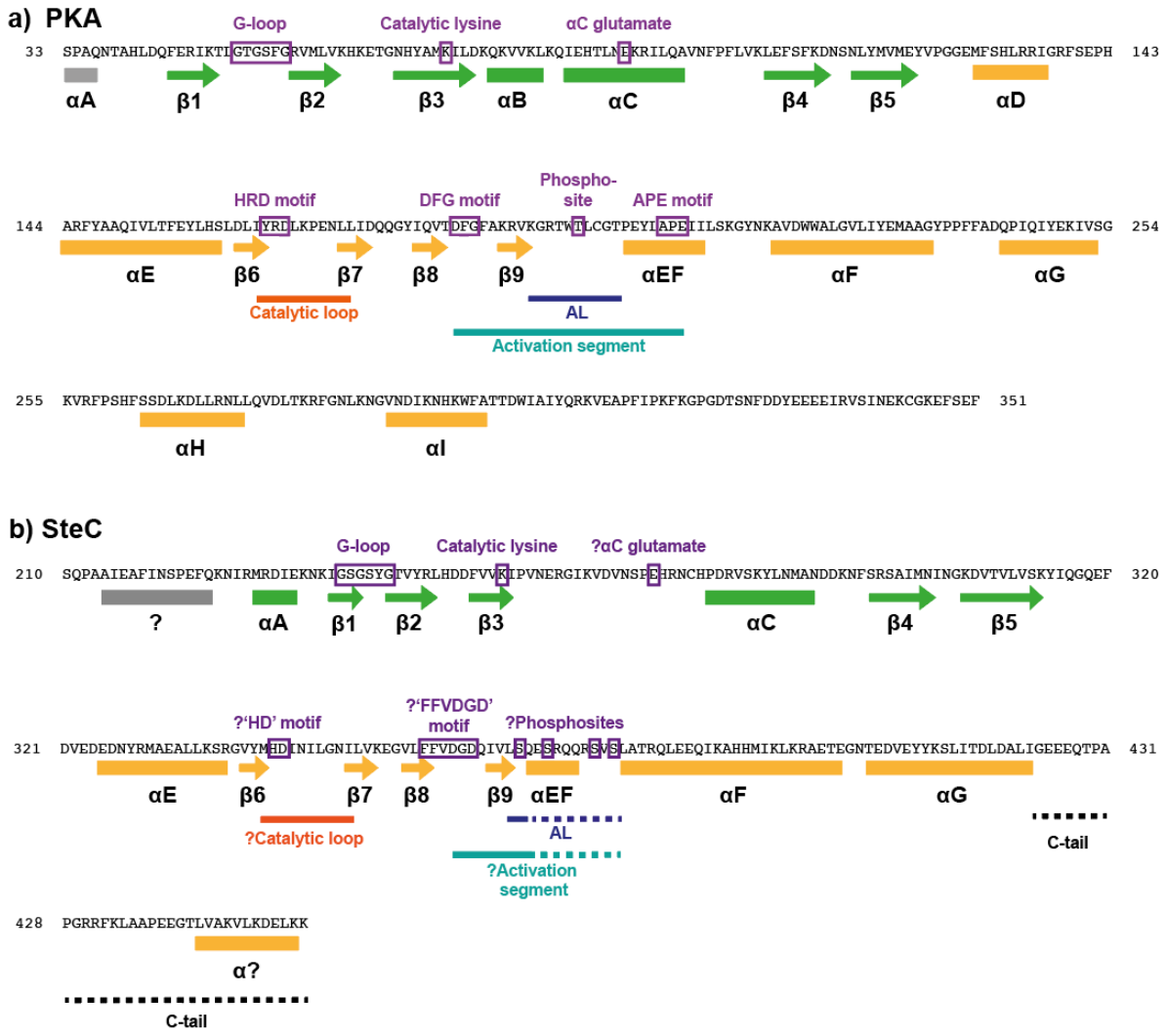


Figure 3. 6: Secondary structural map of canonical kinase features

a) Key sequence motifs (purple) are demonstrated alongside secondary structure elements: α-helices and β-sheets comprising the N-lobe (green) and C-lobe (orange); the catalytic loop, activation segment and activation loop are highlighted. Adapted from Sheetz & Lemmon (2022)

b) SteC is displayed in similar style, with secondary structure elements as per AF prediction in the predicted N-lobe (green) and C-lobe (orange); hypothetical key motifs (purple), and possible catalytic loop, activation segment, activation loop and C-tail are highlighted, with dotted lines referring to non-canonical features.

AL, activation loop; CL, catalytic loop

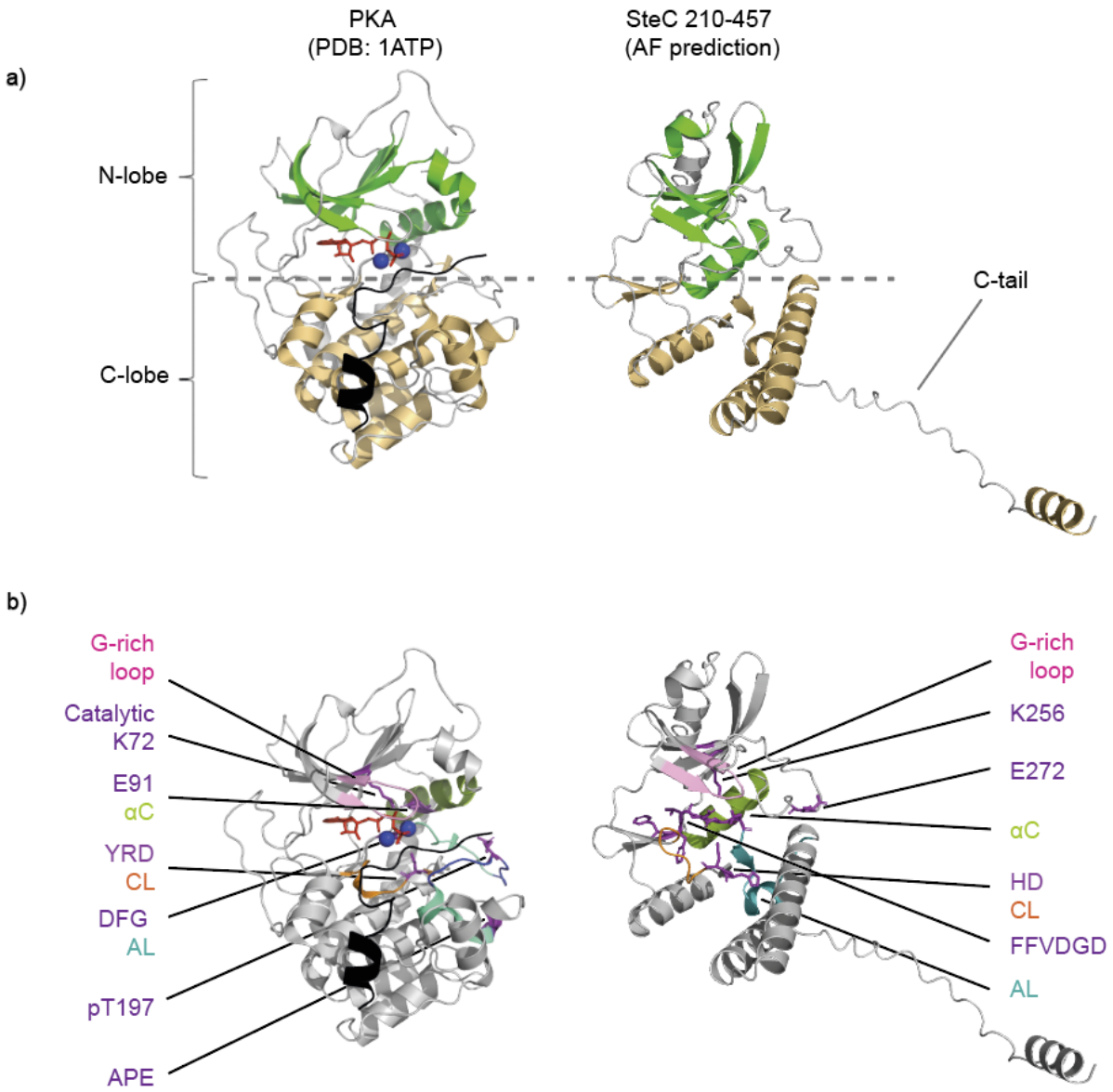


Figure 3. 7: Comparing PKA crystal and predicted SteC kinase domain structures

a) Cartoon representation of PKA (left) in complex with peptide substrate (black), ATP (red) and magnesium (blue) (PDB: 1ATP, left) and SteC kinase domain AF prediction (right) in the same orientation. Secondary structural elements of the N-lobe (green), C-lobes (light orange) and loops (grey) are compared, and the catalytic cleft is indicated by a dotted line. The C-tail is highlighted.

b) Canonical motifs highlighted in PKA (left), and those hypothesised for SteC (right), including catalytic residues (purple), glycine-rich loop (pink), αC helix (green), CL (orange) and activation segment and loop (teal).

PDB, Protein Database.

As well as using BLASTp to identify putative homologues or similar proteins, the AF predicted structure can be used as the input for structural homology searches, either against the PDB or AF databases using FoldSeek (van Kempen et al., 2024). By this method, I identified several proteins from non-*Salmonella* pathogenic bacteria as observed in the BLASTp search,

including *Cedecea* sp. P7760, *Sodalis praecaptivus* and *Erwinia mallotivora* that have predicted structural homology to SteC. When limiting the target database to human, results are dominated by structural homology to the region of SteC between amino acids ~200-360, corresponding to the putative kinase domain N-lobe, also the region of highest pLDDT in the AF structure. The top three structural similarity hits were SBK3, RAF1 and BMPR1B, with probability scores of 1.00, 0.91 and 0.80 respectively, where 1 is the highest confidence and 0 is the least. All aligned to SteC in the region of their kinase domain N-lobes. To identify potential structural homologs to the N-terminal domain of SteC among the human proteome, that may be masked by more significant homology of the kinase domain, the FoldSeek search was limited to amino acids 1-193 in the AF predicted structure. Results are summarised in Table 3.1. Searching by this method yielded top hits IQGAP1, FRMPD1 and IQGAP3 with probability scores of 0.11, 0.1 and 0.1 respectively, much lower than for the kinase domain. On interrogation of the homologous domains in the top 10 target proteins, a common theme was the presence of domains involved in membrane localisation or lipid binding, namely C2, FERM, PH1 and retinol binding domains.

Gene name	UniProt ID	Function	Hom. domain	Score
IQGAP1	P46940	Regulates dynamics and assembly of actin. Binds to activated Cdc42, does not stimulate it. Associates with calmodulin	C2	0.11
FRMPD1	Q5SYB0	Stabilises membrane-bound GPSM1	FERM	0.10
IQGAP3	Q86VI3			0.10
ELAP2	A8MWY0	Regulator of the BMP signalling pathway		0.09
FRMPD4	Q14CM0	Positive regulator of dendritic spine. Required for excitatory synapse transmission. Binds PI(4,5)P	FERM	0.09
RET5	P82980	Intracellular transport of retinol. Fatty-acid binding protein family		0.08
SNHG32	A0A024RCN7			0.08
ELAPOR1	Q6UXG2	May protect from cell death by cytosolic vacuolisation, apoptosis and cell proliferation via HSPA5 interaction		0.06
AFAP1L1	Q8TED9	?Podosome and invadosome formation	PH1	0.06
CRABP2	P29373	Transports retinoic acid to the nucleus. Fatty-acid binding protein family		0.04

Table 3. 1: Top 10 hits searching SteC 1-193 AF predicted structure against human AF proteome with FoldSeek

SteC AF prediction was restricted to amino acids 1-193 and used to perform a structural homology search against the AF human proteome using FoldSeek. The gene name and UniProt ID of the top 10 hits, ranked by FoldSeek score, are listed. Function is summarised, where available in UniProt. Where annotated in UniProt, the domain homologous to SteC 1-193 is named. The score ranges from 1 (identical) to 0 (no similarity).

3.1.3 Conserved sequence motifs across serovars

Using the same database of clinical isolates as described previously (Section 3.1.1), the key catalytic motifs of SteC proposed in Figures 3.6 and 3.7, namely the glycine-rich loop; catalytic lysine; αC glutamate; 'HD' motif; 'FFVDGD' motif; and possible activation loop were aligned by sequence logo (analysis performed by Yan Li, Figure 3.8).

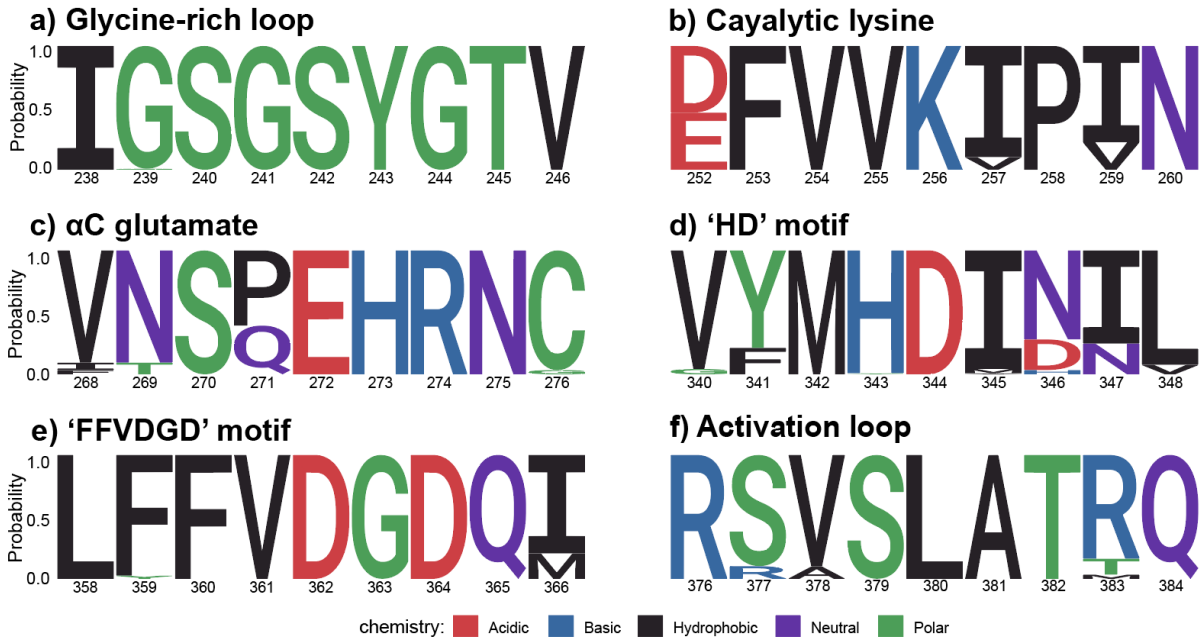


Figure 3. 8: Sequence logos for hypothetical motifs of SteC across *Salmonella* serovars

Regions of interest were identified by comparison of the AF predicted structure with canonical kinase PKA (Figure 3.6 and 3.7). These were compared in the same 879 genomes from *Salmonella* clinical isolates as demonstrated in Figure 3.2. Sequence logos for these motifs were generated using the ggseqlogo package in R. Performed by Yan Li, Hinton Laboratory, University of Liverpool.

The residues proposed to be required for catalytic activity were conserved in this analysis and do not contain any highly variable regions. This provides some evidence of the evolutionary importance of the proposed catalytic residues across *Salmonella* serovars.

3.2 Biophysical and structural features of SteC

3.2.1 SteC expression and purification

To gain understanding of the biophysical and structural properties of SteC, recombinant protein of high quality and purity is needed. Therefore, I wanted to express and purify full-length wild-type and K256H (kinase-dead) mutant SteC (Poh et al., 2008) trialling different solubility and affinity tags and expression systems. First, bacterial expression in pET49b with an N-terminal GST-6His-tag and a 3C cleavage site and NiNTA affinity purification was performed. The recombinant protein from the NiNTA purification was impure and GST and 6His-tag removal with 3C cleavage was inefficient (Figure 3.9a). Furthermore, the protein eluted in the void volume on size exclusion chromatography (SEC) (not shown). Next, SteC was expressed in pET28 with a C-terminal 6His tag and purified to test if the N-terminal position of the tag prevented adequate purification. This procedure produced a larger quantity of protein, however impurities were still present after NiNTA (Figure 3.9b) and the protein eluted largely in the void volume on SEC (Figure 3.9c). Therefore, to improve the purity of the protein, I cloned N-terminally TwinStrep-tagged SteC into pET28. This construct expressed in lower amounts but purified with a resin affinity approach (Figure 3.9d, left). Unfortunately, the product still eluted in the void volume on SEC (not shown). The UV absorption spectrum of this protein demonstrated scattering, suggesting protein aggregation, which improved slightly on performing protein purification at 4 °C. These observed difficulties with producing SteC in *E. coli* indicate that SteC may require a eukaryotic expression system for correct folding, and therefore I tested SteC production in a *Spodoptera frugiperda* (Sf9) insect cell baculovirus expression system. This yielded protein of similar quantity and purity (Figure 3.9d, right), that was also prone to aggregation and eluted in the void on SEC (not shown). For this reason, in all subsequent assays where full-length SteC was required for this chapter, the protein was expressed in *E. coli* or Sf9 and purified with a single affinity resin purification step.

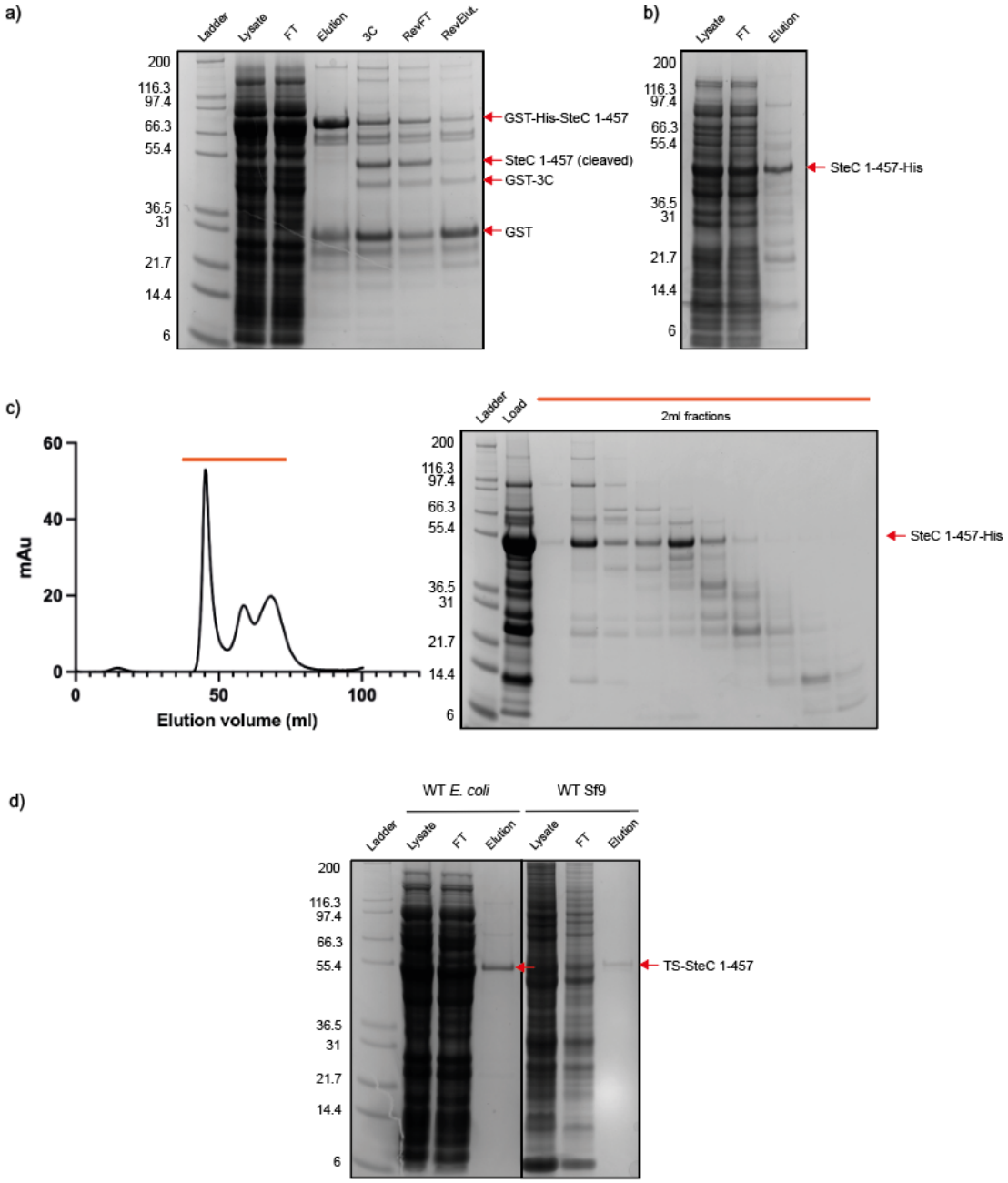


Figure 3. 9: Full-length SteC protein purification attempts

a) Representative gel of GST-His-3C-SteC 1-457 protein expressed in BL21 cells and purified with NiNTA resin. Samples at purification stages were loaded on SDS-PAGE for analysis.

b) Representative gel of SteC 1-457 harbouring a C-terminal His tag (SteC 1-457-His) expressed in BL21 cells and purified with NiNTA resin. Samples of different stages of the purification were taken and loaded on SDS-PAGE for analysis.

c) Left: SEC profile of SteC 1-457-His. Elution from NiNTA was loaded on a Superdex75 16/60 column. Right: Fractions of SEC were analysed on SDS-PAGE.

d) Representative gel of SteC 1-457 with N-terminal Twin-Strep tag expressed in BL21 cells and purified with Streptactin resin. Samples at purification stages were analysed by SDS-PAGE.

FT, flow through; TS, Twin-Strep; 3C, 3C cleavage.

3.2.2 SteC fragment design and purification

As the full-length protein could not be produced at high purity, quality and quantity, the next step was to produce smaller constructs for structural and biophysical analysis (Figure 3.10).

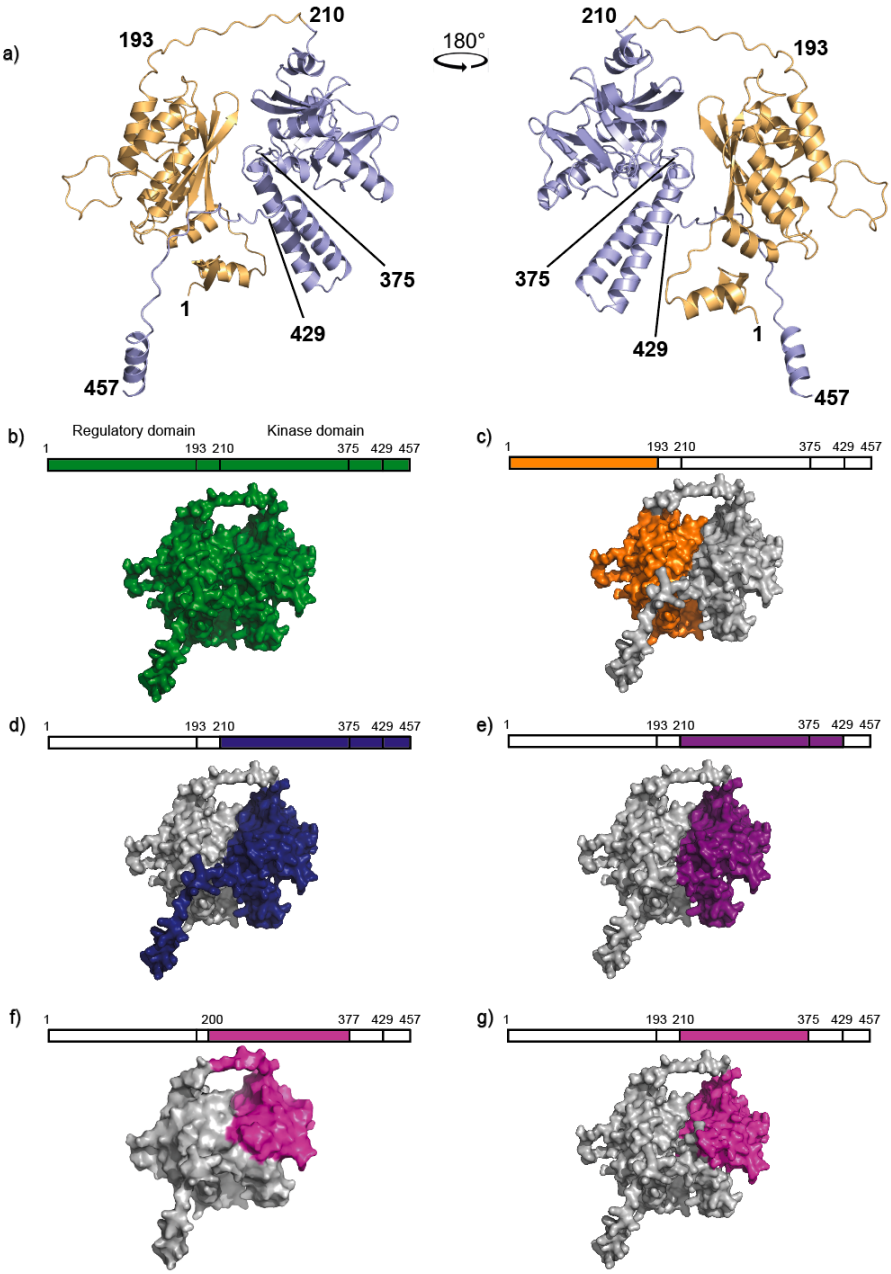


Figure 3. 10: Predicted structure-guided construct design

a) AF predicted structure of SteC coloured light orange (1-209) and light blue (210-457) shown in two orientations, and demonstrating the residues used to create protein fragments of the N-regulatory and C-kinase domains. b) – g) coloured depictions of the surface SteC constructs, including b) the full-length 1-457 and fragments c) 1-193, d) 210-457, e) 210-429 and f) 200-377 and g) 210-375 used going forward. Note: colours will be used to denote relevant protein fragments for the rest of this thesis.

SteC kinase domain WT and K256H mutant constructs 210-457, 210-429, 200-377 and 210-375 were expressed in *E. coli* from the pET28 vector with a non-cleavable C-terminal His tag. They all expressed well and NiNTA resin purification achieved relatively pure protein (Figure 3.11a, left and 11b, left). Kinase domain fragments 210-457 and 210-429 eluted partially in the void volume of the SEC (Figure 3.11a, right). Protein eluted at the expected volume for a monomer. Other fragments eluted (Figure 11b, right) at the expected elution volume with no protein in the void. A gradient of aggregation propensity was seen as detected by UV absorption spectra. Fragment 210-457 displayed a propensity to aggregate at room temperature over minutes to hours, 210-429 less so, and fragments 200-377 and 1-193 were stable. To prevent aggregation, proteins were continuously kept on ice and all purification steps were performed at 4 °C with pre-cooled buffers. For assays performed earlier on in the PhD, SteC 200-377 was used. To make as accurate a comparison between the three kinase domain fragments as possible, I later expressed and purified 210-375 (removing the first 10 amino acids that are predicted to be in the linker region) with the same tag and same method of purification. This fragment behaved similarly to 200-377. For most experiments for biophysical characterisation I used the K256H mutant, reported in the literature to be kinase inactive (Poh et al., 2008) to avoid potential heterogeneity caused by autophosphorylation. A list of all proteins used in this chapter is provided in Table 3.2.

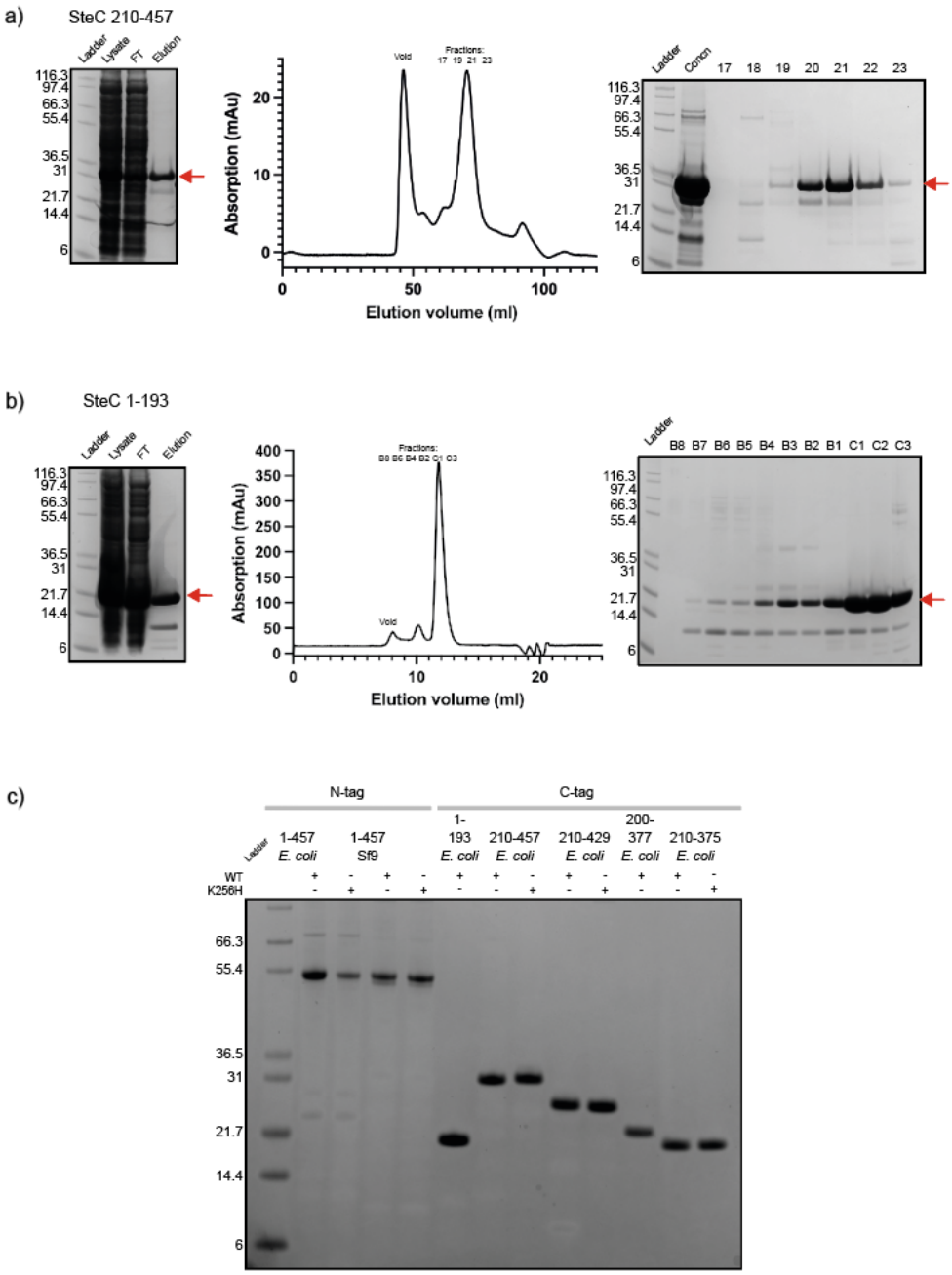


Figure 3. 11: Construct expression and purification

a) SteC 210-457 purification: samples from NiNTA resin purification loaded by SDS-PAGE, gel stained with Coomassie (left). Elution was loaded on s SEC column demonstrating an elution peak (middle). Samples from peaks were analysed by SDS-PAGE, gel stained with Coomassie (right).

b) SteC 1-193 purification following the same approach as for a).

c) All recombinantly expressed and purified SteC proteins used in this chapter: 5 µL of protein at 5 µM analysed by SDS-PAGE and gel stained with Coomassie.

FT, flow through; concn, concentrate.

Amino acid	Variants	Tag	Vector	Expression system	Purification strategy	Molecular weight (Da)	Outcome
1-457	WT K256H	N-term GST-6His-3C	pET49b	*E. coli*	NiNTA	79,812	Impure after NiNTA, eluted in void in SEC, Aggregation prone ++
	WT K256H	C-term 6His	pET28	*E. coli*	NiNTA	53,143	Impure after NiNTA, eluted in void in SEC, Aggregation prone ++
	WT K256H	N-term TS	pET28	*E. coli*	Streptactin resin	56,316	Aggregation prone, effective streptactin purification, eluted in void in SEC
1-457	WT K256H	N-term TS	pACEBac	Sf9	Streptactin resin	56,316	Aggregation prone, effective streptactin purification, eluted in void in SEC
1-193	WT	C-term 6His	pET28	*E. coli*	NiNTA SEC	23,413	Large quantity expressed NiNTA and SEC successful, sable at RT
210-457	WT K256H	C-term 6His	pET28	*E. coli*	NiNTA SEC	29,050	Large quantity expressed NiNTA and SEC successful, aggregation-prone at RT++
210-429	WT K256H	C-term 6His	pET28	*E. coli*	NiNTA SEC	26,005	Large quantity expressed NiNTA and SEC successful, aggregation-prone at RT
200-377	WT	C-term 6His	pET28	*E. coli*	NiNTA SEC	21,022	Large quantity expressed NiNTA and SEC successful, stable at RT
210-375	WT K256H	C-term 6His	pET28	*E. coli*	NiNTA SEC	19,791	Large quantity expressed NiNTA and SEC successful, stable at RT

Table 3. 2: Summary of constructs cloned, expressed and purified for this chapter

3.2.3 Analysis of the secondary structure of SteC

CD spectroscopy monitors the absorbance of circularly polarised light by chiral molecules to determine the presence of secondary structure, or lack thereof, in proteins in solution. I used it to assess the secondary structure of SteC 1-457, 1-193, 210-457, 210-429 and 200-377 (Figure 3.12).

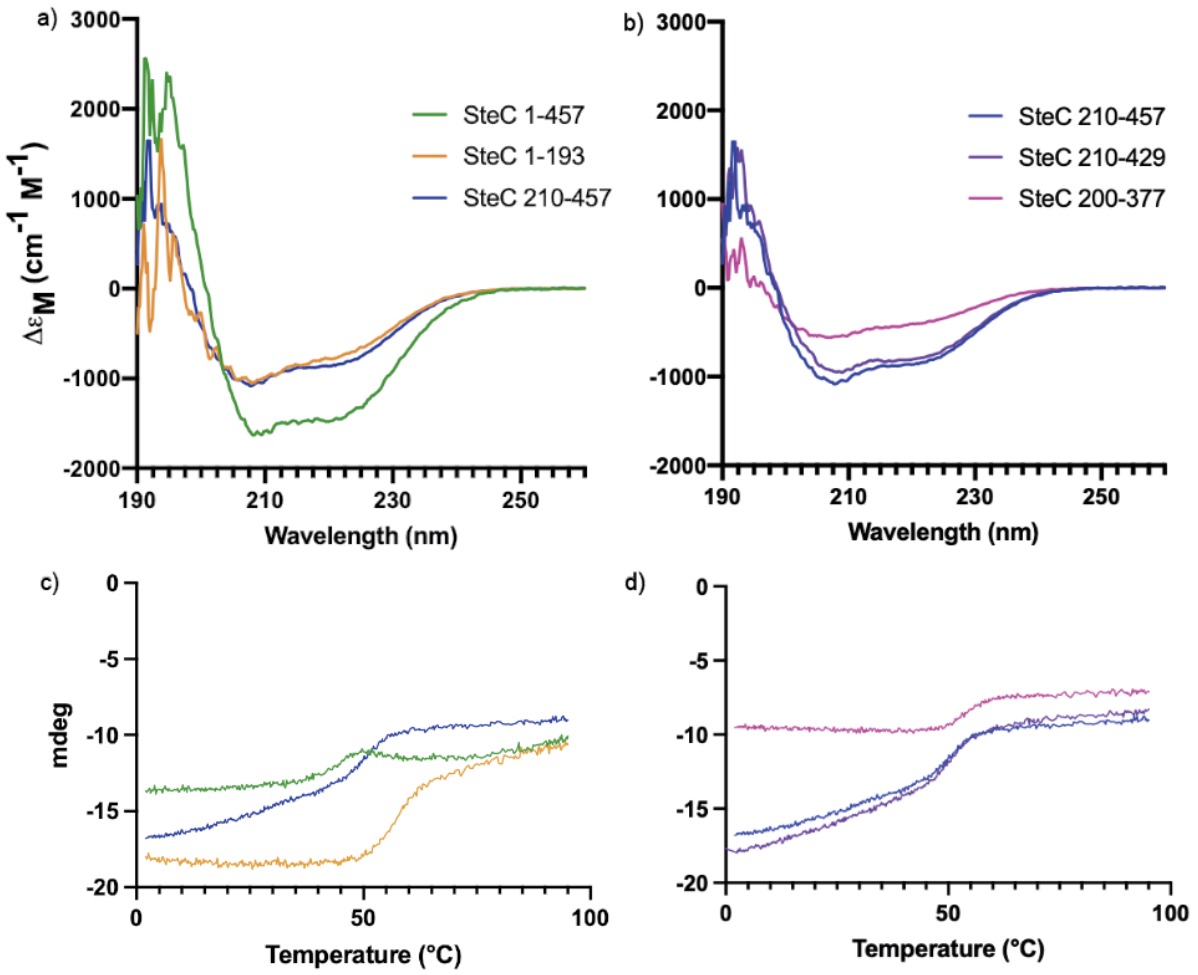


Figure 3. 12: CD analysis of SteC, two domains and kinase domain fragments

CD spectra at 25 °C were recorded between 190 and 260 nm for a) SteC 1-457, two domains and b) three kinase domain fragments. Melting curves were recorded from 2 – 95 °C and the signal at 220 nm was plotted for c) full-length protein and two domains and d) the three kinase domain fragments, demonstrating melting points for each of the proteins. Delta epsilon molar units ($\Delta\varepsilon$) were used to enable comparison between the fragments and assess their relative contribution to secondary structure. SteC K256H mutants were used for full-length and kinase domain fragments. Performed with the supervision of Laura Masino (Structural biology STP).

As expected, given its size, the full-length protein demonstrated a stronger signal at 220 nm than either of the two domains separately (Figure 3.12a). Domains 1-193 and 210-457 demonstrated α-helical content, which classically give a negative inflection at wavelengths 208 and 222 nm. Comparing kinase domain fragments (Figure 3.12b), there was a loss of amplitude of the spectra as the construct reduces in size, suggesting a loss of secondary structure elements for smaller constructs. The most dramatic loss of secondary structure was between SteC 210-429 and 210-375, while 210-457 was similar to 210-429 suggesting 376-429 is alpha-helical and 430-457 may lack secondary structure. These observations support the AF prediction. The proportion of each protein consisting of α-helices, β-sheets, turns and random coils can be estimated by deconvolution of the CD spectra. However, noise in the wavelength range 190-195 precluded accurate estimations.

CD spectroscopy can also be used to analyse protein stability, where a protein in solution is exposed to a temperature gradient and the change in CD amplitude of the sample is used to assess its unfolding (melting curve) and thermal stability. Well-folded globular proteins characteristically have a single cooperative unfolding event while multidomain or less stable proteins or constructs with a tendency to aggregate may show multiple or more gradual melting profiles, as in the case of the full-length SteC (Figure 3.12c). In contrast, the two individual domains encompassing amino acids 1-193 and 210-457 each demonstrated single unfolding events, at 58 °C and 50 °C respectively. The N-domain has a flat signal until 50 °C, suggesting a stable protein in that temperature range. The kinase domain, however, shows a sloping curve prior to an inflection point at 47 °C. This result may reflect aggregation, or else an independent uncooperative unfolding of a protein region prior to the main unfolding event. To interrogate the kinase domain secondary structure further, all three kinase domain fragments were compared (Figure 3.12d). The single cooperative unfolding event of the shortest construct at 54 °C suggests this protein has a stable core region and tertiary structure. The curves for the two larger constructs showed gradual change below 50°C, suggesting the region 376-429 contributes non-cooperative events occurring at a lower temperature and independent of the region 210-375 likely a consequence of protein-protein interaction. The lack of significant difference between SteC 210-429 and 210-457 suggests that the C-terminal region may not contribute to the stability of the longer constructs.

Solution 1D ^{1}H nuclear magnetic resonance (NMR) is a convenient approach to assess overall fold of a protein and in addition can provide some information about potential self-association.

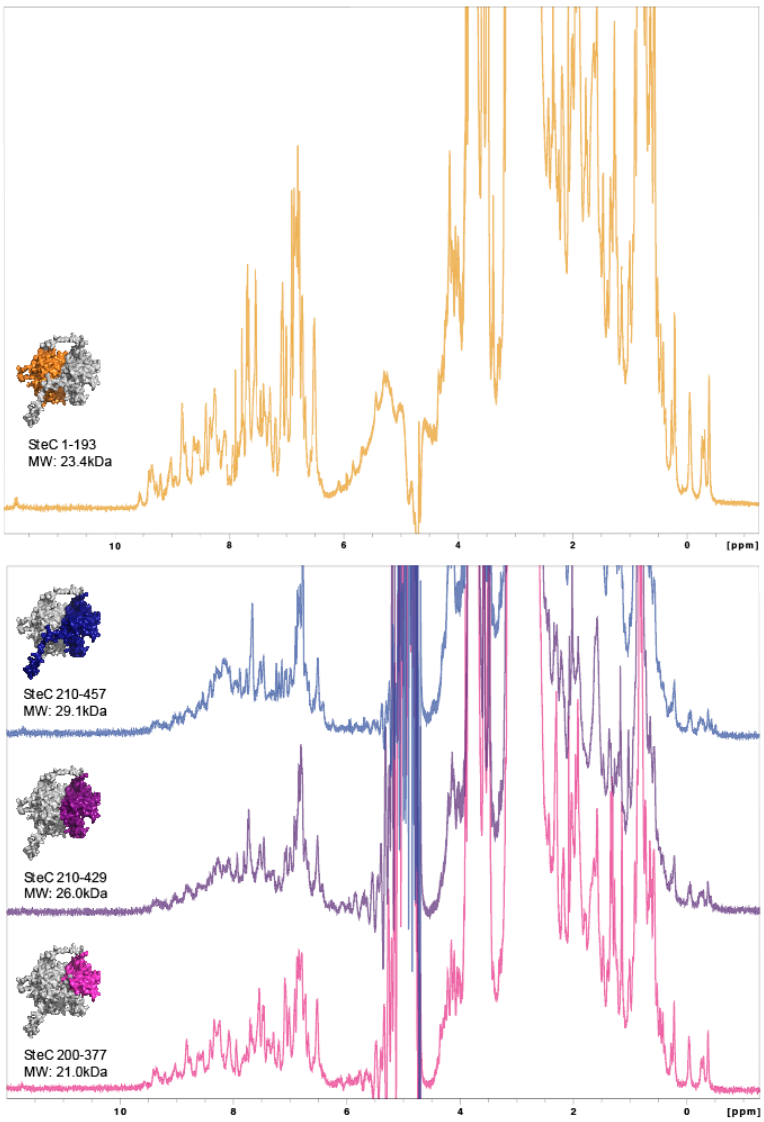


Figure 3. 13: 1D ^{1}H-NMR spectra of SteC constructs

Spectra were collected for SteC 1-193 and K256H mutants of 210-457, 210-429 and 200-377 expressed in *E. coli* with a C-terminal 6His tag at 25 °C on a Bruker AVANCE spectrometer operating at 800 MHz.

The NMR spectra for SteC N-domain and kinase domain fragments (Figure 3.13) have the signature of folded proteins, highlighted by the dispersion of resonances in the amide-proton region (6-10 ppm) and the presence of peaks with chemical shifts below 0 ppm (indicative of the presence of a hydrophobic core). The N-terminal domain has the sharpest resonances in these regions indicating that this construct is well folded and monodisperse in solution. Of the kinase domain fragments, the smaller construct SteC 200-377 has the sharpest spectrum. The other two SteC constructs 210-457 and 210-429 have a similar footprint to SteC 200-377 with a degree of peak broadening partially due to the higher molecular weight as well as a contribution from chemical exchange from inter- or intra-molecular interaction. The 1D ^{1}H-NMR spectrum of SteC 210-457 shows little evidence of an additional inherently disordered

region overlapped to the spectrum of SteC 210-429, suggesting, differently from the AF prediction, that the C-tail may be flexible and in exchange between a free and folded state interacting with the SteC core molecule.

In summary, based on the NMR data, SteC 1-193 is well folded, and within the kinase domain, the region 200-377 is the most folded, while region 378 to 457 may be less well folded and may form dynamic interactions with the rest of the molecule or other molecules in solution.

3.2.4 Analysis of the oligomeric state of the kinase domain using SEC-MALLS

Kinases can auto-regulate through dimerisation in a trans auto-activating or -inhibitory manner (Lavoie et al., 2014b). To assess if SteC exists as a monomer, dimer or higher order oligomer in solution, size exclusion chromatography coupled to multi-angle laser light scattering (SEC-MALLS) was performed, a technique that combines standard SEC with detectors that provide an absolute molecular weight of the species in solution across the elution profile. The SEC-MALLS analysis of kinase domain fragments is shown in Figure 3.14.

The estimated molar mass (dotted line) is flat and at the level expected for a monomer in all three cases (Figure 3.14a-c), suggesting the kinase domain does not dimerise. There is a slight shift in the elution peaks (full lines) for SteC 210-457 (Figure 3.14a) and SteC 210-429 (Figure 3.14b), suggesting SteC 210-457 and SteC 210-429 interact with the column matrix.

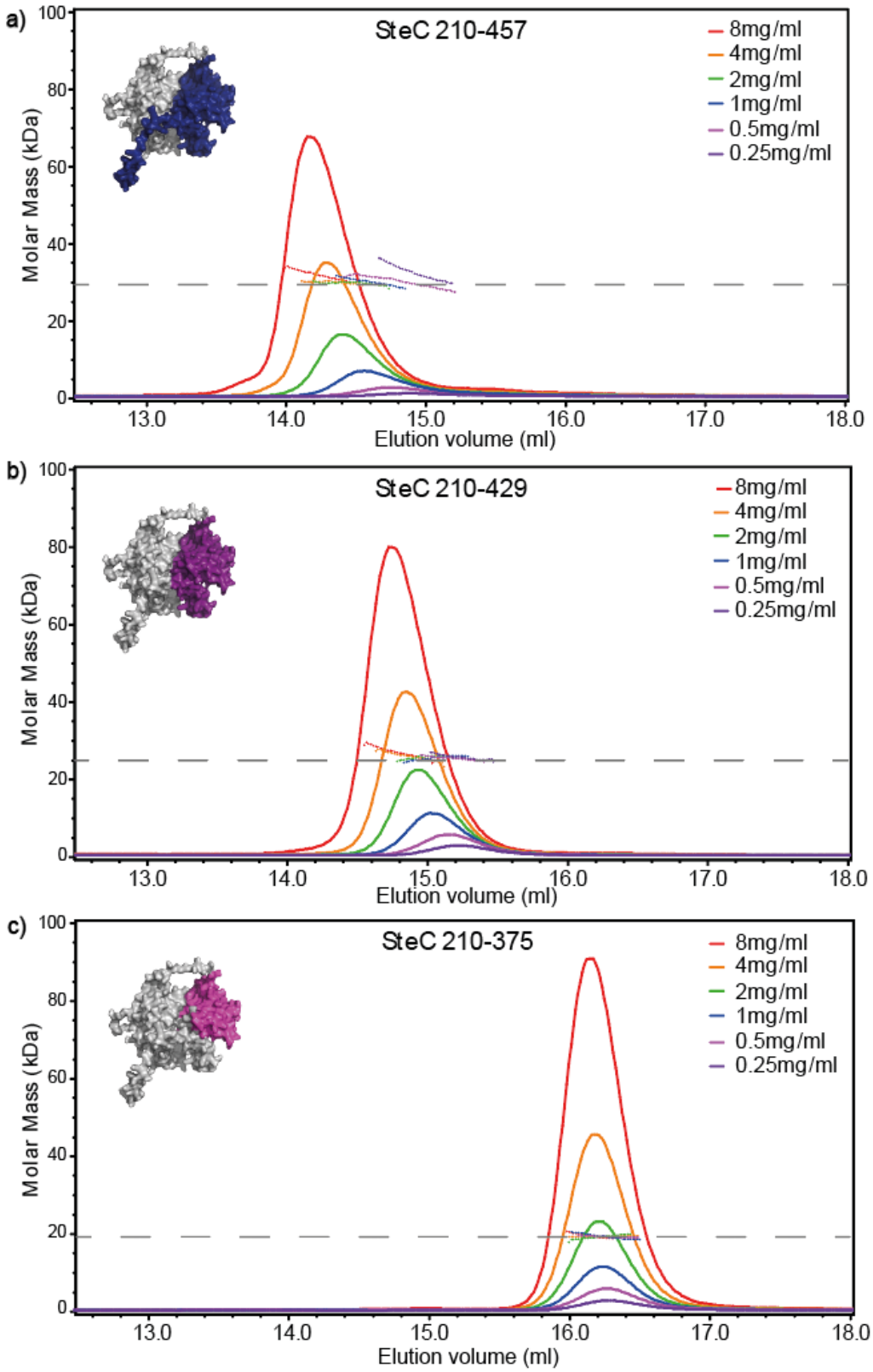


Figure 3. 14: SEC-MALLS of SteC kinase domain fragments

Each full line represents a size exclusion peak at the indicated protein concentration. Each dotted line represents the corresponding estimated molar mass of the eluting species. The fragments SteC 210-457, SteC 210-429, and SteC 210-375 were assessed with K256H mutation to prevent auto-phosphorylation. Performed with Ian Taylor (Francis Crick Institute).

3.2.5 Assessing the solution shape of kinase domain fragments using SAXS

Small angle X-ray scattering (SAXS) can give low-resolution insight into the molecular shape of proteins in solution. A SEC-SAXS analysis was performed for the SteC kinase domain constructs, with the help of Diego Esposito (Francis Crick Institute).

SteC construct	Conc. (mg/ml)	I(0)	Rg	Rc	D_{max} (Å)	MW from seq. (kDa)	MW from scatter (kDa)
210-457 K256H	8	4.8E-02	30.8	15.3	101	29.1	28
210-429 K256H	8	3.7E-02	26.4	14.9	94	26.0	24
210-375 K256H	8	3.6E-02	19.4	15.3	61	19.8	18

Table 3. 3: SEC-SAXS parameters

Data were collected at SOLEIL synchrotron, Gif-sur-Yvette, France, by Diego Esposito, Rittinger Laboratory. I(0) is the forward scattering intensity dependent on the scattering angle. I(0) and the wavelength of the incident radiation can be used to perform a Guinier analysis to calculate estimates of molecular size: Rg, Rc, D_{max} and MW. Rg is the radius of gyration. Rc is the cross-sectional radius of gyration (Rg=Rc for a spherical molecule, Rg>Rc for elongated cylindrical molecules). D_{max} refers to the maximum molecular dimension in angstrom. MW, molecular weight.

One key observation derived from the SAXS measurements is that the molecular weight calculated from the sequence resembles the molecular weight derived from the SAXS data in all three cases (Table 3.3), validating this approach and supporting the claim that all three constructs are monomeric in solution.

The radius of gyration, Rg increases with molecular weight as would be expected. However, the Rc does not. SteC 210-457 and 210-429 have Rg > Rc, suggesting they are both elongated, whereas SteC 210-375, with a Rg similar to Rc, is roughly spherical. In other terms, SteC 210-457 and 210-429 have similar estimates of maximum molecular dimension (101 Å and 94 Å, respectively) while the estimate for SteC 210-375 is much smaller (61 Å) (Table 3.3).

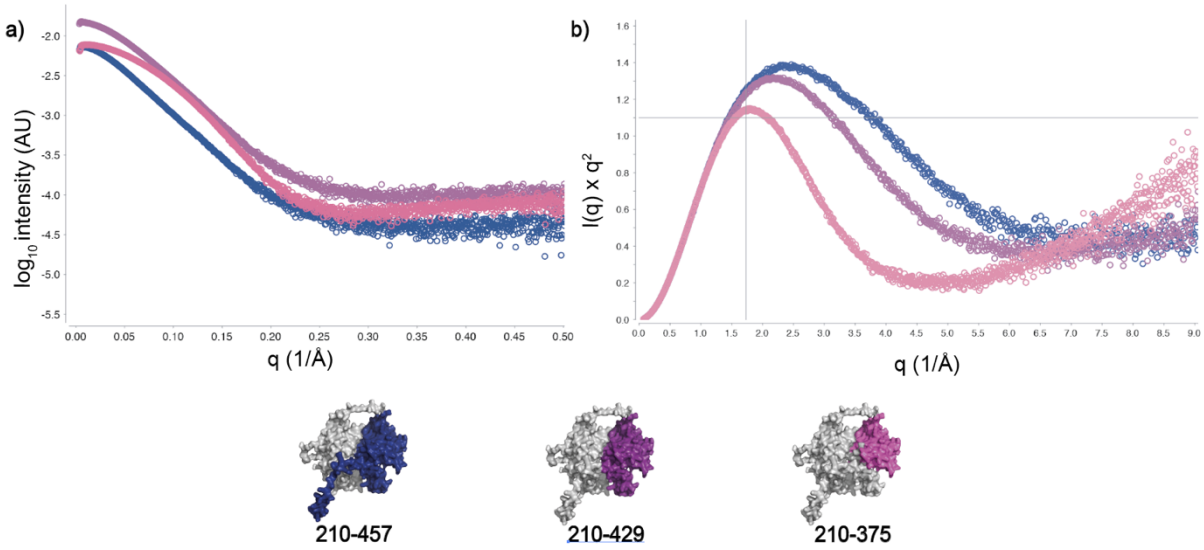


Figure 3. 15: Scattering profiles and Kratky plots for SteC kinase domain fragments

a) X-ray scattering profiles for the SteC K256H 210-457 (blue), 210-429 (purple) and 210-375 (pink), demonstrating the intensity of the scatter as a function of the scattering angle.

b) Kratky plots for the same proteins. In the normalised Kratky plot the position of the cross indicates the theoretical maximum of a globular protein.

The X-ray scattering profiles for the three constructs are shown in Figure 3.15a, from which values in Table 3.3 are derived. A Kratky plot of the scattering profiles can provide further information about the folding state of the protein and the flexibility of multi-domain molecules. This suggests that SteC 210-429 and 210-457 are more elongated than SteC 210-375, as previously described, and demonstrate increased protein flexibility (indicated by the downtrend at higher q values) (Figure 3.15b). SteC 210-375, however, does not show this same flexibility.

SAXS envelopes represent the conformational ensemble average of molecules in solution (Figure 3.16a). This method was used to detect unstructured regions and dynamic conformational changes in different SteC fragments. The model of SteC 210-375 agrees with the globular fold of the AF prediction (Figure 3.16b, left), whereas SteC 210-429 and SteC 210-457 (Figure 3.16, middle and right) are predicted to be more elongated than the smaller construct, in contrast to the AF prediction.

Put together, the SAXS data suggest all three fragments are monomeric, supporting the SEC-MALLS data. SteC 210-375 (the N-lobe) of SteC kinase domain is well-folded and roughly spherical. Region 376-429 adds length to the molecule and may have a degree of flexibility

and is not tightly associated with the N-lobe. The C-tail also adds some length, but does not dramatically change the Rc, suggesting that it is not contributing to the folded domain.

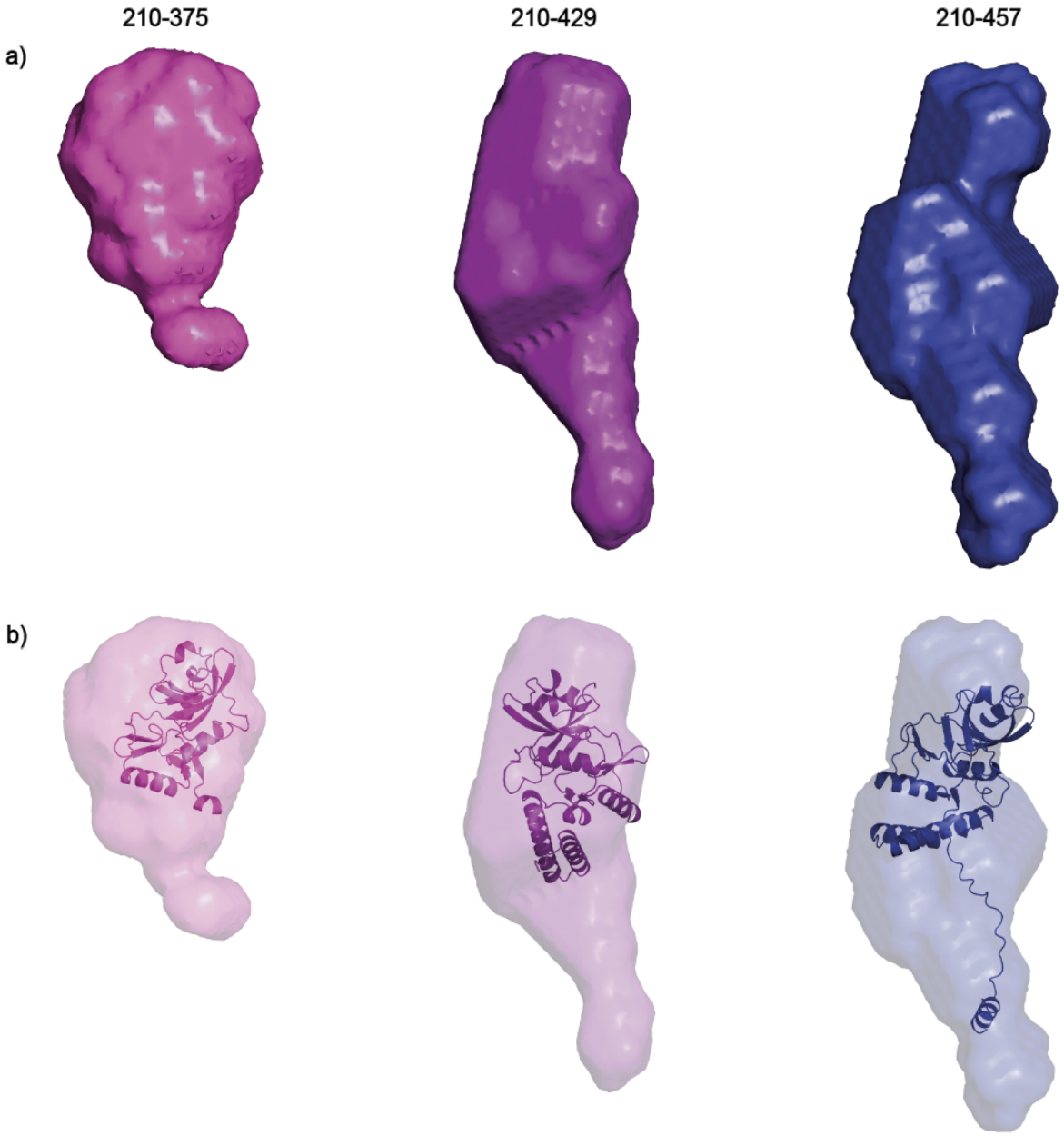


Figure 3. 16: SAXS low-resolution model of the shape the kinase domain fragments

a) Low-resolution *ab initio* models derived from the SAXS data calculated using DAMMIF (Franke et al., 2017). All fragments are SteC K256H mutants and contain a C-terminal 6His tag.

b) Translucent SAXS *ab initio* model with AF-prediction of the same fragment manually superimposed demonstrating extent of agreement.

3.3 SteC interaction with host substrate FMNL1

FMNL1 is a 1100-amino acid, 122 kDa protein that was recently proposed to be a substrate of SteC (Walch et al., 2021). It is the first of a three-member family with intrinsic actin polymerisation activity which is mediated through its C-terminal FH2 domain, and regulated through an inhibitory interaction between the N-terminal domain and the C-terminal DAD domain. This interaction is disrupted by binding of the small GTPase CDC42 to GBD thereby leading to activation (Figure 3.17) (Breitsprecher & Goode, 2013). Phosphorylation of the N-terminal region of FMNL1 by SteC has been recently demonstrated (Walch et al., 2021). Therefore, given that the structure of FMNL1 1-458 has been solved by X-ray crystallography as a dimer interacting via the coiled-coil region 385-458 (Figure 3.17) (Kühn et al., 2015), these domain boundaries were used for recombinant protein expression.

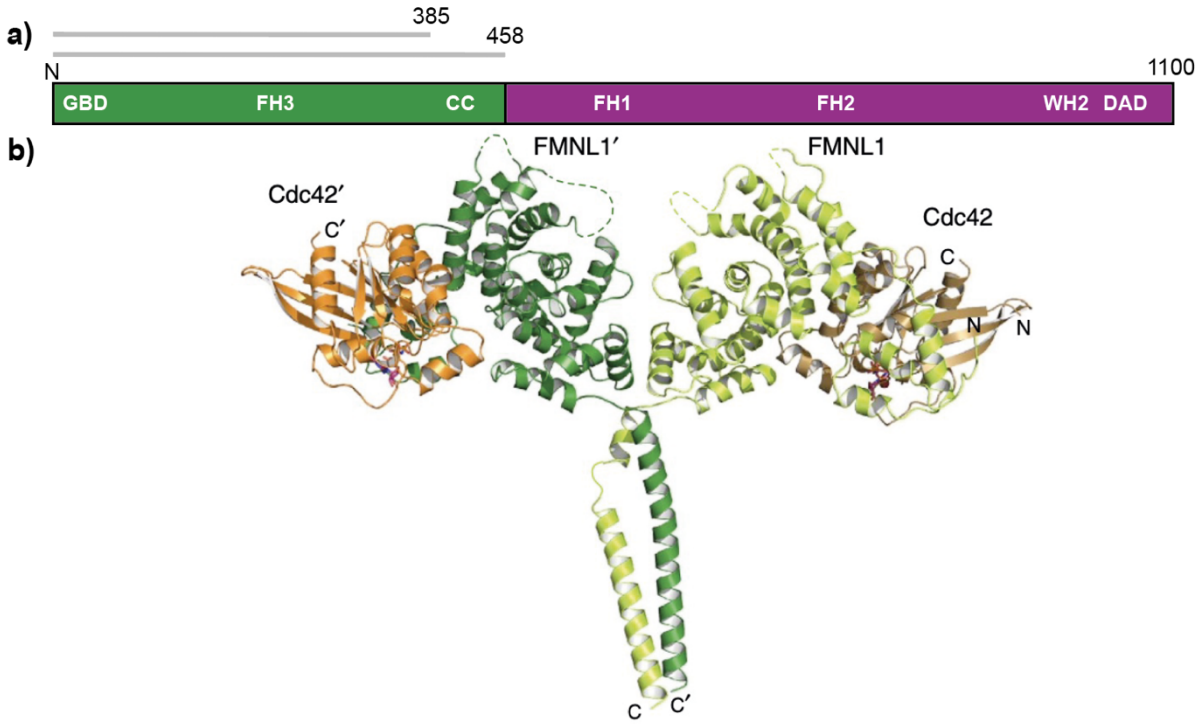


Figure 3. 17: FMNL1 domain map and structure of the N-terminus

a) Domain map of FMNL1. CC, coiled coil; DAD, diaphanous autoregulatory domain; FH, formin homology domain 1,2 or 3; GBD, GTPase protein binding domain; WH2, Wiskott-Aldrich homology domain. Grey bars refer to FMNL1 constructs expressed and purified.

b) Crystal structure from Kühn et al., 2015 (PDB: 4YDH) of a FMNL1 1-458 dimer (green and lime green) in complex with activating Rho GTPase Cdc42 (orange and brown).

3.3.1 FMNL1 expression and purification

Although FMNL1 1-458 was expressed in *E. coli* in the previous study, in this study I only obtained low yield impure protein by the reported method (not shown). Therefore, the construct

was expressed in Sf9 cells instead using a N-terminal 6His-tag. NiNTA affinity purification produced protein of sufficient purity (Figure 3.18a) for further biochemical analysis. FMNL1 1-458 eluted in a broad peak on SEC (Figure 3.18b), which may be due to a monomer-dimer equilibrium of the species in solution, given the same protein appears to be present across the peak on denaturing SDS-PAGE analysis (Figure 3.18c).

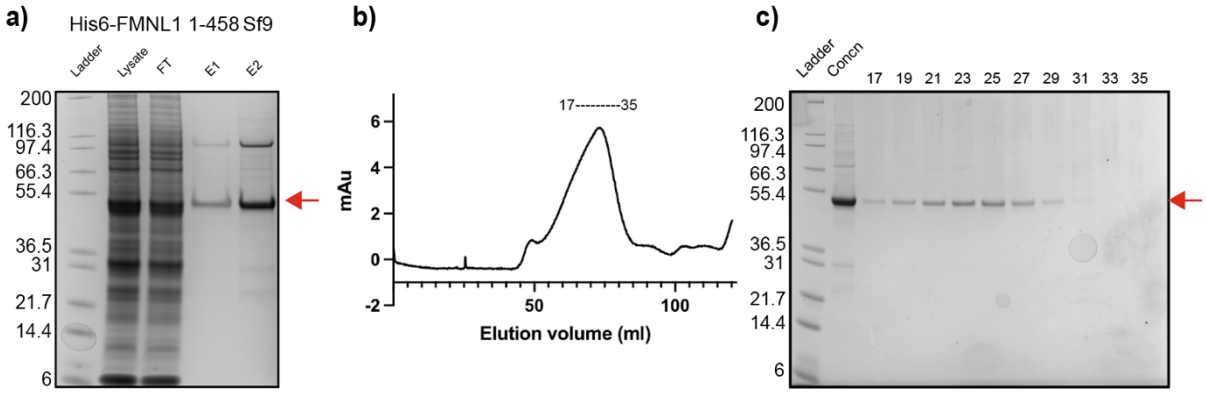


Figure 3. 18: Purification of FMNL1 1-458 recombinantly expressed in Sf9 cells

a) NiNTA resin purification of 6His-FMNL1 1-458 expressed in Sf9 cells. Serial elutions were performed with 300 mM and 500 mM imidazole.

b) FMNL1 1-458 elution profile on SEC.

c) SEC fractions on SDS-PAGE of the marked alternate fractions.

Concn, concentration; E, elution; FT, flow through

3.3.2 Establishing the kinase activity of SteC towards FMNL1

To verify FMNL1 as a substrate of SteC, I performed *in vitro* kinase assays (Figure 3.19). I used Pro-Q Diamond phos-stain, a stain that detects phosphate groups covalently attached to proteins analysed by SDS-PAGE. This method demonstrated phosphorylation of FMNL1 in an SteC- and ATP-dependent manner (Figure 3.19, lane 8). Interestingly, SteC phosphorylation was detected without pre-incubation with ATP suggesting either autophosphorylation or phosphorylation by an insect kinase before the assay. The stronger signal after incubation with ATP suggests either autophosphorylation or alternatively phosphorylation by a contaminating insect cell kinase present in the purified protein sample.

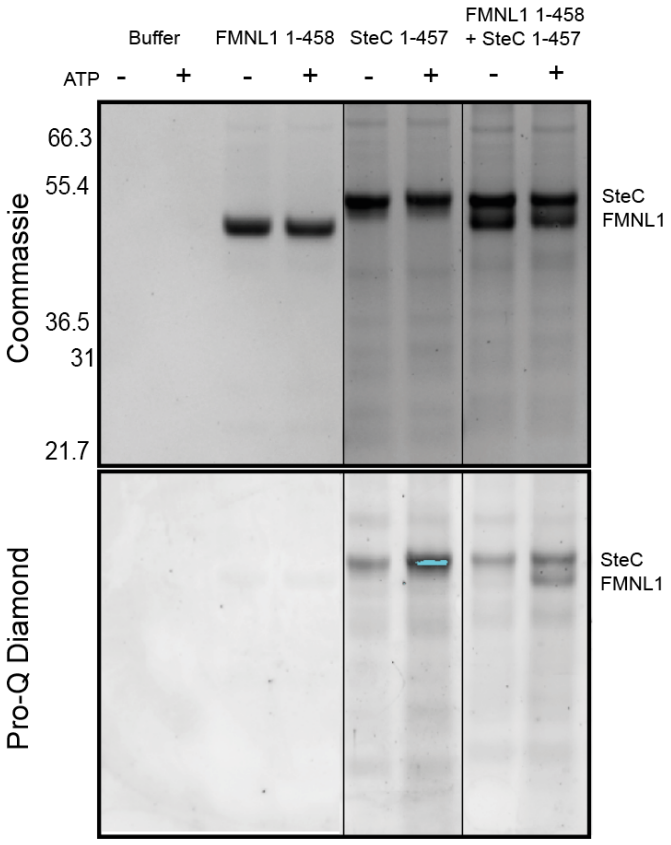


Figure 3. 19: Phos-stain kinase assay with SteC and FMNL1

Kinase assay performed by incubating SteC 1-457 and FMNL1 1-458 both expressed in insect cells alone or together at 10 µM with 100 µM ATP at 30 °C for 30 min. Reactions were quenched by the addition of SDS-loading buffer and analysed with PAGE, Coomassie staining (top) and Pro-Q-Diamond Phos-stain (bottom). Data representative of 3 repeats.

3.3.3 Identification of phosphorylation sites on FMNL1

A previous study performed by Walch et al. (2021) identified phosphorylated residues (phosphosites) on FMNL1 by mass spectrometry of recombinant proteins. To validate these findings and identify a region of FMNL1 that could be used as a substrate binding partner for binding assays, crystallisation and kinase assays, FMNL1 1-458, expressed and purified from insect cells, was incubated alone (to test if insect cell expression induces FMNL1 phosphorylation due to presence of a contaminating kinase) or with Sf9-expressed SteC 1-457 and ATP before being subjected to phospho-mass spectrometry analysis. FMNL1 1-458 was unphosphorylated after expression and purification from insect cells. On incubation with SteC and ATP, 4 phosphorylation sites were identified: S184, S199, T203 and T207 (Table 3.4 & Figure 3.20). These sites agree with those previously reported. However, S99, S194, S195 and T118 as described by Walch et al (2021) were not detected in our hands.

Starting amino acid	Peptide sequence	Mass (Da)	Phosphorylation site
178	NKPLEQ**pS**VEDLSK	1565.73	S185
178	NKPLEQ**pS**VEDLSKGPPSSVPK	2315.14	S185
191	GPPSSVPK**pS**R	1090.56	S199
199	SRHL**pT**IK	933.48	T203
199	SRHL**pT**IKLTPAHSR	1695.89	T203
201	HL**pT**IKLTPAHSRK	1580.86	T203
201	HL**pT**IKL**pT**PAKSR	1532.73	T203 & T207

Table 3. 4: Phosphorylated peptides of FMNL1 after incubation with SteC and ATP

FMNL1 1-458 and SteC 1-457 were expressed in Sf9 cells. FMNL1 was analysed by phospho-MS either alone or after incubation with SteC and ATP. No FMNL1 phosphosites were detected in the control condition. Phosphorylated FMNL1 peptides after incubation with SteC and ATP are reported here. Data analysed with MaxQuant.

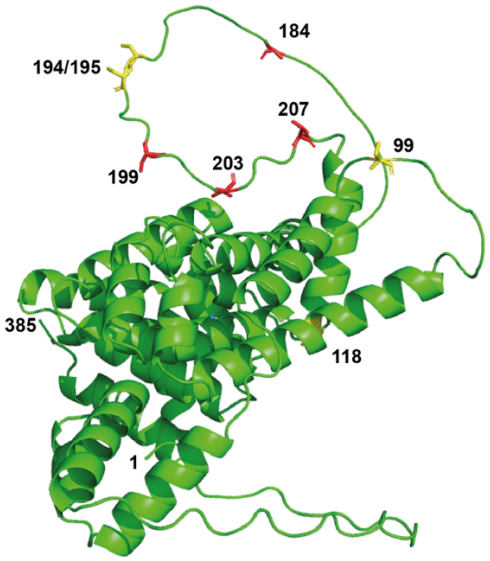


Figure 3. 20: Identifying SteC-induced FMNL1 phosphosites

AF predicted structure of FMNL1 in ribbon representation, used to visualise the flexible loops that were not resolved in crystal structure. Phosphosites are annotated with sticks: red those identified by this study and Walch et al., 2021, yellow those identified only by Walch et al.

3.3.4 Assessing substrate binding using biolayer interferometry

To quantify binding of the SteC kinase domain to the region of FMNL1 encompassing the identified phosphorylation sites, 4 peptides were selected (Table 3.5) based on our and previously published phosphosites (Table 3.5 & Figure 3.20) and synthesised by Dhira Joshi from the Chemical Biology STP.

Name	Starting amino acid	Sequence	Amino acid length
pep_Q85	Q85	QKLKSYVDTGGVSRKVA	17
pep_R113	R113	RVQESTQVLRE	11
pep_N178	N178	NKPLEQSVEDLSKG	14
pep_K190	K190	KGPPSSVPKSRHLTIKLTPAHSRK	24

Table 3. 5: Peptides covering FMNL1 phosphosites synthesised for binding studies

Sites of phosphorylation previously identified in Walch et al (2024) or identified by our assay highlighted in yellow and red, respectively.

SteC kinase domain constructs were expressed in *E. coli*. Biolayer interferometry was used to detect binding, a method whereby a ligand-protein interaction is measured by immobilising one binding partner on a sensor. When the sensor is submerged in a solution of the other, light interference shone across the sensor creates a time-dependent signal of binding. Protein dissociation can be detected concomitantly. In my case, biotinylated peptides were loaded onto streptavidin sensors, then submerged in a solution of SteC at a range of concentrations. K_d values were calculated at steady state with 8-16 concentration points.

No binding was observed between SteC 210-457 and 3 of the 4 FMNL1 peptides: pep_Q85, pep_R113 or pep_N178 (Supplementary Figure 1). Binding of SteC 210-457 was detected to pep_K190, a 24mer peptide spanning 3 of the 4 phosphosites identified by our phosphoMS analysis (Figure 3.21, left). Binding affinity was estimated with a K_d of 180 nM (+/- 20 nM). There was incomplete dissociation on washing the sensor in buffer at higher concentrations, which may represent a degree of aggregation of SteC on the sensor (Supplementary Figure 1). The assay was performed at room temperature, a condition that has previously caused aggregation of the kinase domain fragment (especially 210-457).

SteC 210-429 also showed binding to pep_K190 (Figure 3.21, right) and not to Q85, R113 or N178 (data not shown). However, it required higher protein concentrations to produce a binding curve, ranging from 1.5 to 200 µM. The binding affinity was estimated with a K_d of 102 µM +/- 15 µM, representing a 500-fold weaker affinity than SteC 210-457. These data indicate that SteC kinase domain binds to amino acid region 190-213 of FMNL1, and that the C-tail is required for high-affinity binding.

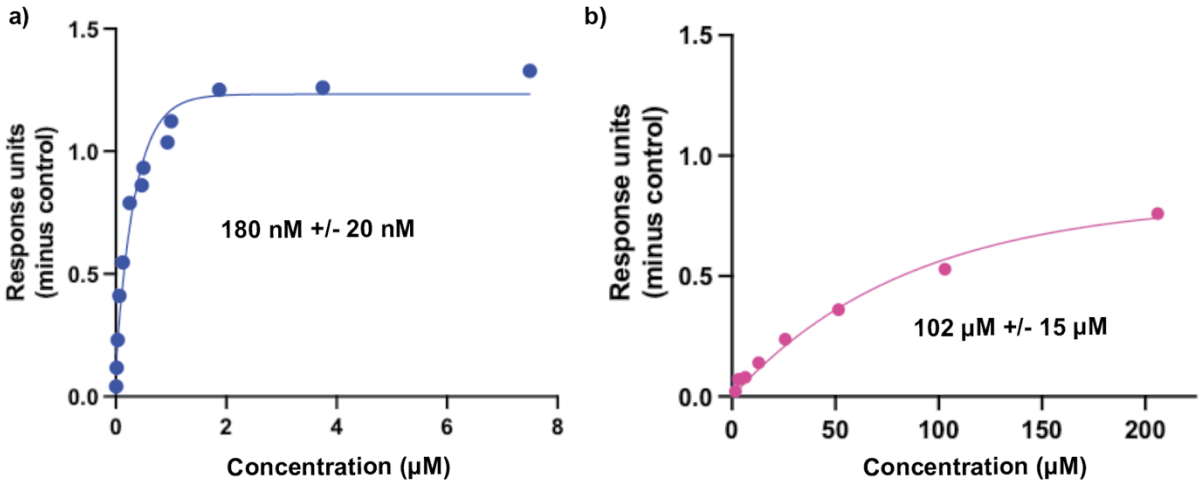


Figure 3. 21: Binding curves with estimated K_d values for SteC 210-457 and 210-429

A timepoint was chosen prior to plateau phase in the BLI signal, and the signal at this point plotted against concentration to achieve the binding curves, for a) SteC 210-457 and b) 210-429. K_d for each protein was calculated using total least squares fit with Prism 10. Experiment and analysis performed with the help of Laura Masino (Structural biology STP).

3.3.5 HDX-MS

Hydrogen-deuterium exchange mass spectrometry (HDX-MS) looks at the exchange of hydrogen molecules in the protein backbone. It can be used to assess the relative exposure of protein residues to solution and hence to infer structural information about a protein and changes in this exposure in the presence of a ligand. To interrogate the SteC kinase domain - FMNL1 peptide interaction structurally, HDX-MS was performed with SteC 210-457 K256H alone and in the presence of the K190 peptide. Peptide coverage of SteC was complete though regions 210-211, 254-261, 315-321, 354-359, 360-361, 366-368, 412-422 and 451-457 – a total of 46/248 residues – were only covered by one peptide (Figure 3.21a). Several regions of SteC undergoing slower deuterium exchange in the presence of the peptide (Figure 3.22b) were observed. Significant regions were mapped onto the AF predicted structure (Figure 3.22c). Interestingly, the differential exchange data in the double alpha helix (388-396) fits with the canonical kinase substrate binding site on the C-lobe with orientation into the catalytic cleft (Hanks & Hunter, 1995). The other regions with differential exchange, 281-284 and 312-320 in the kinase N-lobe, may represent non-canonical substrate binding sites or might undergo allosteric confirmational changes upon peptide binding to the C-lobe region. There is no evidence for a region with an increase in exposure in the presence of the peptide.

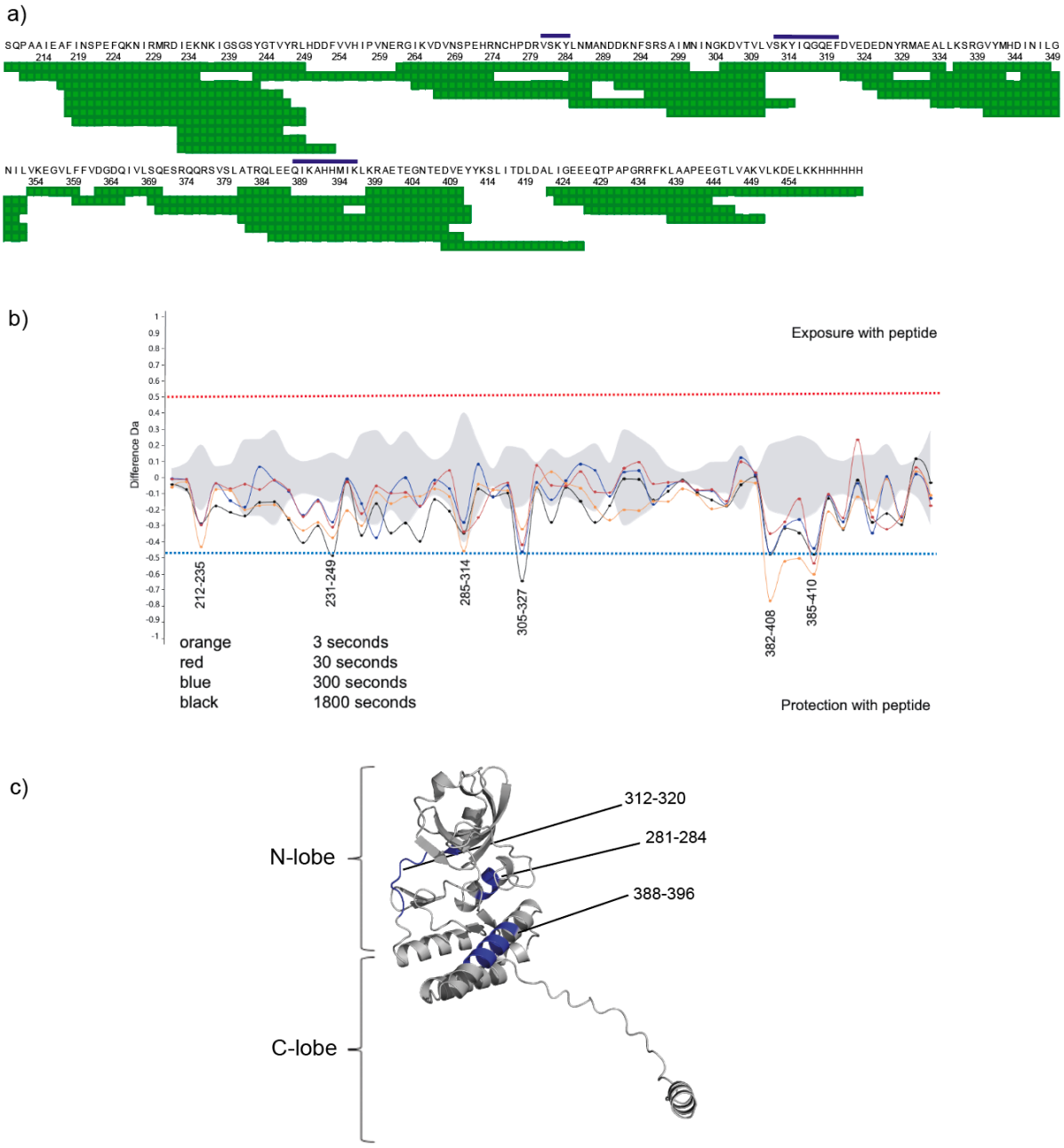


Figure 3. 22: Hydrogen deuterium exchange of SteC 210-457 K256H and FMNL1 peptide

SteC 210-457 K256H with C-tag was expressed in *E. coli* and purified by standard procedures. It was either analysed alone or with FMNL1 pep_K190. Deuterium was added to each of the two samples prior to quenching and trypsin digestion. MS analysis performed by Sarah Maslen (Proteomics STP).

a) Peptide coverage of the experiment by residue (green). Total 53 peptides, 100% coverage, 3.80 redundancy. Regions significantly protected in presence of peptide are highlighted above (blue).

b) Hydrogen deuterium exchange is plotted as a relative difference in deuterium uptake of the SteC 210-457 alone or in the presence of K190 peptide across the amino acid range.

c) Regions of SteC demonstrating protection in the presence of the peptide (at least 0.45 difference in all peptides covering this region) are highlighted on the AF predicted structure in blue.

3.3.6 Crystallisation attempts

Our data indicate that SteC 210-375 is the best-folded, most stable and most globular construct and hence may be amenable to crystallisation. However, the structural features predicted with low confidence by AF and the non-canonical C-lobe of the kinase domain were deemed the most interesting and unexplored, especially given their involvement in binding to the substrate peptide. Therefore, I attempted crystallisation of all three of the kinase domain fragments, under various conditions (Table 3.6). I attempted crystallisation of wild-type SteC kinase domain fragments with non-hydrolysable nucleotide adenylyl-imidodiphosphate (AMPPNP) to stabilise the kinase domain, aid crystallisation, and demonstrate the orientation of the co-factor. I also attempted crystallisation of the same fragments with the K256H mutation to avoid any potential heterogeneity deriving from auto-phosphorylation. Additionally, a non-biotinylated FMNL1 peptide was used as a ligand in co-crystallisation trials with 210-457 and 210-429 fragments.

The most promising result from the crystallisation trials was stacked crystals for smallest kinase domain fragment SteC 210-375 (Figure 3.23a). However, prior to setting of optimisation trays with the aim to improve the crystals for X-ray crystallography, the structure of SteC 202-375 was published by another group (Dai et al., 2024). This structure is interrogated in the discussion (Section 6.2.4).

Construct	Mut.	Concn.	Ligands	Screens	Temp	Findings
210-457 C-6His	K256H	Apo: 360 µM (11 mg/ml) +ligands: 320 µM (10 mg/ml)	+/- 5 mM AMPPNP and 500 µM FMNL1 peptide	JB Kinase, Pi-PEG PEG 4K optimisation (pH 6-8)	4 & 20 °C	Microcrystals in PEG4K
210-457 C-6His	K256H S379D	Apo: 360 µM (11 mg/ml) +ligands: 320 µM (10 mg/ml)	+/- 5 mM AMPPNP and 500 µM FMNL1 peptide	JB Kinase, Pi-PEG PEG 4K optimisation (pH 6-8)	4 & 20 °C	Microcrystals in PEG4K
210-429 C-6His	WT	Apo: 427 µM (10 mg/ml)	+/- 5 mM AMPPNP	Index HT, JCSG Core I-IV PACT Premier HT, PEG Suite I-II PEG Ion HT, Rx HT	4 & 20 °C	~50% aggregation at 5 days
210-429 C-6His	K256H	Apo: 383 µM (9 mg/ml) +ligands: 329 µM (8 mg/ml)	+/- 5 mM AMPPNP + 1 mM FMNL1 peptide	Index HT, JB Kinase, JCSG Core I-IV, Morpheus I-II, PACT Premier, PEG Ion HT, Rx HT, PiPEG, Wizard I-IV	4 °C	~50% aggregation at 5 days
210-429 C-6His	K256H S379D	Apo: 516 µM (15 mg/ml) +ligands: 344 µM (10 mg/ml)	+/- 5 mM AMPPNP + 1 mM FMNL1 peptide	Index HT, JB Kinase, JCSG Core I-IV, Morpheus I-II, PACT Premier, PEG Ion HT, Rx HT, PiPEG, Wizard I-IV	4 & 20 °C	Promising UV-positive microcrystals demonstrate some diffraction using the in-house x-ray machine (see Figure 3.23)
210-375 C-6His	WT	Apo: 733 µM (15 mg/ml), 1018 µM (20 mg/ml), 1272 µM (25 mg/ml)	+ 5 mM AMPPNP	Index HT, JB Kinase, JCSG Core I-IV, Morpheus I-II, PACT Premier, PEG Ion HT, Rx HT, PiPEG, Wizard I-IV	20 °C	Small UV-positive crystals (see Figure 3.23)

Table 3. 6: Summary of attempted crystallisation conditions

Various kinase fragments and mutants were attempted, often with K256H mutation to prevent partial autophosphorylation (see Chapter 4 for more details). Ligands AMPPNP and FMNL1 peptide were sometimes added. Crystallisation screens were available via the Structural Biology STP. Crystallisation plates were stored in Formulatrix imaging machines at 20 and 4 °C. Images were analysed for at least 10 days for each plate. A summary of the results is described in the final column. Crystallisation plates were set up with the help of Andy Purkiss and Roksana Ogrodowicz from the Structural Biology STP. Concn, concentration; Temp, temperature.

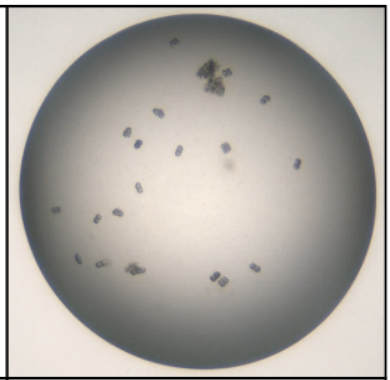
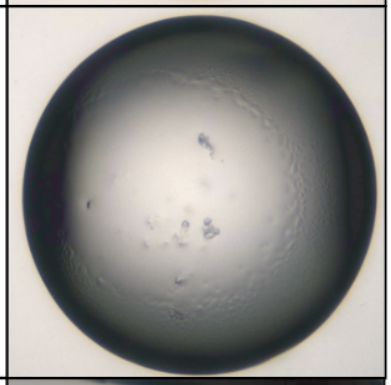
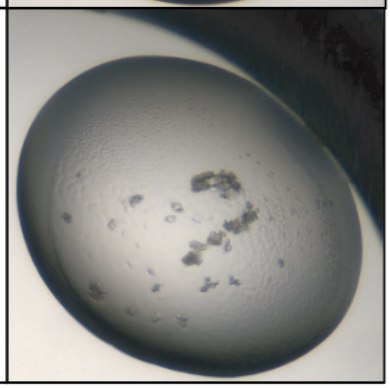
Details of protein, ligand and crystallisation conditions	Visible High Res Image
a) **Protein:** SteC 210-375 **Ligand:** AMP-PNP **Well condition:** 0.1M Tris pH 8.5 25% PEG 4K 20°C **Screen:** NEXTAL PEG **Imaged after:** 24 days	
b) **Protein:** SteC 210-429 K256H S379D **Ligand:** K190 substrate peptide **Screen:** Morpheus **Well condition:** 30% EDO_P8K 0.1M MB3 pH 8.5 0.12M alcohols 20°C **Imaged after:** 24 days	
c) **Protein:** SteC 210-429 K256H S379D **Ligand:** K190 substrate peptide **Well condition:** Na Cacodylate pH 6.5 20% PEG 8K 0.2M Mg Acetate 20°C **Screen:** JCSG Core I **Imaged after:** 17 days	

Figure 3. 23: Promising crystallisation conditions

Crystallisation trials were performed across a range of conditions (Table 3.6). Promising crystallisation conditions are summarised (left), and images of the drops are demonstrated (right). UV imaging demonstrated protein, indicating they are microcrystals. Performed with the supervision of Andy Purkiss (Structural biology STP).

Conclusion

This chapter was directed towards gaining insight into the biophysical properties of SteC, a poorly characterised bacterial effector protein. AF predicts SteC to contain two domains that are connected with a flexible linker region. The kinase domain is made up of a well predicted N-lobe (210-375), and a C-lobe comprising a better predicted double alpha-helix (376-429) and a poorly predicted and largely unstructured C-tail (430-457). Analysis of SteC fragments with biophysical and structural techniques indicated that the N-terminal regulatory domain is well folded. The C-terminal kinase domain is comprised of a well-folded globular N-lobe, that was very recently crystalised (Dai et al., 2024). The C-lobe of the kinase domain is comprised of more flexible α-helical portion and a largely unstructured portion, both of which are dynamic and may transiently interact with the N-lobe. The kinase domain exists as a monomer in solution, there may be transient inter-molecular interactions of the full kinase domain though this could also be attributed to its aggregation-prone nature. I identified key phosphosites in FMNL1 and the peptide encompassing these sites interacted tightly (with nM affinity) by BLI. HDX-MS mapping of the site of binding onto the SteC structure suggested the minimal C-lobe is important for binding. In conclusion, the atypical C-lobe and C-tail of SteC kinase domain interacts dynamically with the rest of the structure and is important for substrate binding. The role of the kinase domain components for enzymatic activity remain to be elucidated.

4. Activity and regulation of SteC

Introduction

Chapter 3 identified important structural features of SteC's N-terminal regulatory domain and C-terminal kinase domain, and demonstrated binding of the kinase domain to a peptide derived from its substrate, FMNL1. In this chapter, I present experiments aimed at characterising its kinase activity. Table 4.1 summarises the methods used to assess kinase activity, to aid interpretation of the data presented in this chapter.

Method	Description	Strengths	Weaknesses
Phos-Stain with Pro-Q-Diamond	Kinase assay samples are analysed by SDS-PAGE followed by specific gel staining, which labels any protein with a phosphate modification.	Can distinguish pre-existing and new phosphosites.	High background.
Autoradiography	Radioactive (γ-P^{32}) ATP is used to spike a kinase assay. Samples are run on a gel, Coomassie-stained and autoradiographic film is used to detect radioactivity of proteins in sample.	Highly specific i.e. minimal background, meaning it can be used to quantify activity. Sensitive to low-level activity.	Only detects phosphorylation that has occurred during the assay.
Phos-Tag gel	Kinase assay samples are denatured with SDS then run on a bespoke acrylamide gel containing phos-tag, a molecule that reacts with phosphate groups and increases the gel running time of phosphorylated proteins, meaning phosphorylation is detected as a shift compared to control a sample.	Theoretically able to distinguish phosphorylated and unphosphorylated species of the same protein in one sample.	Difficult to quantify. Gel shift on phosphorylation of the substrate can vary depending on the protein.
Intact MS	MS can accurately determine the molecular weight of a denatured protein, meaning covalently modified proteins can be detected. Phosphate group adds 80 Da.	Quantitative demonstration of the proportion of phosphorylated versus non-phosphorylated molecules. Identifies multiply phosphorylated species.	Unable to identify site of phosphorylation.
PhosphoMS (recombinant protein)	Upon digestion of a protein, peptides can be detected either with or without phosphosites. The site of phosphorylation on the peptide can also be determined by fragmentation.	Highly specific identification of modified amino acid. Semi-quantitative.	Expensive Not easily scalable Phosphosite identification dependent on peptide detectability.

Table 4. 1: Methods used to assess kinase activity with recombinant expressed protein.

Adapted from (Reinhardt & Leonard, 2023).

4.1 Characterisation of SteC phosphorylation and its effect on SteC enzymatic activity

Kinases often require phosphorylation to become active, and this can be by autophosphorylation or as the substrate of an up-stream kinase (see Section 1.3.3). Studying autophosphorylation *in vitro* can be complicated by auto-phosphorylation of the recombinant protein during protein expression. Kinase expression in a eukaryotic expression system, such as Sf9 insect cells, comes with the added challenge of endogenous kinases, which can also phosphorylate the overexpressed protein. It is widely accepted that *E. coli* strains used for recombinant protein expression do not contain kinases, meaning a protein expressed in *E. coli* will not be phosphorylated, unless it is an autophosphorylating kinase. Therefore, in this section and Section 4.2 I will exploit both systems to a) test if SteC is phosphorylated, b) test if this affects activity, c) test the hypothesis that SteC autophosphorylates and d) test whether SteC might be a substrate for a eukaryotic kinase.

4.1.1 SteC phospho-mass spectrometry identifies two phosphosites

In parallel to the phosphoMS experiment to detect FMNL1 phosphosites (Section 3.3.3), recombinantly expressed SteC 1-457 from Sf9 was assessed before and after incubation with ATP in the absence of FMNL1, performed by Steve Howell (Proteomics STP). In the initial sample, one phosphopeptide was identified, SVSLATR, and using collision induced dissociation the site of phosphorylation was determined to be SV**pS**LATR, i.e. S379 (Figure 4.1a). After incubation with ATP, an additional phoshophosite, S75 was identified. These two residues are highlighted on the AF-predicted structure of SteC (Figure 4.1b). Given the insect cell expression system, it was unclear whether these sites were due to autophosphorylation (S75 during the kinase assay, S379 during expression); an insect cell kinase during expression (S379); or an insect cell kinase contaminating the purified protein (S75). Interestingly, S379 sits in a region predicted to be close to the possible activation loop of the kinase domain (Figure 3.6 and 3.7), and therefore was interrogated further as a potentially biologically relevant site.

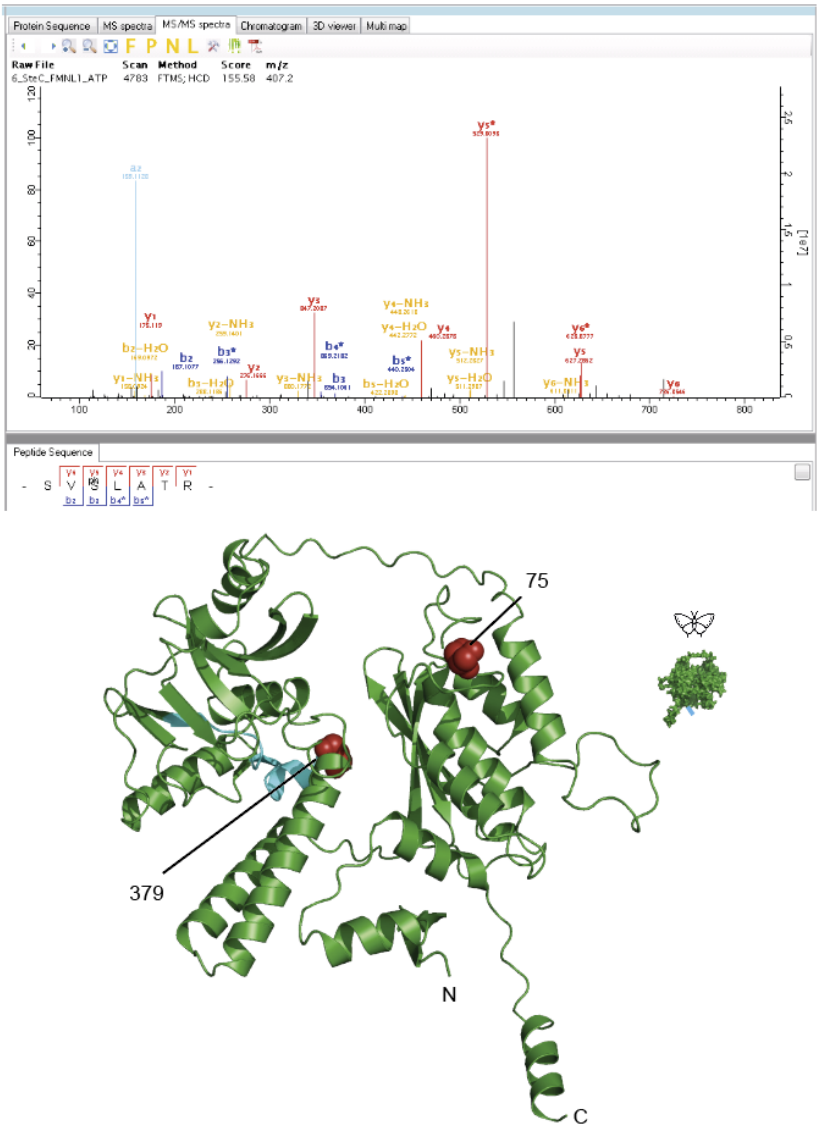


Figure 4. 1: S379 and S75 are phosphosites of SteC

SteC 1-457 was expressed in Sf9 cells and purified with NiNTA affinity resin. Samples of the protein were processed with SDS-PAGE either in this condition or after incubation with ATP for 30 min at 30 °C. The Coomassie-stained gels were cut to isolate the SteC protein bands before dissolving the band and digesting with trypsin. Mass spectrometry was performed by Steve Howell (Proteomics STP).

a) Phosphorylated peptide SVSLATR was identified and phosphorylation site identification was performed by fragmentation with a Lumos Tribrid Orbitrap mass spectrometer obtained for precursor ion at m/z 407.2 processed with Perseus (V1.4.0.2). A screenshot is demonstrated. C-terminal (y) ions and N-terminal (b) ions are presented here. y ions at 627 and 726 and the b3 ion at 354 indicate presence of phosphorylation. 529 represents unphosphorylated y5 minus H_2O.

b) Serine 379, in the kinase domain (left) is highlighted in the AF predicted structure (green) at the C-terminal end of the possible activation loop (blue). On incubation with ATP, a second phosphosite was identified: Serine 75 in the N-terminal domain (right).

m/z, mass to charge ratio in atomic mass units (amu); cps, counts per second.

4.1.2 Protein constructs produced for this chapter

Given S379 of SteC is phosphorylated when the protein is expressed in Sf9 and this site may occupy a key position in a hypothetical activation loop of the kinase domain, I attempted to mimic the unphosphorylated and phosphorylated protein by cloning, expressing and purifying phosphoresistant serine-to-alanine (S379A) and phosphomimetic serine-to-aspartic acid (S379D) mutants of SteC. SteC constructs were also cloned with both N- and C-terminal tags, in case the tag position interfered with kinase activity. Compared to the proteins produced for Chapter 3, no differences were seen in the quantity of expression of these mutants and tagged proteins, or their purity after resin affinity and SEC purification (data not shown). The SteC constructs used in this chapter are listed in Table 4.2.

SteC length	Expression system	Tag	Variants	Purification
1-457	*E. coli*	N-TS	WT K256H S379A/D	Resin affinity
		C-His	WT K256H S379A/D	
	Sf9	N-TS	WT K256H	
210-457	*E. coli*	N-TS	WT K256H S379A/D	Resin affinity + SEC
		C-His	WT K256H S379A/D K256H S379D	
	Sf9	N-TS	WT K256H S379A/D	
210-429	*E. coli*	N-TS	WT K256H S379A/D	
		C-His	WT K256H S379A/D K256H S379D D344A	
	Sf9	N-TS	WT K256H S379A/D	
210-375	*E. coli*	C-His	WT K256H	
200-377	*E. coli*	C-His	WT	

Table 4. 2: Proteins expressed and purified for this chapter

C-His, C-terminal 6His tag; N-TS, N-terminal TwinStrep tag; SEC, size exclusion chromatography; WT, wild-type

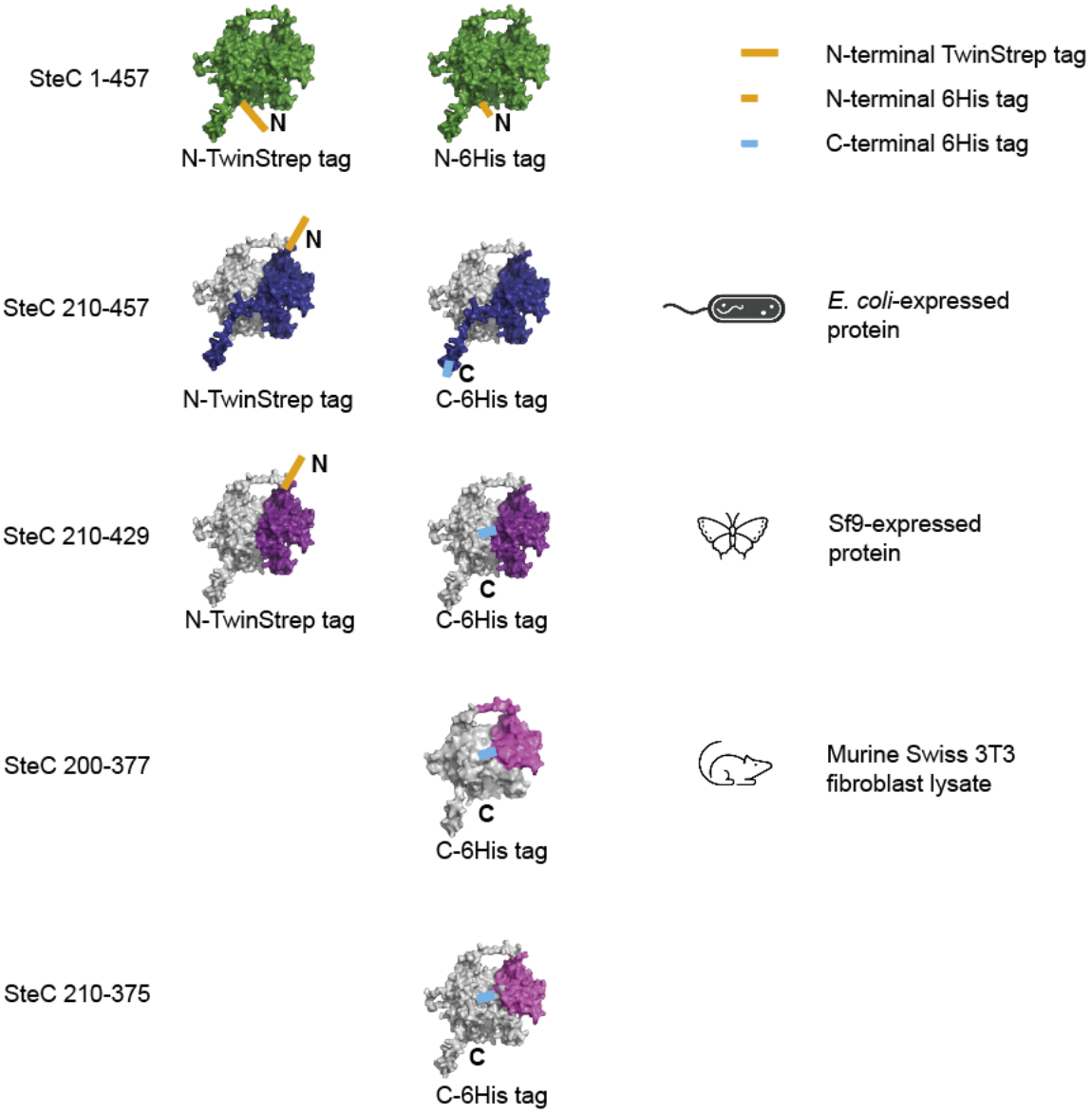


Figure 4. 2: Symbols representing the recombinant protein boundaries, tags, tag positions and expression systems in subsequent figures

To aid interpretation of the figures in Chapters 4 and 5, the following logos have been used to denote the protein construct boundaries (1-457 in green; 210-457 in blue; 210-429 in purple; 210-375 in pink), N-terminal tags (orange), C-terminal tags (light blue), expression systems and use of murine cell lysate. Details are also provided in figure legends.

4.1.3 Role of S379 in SteC activity

To understand the role of S379 phosphorylation I carried out assays to determine the relative kinase activity of SteC WT and mutant variants described in Table 4.2. For convenience, the longest synthesised FMNL1 peptide (24 amino acids K190-K213) was used as the model substrate, given that I have demonstrated binding (Section 3.3.4) and FMNL1 phosphorylation

at 3 sites, S199, T203 and T207 (Section 3.3.3). A radioactive kinase assay was used to determine the activity of the SteC mutants compared to WT (Figure 4.3).

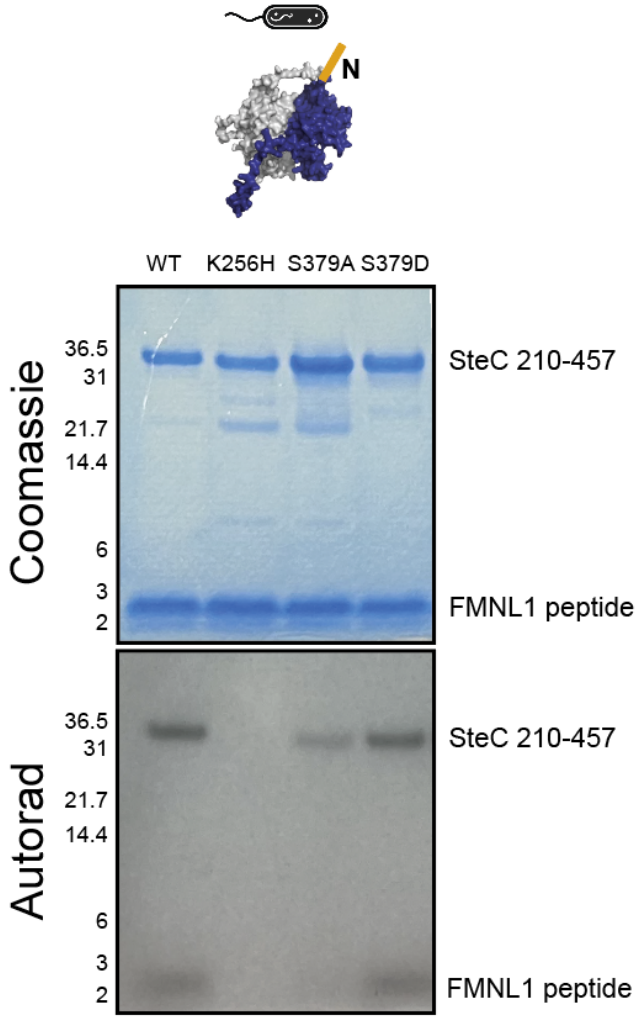


Figure 4. 3: Residue S379 of SteC is required for phosphorylation of FMNL1 peptide SteC

210-457 WT and mutants with N-terminal TwinStrep tag (orange line) expressed in *E. coli* (bacteria symbol) were mixed at a concentration of 5 µM with 100 µM FMNL1 peptide and 100 µM ATP spiked with 20 kBq [γ-P^{32}]ATP before incubation at 30 °C for 30 mins. Mixtures were separated on SDS gels and analysed by Coomassie staining and autoradiography to detect P^{32}. The film was exposed for 4 hours. Result representative of 3 repeats.

SteC activity towards the FMNL1 peptide is demonstrated for the WT. The previously reported kinase-deficient mutant K256H (Poh et al., 2008) was used as a negative control. Additionally, SteC S379A demonstrates negligible substrate phosphorylation relative to WT, suggesting S379 is catalytically important. Interestingly, the S379D mutation rescued activity, suggesting that phosphorylation at this residue is important for activity. In this case, the loss of activity of

S379A compared to WT would suggests WT may autophosphorylate at S379 during expression, given these proteins have been expressed in *E. coli*.

Autophosphorylation is seen for SteC 210-457 WT, S379D and, to a lesser extent S379A and the signal is stronger than for the substrate. In the case of S379D mutant, the autophosphorylation site cannot be at S379, suggesting at least one other site is present in the kinase domain, that was not discovered by the phospho-mass spectrometry experiment. The similar signal of autophosphorylation of WT and S379D suggests that S379 may have already undergone autophosphorylation in the WT, reducing the number of molecules available for phosphorylation during the *in vitro* kinase assay. It is unclear whether the detected autophosphorylation signal is biologically important or simply a bystander effect of catalytic activity in the context of a high concentration of kinase.

4.1.4 Impact of K256 and S379 mutation on protein folding

The differential activity seen in Figure 4.3 could be due to mutations affecting the stability of the protein fold, rather than having a targeted effect on catalytic residues. K256H was originally chosen due to its previous use as an effective kinase-dead mutation. However, given the bulkiness of histidine – less common than arginine or methionine as a replacement residue for catalytic lysine mutation (Reinhardt & Leonard, 2023) – I considered that the K256H mutation may eliminate kinase activity through disrupting N-lobe folding. Similarly, the S379A mutation, located at the start of αF in the predicted structure, may have major structural effects on the kinase domain that could explain its loss of kinase activity, while S379D may rescue this disruptive effect.

To assess this possibility, CD spectroscopy was performed by the same method described in Section 3.2.3, this time on the wild-type protein and mutants. Negligible differences were seen in the secondary structural content of the mutants compared to WT (Figure 4.4), evidence against a change in secondary structure due to K256H or S379A mutations. However, small structural changes – for example in the relative position or α-helical content of the activation loop – would not be detected by this method.

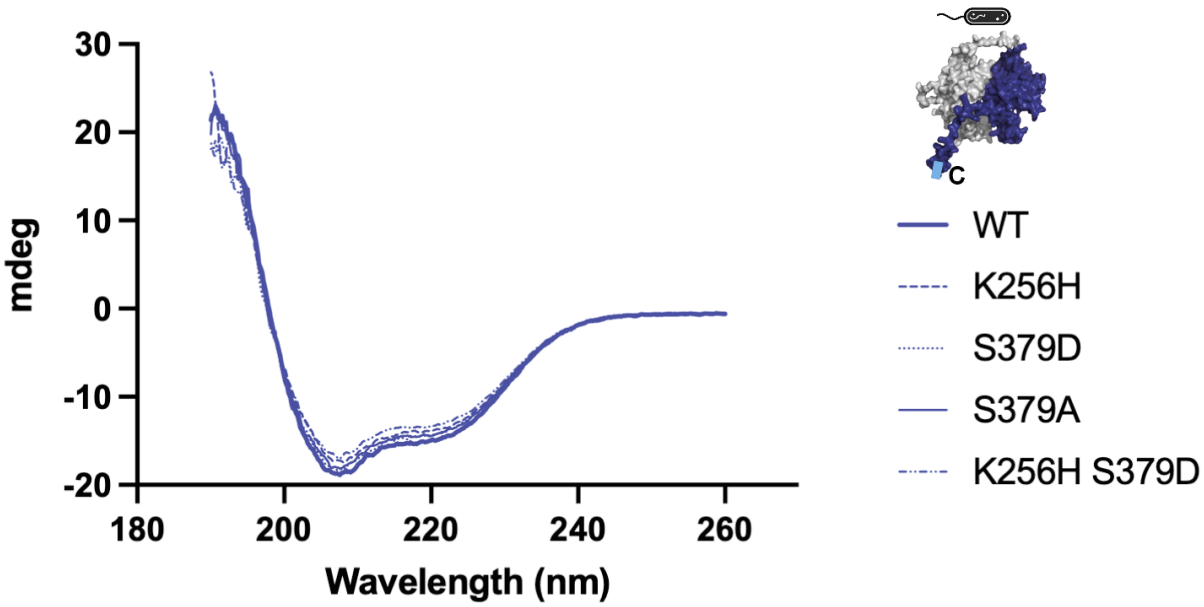


Figure 4. 4: CD spectra of kinase domain WT and mutants are comparable

SteC 210-457 WT and the indicated mutants with C-terminal 6His tag (blue line) expressed in *E. coli* were dialysed into phosphate buffer. CD spectra were collected in the 190-260 nm range at 25 °C. Experiment performed with the support of Laura Masino (Structural biology STP).

4.1.5 Impact of mutations on binding to substrate peptide

Activation loop phosphorylation commonly modulates kinase activity through enhanced binding to the substrate peptide (Adams, 2003). To assess if this was the case for SteC, biolayer interferometry (BLI) was performed again, this time comparing binding of SteC 210-429 WT and mutants to FMNL1 peptide (Figure 4.5). The shorter construct was selected due to the aggregation-prone nature of the full-length protein and the assay being performed at room temperature, therefore making comparisons between constructs unreliable.

Binding curves appeared similar for SteC 210-429 WT and mutants. SteC 210-429 WT had an estimated K_d of 7.9 μM (± 1.9 μM). The K256H, S379A and S379D mutants had K_d estimates of 22 μM (± 3 μM), 21 μM (± 3 μM) and 14 μM (± 2 μM), respectively. To assess if nucleotide may allosterically alter the binding affinity of SteC to the substrate peptide, AMPPNP – a non-hydrolysable ATP analogue – was added to SteC 210-429 WT at a concentration of 1.5 mM and the BLI experiment repeated with the same substrate peptide. This yielded a K_d of 10.3 μM (± 1.9 μM), a slight increase from 7.9 μM for the same protein without AMPPNP i.e. slightly lower affinity binding.

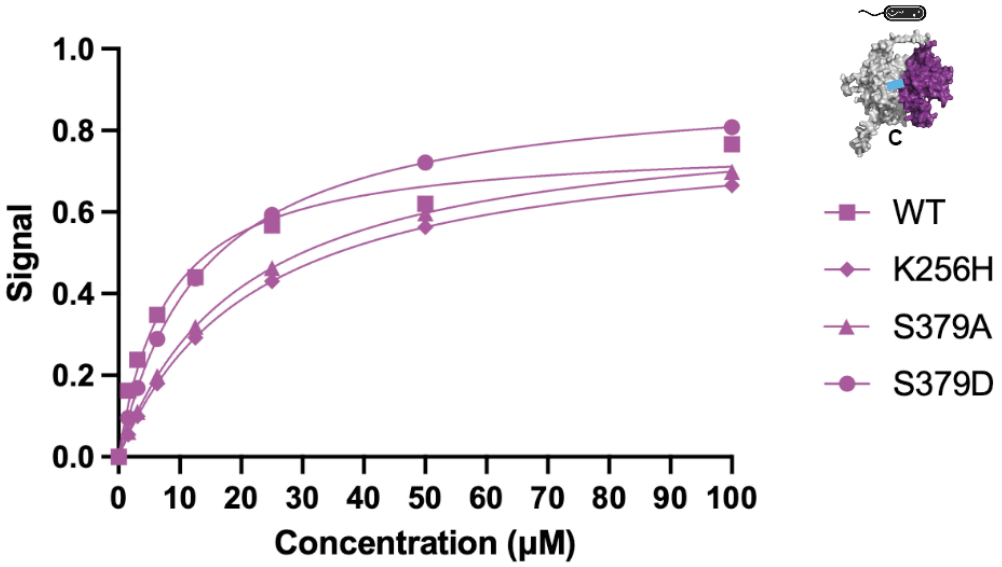


Figure 4. 5: BLI to assess binding of SteC mutants and FMNL1 peptide

SteC 210-429 WT and mutants were expressed in *E. coli* and purified by standard procedures. Peptide pep_K190 was loaded onto Streptavidin sensors at 1 μg/ml and then submerged in a concentration range of SteC proteins. Signal was corrected for a control without peptide. K_d for each protein was calculated using total least squares fit with Prism 10. The experiment was performed with the support of Laura Masino (Structural biology STP).

4.1.6 Impact of mutations on SteC binding to ATP

Lysine has a basic and charged side chain. These properties confer it with the ability to stabilise the α- and β-phosphates of ATP in the active kinase domain. Usually a lysine-to-alanine mutation is used to remove these properties, which would replace the long lysyl side chain with a methyl group thereby making the amino acid non-polar. Hypothetically, a lysine-to-histidine mutation, as previously used for catalytically inactive SteC (Poh et al., 2008), would retain the basic, charged characteristics of lysine but would be replacing the long lysyl group with a bulky imidazole side chain. To further interrogate the mechanism by which SteC K256H loses catalytic activity, the ability of SteC WT and K256H to bind ATP was compared using the fluorescent ATP analogue, Mant-AMPPNP, that increases fluorescence upon protein binding. Furthermore, given the minor differences in peptide binding of the S379 mutants, I hypothesised that S379 phosphorylation may facilitate catalytic activity instead through modulating ATP binding. Therefore, S379A and S379D mutants were also tested. For these experiments the SteC 210-429 fragments were used given that removal of 430-457 makes the protein less aggregation prone at room temperature (Figure 4.6).

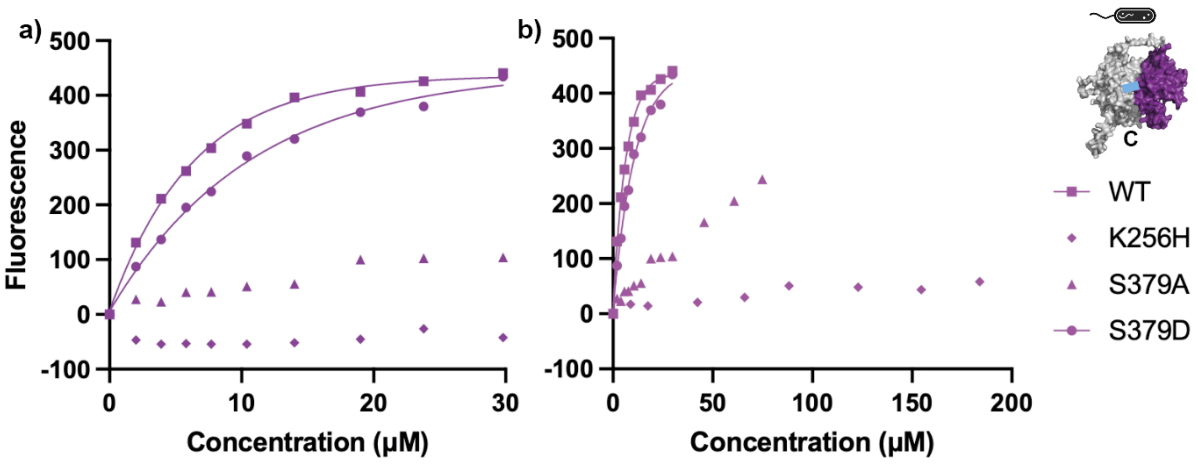


Figure 4. 6: Binding of SteC kinase domain to Mant-AMPPNP

SteC 210-429 WT and mutants expressed in *E. coli* were used to titrate Mant-AMPPNP at 100 μM at 25 °C and the increase in fluorescence at increasing protein concentrations was measured. Data representative of 2 repeats for each construct. Experiment performed with the support of Laura Masino (Structural biology STP)

a) Binding curves over a 0-30 μM range.

b) Titration was continued to explore binding of Mant-AMPPNP to K256H and S379A mutants at higher concentrations.

SteC WT demonstrated concentration-dependent binding to Mant-AMPPNP, with a calculated Kd of 5.4 μM (± 0.4 μM), which is at the stronger end of the AMPPNP affinity range of a kinase (Becher et al., 2013) (Figure 4.6a). No binding to Mant-AMPPNP was detected for SteC K256H, even at high protein concentrations, explaining the catalytically dead nature of this mutant. S379 mutants displayed differential nucleotide binding. S379A mutant displayed some binding above what was seen with the K256H mutant, but this was too weak to fit a binding curve, even at higher protein concentrations (Figure 4.6b). The S379D mutant showed binding to Mant-AMPPNP that was comparable to WT SteC, with a calculated Kd of 11.8 μM (± 1.5 μM) **(**Figure 4.6b). These data imply that S379 phosphorylation is important for nucleotide binding. This is reminiscent of SRC kinase phosphorylation at the activation loop, which leads to increased nucleotide binding affinity, suggesting an allosteric network connecting the activation loop and catalytic cleft (Pucheta-Martínez et al., 2016). The data in this section are incorporated into a model of phospho-activation in the discussion (Chapter 6).

4.2 Studying the origin of SteC S379 phosphorylation

Results from Section 4.1 indicate that autophosphorylation and eukaryotic kinase phosphorylation might both contribute to activation of SteC and were thus explored further.

4.2.1 Intact MS of recombinantly expressed SteC

The kinase activity profile of S379 mutants (Section 4.1.3) provide indirect evidence of autophosphorylation at S379. To explore this more directly, SteC WT, K256H, S379A and S379D C-tagged *E. coli*-expressed protein was first separated by ion exchange chromatography (Figure 4.7). Protein eluted in multiple peaks between 120 mM and 180 mM NaCl, indicating multiple species suggestive of heterogeneously autophosphorylated protein.

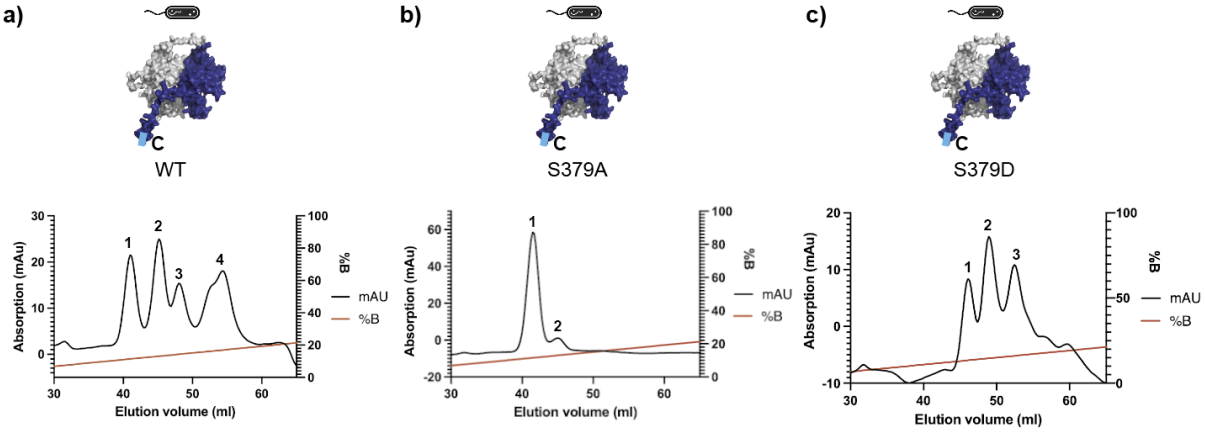


Figure 4. 7: Separating SteC species by AIX

SteC 210-457 with C-tag expressed in *E. coli* was purified by NiNTA, then analysed by anion exchange chromatography (AIX). Samples were loaded onto a Resource Q column and eluted over a salt gradient from 50 mM to 1 M NaCl.

Next, the eluted protein corresponding to these peaks was individually analysed by intact MS (Table 4.3). This demonstrated un-, mono-, di- and tri-phosphorylated species of SteC 210-457 WT. SteC K256H was fully unphosphorylated. SteC S379A was largely unphosphorylated, while SteC S379D was mono- and di-phosphorylated. Without knowing what these phosphosites are, it would follow that S379 is one of them, given S379D only demonstrated di- not tri-phosphorylation.

SteC 210-457	Peak from IEX (Figure 4.7)	Detected MW by Intact Mass (Da)	Expected MW of unmodified protein	Expected MW with N-terminal M cleavage	Predicted phospho-state
WT	1	29,032.60	29,181	29,032	0 phos
WT	2	29,112.60			+1 phos
WT	3	Insufficient sample			-
WT	4	29,192.40 (95%) 29,271.40 (5%)			+2 phos (95%) +3 phos (5%)
K256H	Single peak	29,040.80	29,172	29,041	0 phos
S379A	1	29,016.20	29,147	29,016	0 phos
S379A	2	29,016.20 (60%) 29,096.00 (40%)			0 phos (60%) +1 phos (40%)
S379D	1	Insufficient sample	29,191	29,060	-
S379D	2	29,140.60			+1 phos
S379D	3	29,220.60			+2 phos

Table 4. 3: Intact MS to determine phospho-species separated by AIX

Samples of protein from individual peaks eluted by IEX were each analysed by intact MS to determine covalent modifications. The peak numbers refer to peaks labelled in Figure 4.7. Intact MS was performed by Sarah Maslen (Proteomics STP, Francis Crick Institute). Expected molecular weights were calculated using Expasy. Given all proteins in this experiment had an N-terminal methionine then serine, methionine was expected to be cleaved (Wingfield, 2017), so 149 Da was subtracted from the expected MW.

4.2.2 PhosphoMS of SteC WT and K256H from E. coli and Sf9

To investigate S379 phosphorylation further, phospho-mass spectrometry was performed on different SteC constructs from bacterial and insect cell expression (Figure 4.8).

No S379 phosphorylation is detected in SteC 210-457 K256H expressed in *E.* coli while 40% of this peptide is S379-phosphorylated in the equivalent WT, with no evidence of increase upon incubation with ATP. For SteC 1-457 WT expressed in Sf9, close to 100% of S379 was phosphorylated, even prior to incubation with ATP. This was only reduced by 10% for the equivalent protein with the K256H mutation. These data, combined with the intact MS results (Figure 4.7 and Table 4.3) show that S379 phosphorylation can occur by autophosphorylation, in a K256H-dependent manner, as well as by an insect cell kinase given the kinase dead mutant expressed in insect cells clearly showed S379 phosphorylation. The relative S379

occupancy between Sf9 and *E. coli*-expressed constructs and the presence of unphosphorylated SteC in Figure 4.7a and Table 4.3 suggest that S379 may only be partially autophosphorylated. Furthermore, the signal of S379 phosphorylation seems relatively unchanged upon incubation with ATP in all cases, suggesting that the phosphorylation events primarily occur during the expression and purification process. It seems that a eukaryotic kinase can fully phosphorylate SteC at S379. SteC may also require eukaryotic co-factors for full autophosphorylation, though the high occupancy of S379 in the K256H mutant suggests eukaryotic kinase phosphorylation predominates. The ratios of phosphorylated/unphosphorylated peptide reported here are relative, because the unphosphorylated and phosphorylated peptides may be detected with different efficiencies by the mass spectrometer, meaning the data should be interpreted with caution.

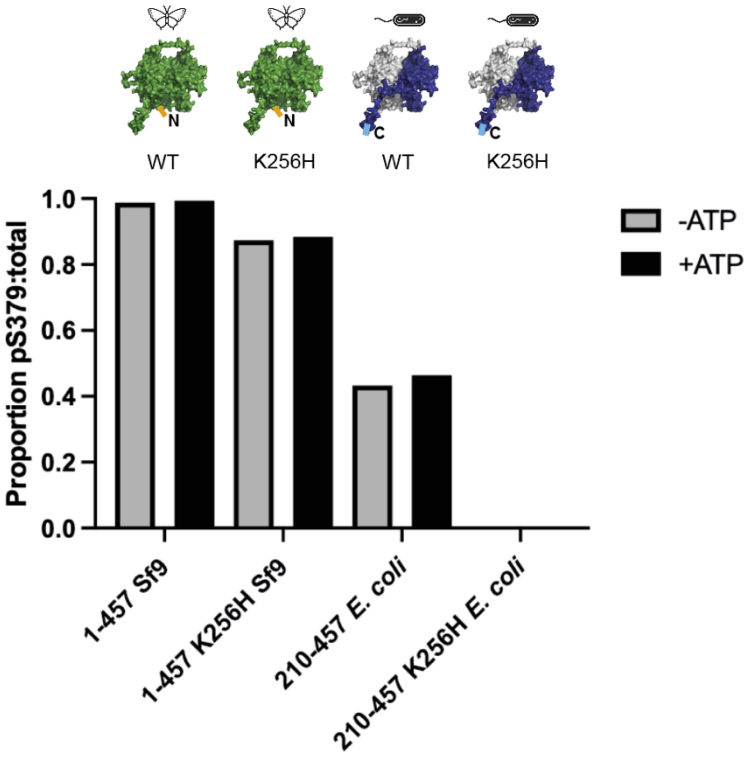


Figure 4. 8: Relative quantification of S379 phosphorylation by phosphoMS

Proteins available at the time of the experiment were full length SteC 1-457 with N-tag from Sf9, and SteC 210-457 with C-tag from *E. coli*, each with corresponding K256H mutations. Proteins were expressed and purified by standard procedures. Samples were either trypsin digested directly (grey bars) or incubated with 100 µM ATP for 30 min at 30 °C prior to trypsin digestion (black bars), and MS analysis, performed by Steve Howell (Proteomics STP). Peptide SVSLATR (S377-R383) was detected in S379-phosphorylated and unphosphorylated forms. For each condition, the proportion of total detected peptide that was phosphorylated was plotted.

4.2.3 Autophosphorylation of SteC fragments

Autophosphorylation in the activation loop of kinases often occurs in trans (Reinhardt & Leonard, 2023), requiring two protomers to associate for phosphorylation to occur. However, a lack of oligomerisation of the SteC kinase domain has been demonstrated in this work (Sections 3.2.4 and 3.2.5) indicating that dimerisation does not induce autophosphorylation. To further explore the mechanism of autophosphorylation further, SteC kinase domains were expressed with tags of different sizes to distinguish WT and kinase-dead (K256H) species by SDS-PAGE, and a kinase assay performed (Figure 4.9).

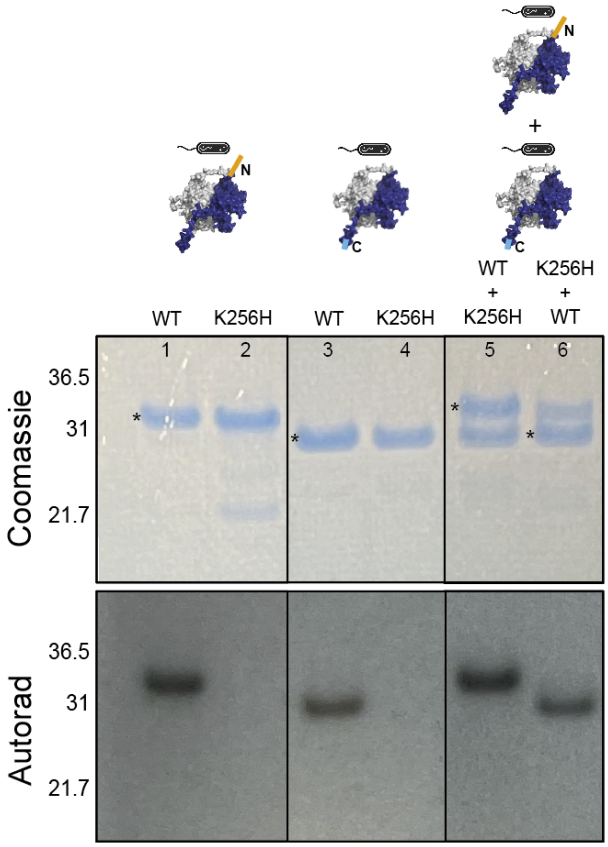


Figure 4. 9: Radioactive kinase assay to distinguish trans- and cis-autophosphorylation

SteC 210-457 WT and K256H with N- and C-tags (orange and light blue, respectively) were expressed in *E. coli*. Protein at 5 µM concentration, either alone or in active-inactive pairs, with 100 µM ATP spiked with 20 kBq [γ-P^{32}]ATP, was incubated for 30 min at 30 °C and analysed on SDS-PAGE followed by Coomassie staining and autoradioagraphy. Film exposed for 4 hours. ***** on Coomassie gel denotes the WT form, while the unstarred bands are K256H mutant.

In the first four wells autophosphorylation can be seen in a K256-dependent manner, matching previous results. This signal may be due to S379 or another phosphosite. The signal for autophosphorylation of the C-tagged protein is slightly weaker than the N-tag. To explore trans- versus cis-autophosphorylation, proteins in wells 1 & 4, and 2 & 3 were combined and loaded into wells 5 and 6, respectively, followed by an *in vitro* kinase assay. Given we know

that SteC 210-457 K256H is unphosphorylated at S379 after expression and purification *in E. coli* (Table 4.3 and Figure 4.8), a trans-autophosphorylation mechanism should lead to SteC 210-457 K256H phosphorylation in the presence of WT SteC and ATP. This was not seen. Phosphorylation only occurred on the WT constructs. One cannot rule out the possibility that the opposing tags prevent dimerisation. However, the lack of trans-autophosphorylation was also observed with the tags in opposing directions (wells 5 &6), suggesting that the tags do not influence this process. Given that the SEC-MALLS and SAXS experiments showed no evidence of homodimer formation (Chapter 3), this supports a cis-autophosphorylation mechanism.

4.2.4 Using murine cell lysate to phosphorylate bacterially expressed SteC

The experiments described thus far demonstrate cis-autophosphorylation as a likely mechanism of SteC S379 phosphorylation. However, phosphoMS in Section 4.2.2 would suggest that a eukaryotic kinase (from insect cells) can also phosphorylate SteC at S379. To explore this further, a kinase assay was performed with SteC constructs (WT and mutants) expressed in *E. coli* without and with cell lysate from Swiss 3T3 mouse fibroblasts (Figure 4.10), a cell line in which SteC-dependent actin polymerisation is observed upon SteC transfection and *Salmonella* infection (Poh et al., 2008).

In this assay, with a shorter exposure time of 1 hour, only weak phosphorylation of SteC S379D was detected in the samples without murine cell lysate, with little phosphorylation detected for the other SteC variants. When murine cell lysate was included, a strong signal was seen for SteC K256H phosphorylation, that was not present or barely detectable for the other mutants, most notably K256H S379D double mutant. This suggests that residue S379 of SteC was phosphorylated by a murine kinase.

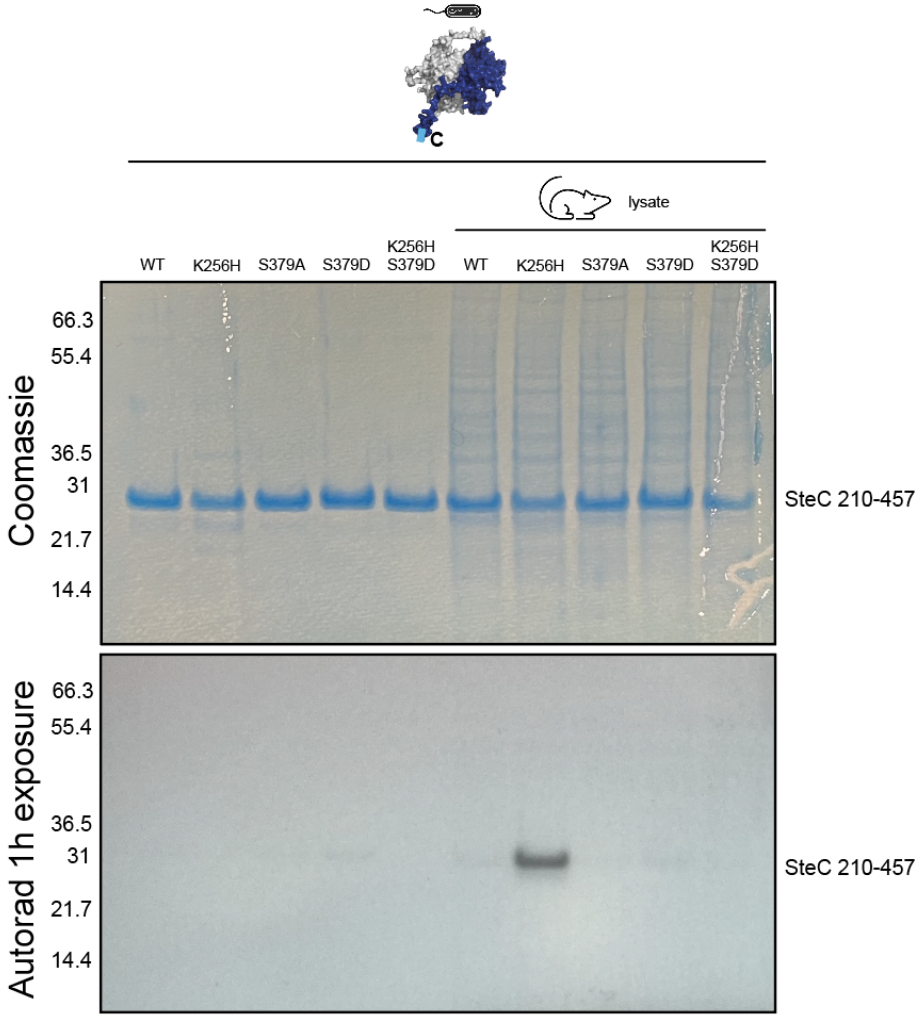


Figure 4. 10: SteC 210-457 K256H phosphorylated in the presence of murine cell lysate

WT and mutant SteC 210-457, expressed in *E. coli* were added at 5 µM to 100 µM ATP spiked with 20 kBq [γ-P^{32}]ATP and 1x phosphatase inhibitor alone (left) or with the 1 µl Swiss 3T3 mouse fibroblast cell lysate (right). Mixtures were incubated for 30 min at 30 °C and analysed on SDS-PAGE followed by Coomassie staining and autoradiography to detect P^{32}. Autoradiography film was exposed for 1 hour. Data representative of 3 repeats.

4.2.5 STK38 as a candidate host kinase phosphorylating SteC S379

STK38 (also known as NDR1) is a mammalian kinase identified in a recent mass spectrometry analysis of SteC binding partners from infection (Walch et al., 2021). It is classified as a basophilic kinase (J. Johnson et al., 2023), with a predilection to basic residues in the -1 to -5 position. Residue S379 of SteC is preceded by QQRSV. Therefore, STK38 was purchased and a kinase assay was used to test whether this enzyme phosphorylates SteC (Figure 4.11). STK38 autophosphorylation activity was observed. STK38 also appeared to phosphorylate SteC K256H, while not phosphorylating SteC K256H S379D, providing evidence that STK38 phosphorylates SteC at S379, at least under *in vitro* conditions.

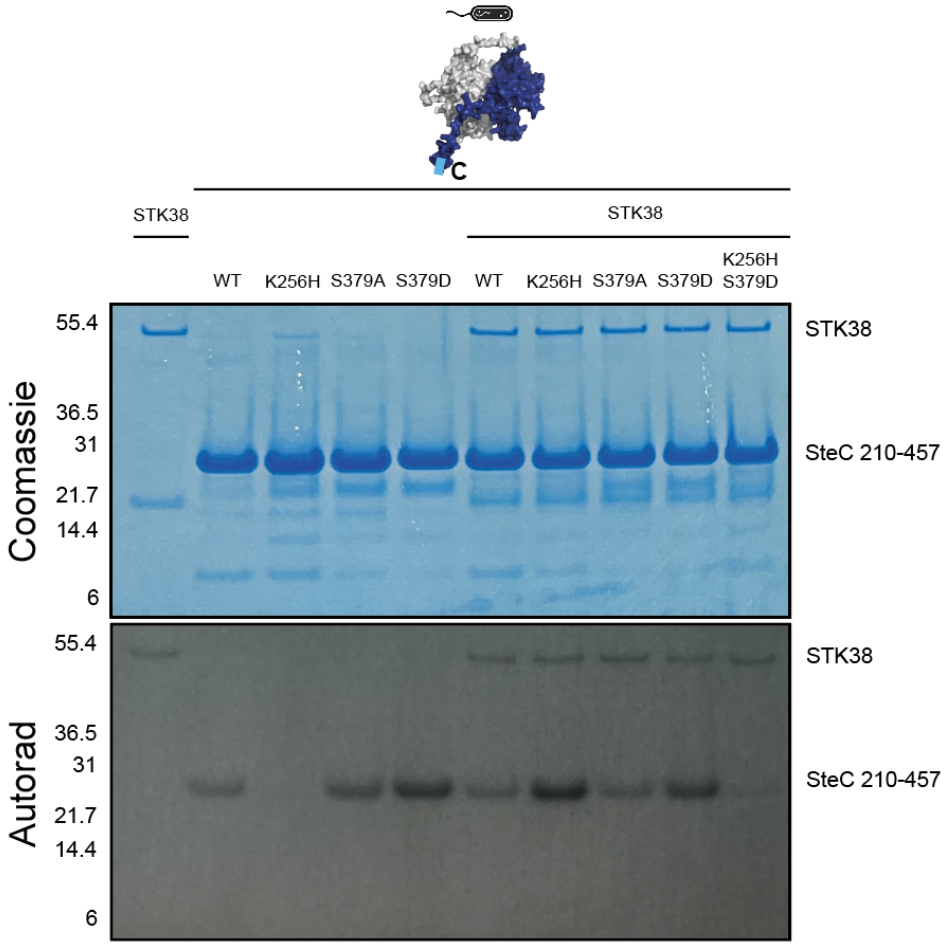


Figure 4. 11: Radioactive kinase assay with SteC and mammalian kinase STK38

SteC 210-457 WT and mutants with C-tag were expressed and purified from *E. coli* and included in the assay at 5 µM with or without addition of STK38 at 2 µM. 100 µM ATP spiked with 20 kBq [γ-P^{32}]ATP was added and the mixtures were incubated for 30 min at 30 °C and analysed on SDS-PAGE followed by Coomassie staining and autoradiography to detect of P^{32}. Film was exposed for 16 hrs. Data representative of 2 repeats.

4.2.6 Relative activity of bacterial- and insect cell-expressed SteC

If SteC is phosphorylated at S379 by a eukaryotic kinase to become active, one may expect the insect cell-expressed kinase to be more active than the equivalent *E. coli*-expressed kinase. If autophosphorylation is the predominant mode of activation, there should be no difference. To test this hypothesis, a kinase assay was performed to compare SteC expressed in different systems (Figure 4.12).

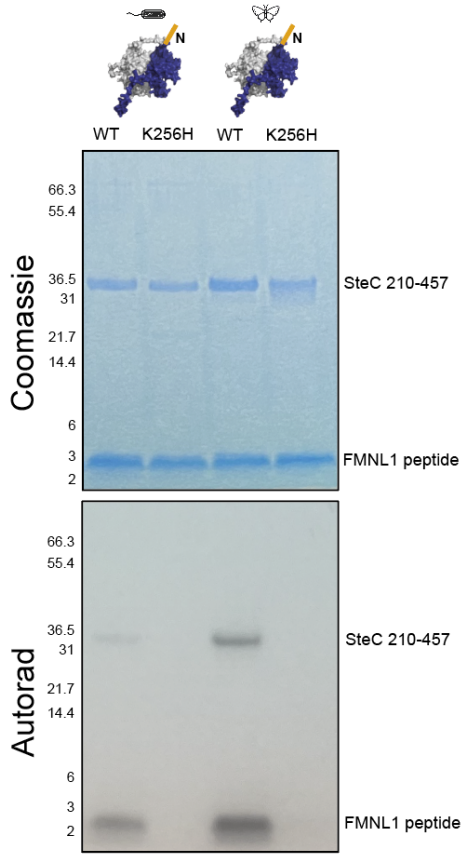


Figure 4. 12: Sf9-expressed SteC 210-457 is more active than the same construct expressed in *E. coli*

SteC 210-457 was expressed with N-tag in *E. coli* and Sf9 expression systems. 5 µM kinase was added to 100 µM FMNL1 peptide and 100 µM ATP spiked with 20 kBq [γ-P^{32}]ATP, then incubated for 30 min at 30 °C and analysed on SDS-PAGE followed by Coomassie staining and autoradiography to detect P^{32}. Film exposed for 2 hours. Representative of 3 repeats.

SteC 210-457 expressed in *E. coli* and Sf9 demonstrate autophosphorylation and substrate peptide phosphorylation, each in a K256-dependent manner. SteC 210-457 expressed in Sf9 demonstrates a stronger autophosphorylation and substrate phosphorylation signal than that expressed in *E. coli.* This suggests that a condition of the Sf9 expression process improves SteC 210-457 catalytic activity.

To summarise, in Section 4.1, data presented showed that SteC is phosphorylated at S379 and its kinase activity was dependent on phosphorylation at S379. In this section, the data presented demonstrates that SteC S379 phosphorylation is more complete in insect cell-expressed protein than *E. coli*-expressed protein. Auto-phosphorylation occurs in cis, not trans, while phosphorylation at S379 can be performed by a kinase in murine cell lysate, and specifically by the eukaryotic kinase STK38. Furthermore, SteC expressed in insect cells was more active in substrate phosphorylation than the same construct expressed in *E. coli*.

4.3 Role of C-lobe elements in kinase activity

In Chapter 3, I explored the structural and biophysical characteristics of different SteC kinase domain fragments. Yet how these segments of the kinase domain relate to catalytic activity has remained unexplored. In this next section I will explore the role of the C-terminal elements of the kinase domain in autophosphorylation and substrate phosphorylation by comparing the kinase activity of various kinase domain fragments.

4.3.1 Comparing the catalytic activity of different kinase domain fragments

To explore their relative catalytic activity, kinase domain fragments were compared in a kinase assay with FMNL1 peptide as the substrate (Figure 4.13).

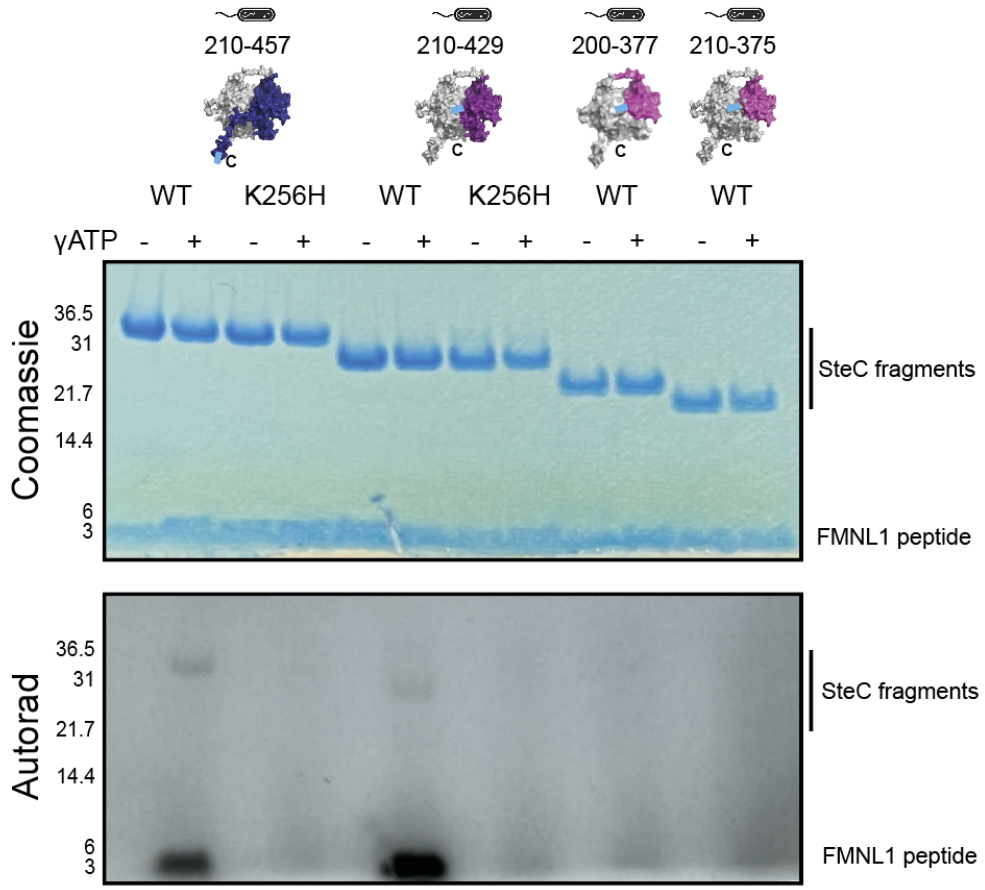


Figure 4. 13: SteC kinase domain fragments 210-457 and 210-429 are active, whereas SteC 200-377 and 210-375 are not

SteC 210-457, 210-429, 200-377 and 210-375 with C-tags expressed in *E. coli* were purified by standard procedures and were mixed at 5 µM with FMNL1 peptide at 100 µM and 100 µM ATP spiked with 20 kBq [γ-P^{32}]ATP, then incubated for 30 min at 30 °C and analysed on SDS-PAGE followed by Coomassie staining and autoradiography to detect P^{32}. Film exposed for 4 hours. Representative of 2 repeats.

Autophosphorylation was seen for the SteC construct consisting of residues 210-457 and slightly less strongly for SteC 210-429. Substrate phosphorylation was also observed, possibly stronger for SteC 210-429. This was in an ATP- and K256-dependent manner, as expected. This suggests that the C-tail is important for auto-phosphorylation. Neither 200-377 nor 210-375 fragments were active in auto- or substrate phosphorylation. This was expected because they lack a) any C-lobe when compared to canonical kinases, meaning substrate binding is unlikely; and b) the putative activation loop containing S379 which we've shown is an essential modulator of kinase activity, as shown in Figures 3.6 and 3.7. This provides evidence that the C-lobe is essential for catalysis.

4.3.2 Comparing SteC activity with and without the C-tail, and with N- or C-tags

To investigate further the role of the C-tail in kinase activity, I next changed the position of the tag, from C terminal to N-terminal and then carried out a kinase assay to interrogate if the tag may influence activity (Figure 4.14).

SteC 210-457 and 210-429 with N- and C-tags all demonstrated auto-phosphorylation and FMNL1 peptide substrate phosphorylation. SteC 210-457 and 210-429 with N-terminal tags have stronger autophosphorylation and substrate phosphorylation signals than their equivalents with C-terminal tags. The effect of tag position is more pronounced for 210-429. This suggests that the 6His-tag after amino acid 457 and 429 may interfere with autophosphorylation capacity. The activity of the SteC kinase domain fragments is K256-dependent as expected.

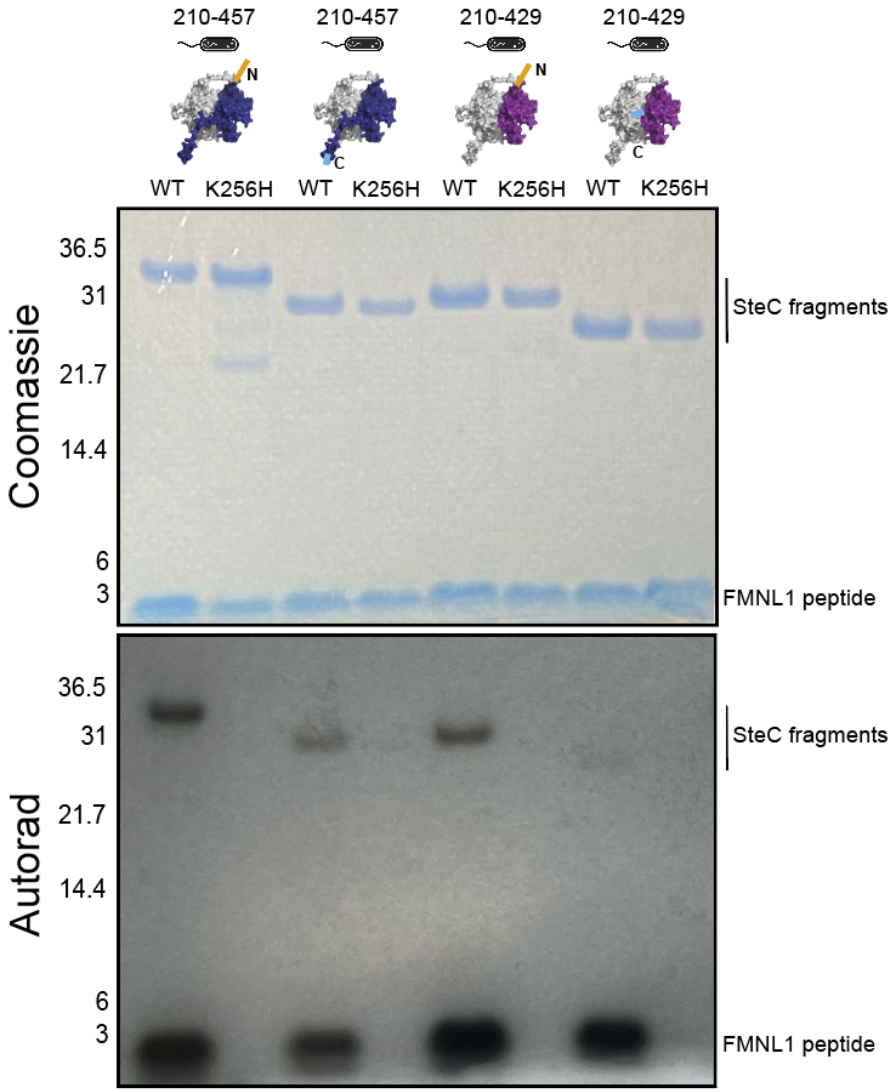


Figure 4. 14: Tag position affects SteC 210-457 and 210-429 autophosphorylation and substrate peptide phosphorylation

SteC 210-457 and 210-429 with either N- or C-tags expressed in *E. coli* and purified with standard procedures, were combined at 5 µM with FMNL1 peptide at 100 µM and 100 µM ATP spiked with 20 kBq [γ-P^{32}]ATP, then incubated for 30 min at 30 °C and analysed on SDS-PAGE followed by Coomassie staining and autoradiography to detect detection of P^{32}. Film exposed for 16 hours.

4.3.3 Comparing kinase activity of SteC constructs and tags with FMNL1 1-385

The kinase recognition motif, made up of the amino acid target residues (classically serine, threonine, or tyrosine) and the neighbouring amino acids in the sequence, is the most studied determinant of kinase substrate specificity (J. Johnson et al., 2023) and is determined by the depth and characteristics of the kinase catalytic cleft and nearby surface residues, usually in the C-lobe. However, other factors also contribute to the substrate specificity of a kinase including docking sites distant from the catalytic cleft which may enforce or prevent proximity to a favourable substrate motif; ternary protein complexes; regulation of expression; cell

localisation; and cell type (Ubersax & Ferrell, 2007). Having assessed 'intrinsic' kinase activity using the FMNL1 peptide, I aimed to determine if the C-tail and tag position may translate into differential kinase activity against a folded protein substrate. I expressed and purified FMNL1 1-385 with an N-terminal GST tag which I cleaved and further purified with SEC. I used this protein as a substrate in a radioactive kinase assay (Figure 4.15).

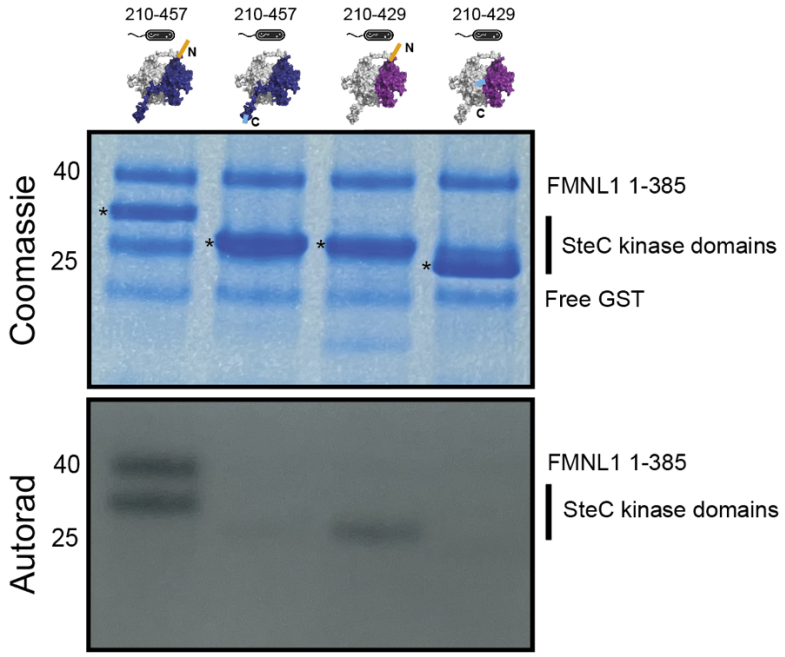


Figure 4. 15: Only N-tagged SteC 210-457 phosphorylates FMNL1 1-185

SteC 210-457 and 210-429 with N- or C-tag were expressed in *E. coli* and purified by standard procedures and GST-His-3C-FMNL1 1-385 was purified by NiNTA, 3C cleavage, reverse NiNTA and SEC. SteC constructs and FMNL1 1-385 were combined at 5 µM each with 100 µM ATP spiked with 20 kBq [γ-P^{32}]ATP, then incubated for 30 min at 30 °C and analysed on SDS-PAGE followed by Coomassie staining and autoradiography to detect P^{32}. Film exposed for 4 hours. * indicates the SteC construct in each well. Representative of 2 repeats.

Coomassie staining shows FMNL1 1-385 (top), SteC constructs (asterix) and a free GST impurity from the FMNL1 purification process (bottom). The autophosphorylation signal can be seen for N-tagged SteC 210-457 and 210-429, but barely for the C-tagged equivalents. This correlates with the previous kinase assay (Figure 4.14). Then, surprisingly, only N-tagged SteC 210-457 phosphorylated the protein substrate, FMNL1 1-385. This suggests that, unlike the substrate peptide, the C-tail of the kinase domain is essential for substrate phosphorylation, and this is disrupted with either removal of the C-tail (430-457) or addition of a 6His tag to the C-terminus.

4.3.4 S379 phosphorylation in kinase domain fragments 210-457 and 210-429

Given the importance of the C-tail and tag position for autophosphorylation (Figure 4.14) and substrate phosphorylation (Figure 4.15) in kinase assays, I hypothesised that S379 of SteC 210-429 may be less well phosphorylated than in SteC 210-457. First, I performed phosphoMS to estimate relative ratios of phosphorylated vs unphosphorylated peptide (Figure 5.16). This revealed that, for the 210-429 fragment (C-tag) 28% of SVSLATR peptides were phosphorylated at S379 (i.e. SVpSLATR), whereas 210-457 (C-tag) demonstrated 43% phosphorylation.

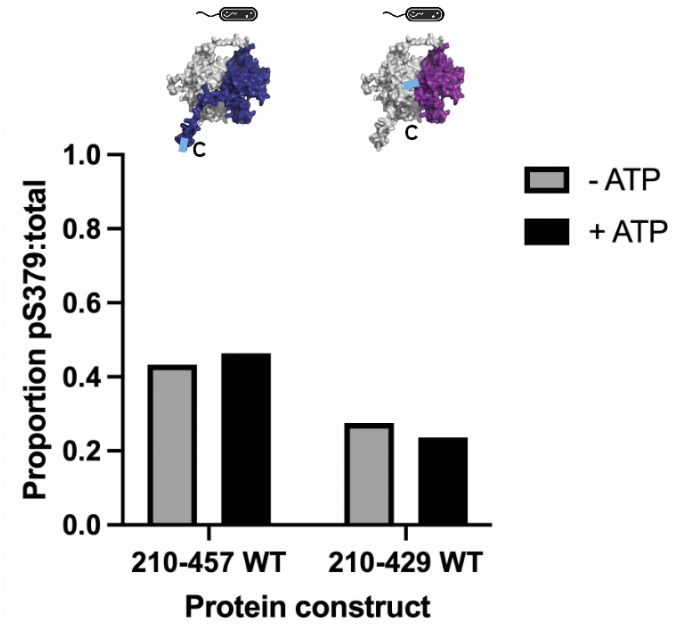


Figure 4. 16: Relative quantification of S379 phosphorylation in SteC 210-457 and 210-429 by phosphoMS

SteC 210-457 and 210-457 WT with C-tag from *E. coli* were purified by standard procedures analysed by phosphoMS in the same experiment as Figure 4.8, by Steve Howell (Proteomics STP). Samples were either trypsin digested directly (grey bars) or incubated with 100 µM ATP for 30 min at 30 °C prior to trypsin digestion (black bars), and MS analysis. Peptide SVSLATR (S377-R383) was detected in S379-phosphorylated and unphosphorylated forms. For each condition, the proportion of total detected peptide that was phosphorylated was plotted.

The proportion of phosphorylated and unphosphorylated SVSLATR peptide is a relative estimate because the two peptides may fly differently in the mass spectrometer. Therefore, an alternative approach was taken to complement these findings. Wild-type and mutant SteC 210-429 kinase fragments were exposed to murine cell lysate in a kinase assay (Figure 4.17). Given the residual possibility that the K256H mutation may unfold the kinase domain, making S379 more accessible for murine kinase phosphorylation, I aimed to abolish kinase activity and hence autophosphorylation capacity by another means that may be less likely to unfold

the protein. Therefore, D344 was selected for mutation, which we hypothesised to be a key residue in the catalytic loop HD motif (equivalent of HRD motif in canonical eukaryotic S/T kinases, Figure 3.6). Briac Lemetais, an MRes student, cloned the SteC 210-429 D344A mutant with a C-tag, expressed the construct in *E. coli* and purified the protein using standard procedures, under my supervision. I performed the radioactive kinase assay.

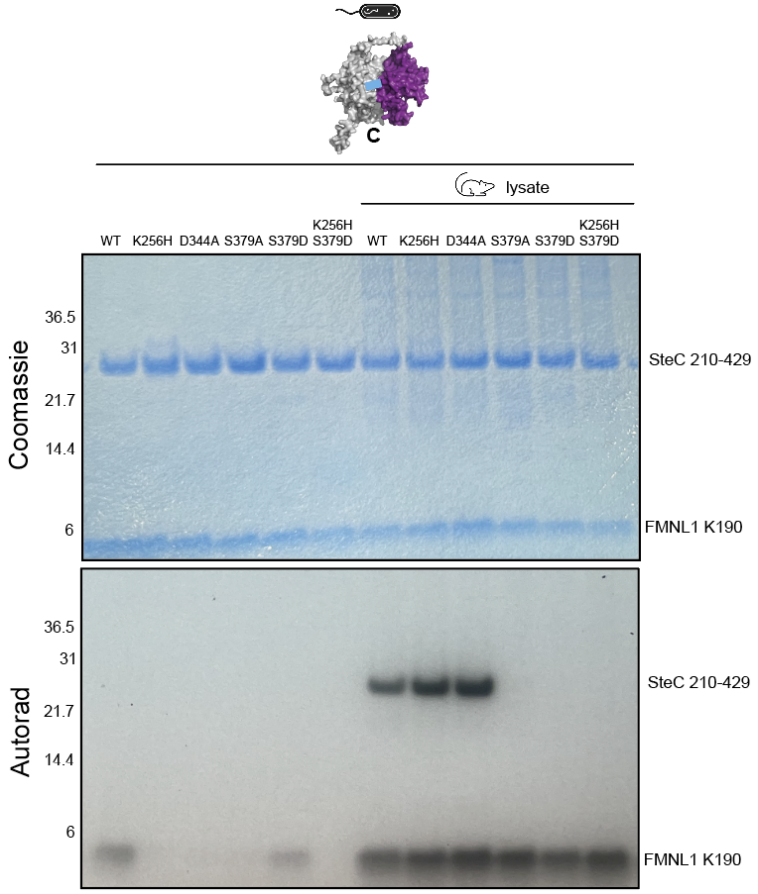


Figure 4. 17: Radioactive kinase assay comparing SteC 210-429 WT and mutants with FMNL1 peptide alone and in the presence of murine cell lysate

SteC 210-429 WT and mutants with C-tag expressed in *E. coli* and purified by standard procedures were combined at 5 µM with FMNL1 peptide at 100 µM, 100 µM ATP spiked with 20 kBq [γ-P^{32}]ATP and 1x phosphatase inhibitor alone (left) or with 1 µL murine cell lysate (right), then incubated for 30 min at 30 °C and analysed on SDS-PAGE followed by Coomassie staining and autoradioagraphy to detect P^{32}. Autoradiography exposure time 1 hour. Representative of 2 repeats.

No autophosphorylation can be seen here due to the short exposure time and the relative lack of autophosphorylation capacity of 210-429. On the left, the kinase activity of various SteC 210-429 mutants towards the FMNL1 peptide was compared. This demonstrated, as expected that SteC K256H was inactive towards the FMNL1 peptide when compared to WT SteC, which caused its phosphorylation. Again, as before, SteC 210-457 S379A was inactive. SteC D344A was also inactive towards the peptide substrate, validating our hypothesis that this residue is

important for the catalytic activity of SteC. As previously, the S379D mutant was active. Finally, the SteC 210-457 K256H S379D double mutant was inactive. Upon addition of murine cell lysate, WT SteC, SteC K256H and SteC D344A all become phosphorylated. However, SteC S379A or SteC S379D did not undergo phosphorylation. This suggests that phosphorylation is occurring at S379. Comparing this profile to 210-457 in a similar assay (Figure 4.10), it appears that S379 is more readily phosphorylated in 210-429 WT than 210-457 WT. The FMNL1 peptide becomes phosphorylated in all cases, independently of SteC, suggesting a murine kinase is responsible (Figure 4.17).

4.3.5 HDX-MS to explore intramolecular interactions of the C-tail

Results above suggest the C-tail of SteC kinase domain is important for auto-phosphorylation and substrate phosphorylation. Results from Chapter 3 suggest that the C-tail does not have significant secondary structure, and may undergo dynamic interaction with the rest of the molecule. Meanwhile, results in this section show that it has a function in auto- and substrate phosphorylation. To explore this further, HDX-MS was performed comparing SteC 210-457 and 210-429 in separate samples. Rather than assessing inter-molecular binding, this experiment aimed to see if any region of the 210-429 molecule was more protected from solution in the presence of the C-tail, which might suggest C-tail folding or transient intra-molecular interaction with the 210-429 core.

Peptide coverage of the experiment was 94.1% (Figure 4.18a). The region 420-429, that might be an important interaction site given its position was not covered by the experiment. This is due to an inherent limitation of the technique when applied to two molecules that share a common region: the same peptides must be detected in both conditions for comparisons to be made, so the boundary region between the sequence covered by both and the sequence covered by one is often poorly assigned. A further 36/220 residues, 210-216, 253-260, 353-358, 360-365, 381 and 412-419 were covered by a single peptide region.

With a difference in deuterium update threshold of 0.45 being defined as significant, one peptide region reached significance: 382-411 (Figure 4.18b). This is a borderline significant result, and so should be treated with caution. The signal above the baseline below the significance threshold for similar peptides 388-408 and 389-411 add some strength to the finding. While this borderline result could be due to chance, it could also represent a relatively weak signal due to the transient or dynamic nature of the C-tail interaction with the alpha-helices in the minimal C-lobe.

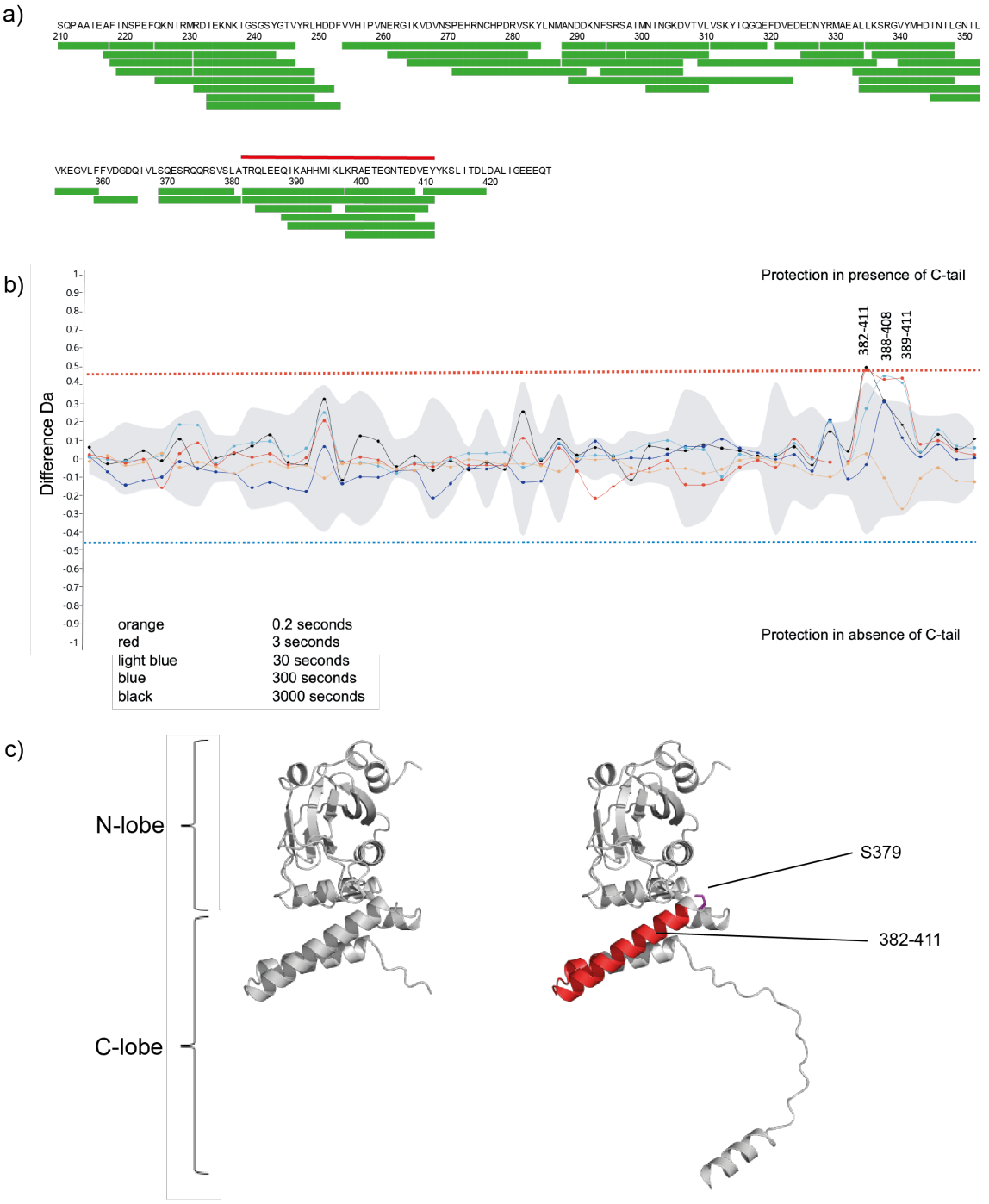


Figure 4. 18: HDX-MS comparing SteC 210-457 and 210-429

SteC 210-457 and 210-429 with C-tag expressed in *E. coli* were purified by standard procedures. Deuterium was added to each of the two protein samples prior to quenching and trypsin digestion. HDX-MS analysis was performed by Sarah Maslen (Proteomics STP).

a) Peptide coverage of the experiment using peptides in the region of 210-429 for both conditions, given the focus of the experiment. Each green box is a residue. Total 50 peptides, 94.1% sequence coverage, 3.78 redundancy score. Regions significantly protected in the presence of the C-tail are highlighted in red above the sequence.

b) Hydrogen deuterium exchange is plotted as a relative difference in deuterium uptake of SteC 210-429 (without C-tail) or SteC 210-457 (with C-tail) across the amino acid range 210-429. Red and blue dotted line highlight the significance threshold of 0.45.
c) The published AF predicted structures of SteC are demonstrated truncated to residues 210-429 (left) and 210-457 (right). Regions of SteC 210-429 that undergo protection in the presence of the C-tail (> 0.45 difference in deuterium uptake) are highlighted in red (right). S379 is highlighted in stick form (purple).

The significant peptide region has been highlighted in red on the AF predicted structure along with S379 in purple (Figure 4.18c**,** right). It is interesting to consider how the C-tail may function to facilitate cis-autophosphorylation of S379, substrate binding and substrate phosphorylation.

4.4 Assessing kinase activity against a novel substrate, MYL12A

Given the lack of mechanistic elaboration of SteC's substrates in the literature to date, and the relatively weak phosphorylation signal for FMNL1 in our hands, I attempted to identify other biologically relevant substrates.

4.4.1 Using murine cell lysate to detect substrates of SteC

Murine cell lysate was used in a kinase assay, this time treating the lysate as a mixture of potential host substrates rather than SteC-activating kinases. The only difference in approach to the previous experiments was the addition of three times more lysate with the aim of identifying faint bands that were differentially phosphorylated between conditions containing SteC WT and K256H mutant proteins (Figure 4.19).

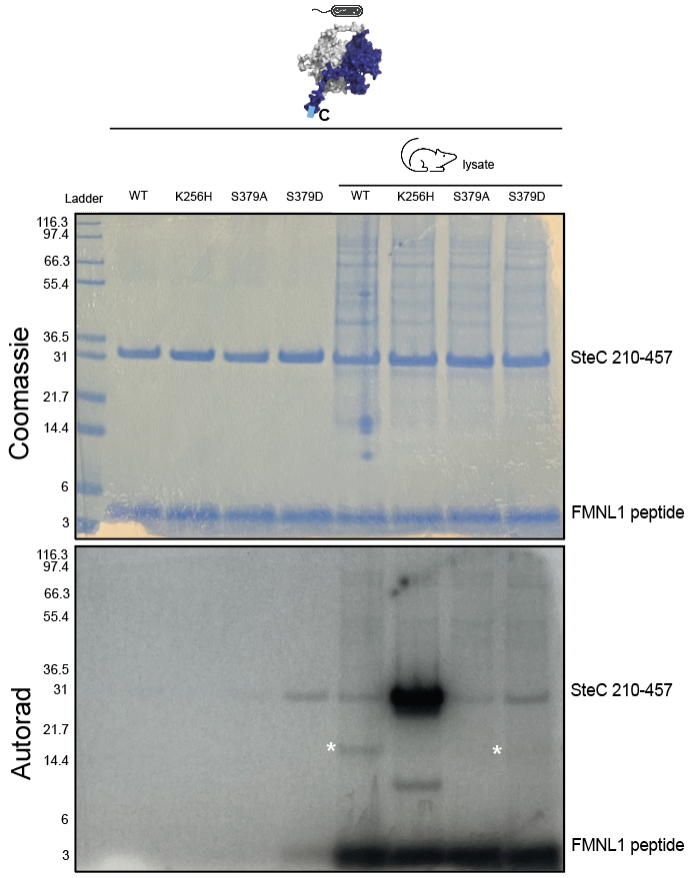


Figure 4. 19: SteC phosphorylates a protein in murine cell lysate

SteC 210-457 WT and mutants with C-tag expressed in *E. coli* were purified by standard procedures. Proteins at 5 µM were added to 100 µM FMNL1 peptide, 100 µM ATP spiked with 20 kBq [γ-P^{32}]ATP and 1x phosphatase inhibitor alone or spiked with 3 µL murine cell lysate, before incubating for 30 min at 30 °C and analysed on SDS-PAGE followed by Coomassie staining and autoradiography to detect P^{32}. Film was exposed for 4 hrs. White * denotes a radioactive signal of interest. Experiment performed once.

In the left four wells, SteC 210-457 WT, K256H, S379A and S379D demonstrate different auto- and FMNL1 peptide phosphorylation, as previously seen. On addition of murine cell lysate, phosphorylation of SteC K256H remains the strongest signal, as previously demonstrated (Figure 4.10). When SteC WT was added to the murine cell lysate, a single band, labelled with a white asterix, and running at approximately 15-20 kDa was detected that was evidently not SteC or the FMNL1 peptide, which was included as an internal phosphorylation control. A band of a similar molecular weight was detected when SteC S379D was included, but not when SteC was inactive due to mutation K256H or S379A. A band at approximately 10-12 kDa was detected only in the condition in which SteC K256H was incubated with the cell lysate. This may be a breakdown product of SteC 210-457 K256H that contains S379. The FMNL1 peptide is phosphorylated in the expected pattern in the SteC-only conditions (left). It is phosphorylated in all lysate conditions (right), suggesting a murine kinase is responsible.

The presence of a phosphorylated band, which was specific to the addition of active but not inactive SteC, and only present upon the inclusion of the 3T3 cell lysate, suggested this may be a substrate of SteC. Therefore, the same experiment was repeated with SteC 210-457 WT and murine cell lysate (one condition) without radioactive material and a band of the Coomassie-stained gel was cut at a level corresponding to the radioactive signal highlighted in Figure 4.19 for mass spectrometry to identify a potential protein substrate (Table 4.4).

FMNL1 was detected with a high abundance score but low exponentially modified protein abundance index (emPAI), meaning coverage of the protein sequence was low. This is due to the presence of the synthesised peptide at a high concentration in the assay. The top two hits of appropriate molecular weight were the MYL12A/MYL12B/MYL9 family and Calmodulin. The proteins could be detected as such but phosphorylated peptides could not be detected, possible due to the relatively low abundance in the sample.

Accession	Description	Score	Mass	emPAI
A0A1W2P6X3	Formin-like protein 1	5498	123418	1.13
Q6ZWY9	Histone H2B type 1-C/E/G	3042	13898	85.85
Q64525	Histone H2B type 2-B	2979	13912	85.85
Q64524	Histone H2B type 2-E	2330	13985	24.75
Q3THE2	Myosin regulatory light chain 12B	2430	19824	34.28
Q6ZWQ9	Myosin, light chain 12A	1913	19940	63.68
Q9CQ19	Myosin regulatory light polypeptide 9	606	19898	17.81
P0DP26	Calmodulin-1	1723	16827	135.94
P05214	Tubulin alpha-3 chain	1719	50612	5.79
P68369	Tubulin alpha-1A chain	1418	50788	4.25
P99024	Tubulin beta-5 chain	1473	50095	8.71
P68372	Tubulin beta-4B chain	1409	50255	8.62
Q7TMM9	Tubulin beta-2A chain	1363	50274	5.32
Q9D6F9	Tubulin beta-4A chain	1127	50010	4.86
Q922F4	Tubulin beta-6 chain	522	50514	1.72
P17742	Peptidyl-prolyl cis-trans isomerase A	1277	18131	61.51
Q60605	Myosin light polypeptide 6	1119	17090	28.96
A0A1W2P6G5	Myosin light polypeptide 6	1045	15609	22.96
Q8CI43	Myosin light chain 6B	120	22849	0.44
C0HKE1	Histone H2A type 1-B	1086	14127	12.98

Table 4. 4: Mass spectrometry identification of band from murine cell lysate showing SteC phosphorylation signal

The sample was cut from Coomassie-stained polyacrylamide gel at the approximate position of the radioactive signal marked with white star in Figure 4.19. Samples were analysed by Lois Kent (Proteomics STP). Accession, UniProt ID. Score, a product of the peptide ion score and used to rank protein hits. Mass, the estimated MW from UniProt. emPAI, exponentially modified protein abundance index, is used to estimate the absolute protein amount in a sample. Hits between 15 and 20 kDa (corresponding to the estimated size of the band) are highlighted in red.

4.4.2 Expression of MYL12A

Myosin light chain 12a/12b and 9 are non-muscle components of myosin. Their phosphorylation by myosin light chain kinase (MLCK) and ROCK at T19 and S20 induces actin rearrangement (Amano et al., 1996; Kamm & Stull, 2001). Given this promising association, the MYL12A gene was bought, and cloned with an N-terminal twin-strep tag in pET28, expressed in *E. coli* and purified by resin affinity (Figure 4.20) and SEC (not shown). Given phosphorylation of MYL12A at T19 and S20 leads to MYL12A-induced actin rearrangement, I also expressed and purified a MYL12A T19A/S20A double mutant. Interestingly, close to the end of my period of study, MYL12A was identified as a novel substrate of SteC (Dai 2024), adding strength to this finding.

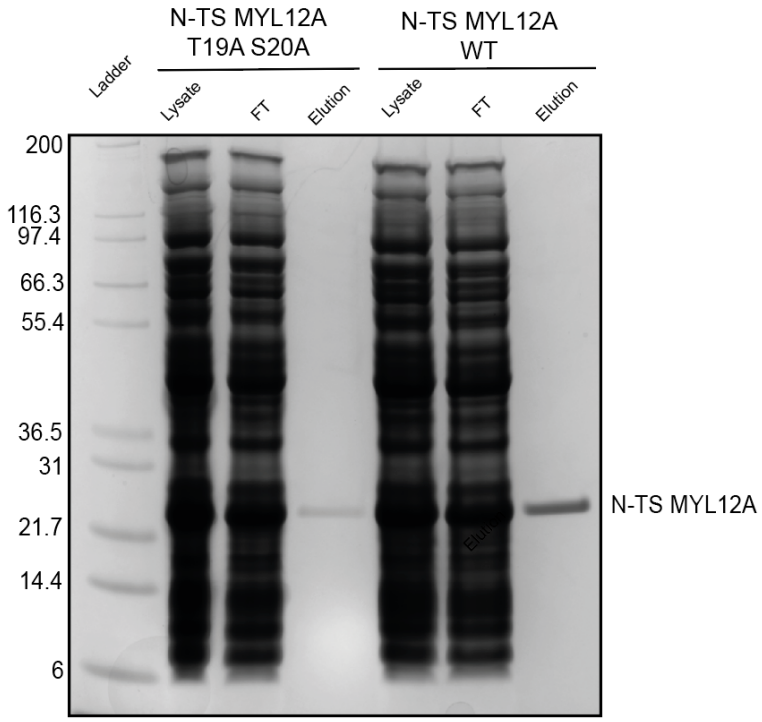


Figure 4. 20: Streptactin resin purification of MYL12A double mutant and WT

MYL12A T19A/S20A and MYL12A WT (Q6ZWQ9) with N-terminal TwinStrep were expressed in *E. coli* and purified with streptactin resin affinity by gravity flow and SEC into standard buffer and flash frozen in liquid nitrogen. FT, flow-through; TS, TwinStrep

4.4.3 Kinase assay with MYL12A

To assess *in vitro* kinase activity of SteC towards MYL12A as a substrate, a radioactive kinase assay was performed (Figure 4.21). Proteins were loaded alone (first 6 columns) or in all possible kinase-substrate pairs (last 8 columns), demonstrating pure protein in all cases. The quantities of MYL12A and MYL12A double mutant protein are different due to a pipetting error, but given the lack of signal for MYL12A DM, this does not interfere with interpretation of the results. Autophosphorylation was seen for both SteC 1-457 and 210-457 WT proteins, but not for the K256H mutants. Then, there was a strong signal representing phosphorylation of MYL12A after only 15 minutes of autoradiographic film exposure when WT SteC (full length of kinase domain) was included in the assay. No phosphorylation of MYL12A was detected when SteC K256H was incubated with the substrate. Of note, WT SteC, was unable to phosphorylate MYL12A T19A/S20A mutant, corroborating the literature that SteC phosphorylates MYL12A at T19 and/or S20. From this, I conclude that *in vitro*, SteC phosphorylates MYL12A and that this requires the catalytic activity of SteC as well as residues T19 and/or S20 of MYL12A.

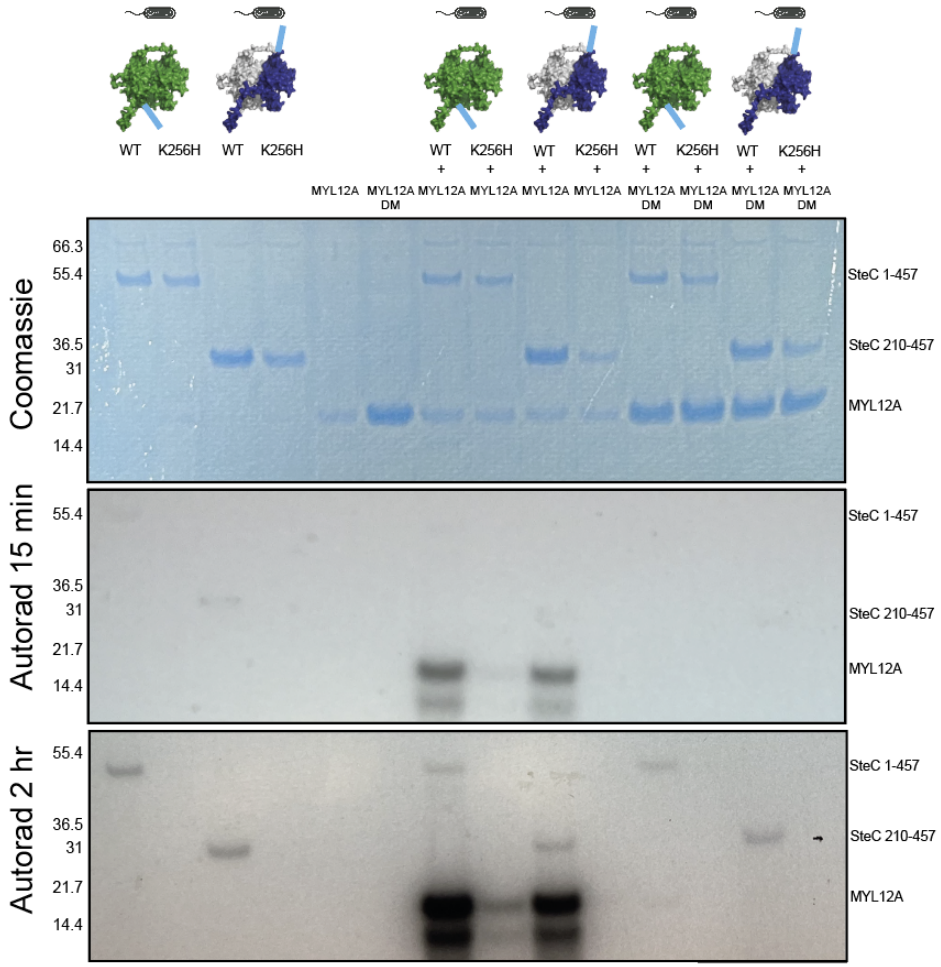


Figure 4. 21: Radioactive kinase assay comparing SteC constructs with MYL12A substrate

SteC 1-457 and 210-457 WT and K256H mutants and MYL12A WT and T19A/S20A double mutant (DM) with N-tags were expressed in *E. coli* and purified by Streptactin resin (all proteins) and SEC (210-457 and MYL12A proteins only). They were added at 5 µM, either alone or combined in kinase-substrate pairs, to 100 µM ATP spiked with 20 kBq [γ-P^{32}]ATP, and mixtures were incubated for 30 min at 30 °C and analysed on SDS-PAGE followed by Coomassie staining and detection of P^{32}. Representative of 2 repeats.

4.4.4 Impact of SteC S379 on MYL12A phosphorylation

To assess if the model of S379-dependent activation holds true in the full-length SteC including its N-terminal regulatory domain, and for the new substrate, MYL12A, a kinase assay was performed combining each of SteC 210-457 and 1-457 WT and mutants with N-tag with MYL12A WT (Figure 4.22). The efficiency of SteC kinase activity towards MYL12A required a reduction of kinase concentration by 50-fold and a reduction in the incubation time for the assay to prevent substrate saturation. A short film exposure time was used to prevent

oversaturated signal. MYL12A was seen in equal amounts in all wells by Coomassie. SteC 1-457 and 210-457 are not visible due to their low concentration in this assay. SteC 1-457 and 210-457 WT demonstrate a strong MYL12A phosphorylation signal, despite the lower kinase concentration and shorter autoradiography exposure time. SteC 1-457 and 210-457 K256H mutants both have no detectable kinase activity, while S379A mutants have a small level of activity. Finally, the SteC 1-457 S379D mutant nearly achieves the same kinase activity as the WT, while the 210-457 S379D mutant shows less activity towards MYL12A than its WT counterpart.

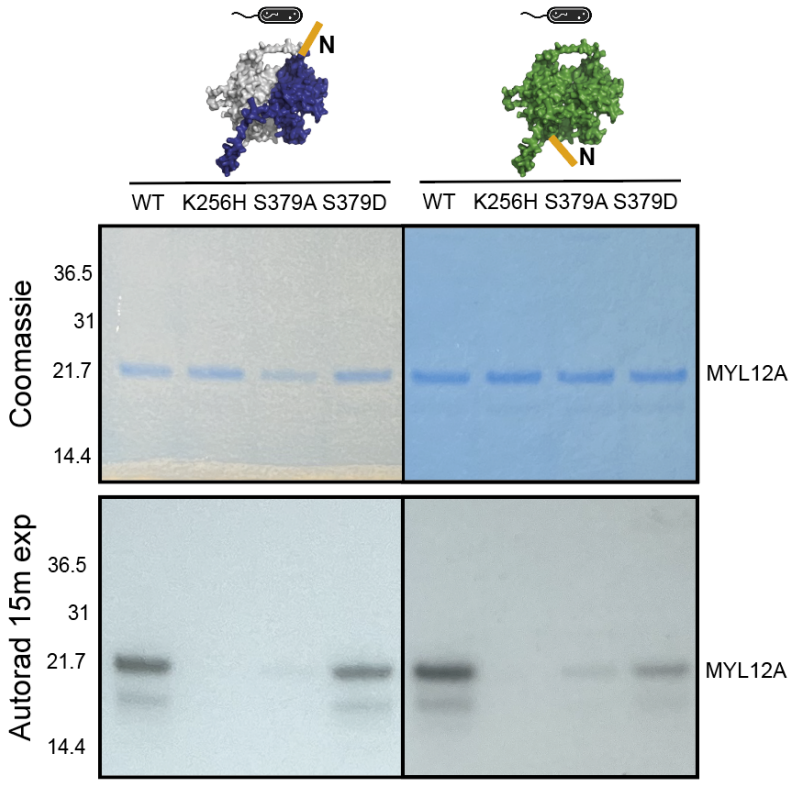


Figure 4. 22: Radioactive kinase assay of SteC 210-457 WT and mutants with MYL12A

SteC 210-457 and 1-457 WT and mutants and MYL12A WT with N-tags were expressed in *E. coli* and purified by standard procedures. 100 nM kinase constructs were combined with 5 µM MYL12A and 100 µM ATP spiked with 20 kBq [γ-P^{32}]ATP, and the mixtures incubated for 5 minutes at 30 °C before SDS-PAGE, Coomassie staining and autoradioagraphy to detect P^{32}. Film exposed for 15 min. Representative of 3 repeats

4.4.5 Assessing kinase activity with Michaelis-Menten kinetics

The strong kinase signal for SteC phosphorylation of MYL12A provides the opportunity to quantify kinase activity kinetically. I plotted a Michaelis-Menten curve where substrate concentration range is subjected to a fixed kinase concentration (Figure 4.23).

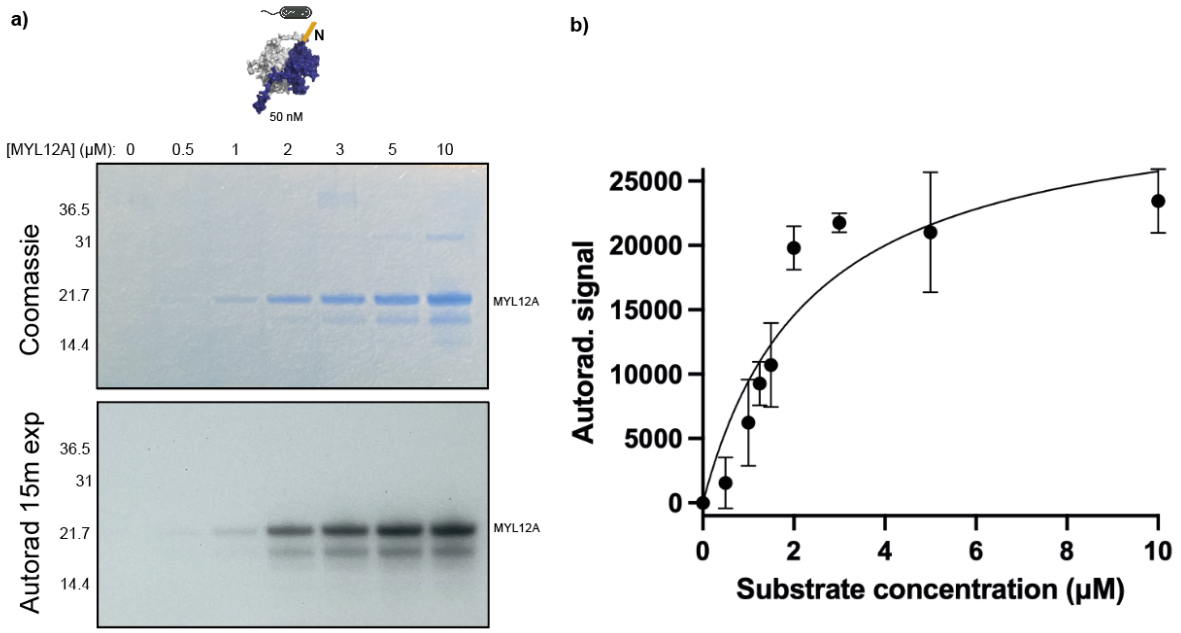


Figure 4. 23: Radioactive kinase assay with fixed SteC concentration and increasing MYL12A concentration

a) SteC 210-457 WT with N-tag expressed in *E. coli* at 50 nM was added to MYL12A at concentrations of 0 to 10 μM and incubated at 30 °C for 1 min prior to quenching.

b) This assay was repeated for the concentrations plotted 3-4 times. ImageJ was used to quantify the intensity of bands and the means and standard errors of the intensities were plotted on a graph of signal against substrate concentration. Michaelis-Menten curve and Km were calculated with Prism 10.

The Coomassie stained gel demonstrates MYL12A at increasing concentrations from left to right (Figure 4.23a, top). The autoradiography signal increases with the substrate concentration (Figure 4.23a, bottom). The strength of this autoradiography signal was estimated in ImageJ and plotted on a graph across 9 concentration points with 3-4 repeats (Figure 4.23b). The data points demonstrate a steep increase then plateau in signal intensity as substrate concentration increases, suggesting SteC kinase activity, which is at fixed concentration, plateaus as it reaches its peak enzymatic efficiency. Using the Michaelis-Menten equation, a curve can be fitted to these data (line). By this method a K_m of SteC kinase domain for MYL12A substrate is calculated at 2.38 μM (95% confidence interval, 1.44 – 3.88 μM and R^2 of 0.813). The Michaelis-Menten curve is classically hyperbolic, as plotted. However, the data may suggest a sigmoidal pattern. If this were the case, it would mean the substrate has an allosteric effect on SteC activity, i.e. that substrate binding leads to a conformational change in the enzyme that is more catalytically active.

4.4.6 Kinase assay with calmodulin

Calmodulin was identified as a second potential substrate of SteC in the mass spectrometry experiment (Table 4.4). Calmodulin binds calcium and mediates the control of many enzymes, ion channels and other proteins in the cell (Reichow et al., 2013). Interestingly, myosin light-chain kinases are stimulated by the calmodulin-calcium complex. To assess SteC kinase activity against this putative substrate, calmodulin was purchased commercially and assessed with a radioactive kinase assay (Figure 4.24).

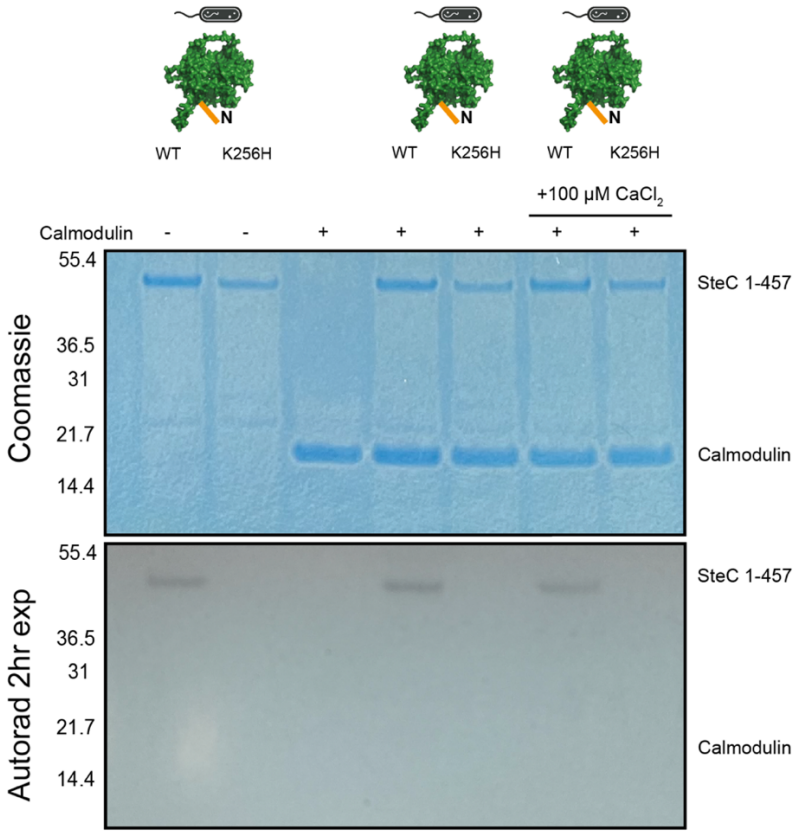


Figure 4. 24: Radioactive kinase assay of SteC and calmodulin

SteC 1-457 WT and K256H with N-tag were expressed from *E. coli* and purified by standard procedures. Recombinant calmodulin was purchased (Merck). SteC at 2.5 µM and calmodulin at 5 µM were incubated alone or in kinase-substrate pairs with or without 100 µM $CaCl_2$. 100 µM ATP spiked with 20 kBq [γ-P^{32}]ATP was added to all samples before incubation for 30 mins at 30 °C prior to SDS-PAGE, Coomassie and autoradiography to detect P^{32}. Film exposed for 2 hours. Representative of 2 repeats.

SteC 1-457 and calmodulin can be seen by Coomassie at the expected molecular weight. SteC 1-457 WT demonstrates autophosphorylation signal, whereas the K256H mutant does not, as previously shown. No signal of phosphorylation was seen at the expected position for calmodulin for SteC 1-457 WT or K256H, and the addition of calcium had no effect. This suggests calmodulin is not a substrate of SteC in these conditions, and MYL12A is more likely to be the substrate identified in the murine cell lysate assay.

Conclusion

Prior to this thesis, the catalytic activity of recombinantly expressed SteC had been demonstrated against the model substrate MBP (Poh et al., 2008) and proposed biologically relevant substrates MEK1 (Odendall et al., 2012) and FMNL1 (Walch et al., 2021). The K256H mutation was shown to result in loss of kinase activity (Poh et al., 2008). In this chapter, I have validated FMNL1 as a substrate of SteC. Furthermore, I have found the K256H mutation to be an effective kinase-dead mutation that prevents auto- and substrate phosphorylation by preventing nucleotide binding without interfering significantly with the secondary structure of the kinase domain.

The data presented here also show that SteC is phosphorylated at S379, which is located in the proposed activation loop of the kinase domain. Mutagenesis demonstrates that a negative charge at this residue is required for kinase activity, both in the case of substrate peptides and fully folded substrate proteins, confirming that the activation of this kinase is dependent on S379 phosphorylation. S379 phosphorylation may activate SteC by improving ATP binding affinity, as we show that the S379A mutant had minimal Mant-AMPNP binding at the concentrations tested. S379 phosphorylation also triples substrate binding affinity as demonstrated by BLI with the substrate peptide. SteC S379 phosphorylation is more complete when SteC is expressed in Sf9 cells rather than *E. coli*, suggesting a eukaryotic upstream kinase is more likely to be responsible for SteC activation than autophosphorylation, and SteC S379 can be phosphorylated by the mammalian kinase STK38 (NDR1). I also show that SteC can autophosphorylate at S379 in cis but not in trans, a process facilitated by the C-tail (430-457). SteC kinase activity against FMNL1 is higher when expressed from Sf9 than from *E. coli*, again suggesting a eukaryotic activating mechanism predominates.

Using murine cell lysate, I identified MYL12A as a potential substrate of SteC, corroborating data from another group that was published towards the end of this thesis (Dai et al., 2024), and demonstrating that SteC phosphorylates MYL12A exclusively at the canonical activating sites, T19 and S20. The efficiency of this reaction, which greatly surpasses the phosphorylation of FMNL1, meant that enzyme kinetics could be performed and a Km for SteC with MYL12A as a substrate was estimated to be 2.38 μM.

These findings pose further questions. First and foremost: what is the biological relevance of S379 phosphorylation for SteC's actin polymerisation activity in cells (see Chapter 5)? Secondly, what is the biological relevance of N-domain phosphorylation at S75, or kinase domain phosphorylation in sites other than S379? Thirdly, what are the allosteric steps from

S379 phosphorylation to increased kinase activity? Attempts are made to answer this question with mutagenesis in Chapter 5, though further structural and biochemical work would be required to tease this apart. Fourthly, what is the relevance of the C-tail in cells? The aggregation-prone nature of the recombinantly expressed constructs including this C-tail makes activity hard to compare against the more stable SteC 210-429 construct lacking the C-tail. The relevance of the C-tail in cells is also explored in Chapter 5. Finally, the biological significance of the novel substrate, MYL12A, requires further study. Colleagues in the field propose its importance for the migration of infected macrophages (Dai et al., 2024). This is also interrogated in Chapter 5.

5. Function of SteC in mammalian cells

Introduction

The biochemical data presented in Chapter 4 suggest that SteC requires phosphoactivation, by autophosphorylation or by a host kinase, and that it has a C-terminal tail (residues 429-457 and referred to as the C-tail) important for substrate binding. In this chapter I aimed to explore the relevance of these findings to the function of SteC in mammalian cells. SteC is responsible for the polymerisation of actin both in transfected, and *Salmonella*-infected, mammalian cells, and this is kinase-dependent (Poh et al., 2008). Reproducing these actin polymerisation phenotypes and testing the relevance of S379 phosphorylation for these was the main focus of further study.

5.1 Screening for key catalytic residues by transfection

Given SteC is associated with actin polymerisation in transfected mammalian cells in a K256-dependent manner (Poh et al., 2008), i.e. likely to be based on its kinase activity, we hypothesised that other key residues in the kinase domain mutated to alanine would lose the ability to polymerise actin. Experiments in this section were performed in collaboration with Briac Lemetais, an MRes student supervised by Dr Ines Diaz Del Olmo (Post-Doc, Thurston laboratory) and me.

5.1.1 SteC-mediated actin polymerisation upon ectopic expression of SteC in HEK 293ET cells is kinase-dependent

HEK 293ET cells transfected with a pTCMV vector ectopically expressing GFP-SteC induced actin polymerisation in dense F-actin foci visualised by phalloidin stain (Figure 5.1a), as previously reported (Poh et al., 2008). Meanwhile, ectopically expressed GFP-SteC K256H demonstrated no actin foci (Figure 5.1b), again corroborating the existing literature. Cells expressing GFP-SteC S379A, the phosphoresistant mutant presented in Chapter 4, also demonstrated no actin foci (Figure 5.1c), while those expressing GFP-SteC S379D, the phosphomimetic mutant, demonstrated actin foci (Figure 5.1d) supporting the model of kinase phospho-activation and its importance in cells. Combined, these data demonstrate an actin polymerisation effect that is dependent on kinase activity (Chapter 4). Interestingly, the actin foci seen in cells expressing GFP-SteC S379D were less dense, larger and appeared to be associated with membrane blebs, unlike cells expressing WT GFP-SteC (Figure 5.1). Furthermore, the cellular localisation of the GFP signal differs between mutants. GFP-SteC WT appears to concentrate at the site of F-actin; GFP-SteC K256H and S379A have a

perinuclear predominance; and GFP-SteC S379D appears diffuse throughout the cytoplasm (Figure 5.1). This suggests that the S379D mutant imperfectly mimics phosphorylated serine.

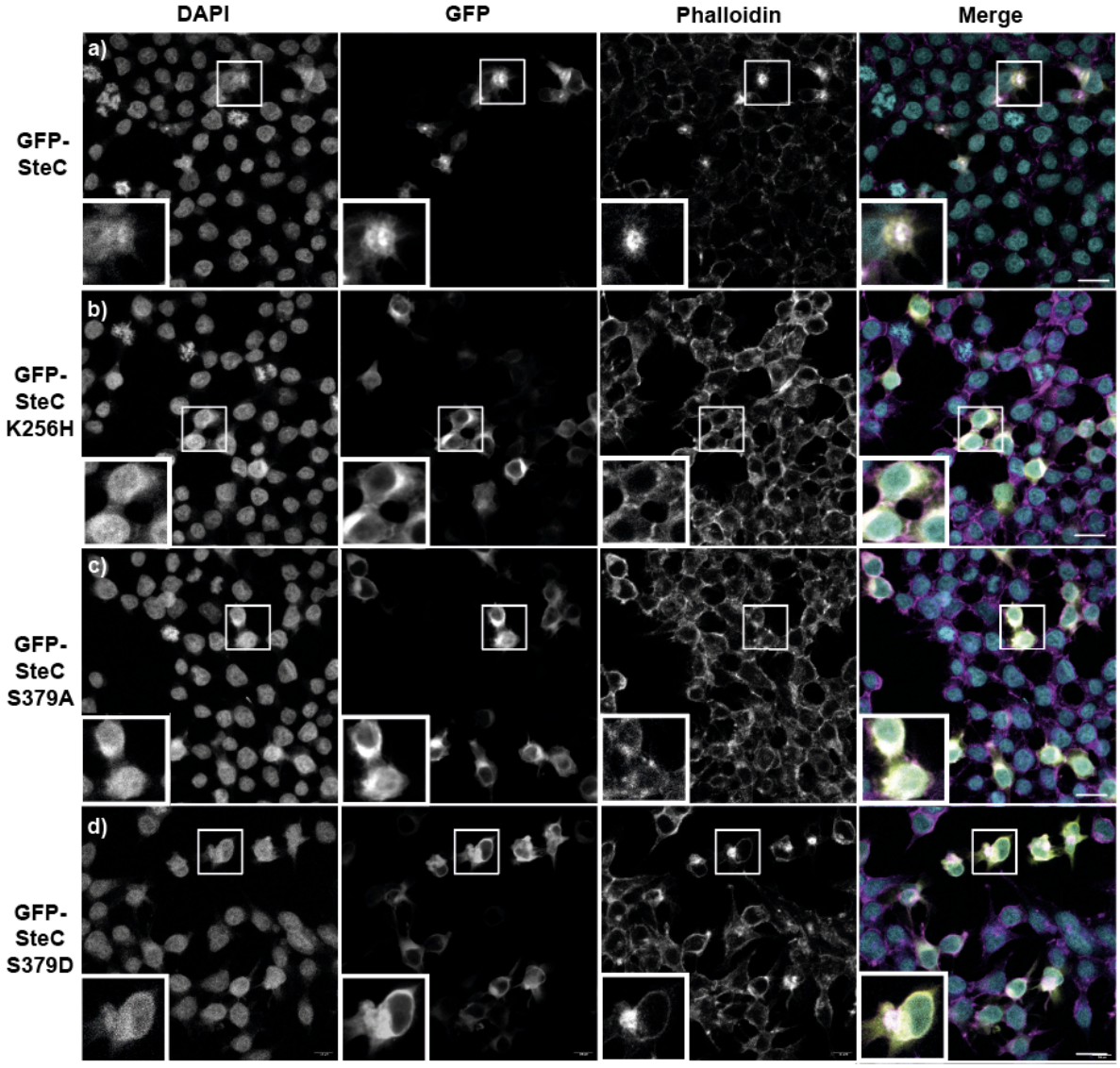


Figure 5. 1: HEK 293ET cells ectopically expressing GFP-SteC WT, K256H, S379A or S379D

HEK 293ETs were transfected with 600 ng of a) pTCMV GFP-SteC, b) pTCMV GFP-SteC K256H, c) pTCMV GFP-SteC S379A or d) pTCMV GFP-SteC S379D. After 24 hr, cells were fixed, stained with DAPI (4',6-diamidino-2-phenylindole, coloured blue in the merged image) and phalloidin AF647 (pink) and imaged. GFP is false coloured in yellow in the merge. Scale bar represents 20 µm. Boxed insets are 2-fold magnifications of the indicated regions. Images were selected at random and are representative of 3 separate experiments.

Canonical kinase motif	(Proposed) Role	Candidate mutants	Rationale for mutation
Catalytic lysine	Involved in favourable ATP positioning.	K256H (K265A)	Established catalytic residue *in vitro*, prevents ATP binding (Chapter 3). Close-by lysine, may be alternative to K256.
αC glutamate	Salt bridge with catalytic lysine to orient it correctly for catalysis.	E272A (E261A D279A)	Seems most likely to play this role from structure. Two negatively charged amino acids in the αC helix that may form salt bridge with K, though less likely based on AF.
αC basic residue	Stabilises a phosphorylated activation loop.	R280A (R274A K283A)	Primary candidate given orientation towards S379. Alternative basic residues in αC helix that may play role.
HRD-like motif	Stabilises phosphorylated activation loop and mediates γ-phosphate transfer onto substrate.	H343A D344A	Canonically, this motif is either HRD if a phosphorylated activation loop is required for activity, or H*D if not. No R seen in SteC. H may mimic the basic features of R. D344A mutant also expressed recombinantly and is kinase-inactive *in vitro* (Figure 4.18).
DFG-like motif	Important to form the R-hydrophobic spine and coordinate the divalent cation.	D362A D364A (F359A& F360A)	Similar residues are seen in the equivalent part of the AF structure, though in a different order: 'FFVDGD'. Two aspartic acid residues most likely. FF double mutation was performed because we suspected that they were not key residues and could explore further if they were essential for function.
pAL site	Many kinases are phosphorylated in the activation loop to become active.	S379A S379D	As shown in Chapter 4 with recombinantly expressed protein, S379 is phosphorylated, and is catalytically important.
Non-canonical C-tail		1-429	Given the unusual nature of the C-tail, and findings in Chapter 4, a truncated version of SteC was cloned

Table 5. 1: List of 16 SteC kinase domain mutations used to screen for key catalytic residues by transfection, with rationale

Mutations in brackets are alternatives to the primary candidate(s). AF, Alphfold2; pAL, phosphorylated activation loop. Relevant residues are highlighted in Figures 3.6 and 3.7.

5.1.2 Mutagenesis to identify functionally relevant residues in SteC

Given that transfected SteC induced the previously reported actin foci phenotype, this experimental set up allowed for the facile screening of further SteC mutants. Key residues were selected based on knowledge of the literature on kinase activity (Section 1.3); the modelling of SteC kinase domain alongside canonical PKA (Figures 3.6 and 3.7); and features of SteC kinase activity presented in Chapter 4. The 16 mutants chosen for testing are listed in Table 5.1 with their rationale. These mutants were cloned into the pTCMV vector for expression in mammalian cells with an N-terminal GFP tag.

5.1.3 Novel SteC mutants R280A, H343A, D344A, F359A-F360A, D362A, D364A and S379A lost the ability to polymerise actin when ectopically expressed in HEK 293ET cells

The results in Section 5.1.2 provide evidence that HEK 293ET cells ectopically expressing GFP-SteC in a kinase-active state is sufficient to induce actin polymerisation, while those expressing GFP-SteC with kinase-inactivating mutations do not. Therefore, transfection experiments were performed for the remaining mutants in Table 5.1, to explore which other residues were important for the observed actin polymerisation phenotype (Figure 5.2a).

The actin polymerisation effect of mutants can be broadly divided into those that retain actin polymerisation ability (Figure 5.2a and c-left) and those that lose activity (Figure 5.2b and c-right). The following mutants lose actin polymerisation capacity when ectopically expressed in HEK 293ET: SteC K256H, R280A, H343A, D344A, F359A-F360A, D362A and D364A and S379A. This suggests that each of these residues is likely to be required for catalytic activity of substrate phosphorylation. The expression levels of GFP-SteC mutants were analysed by immunoblot (Figure 5.2d) and appear comparable.

Interestingly, even though recombinantly expressed SteC 210-429 showed lower affinity binding to the substrate peptide than SteC 210-457 (Section 3.3.4) and was unable to phosphorylate FMNL1 1-385 while SteC 210-457 was able to (Section 4.3.3), ectopic expression of SteC 1-429 induced levels of actin polymerisation that were comparable to WT SteC (Figure 5.2c). This raises the possibility that the C-tail is not important for polymerisation of actin in cells or that FMNL1 phosphorylation is not essential for the formation of actin foci upon the overexpression of SteC.

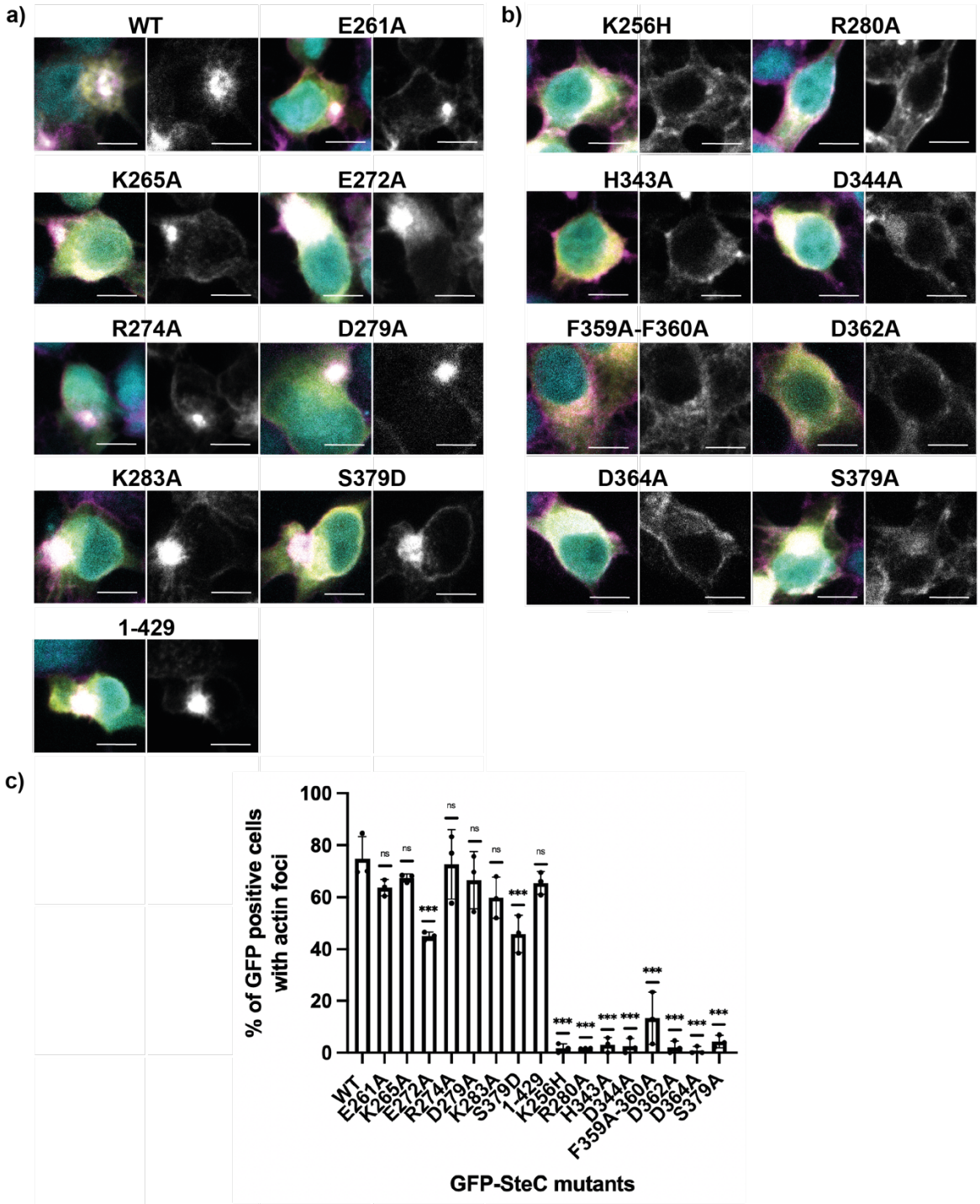

a)
WT
E261A
K265A
E272A
R274A
D279A
K283A
S379D
1-429
b)
K256H
R280A
H343A
D344A
F359A-F360A
D362A
D364A
S379A
c)
% of GFP positive cells with actin foci
100
80
60
40
20
0
ns
***
WT
E261A
K265A
E272A
R274A
D279A
K283A
S379D
1-429
K256H
R280A
H343A
D344A
F359A-360A
D362A
D364A
S379A
GFP-SteC mutants
d)
WT
GFP
NT
R280A
S379D
1-429
K283A
D362A
D364A
H343A
F359A-F360A
K256H
S379A
D344A
D279A
E261A
E272A
K265A
R274A
75
50
25
GFP-SteC
Tubulin
GFP

Figure 5. 2: Quantification of actin polymerisation in HEK 293ET ectopically expressing GFP-SteC WT and mutants

HEK 293ETs were transfected with pTCMV GFP-SteC WT and mutants using Lipofectamine. After 24 hr, cells were fixed, stained with 4'6-diamidino-2-phenylindole (DAPI) and phalloidin AF647 and imaged.

a) Representative merged images with DAPI (cyan), phalloidin AF647 (magenta) and GFP (yellow) and a grayscale for the phalloidin channel (right) are demonstrated for each active mutant, i.e. for conditions in which actin polymerisation was retained. Representative cells were chosen randomly. Scale bar represents 5 µm.
b) As for a), the mutants that lost actin polymerisation are demonstrated.
c) For each condition, 100 transfected cells from each of 3 coverslips were blind counted. Each dot represents a coverslip and bars represent mean and standard deviation of 3 coverslips. Statistical analysis: One-way ANOVA (***, $p < 0.001$) with Dunnett's multiple comparisons test. Assumption of normally distributed data was accepted using a Shapiro-Wilk and Kolmogoro-Smirnov test.
d) Cells transfected with GFP-SteC mutants were lysed and immunoblotted for GFP and tubulin. Representative of 2 experiments.

Exceptions to the binary classification (actin foci / no actin foci) include SteC E272A. This mutant polymerised actin in significantly less transfected cells than SteC WT (45.1% vs 74.8%, respectively $p<0.0001$) (Figure 5.2c). Given E261A (63.7%, $p=0.274$ vs. WT) and D279A (66.6%, $p=0.636$) mutations had no significant effect on actin polymerisation, E272 may be the most important, albeit not essential, charged αC residue for kinase activity. Cells expressing SteC S379D also demonstrated significantly less actin polymerisation compared to cells expressing WT SteC (45.9%, $p<0.0001$). This may be due to aspartate being unable to fully mimic a phosphate at that site, as seen with *in vitro* kinase assays with full-length SteC (Figure 4.22). Finally, ectopic expression of an SteC F359A F360A double mutant retained actin polymerisation capacity in only 13.3% of transfected cells ($p<0.0001$).

To further test for correlation between actin polymerisation capacity upon SteC transfection and kinase activity *in vitro*, mutations of residues E272, D279 and D344 were introduced into SteC 210-429, which readily expresses. Following recombinant protein expression and purification from *E. coli* with a C-terminal tag, a radioactive kinase assay was performed (Figure 5.3). Coomassie stain demonstrated similar loading of the SteC 210-429 WT and mutant variants and FMNL1 peptide (Figure 5.3). SteC 210-429 WT demonstrated weak auto-phosphorylation signal and strong FMNL1 peptide phosphorylation, both in an ATP-dependent manner, and these signals were completely lost with the K256H mutation (Figure 5.3), as previously shown (Figure 4.3). SteC 210-429 E272A and SteC 210-429 D279A demonstrated auto-phosphorylation and substrate phosphorylation signals similar to WT SteC 210-429. In contrast, mutation of D334A resulted in no autophosphorylation signal for SteC 210-429 and

only a small amount of substrate phosphorylation (Figure 5.3), which is in line with data presented previously in Figure 4.17. In summary, E272A and D279A mutants retained kinase activity against FMNL1 peptide, while D344A lost most of its activity (Figure 5.3). These findings broadly correlate with their observed effect on actin polymerisation (Figure 5.2), providing more evidence that actin polymerisation in HEK 293ET cells ectopically expressing SteC is dependent on its kinase activity.

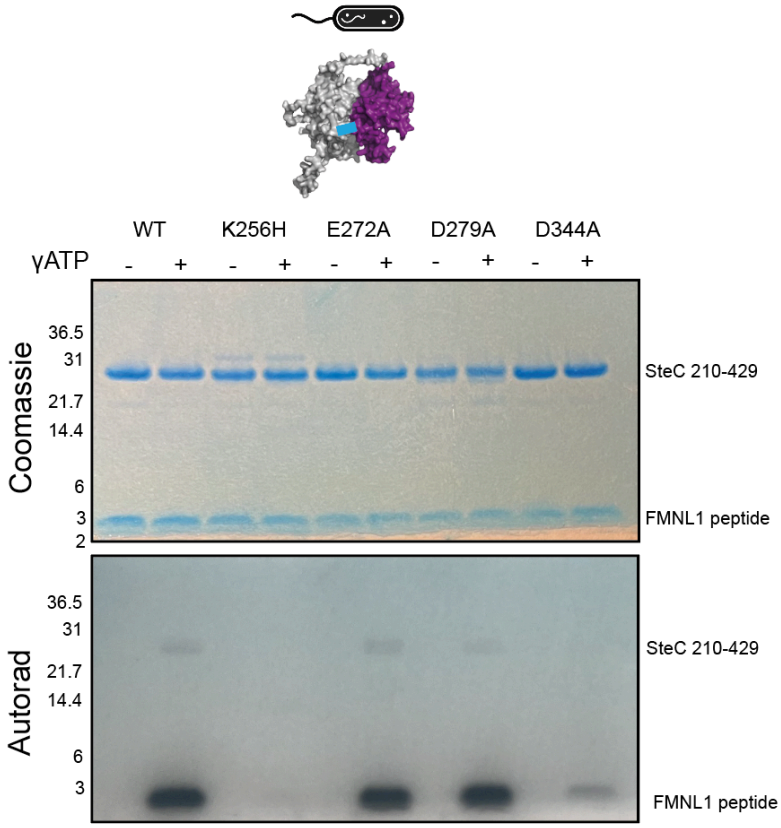


Figure 5. 3: SteC mutants E272A and D279A retain the ability to phosphorylate an FMNL1 peptide

SteC 210-457 WT and mutants with C-tag expressed in *E. coli* and purified by standard procedures were combined at 5 µM with 100 µM FMNL1 peptide and 100 µM ATP spiked with 20 kBq [γ-P^{32}]ATP before incubation at 30 °C for 30 mins. Mixtures were analysed by SDS-PAGE and gels assayed by Coomassie staining and autoradiography to detect P^{32}. Radioactive film was exposed for 4 hrs. Representative of 2 repeats.

In this section, many essential amino acids of SteC were identified based on the ability of ectopically expressed SteC to induce actin polymerisation in HEK 293ET cells. Where tested, the polymerisation capacity correlated with kinase activity against FMNL1 as a substrate. These essential amino acids for SteC function, namely SteC K256, R280, H343, D344, F359-F360, D362, D364 and S379, will be integrated with results from Chapters 3 and 4 into a model of SteC kinase activation and activity in the discussion (Chapter 6).

5.2 The role of SteC phosphorylation at S379 and the C-tail in *Salmonella* infection

Data in Section 5.1 suggest that *in vitro* kinase activity and actin polymerisation activity of SteC in transfected cells is dependent on SteC residues K256 and S379, among other newly identified putative catalytic residues. Furthermore, the C-tail, while important for FMNL1 substrate binding to (Section 3.3.4) or kinase activity against FMNL1 (Section 4.3.3), is not absolutely required for actin polymerisation in transfected cells (Figure 5.2a-b). One possible explanation for this disconnect is the non-physiological concentration of ectopically expressed SteC following mammalian cell transfection. As actin polymerisation during *Salmonella* infection of mammalian cells is dependent on the kinase activity of SteC (Odendall et al., 2012; Poh et al., 2008), in this section I aimed to assess the contributions of K256, S379 phosphorylation and the C-tail to SteC-induced actin foci formation following translocation of SteC from bacteria into the mammalian host cell.

5.2.1 Designing plasmids to complement a Salmonella *knock-out strain*

The *Salmonella* strains produced for this section are listed in Table 2.2. The wild-type (WT) *Salmonella* Typhimurium 14028s strain and equivalent strain with genomic *steC* replaced with a kanamycin resistance cassette were available from the Thurston laboratory. The knockout strain was transformed with pWSK29 plasmids containing the gene for *steC* WT, or *steC* variants with the following mutations: K256H, S379A, S379D and 1-429 truncation. *steC* was driven from its endogenous promoter region and Shine-Dalgarno sequence and followed with a C-terminal 2HA tag. The endogenous promoter was retained to mimic the *in cellulo* expression of SteC. The Shine-Dalgarno sequence facilitates *Salmonella* ribosome binding to the messenger RNA. The 2HA-tag was placed on the C-terminus because the N-terminus of the open reading frame (ORF) is believed to hold an amino acid sequence that signals translocation of the effector via the T3SS (Y. Wang et al., 2013), that would likely be disrupted with the addition of an N-terminal tag. For the same reason, constructs of the *steC* kinase domain alone retained the first 25 amino acids of the N-terminal regulatory domain, corresponding to a region of the AF prediction containing a small β-sheet and α-helix and predicted with a low pLDDT score (see Figure 3.5). A *Salmonella* strain transformed with an empty pWSK29 plasmid was also produced as a control for plasmid burden within the strains.

5.2.2 SteC is translocated into the host cytosol if its N-terminus is intact

To assess if WT and SteC variants are delivered across the bacterial and SCV membranes into the host cytosol by *Salmonella*, a translocation assay was performed (Figure 5.4). A translocation assay involves infecting host cells with *Salmonella* for a defined period of time,

then lysing the host cells and SCV without lysing the bacterial membrane, separating the pellet and post-nuclear supernatant (PNS) fractions by centrifugation, and immunoblotting each to assess whether a) an effector protein is expressed in *Salmonella* (present in the pellet fraction) and b) whether it is translocated into the host (present in the PNS fraction). Swiss 3T3 mouse fibroblasts (3T3 cells) were used given the actin phenotype has been demonstrated in this cell line from infection previously (Poh et al., 2008).

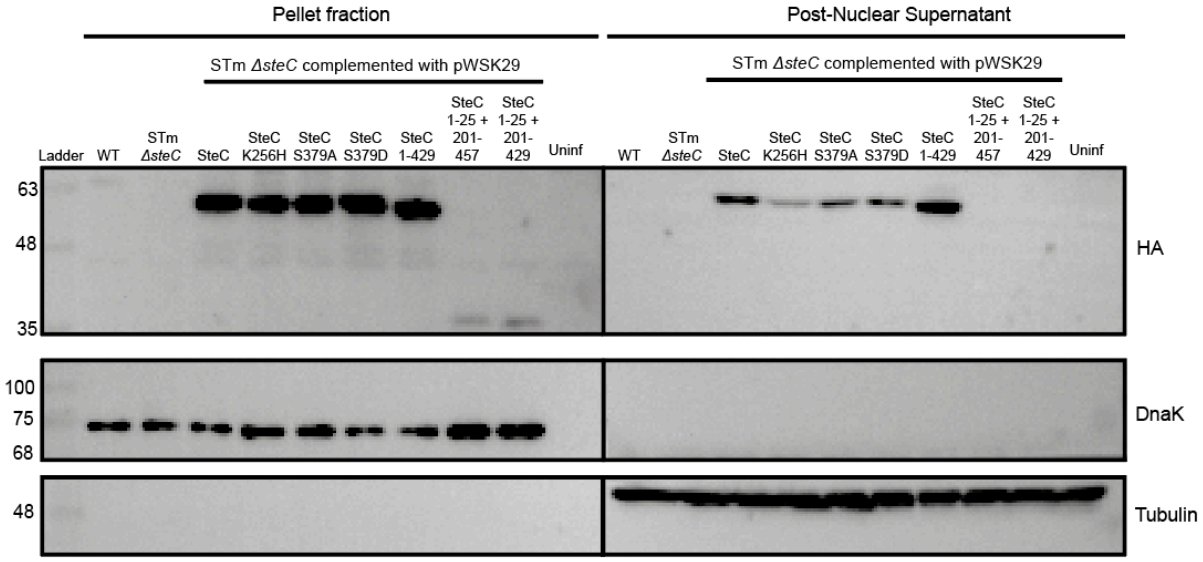


Figure 5. 4: Translocation assay in 3T3 cells by western immunoblot

Swiss 3T3 mouse fibroblasts (3T3 cells) were infected with *Salmonella* WT, *ΔsteC* and *ΔsteC psteC* strains for 20 minutes prior to gentamicin treatment. At 8 hours post invasion, cells were lysed and pellet and post-nuclear fractions analysed by immunoblotting with antibodies against HA, DnaK and Tubulin. Images representative of 2 independent repeats with 1 repeat performed with the help of Mei Liu from the Thurston laboratory.

In this assay, intracellular *Salmonella* bacteria remain intact after mammalian cell lysis and become pelleted. This is evidenced by the detection of the *Salmonella* DnaK protein in the pellet sample, but not the PNS sample. In contrast, tubulin, a cytosolic mammalian protein, is absent from the pellet but present in the PNS fraction (Figure 5.4). Having established the anticipated fractionation of the infected cells, analysis of HA-tagged SteC constructs demonstrated that HA-tagged SteC proteins were detected in the pellet fraction in all cases. This suggested that all the proteins were expressed in *Salmonella*. However, the signal was weaker for SteC 1-25+210-457 and SteC 1-25+210-429. The PNS fraction was confirmed as clear of *Salmonella* bacteria by the lack of DnaK signal. HA tagged SteC WT, K256H, S379A, S379D and C-tail truncation mutants were detected in the post-nuclear supernatant, revealing that they were translocated into the cell cytosol (Figure 5.4). SteC K256H was reproducibly

translocated to a lower degree than WT SteC whereas SteC 1-429 was translocated in higher amounts despite similar expression observed in the *Salmonella* positive pellet fraction. In contrast, no signal was detected for HA tagged SteC 1-25+210-457 and SteC 1-25+210-429 suggesting that they were not translocated. The lower expression of these constructs could contribute to the failure of translocation, and it is likely that the putative N-terminal translocation signal is longer than the 25 amino acids I predicted based on the AF model. Given these two *Salmonella* strains (SteC 1-25+210-457 and SteC 1-25+210-429) failed to translocate the truncated SteC, subsequent experiments focused on the analysis of those strains that were translocated: HA tagged SteC WT, K256H, S379A, S379D and C-tail truncation mutants.

5.2.3 Actin polymerisation in infected 3T3 cells is SteC-dependent

To explore the role of translocated SteC WT and mutants in actin polymerisation during infection, 3T3 cells were infected with *Salmonella* strains for 8 hours prior to fixing, permeabilising and staining.

Prior to analysing the SteC complemented strains, I first established that actin foci occurred in an SteC dependent manner. 3T3 cells infected with WT *Salmonella* demonstrated F-actin in a dense focus within the *Salmonella* microcolony with filamentous or tubular structures emanating from the centre, shown in two separate images (Figure 5.5a-b). Meanwhile, in cells infected with *Salmonella ΔsteC*, no discernible F-actin was seen associated with the microcolony (Figure 5.5c). Cells infected with *Salmonella ΔsteC psteC* recapitulated the WT-pattern of F-actin foci (Figure 5.5d). More abundant and less focused actin polymerisation was observed in the complemented strain (*ΔsteC psteC*), which likely reflects the slightly elevated expression of SteC from pWSK29, even though it is a low copy number plasmid (Rong & Kushner, 1991)(Figure 5.5d).

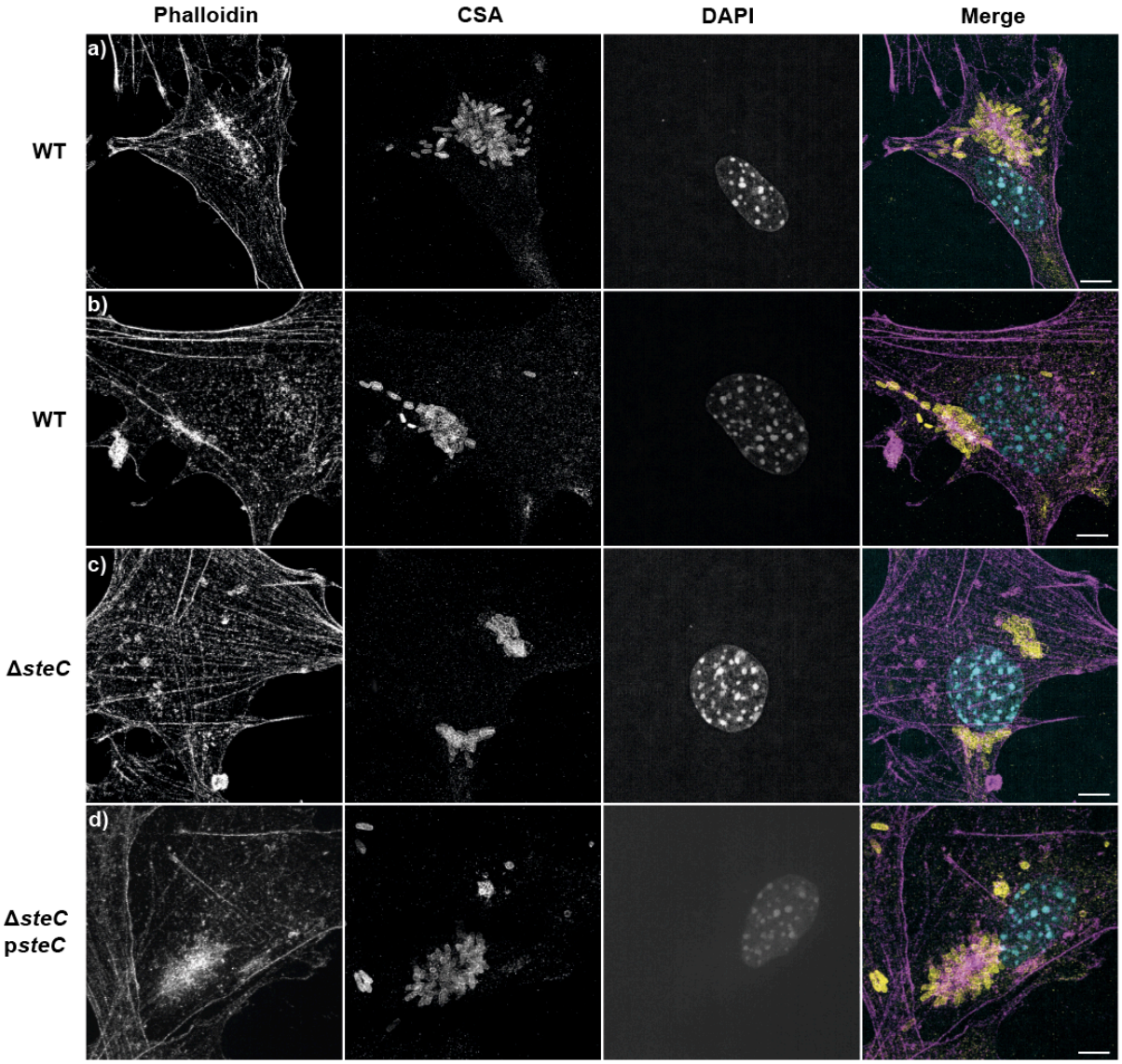


Figure 5. 5: Immunofluorescence of 3T3 cells infected with *Salmonella* WT, *ΔsteC* and *ΔsteC psteC* strains

Swiss 3T3 mouse fibroblasts (3T3 cells) seeded on cover slips were infected with a) and b) *Salmonella* WT, c) *ΔsteC* and d) *ΔsteC psteC* strains for a total of 8 hours, after which the cells were fixed, permeabilised and stained with phalloidin (represented in pink in the merge), CSA-1 (yellow) and DAPI (cyan). Super-resolution images were collected with an Olympus CSU-W1 SoRa spinning disk with 60X lens and deconvolved with the Olympus CellSens software under the supervision of Matt Renshaw and Donald Bell (Light Microscopy STP). Images are representative of data obtained from three independent experiments. Scale bars represent 5 μm.

5.2.4 SteC produces two actin phenotypes in infection, one of which is kinase activity-independent

Next, the *Salmonella ΔsteC psteC,* WT and SteC mutant expressing strains were compared by immunofluorescence microscopy.

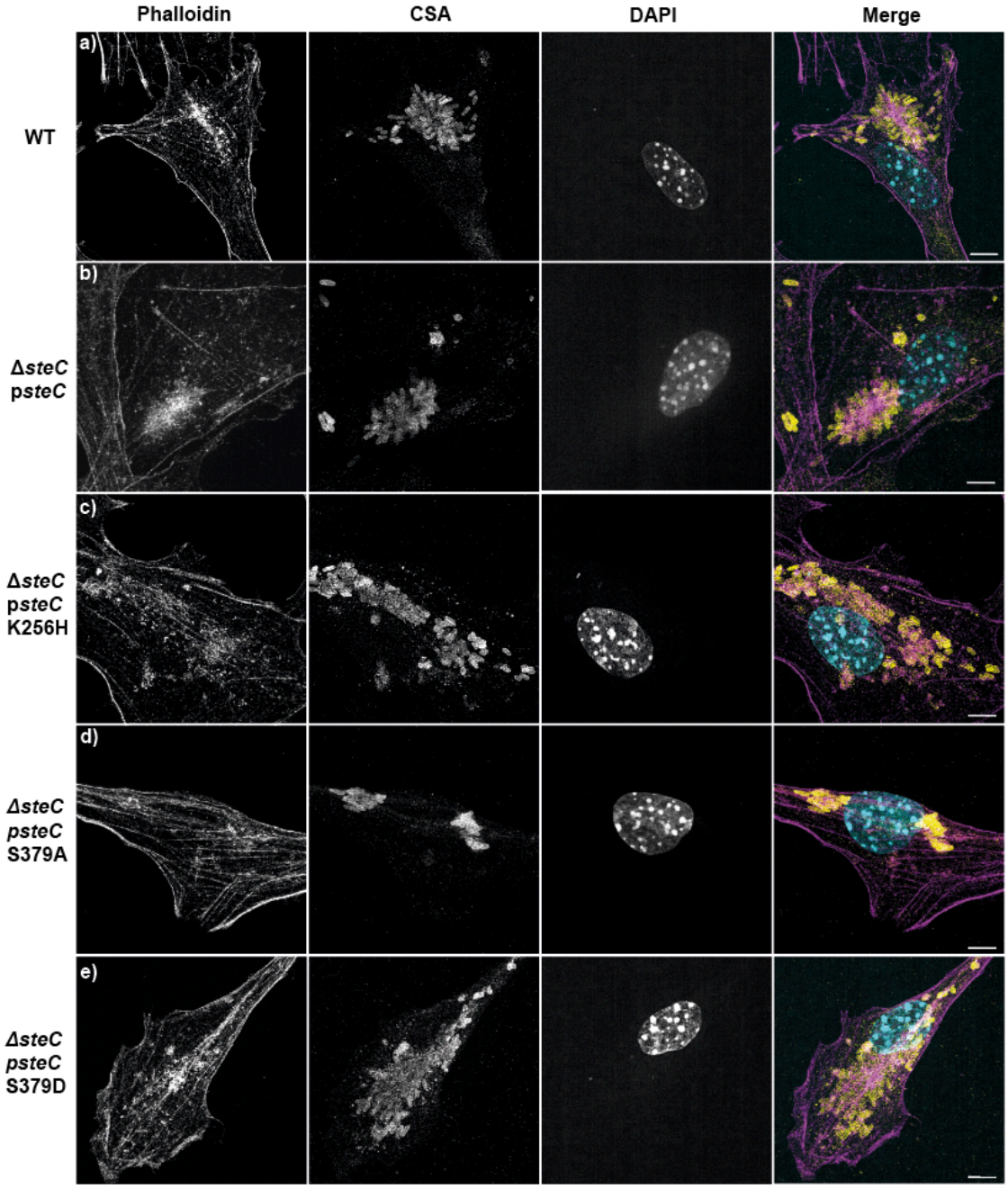


Figure 5. 6: Immunofluorescence of 3T3 cells infected with *Salmonella* WT, *ΔsteC psteC*, K256H, S379A and S379D mutant strains

Swiss 3T3 mouse fibroblasts (3T3 cells) seeded on cover slips were infected with a) *Salmonella* WT (same image as Figure 5.5a for comparison), b) *Salmonella ΔsteC psteC* (same image as Figure 5.5d), c) *Salmonella ΔsteC psteC K256H*, d) *Salmonella ΔsteC psteC S379A* and e) *Salmonella ΔsteC psteC S379D* strains for 8 hours. Following fixation, the cells were stained with phalloidin (represented in pink in the merge), CSA-1 (yellow) and DAPI (cyan). Super-resolution images were collected with an Olympus CSU-W1 SoRa spinning disk with 60X lens and deconvolved with Olympus CellSens software under the supervision of Matt Renshaw and Donald Bell (Light Microscopy STP). Scale bars represent 5 µm.

For *Salmonella* WT and *ΔsteC psteC* strains (Figure 5.6a and b), the same image is displayed as in Figure 5.5a and d to aid direct comparison. In the cells infected with *Salmonella* strains expressing kinase-inactive mutants K256H and S379A (Figure 5.6b and c), a less dense actin polymerisation was seen than in cells infected with *Salmonella ΔsteC psteC*, as if a thin F-actin capsule surrounded groups of bacteria (Figure 5.6a). When observed through an Axio Imager Upright Microscope (Zeiss) passing the field of focus through the cell monolayer, the actin had a bubble-like appearance, further suggestive that clusters of bacteria were partially encapsulated by thin actin structures. I refer to this phenotype as 'capsular actin polymerisation', which is reminiscent of the ring-like structures enclosing multiple bacterial vacuoles displayed in *Salmonella* infection of 3T3 cells previously (Méresse et al., 2001b). 3T3 cells infected with *Salmonella* expressing SteC S379D demonstrated a similar dense focal pattern to the wild type (Figure 5.6e). This analysis suggests that SteC mediates two actin phenotypes that warranted further investigation.

5.2.5 Quantifying actin phenotypes in infection

To quantify these findings across all *Salmonella* strains analysed here, 3-5 coverslips, from 3 independent infections of 3T3 cells infected with each *Salmonella* strain for 8 hours were prepared, and the counter blinded to the bacterial strain. 100 infected cells were counted for each coverslip. The three phenotypes allocated for each infected cell were F-actin in a dense focal pattern associated with *Salmonella* microcolonies (dense focal, Figure 5.7a); F-actin in a capsular pattern associated with *Salmonella* microcolonies (capsular, Figure 5.7b); and no F-actin polymerisation or F-actin not associated with *Salmonella* microcolonies (Figure 5.7c). Infected cells where dense focal actin polymerisation *and* a capsular pattern was seen were classed as dense focal, i.e. the categorisation was treated as mutually exclusive with the dense focal phenotype as "dominant".

Cells infected with *Salmonella* WT, *ΔsteC psteC* and *ΔsteC psteC S379D* strains were associated with actin polymerisation in dense focal pattern in 38.8%, 65.2% and 45.5% of infected cells, respectively (Figure 5.8a). The increased percentage of infected cells showing the dense focal pattern when comparing *ΔsteC psteC* to the WT-infected cells was significant ($p=0.0038$) likely representing elevated levels of SteC expressed from the plasmid (Rong Fu Wang & Kushner, 1991). Cells infected with strains that did not express SteC demonstrated significantly less focal actin polymerisation, i.e. *ΔsteC* (5.97% of infected cells, $p=0.0035$ compared to WT) and *ΔsteC pE* (7.34%, $p=0.0052$). When cells were infected with the *ΔsteC* strain expressing SteC K256H (9.94%, $p=0.0076$), SteC S379A (11.5%, $p=0.0272$) or SteC 1-429 (6.98%, $p=0.0047$) there was also a significant reduction in formation of the dense actin

pattern when compared to WT infected cells (Figure 5.8a). This suggests that dense focal actin polymerisation is dependent on the kinase activity of SteC and that the phosphorylation of S379 is important for this SteC-dependent phenotype in *Salmonella*-infected 3T3 cells.

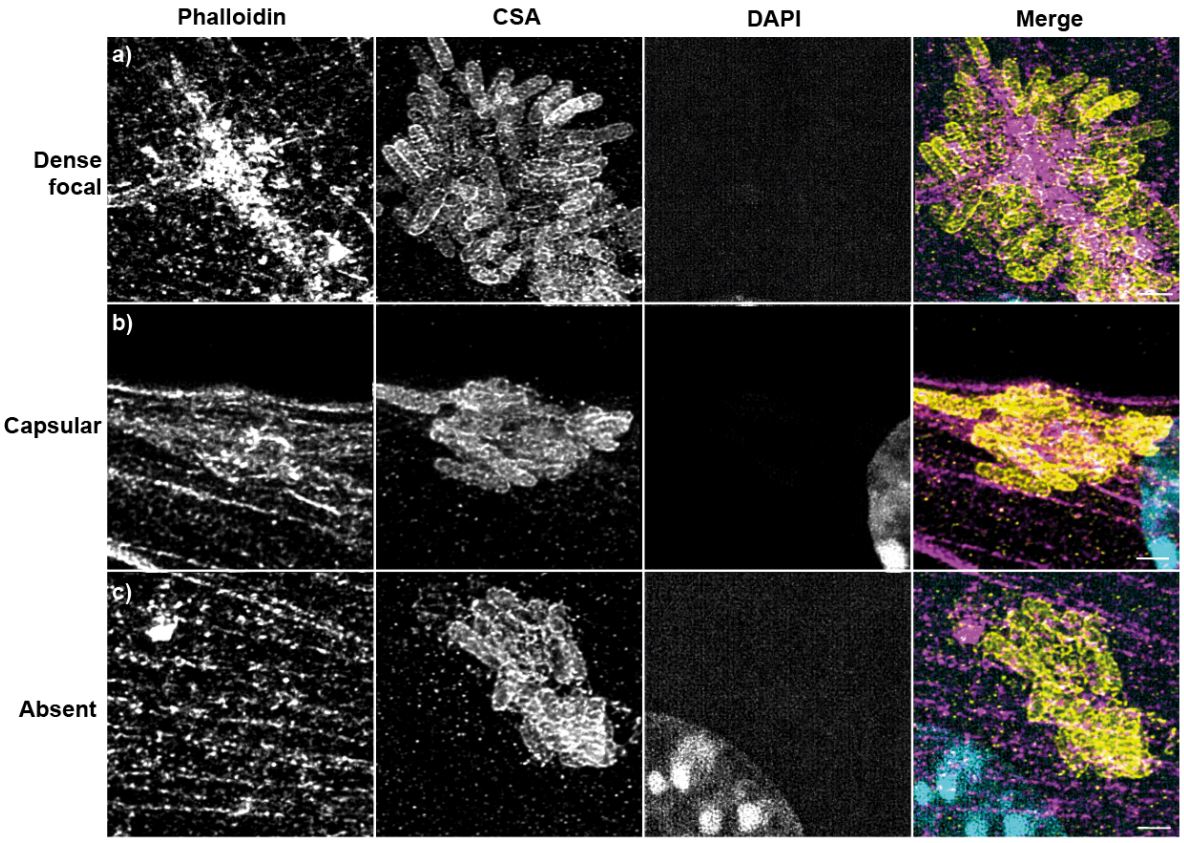


Figure 5. 7: Demonstration of the three actin phenotypes used for quantification

Data as in Figures 5.5 and 5.6 is shown with a further 4 times zoom to demonstrate the three F-actin phenotypes used for quantification of the actin phenotype (Figure 5.8). Scale bars represent 1.25 µm.

a) Dense focal actin polymerisation in 3T3s infected with *Salmonella* WT, as presented in Figures 5.5a and 5.6a

b) Capsular actin polymerisation in 3T3s infected with *Salmonella ΔsteC psteC S379A*, as presented in Figure 5.6d

c) Absent actin polymerisation or F-actin not associated with *Salmonella*, in 3T3s infected with *Salmonella ΔsteC*, as presented in Figure 5.5c

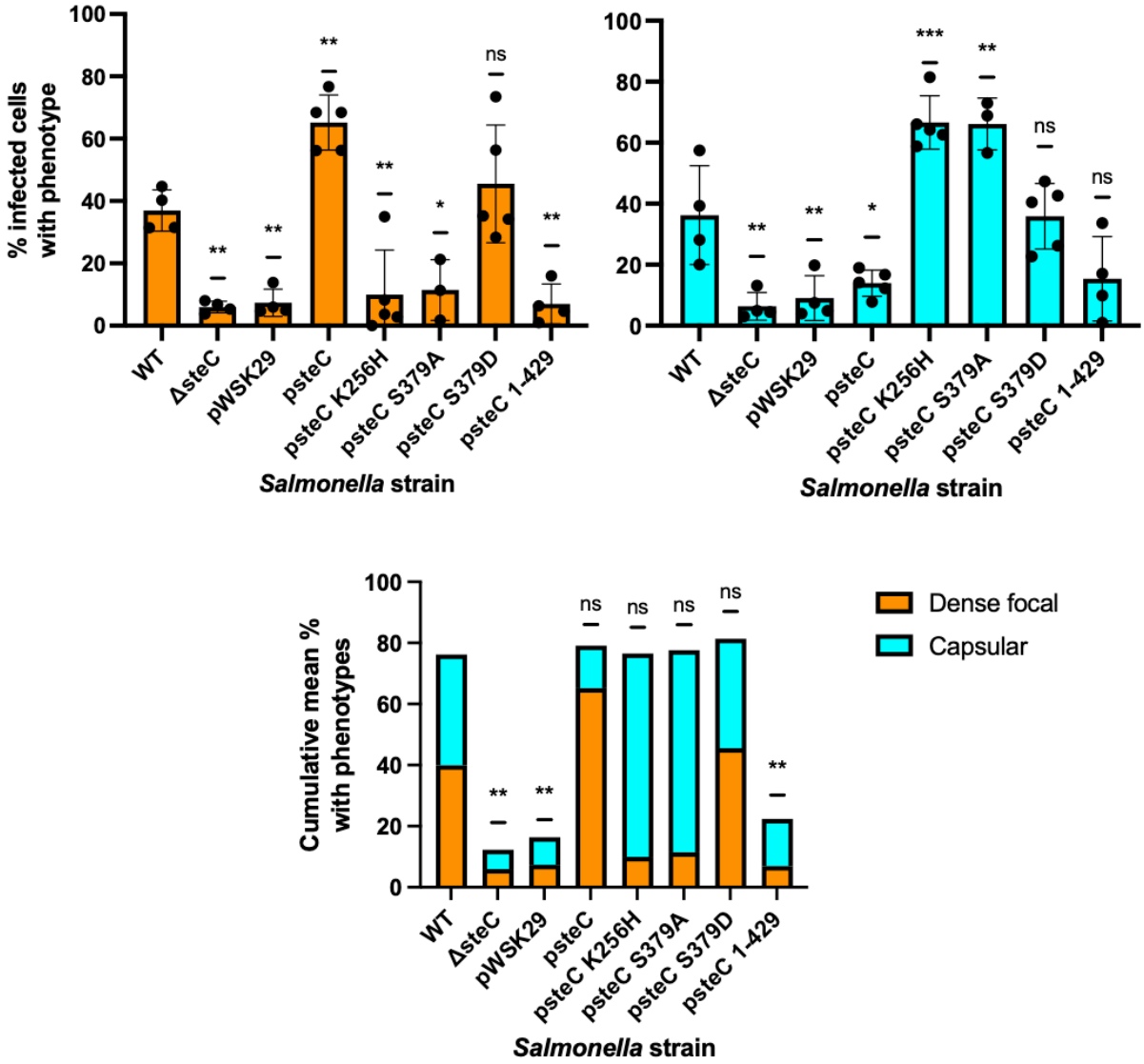


Figure 5. 8: proportion of infected cells with actin polymerisation

For each *Salmonella* strain, 100 infected 3T3 cells from each of 3-5 coverslips were prepared as reported in Figures 5.5 and 5.6 and counted with the counter blinded to the strain using the categorisation introduced in Figure 5.7.

a) Proportion of infected cells displaying dense focal actin polymerisation associated with *Salmonella* microcolonies. Each dot represents a coverslip and bars represent mean and standard deviation of 3 coverslips. Statistical analysis: One-way ANOVA (*, $p < 0.05$; **, $p < 0.01$) comparing each strain to WT, with Dunnett's multiple comparisons test. Assumption of normally distributed data was accepted using a Shapiro-Wilk and Kolmogoro-Smirnov test.

b) Proportion of infected cells displaying capsular actin polymerisation associated with microcolonies, analysed as above.

c) Cumulative percentage of cells with each actin polymerisation pattern – using the means from the coverslips for each condition. For this graph, statistical analysis was performed on the disaggregated values for each coverslip and analysed as above.

Capsular actin polymerisation, on the other hand, was seen in a different pattern according to the *Salmonella* strain. This phenotype was seen in 3T3s infected with *Salmonella* WT (36.2% of infected cells) to a similar amount as the dense focal pattern. Capsular pattern was seen in a significantly higher proportion of cells infected with *Salmonella* strains expressing kinase inactive SteC i.e. *ΔsteC pseC* K256H (70.1%, p=0.0003 compared to WT) and *ΔsteC pseC* S379A (66.2%, p=0.0051) (Figure 5.8b) than WT and seen in a significantly larger proportion of cells infected with *Salmonella* WT than *Salmonella* strains not expressing SteC i.e. *ΔsteC* (6.35%, p=0.0024) and *ΔsteC pE* (9.04%, p=0.0062), or *ΔsteC* strains expressing *pseC* WT (13.9%, p=0.0223). Finally, cells infected with *Salmonella ΔsteC pseC* 1-429 demonstrated a borderline significant reduction in the phenotype compared to WT (15.4%, p=0.0538) and cells infected with *Salmonella ΔsteC pseC* S379D demonstrated a similar amount of capsular actin as WT *Salmonella* infected cells (35.9%, p>0.9999).

Combining the two actin phenotypes, the proportion of infected cells demonstrating any actin polymerisation is similar for *Salmonella* WT (73.2%), *ΔsteC pseC* (79.1%, p=0.9369 vs. WT), *ΔsteC pseC K256H* (76.6%, p=0.9968), *ΔsteC pseC S379A* (77.6%, p=0.9928) and *ΔsteC pseC S379D* (81.4%, p=0.7809). In contrast, only a minority of cells infected with *Salmonella ΔsteC* (12.3%, p<0.0001), *ΔsteC pE* (16.4%, p<0.0001) and *ΔsteC pseC 1-429* (22.4%, p<0.0001) demonstrated any form of actin polymerisation (Figure 5.8c). Given the method of counting, i.e. dense focal actin pattern was prioritised over a capsular one if both were present, this combination compared to Figure 5.8a is a comparison that demonstrates the distribution of the two phenotypes more clearly than interpreting Figure 5.8b alone.

In summary, SteC is responsible for two actin polymerisation phenotypes in infected mammalian cells at 8 hours post invasion. Strains lacking SteC are unable to produce either actin foci or actin capsules around the bacteria. The dense focal actin polymerisation is seen more frequently in cells infected with *Salmonella* WT, *ΔsteC pseC*, and *ΔsteC pseC S379D* conditions and therefore I conclude it is dependent on the kinase activity of SteC. Intriguingly, the actin capsule phenotype is predominantly seen in cells infected with *Salmonella ΔsteC pseC K256H* and *ΔsteC pseC S379A* and therefore appears to be dependent on SteC but independent of kinase activity. Meanwhile, *Salmonella ΔsteC pseC 1-429* was unable to induce either actin phenotype in infected cells, suggesting that the C-tail is important not just for SteC kinase activity but also its non-kinase activity in the infection context.

5.3 Understanding the kinase-independent activity of SteC in infection

The kinase-independent actin polymerisation capacity of SteC in infection is mechanistically unexplained. Without catalytic capacity, SteC may recruit actin directly or indirectly to the site of *Salmonella* microcolonies in the cell, perform an adaptor function mediating the regulation of actin polymerising machinery in the cell, or perform an adaptor function for a host enzyme itself a regulator of actin polymerisation. This section looks at the localisation of SteC in infected host cells and describes assessment of SteC interaction with actin and phospholipids.

5.3.1 The cellular localisation of SteC in infection

To understand how SteC may lead to the polymerisation of actin, even when not enzymatically active, cell localisation of SteC in the cell during infection was explored with the same *Salmonella* strains: *Salmonella* WT, *ΔsteC*, *ΔsteC, ΔsteC pWSK29* empty, *ΔsteC psteC*, *ΔsteC psteC K256H, ΔsteC psteC S379A, ΔsteC psteC S379D* and *ΔsteC psteC 1-429* (Figures 5.9 and 5.10). SteC has been reported to localise to *Salmonella*-induced filaments (Sifs) (Poh et al., 2008), which are membrane protrusions from the *Salmonella* containing vacuole (SCV) that develop 5-8 hours after *Salmonella* infection of the host cell (Knuff & Finlay, 2017). No signal was seen in the empty vector control (Figure 5.9a). When focusing on the microcolony, with the highest HA signal, the main observation was the presence of SteC-2HA in puncta localised to filaments emanating from bacterial microcolonies (Figure 5.9b). The filament was seen in SteC K256H, S379A and S379D suggesting the localisation was independent of SteC kinase activity (Figure 5.9c-e). The HA signal was associated with actin filaments in most cases, though in some the filament appeared to follow a separate route, suggesting SteC-HA may not directly associate with actin, but rather another filamentous structure, likely Sifs as previously described. However, SteC 1-429 failed to localise to any discernible filament, instead concentrating in puncta around the microcolony (Figure 5.9f).

A second SteC-2HA localisation was observed associated with solitary bacteria within the host cell (Figure 5.10). SteC appeared to associate close to one tip of individual bacilli (Figure 5.10b). This was also seen for SteC K256H and S379D (Figure 5.10c and e). It was less clear for S379A (Figure 5.10d) and was not observed for SteC 1-429 (Figure 5.10f). This observation suggests SteC uses its C-tail to localise to the outside of the SCV of some bacteria, possibly close to the translocation tip of the T3SS, given they tend to be found at one pole of the bacterium.

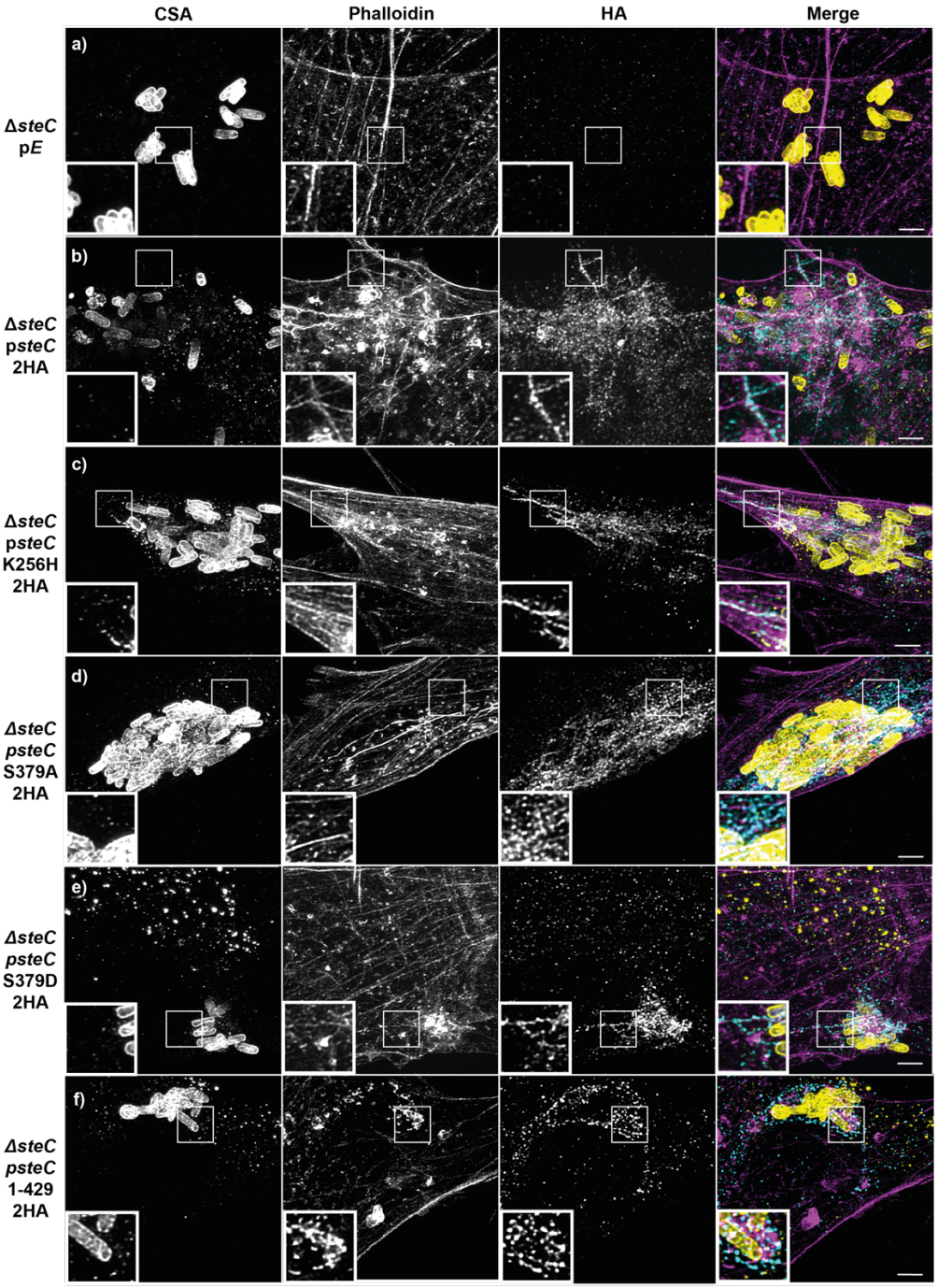


Figure 5. 9: Immunofluorescence of 3T3 cells infected with *Salmonella* strains shows SteC-2HA localised to filamentous structures

Swiss 3T3 mouse fibroblasts (3T3 cells) seeded on cover slips were infected with a) *Salmonella ΔsteC* pWSK29 empty plasmid (pE), b) *ΔsteC psteC* WT, c) *ΔsteC psteC* K256H, d) *ΔsteC psteC* S379A, e) *ΔsteC psteC* S379D, and f) *ΔsteC psteC* 1-429 strains for 8 hours. Fixed cells were stained with phalloidin (represented in pink in the merge), CSA-1 (yellow) and HA (cyan) and cover slips mounted on slides. Super-resolution images were collected with an Olympus CSU-W1 SoRa spinning disk with 60X lens and deconvolved with the Olympus CellSens software under the supervision of Matt Renshaw and Donald Bell (Light Microscopy STP). Scale bars represent 5 µm. Boxed insets are 2-fold magnifications of the indicated regions.

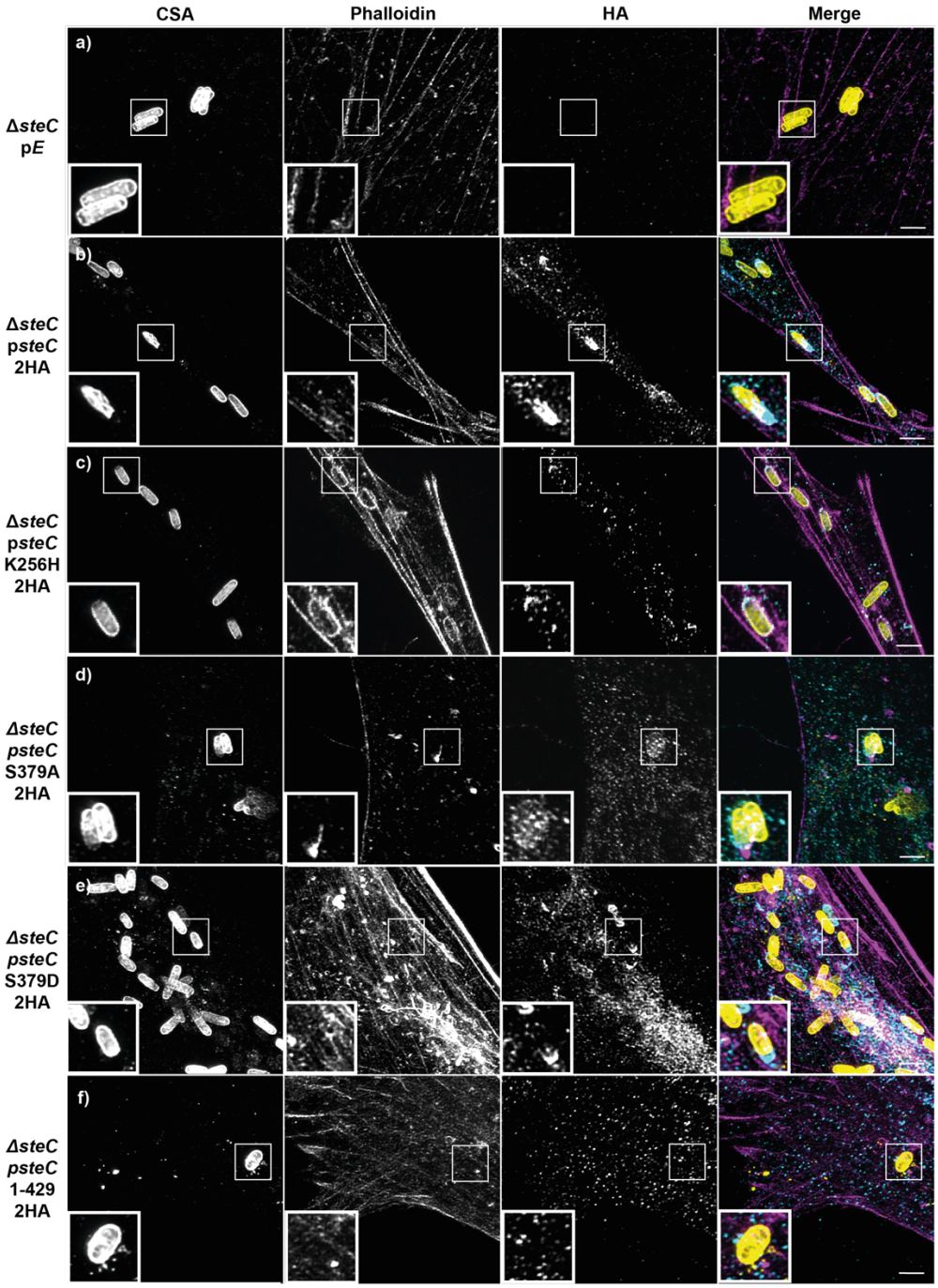


Figure 5. 10: Immunofluorescence of 3T3 cells infected with *Salmonella* strains shows SteC-2HA localised to the tip of solitary bacteria in the host cell

Swiss 3T3 mouse fibroblasts (3T3 cells) seeded on cover slips were infected with a) *Salmonella ΔsteC* pWSK29 empty plasmid (pE), b) *ΔsteC psteC* WT, c) *ΔsteC psteC* K256H, d) *ΔsteC psteC* S379A, e) *ΔsteC psteC* S379D, and f) *ΔsteC psteC* 1-429 strains for 8 hours. Fixed cells were stained with phalloidin (represented in pink in the merge), CSA-1 (yellow) and HA (cyan) and cover slips mounted on slides. Super-resolution images were collected with an Olympus CSU-W1 SoRa spinning disk with 60X lens and deconvolved with the Olympus CellSens software under the supervision of Matt Renshaw and Donald Bell (Light Microscopy STP). Scale bars represent 5 μm. Boxed insets are 2-fold magnifications of the indicated regions.

5.3.2 SteC does not bind to or phosphorylate actin

SteC can trigger actin polymerisation non-enzymatically (Figure 5.8). To test whether SteC binds actin directly, the two proteins were purified and analysed by SEC-MALLS alone and together (Figure 5.11).

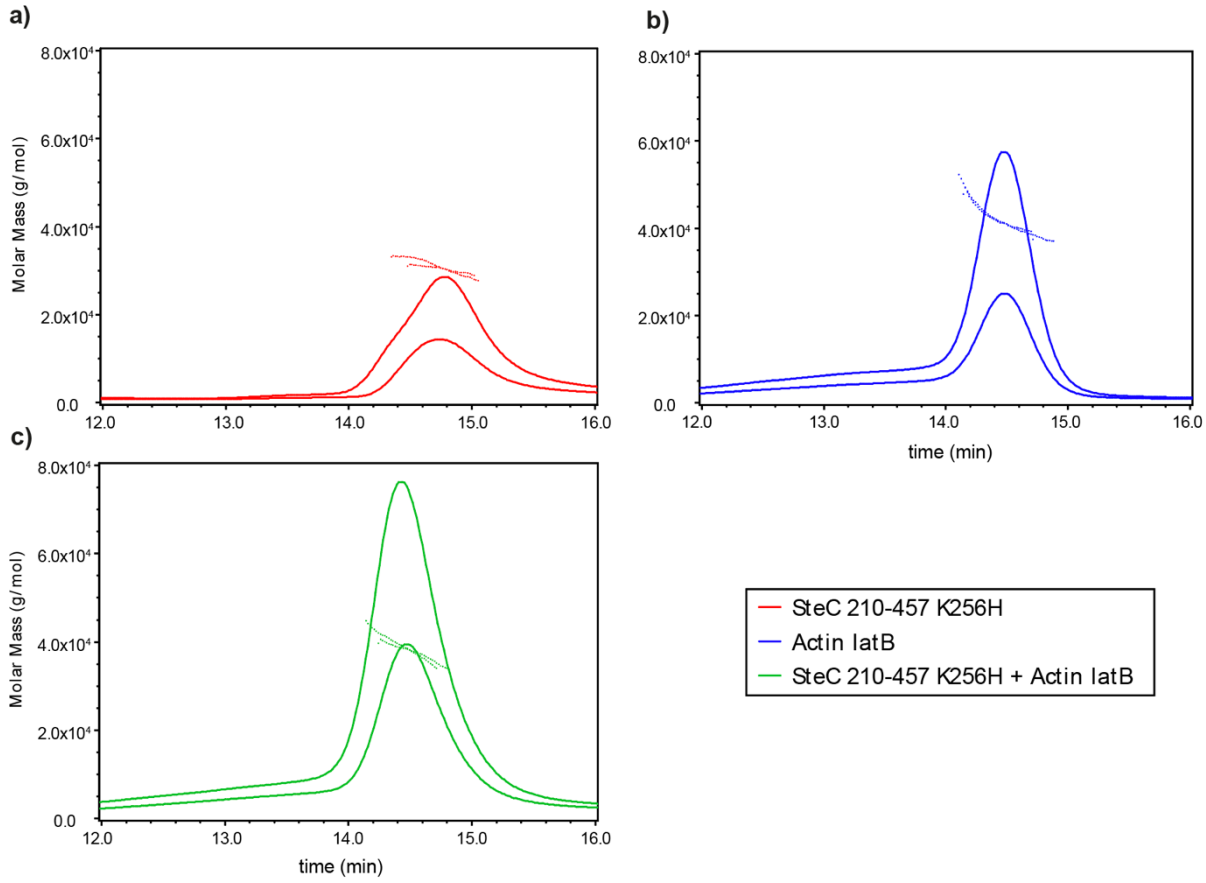


Figure 5. 11: SEC-MALLS of SteC 210-457 and Actin LatB

SteC 210-457 K256H with C-terminal tag was expressed in *E. coli* and purified by standard procedures. Actin, provided as a gift from Aleksandra Michrowska (Triesman laboratory, Francis Crick Institute), was purified from rabbit skeletal muscle and stabilised to prevent polymerisation with latrunculin B (LatB). 100 µL of each protein at 26.4 µM (higher curve) or 14 µM (lower curve) alone or mixed, were loaded onto the SEC column. Each full line represents a size exclusion peak at the corresponding protein concentration. Each dotted line represents the corresponding estimated molar mass of the eluting species (y axis). Experiment performed with supervision from Ian Taylor (Francis Crick Institute).

SteC elutes in a single peak (full line) at approximately 30 kDa (dotted line), i.e. as a monomer at both concentrations (Figure 5.11a), as previously shown (Section 3.2.4). Actin LatB elutes at approximately 40 kDa, with an upward trend in in the molar mass estimate suggesting a tendency to form higher order oligomers (Figure 5.11b). SteC and actin LatB eluted at similar times. When combined, the estimated molar mass was 35 kDa, i.e. the mean of the two monomers, rather than 70 kDa which would have suggested association of the two molecules

in solution a 1:1 heterodimer. This means no discernible interaction was seen between SteC 210-457 K256H and Actin LatB by this method, suggesting that this portion of SteC does not interact directly with Actin LatB in solution at these concentrations.

The bacterial effector in *Yersinia* spp, YpkA, requires binding to actin via its C-terminal 20 amino acids to become active (Juris et al., 2000). To test if SteC activity may be modulated by actin in a similar way, SteC 210-457 kinase activity towards FMNL1 peptide was assessed with and without actin LatB (Figure 5.12).

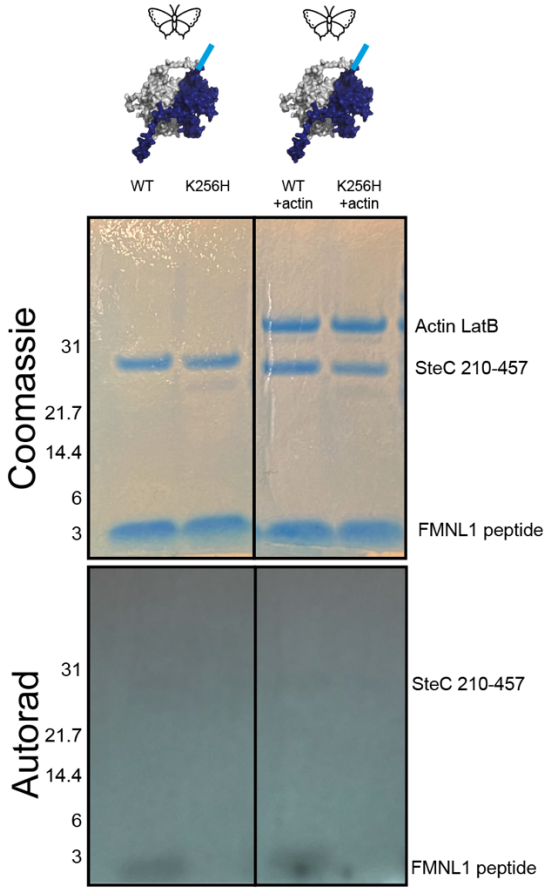


Figure 5. 12: SteC 210-457 does not demonstrate kinase activity towards Actin LatB

SteC 210-457 WT and K256H with C-tag expressed in Sf9 cells and purified by standard procedures was combined at 5 µM with FMNL1 peptide at 100 µM alone (left) or with Actin LatB at 5 µM (right), with 100 µM ATP spiked with 20 kBq [γ-P^{32}]ATP. The mixtures were incubated for 30 min at 30 °C and analysed on SDS-PAGE followed by Coomassie staining and autoradioagraphy to detect P^{32}. Autoradiography exposure time was 4 hr.

SteC 210-457 demonstrated the expected phosphorylation signal with FMNL1 peptide in a K256-dependent manner. The phosphorylation of FMNL1 peptide was unchanged on addition of Actin LatB, suggesting there is no potentiation or inhibition of SteC by Actin LatB. Furthermore, no phosphorylation signal was seen for Actin LatB, revealing that it is not a direct

substrate of SteC. Together, I conclude that the SteC kinase domain does not interact directly with actin, nor directly mediate its phosphorylation and the activity of SteC itself is not directly modulated by actin *in vitro.*

5.3.3 SteC 1-193 binds to phospholipids

The AF predicted structure of SteC 1-193 demonstrates homology to membrane-localising domains by FoldSeek (Table 3.1). Furthermore, SteC can localise in infection to structures that resemble Sifs (Poh et al., 2008) with our own findings corroborating this (Figure 5.9) and revealing a possible enhanced association to the pole of the SCV (Figure 5.10).

Membrane binding can be estimated from a PDB file using Membrane Optimal Docking Area (MODA). This computational tool was developed with known membrane binding domains but does not rely on homology but rather the surface characteristics of a protein based on the PDB file and has identified novel protein folds that bind membrane (Kufareva et al., 2014). Therefore, the AF-predicted structure of SteC was used to screen for regions of the protein that may interact with membrane.

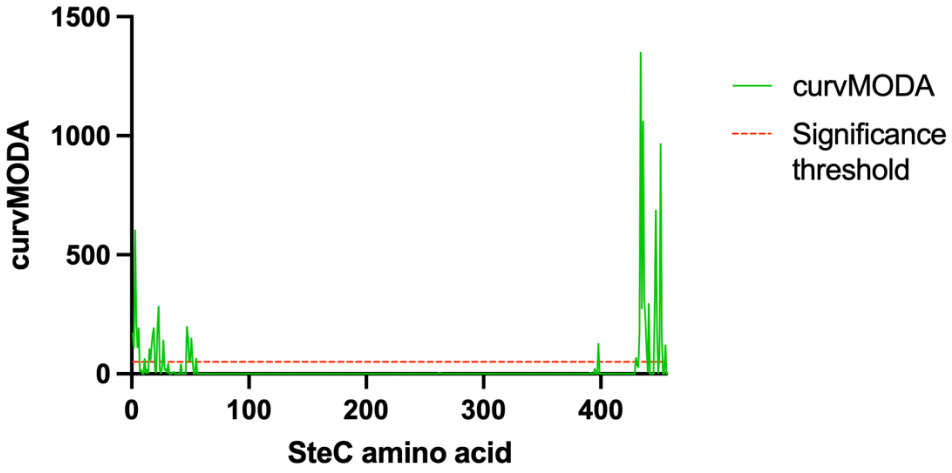


Figure 5. 13: Predicting SteC-membrane association using the AF prediction and MODA

Membrane optimal docking area (MODA) is a validated bioinformatic tool for predicting likely membrane-associated regions of protein structures, requiring a PDB file for analysis. AF predicted structure of SteC was submitted to the platform (available here). CurvMODA score is plotted per SteC residue. Values above 50 are significant.

SteC amino acids 1-6, 11, 13, 15-19, 22-23, 27, 47-52, 55, 398, 430-439, 441, 446-448, 450-452 and 455 demonstrate significance on analysis of the predicted structure of SteC by MODA (Figure 5.13), suggesting the N-terminus of the protein and the C-tail may associate with membranes.

In light of these observations, the lipid binding activity of SteC 1-457, SteC 1-193 and SteC 210-457 was tested using phosphoinositide (PIP) strip membranes (ThermoFisher). These strips contain several PIPs so can identify proteins with PIP recognition domains as well as determination of any preference (Figure 5.14).

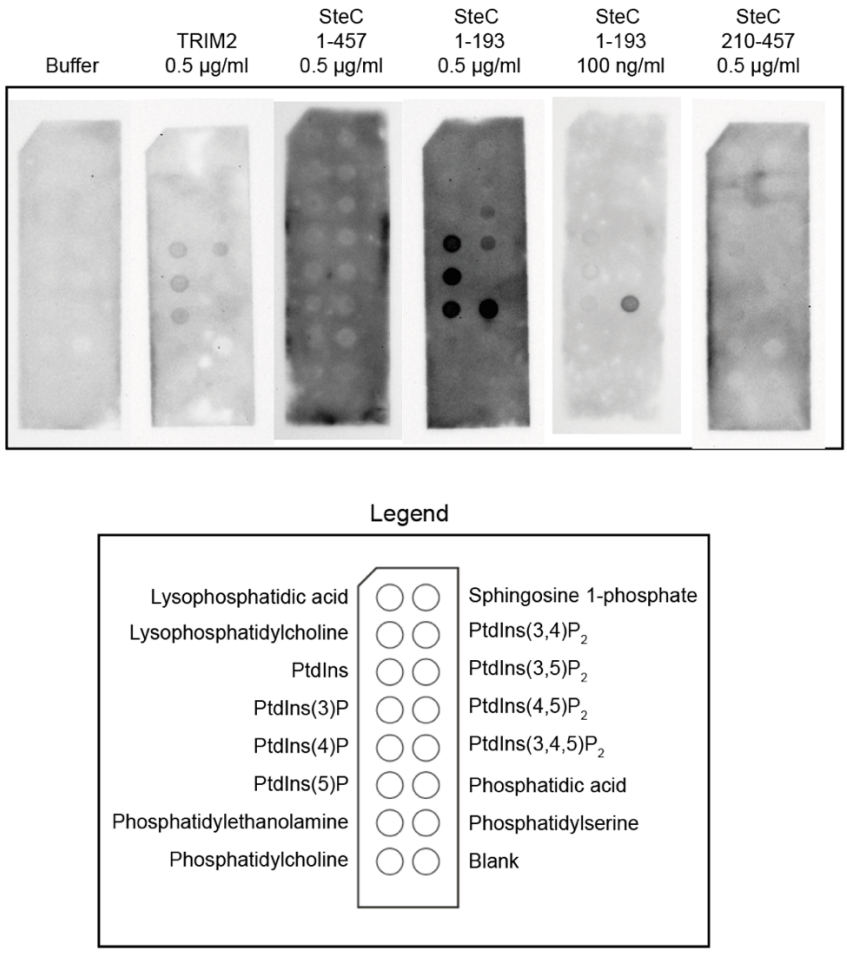


Figure 5. 14: SteC 1-193 binds phosphatidic acid

PIP strips were blocked in PBS with 3% fat-free BSA then exposed to recombinantly expressed and purified N-terminal TwinStrep-tagged SteC 1-457, SteC 1-193 C-terminal His or SteC 210-457 C-terminal His in PBS at the concentration stated. Anti-His antibody was used to detect His-tagged protein bound to the strip. C-terminally His6-tagged TRIM2, available in the lab and prepared as previously reported (Esposito et al., 2022) was used as a positive control given work in the lab has established it binds to PtdInsP and PtdIns(3,5)P_2 (unpublished). Imaging with chemiluminescence was performed. Spots contain 100 picomoles of lipid. Data representative of 2 repeats. PtdIns, phosphatidylinositol.

The negative control, exposed to PBS with 3% fat free BSA only, showed no interaction with any lipid as expected. TRIM2, analysed as a positive control, bound to phosphatidylinositol monophosphate (PtdInsP) and phosphatidylinositol 4,5-bisphosphate (PtdIns(4,5)P_2), as

previously demonstrated in the laboratory (unpublished data, Rittinger laboratory). SteC 1-457 and SteC 210-457 did not bind lipids at 0.5 µg/ml whereas SteC 1-193 at this concentration demonstrated a strong signal for phosphoinositol monophosphates, i.e. PtdIns(3)P, PtdIns(4)P, and PtdIns(5)P, and phosphatidic acid and a weaker signal for phosphoinositol diphosphates i.e. PtdIns(3,4)P, PtdIns(3,5)P and PtdIns(4,5)P. When reducing the concentration of SteC 1-193 to 100 ng/ml, the binding was specific for phosphatidic acid. The common feature of these targets is a single phosphate head (Figure 5.15).

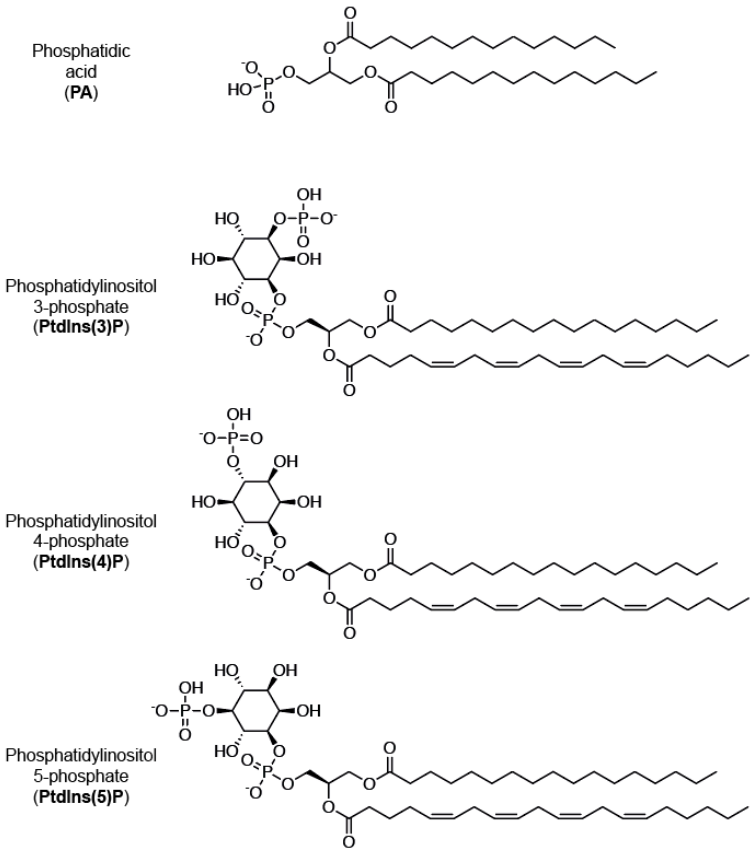


Figure 5. 15: Skeletal formulae of the phospholipids that SteC 1-193 binds to

Demonstration of the phospholipids with polar head (left) and fatty acid tail (right). Produced in Chemdraw with the help of Ben Mortishire-Smith (Rittinger laboratory).

Together with the cell localisation data, the PIP strip assay results suggest that SteC may localise to membrane structures associated with *Salmonella* microcolonies within the host cell through direct interaction with the N-terminal regulatory domain. The relevance of the PtdIns specificity is unclear at this point.

5.3.4 Co-immunoprecipitation and mass spectrometry to identify stable binding partners of SteC in infected cells

The kinase-independent effects of SteC in infected cells could be due to its interaction with host proteins. Furthermore, some bacterial kinases, such as OspG (Pruneda et al., 2014), require a stabilising binding partner to become an active kinase (see Section 1.3.7 for a summary). To identify any putative stable interaction partners of SteC during infection, and to identify if SteC is phosphorylated at S379 in an infection context and in an SteC kinase-activity dependent manner, 3T3 cells infected with *Salmonella ΔsteC* strains expressing SteC-2HA, SteC K256H-2HA or *Salmonella ΔsteE* expressing SteE-2HA (positive control) and uninfected 3T3 cells (negative control) were compared. The cells were lysed 8 hours post invasion and the HA-tagged proteins immunoprecipitated prior to gentle washing, denaturing and digestion for detection of putative binding partners by mass spectrometry.

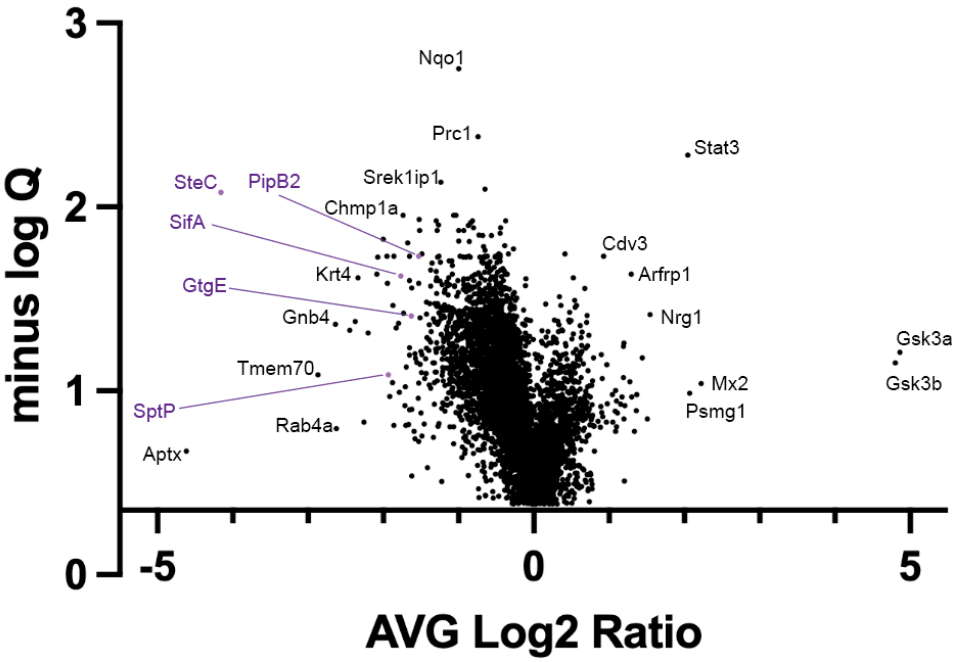


Figure 5. 16: Significantly enriched SteC-2HA binding partners from infected 3T3 cells

3T3 cells were infected with mutant *Salmonella* strains expressing SteC-2HA or SteE-2HA. After 8 hours, cells were lysed and co-immunoprecipitation was performed with HA-beads. Proteins were denatured with SDS and digested with trypsin prior to mass spectrometry analysis. Peptides were searched against *Mus musculus* and *Salmonella* 14028s strain proteomes. Differentially identified binding partners are displayed in a volcano plot, with those more significantly detected in the SteC condition on the left, and those more significantly detected in the SteE condition on the right. Identified *Salmonella* proteins are highlighted in purple. Data is the sum of 3 biological repeats. Experiment performed with Briac Lemetais, MRes student, Thurston laboratory and samples analysed by MS by Tania Auchynnikava (Proteomics STP).

Comparing SteC WT and SteE conditions to identify host binding partners, the previously reported binding partners of SteE, namely Gsk3α, Gsk3β and STAT3 (Panagi et al., 2020) were identified (Figure 5.16). Many more significant binding partners were identified for SteC. Reassuringly, SteC was identified as significantly enriched, along with several host binding partners (black) and other *Salmonella* effectors (purple). None of the previously reported binding partners or substrates of SteC were outliers in the plot. However, of the 3652 proteins significantly enriched in the SteC condition, Fmnl2, Fmnl3, Myl1, Myl4, Myl6, Myl6b, Myl9, Myl10, Myl12a and Myl12b were detected.

Finally, the data were searched for phosphopeptides. Due to no enrichment step for phosphopeptides, this did not provide the platform to perform a rigorous phosphoproteomics experiment of potential substrates. Of 63,776 peptides identified in the experiment across the four conditions, 117 were phosphopeptides from 83 proteins (Supplementary Table 1).



Figure 5. 17: Peptide coverage of SteC-2HA and SteC K256H-2HA from co-immunoprecipitation mass spectrometry during infection

Peptides identified by mass spectrometry of SteC-2HA and SteC K256H-2HA immunoprecipitated following the same infection experiment as Figure 5.16. The N-terminal regulatory domain (1-193) is highlighted in light pink and the C-terminal kinase domain (210-457) is highlighted in light blue. Total peptide coverages were 68% (SteC) and 65% (SteC K256H). Peptides identified in both SteC and SteC K256H conditions (teal), only identified in the SteC condition (yellow) and phosphopeptide identified in both conditions (red) are demonstrated. Phosphosite S379 is in red and bold. No unphosphorylated form of this phosphopeptide was identified. Arrows represent one continuous peptide.

Given the enrichment of SteC and SteC K256H with HA beads, the peptides included a coverage of SteC itself of 68% in the SteC condition, and 65% in the SteC K256H condition, only lacking the peptides containing K256 as expected given the mutation (Figure 5.17). The equivalent coverages for the kinase domain were higher: 90% and 84%, respectively. The only phosphopeptide of SteC identified was SV**pS**LATR, and this peptide was not detected in the unphosphorylated form. This phosphosite was identified in both the SteC and SteC K256H conditions, confirming the presence of S379 in a biologically relevant model and independent of SteC kinase activity, suggesting host kinase phosphorylation.

5.4 MYL12A phosphorylation in infection

MYL12A was identified as a substrate of SteC in 3T3 cell lysate, and was confirmed to be phosphorylated on T19 and/or S20 by SteC *in vitro* (Section 4.4). MYL12A has recently been independently described to be a substrate of SteC implicated in macrophage migration of infected cells (Dai et al., 2024). I therefore aimed to explore the SteC-dependent nature of MYL12A phosphorylation in infection, testing whether S379 was important for phosphorylation of this substrate.

5.4.1 Assessing phospho-MYL12A antibodies

To test the relative specificity of available antibodies against T19/S20-phosphorylated MYL12A two were purchased: AB126739 (Abcam, monoclonal), used in the relevant publication (Dai et al., 2024) and AB2480 (Abcam, polyclonal). First, to test the specificity of the antibodies for phosphorylated T19-S20 over unphosphorylated MYL12A, recombinantly expressed MYL12A incubated with or without SteC 210-457 was tested by western blot (Figure 5.18).

Loading of the protein was similar between conditions as judged by Coomassie staining. Western blot with ab2480 detected a strong MYL12A band for the condition containing MYL12A WT and SteC 210-457 WT, suggesting that MYL12A has been phosphorylated by SteC 210-457 at T19 and/or S20. There was a weak signal in the MYL12A WT-only and MYL12A WT + SteC 210-457 K256H condition, suggesting this antibody detects either phosphorylation of MYL12A at T19A/S20A that is mediated by a contaminating kinase, or that it can detect unphosphorylated MYL12A but is unable to detect the site if MYL12A T19A/S20A mutations are present. Western blotting of the same samples with ab126739 demonstrated a strong band specific for the sample containing both MYL12A WT and SteC 210-457 WT. With this antibody, no signal was detected for the MYL12A WT-only sample or MYL12A WT + SteC 210-457 K256H conditions, suggesting ab126739 may be more specific for SteC induced phosphorylation of MYL12A. Therefore, ab126739 was used going forward.

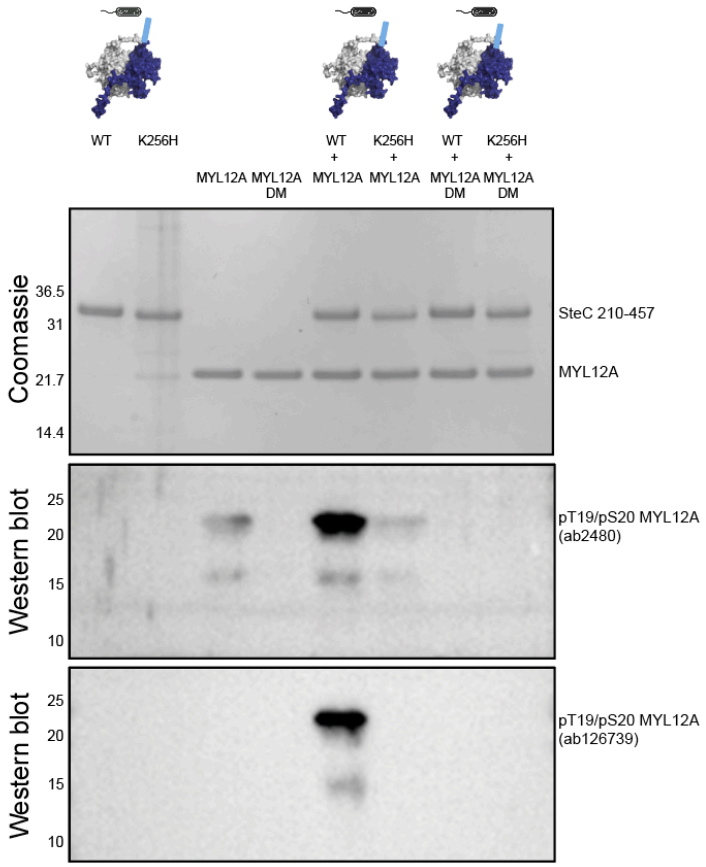


Figure 5. 18: Testing anti-phosphoMYL12A antibodies with recombinantly expressed phosphorylated and unphosphorylated protein

SteC 210-457 WT, SteC 210-457 K256H, MYL12A WT and MYL12A T19A/S19A (DM) all with N-terminal TwinStrep tag were expressed in *E. coli* and purified by standard procedures. 5 µM of each protein alone and SteC-MYL12A pairs were mixed, 100 µM ATP was added, and the mixtures were incubated for 30 min at 30 °C. The samples were analysed by SDS-PAGE before Coomassie staining and Western Immunoblotting with each of the pT19/S20-MYL12A antibodies. Data representative of 2 independent experiments.

5.4.2 No SteC-dependent MYL12A phosphorylation was detected during infection

To test if MYL12A was phosphorylated by SteC during infection, 3T3 cells and RAW 264.7 macrophages (the cell line reported to demonstrate a MYL12A-dependent macrophage migration phenotype) were infected with *Salmonella* WT, *ΔsteC*, *ΔinvG* (a SPI-1 mutant), *ΔssaV* (a mutant unable to translocate effectors via the SPI-2 T3SS) and compared to uninfected cells. Then western blots of cell lysates were performed to analyse protein phosphorylation (Figure 5.19).

In 3T3 cells, a faint band of phosphorylated MYL12A was seen in all conditions at all timepoints, regardless of *Salmonella* strain, and including in uninfected cells. This was below

the level expected for MYL12A, which is ~20 kDa, possibly representing a truncation or degradation band. In RAW 264.7 macrophages, phosphorylated MYL12A was absent at 1 hr post infection but became visible at 17 hours, though this was also visible in the uninfected cells. Based on this result, there is no evidence that MYL12A is phosphorylated during *Salmonella* infection, with no effect of SteC on this process. This could be due to the relatively small quantity of MYL12A phosphorylated locally to the microcolony.

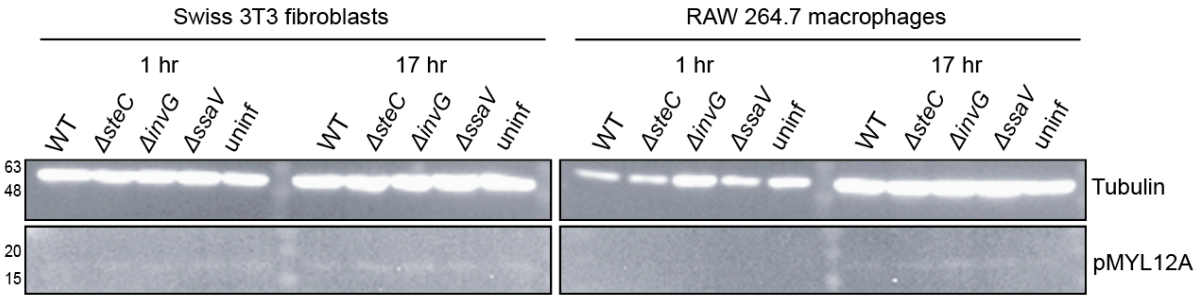


Figure 5. 19: Detecting phosphorylated MYL12A in mammalian cells infected with *Salmonella* mutants

3T3 cells and RAW 264.7 macrophages were infected with the *Salmonella* WT, *ΔsteC*, *ΔinvG*, *ΔssaV* strains and compared to uninfected cells. After 1 and 17 hours, whole cell lysates were prepared prior to western blotting for tubulin and Phospho-MYL12A (pT19/S20, ab126739). Blot is inverted to make faint bands clearer. Data representative of 2 repeats.

To check if MYL12A is expressed in the cell lines used in these experiments, and whether the method of infection used in the literature may be responsible for the phosphorylation of MYL12A, the same experiment was performed, but this time preparing the subculture of *Salmonella* as described by Dai et al. (2024), in the RAW cells for 1 hr as they did.

MYL12A was detected in all conditions, possibly slightly less in the uninfected cells. However, minimal phospho-MYL12A was detected, and this appeared to be independent of the *Salmonella* strain. Tubulin demonstrates similar loading and DnaK was present in all conditions except the uninfected, as would be expected (Figure 5.20).

In conclusion, antibodies used to detect T19/S20-phosphorylated MYL12A were specific for the phosphorylated recombinantly expressed protein, but did not detect SteC-dependent MYL12A phosphorylation in infected 3T3 cells or macrophages either using the Thurston laboratory infection method or following the method reported in the recent publication (Dai et al., 2024). Therefore, I was unable to corroborate the reported importance of MYL12A as a substrate of SteC mediating its cellular effects.

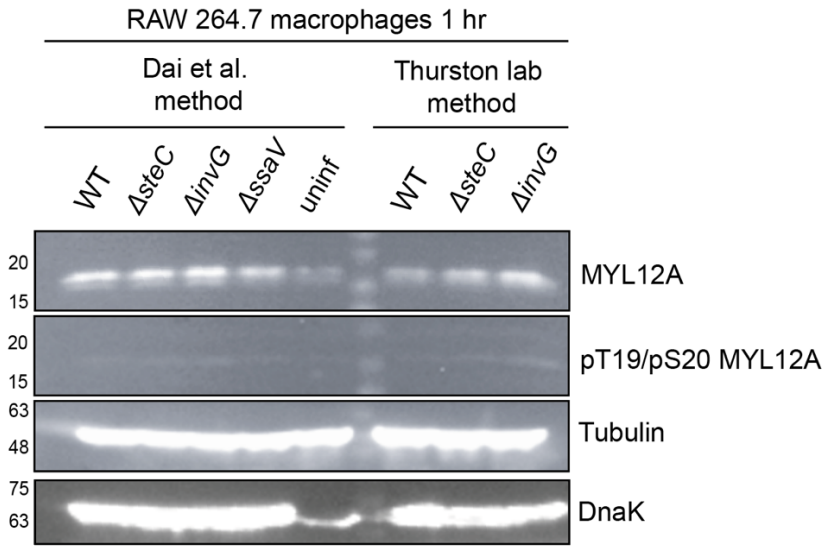


Figure 5. 20: Immunoblot analysis of infected RAW cells demonstrates no phosphorylation of MYL12A

RAW 264.7 macrophages were infected with the *Salmonella* WT, *ΔsteC*, *ΔinvG*, *ΔssaV* strains by growing a subculture to OD_{600} of 0.5 (Dai et al method) or for 3.5 hr (OD_{600} of approximately 1.2 – 1.5, Thurston laboratory method) and compared to uninfected cells. After 1 hr, whole cell lysates were prepared prior to western blotting for tubulin, DnaK, MYL12A and Phospho-MYL12A (pT19/S20, ab126739). Experiment performed with the help of Xiujun Yu, Thurston Laboratory. Blot is inverted to make faint bands clearer. Data representative of 2 repeats.

Conclusion

This chapter explored the effects of SteC in mammalian cells. Firstly, in HEK 293ET cells ectopically expressing GFP-SteC WT and mutants, actin polymerisation followed a kinase-dependent pattern. WT SteC and the S379D mutant demonstrated an ability to polymerise actin while SteC K256H and SteC S379A lacked this ability, validating in cells the model generated *in vitro* (Chapter 4): that S379 is an essential amino acid and phosphorylation at this site has an activating role for SteC function. Furthermore, this method allowed identification of another 7 amino acids that are likely to be catalytically essential. In this model, loss of the C-tail had no effect on the capacity of SteC to polymerise actin.

In infection, a more complex pattern was seen. 3T3s cells infected with various *Salmonella* strains expressing SteC and mutants demonstrated dense focal actin polymerisation at 8 hours post invasion in a distinctly SteC kinase-dependent manner, again validating the importance of S379 phosphorylation in infection. Immunoprecipitation mass spectrometry identified S379 as the only SteC phosphosite in infected 3T3 cells both in the WT and K256H-complemented strains, confirming its presence in a biologically relevant context and its phosphorylation by a host kinase.

Another SteC-dependent phenotype became apparent that was kinase-independent, suggesting that SteC has at least one kinase-independent mechanism of polymerising actin alongside its phosphorylation of a host substrate. Even more unexpectedly, SteC's C-tail (430-457) was essential for both actin polymerisation phenotypes in infect cells, contrasting with a) SteC 210-429 retaining *in vitro* kinase activity, and b) SteC 1-429 retaining actin polymerisation capacity when ectopic expressed in HEK 293ET cells. This apparent contradiction is discussed further in Chapter 6.

Thirdly, a mechanism for SteC's kinase-independent activity was explored. Immunofluoresence microscopy showed that SteC-2HA localised to the pole of *Salmonella*, presumably to the SCV given the method of permeabilisation left the SCV intact. Given what is known of the distribution of the T3SS needle – to one pole of *Salmonella* within the SCV – one explanation for this phenomenon is that SteC, on translocation, localises to the outer surface of the SCV close to the pore formed by the T3SS. In light of this finding, and the weak but consistent homology between AF predicted structure of SteC and human proteins that have lipid-binding or membrane-localising domains, SteC binding to phospholipids was assessed. Lipid binding assays using PIP strips revealed that recombinant SteC 1-193 (N-terminal regulatory domain) bound to phosphatidic acid. Phosphatidic acid is a key component

of cell membranes and many intracellular organelles (Zhukovsky et al., 2019), and so is likely to exist in the SCV as well, given its formation through phagocytosis.

Finally, a recently published substrate, MYL12A, was not found to be phosphorylated in infection in two cell lines despite using antibodies able to detect recombinantly expressed and phosphorylated MYL12A. This could be due to below-detection levels but means the mechanism by which SteC caused actin polymerisation remains undetermined. The results of Chapters 3, 4 and this chapter are incorporated into a model of SteC's activity in Chapter 6.

6. Discussion

The mechanisms by which bacterial effector proteins manipulate their host are still poorly understood. Work presented in this thesis attempted to reveal the mechanism by which the *Salmonella* effector kinase, SteC, functions to manipulate the host environment.

In Chapter 3, I describe the analysis of the AlphaFold2-predicted structure to identify residues and motifs of SteC that may perform canonical or non-canonical roles in its kinase activity. Structural and biophysical studies revealed that SteC has a well-folded N-terminal regulatory domain, and that the C-terminal kinase domain has a globular N-lobe and flexible and minimal C-lobe. The kinase domain is monomeric in solution. It phosphorylates FMNL1, a substrate recently reported in the literature, at 4 sites and binds to a peptide region containing 3 of these sites. The full-length kinase domain protein, including the flexible 'C-tail' (αα 430-457) binds with highest affinity. In Chapter 4, I interrogated the kinase activity of SteC and showed that SteC is phosphorylated on S379 and this is important for SteC activity towards substrates. SteC is phosphorylated at S379 by the eukaryotic kinase STK38. It can also auto-phosphorylate at S379, though less efficiently, and my experiments showed that this occurs in cis, not in trans. The C-tail was important for auto- and substrate phosphorylation, though the kinase domain without the C-tail retained some activity. Furthermore, I identified a new substrate, MYL12A, which has also very recently been reported in the literature (Dai et al., 2024) for which I determined an estimated K_m of SteC towards MYL12A of 2.4 μM. In Chapter 5, I investigated the role of SteC in mammalian cells through ectopic expression of SteC and mutants in HEK 293ET cells to identify further key residues for its function. In addition, I engineered *Salmonella* strains expressing mutants of SteC to perform infection experiments, which revealed that kinase-active SteC produces a dense focal actin polymerisation pattern, while kinase-inactive SteC polymerises actin in a capsular pattern around the *Salmonella* containing vacuole (SCV). SteC localised to filament structures associated with actin – possibly *Salmonella*-induced filaments (Sifs) – and to the tip of the SCV suggesting SteC associates with membrane structures. In support of these observations, I was able to show that the N-terminal domain of SteC binds to phosphatidic acid *in vitro*.

By synthesising the data from Chapters 3, 4 and 5, the literature and novel predictive tools AlphaFold multimer and AlphaFold 3, I developed a novel model of the mechanism of SteC's activity, which is described in three parts: how SteC becomes phospho-activated, how kinase-substrate binding occurs, and how SteC has its actin polymerising effects in host cells. Alongside these, I present avenues for further study.

6.1 A model of SteC activation through phosphorylation

6.1.1 What are the key catalytic residues for SteC kinase activity?

The canonical kinase catalytic motifs are highlighted on PKA in Figure 6.1a. SteC shares some of these whilst also having novel motifs. Through analysing the sequence and predicted structure, key residues were proposed (Chapter 3), then tested through mutagenesis of recombinantly expressed protein (Chapter 4) and ectopic expression in mammalian cells (Section 5.1). Confirmed key residues are summarised on the AlphaFold2-predicted structure in Figure 6.1b. K256 in the predicted β3, previously demonstrated to be important for kinase function (Poh et al., 2008), was required for kinase activity and mutation to His eliminated binding to Mant-AMPPNP whilst not significantly disrupting secondary structure, confirming it is the catalytic lysine. S379, predicted to be in or near a putative activation loop, was shown to become phosphorylated and this modification was essential for kinase activity. The DFG-like 'FFVDGD' motif, which canonically coordinates magnesium ions in the catalytic cleft, was also found to be important: each of the aspartic acid residues D362 and D364 were essential for activity of ectopically expressed SteC in human cells, though whether they coordinate divalent cations has not been tested. R280, in the predicted αC helix, was also essential for SteC activity, suggesting it could be the basic residue at this site that stabilises a phosphorylated activation loop. Instead of an HRD motif, SteC contains an H(343)D(344) motif, with both residues essential for SteC's activity when expressed in mammalian cells and D344 essential for *in vitro* kinase activity. This motif is thought to mediate γ-phosphate transfer onto the substrate. I propose D344 as an appropriate alternative residue to K256 to abrogate kinase activity given it is not likely to disrupt the secondary structure of the kinase (Reinhardt & Leonard, 2023). Interestingly, each of these residues are conserved in the 4 bacterial homologs of SteC identified by BLAST analysis of the amino acid sequence (Figure 3.3).

When compared to canonical kinases, SteC's catalytic residues are just as interesting for what they are lacking. The αC glutamate is a highly conserved and catalytically essential residue in eukaryotic kinases (Soudah et al., 2023). E272, E261 and D279 are all acidic residues close to the AF-predicted αC helix. Interestingly, the SteC E272A mutant demonstrated no discernible difference in activity to WT *in vitro* (Figure 5.3), while demonstrating a partial but significant reduction in actin polymerisation capacity when ectopically expressed in HEK 293ET cells (Figure 5.2). This may be due to E272A mutation affecting SteC catalytic activity against the true substrate that is not detected against FMNL1 or MYL12A, or else that it has a non-enzymatic role. Meanwhile, E261 and D279 appear dispensable when mutated to alanine in ectopically expressed cells, suggesting that SteC does not require an αC glutamate or equivalent for activity. This should be tested with further studies.

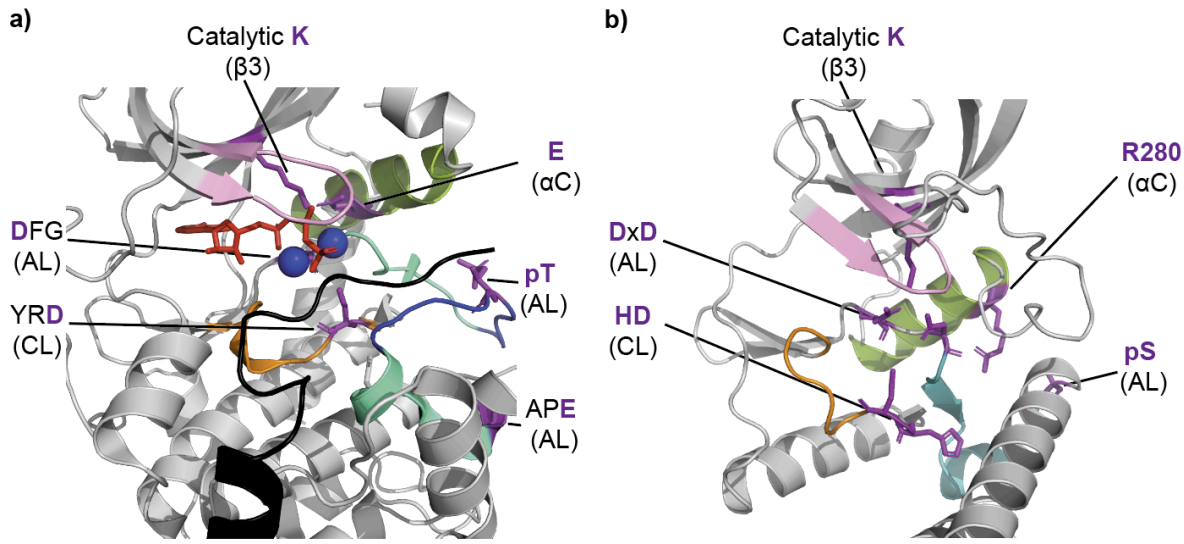


Figure 6. 1: Highlighting SteC's experimentally determined key residues compared to PKA

a) Catalytic site of the PKA crystal structure (PDB: 1ATP) with key catalytic residues highlighted (purple), glycine-rich loop (pink), αC helix (green), catalytic loop (CL) (orange), activation loop (AL) (turquoise)

b) Experimentally confirmed key residues highlighted in purple in the catalytic site of SteC AlphaFold2 prediction. S379 is 4.5 Å from R280.

Furthermore, from interrogation of the sequence and AF-predicted structure, it is apparent that other canonical kinase features are lacking. SteC does not have an APE motif at the end of the activation loop preceding αF. This motif is thought to anchor the activation loop structurally (Gógl et al., 2019), so its absence may explain the relative flexibility of the C-lobe of SteC compared to canonical kinases. Furthermore, it has a proposed role in coordinating homodimerisation for trans-autophosphorylation (Marcotte et al., 2017), which based on my data does not occur for SteC. The APE motif also acts as a core contact in crystal structures between the kinase and substrate or upstream kinases for trans-phosphorylation (Gógl et al., 2019). An 'APE' can be found towards the C-terminal end of the C-tail (Figure 3.6), and it is interesting to consider that when the C-tail folds back, it may bring the APE motif into a position to facilitate substrate binding (an explanation for the importance of the C-tail for substrate binding and activity) or upstream kinase docking. This could be tested by mutagenesis.

Another canonical feature lacking in SteC is the αD helix, which is usually a central component of the hydrophobic C-spine and necessary for ATP binding (Taylor & Kornev, 2011). How ATP binding occurs in its absence would require further study.

Finally, the lack of an R in the HD motif would usually denote a kinase that does not require phosphoactivation, yet SteC does. The role of the R would be to combine with a K close to the DFG motif (also lacking from SteC), to create a basic pocket to house the phosphorylated activation loop residue, thereby stabilising the activation loop in the active position. This makes the HD motif more intriguing, and begs the question: how does S379 phosphorylation promote activity?

6.1.2 How is S379 phosphorylated?

Using mass spectrometry, S379 was identified as phosphorylated and functionally important (Section 4.1). Based on analysis of *E. coli* recombinantly expressed protein by mass spectrometry, SteC can autophosphorylate at S379 with the help of its C-tail. However, only a portion of SteC is phosphorylated at S379 and incubation of the protein with ATP did not increase the proportion that was phosphorylated, suggesting SteC autophosphorylates in an intermediate state during expression and loses this capacity once fully folded. The SteC autophosphorylation signal in kinase assays appears similar for the WT and S379D mutant (Figure 4.3), suggesting that S379 phosphorylation is not a contributor to this signal, rather autophosphorylation occurs at other residues. This proposed model is supported by the site of S379 in the predicted structure being distant from the catalytic site, suggesting the need for a significant change in conformation for SteC to be able to autophosphorylate at that site (Figure 6.1b). This intermediate structural state has been referred to as a ‘prone to autophosphorylation’ conformation in eukaryotic kinases, where it may require additional binding partners, or chaperones.

Many kinases trans-autophosphorylate, which requires dimerisation, and this is proposed as a mechanism of kinase regulation since dimerisation may be dependent on kinase concentration or the presence of other binding partners. My data show that the SteC kinase domain does not dimerise (Section 3.2) and does not autophosphorylate in trans (Figure 4.9), meaning SteC must autophosphorylate in cis, as has been described for a small number of eukaryotic kinases (Beenstock et al., 2016; Lochhead et al., 2005).

Despite being, in principle, able to autophosphorylate, SteC is more efficiently phosphorylated by a eukaryotic kinase. Three findings point towards host kinase phosphorylation at S379 being the predominant mode of activation. Firstly, the proportion of phosphorylated S379 in SteC recombinantly expressed in Sf9 is much higher than when expressed in *E. coli*, and this occurs even in the SteC K256H mutant which strongly suggests a eukaryotic kinase can efficiently phosphorylate at that site (Figure 4.8), though this does not rule out a host factor also facilitating auto-phosphorylation. Secondly, the activity of Sf9-expressed SteC was

higher, relative to *E. coli* expressed proteins. The mechanism by which S379 phosphorylation activates kinase activity will be discussed below (Section 6.1.3).

Thirdly, SteC 210-429 WT and S379D bind the ATP analogue Mant-AMPPNP with affinities of 5.4 μM and 11.8 μM respectively (Figure 4.6), while S379A showed barely detectable binding and a binding affinity could not be estimated, suggesting that S379 phosphorylation leads to a conformational change in the kinase domain to promote ATP binding. However, this argument has two caveats: the cellular ATP concentration, estimated to be 4.4 mM (Greiner & Glonek, 2021) should be able to overcome the low affinity given binding was tested at a Mant-AMPPNP concentration of 100 μM. Furthermore, as mentioned, the "prone-to-autophosphorylation" state could be transient during the expression process and able to bind ATP in the unphosphorylated state, unlike the fully folded version. Prone-to-autophosphorylation states are less structurally stable so less amenable to crystallisation, meaning their mechanism remains poorly understood (Soudah et al., 2023).

The dual mechanism of SteC activation at S379, either requiring an intermediate state during expression (with or without a bacterial or host chaperone) or host kinase could present a evolved two-tier regulatory mechanism where *Salmonella* is able to pre-activate in certain conditions hence translocating the active kinase, or else translocate the inactive kinase that is able to perform its kinase-independent functions (See Sections 5.2) until the appropriate host kinase is present to activate its kinase activity. Either way, it is likely that SteC is under exquisite control for its catalytic activity. Complicating matters further, SteC is unlikely to be folded in the bacteria but translocated in a semi-folded state through the T3SS, possibly with a bacterial chaperone. Whether this semi-folded state is a prone-to-phosphorylation intermediate or autophosphorylation is only achieved after translocation, with or without a host chaperone, is unclear. The diameter of the T3SS is 2-2.5 nm (Lombardi et al., 2019) which is on the limit, according to the AlphaFold2 prediction of SteC, of being able to accommodate the fully folded ATP-binding N-lobe (with minimum diameter of approximated 2.5 x 2.5 nm) alongside a more flexible C-lobe.

The question then arises which host kinase is responsible for activation of SteC. Kinases that recognise a given serine/threonine motif (-5 to +4) can be searched using a recently published PhosphoSitePlus serine-threonine kinase library (J. Johnson et al., 2023). Using the motif surrounding S379, "QQRSV**S**LATR", as the search term, I identified PHKG2 as the top hit, with a site percentile score of 98.7%. PHKG2 is in the CAMK group with preference for motifs containing basic residues in the -5 to +4 region, which aligns well with R in the -3 and +4 positions in relation to S379. Experimentally, I demonstrated that the mammalian kinase

STK38 (NDR1), which was an SteC binding partner in a previous co-immunoprecipitation mass spectrometry experiment (Walch et al., 2021), was able to phosphorylate SteC at S379. STK38 ranks 72/303 in the serine-threonine kinase library approach, with a score of 87.8%. It is an AGC kinase and has a basophilic recognition motif in the -1 to -5 positions, again fulfilled by the -3R in this case.

The -2S and +3T either side of S379 present the possibility that a priming kinase phosphorylates close to S379, changing the chemical characteristics of the motif to attract an acidophilic kinase for second phosphorylation event at S379 itself. S377 and T382 were two of a further nine phosphosites identified as phosphorylated in SteC in a previous study (Walch et al., 2021), so could theoretically be important. However, further phosphosites were not detected in our mass spectrometry experiments either from recombinantly expressed protein from *E. coli* or Sf9 (Section 4.1.1) or from infection (Section 5.3.4), suggesting S379 is the only biologically relevant phosphosite. Studying sequence similarity with the bacterial homologs (Figure 3.3) reveals that S379 is conserved, as well as the -3R, +2A, +3T, and other Arg in the -1 to -6 positions. This hints at a possible conserved recognition motif for an activating host kinase (Figure 6.2). Mutagenesis focusing on these residues combined with the techniques employed above would help refute or confirm this theory.



Figure 6. 2: Sequence logo comparing the S379 region of SteC with 4 homologs

S379 is conserved in the first four non-*Salmonella* bacterial homologs of SteC identified by BLAST (Figure 3.3) from *Yokenella regensburgei, Cedecea neteri, Sodalis praecaptivus* and *Erwinia mallotivora*. The 6 amino acids either side of S379 were compared using WebLogo to produce a frequency plot (https://weblogo.berkeley.edu/logo.cgi).

6.1.3 How does S379 phosphorylation promote SteC-ATP binding?

Based on crystal structures of kinases in the active conformation, four intra-molecular interactions between the activation loop phosphate and surrounding residues have been described that could lead to allosteric activation of a kinase and could therefore explain the activation of SteC following S379 phosphorylation. The phosphate can interact with a) the R from the HRD motif, usually present in phospho-activated kinases due to its stabilising effect on the phosphorylated activation loop (L. Johnson et al., 1996), which has a number of possible downstream allosteric effects (Nolen et al., 2004), b) a basic residue in β9 that leads

to the alignment of β6 and β9, referred to as a β6-β9 zip (Zanela et al., 2023); or c) a positively charged residue in αC leading to anchoring of the N- and C-lobes (Nolen et al., 2004). Alternatively, activation loop phosphorylation leads to d) a disorder-to-order transition in the loop itself that facilitates substrate binding, though the active conformation can significantly vary (Nolen et al., 2004). I will discuss the relevance of these mechanisms to SteC, in reference to the AF-predicted structure of SteC alongside PKA in Figure 6.1.

SteC lacks an R from the HRD motif, instead having an "HD" motif in its catalytic loop. The H from this motif, also being positively charged, could hypothetically perform a similar role to the canonical R. However, S379 is ~10 Å from H343 in the predicted structure, making this less likely. The β9 of SteC, which is predicted to be structured and in the active 'zip' formation with β6, lacks a basic residue for interaction with phosphorylated S379, again making this an unlikely allosteric contact site with pS379. The positively charged residue in αC, R280, is just 4.5 Å from S379 (Figure 6.1), so may provide the key allosteric interaction site for S379 to stabilise the N- and C-lobes for activity. Alternatively, SteC's activation loop itself may gain secondary structure through phosphorylation. CD spectroscopy revealed no detectable difference in secondary structure of the recombinantly expressed SteC WT, K256H, S379A and S379D. However, this technique is not suitable for detecting smaller changes in the overall structure and a high-resolution structure is required to provide molecular insight into the structural changes induced by phosphorylation of S379. Therefore, how S379 phosphorylation, despite its distance from the ATP binding pocket, can so dramatically increase ATP-binding affinity remains unclear.

6.1.4 Modelling S379 phosphorylation, ATP and magnesium binding using AlphaFold3

AlphaFold3 is a recently published machine learning algorithm that has been developed to predict the structural effects of post-translational modifications, nucleic acid and lipid binding and small molecule binding (Abramson et al., 2024), though was not available for the majority of my PhD. It enables modelling of ATP binding, magnesium binding and phosphorylation of amino acid residues, providing the opportunity to explore the *in silico* effect of adding these ligands and modifications to SteC's predicted structure to explore allosteric changes that could uncover an activation mechanism.

SteC's full sequence was submitted to the AlphaFold3 online server (https://alphafoldserver.com/) selecting a stoichiometry of one protein molecule to one ATP molecule and two magnesium ions, either in the wild-type protein form or with phosphorylation at S379. Unexpectedly, the top-ranked prediction for wild-type and phosphorylated SteC kinase domain overlaid with a RMSD of 0.266, i.e. there was negligible difference between

the unphosphorylated and phosphorylated forms. Furthermore, the predicted activation loop and orientation of the experimentally determined key residues were unchanged (Figure 6.3a).

Of note in these predicted structures is the pincer formation of the two aspartic acid residues within the FFVDGD (DFG-like) motif (Figure 6.3a). This is reminiscent of the DxD motif essential for the catalytic activity of bacterial and mammalian glycosyltransferases, that coordinates Mn^{2+}-binding motif in a pincer formation (Esposito et al., 2018). SteC may have adopted this approach to divalent cation coordination in its kinase active site, rather than the canonical DFG. This requires experimental validation. Structural alignment of the AF2 and two AF3 predictions of SteC 210-429 showed negligible differences (Figure 6.3b).

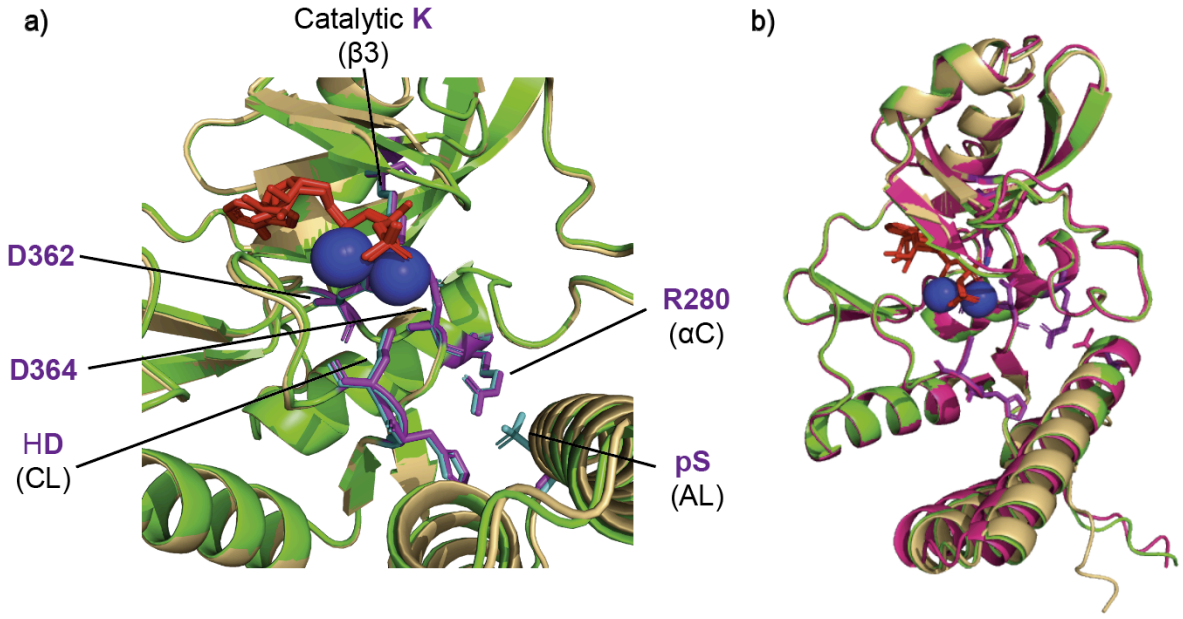


Figure 6. 3: Predicting the effect of S379 phosphorylation using AlphaFold3

a) Top-ranked AF3 predictions SteC 210-429 WT (green) and pS379 (sand) each with ATP (red) and magnesium (blue) were aligned with super command in Pymol. Experimentally determined key catalytic residues of WT and pS379 are highlighted in purple and teal, respectively.

b) Superposition of the two AF3 predictions and the AF2 published prediction (pink).

Isomorphic Labs, a subsidiary company of Alphabet Inc. (formerly Google), has come under criticism for not publishing its code for AlphaFold3 open-source (Wankowicz et al., 2024), and for having weaknesses in predictive accuracy for binding of certain interactions, especially protein-RNA interactions (Desai et al., 2024). The AF3 algorithm may also not perform well at predicting conformational changes of a protein structure on modification or binding of a cofactor or ligand (compared to the unmodified form). Kinases are recognised to undergo allosteric conformational changes on phosphorylation (as described above) and ATP binding (Meharena et al., 2013), represented in a diversity of states in the PDB. However, because

these PDB structures have been incorporated into the training data, a prediction of the apo-structure would potentially predict the protein structure as if these common features were present, rather than specifically predict the conformation in their absence. This challenge is already being addressed with algorithms designed to predict a protein in different conformational states (Bryant & Noé, 2024). Kinases may provide a useful test case for this, given their abundance in the PDB, the number of inactive and active conformations solved, and their annotation with accompanying biochemical activity data.

6.1.5 Summary

SteC is phosphorylated by a host kinase at S379 to become active. It can also autophosphorylate at S379 though appears unable to do so once in the fully folded form, suggesting it adopts an intermediate 'prone-to-autophosphorylation' state either in *Salmonella* or after translocation into the host cytosol. This second S379 phosphorylation mechanism appears less efficient. Phosphorylated S379 may interact with R280 in the αC helix to stabilise the N- and C-lobes, leading to a substantial increase in affinity for ATP, a tripling in substrate binding affinity, and activation of the kinase's enzymatic activity.

6.2 What is the role of the C-tail?

According to AlphaFold2, the minimal C-lobe of SteC's kinase domain is largely comprised of αE (αα 324-338), catalytic and activation loops (339-379), αF and αG (380-429) and a C-tail (430-457). The C-tail is predicted to be dissociated from the rest of the kinase domain and largely lacking secondary structure.

6.2.1 The role of the C-tail in substrate binding and kinase activity

Work presented here suggests the C-tail does not contribute to the protein's globular structure (Section 3.2) and interacts transiently with the αF in region 382-411 (Figure 4.18). The C-tail is important for catalytic activity. It increases S379 autophosphorylation two-fold compared to the kinase domain lacking the C-tail (Figure 4.16). It improves the affinity of SteC kinase domain to a FMNL1-derived substrate peptide from approximately 10 μM (210-429) to 180 nM, a factor of ~100x. Kinase activity against the recombinant FMNL1 1-385 protein substrate is only demonstrated for the N-terminally tagged full kinase domain 210-457. Either removal of the C-tail or positioning of the tag on the C-terminus of the C-tail prevent activity (Figure 4.15).

AlphaFold Multimer was used to model the SteC kinase domain (210-457) in complex with the FMNL1 K190-K213 24mer peptide. Of 10 models produced, all but the lowest ranked closely align with the AlphaFold2-predicted published structure of SteC, or 'apo-prediction', from residues 210-423. However, all predictions differed in the position of the C-tail. The top 9 models and apo-structure (pink) are shown with the predicted orientation of the peptide (black) (Figure 6.4a). The 1st (green), 5th (sky blue) and 6th (grey)-ranked predictions show the C-tail to form an α-helix from residues 446-456 and pack into the C-lobe next to αF and close to the C-terminus of the peptide, while the others showed the C-tail not interacting with the rest of the kinase domain (Figure 6.4a). Given I have shown the importance of the C-tail for binding to the peptide, the 1st, 5th and 6th models were enlarged in Figure 6.4b. The interaction of the C-tail with αF closely resembles the location of C-tail interaction identified by my HDX-MS data (Figure 4.18). This could explain the improved binding affinity of the full-length kinase domain. The reason for the disagreement between predictions could be due to the inherently dynamic nature of the C-tail.

The orientation of the peptide in the catalytic cleft was similar for all predictions. Interestingly, the peptide appeared sandwiched between the N-lobe and minimal C-lobe in all cases, with possible points of contact with the HDI**NI** motif in the catalytic loop ('HRD'-like motif), the αF and αG below and the C-tail (Figure 6.4b). Interestingly, the αG contacts correspond to the

region identified by HDX-MS as becoming protected in the presence of FMNL1 peptide (Figure 3.22). The FMNL1 peptide region TIKLT from T203 to T207 (two identified phosphosites on the peptide) sit close to the proposed ATP and magnesium binding sites in the AlphaFold3 prediction (Figure 6.3) as would be expected for a substrate primed for transfer of the γ-phosphate of ATP.

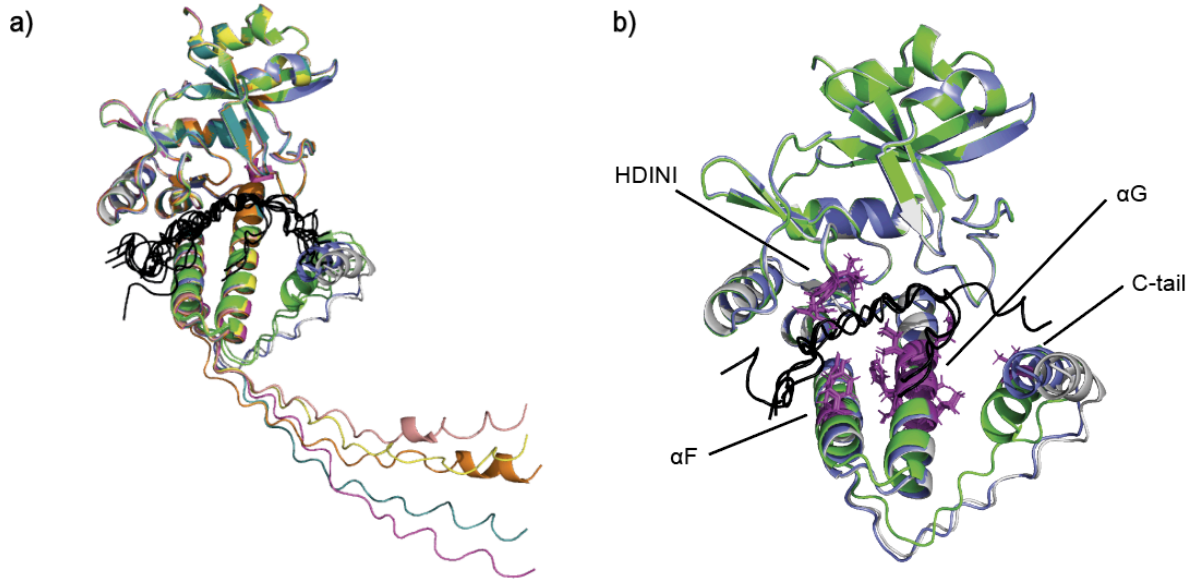


Figure 6. 4: AlphaFold multimer predicts the site of FMNL1 peptide binding in SteC kinase domain

a) Published SteC AF2 prediction and 9 multimer predictions with peptide (black) are aligned using 'super' command in Pymol.

b) Zoom of the three multimer predictions that demonstrate the C-tail folded into the C-lobe and interacting with the substrate peptide. Possible contact amino acids for peptide binding are highlighted.

The models showing FMNL1 pep_K190 docked in the kinase domain catalytic cleft prompts the question: what is SteC's substrate specificity and how is it determined? Just as SteC's upstream kinase may require a -3R for example, SteC may have its own kinase-recognition motif. In the AF multimer prediction, two possible target residues are docked close to the ATP-binding site, T203 and T207. When aligning the residues of the two substrates studied here (4 sites in FMNL1 and 2 sites in MYL12A) (Figure 6.5), no motif is fully conserved though a predilection for basic residues in the minus positions is observed. SteC may also be processive, i.e. able to phosphorylate at multiple sites on the same substrate molecule rather than releasing the substrate after one phosphorylation event before binding to the next, for example recognising T203 in FMNL1 by the -3 R then phosphorylating T207 subsequently, or T19 of MYL12A followed by S20. SteC may acquire substrate specificity by other means too, such as timing of phosphoactivation by a host kinase, host binding partners, or localisation in

the host cell. Furthermore, the substrates represented here may not be biologically relevant. Regardless, SteC does not phosphorylate all proteins tested *in vitro*, so a positional scanning peptide array could be performed to interrogate SteC's kinase-recognition motif experimentally.



Figure 6. 5: Sequence motif analysis of SteC's substrates

Four phosphosites of FMNL1 and two phosphosites of MYL12A and their surrounding 12 amino acids (-6 to +6) were analysed for consensus motifs using WebLogo to produce a frequency plot (https://weblogo.berkeley.edu/logo.cgi).

6.2.2 The role of the C-tail in mammalian cells

In cells, the C-tail was dispensable in polymerising actin when SteC was ectopically expressed in HEK 293ETs (Figure 5.2). However, in infection, the C-tail was vital both for the kinase-dependent actin polymerisation seen and the kinase-independent phenotype (Figure 5.8), demonstrating the importance of the infection model for understanding the complexity of an effector protein's function *in vivo*. Furthermore, when SteC lacks the C-tail it is not seen localising to filamentous structures or the cap of the SCV (Figures 5.9 and 5.10). This is discussed further in Section 6.3.

6.2.3 Challenges with studying recombinantly expressed and purified SteC

The structurally dynamic nature of SteC, that makes the protein so interesting, also makes it hard to study with conventional structural and biophysical techniques such as X-ray crystallography. This complicated the assessment of substrate binding and activity assays. The tight binding data for freshly purified SteC 210-457 was not reproduced with protein that had been frozen and thawed, further evidence that this construct is susceptible to aggregation and/or unfolding. Comparisons between 210-457 and 210-429 in catalytic activity were also difficult because 210-457 was susceptible to aggregation at 30 °C. Ironically, SteC 210-375, comprised largely of the canonical kinase N-lobe and the portion most confidently predicted by AlphaFold2, was not kinase active, lacking key components that we have shown are important for kinase activity (S379), substrate binding (αF and the C-tail) and function during infection in cells (S379 and the C-tail).

6.2.4 A high-resolution structure of SteC

A key aim of my PhD was to solve the crystal structure of SteC (Section 3.3.6). However, crystallisation attempts were unsuccessful likely due to the flexibility of the C-terminal region of the protein. Recently, the crystal structure of an SteC fragment encompassing 202-375 was reported by another group (Dai et al., 2024), (Figure 6.6). Interestingly, this SteC kinase N-lobe fragment crystallised as a dimer (Figure 6.6). This observed dimerisation disagrees with our solution studies by SEC-MALLS, NMR and SAXS which showed that a similar fragment (210-375) exists as a monomer in solution. The authors of that study did not interrogate dimerisation in solution and we speculate that the dimerisation observed in the crystal structure is a crystallisation artifact due to a domain swap. Furthermore, this construct does not possess kinase activity, and so does not provide much insight into the enzymatic mechanism of SteC.

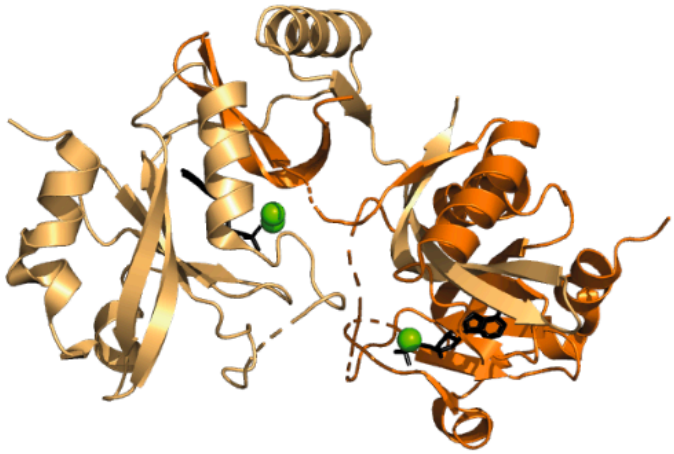


Figure 6. 6: Published crystal structure of SteC 202-375

Structure of SteC 202-375 C276S (Dai et al., 2024) (PDB: 8JBI) crystallised with adenosine monophosphate (AMP) and magnesium. Amino acids 211-261, 263-4 and 272-371 are resolved. The structure shows a dimer, with protomers coloured light orange and orange, AMP in black and magnesium ions in green. The domain swap involved the C-terminal beta-sheet 345-371 appearing to fold onto the opposite domain.

Crystallisation of the entire kinase domain, or full-length protein, may require stabilisation with a binding partner, not just substrate peptide, or may not be possible if SteC does not exist in a stable complex with another protein. This interaction will be discussed further in Section 6.3.

6.3 A model for the roles of SteC in actin polymerisation

SteC is known to create an actin polymerisation phenotype in mammalian cells (Odendall et al., 2012; Poh et al., 2008). This is most commonly seen in 3T3 cells and therefore the physiological importance of this phenotype is not clear. However, a lack of effect in replication studies (Odendall et al., 2012; Poh et al., 2008) including our own data (not shown), leaves actin polymerisation in *Salmonella*-infected 3T3 cells as a useful proxy read-out.

I found that S379 is essential for a dense focal kinase-dependent actin phenotype in *Salmonella*-infected 3T3 cells (Figure 5.8). I observed a second SteC-dependent but kinase-*independent* actin polymerisation phenotype. Furthermore, SteC was found not just associated with filamentous structures, as previously described (Poh et al., 2008), but also at the caps of some SCVs. Finally, I identified MYL12A as a substrate from 3T3 cell lysate that was also very recently reported in the literature (Dai et al., 2024). However, its phosphorylation could not be detected in an SteC-dependent manner from *Salmonella*-infected cells.

6.3.1 SteC has a kinase-independent ability to polymerise actin

While actin polymerisation due to ectopic overexpression of SteC in mammalian cells appears strictly SteC kinase activity-dependent, in infection, there is a parallel kinase-independent mechanism. This discrepancy may be due to technical issues with ectopic expression. The N-terminal GFP tag on ectopically expressed SteC could disrupt its localisation to membranes via the N-terminus, domain-domain interactions or some other aspect of SteC's kinase-independent activity. An experiment using C-terminally tagged ectopic SteC would be required to interrogate this. However, the discrepancy may also be due to the unique cellular environment encountered by SteC in infection, for example the presence of the SCV and Sifs, or the expression of host proteins specific to the infection response. Alternatively, another *Salmonella* effector(s) might support a kinase-independent function of SteC. Either way, my findings suggest SteC promotes actin polymerisation by multiple mechanisms.

The relationship between cell membranes and actin polymerisation in eukaryotic cells is multifaceted. Protein complexes mediate actin-plasma membrane interactions, a necessity for the structural integrity of the cell and for membrane movement (Carlsson, 2018), for example in phagocytosis (Section 1.2.2). Furthermore, plasma 'nanodeformations' promote actin polymerisation through type II interferon and CDC42 signalling, suggesting their interplay has a role in the innate immune response to external threats (Ledoux et al., 2024). Within the cell, membrane tubules important for intracellular trafficking and organelle reshaping are formed and regulated by dynamic actin polymerisation (Allard et al., 2021). This context provides rich

background for an ongoing mystery: what purpose does SteC, which sits at the interface of the SCV, Sifs and F-actin, have in the survival and replication of *Salmonella* in the mammalian host?

The observation that the N-terminal domain of SteC can bind to phospholipid (Figure 5.14), could suggest that SteC plays a scaffolding role to mediate association with phospholipid structures via its N-terminal domain and with other host factors via its kinase domain. The phospholipid preference of the N-terminal domain, primarily phosphatidic acid (PA), is interesting. PA is found in cell membranes. It only accounts for a small proportion of the total phospholipid content but is a key precursor for several glycerophopsholipids as well as having its own signalling functions (Kooijman & Burger, 2009). SteC interaction with PA is therefore consistent with its localisation to SCV by microscopy, given that the SCV is formed from the host cell membrane.

The kinase independent phenotype observed upon overexpression of SteC was not seen in *Salmonella* infection assays with complemented SteC 1-429, suggesting that N-terminal lipid binding is not sufficient for its kinase-independent activity, and highlighting the pivotal role of its C-tail in binding to host factors, not just in kinase activity itself. In fact, the C-tail is also predicted to bind membrane according to the predictive tool Membrane Optimal Docking Area (Figure 5.13), though PIP strip analysis did not show SteC 210-457 (the kinase domain including the C-tail) to specifically bind a phospholipid target (Figure 5.14). At least 54 of the 518 human protein kinases contain at least one lipid-binding module, including C1, C2, FERM, PH and PX domains (Lemmon, 2008), demonstrating the importance of lipid's role in cell localisation of eukaryotic protein kinases. Structural homology searching of the AF prediction of SteC's N-terminal domain shows that it has low structural homology to three such domains, C2, FERM and PH (Table 3.1). Interestingly, the main class of human serine/threonine kinases containing lipid-binding modules are AGC kinases (of which STK38 is one), that characteristically have a C-terminal extension that wraps around the N-lobe of the catalytic domain as a mechanism of kinase regulation (Leonard & Hurley, 2011). Perhaps SteC's C-tail regulates its kinase activity in a lipid-dependent manner.

Also, though SteC's kinase domain did not binding G-actin according to SEC-MALLS (Figure 5.11), it may bind to F-actin. SteC may also enact its non-enzymatic function through binding to FMNL1, which is associated with membranes (Kühn et al., 2015), as an adaptor to promote its actin polymerisation potential, rather than or as well as phosphorylating it as a substrate. The kinase-independent SteC-dependent actin phenotype also raises the question of whether SteC can itself nucleate or propagate F-actin chains from G-actin. *In vitro* actin polymerisation

assays could explore some of these questions further. Also, given the relative stability of the N-terminal domain or crystallisation trials with a soluble phosphatidic acid ligand may help understand some of SteC's mechanisms of action.

Finally, though SteC most closely resembles protein kinases, its biologically relevant role may be as a lipid kinase. The relatively well characterised eukaryotic phosphatidylinositol 3-kinases can resemble their protein kinase counterparts, sharing N- and C-lobes, a DFG motif and an activation loop (M. Zhang et al., 2020). Whether SteC phosphorylates phospholipids directly or modulates their metabolism indirectly, lipidomics may help to reveal insights into SteC's function in cells where protein substrate analysis has so far failed.

6.3.2 Evidence for substrates of SteC that mediate actin polymerisation

None of the substrates of SteC described in the literature has been found to be essential for SteC's actin polymerisation effects in transfected or infected cells (Odendall et al., 2012; Walch et al., 2021). Kinase substrates remain difficult to identify due to the low affinity of substrate binding, meaning using the kinase as bait in immunoprecipitation does not reliably identify them (St Louis et al., 2023). Furthermore, confirming biochemical kinase activity *in vitro* is not always relevant *in vivo* because kinase recognition motifs can be non-specific and kinases may exist as part of multi-protein complexes or in a restricted range of subcellular location, cell type and/or cell state (St Louis et al., 2023). MYL12A, a substrate I identified in parallel with another group (Dai et al., 2024), appeared promising due to the strong SteC-dependent phosphorylation signal observed *in vitro*, with a calculated Km of SteC 210-457 towards MYL12A of 2.4 µM in the conditions tested, which is at the low (higher activity) end of the expected range for an active enzyme (Stryer et al., 2002). I then attempted to corroborate the recently published findings that link the kinase activity of SteC towards MYL12A with a phenotype of infected RAW macrophages. Using antibodies that I showed were able to detect recombinant expressed and T19/S20-phosphorylated MYL12A, I failed to identify MYL12A becoming phosphorylated in 3T3 cells or RAW cells in either an infection-dependent, SPI-1 dependent, T3SS-dependent or SteC-dependent manner. The recent report demonstrated SteC-dependent phosphorylation of MYL12A in cells infected with *Salmonella* for 1 hour. This seemed to contradict the fact that SteC is translocated through the SPI-2-encoded T3SS, normally evident with delayed kinetics compared to effectors delivered by the SPI-1-encoded T3SS (Geddes et al., 2005). Regardless, even when using the conditions reported in the recent publication, I was unable to demonstrate SteC-dependent phosphorylation of MYL12A.

6.4 Summary

The findings described in this thesis, combined with data from the literature, can be put together into a model of SteC activity (Figure 6.7). SteC is translated in the *Salmonella* bacterium (yellow) and then translocated into the host cytosol (grey) via the type 3 secretion system (T3SS, orange), most likely in an unphosphorylated inactive state (purple). It is associated with the *Salmonella*-containing vacuole (SCV, thin grey line) at one cap, possibly close to the outer pore of the T3SS, as well as Sifs. Here it can mediate the polymerisation of actin in a capsular pattern around the SCV in a kinase-independent manner and requiring the C-tail. Alternatively, on phosphorylation at S379 by a host kinase (green), SteC becomes active (red) and able to trigger dense focal actin polymerisation associated with the SCV inside the cell. The substrates of SteC kinase activity, the non-kinase mechanism of actin polymerisation and the biological role of SteC in *Salmonella* survival and replication remain unknown. Furthermore, there are still big gaps in our understanding of what role it plays in the *Salmonella* lifecycle.

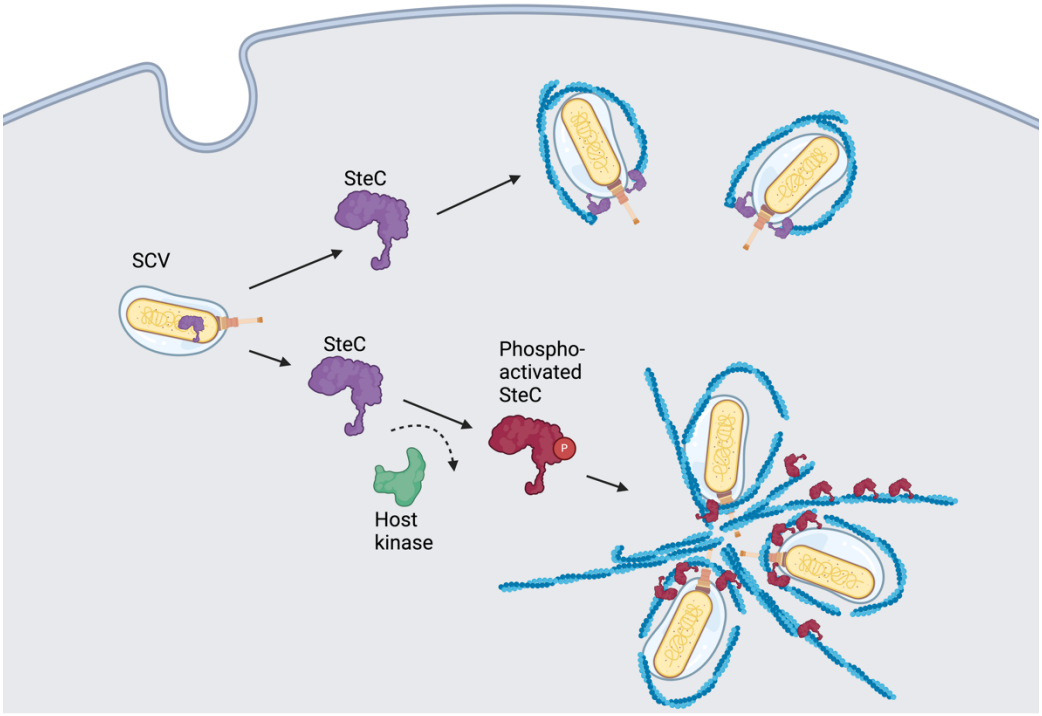


Figure 6. 7: Schematic of SteC's regulation and roles in the host cell

A schematic of the mechanism of SteC-mediated actin polymerisation in host cells, comprising the host cytosol (grey), *Salmonella* (yellow), *Salmonella*-containing vacuole (light blue line), unphosphorylated SteC (purple), phosphorylated SteC (red), host activating kinase (green), F-actin capsular pattern around the SCV or dense focal pattern (blue). Figure made in BioRender.

Conclusion

Studying SteC, the *Salmonella* effector kinase, with a multi-disciplinary approach combining biochemical and biophysical techniques with cell biology and infection studies has provided a rich understanding of the structural and functional features of this effector protein. Just as importantly, this PhD has provided a perspective on the difficulties of gaining structural insight into proteins with intrinsic conformational flexibility. The many questions that remain regarding the complex and multifaceted mechanisms of SteC activity warrant further study.

7. Appendices

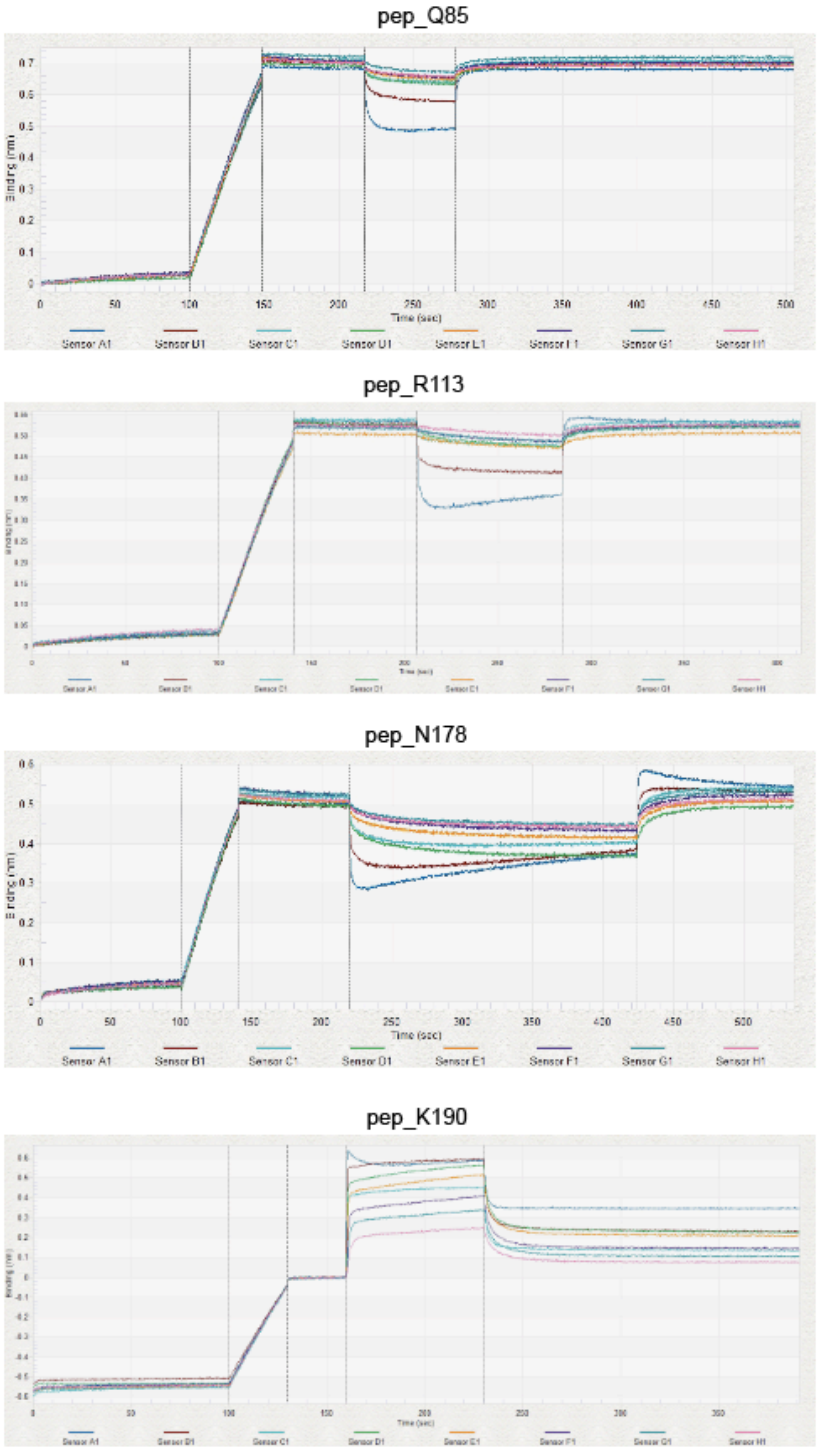


Figure S 1: Binding curves for SteC 210-457 with FMNL1 peptides

Traces from the Octet machine are displayed, representing the binding curves when pep_Q85, pep_R113, pep_N178 and pep_K190 were loaded onto the sensor before exposure to SteC 210-457. Only K190 demonstrates binding with SteC. See Section 3.3.4 for more details.

UniProt ID	Protein	Organism	Peptide sequence and phosphosite	mean uninf	mean SteC	mean K256H	mean SteE ([10] removed)
D0ZIB5	steC	Salmonella 14028s	_SVS[Phospho (STY)]LATR_.1	0	81.8007851	21.3608691	1.74395061
D0ZIB5	steC	Salmonella 14028s	_SVS[Phospho (STY)]LATR_.2	0	131.543498	0	0
A0A087WR50	Fn1	Mus musculus	_KT[Phospho (STY)]DELPQLVT[Phospho (STY)]LPHPNLHGPEILDVPSTVQK_.4	0	106.616945	117.465429	72.0267715
A0A087WR50	Fn1	Mus musculus	_KT[Phospho (STY)]DELPQLVT[Phospho (STY)]LPHPNLHGPEILDVPSTVQK_.5	0	19.0423762	56.3862852	27.2212143
A0A0A6YW06; A0A0A6YXC8; E9QKQ9; E9QLZ9; J3QNM3; Q03173	Enah	Mus musculus	_TNTMNGSKS[Phospho (STY)]PVISRPK_.3	0	0	4.21635564	0
A0A0J9YVG0; Q61074	Ppm1g	Mus musculus	_NLPPQEQM[Oxidation (M)]IS[Phospho (STY)]ALPDIK_.2	0.33333333	3.00200399	2.8280131	1.10903096
A0A0N5E9G7; G3X8Z6; P35601; Q5U4B1	Rfc1	Mus musculus	_QKS[Phospho (STY)]PNKAELFSTAR_.3	0	8.60378218	13.0051797	8.45609093
A0A0R4J0U8; A2RSY6	Trmt1l	Mus musculus	_VT[Phospho (STY)]INDLNENSVTLIQK_.4	53.9593722	30.0033333	25.6132507	70.5781555
A0A0U1RPL0	Atxn2l	Mus musculus	_STSTPTS[Phospho (STY)]PGPR_.2	0	130.087906	194.032476	0
A0A0U1RPL0	Atxn2l	Mus musculus	_GINGGPS[Phospho (STY)]RMSPK_.2	0	0	19.4815979	0
A0A0U1RPL0	Atxn2l	Mus musculus	_GINGGPSRMS[Phospho (STY)]PK_.2	0	27.8000717	33.1558011	22.7574902
A0A0U1RPL0	Atxn2l	Mus musculus	_QGSGRES[Phospho (STY)]PSLVSR_.2	0	27.9666667	0	0
A0A1B0GT26; O55003	Bnip3	Mus musculus	_REVES[Phospho (STY)]ILKK_.3	2070.83302	2352.85066	1551.55247	1703.6391
A0A1L1STC6; Q6ZWR6	Syne1	Mus musculus	_KIQVS[Phospho (STY)]VT[Phospho (STY)]NLEELNVIQSR_.3	720.208923	816.419464	442.242968	589.6931
A0A2R8VHU8; E9Q7U2; Q8CGU1	Calcoco1	Mus musculus	_QELLEY[Phospho (STY)]M[Oxidation (M)]RKLEAR_.3	0	292.39497	362.199001	494.610498
A2AQD5; Q922B9	Itprid2	Mus musculus	_LLGLDEQLRAVRVPS[Phospho (STY)]PFR_.3	26.1193266	36.6543735	39.0960548	23.8641629
A2AQW0	Map3k15	Mus musculus	_Y[Phospho (STY)]LHENQIVHR_.3	27.3432821	40.3166962	54.6072108	53.376049
A2AR02	Ppig	Mus musculus	_ILS[Phospho (STY)]C[Carbamidomethyl (C)]GELIPKSK_.3	3838.44419	1289.06333	551.716064	2129.945
A2BFF5; A2BFF8; A2BFF9; O88487; Q3TPJ8	Dync1i2	Mus musculus	_MS[Phospho (STY)]DKS[Phospho (STY)]DLKAELER_.3	129.309792	132.350422	153.970209	106.918251

A2CEK1; Q8VDY4	Efcab7	Mus musculus	_KET[Phospho (STY)]KPVTEEAQLVHR_.4	70.8210602	40.0763971	54.32	72.3950329
B2RRE2; E9QAX2; K3W4L0; Q9JMH9	Myo18a	Mus musculus	_DMLLAEAFS[Phospho (STY)]LKQQMEEK_.3	16.8915602	3.86936049	8.20810256	10.9001175
B9EKK6	Nxpe3	Mus musculus	_LQEEKPDRVY[Phospho (STY)]FK_.2	5.43830721	89.9729668	191.476487	61.8943644
B9EKK6	Nxpe3	Mus musculus	_LQEEKPDRVY[Phospho (STY)]FK_.3	11.8251114	277.191796	484.951733	283.332472
D3Z5G0	Ifi213	Mus musculus	_KLT[Phospho (STY)]REAEGLVHR_.4	0	1347.41846	1026.08965	1848.96975
D3Z7E5; Q2NL51	Gsk3a	Mus musculus	_ARTSSFAEPGGGGGGGGGGPGGSAS[Phospho (STY)]GPGGTGGGK_.3	0	0.44830251	0	59.8034515
E9PUT0	Zfp983	Mus musculus	_SFLELSHLKRHS[Phospho (STY)]R_.3	0	56.6920827	200.211288	91.1562119
E9Q1Z0	Krt90	Mus musculus	_TAAENEFVVLKKDVDAAY[Phospho (STY)]M[Oxidation (M)]NK_.3	254.322103	126.207123	154.046884	226.247093
E9Q616	Ahnak	Mus musculus	_FS[Phospho (STY)]VPGVK_.1	0	69.6279093	359.831731	0
E9QAQ5	Gsk3b	Mus musculus	_GEPNVSYIC[Carbamidomethyl (C)]S[Phospho (STY)]R_.2	0	0	0	181.765884
E9QN08; P57776; Q80T06	Eef1d	Mus musculus	_ATAPQTQHVS[Phospho (STY)]PMR_.2	0	0	36.5773185	0
F8WJG3; P62996	Tra2b	Mus musculus	_RRS[Phospho (STY)]PSPYYSR_.3	0	7.13021215	0	0
H3BL37; O08784	Tcof1	Mus musculus	_ASAVS[Phospho (STY)]PEKAPMTSK_.2	21.469415	22.62273	16.9415296	25.643383
K4DID3; Q8QZY3	Dppa3	Mus musculus	_KLT[Phospho (STY)]LNPGVKR_.3	821.382026	517.932196	367.835562	609.541199
P10922	H1-0	Mus musculus	_[Acetyl (Protein N-term)]T[Phospho (STY)]ENS[Phospho (STY)]TSAPAAKPK_.2	125.553162	0	34.6950175	0
P11499	Hsp90ab1	Mus musculus	_VILHLKEDQT[Phospho (STY)]EY[Phospho (STY)]LEER_.2	13.9878756	10.2498112	25.1996291	5.11903572
P20152	Vim	Mus musculus	_SLYSSS[Phospho (STY)]PGGAYVTR_.2	301.442535	129.862594	93.454127	161.668274
P21278	Gna11	Mus musculus	_IAT[Phospho (STY)]VGYLPTQQDVLR_.2	0	179.943542	0	0
P23950	Zfp36l1	Mus musculus	_RHS[Phospho (STY)]VTLPSSK_.2	0	7.11303266	5.49205144	0
P23950	Zfp36l1	Mus musculus	_RHS[Phospho (STY)]VTLPSSK_.3	0	4.19500732	0	15.9046335
P26369; Q80XR5	U2af2	Mus musculus	_GAKEEHGGLIRS[Phospho (STY)]PR_.3	144.343719	65.7591953	56.5202204	78.5582619
P35486	Pdha1	Mus musculus	_YHGHS[Phospho (STY)]MSDPGVSYR_.3	0	33.06	10.2525342	0
P35486	Pdha1	Mus musculus	_Y[Phospho (STY)]HGHSMSDPGVSYR_.3	4.35174815	3.50068728	7.52139854	24.7424812
P35486	Pdha1	Mus musculus	_YHGHS[Phospho (STY)]MSDPGVSYR_.2	0	3.48509534	5.8715814	0
P39428	Traf1	Mus musculus	_VVELQQT[Phospho (STY)]LAQK_.2	34.773524	263.234309	625.894697	368.977455

P43274	H1-4	Mus musculus	_[Acetyl (Protein N-term)]SETAPAAPAAPAPAEKT[Phospho (STY)]PVKK_.3	143.924133	43.6581837	42.4972054	116.539042
P43277	H1-3	Mus musculus	_S[Phospho (STY)]ETAPAAPAAPAPVEK_.2	0	0	14.8966242	299.544983
P47963	Rpl13	Mus musculus	_NKS[Phospho (STY)]TESLQANVQR_.2	1.92333333	2.19414918	2.25507593	0
P52480	Pkm	Mus musculus	_EATESFASDPILYRPVAVALDT[Phospho (STY)]K_.3	0	19.2106781	0	35.0989532
P60710	Actb	Mus musculus	_M[Oxidation (M)]QKEIT[Phospho (STY)]ALAPSTMK_.3	183.2544	102.889679	54.7542585	59.6726151
P60710	Actb	Mus musculus	_EKLC[Carbamidomethyl (C)]Y[Phospho (STY)]VALDFEQEMATAASSSSLEK_.4	68.6747824	91.3318977	19.3464686	20.0831757
P62737	Acta2	Mus musculus	_LC[Carbamidomethyl (C)]Y[Phospho (STY)]VALDFENEMATAASSSSLEK_.3	1815.28188	1106.58205	1193.9127	1379.84218
P62754	Rps6	Mus musculus	_LSS[Phospho (STY)]LRASTSKSESSQK_.3	113.65763	101.499977	92.5132179	110.272957
P62754	Rps6	Mus musculus	_LSSLRAS[Phospho (STY)]TSKSESSQK_.3	0	25.2633333	0	0
P68033	Actc1	Mus musculus	_YPIEHGIIT[Phospho (STY)]NWDDMEKIWHHTFYNELR_.5	520.055944	313.510885	0	44.4726372
P97393	Arhgap5	Mus musculus	_IRT[Phospho (STY)]EAS[Phospho (STY)]QIRK_.3	523.248886	775.503902	682.546275	604.25914
P97434	Mprip	Mus musculus	_SKS[Phospho (STY)]NPDFLKK_.2	48.4569194	8.58004761	42.8377622	7.96092319
P97434	Mprip	Mus musculus	_SKS[Phospho (STY)]NPDFLKK_.3	31.3	6.89890544	11.9388924	0
P97434	Mprip	Mus musculus	_S[Phospho (STY)]KSNPDFLKK_.3	0	14.5873782	0	0
P97855	G3bp1	Mus musculus	_STS[Phospho (STY)]PAPADVAPAQEDLR_.2	0	0	40.8365402	0
Q3TXV4; Q921E2	Rab31	Mus musculus	_GSAAAVIVYDITKQDS[Phospho (STY)]FHT[Phospho (STY)]LKK_.3	0	5.9920845	0	0
Q3UL36	Arglu1	Mus musculus	_ASS[Phospho (STY)]PPDRIDIFGR_.2	789.72703	111.663086	72.2577871	220.994114
Q4V9W2	Srek1ip1	Mus musculus	_RS[Phospho (STY)]SPYHSELTK_.2	0	0	10.6113122	0
Q4V9W2	Srek1ip1	Mus musculus	_RSS[Phospho (STY)]PYHSELTK_.3	106.47907	23.6361497	19.3489685	11.4935513
Q4V9W2	Srek1ip1	Mus musculus	_RSS[Phospho (STY)]PYHSELTK_.2	29.2202937	31.5824038	4.96361001	18.9739761
Q4V9W2	Srek1ip1	Mus musculus	_RS[Phospho (STY)]SPYHSELTK_.3	25.3066667	0	0	21.030653
Q4V9W2	Srek1ip1	Mus musculus	_RSS[Phospho (STY)]PYHSELTKK_.3	465.567708	160.9387	10.6458136	0
Q4V9W2	Srek1ip1	Mus musculus	_S[Phospho (STY)]SPYHSELTK_.2	38.9450556	12.1069489	18.9054667	31.2561111
Q4V9W2	Srek1ip1	Mus musculus	_SS[Phospho (STY)]PYHSELTKK_.2	72.6759961	19.2447446	6.13398425	0
Q4V9W2	Srek1ip1	Mus musculus	_S[Phospho (STY)]SPYHSELTKK_.3	0	8.95103645	0	0
Q4V9W2	Srek1ip1	Mus musculus	_SS[Phospho (STY)]PYHSELTKK_.3	10.0350927	6.57314428	0	0
Q4V9W2	Srek1ip1	Mus musculus	_S[Phospho (STY)]SPYHSELTKK_.2	0	0	0	8.95441055

Q5F2E7	Nufip2	Mus musculus	_RNS[Phospho (STY)]AKGC[Carbamidomethyl (C)]ENLNLVQDK_.3	0	4.59019248	4.99733639	6.47767019
Q5SV64	Myh10	Mus musculus	_DHNIPQES[Phospho (STY)]PKPVKPQGELER_.4	860.804179	2842.3931	2144.92511	2704.47736
Q5SWZ5	Mprip	Mus musculus	_SKS[Phospho (STY)]LKEGLTVQER_.3	11.2721113	71.2248866	39.8320732	46.5912027
Q61033	Tmpo	Mus musculus	_LSQSSYQDSESLS[Phospho (STY)]PPRKVPR_.3	140.390289	38.1627655	24.0908496	26.8560162
Q61033	Tmpo	Mus musculus	_SSS[Phospho (STY)]PSSQREFAAR_.2	0	0	13.3477071	0
Q61879	Myh10	Mus musculus	_ALEQQVEEM[Oxidation (M)]RT[Phospho (STY)]QLEELEDELQATEDAK_.3	20.1	31.8879467	51.3607699	46.9813328
Q61879	Myh10	Mus musculus	_KMEEEVLLLEDQNS[Phospho (STY)]K_.2	0	46.9474462	70.2756399	14.9899111
Q61879	Myh10	Mus musculus	_FSKVEDMAELTC[Carbamidomethyl (C)]LNEAS[Phospho (STY)]VLHNLKDR_.4	0	94.26216	81.0524089	86.1011581
Q61879	Myh10	Mus musculus	_VEDMAELT[Phospho (STY)]C[Carbamidomethyl (C)]LNEASVLHNLKDR_.4	718.209066	0	0	0
Q6NZC7	Sec23ip	Mus musculus	_KLS[Phospho (STY)]VGAYVSSVR_.2	0	18.6657257	23.5728124	0
Q6Y7W8	Gigyf2	Mus musculus	_WRPHS[Phospho (STY)]PDGPR_.3	20.441452	96.1528447	95.9535433	40.8598919
Q8BIJ7	Rufy1	Mus musculus	_ITDVLDQKNY[Phospho (STY)]VEELNR_.3	0	51.2613729	32.8487508	50.1815214
Q8BU30	Iars1	Mus musculus	_T[Phospho (STY)]ENAVS[Phospho (STY)]RMQS[Phospho (STY)]VIELGR_.3	19.8513209	47.0455507	57.7774811	54.8037849
Q8BU30	Iars1	Mus musculus	_TENAVSRMQS[Phospho (STY)]VIELGR_.4	0	74.677901	4.40611808	56.1258354
Q8C129	Lnpep	Mus musculus	_VT[Phospho (STY)]FMSAVSSQEK_.2	776.187231	2177.39	1223.42	790.525
Q8C1S0	Med19	Mus musculus	_NRHS[Phospho (STY)]PDHPGMGSSQASSSSSLR_.4	0	20.4920673	17.1025912	49.2840729
Q8C2Q3	Rbm14	Mus musculus	_TRLS[Phospho (STY)]PPR_.2	10.8952624	17.8960609	61.5735283	30.7092171
Q8C2Q3	Rbm14	Mus musculus	_LSESQLSFRRS[Phospho (STY)]PTK_.3	0	28.4066772	0	35.71
Q8CGC7	Eprs1	Mus musculus	_AS[Phospho (STY)]EDLKT[Phospho (STY)]HM[Oxidation (M)]VVSNTLEDFQK_.3	0	6.16545105	0	0
Q8CH72	Trim32	Mus musculus	_LRELT[Phospho (STY)]GELQR_.3	0	0	9.72268073	13.0449533
Q8K363	Ddx18	Mus musculus	_NGTGVLILS[Phospho (STY)]PT[Phospho (STY)]R_.2	34.2767461	305.861547	183.617404	164.964996
Q8R3F9	Tut1	Mus musculus	_IKDS[Phospho (STY)]PLGGVNK_.3	155.442347	45.9273726	88.2838826	83.5476837
Q8VDD5	Myh9	Mus musculus	_LQKDLEGLS[Phospho (STY)]QRLEEK_.4	0	0.33333333	1	0
Q8VDD5	Myh9	Mus musculus	_KGTGDC[Carbamidomethyl (C)]S[Phospho (STY)]DEEVDGKADGADAK_.3	0	8.98010381	5.1577603	0
Q8VEK3	Hnrnpu	Mus musculus	_[Acetyl (Protein N-term)]SSS[Phospho (STY)]PVNVKKLK_.2	0	0	1.4061656	0
Q99M28	Rnps1	Mus musculus	_RFS[Phospho (STY)]PPRR_.3	116.211149	163.821607	94.7541226	110.243106
Q9CX86	Hnrnpa0	Mus musculus	_LFIGGLNVQT[Phospho (STY)]SESGLR_.2	0	76.6613702	62.3419374	43.2730751

Q9CZ44	Nsfl1c	Mus musculus	_RRHS[Phospho (STY)]GQDVHVVLK_.4	0	3.88119614	15.5526413	0
Q9D662	Sec23b	Mus musculus	_AES[Phospho (STY)]EEGPDVLR_.3	107.577204	111.636347	106.514303	70.5962906
Q9D824	Fip1l1	Mus musculus	_ERDHS[Phospho (STY)]PTPSVFNSDEERYR_.4	79.7252604	7.6678435	0	20.3344688
Q9D824	Fip1l1	Mus musculus	_YGRAES[Phospho (STY)]PDLR_.2	0	0	18.2352028	0
Q9D898	Arpc5l	Mus musculus	_YIY[Phospho (STY)]KGFEKPTENSSAVLLQWHEK_.4	0	0	14.7	0
Q9DAW9	Cnn3	Mus musculus	_LT[Phospho (STY)]LQPVDNSTISLQMGTNK_.2	0	0	31.0933333	0
Q9DBR7	Ppp1r12a	Mus musculus	_DKKS[Phospho (STY)]PLIESTANMENNQPQK_.3	0	2.26985534	0	0
Q9EP71	Rai14	Mus musculus	_ISQDADLKT[Phospho (STY)]PTKPK_.3	2.53	0	0	8.0592947
Q9ERG0	Lima1	Mus musculus	_TSSLPESS[Phospho (STY)]PSKTAK_.2	0	0	13.5921809	0
Q9ERG0	Lima1	Mus musculus	_SEAQQPMHPKPLS[Phospho (STY)]PDAR_.3	78.4870602	117.357796	135.323209	68.4768143
Q9ERG0	Lima1	Mus musculus	_SEAQQPM[Oxidation (M)]HPKPLS[Phospho (STY)]PDAR_.3	12.9841373	15.2650522	17.5686935	16.7030597
Q9EST3	Eif4enif1	Mus musculus	_APS[Phospho (STY)]PPMSQVFR_.2	0	12.8581314	0	32.23
Q9ESX5	Dkc1	Mus musculus	_LVKS[Phospho (STY)]QQS[Phospho (STY)]AGKEYVGIVR_.4	142.903503	28.960879	13.1801631	162.970383
Q9QXS1	Plec	Mus musculus	_SSS[Phospho (STY)]VGSSSSYPISSAGPR_.2	203.872416	171.965673	187.949331	398.198029
Q9QXS1	Plec	Mus musculus	_AQLEPVAS[Phospho (STY)]PAKKPK_.2	9.40273412	7.2236865	13.1665273	4.16184841
Q9QYC0	Add1	Mus musculus	_WLHSGRGDDASEEGQNGSSPKS[Phospho (STY)]K_.4	0	61.6815313	0	0
Q9R0L6	Pcm1	Mus musculus	_SNRPVS[Phospho (STY)]ADGNYRPLAK_.3	0	25.6004155	7.28654099	13.3163109
Q9WTI7	Myo1c	Mus musculus	_IIAKGEELLS[Phospho (STY)]PLNLEQAAYAR_.3	0	84.9679311	75.9916255	93.0387726
Q9WTI7	Myo1c	Mus musculus	_S[Phospho (STY)]RVVHQNHGERNFHVFY[Phospho (STY)]QLLEGGEEETLR_.5	14.8600553	153.614206	95.4136683	121.989964

Table S 1: Phosphopeptides identified by co-immunoprecipitation mass spectrometry

Co-immunoprecipitation mass spectrometry was performed as described in Section 5.3.4. Phosphopeptides identified when searched against *Salmonella* Typhimurium serovar 14028s and *Mus musculus* are listed. The scores represent a mean of the total peptide quantity score of three biological repeats for each condition. Due to sample 10 containing very few peptides, suggesting a technical failure, the results for the SteE condition are a mean of 2 biological repeats only.

References

Abramson, J., Adler, J., Dunger, J., Evans, R., Green, T., Pritzel, A., Ronneberger, O., Willmore, L., Ballard, A. J., Bambrick, J., Bodenstein, S. W., Evans, D. A., Hung, C.-C., O'Neill, M., Reiman, D., Tunyasuvunakool, K., Wu, Z., Žemgulytė, A., Arvaniti, E., … Jumper, J. M. (2024). Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature*, *630*(8016), 493–500. https://doi.org/10.1038/s41586-024-07487-w

Achtman, M., Zhou, Z., Charlesworth, J., & Baxter, L. (2022). EnteroBase: Hierarchical clustering of 100 000s of bacterial genomes into species/subspecies and populations. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *377*(1861). https://doi.org/10.1098/rstb.2021.0240

Adams, J. A. (2003). Activation loop phosphorylation and catalysis in protein kinases: Is there functional evidence for the autoinhibitor model? *Biochemistry*, *42*(3), 601–607. https://doi.org/10.1021/bi020617o

Albataineh, M. T., & Kadosh, D. (2016). Regulatory roles of phosphorylation in model and pathogenic fungi. *Medical Mycology*, *54*(4), 333–352. https://doi.org/10.1093/mmy/myv098

Alix, E., Godlee, C., Cerny, O., Blundell, S., Tocci, R., Matthews, S., Liu, M., Pruneda, J. N., Swatek, K. N., Komander, D., Sleap, T., & Holden, D. W. (2020). The Tumour Suppressor TMEM127 Is a Nedd4-Family E3 Ligase Adaptor Required by Salmonella SteD to Ubiquitinate and Degrade MHC Class II Molecules. *Cell Host & Microbe*, *28*(1), 54-68.e7. https://doi.org/https://doi.org/10.1016/j.chom.2020.04.024

Allard, A., Lopes dos Santos, R., & Campillo, C. (2021). Remodelling of membrane tubules by the actin cytoskeleton. *Biology of the Cell*, *113*(8), 329–343. https://doi.org/https://doi.org/10.1111/boc.202000148

Amano, M., Ito, M., Kimura, K., Fukata, Y., Chihara, K., Nakano, T., Matsuura, Y., & Kaibuchi, K. (1996). Phosphorylation and activation of myosin by Rho-associated kinase (Rho-kinase). *Journal of Biological Chemistry*, *271*(34), 20246–20249. https://doi.org/10.1074/jbc.271.34.20246

Amano, M., Nakayama, M., & Kaibuchi, K. (2010). Rho-kinase/ROCK: A key regulator of the cytoskeleton and cell polarity. In *Cytoskeleton* (Vol. 67, Issue 9, pp. 545–554). https://doi.org/10.1002/cm.20472

Ao, T. T., Feasey, N. A., Gordon, M. A., Keddy, K. H., Angulo, F. J., & Crump, J. A. (2015). Global burden of invasive nontyphoidal salmonella disease, 2010. *Emerging Infectious Diseases*, *21*(6), 941–949. https://doi.org/10.3201/eid2106.140999

Ardito, F., Giuliani, M., Perrone, D., Troiano, G., & Muzio, L. Lo. (2017). The crucial role of protein phosphorylation in cell signalingand its use as targeted therapy (Review). In *International Journal of Molecular Medicine* (Vol. 40, Issue 2, pp. 271–280). Spandidos Publications. https://doi.org/10.3892/ijmm.2017.3036

Bakkeren, E., Gül, E., Huisman, J. S., Steiger, Y., Rocker, A., Hardt, W. D., & Diard, M. (2022). Impact of horizontal gene transfer on emergence and stability of cooperative virulence in Salmonella Typhimurium. *Nature Communications*, *13*(1). https://doi.org/10.1038/s41467-022-29597-7

Bamburg, J. R., Minamide, L. S., Wiggan, O., Tahtamouni, L. H., & Kuhn, T. B. (2021). Cofilin and actin dynamics: Multiple modes of regulation and their impacts in neuronal development and degeneration. In *Cells* (Vol. 10, Issue 10). MDPI. https://doi.org/10.3390/cells10102726

Barta, M. L., Dickenson, N. E., Patil, M., Keightley, A., Wyckoff, G. J., Picking, W. D., Picking, W. L., & Geisbrecht, B. V. (2012). The Structures of Coiled-Coil Domains from Type III Secretion System Translocators Reveal Homology to Pore-Forming Toxins. *Journal of Molecular Biology*, *417*(5), 395–405. https://doi.org/https://doi.org/10.1016/j.jmb.2012.01.026

Batool, R., Qamar, Z. H., Salam, R. A., Yousafzai, M. T., Ashorn, P., & Qamar, F. N. (2024). Efficacy of typhoid vaccines against culture-confirmed Salmonella Typhi in typhoid endemic countries: a systematic review and meta-analysis. *The Lancet Global Health*, *12*(4), e589–e598. https://doi.org/10.1016/S2214-109X(23)00606-X

Becher, I., Savitski, M. M., Savitski, M. F., Hopf, C., Bantscheff, M., & Drewes, G. (2013). Affinity profiling of the cellular kinome for the nucleotide cofactors ATP, ADP, and GTP. *ACS Chemical Biology*, *8*(3), 599–607. https://doi.org/10.1021/cb3005879

Beenstock, J., Mooshayef, N., & Engelberg, D. (2016). How Do Protein Kinases Take a Selfie (Autophosphorylate)? *Trends in Biochemical Sciences*, *41*(11), 938–953. https://doi.org/10.1016/j.tibs.2016.08.006

Beuzón, C. R., Méresse, S., Unsworth, K. E., Ruz-Albert, J., Garvis, S., Waterman, S. R., Ryder, T. A., Boucrot, E., & Holden, D. W. (2000). Salmonella maintains the integrity of its intracellular vacuole through the action of SifA. *The EMBO Journal*, *19*(13), 3235–3249. https://doi.org/https://doi.org/10.1093/emboj/19.13.3235

Bhaskaran, S. S., & Stebbins, C. E. (2012). Structure of the catalytic domain of the Salmonella virulence factor SseI. *Acta Crystallographica Section D*, *68*(12), 1613–1621. https://doi.org/10.1107/S0907444912039042

Birk, M. S., Walch, P., Baykara, T., Sefried, S., Amelang, J., Buerova, E., Breuer, I., Vervoorts, J., Typas, A., Savitski, M. M., Mateus, A., & Selkrig, J. (2024). Salmonella infection

impacts host proteome thermal stability. *European Journal of Cell Biology*, *103*(4), 151448. https://doi.org/https://doi.org/10.1016/j.ejcb.2024.151448

Brawn, L. C., Hayward, R. D., & Koronakis, V. (2007). Salmonella SPI1 Effector SipA Persists after Entry and Cooperates with a SPI2 Effector to Regulate Phagosome Maturation and Intracellular Replication. *Cell Host & Microbe*, *1*(1), 63–75. https://doi.org/https://doi.org/10.1016/j.chom.2007.02.001

Breitsprecher, D., & Goode, B. L. (2013). Formins at a glance. *Journal of Cell Science*, *126*(1), 1–7. https://doi.org/10.1242/jcs.107250

Brink, T., Leiss, V., Siegert, P., Jehle, D., Ebner, J. K., Schwan, C., Shymanets, A., Wiese, S., Nürnberg, B., Hensel, M., Aktories, K., & Orth, J. H. C. (2018). Salmonella Typhimurium effector SseI inhibits chemotaxis and increases host cell survival by deamidation of heterotrimeric Gi proteins. *PLOS Pathogens*, *14*(8), e1007248-. https://doi.org/10.1371/journal.ppat.1007248

Bryant, P., & Noé, F. (2024). Structure prediction of alternative protein conformations. *Nature Communications*, *15*(1), 7328. https://doi.org/10.1038/s41467-024-51507-2

Buchwald, G., Friebel, A., Galán, J. E., Hardt, W., Wittinghofer, A., & Scheffzek, K. (2002). Structural basis for the reversible activation of a Rho protein by the bacterial toxin SopE. *The EMBO Journal*, *21*(13), 3286-3295–3295. https://doi.org/https://doi.org/10.1093/emboj/cdf329

Burkinshaw, B. J., Prehna, G., Worrall, L. J., & Strynadka, N. C. J. (2012). Structure of Salmonella Effector Protein SopB N-terminal Domain in Complex with Host Rho GTPase Cdc42*. *Journal of Biological Chemistry*, *287*(16), 13348–13355. https://doi.org/https://doi.org/10.1074/jbc.M111.331330

Burnett, G., & Kennedy, E. P. (1954). THE ENZYMATIC PHOSPHORYLATION OF PROTEINS. *Journal of Biological Chemistry*, *211*(2), 969–980. https://doi.org/https://doi.org/10.1016/S0021-9258(18)71184-8

Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., & Madden, T. L. (2009). BLAST+: architecture and applications. *BMC Bioinformatics*, *10*(1), 421. https://doi.org/10.1186/1471-2105-10-421

Carden, S. E., Walker, G. T., Honeycutt, J., Lugo, K., Pham, T., Jacobson, A., Bouley, D., Idoyaga, J., Tsolis, R. M., & Monack, D. (2017). Pseudogenization of the Secreted Effector Gene sseI Confers Rapid Systemic Dissemination of S. Typhimurium ST313 within Migratory Dendritic Cells. *Cell Host & Microbe*, *21*(2), 182–194. https://doi.org/10.1016/j.chom.2017.01.009

Carlsson, A. E. (2018). Membrane bending by actin polymerization. *Current Opinion in Cell Biology*, *50*, 1–7. https://doi.org/https://doi.org/10.1016/j.ceb.2017.11.007

Carrera, A. C., Alexandrov, K., & Roberts, T. M. (1993). The conserved lysine of the catalytic domain of protein kinases is actively involved in the phosphotransfer reaction and not required for anchoring ATP (enzyme mechanism/protein-tyroslne kinase/phosphorylation). In *Proc. Nati. Acad. Sci. USA* (Vol. 90). https://www.pnas.org

Céline, M., Daniel, S., Nathalie, B., Javier, P.-C., Lawrence, L., Elise, C.-A., Isabel, P.-G., Eva, H., Anne, V., Christophe, G., Mathias, F., Pascale, C., & Patricia, D. (2015). The Legionella Kinase LegK2 Targets the ARP2/3 Complex To Inhibit Actin Nucleation on Phagosomes and Allow Bacterial Evasion of the Late Endocytic Pathway. *MBio*, *6*(3), 10.1128/mbio.00354-15. https://doi.org/10.1128/mbio.00354-15

Chang, J., Myeni, S. K., Lin, T. L., Wu, C. C., Staiger, C. J., & Zhou, D. (2007). SipC multimerization promotes actin nucleation and contributes to Salmonella-induced inflammation. *Molecular Microbiology*, *66*(6), 1548–1556. https://doi.org/https://doi.org/10.1111/j.1365-2958.2007.06024.x

Chattaway, M. A., Langridge, G. C., & Wain, J. (2021). Salmonella nomenclature in the genomic era: a time for change. *Scientific Reports*, *11*(1). https://doi.org/10.1038/s41598-021-86243-w

Chatterjee, S., Zhong, D., Nordhues, B. A., Battaile, K. P., Lovell, S., & De Guzman, R. N. (2011). The crystal structures of the Salmonella type III secretion system tip protein SipD in complex with deoxycholate and chenodeoxycholate. *Protein Science*, *20*(1), 75–86. https://doi.org/https://doi.org/10.1002/pro.537

Chen, L., Wang, H., Zhang, J., Gu, L., Huang, N., Zhou, J.-M., & Chai, J. (2008). Structural basis for the catalytic mechanism of phosphothreonine lyase. *Nature Structural & Molecular Biology*, *15*(1), 101–102. https://doi.org/10.1038/nsmb1329

Cobbaut, M., Derua, R., Parker, P. J., Waelkens, E., Janssens, V., & Van Lint, J. (2018). Protein kinase D displays intrinsic Tyr autophosphorylation activity: insights into mechanism and regulation. *FEBS Letters*, *592*(14), 2432–2443. https://doi.org/https://doi.org/10.1002/1873-3468.13171

Cohen, P., Cross, D., & Jänne, P. A. (2021). Kinase drug discovery 20 years after imatinib: progress and future directions. *Nature Reviews Drug Discovery*, *20*(7), 551–569. https://doi.org/10.1038/s41573-021-00195-4

Colonne, P. M., Winchell, C. G., & Voth, D. E. (2016). Hijacking Host Cell Highways: Manipulation of the Host Actin Cytoskeleton by Obligate Intracellular Bacterial Pathogens. *Frontiers in Cellular and Infection Microbiology*, *6*. https://doi.org/10.3389/fcimb.2016.00107

Courtemanche, N. (2018). Mechanisms of formin-mediated actin assembly and dynamics. In *Biophysical Reviews* (Vol. 10, Issue 6, pp. 1553–1569). Springer Verlag. https://doi.org/10.1007/s12551-018-0468-6

Dai, Y., Zhang, M., Liu, X., Sun, T., Qi, W., Ding, W., Chen, Z., Zhang, P., Liu, R., Chen, H., Chen, S., Wang, Y., Yue, Y., Song, N., Wang, W., Jia, H., Ma, Z., Li, C., Chen, Q., & Li, B. (2024). Salmonella manipulates macrophage migration via SteC-mediated myosin light chain activation to penetrate the gut-vascular barrier. *The EMBO Journal*, *43*(8), 1499-1518–1518. https://doi.org/https://doi.org/10.1038/s44318-024-00076-7

D'Costa, V. M., Braun, V., Landekic, M., Shi, R., Proteau, A., McDonald, L., Cygler, M., Grinstein, S., & Brumell, J. H. (2015). Salmonella Disrupts Host Endocytic Trafficking by SopD2-Mediated Inhibition of Rab7. *Cell Reports*, *12*(9), 1508–1518. https://doi.org/https://doi.org/10.1016/j.celrep.2015.07.063

de Souza Santos, M., & Orth, K. (2015). Subversion of the cytoskeleton by intracellular bacteria: lessons from Listeria, Salmonella and Vibrio. *Cellular Microbiology*, *17*(2), 164–173. https://doi.org/https://doi.org/10.1111/cmi.12399

Deng, W., Marshall, N. C., Rowland, J. L., McCoy, J. M., Worrall, L. J., Santos, A. S., Strynadka, N. C. J., & Finlay, B. B. (2017). Assembly, structure, function and regulation of type III secretion systems. In *Nature Reviews Microbiology* (Vol. 15, Issue 6, pp. 323–337). Nature Publishing Group. https://doi.org/10.1038/nrmicro.2017.20

Desai, D., Kantliwala, S. V, Vybhavi, J., Ravi, R., Patel, H., & Patel, J. (2024). Review of AlphaFold 3: Transformative Advances in Drug Design and Therapeutics. *Cureus*, *16*(7), e63646. https://doi.org/10.7759/cureus.63646

Diao, J., Zhang, Y., Huibregtse, J. M., Zhou, D., & Chen, J. (2008). Crystal structure of SopA, a Salmonella effector protein mimicking a eukaryotic ubiquitin ligase. *Nature Structural & Molecular Biology*, *15*(1), 65–70. https://doi.org/10.1038/nsmb1346

Domingues, L., Ismail, A., Charro, N., Rodríguez-Escudero, I., Holden, D. W., Molina, M., Cid, V. J., & Mota, L. J. (2016). The Salmonella effector SteA binds phosphatidylinositol 4-phosphate for subcellular targeting within host cells. *Cellular Microbiology*, *18*(7), 949–969. https://doi.org/https://doi.org/10.1111/cmi.12558

Doshi, B. M., Hightower, L. E., & Lee, J. (2009). The role of Hsp27 and actin in the regulation of movement in human cancer cells responding to heat shock. *Cell Stress and Chaperones*, *14*(5), 445–457. https://doi.org/10.1007/s12192-008-0098-1

Esposito, D., Dudley-Fraser, J., Garza-Garcia, A., & Rittinger, K. (2022). Divergent self-association properties of paralogous proteins TRIM2 and TRIM3 regulate their E3 ligase activity. *Nature Communications*, *13*(1), 7583. https://doi.org/10.1038/s41467-022-35300-7

Esposito, D., Günster, R. A., Martino, L., El Omari, K., Wagner, A., Thurston, T. L. M., & Rittinger, K. (2018). Structural basis for the glycosyltransferase activity of the Salmonella effector SseK3. *Journal of Biological Chemistry*, *293*(14), 5064–5078. https://doi.org/https://doi.org/10.1074/jbc.RA118.001796

Fang, X., Yu, S. X., Lu, Y., Bast, R. C., Woodgett, J. R., & Mills, G. B. (2000). Phosphorylation and inactivation of glycogen synthase kinase 3 by protein kinase A. *Proceedings of the National Academy of Sciences*, *97*(22), 11960–11965. https://doi.org/10.1073/pnas.220413597

Feasey, N. A., Dougan, G., Kingsley, R. A., Heyderman, R. S., & Gordon, M. A. (2012). Invasive non-typhoidal salmonella disease: an emerging and neglected tropical disease in Africa. *The Lancet*, *379*(9835), 2489–2499. https://doi.org/10.1016/S0140-6736(11)61752-2

Feng, Z.-Z., Jiang, A.-J., Mao, A.-W., Feng, Y., Wang, W., Li, J., Zhang, X., Xing, K., & Peng, X. (2018). The Salmonella effectors SseF and SseG inhibit Rab1A-mediated autophagy to facilitate intracellular bacterial survival and replication. *Journal of Biological Chemistry*, *293*(25), 9662–9673. https://doi.org/https://doi.org/10.1074/jbc.M117.811737

Fernandez-Piñar, P., Alemán, A., Sondek, J., Dohlman, H. G., Molina, M., & Martín, H. (2012). The Salmonella typhimurium effector SteC inhibits Cdc42-mediated signaling through binding to the exchange factor Cdc24 in Saccharomyces cerevisiae. *Molecular Biology of the Cell*, *23*(22), 4430–4443. https://doi.org/10.1091/mbc.E12-03-0243

Fiskin, E., Bhogaraju, S., Herhaus, L., Kalayil, S., Hahn, M., & Dikic, I. (2017). Structural basis for the recognition and degradation of host TRIM proteins by Salmonella effector SopA. *Nature Communications*, *8*(1), 14004. https://doi.org/10.1038/ncomms14004

Fookes, M., Schroeder, G. N., Langridge, G. C., Blondel, C. J., Mammina, C., Connor, T. R., Seth-Smith, H., Vernikos, G. S., Robinson, K. S., Sanders, M., Petty, N. K., Kingsley, R. A., Bäumler, A. J., Nuccio, S.-P., Contreras, I., Santiviago, C. A., Maskell, D., Barrow, P., Humphrey, T., … Thomson, N. R. (2011). Salmonella bongori Provides Insights into the Evolution of the Salmonellae. *PLOS Pathogens*, *7*(8), e1002191-. https://doi.org/10.1371/journal.ppat.1002191

Franke, D., Petoukhov, M. V, Konarev, P. V, Panjkovich, A., Tuukkanen, A., Mertens, H. D. T., Kikhney, A. G., Hajizadeh, N. R., Franklin, J. M., Jeffries, C. M., & Svergun, D. I. (2017). ATSAS 2.8: A comprehensive data analysis suite for small-angle scattering from macromolecular solutions. *Journal of Applied Crystallography*, *50*, 1212–1225. https://doi.org/10.1107/S1600576717007786

Frazer, C., & Young, P. G. (2012). Carboxy-terminal phosphorylation sites in Cdc25 contribute to enforcement of the DNA damage and replication checkpoints in fission yeast. *Current Genetics*, *58*(4), 217–234. https://doi.org/10.1007/s00294-012-0379-1

Friebel, A., & Hardt, W.-D. (2000). Purification and biochemical activity of Salmonella exchange factor SopE. In W. E. Balch, C. J. Der, & A. Hall (Eds.), *Regulators and Effectors of Small GTPases - Part D: Rho Family* (Vol. 325, pp. 82–91). Academic Press. https://doi.org/https://doi.org/10.1016/S0076-6879(00)25433-5

Friebel, A., Ilchmann, H., Aepfelbacher, M., Ehrbar, K., Machleidt, W., & Hardt, W.-D. (2001). SopE and SopE2 from Salmonella typhimurium Activate Different Sets of RhoGTPases of the Host Cell*. *Journal of Biological Chemistry*, *276*(36), 34035–34040. https://doi.org/https://doi.org/10.1074/jbc.M100609200

Fu, Y., & Galán, J. E. (1999). A Salmonella protein antagonizes Rac-1 and Cdc42 to mediate host-cell recovery after bacterial invasion. *Nature*, *401*(6750), 293–297. https://doi.org/10.1038/45829

García-García, A., Hicks, T., El Qaidi, S., Zhu, C., Hardwidge, P. R., Angulo, J., & Hurtado-Guerrero, R. (2021). NleB/SseK-catalyzed arginine-glycosylation and enteropathogen virulence are finely tuned by a single variable position contiguous to the catalytic machinery. *Chemical Science*, *12*(36), 12181–12191. https://doi.org/10.1039/D1SC04065K

Ge, J., Xu, H., Li, T., Zhou, Y., Zhang, Z., Li, S., Liu, L., & Shao, F. (2009). A Legionella type IV effector activates the NF-B pathway by phosphorylating the IB family of inhibitors. *Proceedings of the National Academy of Sciences*. www.pnas.orgcgidoi10.1073pnas.0907200106

Ge, P., Lei, Z., Yu, Y., Lu, Z., Qiang, L., Chai, Q., Zhang, Y., Zhao, D., Li, B., Pang, Y., Liu, C. H., & Wang, J. (2022). M. tuberculosis PknG manipulates host autophagy flux to promote pathogen intracellular survival. *Autophagy*, *18*(3), 576–594. https://doi.org/10.1080/15548627.2021.1938912

Geddes, K., Worley, M., Niemann, G., & Heffron, F. (2005). Identification of new secreted effectors in Salmonella enterica serovar typhimurium. *Infection and Immunity*, *73*(10), 6260–6271. https://doi.org/10.1128/IAI.73.10.6260-6271.2005

Gibson, D. G., Young, L., Chuang, R. Y., Venter, J. C., Hutchison, C. A., & Smith, H. O. (2009). Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nature Methods*, *6*(5), 343–345. https://doi.org/10.1038/nmeth.1318

Global Burden of Disease Collaborators. (2018). *Global, regional, and national age-sex-specific mortality for 282 causes of death in 195 countries and territories, 1980-2017: a systematic analysis for the Global Burden of Disease Study 2017*. https://vizhub.health

Gógl, G., Kornev, A. P., Reményi, A., & Taylor, S. S. (2019). Disordered Protein Kinase Regions in Regulation of Kinase Domain Cores. *Trends in Biochemical Sciences*, *44*(4), 300–311. https://doi.org/https://doi.org/10.1016/j.tibs.2018.12.002

Gong, H., Vu, G.-P., Bai, Y., Yang, E., Liu, F., & Lu, S. (2010). Differential expression of Salmonella type III secretion system factors InvJ, PrgJ, SipC, SipD, SopA and SopB in cultures and in mice. *Microbiology*, *156*(1), 116–127. https://doi.org/10.1099/mic.0.032318-0

Grabe, G. J., Zhang, Y., Przydacz, M., Rolhion, N., Yang, Y., Pruneda, J. N., Komander, D., Holden, D. W., & Hare, S. A. (2016). The Salmonella Effector SpvD Is a Cysteine Hydrolase with a Serovar-specific Polymorphism Influencing Catalytic Activity, Suppression of Immune Responses, and Bacterial Virulence*. *Journal of Biological Chemistry*, *291*(50), 25853–25863. https://doi.org/https://doi.org/10.1074/jbc.M116.752782

Greiner, J. V, & Glonek, T. (2021). Intracellular ATP Concentration and Implication for Cellular Evolution. *Biology*, *10*(11). https://doi.org/10.3390/biology10111166

Gunster, R., Matthews, S., Holden, D. W., & Thurston, T. (2017). SseK1 and SseK3 Type III Secretion System Effectors Inhibit NF-κB Signaling and Necroptotic Cell Death in Salmonella-Infected Macrophages. *Infection and Immunity*, *85*(3), 10.1128/iai.00010-17. https://doi.org/10.1128/iai.00010-17

Guo, H.-B., Perminov, A., Bekele, S., Kedziora, G., Farajollahi, S., Varaljay, V., Hinkle, K., Molinero, V., Meister, K., Hung, C., Dennis, P., Kelley-Loughnane, N., & Berry, R. (2022). AlphaFold2 models indicate that protein sequence determines both structure and dynamics. *Scientific Reports*, *12*(1), 10696. https://doi.org/10.1038/s41598-022-14382-9

Hanks, S. K., & Hunter, T. (1995). The eukaryotic protein kinase superfamily: kinase (catalytic) domain structure and classification 1 . *The FASEB Journal*, *9*(8), 576–596. https://doi.org/10.1096/fasebj.9.8.7768349

Hari, S. B., Merritt, E. A., & Maly, D. J. (2013). Sequence Determinants of a Specific Inactive Protein Kinase Conformation. *Chemistry & Biology*, *20*(6), 806–815. https://doi.org/https://doi.org/10.1016/j.chembiol.2013.05.005

Heggie, A., Cerny, O., & Holden, D. W. (2021). SteC and the intracellular Salmonella-induced F-actin meshwork. In *Cellular Microbiology* (Vol. 23, Issue 4). Blackwell Publishing Ltd. https://doi.org/10.1111/cmi.13315

Heissler, S. M., & Sellers, J. R. (2014). Myosin light chains: Teaching old dogs new tricks. In *Bioarchitecture* (Vol. 4, Issue 6, pp. 169–188). https://doi.org/10.1080/19490992.2015.1054092

Higgs, H. N., & Peterson, K. J. (2004). Phylogenetic Analysis of the Formin Homology 2 Domain. *Molecular Biology of the Cell*, *16*(1), 1–13. https://doi.org/10.1091/mbc.e04-07-0565

Houles, T., Yoon, S.-O., & Roux, P. P. (2024). The expanding landscape of canonical and non-canonical protein phosphorylation. *Trends in Biochemical Sciences*. https://doi.org/10.1016/j.tibs.2024.08.004

Huse, M., & Kuriyan, J. (2002). The Conformational Plasticity of Protein Kinases. *Cell*, *109*(3), 275–282. https://doi.org/https://doi.org/10.1016/S0092-8674(02)00741-9

Imami, K., Bhavsar, A. P., Yu, H., Brown, N. F., Rogers, L. D., Finlay, B. B., & Foster, L. J. (2013). Global impact of Salmonella pathogenicity island 2-secreted effectors on the host phosphoproteome. *Molecular and Cellular Proteomics*, *12*(6), 1632–1643. https://doi.org/10.1074/mcp.M112.026161

Jennings, E., Esposito, D., Rittinger, K., & Thurston, T. L. M. (2018). Structure–function analyses of the bacterial zinc metalloprotease effector protein GtgA uncover key residues required for deactivating NF-κB. *Journal of Biological Chemistry*, *293*(39), 15316–15329. https://doi.org/https://doi.org/10.1074/jbc.RA118.004255

Jennings, E., Thurston, T. L. M., & Holden, D. W. (2017). Salmonella SPI-2 Type III Secretion System Effectors: Molecular Mechanisms And Physiological Consequences. In *Cell Host and Microbe* (Vol. 22, Issue 2, pp. 217–231). Cell Press. https://doi.org/10.1016/j.chom.2017.07.009

Johnson, J., Yaron, T. M., Huntsman, E. M., Kerelsky, A., Song, J., Regev, A., Lin, T. Y., Liberatore, K., Cizin, D. M., Cohen, B. M., Vasan, N., Ma, Y., Krismer, K., Robles, J. T., van de Kooij, B., van Vlimmeren, A. E., Andrée-Busch, N., Käufer, N. F., Dorovkov, M. V, … Cantley, L. C. (2023). An atlas of substrate specificities for the human serine/threonine kinome. *Nature*, *613*(7945), 759–766. https://doi.org/10.1038/s41586-022-05575-3

Johnson, L., & Lewis, R. (2001). Structural Basis for Control by Phosphorylation. *Chemical Reviews*, *101*(8), 2209–2242. https://doi.org/10.1021/cr000225s

Johnson, L., Noble, M. E. M., & Owen, D. J. (1996). Active and Inactive Protein Kinases: Structural Basis for Regulation. *Cell*, *85*(2), 149–158. https://doi.org/10.1016/S0092-8674(00)81092-2

Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Žídek, A., Potapenko, A., Bridgland, A., Meyer, C., Kohl, S. A. A., Ballard, A. J., Cowie, A., Romera-Paredes, B., Nikolov, S., Jain, R., Adler, J., … Hassabis, D. (2021). Highly accurate protein structure prediction with AlphaFold. *Nature*, *596*(7873), 583–589. https://doi.org/10.1038/s41586-021-03819-2

Juris, S. J., Rudolph, A. E., Huddler, D., Orth, K., & Dixon, J. E. (2000). A distinctive role for the Yersinia protein kinase: Actin binding, kinase activation, and cytoskeleton disruption. *Proceedings of the National Academy of Sciences*. www.pnas.orgcgidoi10.1073pnas.170281997

Juris, S. J., Shah, K., Shokat, K., Dixon, J. E., & Vacratsis, P. O. (2006). Identification of otubain 1 as a novel substrate for the Yersinia protein kinase using chemical genetics and mass spectrometry. *FEBS Letters*, *580*(1), 179–183. https://doi.org/https://doi.org/10.1016/j.febslet.2005.11.071

Kamanova, J., Sun, H., Lara-Tejero, M., & Galán, J. E. (2016). The Salmonella Effector Protein SopA Modulates Innate Immune Responses by Targeting TRIM E3 Ligase Family

Members. *PLOS Pathogens*, *12*(4), e1005552-. https://doi.org/10.1371/journal.ppat.1005552

Kamm, K. E., & Stull, J. T. (2001). Dedicated Myosin Light Chain Kinases with Diverse Cellular Functions. In *Journal of Biological Chemistry* (Vol. 276, Issue 7, pp. 4527–4530). https://doi.org/10.1074/jbc.R000028200

Kaniga, K., Trollinger, D., & Galán, J. E. (1995). Identification of two targets of the type III protein secretion system encoded by the inv and spa loci of Salmonella typhimurium that have homology to the Shigella IpaD and IpaA proteins. *Journal of Bacteriology*, *177*(24), 7078–7085. https://doi.org/10.1128/jb.177.24.7078-7085.1995

Kannan, N., Taylor, S. S., Zhai, Y., Venter, J. C., & Manning, G. (2007). Structural and functional diversity of the microbial kinome. *PLoS Biol*, *5*(3), 17. https://doi.org/10.1371/journal.pbio

Kariuki, S., Gordon, M. A., Feasey, N., & Parry, C. M. (2015). Antimicrobial resistance and management of invasive Salmonella disease. *Vaccine*, *33*, C21–C29. https://doi.org/https://doi.org/10.1016/j.vaccine.2015.03.102

Keszei, A. F. A., Tang, X., McCormick, C., Zeqiraj, E., Rohde, J. R., Tyers, M., & Sicheri, F. (2014). Structure of an SspH1-PKN1 Complex Reveals the Basis for Host Substrate Recognition and Mechanism of Activation for a Bacterial E3 Ubiquitin Ligase. *Molecular and Cellular Biology*, *34*(3), 362–373. https://doi.org/10.1128/MCB.01360-13

Kidwai, A. S., Mushamiri, I., Niemann, G. S., Brown, R. N., Adkins, J. N., & Heffron, F. (2013). Diverse Secreted Effectors Are Required for Salmonella Persistence in a Mouse Infection Model. *PLOS ONE*, *8*(8), e70753-. https://doi.org/10.1371/journal.pone.0070753

Kim, D. W., Lenzen, G., Page, A.-L., Legrain, P., Sansonetti, P. J., & Parsot, C. (2005). The Shigella flexneri effector OspG interferes with innate immune responses by targeting ubiquitin-conjugating enzymes. *Proceedings of the National Academy of Sciences*, *102*(39), 14046–14051. https://doi.org/10.1073/pnas.0504466102

Kim, J., Ahuja, L. G., Chao, F.-A., Xia, Y., McClendon, C. L., Kornev, A. P., Taylor, S. S., & Veglia, G. (2017). A dynamic hydrophobic core orchestrates allostery in protein kinases. *Science Advances*, *3*(4), e1600663. https://doi.org/10.1126/sciadv.1600663

King, A., & Blackledge, M. S. (2021). Evaluation of small molecule kinase inhibitors as novel antimicrobial and antibiofilm agents. *Chemical Biology & Drug Design*, *98*(6), 1038–1064. https://doi.org/https://doi.org/10.1111/cbdd.13962

Knighton, D. R., Zheng, J., Eyck, L. F. Ten, Ashford, V. A., Xuong, N.-H., Taylor, S. S., & Sowadski, J. M. (1991). Crystal Structure of the Catalytic Subunit of Cyclic Adenosine Monophosphate-Dependent Protein Kinase. *Science*, *253*(5018), 407–414. https://doi.org/10.1126/science.1862342

Knuff, K., & Finlay, B. B. (2017). What the SIF Is happening—The role of intracellular Salmonella-induced filaments. *Frontiers in Cellular and Infection Microbiology*, *7*(JUL). https://doi.org/10.3389/fcimb.2017.00335

Knuff-Janzen, K., Tupin, A., Yurist-Doutsch, S., Rowland, J. L., & Finlay, B. B. (2020). Multiple Salmonella-pathogenicity island 2 effectors are required to facilitate bacterial establishment of its intracellular niche and virulence. *PLOS ONE*, *15*(6), e0235020-. https://doi.org/10.1371/journal.pone.0235020

Kooijman, E. E., & Burger, K. N. J. (2009). Biophysics and function of phosphatidic acid: A molecular perspective. *Biochimica et Biophysica Acta (BBA) - Molecular and Cell Biology of Lipids*, *1791*(9), 881–888. https://doi.org/https://doi.org/10.1016/j.bbalip.2009.04.001

Kornev, A. P., Taylor, S. S., & Ten Eyck, L. F. (2008). A helix scaffold for the assembly of active protein kinases. *Proceedings of the National Academy of Sciences*, *105*(38), 14377–14382. https://doi.org/10.1073/pnas.0807988105

Kubori, T., & Galán, J. E. (2003). Temporal Regulation of Salmonella Virulence Effector Function by Proteasome-Dependent Protein Degradation. *Cell*, *115*(3), 333–342. https://doi.org/https://doi.org/10.1016/S0092-8674(03)00849-3

Kufareva, I., Lenoir, M., Dancea, F., Sridhar, P., Raush, E., Bissig, C., Gruenberg, J., Abagyan, R., & Overduin, M. (2014). Discovery of novel membrane binding structures and functions. *Biochemistry and Cell Biology*, *92*(6), 555–563. https://doi.org/10.1139/bcb-2014-0074

Kühn, S., Erdmann, C., Kage, F., Block, J., Schwenkmezger, L., Steffen, A., Rottner, K., & Geyer, M. (2015). The structure of FMNL2-Cdc42 yields insights into the mechanism of lamellipodia and filopodia formation. *Nature Communications*, *6*. https://doi.org/10.1038/ncomms8088

Labriola, J. M., Zhou, Y., & Nagar, B. (2018). Structural Analysis of the Bacterial Effector AvrA Identifies a Critical Helix Involved in Substrate Recognition. *Biochemistry*, *57*(33), 4985–4996. https://doi.org/10.1021/acs.biochem.8b00512

Lauber, B. S., Hardegger, L. A., Asraful, A. K., Lund, B. A., Dumele, O., Harder, M., Kuhn, B., Engh, R. A., & Diederich, F. (2016). Addressing the Glycine-Rich Loop of Protein Kinases by a Multi-Facetted Interaction Network: Inhibition of PKA and a PKB Mimic. *Chemistry – A European Journal*, *22*(1), 211–221. https://doi.org/https://doi.org/10.1002/chem.201503552

Lavoie, H., Li, J. J., Thevakumaran, N., Therrien, M., & Sicheri, F. (2014a). Dimerization-induced allostery in protein kinase regulation. *Trends in Biochemical Sciences*, *39*(10), 475–486. https://doi.org/https://doi.org/10.1016/j.tibs.2014.08.004

Lavoie, H., Li, J. J., Thevakumaran, N., Therrien, M., & Sicheri, F. (2014b). Dimerization-induced allostery in protein kinase regulation. *Trends in Biochemical Sciences*, *39*(10), 475–486. https://doi.org/https://doi.org/10.1016/j.tibs.2014.08.004

Ledoux, B., Zanin, N., Yang, J., Mercier, V., Coster, C., Dupont-Gillain, C., Alsteens, D., Morsomme, P., & Renard, H.-F. (2024). Plasma membrane nanodeformations promote actin polymerization through CIP4/CDC42 recruitment and regulate type II IFN signaling. *Science Advances*, *9*(50), eade1660. https://doi.org/10.1126/sciadv.ade1660

Lee, P.-C., & Machner, M. P. (2018). The Legionella Effector Kinase LegK7 Hijacks the Host Hippo Pathway to Promote Infection. *Cell Host & Microbe*, *24*(3), 429-438.e6. https://doi.org/https://doi.org/10.1016/j.chom.2018.08.004

Lemmon, M. A. (2008). Membrane recognition by phospholipid-binding domains. *Nature Reviews Molecular Cell Biology*, *9*(2), 99–111. https://doi.org/10.1038/nrm2328

Leonard, T. A., & Hurley, J. H. (2011). Regulation of protein kinases by lipids. *Current Opinion in Structural Biology*, *21*(6), 785–791. https://doi.org/https://doi.org/10.1016/j.sbi.2011.07.006

Lesnick, M. L., Reiner, N. E., Fierer, J., & Guiney, D. G. (2001). The Salmonella spvB virulence gene encodes an enzyme that ADP-ribosylates actin and destabilizes the cytoskeleton of eukaryotic cells. *Molecular Microbiology*, *39*(6), 1464–1470. https://doi.org/https://doi.org/10.1046/j.1365-2958.2001.02360.x

Li, H., Xu, H., Zhou, Y., Zhang, J., Long, C., Li, S., Chen, S., Zhou, J.-M., & Shao, F. (2007). The Phosphothreonine Lyase Activity of a Bacterial Type III Effector Family. *Science*, *315*(5814), 1000–1003. https://doi.org/10.1126/science.1138960

Li, J., Gong, C., Zhou, H., Liu, J., Xia, X., Ha, W., Jiang, Y., Liu, Q., & Xiong, H. (2024). Kinase Inhibitors and Kinase-Targeted Cancer Therapies: Recent Advances and Future Perspectives. In *International Journal of Molecular Sciences* (Vol. 25, Issue 10). Multidisciplinary Digital Publishing Institute (MDPI). https://doi.org/10.3390/ijms25105489

Li, S., Zhang, L., Yao, Q., Li, L., Dong, N., Rong, J., Gao, W., Ding, X., Sun, L., Chen, X., Chen, S., & Shao, F. (2013). Pathogen blocks host death receptor signalling by arginine GlcNAcylation of death domains. *Nature*, *501*(7466), 242–246. https://doi.org/10.1038/nature12436

Lilic, M., Vujanac, M., & Stebbins, C. E. (2006). A Common Structural Motif in the Binding of Virulence Factors to Bacterial Secretion Chaperones. *Molecular Cell*, *21*(5), 653–664. https://doi.org/https://doi.org/10.1016/j.molcel.2006.01.026

Lochhead, P. A., Sibbet, G., Morrice, N., & Cleghon, V. (2005). Activation-Loop Autophosphorylation Is Mediated by a Novel Transitional Intermediate Form of DYRKs. *Cell*, *121*(6), 925–936. https://doi.org/10.1016/j.cell.2005.03.034

Loeffler, F. (1892). Epidemieen unter den im hygienishcen Institute zu Greifswald gehaltenen Mausen und uber die BekaUeber mpfung der Feldmausplage. *Zbl. Bakt. Parasitenkunde*, *11*.

Lombardi, C., Tolchard, J., Bouillot, S., Signor, L., Gebus, C., Liebl, D., Fenel, D., Teulon, J.-M., Brock, J., Habenstein, B., Pellequer, J.-L., Faudry, E., Loquet, A., Attrée, I., Dessen, A., & Job, V. (2019). Structural and Functional Characterization of the Type Three Secretion System (T3SS) Needle of Pseudomonas aeruginosa. *Frontiers in Microbiology*, *10*. https://www.frontiersin.org/journals/microbiology/articles/10.3389/fmicb.2019.00573

Lopez, V. A., Park, B. C., Nowak, D., Sreelatha, A., Zembek, P., Fernandez, J., Servage, K. A., Gradowski, M., Hennig, J., Tomchick, D. R., Pawłowski, K., Krzymowska, M., & Tagliabracci, V. S. (2019). A Bacterial Effector Mimics a Host HSP90 Client to Undermine Immunity. *Cell*, *179*(1), 205-218.e21. https://doi.org/https://doi.org/10.1016/j.cell.2019.08.020

Manning, G., Whyte, D. B., Martinez, R., Hunter, T., & Sudarsanam, S. (2002). The Protein Kinase Complement of the Human Genome. *Science*, *298*(5600), 1912–1934. https://doi.org/10.1126/science.1075762

Marchello, C. S., Birkhold, M., Crump, J. A., Martin, L. B., Ansah, M. O., Breghi, G., Canals, R., Fiorino, F., Gordon, M. A., Kim, J. H., Hamaluba, M., Hanumunthadu, B., Jacobs, J., Kariuki, S., Malvolti, S., Mantel, C., Marks, F., Medaglini, D., Mogasale, V., … Tack, B. (2022). Complications and mortality of non-typhoidal salmonella invasive disease: a global systematic review and meta-analysis. *The Lancet Infectious Diseases*. https://doi.org/10.1016/S1473-3099(21)00615-0

Marcotte, D., Rushe, M., M. Arduini, R., Lukacs, C., Atkins, K., Sun, X., Little, K., Cullivan, M., Paramasivam, M., Patterson, T. A., Hesson, T., D. McKee, T., May-Dracka, T. L., Xin, Z., Bertolotti-Ciarlet, A., Bhisetti, G. R., Lyssikatos, J. P., & Silvian, L. F. (2017). Germinal-center kinase-like kinase co-crystal structure reveals a swapped activation loop and C-terminal extension. *Protein Science*, *26*(2), 152–162. https://doi.org/https://doi.org/10.1002/pro.3062

Margarit, S. M., Davidson, W., Frego, L., & Stebbins, C. E. (2006). A Steric Antagonism of Actin Polymerization by a Salmonella Virulence Protein. *Structure*, *14*(8), 1219–1229. https://doi.org/https://doi.org/10.1016/j.str.2006.05.022

Mascher, T., Helmann, J. D., & Unden, G. (2006). Stimulus Perception in Bacterial Signal-Transducing Histidine Kinases. *Microbiology and Molecular Biology Reviews*, *70*(4), 910–938. https://doi.org/10.1128/mmbr.00020-06

Matsumoto, H., & Young, G. M. (2009). Translocated effectors of Yersinia. *Current Opinion in Microbiology*, *12*(1), 94–100. https://doi.org/https://doi.org/10.1016/j.mib.2008.12.005

McGhie, E. J., Hayward, R. D., & Koronakis, V. (2004). Control of Actin Turnover by a Salmonella Invasion Protein. *Molecular Cell*, *13*(4), 497–510. https://doi.org/https://doi.org/10.1016/S1097-2765(04)00053-X

McLaughlin, L. M., Govoni, G. R., Gerke, C., Gopinath, S., Peng, K., Laidlaw, G., Chien, Y.-H., Jeong, H.-W., Li, Z., Brown, M. D., Sacks, D. B., & Monack, D. (2009). The Salmonella SPI2 Effector SseI Mediates Long-Term Systemic Infection by Modulating Host Cell Migration. *PLOS Pathogens*, *5*(11), e1000671-. https://doi.org/10.1371/journal.ppat.1000671

McLaughlin, L. M., Xu, H., Carden, S. E., Fisher, S., Reyes, M., Heilshorn, S. C., & Monack, D. M. (2014). A microfluidic-based genetic screen to identify microbial virulence factors that inhibit dendritic cell migration. *Integrative Biology*, *6*(4), 438–449. https://doi.org/10.1039/c3ib40177d

Meharena, H. S., Chang, P., Keshwani, M. M., Oruganty, K., Nene, A. K., Kannan, N., Taylor, S. S., & Kornev, A. P. (2013). Deciphering the Structural Basis of Eukaryotic Protein Kinase Regulation. *PLOS Biology*, *11*(10), e1001680-. https://doi.org/10.1371/journal.pbio.1001680

Meier, F., Brunner, A.-D., Frank, M., Ha, A., Bludau, I., Voytik, E., Kaspar-Schoenefeld, S., Lubeck, M., Raether, O., Bache, N., Aebersold, R., Collins, B. C., Röst, H. L., & Mann, M. (2020). diaPASEF: parallel accumulation–serial fragmentation combined with data-independent acquisition. *Nature Methods*, *17*(12), 1229–1236. https://doi.org/10.1038/s41592-020-00998-0

Méresse, S., Unsworth, K. E., Habermann, A., Griffiths, G., Fang, F., Martínez-Lorenzo, M. J., Waterman, S. R., Gorvel, J. P., & Holden, D. W. (2001a). Remodelling of the actin cytoskeleton is essential for replication of intravacuolar Salmonella. *Cellular Microbiology*, *3*(8), 567–577. https://doi.org/10.1046/j.1462-5822.2001.00141.x

Méresse, S., Unsworth, K. E., Habermann, A., Griffiths, G., Fang, F., Martínez-Lorenzo, M. J., Waterman, S. R., Gorvel, J. P., & Holden, D. W. (2001b). Remodelling of the actin cytoskeleton is essential for replication of intravacuolar Salmonella. *Cellular Microbiology*, *3*(8), 567–577. https://doi.org/10.1046/j.1462-5822.2001.00141.x

Michel, D., Shaynoor, D., Olivier, D., Marie-Anne, N., Elisabeth, V., Grégory, J., Alain, C., Tarek, M., & Bertrand, D. (2009). Characterization of a Serine/Threonine Kinase Involved in Virulence of Staphylococcus aureus. *Journal of Bacteriology*, *191*(13), 4070–4081. https://doi.org/10.1128/jb.01813-08

Modi, V., & Dunbrack, R. L. (2019). A Structurally-Validated Multiple Sequence Alignment of 497 Human Protein Kinase Domains. *Scientific Reports*, *9*(1). https://doi.org/10.1038/s41598-019-56499-4

Morgan, E., Campbell, J. D., Rowe, S. C., Bispham, J., Stevens, M. P., Bowen, A. J., Barrow, P. A., Maskell, D. J., & Wallis, T. S. (2004). Identification of host-specific colonization factors of Salmonella enterica serovar Typhimurium. *Molecular Microbiology*, *54*(4), 994–1010. https://doi.org/10.1111/j.1365-2958.2004.04323.x

Moss, S. M., Taylor, I. R., Ruggero, D., Gestwicki, J. E., Shokat, K. M., & Mukherjee, S. (2019). A Legionella pneumophila Kinase Phosphorylates the Hsp70 Chaperone Family to Inhibit Eukaryotic Protein Synthesis. *Cell Host & Microbe*, *25*(3), 454-462.e6. https://doi.org/https://doi.org/10.1016/j.chom.2019.01.006

Myeni, S. K., & Zhou, D. (2010). The C Terminus of SipC Binds and Bundles F-actin to Promote Salmonella Invasion. *Journal of Biological Chemistry*, *285*(18), 13357–13363. https://doi.org/https://doi.org/10.1074/jbc.M109.094045

Navarro, L., Koller, A., Nordfelth, R., Wolf-Watz, H., Taylor, S., & Dixon, J. E. (2007). Identification of a Molecular Target for the Yersinia Protein Kinase A. *Molecular Cell*, *26*(4), 465–477. https://doi.org/https://doi.org/10.1016/j.molcel.2007.04.025

Nolen, B., Taylor, S., & Ghosh, G. (2004). Review Regulation of Protein Kinases: Controlling Activity through Activation Segment Conformation. In *Molecular Cell* (Vol. 15).

Odendall, C., Rolhion, N., Förster, A., Poh, J., Lamont, D. J., Liu, M., Freemont, P. S., Catling, A. D., & Holden, D. W. (2012). The salmonella kinase SteC targets the MAP kinase MEK to regulate the host actin cytoskeleton. *Cell Host and Microbe*, *12*(5), 657–668. https://doi.org/10.1016/j.chom.2012.09.011

Ohlson, B., Kerry, F., Birmingham, L., & Brumell, J. (2005). SseJ Deacylase Activity by Salmonella enterica Serovar Typhimurium Promotes Virulence in Mice. *Infection and Immunity*, *73*(10), 6249–6259. https://doi.org/10.1128/iai.73.10.6249-6259.2005

Ohlson, M. B., Huang, Z., Alto, N. M., Blanc, M.-P., Dixon, J. E., Chai, J., & Miller, S. I. (2008). Structure and Function of Salmonella SifA Indicate that Its Interactions with SKIP, SseJ, and RhoA Family GTPases Induce Endosomal Tubulation. *Cell Host & Microbe*, *4*(5), 434–446. https://doi.org/https://doi.org/10.1016/j.chom.2008.08.012

Olsen, J. V, Blagoev, B., Gnad, F., Macek, B., Kumar, C., Mortensen, P., & Mann, M. (2006). Global, In Vivo, and Site-Specific Phosphorylation Dynamics in Signaling Networks. *Cell*, *127*(3), 635–648. https://doi.org/10.1016/j.cell.2006.09.026

Ono, S. (2007). Mechanism of Depolymerization and Severing of Actin Filaments and Its Significance in Cytoskeletal Dynamics. In *International Review of Cytology* (Vol. 258, pp. 1–82). https://doi.org/10.1016/S0074-7696(07)58001-0

Ørskov, J. V., & Moltke, O. (1929). Studien über den Infektions-mechanismus bei verschiedenen Paratyphus-infektionen aus weissen Mäusen. . . *Z Immunitaetsforsch Exp Ther, 59, 357.*

Panagi, I., Jennings, E., Zeng, J., Günster, R. A., Stones, C. D., Mak, H., Jin, E., Stapels, D. A. C., Subari, Nur. Z., Pham, T. H. M., Brewer, S. M., Ong, S. Y. Q., Monack, D. M., Helaine, S., & Thurston, T. L. M. (2020). Salmonella Effector SteE Converts the Mammalian Serine/Threonine Kinase GSK3 into a Tyrosine Kinase to Direct Macrophage Polarization. *Cell Host & Microbe*, *27*(1), 41-53.e6. https://doi.org/10.1016/j.chom.2019.11.002

Papalazarou, V., & Machesky, L. M. (2021). The cell pushes back: The Arp2/3 complex is a key orchestrator of cellular responses to environmental forces. In *Current Opinion in Cell Biology* (Vol. 68, pp. 37–44). Elsevier Ltd. https://doi.org/10.1016/j.ceb.2020.08.012

Park, E. R., Eblen, S. T., & Catling, A. D. (2007). *MEK1 activation by PAK: A novel mechanism*.

Park, J. B., Kim, Y. H., Yoo, Y., Kim, J., Jun, S.-H., Cho, J. W., El Qaidi, S., Walpole, S., Monaco, S., García-García, A. A., Wu, M., Hays, M. P., Hurtado-Guerrero, R., Angulo, J., Hardwidge, P. R., Shin, J.-S., & Cho, H.-S. (2018). Structural basis for arginine glycosylation of host substrates by bacterial effector proteins. *Nature Communications*, *9*(1), 4283. https://doi.org/10.1038/s41467-018-06680-6

Patel, S., Wall, D. M., Castillo, A., & McCormick, B. A. (2019). Caspase-3 cleavage of Salmonella type III secreted effector protein SifA is required for localization of functional domains and bacterial dissemination. *Gut Microbes*, *10*(2), 172–187. https://doi.org/10.1080/19490976.2018.1506668

Pearson, J. S., Giogha, C., Ong, S. Y., Kennedy, C. L., Kelly, M., Robinson, K. S., Lung, T. W. F., Mansell, A., Riedmaier, P., Oates, C. V. L., Zaid, A., Mühlen, S., Crepin, V. F., Marches, O., Ang, C.-S., Williamson, N. A., O'Reilly, L. A., Bankovacki, A., Nachbur, U., … Hartland, E. L. (2013). A type III effector antagonizes death receptor signalling during bacterial gut infection. *Nature*, *501*(7466), 247–251. https://doi.org/10.1038/nature12524

Petri, E., Errico, A., Escobedo, L., Hunt, T., & Basavappa, R. (2007). The Crystal Structure of Human Cyclin B. *Cell Cycle*, *6*(11), 1342–1349. https://doi.org/10.4161/cc.6.11.4297

Pham, T. H., Gao, X., Singh, G., & Hardwidge, P. R. (2013). Escherichia coli Virulence Protein NleH1 Interaction with the v-Crk Sarcoma Virus CT10 Oncogene-like Protein (CRKL) Governs NleH1 Inhibition of the Ribosomal Protein S3 (RPS3)/Nuclear Factor κB (NF-κB) Pathway*. *Journal of Biological Chemistry*, *288*(48), 34567–34574. https://doi.org/https://doi.org/10.1074/jbc.M113.512376

Philpott, J. M., Narasimamurthy, R., Ricci, C. G., Freeberg, A. M., Hunt, S. R., Yee, L. E., Pelofsky, R. S., Tripathi, S., Virshup, D. M., & Partch, C. L. (2020). Casein kinase 1 dynamics underlie substrate selectivity and the PER2 circadian phosphoswitch. *ELife*, *9*, e52343. https://doi.org/10.7554/eLife.52343

Pilar, A. V. C., Reid-Yu, S. A., Cooper, C. A., Mulder, D. T., & Coombes, B. K. (2012). GogB Is an Anti-Inflammatory Effector that Limits Tissue Damage during Salmonella Infection

through Interaction with Human FBXO22 and Skp1. *PLOS Pathogens*, *8*(6), e1002773-. https://doi.org/10.1371/journal.ppat.1002773

Pillay, T. D., Hettiarachchi, S. U., Gan, J., Diaz-Del-Olmo, I., Yu, X. J., Muench, J. H., Thurston, T. L. M., & Pearson, J. S. (2023). Speaking the host language: how Salmonella effector proteins manipulate the host. *Microbiology (Reading, England)*, *169*(6). https://doi.org/10.1099/mic.0.001342

Pinna, L. A. (2003). The Raison D'Être of Constitutively Active Protein Kinases: The Lesson of CK2. *Accounts of Chemical Research*, *36*(6), 378–384. https://doi.org/10.1021/ar020164f

Plaza, N., Urrutia, I. M., Garcia, K., Waldor, M. K., & Blondel, C. J. (2021). Identification of a Family of Vibrio Type III Secretion System Effectors That Contain a Conserved Serine/Threonine Kinase Domain. *MSphere*, *6*(4), 10.1128/msphere.00599-21. https://doi.org/10.1128/msphere.00599-21

Poh, J., Odendall, C., Spanos, A., Boyle, C., Liu, M., Freemont, P., & Holden, D. W. (2008). SteC is a Salmonella kinase required for SPI-2-dependent F-actin remodelling. *Cellular Microbiology*, *10*(1), 20–30. https://doi.org/10.1111/j.1462-5822.2007.01010.x

Pollard, T. D. (2016). Actin and actin-binding proteins. *Cold Spring Harbor Perspectives in Biology*, *8*(8). https://doi.org/10.1101/cshperspect.a018226

Pollock, G. L., Grishin, A. M., Giogha, C., Gan, J., Oates, C. V, Mcmillan, P. J., Gaeta, I., Tyska, M. J., Pearson, J. S., Scott, N. E., Cygler, M., & Hartland, E. L. (2022). Targeting of microvillus protein Eps8 by the NleH effector kinases from enteropathogenic E. coli. *Proceedings of the National Academy of Sciences*. https://doi.org/10.1073/pnas

Pruneda, J. N., Durkin, C. H., Geurink, P. P., Ovaa, H., Santhanam, B., Holden, D. W., & Komander, D. (2016). The Molecular Basis for Ubiquitin and Ubiquitin-like Specificities in Bacterial Effector Proteases. *Molecular Cell*, *63*(2), 261–276. https://doi.org/https://doi.org/10.1016/j.molcel.2016.06.015

Pruneda, J. N., Smith, F. D., Daurie, A., Swaney, D. L., Villén, J., Scott, J. D., Stadnyk, A. W., Trong, I. Le, Stenkamp, R. E., Klevit, R. E., Rohde, J. R., & Brzovic, P. S. (2014). E2~Ub conjugates regulate the kinase activity of Shigella effector OspG during pathogenesis. *EMBO Journal*, *33*(5), 437–449. https://doi.org/10.1002/embj.201386386

Pucheta-Martínez, E., Saladino, G., Morando, M. A., Martinez-Torrecuadrada, J., Lelli, M., Sutto, L., D'Amelio, N., & Gervasio, F. L. (2016). An allosteric cross-talk between the activation loop and the ATP binding site regulates the activation of Src kinase. *Scientific Reports*, *6*. https://doi.org/10.1038/srep24235

Quezada, C. M., Hicks, S. W., Galá, J. E., & Stebbins, C. E. (2008). A family of Salmonella virulence factors functions as a distinct class of autoregulated E3 ubiquitin ligases.

Proceedings of the National Academy of Sciences. www.pnas.orgcgidoi10.1073pnas.0811058106

Rambo, R. P., & Tainer, J. A. (2011). Characterizing flexible and intrinsically unstructured biological macromolecules by SAS using the Porod-Debye law. *Biopolymers*, *95*(8), 559–571. https://doi.org/10.1002/bip.21638

Reichow, S. L., Clemens, D. M., Freites, J. A., Németh-Cahalan, K. L., Heyden, M., Tobias, D. J., Hall, J. E., & Gonen, T. (2013). Allosteric mechanism of water-channel gating by Ca 2+ -calmodulin. *Nature Structural and Molecular Biology*, *20*(9), 1085–1092. https://doi.org/10.1038/nsmb.2630

Reinhardt, R., & Leonard, T. A. (2023). A critical evaluation of protein kinase regulation by activation loop autophosphorylation. In *eLife* (Vol. 12). eLife Sciences Publications Ltd. https://doi.org/10.7554/eLife.88210

Rong, F., & Kushner, S. R. (1991). Construction of versatile low-copy-number vectors for cloning, sequencing and gene expression in Escherichia coli. *Gene*, *100*, 195–199. https://doi.org/https://doi.org/10.1016/0378-1119(91)90366-J

Roskoski, R. (2015). A historical overview of protein kinases and their targeted small molecule inhibitors. *Pharmacological Research*, *100*, 1–23. https://doi.org/https://doi.org/10.1016/j.phrs.2015.07.010

Rottner, K., Faix, J., Bogdan, S., Linder, S., & Kerkhoff, E. (2017). Actin assembly mechanisms at a glance. *Journal of Cell Science*, *130*(20), 3427–3435. https://doi.org/10.1242/jcs.206433

Rotty, J. D., Wu, C., & Bear, J. E. (2013). New insights into the regulation and cellular functions of the ARP2/3 complex. *Nature Reviews Molecular Cell Biology*, *14*(1), 7–12. https://doi.org/10.1038/nrm3492

Russell, P., & Nurse, P. (1987). Negative regulation of mitosis by wee1, a gene encoding a protein kinase homolog. *Cell*, *49*(4), 559–567. https://doi.org/10.1016/0092-8674(87)90458-2

Scott, N. E., Giogha, C., Pollock, G. L., Kennedy, C. L., Webb, A. I., Williamson, N. A., Pearson, J. S., & Hartland, E. L. (2017). The bacterial arginine glycosyltransferase effector NleB preferentially modifies Fas-associated death domain protein (FADD). *Journal of Biological Chemistry*, *292*(42), 17337–17350. https://doi.org/10.1074/jbc.M117.805036

Sessa, F., Mapelli, M., Ciferri, C., Tarricone, C., Areces, L. B., Schneider, T. R., Stukenberg, P. T., & Musacchio, A. (2005). Mechanism of Aurora B activation by INCENP and inhibition by hesperadin. *Molecular Cell*, *18*(3), 379–391. https://doi.org/10.1016/j.molcel.2005.03.031

Shankar, A., Agrawal, N., Sharma, M., Pandey, A., & Pandey, G. K. (2015). Role of Protein Tyrosine Phosphatases in Plants. In *Current Genomics* (Vol. 16).

Sheetz, J. B., & Lemmon, M. A. (2022). Looking lively: emerging principles of pseudokinase signaling. *Trends in Biochemical Sciences*, *47*(10), 875–891. https://doi.org/https://doi.org/10.1016/j.tibs.2022.04.011

Shekhar, S., Pernier, J., & Carlier, M.-F. (2016). Regulators of actin filament barbed ends at a glance. *Journal of Cell Science*, *129*(6), 1085–1091. https://doi.org/10.1242/jcs.179994

Shen, W., Le, S., Li, Y., & Hu, F. (2016). SeqKit: A Cross-Platform and Ultrafast Toolkit for FASTA/Q File Manipulation. *PLOS ONE*, *11*(10), e0163962-. https://doi.org/10.1371/journal.pone.0163962

Sievers, F., Wilm, A., Dineen, D., Gibson, T. J., Karplus, K., Li, W., Lopez, R., McWilliam, H., Remmert, M., Söding, J., Thompson, J. D., & Higgins, D. G. (2011). Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Molecular Systems Biology*, *7*(1), 539. https://doi.org/https://doi.org/10.1038/msb.2011.75

Silva, J. C., Denny, R., Dorschel, C. A., Gorenstein, M., Kass, I. J., Li, G. Z., McKenna, T., Nold, M. J., Richardson, K., Young, P., & Geromanos, S. (2005). Quantitative proteomic analysis by accurate mass retention time pairs. *Analytical Chemistry*, *77*(7), 2187–2200. https://doi.org/10.1021/ac048455k

Soudah, N., Baskin, A., Smorodinsky-Atias, K., Beenstock, J., Ganon, Y., Hayouka, R., Aboraya, M., Livnah, O., Ilouz, R., & Engelberg, D. (2023). A conserved arginine within the αC-helix of Erk1/2 is a latch of autoactivation and of oncogenic capabilities. *Journal of Biological Chemistry*, *299*(9). https://doi.org/10.1016/j.jbc.2023.105072

Sreelatha, A., Nolan, C., Park, B. C., Pawłowski, K., Tomchick, D. R., & Tagliabracci, V. S. (2020). A Legionella effector kinase is activated by host inositol hexakisphosphate. *Journal of Biological Chemistry*, *295*(18), 6214–6224. https://doi.org/https://doi.org/10.1074/jbc.RA120.013067

St Louis, B. M., Quagliato, S. M., & Lee, P. C. (2023). Bacterial effector kinases and strategies to identify their target host substrates. In *Frontiers in Microbiology* (Vol. 14). Frontiers Media S.A. https://doi.org/10.3389/fmicb.2023.1113021

Stanaway, J. D., Parisi, A., Sarkar, K., Blacker, B. F., Reiner, R. C., Hay, S. I., Nixon, M. R., Dolecek, C., James, S. L., Mokdad, A. H., Abebe, G., Ahmadian, E., Alahdab, F., Alemnew, B. T. T., Alipour, V., Bakeshei, F. A., Animut, M. D., Ansari, F., Arabloo, J., … Crump, J. A. (2019). The global burden of non-typhoidal salmonella invasive disease: a systematic analysis for the Global Burden of Disease Study 2017. *The Lancet Infectious Diseases*, *19*(12), 1312–1324. https://doi.org/10.1016/S1473-3099(19)30418-9

Stebbins, C. E., & Galán, J. E. (2000). Modulation of Host Signaling by a Bacterial Mimic: Structure of the Salmonella Effector SptP Bound to Rac1. *Molecular Cell*, *6*(6), 1449–1460. https://doi.org/https://doi.org/10.1016/S1097-2765(00)00141-6

Stebbins, C. E., & Galán, J. E. (2001). Maintenance of an unfolded polypeptide by a cognate chaperone in bacterial type III secretion. *Nature*, *414*(6859), 77–81. https://doi.org/10.1038/35102073

Stryer, L., Tymoczko, J., & Berg, J. (2002). *Biochemistry* (5th ed.). W.H.Freeman.

Svitkina, T. (2018). The actin cytoskeleton and actin-based motility. *Cold Spring Harbor Perspectives in Biology*, *10*(1). https://doi.org/10.1101/cshperspect.a018267

Tamizi, A. A., Abu-Bakar, N., Samsuddin, A. F., Rozano, L., Ahmad-Redzuan, R., & Abdul-Murad, A. M. (2020). Characterisation and mutagenesis study of an alternative sigma factor gene (Hrpl) from erwinia mallotivora reveal its central role in papaya dieback disease. *Biology*, *9*(10), 1–16. https://doi.org/10.3390/biology9100323

Taylor, S. S., & Kornev, A. P. (2011). Protein kinases: evolution of dynamic regulatory proteins. *Trends in Biochemical Sciences*, *36*(2), 65–77. https://doi.org/10.1016/j.tibs.2010.09.006

Thompson, D. K., & Sharkady, S. M. (2021). Genomic insights into drug resistance determinants in cedecea neteri, a rare opportunistic pathogen. *Microorganisms*, *9*(8). https://doi.org/10.3390/microorganisms9081741

Ubersax, J. A., & Ferrell, J. E. (2007). Mechanisms of specificity in protein phosphorylation. In *Nature Reviews Molecular Cell Biology* (Vol. 8, Issue 7, pp. 530–541). Nature Publishing Group. https://doi.org/10.1038/nrm2203

van Kempen, M., Kim, S. S., Tumescheit, C., Mirdita, M., Lee, J., Gilchrist, C. L. M., Söding, J., & Steinegger, M. (2024). Fast and accurate protein structure search with Foldseek. *Nature Biotechnology*, *42*(2), 243–246. https://doi.org/10.1038/s41587-023-01773-0

Wachtel, R., Bräuning, B., Mader, S. L., Ecker, F., Kaila, V. R. I., Groll, M., & Itzen, A. (2018). The protease GtgE from Salmonella exclusively targets inactive Rab GTPases. *Nature Communications*, *9*(1), 44. https://doi.org/10.1038/s41467-017-02110-1

Wagih, O. (2017). Ggseqlogo: A versatile R package for drawing sequence logos. *Bioinformatics*, *33*(22), 3645–3647. https://doi.org/10.1093/bioinformatics/btx469

Walch, P., Selkrig, J., Knodler, L. A., Rettel, M., Stein, F., Fernandez, K., Viéitez, C., Potel, C. M., Scholzen, K., Geyer, M., Rottner, K., Steele-Mortimer, O., Savitski, M. M., Holden, D. W., & Typas, A. (2021). Global mapping of Salmonella enterica-host protein-protein interactions during infection. *Cell Host and Microbe*, *29*(8), 1316-1332.e12. https://doi.org/10.1016/j.chom.2021.06.004

Wang, S., Li, S., Wang, J., Li, Q., Xin, X.-F., Zhou, S., Wang, Y., Li, D., Xu, J., Luo, Z.-Q., He, S. Y., & Sun, W. (2021). A bacterial kinase phosphorylates OSK1 to suppress stomatal immunity in rice. *Nature Communications*, *12*(1), 5479. https://doi.org/10.1038/s41467-021-25748-4

Wang, Y., Sun, M., Bao, H., Zhang, Q., & Guo, D. (2013). Effective Identification of Bacterial Type III Secretion Signals Using Joint Element Features. *PLOS ONE*, *8*(4), e59754-. https://doi.org/10.1371/journal.pone.0059754

Wankowicz, S., Beltrao, P., Cravatt, B., Dunbrack, R., Gitter, A., Lindorff-Larsen, K., Ovchinnikov, S., Polizzi, N., Shoichet, B., & Fraser, J. (2024). AlphaFold3 Transparency and Reproducibility. . *Zenodo*.

Williams, C., Galyov, E. E., & Bagby, S. (2004). Solution Structure, Backbone Dynamics, and Interaction with Cdc42 of Salmonella Guanine Nucleotide Exchange Factor SopE2,. *Biochemistry*, *43*(38), 11998–12008. https://doi.org/10.1021/bi0490744

Winfield, M. D., & Groisman, E. A. (2004). Evolution and Ecology of Salmonella. *EcoSal Plus*, *1*(1), 10.1128/ecosalplus.6.4.6. https://doi.org/10.1128/ecosalplus.6.4.6

Wingfield, P. T. (2017). N-Terminal Methionine Processing. *Current Protocols in Protein Science*, *88*(1), 6.14.1-6.14.3. https://doi.org/https://doi.org/10.1002/cpps.29

World Health Organization. (2024). *WHO bacterial priority pathogens list, 2024: Bacterial pathogens of public health importance to guide research, development and strategies to prevent and control antimicrobial resistance*.

Worley, J., Meng, J., Allard, M. W., Brown, E. W., & Timme, R. E. (2018). Salmonella enterica Phylogeny Based on Whole-Genome Sequencing Reveals Two New Clades and Novel Patterns of Horizontally Acquired Genetic Elements. *MBio*, *9*(6), 10.1128/mbio.02303-18. https://doi.org/10.1128/mbio.02303-18

Xu, C., Kozlov, G., Wong, K., Gehring, K., & Cygler, M. (2016). Crystal Structure of the Salmonella Typhimurium Effector GtgE. *PLOS ONE*, *11*(12), e0166643-. https://doi.org/10.1371/journal.pone.0166643

Xu, Y., Cheng, S., Zeng, H., Zhou, P., Ma, Y., Li, L., Liu, X., Shao, F., & Ding, J. (2022). ARF GTPases activate Salmonella effector SopF to ADP-ribosylate host V-ATPase and inhibit endomembrane damage-induced autophagy. *Nature Structural & Molecular Biology*, *29*(1), 67–77. https://doi.org/10.1038/s41594-021-00710-6

Yaffe, M. B., & Smerdon, S. J. (2001). PhosphoSerine/Threonine Binding Domains: You Can't pSERious? *Structure*, *9*(3), R33–R38. https://doi.org/10.1016/S0969-2126(01)00580-9

Yeon-Gil, K., Jin-Hong, K., & Kyung-Jin, K. (2009). Crystal Structure of the Salmonella enterica Serovar Typhimurium Virulence Factor SrfJ, a Glycoside Hydrolase Family Enzyme. *Journal of Bacteriology*, *191*(21), 6550–6554. https://doi.org/10.1128/jb.00641-09

Yuan, Y., Wang, X., Jin, J., Tang, Z., Xian, W., Zhang, X., Fu, J., He, K., & Liu, X. (2023). The Salmonella Typhimurium Effector SpvB Subverts Host Membrane Trafficking by Targeting Clathrin and AP-1. *Molecular & Cellular Proteomics*, *22*(12), 100674. https://doi.org/https://doi.org/10.1016/j.mcpro.2023.100674

Zanela, T., Woudenberg, A., Romero Bello, K. G., & Underbakke, E. S. (2023). Activation loop phosphorylation tunes conformational dynamics underlying Pyk2 tyrosine kinase activation. *Structure*, *31*(4), 447-454.e5. https://doi.org/10.1016/j.str.2023.02.003

Zhang, M., Jang, H., & Nussinov, R. (2020). Structural Features that Distinguish Inactive and Active PI3K Lipid Kinases. *Journal of Molecular Biology*, *432*(22), 5849–5859. https://doi.org/https://doi.org/10.1016/j.jmb.2020.09.002

Zhang, Y., Higashide, W. M., McCormick, B. A., Chen, J., & Zhou, D. (2006). The inflammation-associated Salmonella SopA is a HECT-like E3 ubiquitin ligase. *Molecular Microbiology*, *62*(3), 786–793. https://doi.org/https://doi.org/10.1111/j.1365-2958.2006.05407.x

Zhou, Z., Alikhan, N. F., Sergeant, M. J., Luhmann, N., Vaz, C., Francisco, A. P., Carriço, J. A., & Achtman, M. (2018). Grapetree: Visualization of core genomic relationships among 100,000 bacterial pathogens. *Genome Research*, *28*(9), 1395–1404. https://doi.org/10.1101/gr.232397.117

Zhu, Y., Li, H., Long, C., Hu, L., Xu, H., Liu, L., Chen, S., Wang, D.-C., & Shao, F. (2007). Structural Insights into the Enzymatic Mechanism of the Pathogenic MAPK Phosphothreonine Lyase. *Molecular Cell*, *28*(5), 899–913. https://doi.org/https://doi.org/10.1016/j.molcel.2007.11.011

Zhukovsky, M. A., Filograna, A., Luini, A., Corda, D., & Valente, C. (2019). Phosphatidic acid in membrane rearrangements. *FEBS Letters*, *593*(17), 2428–2451. https://doi.org/https://doi.org/10.1002/1873-3468.13563

Zouhir, S., Bernal-Bayard, J., Cordero-Alba, M., Cardenal-Muñoz, E., Guimaraes, B., Lazar, N., Ramos-Morales, F., & Nessler, S. (2014). The structure of the Slrp–Trx1 complex sheds light on the autoinhibition mechanism of the type III secretion system effectors of the NEL family. *Biochemical Journal*, *464*(1), 135–144. https://doi.org/10.1042/BJ20140587