Imperial College London

IMPERIAL COLLEGE LONDON

DOCTORAL THESIS

Atomistic Simulations of Cardiac Troponin with Phosphorylation and Mutation Effects

Author:
Zeyu YANG

Supervisors:
Prof. Ian R. GOULD
Prof. Steven B. MARSTON

A thesis submitted in fulfillment of the requirements
for the degree of Doctor of Philosophy
in the
Department of Chemistry

August 9, 2023

Abstract

Cardiac troponin, the regulator of the striated muscle contraction process in sarcomere, plays an important role in the heart's circulation function. Phosphorylation of the cardiac troponin I by protein kinase A decreases troponin's calcium sensitivity and reduces the muscle relaxation time constant under β-adrenergic signalling. This phosphorylation induced lusitropic effect is disrupted by single amino-acid mutations in the thin filament which are the main causes of many genetic cardiomyopathies that can lead to unexpected heart failures or sudden death.

Experimentally, epigallocatechin-3 gallate (EGCG) and its analogues were shown to be able to restore the coupled relationship between Ca^{2+} sensitivity and troponin's phosphorylation states in the diseased mutants. It is therefore of great interest to elucidate how phosphorylation at serines 22 and 23 of troponin I modulates muscle contractility and discover potential treatments for cardiomyopathies caused by mutations. However, this has been difficult because the key regulatory segments of cTn, especially cTnI 1-33, are disordered and absent from the crystal structures. A complete molecular mechanism of phosphorylation dependent troponin dynamics is still unknown owing to the current limitations of experiments where only a partial structure or a static structure can be obtained.

Molecular dynamics (MD) simulation has gained popularity in recent years and the rapid increase in available computational resources has allowed the simulations of a more complete structure of troponin over a longer timescale. This study aims to discover the molecular mechanisms of phosphorylation effects in the wild-type and disease-causing mutations with MD simulation. Here I present the key results discovered from our all-atom MD simulations of the core of the troponin complex including the major changes in troponin's molecular interactions caused by phosphorylation and the TNNC1 G159D mutation. Additional work was conducted to include small molecules in the simulations to study the structural and dynamical changes in troponin due to ligand interactions. Finally, I propose a new method in analysing residue interaction networks based on the concept of Granger causality.

Statement of Originality

I hereby confirm that everything presented in this thesis is the product of my own work unless otherwise stated. Where I have consulted or used any material from the work of others, these have been clearly cited and referenced. I certify that no parts of this thesis has been previously submitted for any other academic degrees.

Declaration of Copyright

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.

For my 好婆

Acknowledgement

I am grateful to Prof. Ian Gould for introducing me to the wonderful world of MD simulations, and for his constant guidance and trust. I would like to thank Prof. Steven Marston for his intellectual discussions and providing the experimental data which we rely on. Gratitude is also due to previous group members, Dr Juan Zamora and Michael Cognet, for laying the cornerstone of this project, and other past and present group members for providing support and help when needed, namely Hisham, Sara, Zhili, and Astrid. Additionally, Prof. Phil Parsons and Dr Lewis Allen have been my guiding stars in my research journey, and without a doubt, I would not have completed this work without their support.

I am indebted to the open source community for providing the tools and support that this project depends on. You showed me what is achievable by fostering an open and welcoming community. Although late to the club, ICMC has welcomed me with open arms and provided the much needed trips to crags in the UK and beyond, thank you Miguel, Amy, Sofia, Wilson, Tom, and Spencer. Acknowledgement is also owed to my friends and colleagues from Arabesque (Matthias and Ngoc) and Ro5. All of you were influential in building my professional career.

Finally, it goes without saying that my parents and Josie all supported me immensely over the past few year. It has been a journey and I could not have done it without all of your support.

Table of contents

Abstract 1

Statement of Originality 2

Copyright Declaration 3

Acknowledgement 5

Table of contents 6

List of Figures 13

List of Tables 21

Abbreviation 25

1 Introduction 27

1.1 The Anatomy of Cardiac Muscles 28

1.2 The Role of Cardiac Troponin 32

1.3 The Structure of Troponin 33

1.3.1 Troponin C 35

1.3.2 Troponin I 37

1.3.3 Troponin T 38

1.4 Post-translational Modifications on Troponin 40

1.5 Troponin Mutation and Genetic Cardiomyopathies . . . 41
1.5.1 The TNNC1 G159D Mutation . . . 43
1.6 Therapeutic Small Molecules for Cardiomyopathies . . . 45
1.7 Computational Simulations of Cardiac Muscle . . . 48
1.8 Thesis Outline . . . 50

2 Computational Methods **51**

2.1 Molecular Dynamics . . . 51
2.1.1 AMBER force-field . . . 52
2.1.2 Velocity Verlet Algorithm . . . 52
2.1.3 Molecular Dynamics Simulation Protocol . . . 53
2.1.4 Ligand Parameterisation . . . 55
2.1.5 Data validation . . . 56
2.1.6 Representative Structures . . . 56
2.1.7 Binding Region Clustering . . . 56
2.2 Binding Energy Estimation . . . 57
2.3 Residue Flexibility . . . 57
2.4 Interhelical Angle Analysis . . . 58
2.5 Hinge Angle Analysis . . . 58
2.6 Effect Size . . . 59
2.7 Atomic Interaction Profiling . . . 60

2.8 Granger Causality Analysis 61

3 Effects of Phosphorylation and Mutation on Troponin 63

3.1 Disordered Regions Have Higher Flexibility 64

3.2 Global Structure Changes from Principal Component Analysis 66

3.3 Changes in Regional Structural Features 67

3.3.1 Interdomain Hinge Angle 67

3.3.2 Contraction of The Regulatory Domain 69

3.3.3 Interhelical A-B Angle 71

3.4 Differences in Peptide Binding Affinity 74

3.4.1 Switch Peptide 75

3.4.2 Inhibitory Peptide 77

3.4.3 TnI_{34-71} 77

3.4.4 NcTnI 78

3.5 Residue Interaction Profiling 79

3.5.1 CcTnC and TnI Anchor Region 80

3.5.2 Switch Peptide Region 81

3.5.3 Inhibitory Peptide Region 82

3.5.4 Catalytic Calcium Binding Site with TnI 83

3.5.5 NcTnC and NcTnI 85

3.5.6 TnC and TnT Interface 86

3.5.7 Phosphorylation Induced Changes . . . 88
3.5.8 Mutation Induced Changes . . . 91
3.5.9 TnC self-interactions . . . 92
3.5.10 IT Arm Interface . . . 93
3.6 Granger Causality Analysis . . . 94
3.6.1 Phosphorylation Induced Perturbations in the Wild Type . . . 94
3.6.2 Mutation Induced Perturbations . . . 96
3.6.3 Molecular Interactions in G159D SEP . . . 98
3.6.4 Limitations of the Granger Causality Analysis . . . 100
3.7 Discussion . . . 101
3.7.1 Phosphorylation and Mutation Lead to Rigidification . . . 101
3.7.2 Mutation Increases Inhibitory Peptide Binding Energy . . . 101
3.7.3 Mutation and Phosphorylation Change TnC/TnI Anchoring Stability . . . 102
3.7.4 A-B Angle Dynamics are Linked to Interactions Between TnI and TnC . . . 103
3.7.5 Phosphorylation and Mutation Cause Reposition of NcTnI . . . 103
3.8 Conclusion . . . 104

4 Study of EGCG, SilybinA, SilybinB, and Resveratrol 105

4.1 Global Structure Changes from Principal Component Analysis . . . 108
4.2 Changes in Regional Structural Features . . . 110
4.2.1 Interhelical A-B Angle . . . 110

4.2.2 Interdomain Hinge Angle . . . 113
4.2.3 TnC Domain Distance . . . 116
4.3 Changes in Peptide Binding Affinity . . . 118
4.3.1 Switch Peptide . . . 119
4.3.2 Inhibitory Peptide . . . 120
4.3.3 TnI_{34-71} . . . 121
4.3.4 NcTnI . . . 123
4.4 Change in Residue Interactions . . . 124
4.5 Binding Region Clustering . . . 126
4.5.1 EGCG . . . 127
4.5.2 RVL . . . 129
4.5.3 Silybin A . . . 131
4.5.4 Silybin B . . . 133
4.6 Ligand binding pattern . . . 135
4.6.1 EGCG . . . 135
4.6.2 RVL . . . 136
4.6.3 Silybin A . . . 136
4.6.4 Silybin B . . . 141
4.7 Ligand Binding Energy . . . 141
4.8 Discussion . . . 148
4.8.1 Recouplers Restored Key Metrics . . . 148

4.8.2 Common Binding Regions Present in All Simulated Systems 149

4.8.3 Ligand Binding Disrupts Ionic Interactions from the mutation site 151

5 Study of ECG, DHSA, DHSB, and r07 152

5.1 Global Structure Changes from Principal Component Analysis 154

5.2 Changes in Regional Structural Features 156

5.2.1 Interhelical A-B Angle 156

5.2.2 Interdomain Hinge Angle 159

5.2.3 TnC Domain Distance 162

5.3 Peptide Binding Energy 164

5.3.1 Switch Peptide 164

5.3.2 Inhibitory Peptide 166

5.3.3 TnI_{34-71} 167

5.3.4 NcTnI 168

5.4 Change in Residue Interactions 169

5.5 Binding Region Clustering 171

5.5.1 DHSA 171

5.5.2 DHSB 173

5.5.3 r07R 176

5.5.4 r07S 178

5.5.5 ECG 178

5.6 Ligand binding pattern . . . 180
5.6.1 DHSA . . . 180
5.6.2 DHSB . . . 181
5.6.3 r07R . . . 181
5.6.4 r07S . . . 188
5.6.5 ECG . . . 188
5.7 Ligand Binding Energy . . . 193
5.8 Discussion . . . 195
5.8.1 Changes in Key Metrics with Phosphorylation . . . 195
5.8.2 Ligand Binding Regions . . . 196

6 Conclusion **198**

6.1 Future Work . . . 199
6.1.1 A More Complete Model . . . 199
6.1.2 Further Development on Causal Inference Methods . . . 199

7 Appendix **200**

7.1 Chapter 2 . . . 200
7.2 Chapter 3 . . . 201

References **223**

List of Figures

1 Schematic illustration of the cardiac muscle showing that cardiomyocytes consist of myofibrils. Each strand of myofibril consists of regularly arranged sarcomeres, and sarcoplasmic reticulum. Created with BioRender.com. 29

2 Cross-bridge cycle showcases the interactions between thin filament and thick filament that generates force in the expense of ATP. The sarcoplasmic calcium ions play an important role in this process. Created with BioRender.com. 30

3 With cross-bridge cycles, the thin and thick filaments slide along each other and shorten the sarcomere length which is the basis of muscle contraction. Created with BioRender.com. 31

4 Illustrations of thin filament and troponin subunits. 32

5 The cTnC molecule with its nine helices labelled. 35

6 The cTnI molecule with relevant structural regions labelled. Residues 172 to 210 of cTnI are not present in this model. 37

7 The cTnT molecule with relevant structural regions labelled. Residues 1 to 201 of cTnI are not present in this model. 39

8 Measurements of Ca^{2+}-activated development of steady-state tension. Left graph shows that Tn containing the TnC G159D mutation does not alter trabeculae function without phosphorylation. Right graph compares pseudo-phosphorylated WT and mutant which shows that exchange of Tn containing the cTnC G159D and Ser22/23 pseudo-phosphorylated cTnI into trabeculae blunts the response of cTnI phosphorylation. Reproduced with permission.[48] 44

9 The four proposed functional states of troponin.[67] The wild-type troponin transitions from the high affinity state to the low affinity state upon phosphorylation while the pathway is blocked with mutantions. By stabilising the phosphorylated low affinity state, small molecules can recouple troponin mutants. ©2018 Sheehan, Messer, Papadaki, Choudhry, Kren, Biedermann, Blagg, Khandelwal and Marston. Licensed under the Creative Commons Attribution License (CC BY). 48

10 An illustration of the interhelical A/B angle on troponin. The three subunits of troponin are shown in white (TnC), light green (TnI), and light pink (TnT). Helix A and B of TnC are coloured in ligh blue. The two yellow arrows represent the vectors used to calculate the angle. 58

11 An illustration of the hinge angle on troponin. The two arrows represent the vectors used to calculate the angle. 59

12 Root mean squared fluctuation for TnC. Shaded area represents the 95% confidence interval. 64

13 Root mean squared fluctuation for TnI. Shaded area represents the 95% confidence interval. 65

14 Root mean squared fluctuation for TnT. Shaded area represents the 95% confidence interval. 66

15 Scatter plot of troponin conformations in the first two principal components. . 68

16 Depiction of troponin structures at the two diagonal ends (bottom left and top right) of the PCA plot. The structure coloured in blue is from WT uP simulations and represents the low PC1 and low PC2 state. The structure coloured in red is from G159D SEP simulations and represents the high PC1 and high PC2 state. . 69

17 The hinge angle distributions as KDE plots. 70

18 The TnC domain distance distributions as KDE plots. 71

19 The interhelical A-B angle distributions for all apo systems. 72

20 Illustration of the TnI peptide segments on troponin. The key segments are switch peptide (red), inhibitory peptide (green), TnI_{34-71} (cyan), and NcTnI (magenta). 75

21 Depiction of the charged residues near the catalytic calcium binding site that interacted with TnI Lys35. TnI subunit is coloured in beige and TnC subunit is coloured in blue. 86

22 Heatmap showing the difference in hydrogen bond interaction percentage upon phosphorylation for WT and SEP. The graph focuses on the NcTnI and C-terminal end of TnI of the model used in this work. 88

23 Granger causality inferred network for direct ionic and hydrogen bond interactions between residues for WT SEP. This network shows the perturbation effects of additional interactions introduced by the phosphoserines. 95

24 Granger causality inferred network for direct ionic and hydrogen bond interactions between residues for G159D uP. This network shows the perturbation effects of additional interactions introduced by the mutated residue. 97

25 Granger causality inferred network for direct ionic and hydrogen bond interactions between residues for G159D SEP. This network shows the perturbation effects of additional interactions introduced by TnC Asp159 and TnI Sep22/23. 99

26 Chemical structure of EGCG, RVL, silybin A, and silybin B. 106

27 Scatter plot of troponin conformations in the apo systems' first two principal components. The faded circles represent conformations from the apo system. The coloured stars represent the ligand systems. 108

28 The interhelical angle distributions for EGCG, RVL, silybin A, and silybin B systems. 110

29 The interdomain angle distributions for EGCG, RVL, silybin A, and silybin B systems. . . . 113

30 The TnC domain distance distributions for the systems with EGCG, RVL, silybin A, and silybin B. . . . 116

31 Binding site clusters for systems with EGCG. . . . 128

32 Binding site clusters for systems with RVL. . . . 130

33 Binding site clusters for systems with Silybin A 132

34 Binding site clusters for systems with Silybin B 134

35 Interaction counts of EGCG with residues in WT systems. . . . 137

36 Interaction counts of EGCG with residues in G159D systems. . . . 138

37 Interaction counts of RVL with residues in WT systems. . . . 139

38 Interaction counts of RVL with residues in G159D systems. . . . 140

39 Interaction counts of silybin A with residues in WT systems. . . . 142

40 Interaction counts of silybin A with residues in G159D systems. . . . 143

41 Interaction counts of silybin B with residues in WT systems. . . . 144

42 Interaction counts of silybin B with residues in G159D systems. . . . 145

43 Comparison of changes in key metrics upon phosphorylation in ligand systems with apo systems. . . . 150

44 Chemical structure of DHSA, DHSB, r07R, r07S, and ECG. . . . 152

45 Scatter plot of troponin conformations in the apo systems' first two principal components. The faded circles represent conformations from the apo system. The coloured stars represent the ligand systems. . . . 155

46 The interhelical angle distributions for DHSA, DHSB, r07R, r07S, and ECG systems. Compare to apo figures in Figure 19. 156

47 The interdomain angle distributions for DHSA, DHSB, r07R, r07S, and ECG systems. Compare to apo figures in Figure 17. 159

48 The TnC domain distance distributions for the systems with DHSA, DHSB, r07R, r07S, and ECG. Compare to apo figures in Figure 18. 162

49 Binding site clusters for systems with DHSA. 172

50 Binding site clusters for systems with DHSB. 174

51 Binding site clusters for systems with r07R. 175

52 Binding site clusters for systems with r07S. 177

53 Binding site clusters for systems with ECG. 179

54 Interaction counts of DHSA with residues in WT systems. 182

55 Interaction counts of DHSA with residues in G159D systems. 183

56 Interaction counts of DHSB with residues in WT systems. 184

57 Interaction counts of DHSB with residues in G159D systems. 185

58 Interaction counts of r07R with residues in WT systems. 186

59 Interaction counts of r07R with residues in G159D systems. 187

60 Interaction counts of r07S with residues in WT systems. 189

61 Interaction counts of r07S with residues in G159D systems. 190

62 Interaction counts of ECG with residues in WT systems. 191

63 Interaction counts of ECG with residues in G159D systems. 192

64 Comparison of changes in key metrics upon phosphorylation in ligand systems with apo systems. 197

S1 Pairwise protein backbone root-mean-square deviation (RMSD) between all frames of each trajectory for all apo systems. No obvious burn-in period was observed. 200

S2 Distributions of simulated conformations for the apo systems along the first ten principal components as KDE plots. 202

S3 Heatmap showing the ionic interaction pattern between TnC and TnI for WT uP apo system. The interaction count is expressed as a percentage of the total number of frames. 203

S4 Heatmap showing the ionic interaction pattern between TnC and TnI for WT SEP apo system. The interaction count is expressed as a percentage of the total number of frames. 204

S5 Heatmap showing the ionic interaction pattern between TnC and TnI for G159D uP apo system. The interaction count is expressed as a percentage of the total number of frames. 205

S6 Heatmap showing the ionic interaction pattern between TnC and TnI for G159D SEP apo system. The interaction count is expressed as a percentage of the total number of frames. 206

S7 Heatmap showing the ionic interaction pattern between TnC and TnT for WT uP apo system. The interaction count is expressed as a percentage of the total number of frames. 207

S8 Heatmap showing the ionic interaction pattern between TnC and TnT for WT SEP apo system. The interaction count is expressed as a percentage of the total number of frames. 208

S9 Heatmap showing the ionic interaction pattern between TnC and TnT for G159D uP apo system. The interaction count is expressed as a percentage of the total number of frames. 209

S10 Heatmap showing the ionic interaction pattern between TnC and TnT for G159D SEP apo system. The interaction count is expressed as a percentage of the total number of frames. 210

S11 Heatmap showing the ionic interaction pattern between TnC and TnC for WT uP apo system. The interaction count is expressed as a percentage of the total number of frames. 211

S12 Heatmap showing the ionic interaction pattern between TnC and TnC for WT SEP apo system. The interaction count is expressed as a percentage of the total number of frames. 212

S13 Heatmap showing the ionic interaction pattern between TnC and TnC for G159D uP apo system. The interaction count is expressed as a percentage of the total number of frames. 213

S14 Heatmap showing the ionic interaction pattern between TnC and TnC for G159D SEP apo system. The interaction count is expressed as a percentage of the total number of frames. 214

S15 Heatmap showing the ionic interaction pattern between TnI and TnI for WT uP apo system. The interaction count is expressed as a percentage of the total number of frames. 215

S16 Heatmap showing the ionic interaction pattern between TnI and TnI for WT SEP apo system. The interaction count is expressed as a percentage of the total number of frames. 216

S17 Heatmap showing the ionic interaction pattern between TnI and TnI for G159D uP apo system. The interaction count is expressed as a percentage of the total number of frames. 217

S18 Heatmap showing the ionic interaction pattern between TnI and TnI for G159D SEP apo system. The interaction count is expressed as a percentage of the total number of frames. 218

S19 Heatmap showing the ionic interaction pattern between TnT and TnI for WT uP apo system. The interaction count is expressed as a percentage of the total number of frames. 219

S20 Heatmap showing the ionic interaction pattern between TnT and TnI for WT SEP apo system. The interaction count is expressed as a percentage of the total number of frames. 220

S21 Heatmap showing the ionic interaction pattern between TnT and TnI for G159D uP apo system. The interaction count is expressed as a percentage of the total number of frames. 221

S22 Heatmap showing the ionic interaction pattern between TnT and TnI for G159D SEP apo system. The interaction count is expressed as a percentage of the total number of frames. 222

List of Tables

1 The standard nomenclature of the different structural regions of Troponin C. . . 36

2 The standard nomenclature of the different structural regions of Troponin I. . . 38

3 The standard nomenclature of the different structural regions of Troponin T. . . 40

4 The hinge angle metrics for the four apo systems. 69

5 The TnC domain distance metrics for the four apo systems. 70

6 The interhelical angle metrics for the four apo systems. 73

7 The interhelical A-B angle pairwise effect sizes for different systems. 74

8 Switch peptide binding energy to the rest of the troponin complex for the apo systems. Unit: $\mathrm{kcal\,mol^{-1}}$ 76

9 Inhibitory peptide binding energy to the rest of the troponin complex for the apo systems. Unit: $\mathrm{kcal\,mol^{-1}}$ 77

10 $\mathrm{TnI_{34-71}}$ binding energy to the rest of the troponin complex for the apo systems. Unit: $\mathrm{kcal\,mol^{-1}}$ 77

11 Binding energy of NcTnI to the rest of troponin for the apo systems. Unit: $\mathrm{kcal\,mol^{-1}}$ 78

12 Ionic interactions at the CcTnC and TnI interface that anchored TnC to TnI. . . 81

13 Ionic interactions at the hydrophobic patch and switch peptide region. 82

14 Ionic interactions between inhibitory peptide and TnC. 84

15 Ionic interactions between TnI and the catalytic calcium binding site. 85

16 Ionic interactions between NcTnI and NcTnC. 87

17 Ionic interactions at the TnC and TnT interface. 87

18 Hydrogen bonding between NcTnI and TnI_{IP}. 89

19 Ionic interactions from TnI phosphoserines 22 and 23. 89

20 Ionic interactions formed from TnC Asp159 in the mutant. 91

21 Ionic interactions formed within TnC. 93

22 The interhelical angle metrics for the systems with EGCG, RVL, silybin A, and silybin B. 111

23 The interdomain angle metrics for the systems with EGCG, RVL, silybin A, and silybin B. 114

24 The TnC domain distance metrics for the systems with EGCG, RVL, silybin A, and silybin B. 117

25 Switch peptide binding energy to the rest of the troponin complex for troponin systems with EGCG, RVL, silybin A, and silybin B. Unit: $\mathrm{kcal\,mol^{-1}}$ 119

26 Inhibitory peptide binding energy to the rest of the troponin complex for troponin systems with EGCG, RVL, silybin A, and silybin B. Unit: $\mathrm{kcal\,mol^{-1}}$ 120

27 TnI_{34-71} peptide binding energy to the rest of the troponin complex for troponin systems with EGCG, RVL, silybin A, and silybin B. Unit: $\mathrm{kcal\,mol^{-1}}$ 121

28 Binding energy of NcTnI to the rest of troponin for troponin systems with EGCG, RVL, silybin A, and silybin B. Unit: $\mathrm{kcal\,mol^{-1}}$ 123

29 Ionic interaction percentages between TnC Asp159 and TnC residues in simulations with EGCG, RVL, silybin A, and silybin B. 125

30 Ionic interaction percentages between TnC Asp159 and TnI residues in simulations with EGCG, RVL, silybin A, and silybin B. 126

31 Binding region cluster percentage for EGCG systems. 127

32 Binding region cluster percentages for RVL systems. 129

33 Binding region cluster percentages for Silybin A systems. 131

34 Binding region cluster percentages for silybin B systems. 133

35 Mean ligand binding energy in troponin systems with EGCG, RVL, silybin A, and silybin B. Unit: $\mathrm{kcal\,mol^{-1}}$ 146

36 Ligand binding energy in troponin systems with EGCG, RVL, silybin A, and silybin B separated by clusters. Unit: $\mathrm{kcal\,mol^{-1}}$ 147

37 The interhelical angle metrics for the systems with DHSA, DHSB, r07R, r07S, and ECG. 158

38 The interdomain angle metrics for the systems with DHSA, DHSB, r07R, r07S, and ECG. 161

39 The TnC domain distance metrics for the systems with DHSA, DHSB, r07R, r07S, and ECG. 163

40 Switch peptide binding energy to the rest of the troponin complex for troponin systems with DHSA, DHSB, r07R, r07S, and ECG. Unit: $\mathrm{kcal\,mol^{-1}}$ 165

41 Inhibitory peptide binding energy to the rest of the troponin complex for troponin systems with DHSA, DHSB, r07R, r07S, and ECG. Unit: $\mathrm{kcal\,mol^{-1}}$ 166

42 $\mathrm{TnI_{34-71}}$ peptide binding energy to the rest of the troponin complex for troponin systems with DHSA, DHSB, r07R, r07S, and ECG. Unit: $\mathrm{kcal\,mol^{-1}}$ 167

43 Binding energy of NcTnI to the rest of troponin for troponin systems with DHSA, DHSB, r07R, r07S, and ECG. Unit: $\mathrm{kcal\,mol^{-1}}$ 168

44 Ionic interaction percentages between TnC Asp159 and TnC residues in simulations with DHSA, DHSB, r07R, r07S, and ECG. 170

45 Ionic interaction percentages between TnC Asp159 and TnI residues in simulations with DHSA, DHSB, r07R, r07S, and ECG. 171

46 Binding region cluster percentage for DHSA systems. 173

47 Binding region cluster percentage for DHSB systems. 176

48 Binding region cluster percentage for r07R systems. 176

49 Binding region cluster percentage for r07S systems. 178

50 Binding region cluster percentage for ECG systems. 180

51 Mean ligand binding energy in troponin systems with DHSA, DHSB, r07R, r07S, and ECG. Unit: $\mathrm{kcal\,mol^{-1}}$ 193

52 Ligand binding energy in troponin systems with DHSA, DHSB, r07R, r07S, and ECG separated by clusters. Unit: $\mathrm{kcal\,mol^{-1}}$ 194

S1 Ionic interactions between TnT and TnI. 201

Abbreviation

ARVC	Arrhythmogenic Right Ventricular Cardiomyopathy
cTn	cardiac troponin
CUDA	Compute Unified Device Architecture
DCM	Dilated Cardiomyopathy
EGCG	Epigallocatechin-3 gallate
ECG	Epicatechin gallate
HCM	Hypertrophic Cardiomyopathy
IVMA	*in vitro* motility assay
MD	Molecular Dynamics
MM	Molecular Mechanics
MMPBSA	Molecular Mechanics Poisson Boltzmann and Surface Area
PKA	Protein Kinase A
PLIP	Protein–Ligand Interaction Profiler
PMEMD	Particle Mesh Ewald Molecular Dynamics
KDE	Kernel Density Estimation
PCA	Principal Component Analysis
RCM	Restrictive Cardiomyopathy
RESP	Restrained ElectroStatic Potential
RMSF	Root Mean Square Fluctuation

RMSD	Root Mean Square Deviation
Tm	Tropomyosin
Tn	Troponin
TnC	Troponin C
TnI	Troponin I
TnT	Troponin T
TNNC1	Troponin C type-1 gene
TPM1	Tropomyosin alpha-1 gene
SR	Sarcoplasmic Reticulum
WT	wild type

1 Introduction

Troponin (Tn) plays an essential role in the skeletal and cardiac muscle contraction process which is mediated by sarcomeric Ca^{2+} concentration. Troponin forms a complex with actin and tropomyosin (Tm) on the thin filament and regulates the myosin-actin binding process by inhibiting the myosin binding site on actin at low sarcoplasmic Ca^{2+}-concentration.[1] Upon binding with Ca^{2+} ion, the troponin-tropomyosin complex undergoes a conformational change and exposes the myosin binding site on actin, and initiates the muscle contraction process. Additionally, cardiac troponin (cTn) is phosphorylated during periods of increased physical activity or stress which reduces its Ca^{2+} affinity and increases heart muscle relaxation rate. Mutations of cTn change the dynamics of the molecular switch and disrupt the phosphorylation-dependent regulatory mechanism which eventually lead to genetic cardiomyopathies. Although the effects of phosphorylation and mutations of troponin are well-studied through *in vitro* and *in vivo* experiments, the detailed molecular mechanism of action is still lacking. In this project, I aim to shine more light on the structural effects of post-translational modification as well as mutation through all-atom molecular dynamics simulations.

This introductory chapter will guide the reader though the general background in understanding cardiac troponin's functional importance and the approaches I have taken to study it. The chapter is organised as the following sections:

- Sections 1.1 to 1.4: the structure of cardiac troponin and its role in the muscle contraction process
- Section 1.5: cardiomyopathy and the disease-causing mutation studied in this work
- Section 1.6: the therapeutic small molecules for treating abnormal heart muscles
- Section 1.7: recent advances in computational study of cardiac muscles

1.1 The Anatomy of Cardiac Muscles

The heart contracts and relaxes to pump blood around the circulation system and the cardiac muscles generate the contractile force that causes the propulsion of blood. Cardiomyocytes, the heart muscle cells, consist of a high density of contraction organelles, myofibrils (Figure 1). On the molecular level, cardiac muscle consists of regularly arranged actin and myosin which form a striated pattern. A force generation unit in the muscle fibre is called sarcomere which makes up myofibrils, and consists of myosin based thick filaments and actin based thin filaments that interact to generate molecular motions while consuming ATP via hydrolysis. The thick filaments are made up from titin and motor protein myosin. The thin filament consist of G-actins, a dimer strand of tropomyosin, and the troponin ternary complex in 7:1:1 stoichiometry.[2] Tropomyosin is an alpha-helical, coiled-coil protein that forms a head-to-tail polymeric cable wrapped helically around the central actin core of the thin filament. The interaction between thin filament and thick filament is modulated by both cytoplasmic Ca^{2+} binding to troponin in the thin filament as well as post-translational modifications of key regulatory proteins, including the phosphorylation of troponin I.

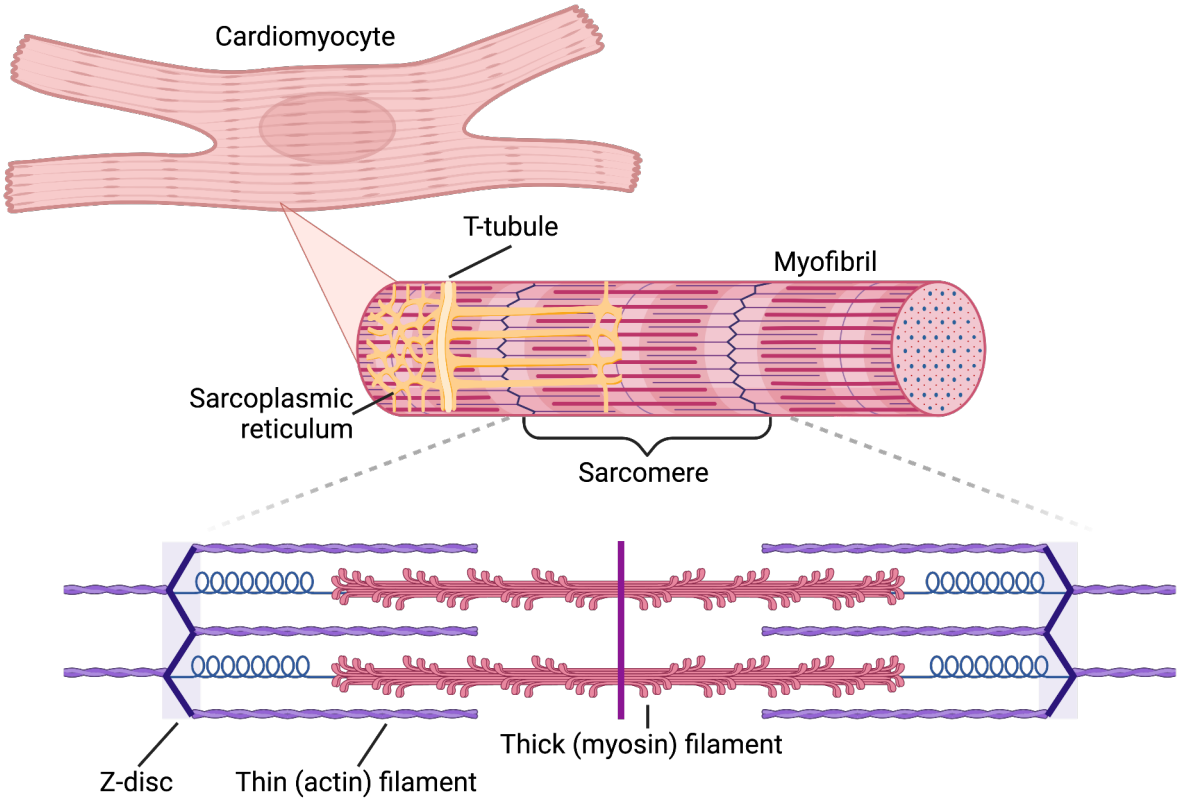


Figure 1: Schematic illustration of the cardiac muscle showing that cardiomyocytes consist of myofibrils. Each strand of myofibril consists of regularly arranged sarcomeres, and sarcoplasmic reticulum. Created with BioRender.com.

Sarcoplasmic reticulum (SR) is the smooth endoplasmic reticulum in the muscle cells and responsible for the storage and modulation of cytoplasmic Ca^{2+} concentration in cardiomyocytes.[3] During the excitation-contraction coupling process, action potentials conduct along the sarcolemma and into the transverse tubules (T-tubules) located between SRs.[4] L-type calcium channels, part of the voltage-dependent dihydropyridine (DHP) receptors, are opened by the depolarisation event to allow calcium ions into the cell. The Ca^{2+} influx subsequently triggers the release of calcium ions stored in the SR through ryanodine receptors (RyR), termed the calcium-induced calcium release (CICR), and this increases intracellular calcium concentration.[5]

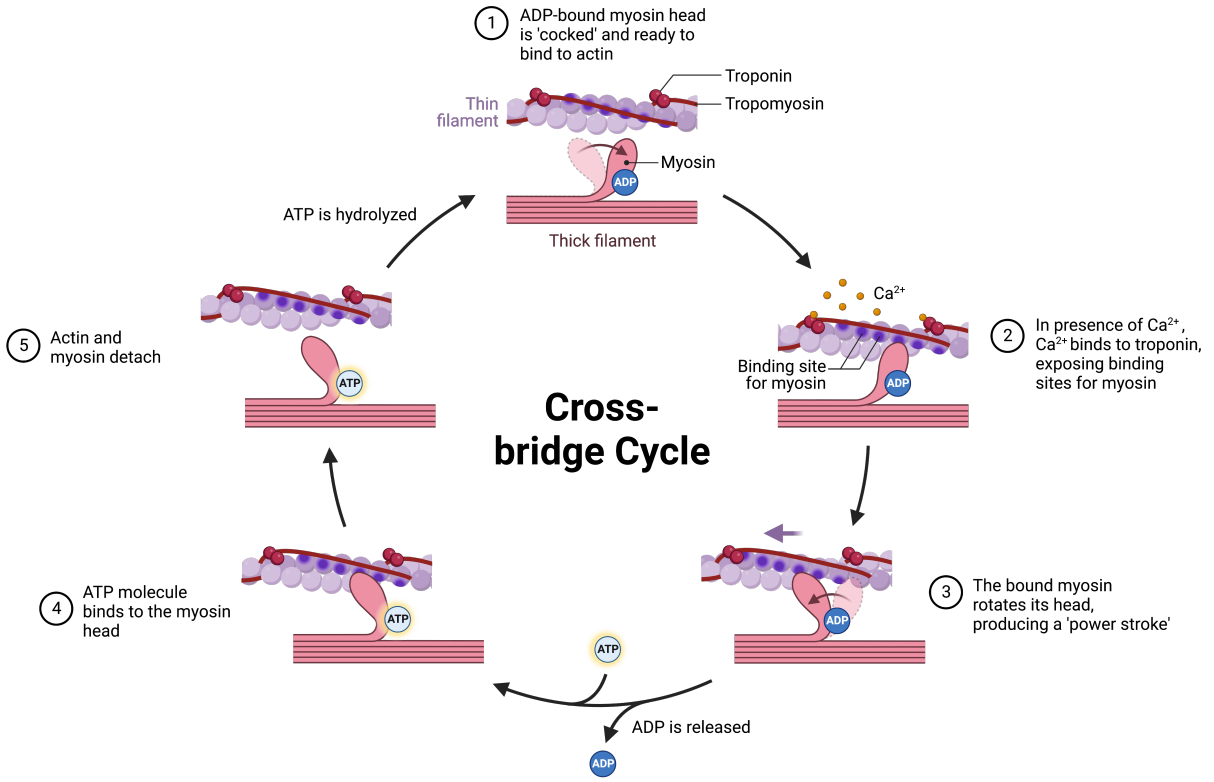


Figure 2: Cross-bridge cycle showcases the interactions between thin filament and thick filament that generates force in the expense of ATP. The sarcoplasmic calcium ions play an important role in this process. Created with BioRender.com.

Muscle contraction begins when actin is pulled by myosin. This occurs when the myosin binding sites are exposed in response to the binding event between free calcium ions and troponin C, which introduces a conformational shift in the troponin-tropomyosin complex that shields the myosin binding sites under low sarcoplasmic Ca^{2+}-concentration (Figure 2). After the myosin head power stroke motion as part of the cross-bridge cycle, the thin and thick filaments slide along each other hence shorten the sarcomere length and contract muscle fibres (Figure 3). The cross-bridge cycle proceeds under elevated sarcoplasmic Ca^{2+}-concentration. When cardiomyocyte repolarises, the L-type calcium channels close and sarcoplasmic calcium ions are pumped back into SR by Ca^{2+}-ATPase. Calcium ion dessociates from troponin C with reduced Ca^{2+}-concentration and restores the troponin-tropomyosin conformation to block the myosin binding site hence preventing cross-bridge formation thus relaxes the muscle.

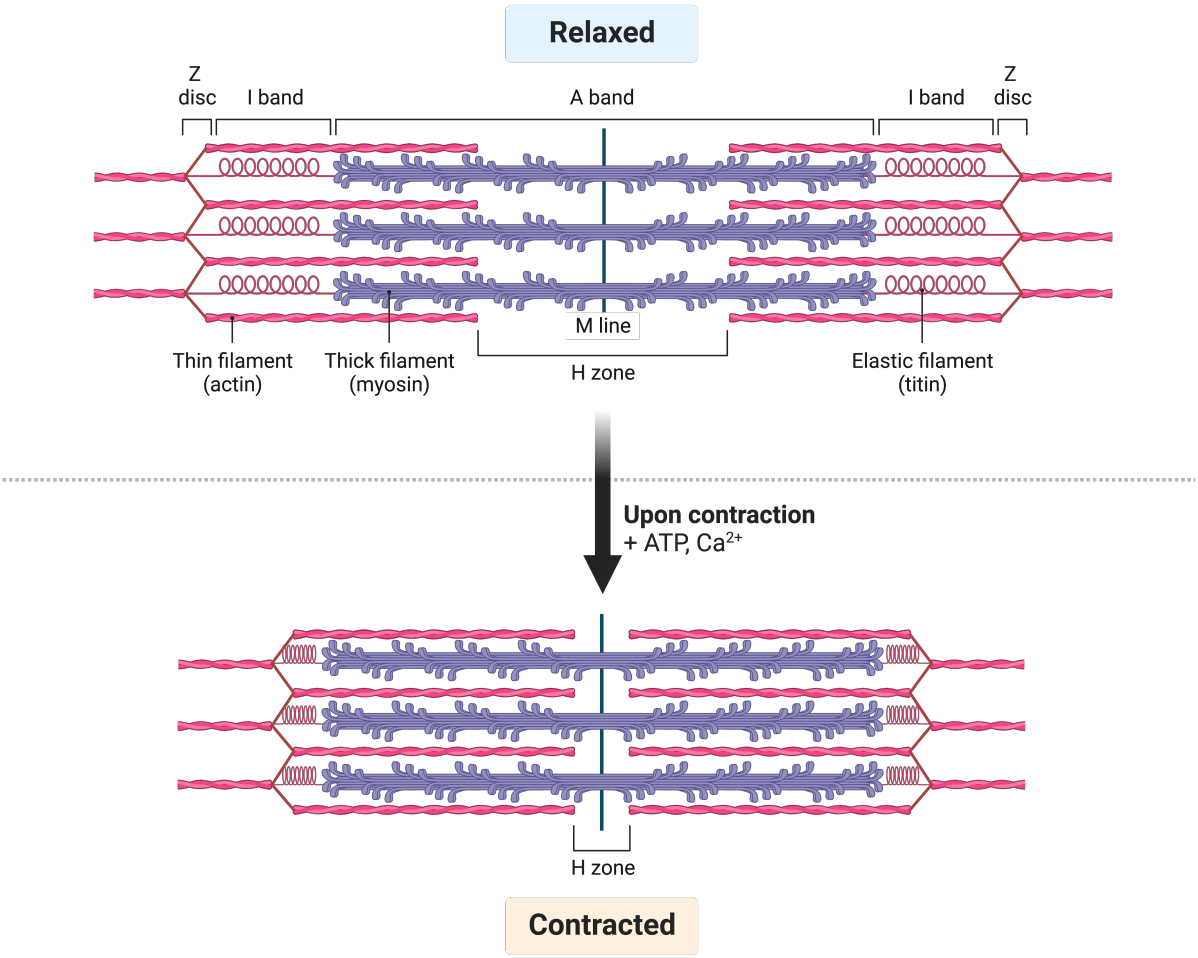
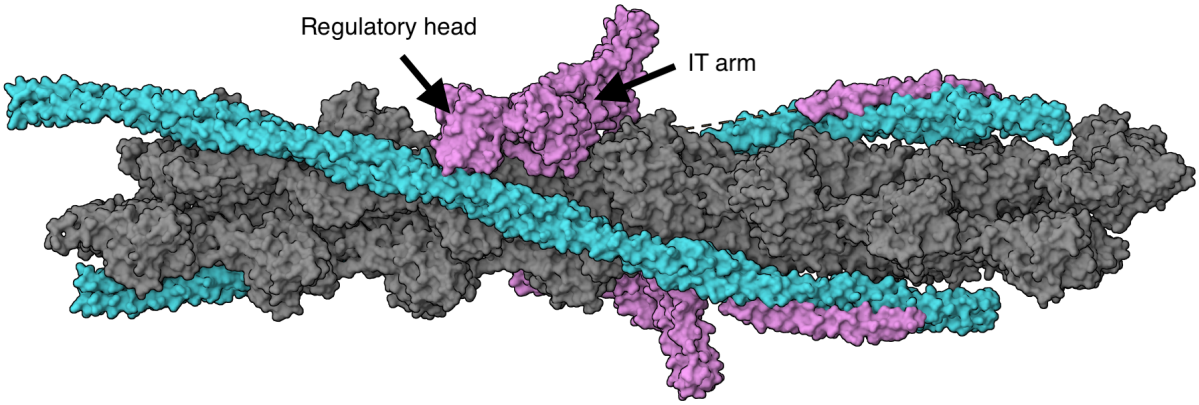
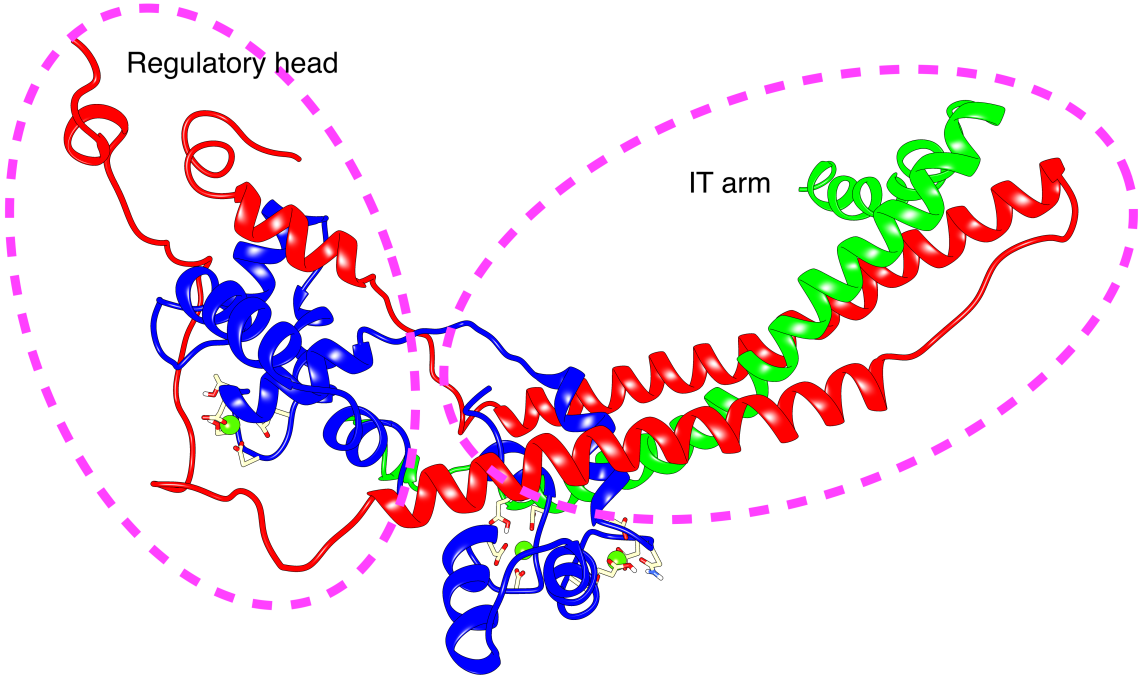


Figure 3: With cross-bridge cycles, the thin and thick filaments slide along each other and shorten the sarcomere length which is the basis of muscle contraction. Created with BioRender.com.

1.2 The Role of Cardiac Troponin



(a) Structure of a human cardiac thin filament in the calcium bound state from cryo-EM. The thin filaments include actin (grey), tropomyosin (blue), and troponin (pink). PDB ID: 6KN8.



(b) A model of the troponin core with the regulatory head and the IT arm labelled. Troponin consists of TnC (blue), TnI (red), and TnT (green).

Figure 4: Illustrations of thin filament and troponin subunits.

Troponin functions as a molecular switch that is activated by Ca^{2+} to undergo a sequence of conformational changes which propogates to the thin filament and reveals the myosin binding site on actin. The regulatory head of troponin (Figure 4) is the Ca^{2+}-sensing part and consists of the N-terminal lobe of cTnC, NcTnI, the switch peptide and inhibitory peptide. At

the resting state, the calcium binding site (EF-hand II) on the N-terminal regulatory domain is unoccupied and the cTnI subunit inhibits the myosin binding site on the actin groove together with tropomyosin.[6,7] The muscle contraction process starts when sarcoplasmic reticulum releases Ca^{2+} through voltage-gated calcium channels. During cardiac contraction, the cytosolic Ca^{2+} concentration fluctuates between 100 nM in diastole and 400-1000 nM during systole.[8,9] With increased Ca^{2+} concentration, the empty EF-hand II becomes populated with one Ca^{2+} ion and motivate the first EF-hand to change from a closed conformation to open, hence exposes the hydrophobic patch.[10] The opening event is characterised by the interhelical angle between helix A and B of EF-hand I to shift from greater than 110° to less than 100°. Note that the angle is reported as the supplement of the interhelical angle *between* the helices as detailed in Section 2.4. The hydrophobic patch, formed by helix A and B of cTnC, binds to the switch peptide on the cTnI and pulls away the C-terminal domain of cTnI from the actin groove. Subsequently, tropomyosin shifts into the actin groove and moves from the blocked state (B) to the closed state (C), therefore partially exposing the myosin binding sites. The binding of myosin to the exposed binding sites on actin leads to a further translocation of tropomyosin to the open state (M) hence fully expose the myosin binding sites.[11] Myosin, together with actin, forms the actomyosin complex which initiates the cross-bridge force generation cycle and leads to muscle contraction.

1.3 The Structure of Troponin

Cardiac troponin contains three intercalated subunits (TnC, TnI and TnT). Specifically, TnC is the calcium ion binding subunit which undergoes and initiates a broader conformational change when binding to Ca^{2+}; TnI is the putative actin inhibitory unit and inhibits the myosin binding site when the Tn complex is at the Ca^{2+}-free state; TnT serves a structural role and interlocks troponin with tropomyosin to form a complex.[12] Although the physiology of cTn has been extensively studied and well characterised, the structure of cTn remained unclear for a long time. The last 30 years' intense research effort by biochemists and structural biologists has cumulated in detailed X-ray crystal structures and cryoelectron microscopy (cryo-EM) structures. Several crystallography studies have resolved partial structures of troponin.[13–15] Tadaka *et al.* resolved the core domain of the troponin complex in the Ca^{2+}-saturated state at

2.6 Å in 2003.[16] This study showed that cTn in the Ca^{2+}-saturated state adopts an L-shaped conformation with two subdomains (regulatory head and the IT arm) connected by flexible linkers and suggested that troponin might adopt multiple conformations with variable subdomain orientations. Recent advances in cryo-EM technology has afforded the structures of the full thin filaments in both Ca^{2+}-bound and -free state at 3-5 Å resolution from Oda *et al.* and Yamada *et al.*[17,18] Both studies demonstrated the entire elongated structure of cTn along the thin filament. Comparative analysis of Ca^{2+}-free and -bound thin filament structures provided further clarity on the Ca^{2+}-dependent regulation of myosin binding site on actin. These structures have informed much of recent discoveries in cardiac muscle research and served as the basis for many computational modelling studies.[19–21] However, some key regulatory segments of troponin could not be resolved in those structures owing to their disordered nature. On the other hand, static structures only inform us about the structures at either high calcium or calcium-free state while the whole conformational transition upon calcium binding is of interest.

Nuclear magnetic resonance (NMR) in solution provides another technique in gaining insights into protein's local and global geometric information. NMR structure determination allows the elucidation of solution phase structure and provides an ensemble of models that are consistent with experimental data. Several NMR studies focused on revealing the missing regions from the crystal structure which were thought to be more flexible hence could not crystallise.[22,23] Attention in recent years turned to study how small molecules interact with troponin as well as diseased mutants.[24] These results have revealed important interaction patterns between key segments of troponin and small molecules.[25,26] However, NMR study is again limited by the size of the system which currently could not cover the full core domain of the troponin complex. This again highlights the complementary nature of multiple experimental and computational techniques in studying cardiac troponin.

The following sections detail the structures of the three subunits of troponin.[12]

1.3.1 Troponin C

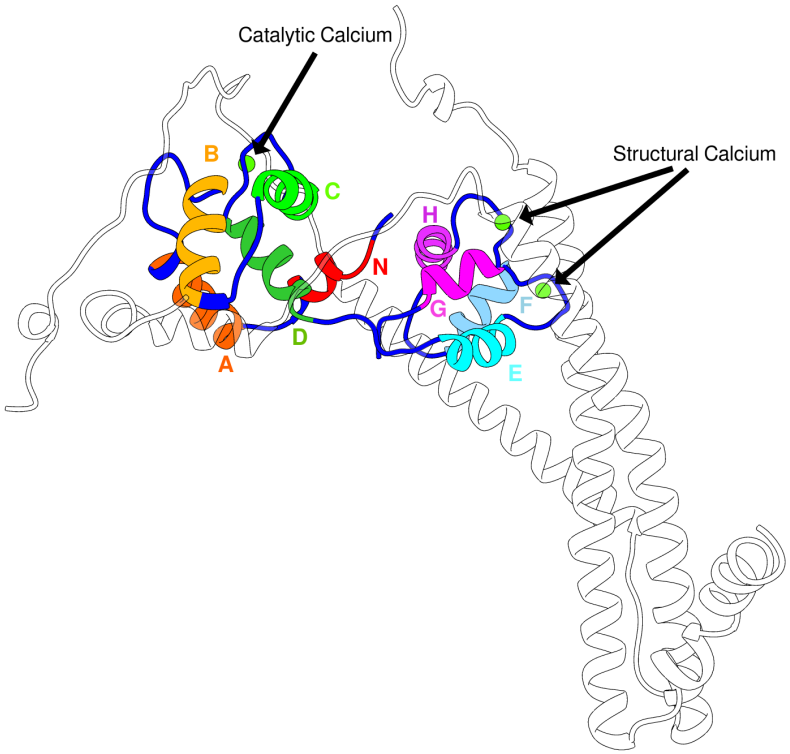


Figure 5: The cTnC molecule with its nine helices labelled.

Troponin C (TnC) is the Ca^{2+}-sensing portion of the troponin complex (Figure 5). Human cardiac isoform of TnC is 161-residue long and contains two globular domains (Table 1), the N-terminal domain (NcTnC) which contains the catalytic Ca^{2+} and the C-terminal domain (CcTnC) which contains the two structural Ca^{2+} ions. The two terminal domains are joined by a flexible hydrophobic linker. The human cTnC has nine α-helices where one of the helices is the N-terminal helix (helix N) and the other eight helices (helix A-H) form four EF-hands. EF-hands are known as the cation binding motif and each one consists of two helices joined by a loop region which coordinates with positively charged ions. No binding was observed at EF-hand I in cardiac troponin at any ion concentrations while EF-hand III and IV are always occupied under physiological conditions hence termed the structural Ca^{2+} binding sites.[27] EF-hand II shows a lower affinity ($K_d = 0.3\,\mu M$) for Ca^{2+} ion but binds to Ca^{2+} exclusively while EF-hand III and IV can bind to other divalent ions with much higher affinity ($K_d = 3\,nM$). In the absence of Ca^{2+} in diastole, cTnC is in a "closed" conformation and the cTnI interacts with

actin, inhibiting the actin–myosin interaction. In systole, upon Ca^{2+} binding to EF-hand II, cTnC adopts an "open" conformation that exposes a hydrophobic patch for interaction with cTnI switch peptide and removing cTnI inhibition from actin. The binding event of Ca^{2+} at EF-hand II initiates the conformational changes in Tn complex that forms a crucial part of the force generation process hence the Ca^{2+} at EF-hand II is termed the catalytic Ca^{2+}.

Troponin C

Residues	Region	Residues	Region	Residues	Region
3 - 9	Helix N	16 - 51	EF-hand I	1 - 87	N-terminal domain
14 - 25	Helix A	52 - 87	EF-hand II	88 - 91	Flexible linker
38 - 47	Helix B	92 - 127	EF-hand III	92 - 161	C-terminal domain
54 - 64	Helix C	128 - 161	EF-hand IV		
74 - 85	Helix D				
94 - 104	Helix E				
114 - 123	Helix F				
130 - 140	Helix G				
150 - 157	Helix H				

Table 1: The standard nomenclature of the different structural regions of Troponin C.

1.3.2 Troponin I

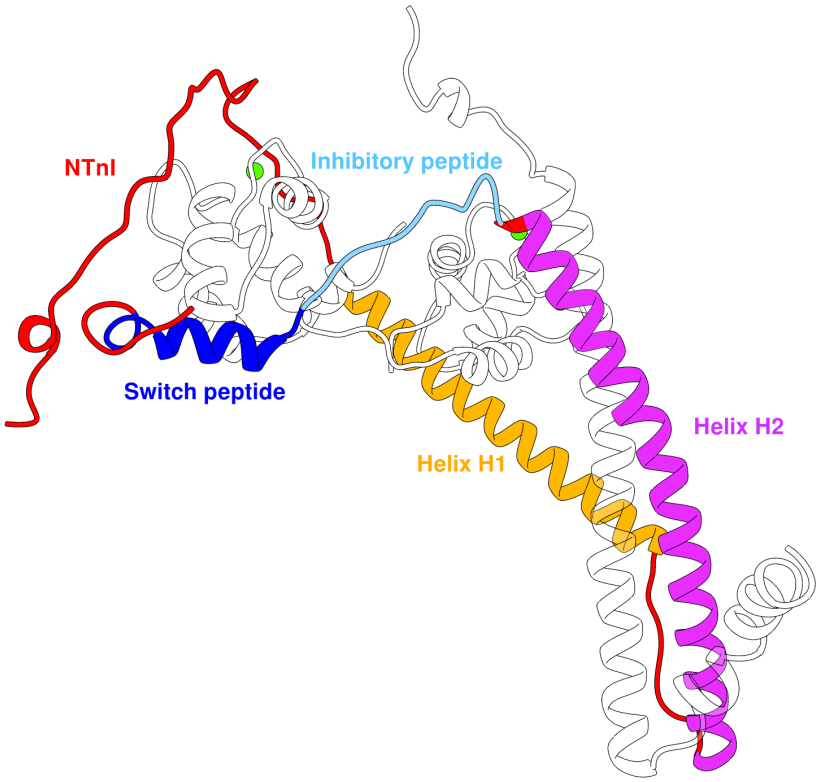


Figure 6: The cTnI molecule with relevant structural regions labelled. Residues 172 to 210 of cTnI are not present in this model.

In human cTn complex, cTnI is expressed as a 210-residue long protein which consists of four α-helices interconnected by flexible linkers. From N-terminal to C-terminal, cTnI is classified into four domains (Table 2). The N-terminal domain (NcTnI), which is absent in skeletal muscle, contains a short and unstable α-helix and a disordered region. Two adjacent serine residues (Ser 22 and Ser 23) are located in the disordered region of NcTnI. Phosphorylation of Ser 22/23 by PKA reduces the Ca^{2+}-affinity at EF-hand II of cTnC and increase the cardiac muscle contraction rate when the body is under stress or intense physical activity. Although some literature numbers these two serines as 23 and 24, recent mass spectrometry studies have shown that the initial methionine of TnI is cleved off.[28] Hence I adopt the numbering of 22/23 in this work. The next two α-helices (helix H1 and H2) are joint by a flexible U-turn and this domain forms part of the IT arm of the troponin complex which serves a structural role. The short disordered region after helix H2 is termed the inhibitory peptide which contains

six positively charged residues. At low Ca^{2+} state, this region plays a crucial role in inhibiting the myosin binding site on actin, hence preventing the formation of the actomyosin complex which is critical to the muscle contraction process thus causing muscle relaxation.[29] The short helical section after the inhibitory peptide is the switch peptide which interacts with EF-hand I of cTnC to form a hydrophobic cleft under high Ca^{2+} concentration, and pulls away the inhibitory peptide from its inhibitory position hence exposes the myosin binding site and initiates the muscle contraction cycle. The final C-terminal domain of cTnI did not have a clearly resolved structure until the recent cryo-EM thin filament structures elucidation. It is shown to interact with tropomyosin and actin under low Ca^{2+} concentration.

Troponin I			
Residues	Region	Residues	Region
43 - 79	Helix H1	2 - 32	N-terminal domain
80 - 89	Flexible U-turn	42 - 136	IT arm
90 - 135	Helix H2	137 - 148	Inhibitory peptide
151 - 162	Helix H3	149 - 163	Switch peptide
163 - 188	Helix H4	164 - 210	C-terminal domain

Table 2: The standard nomenclature of the different structural regions of Troponin I.

1.3.3 Troponin T

The human cardiac troponin T (cTnT) is the largest subunit of the cTn complex (Figure 7). It mainly serves a structural role that pivots the cTn on the actin surface and binds to tropomyosin. There are three domains on cTnT, the N-terminal domain, the central domain, and the C-terminal domain (Table 3). The N-terminal domain is a 68-residue region and there is no clear understanding of its interaction with other components of the thin filament. The central domain contains T1 site which is known to form interactions with tropomyosin, and a

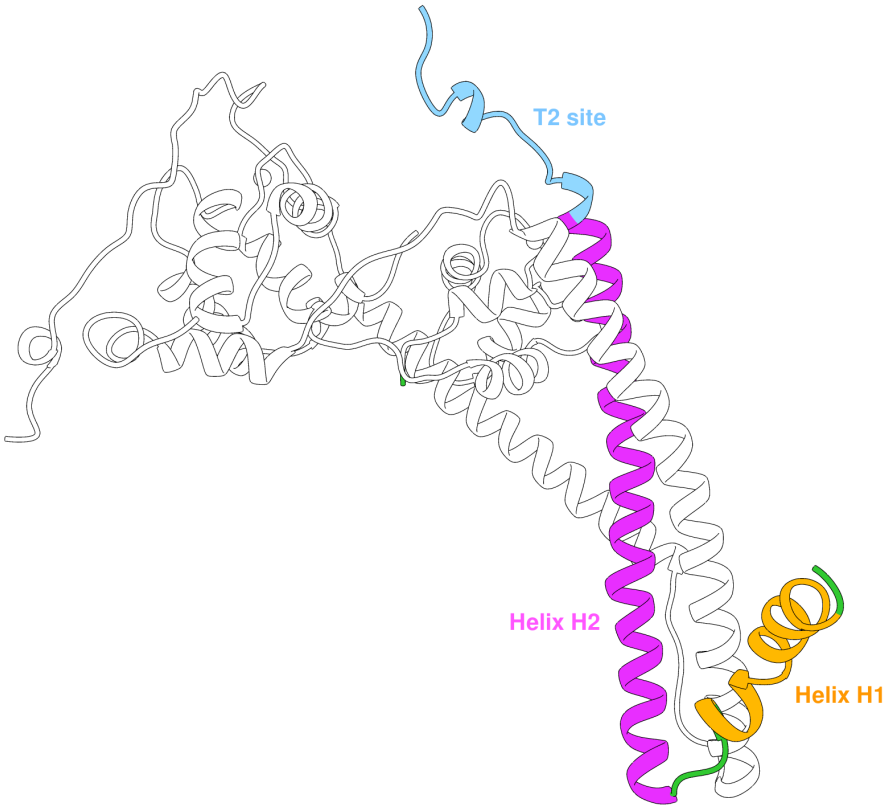


Figure 7: The cTnT molecule with relevant structural regions labelled. Residues 1 to 201 of cTnI are not present in this model.

flexible linker region. The C-terminal domain is formed by two α-helices, H1 and H2. Helix H2 of cTnT, together with helix H1 and H2 of cTnI, forms the coiled-coil IT arm of cTn which provides a rigid structural support for the complex. The structural C-terminal domain of cTnC also interacts with the end of helix H2 of cTnT. The final residues after helix H2 form the T2 site which is proposed to be another interaction site between cTn and Tm.[30]

Troponin T					
Residues	Region	Residues	Region	Residues	Region
204 - 220	Helix H1	98 - 136	T1 site	2 - 68	N-terminal domain
226 - 271	Helix H2	197 - 239	T2 site	69 - 200	Central domain
38 - 47	Helix B	272 - 288		201 -288	C-terminal domain

Table 3: The standard nomenclature of the different structural regions of Troponin T.

1.4 Post-translational Modifications on Troponin

In addition to the Ca^{2+}-activated switching event, cardiac troponin possesses a unique mechanism that modulates its Ca^{2+}-sensitivity by phosphorylation. The β_1-adrenergic receptors are vital G-protein coupled receptors located in the cardiomyocytes. When the body is under stress or demanding physical activities, the β_1-adrenergic receptors are triggered by catecholamines which leads to the stimulation of the cAMP/PKA pathway as well as the opening of the L-type calcium channel that directly increases the cytosolic Ca^{2+}-concentration. β-Blockers which inhibit this pathway have been used for five decades to treat a variety of cardiac diseases. The result of the catecholamine uptake is an increased heart rate (chronotropic effect), an increase in the force of muscular contraction (inotropic effect), and an increased rate of muscle relaxation (lusitropic effect).

Protein kinase A (PKA), together with several other kinases, can lead to the phosphorylation of a number of residues on myosin binding protein C and cTn. The two adjacent serine residues on cTnI, serine 22 and 23, are of primary focus as they are located on the cardiac specific N-terminal domain of cTnI (NcTnI) which interact with the regulatory domain of cTnC directly. Top down mass spectrometry (MS) of diphosphorylated cTnI showed that the phosphorylation of the Ser22/23 pair is sequential and Ser23 is only phosphorylated after Ser22.[28] This post-translational modification is crucial to the normal regulatory function of the troponin complex. A study using transgenic mice that contain the non-phosphorylatable, slow skeletal

isoform of TnI has proven that the acceleration of myofibrillar relaxation is due to phosphorylation of the N-terminal domain of cTnI.[31] Phosphorylation of Ser22/23 results in a 2-3 fold decrease in myofilament Ca^{2+} sensitivity and an increased Ca^{2+} dissociation rate which determines myocyte relaxation rate. This change equates to a less than $5\,\mathrm{kcal\,mol^{-1}}$ difference in interaction energy which is beyond the accuracy of the computational methods used in this work. On the molecular level, phosphorylation is thought to decrease the catalytic Ca^{2+}-binding affinity, reduce the strength of cTnI interaction with cTnC, and increase the rate of muscle relaxation which leads to an increased heart rate thus supplying more oxygen to the body.[32]

2D NMR study by Hwang *et al.* asserted that NcTnI interacts with NcTnC in the unphosphorylated state which fixed the orientation of NcTnC with respect to the rest of the troponin complex.[23] In contrast, phosphorylation of Ser22/23 in cTnI has been shown to disrupt the interaction with NcTnC, with the two phosphate groups neutralizing the positive charges contributed by Arg19, Arg20, Arg21, and Arg26 located in close proximity.

In addition to the Ser22/23 phosphorylation by PKA investigated in this project, other phosphorylation sites have also been studied in recent years. When stuyding monophosphorylated cTnI with top down MS, phosphorylation site Ser76/Thr77 was found with partial occupancy 36% in addition to Ser22 at 53%.[28] Notably, Cool *et al.* computationally investigated the PKC initiated phosphorylation on TnI T143 which is located on the inhibitory peptide. This work showed that the effective concentration of switch peptide to the hydrophobic patch is significantly increased for the phosphothreonine models.[21] Phosphorylation of TnI Tyr25 was studied experimentally with tyrosine kinase as well as pseudo-phosphorylation mutations, Y25E and Y25D.[33] It shows that filaments containing Tyr25 phosphorylated TnI accelerate the rate of calcium dissociation from TnC similar to that of TnI Ser22/23. This represents the first known myofilament protein tyrosine phosphorylation to directly modulate cardiac function.

1.5 Troponin Mutation and Genetic Cardiomyopathies

Cardiomyopathies, a diverse group of heart muscle diseases that frequently lead to heart failure and death, can be caused by many factors including viral infections, alcoholism, genet-

ics, or other diseases. Familial cardiomyopathies are potentially life-threatening heart diseases caused by inherited mutations within the troponin complex and other sarcomeric proteins in the cardiomyocytes.[34] There are several types of cardiomyopathies which are characterised by their clinical characteristics: dilated cardiomyopathy (DCM), hypertrophic cardiomyopathy (HCM), restrictive cardiomyopathy (RCM), and arrhythmogenic right ventricular cardiomyopathy (ARVC).[35,36] Many mutations within the thin filament has been associated with inherited cardiomyopathies, in particular, DCM and HCM. Commonly, these mutations were found to affect the PKA-dependent modulation in cTn dynamics by either mimicking the change in Ca^{2+} sensitivity induced by phosphorylation or abolishing the relationship between phosphorylation and Ca^{2+}-sensitivity completely.

HCM is commonly caused by genetic mutations in the sarcomeric proteins and is the most common cause of sudden cardiac death in young athletes.[37] The disease has an estimated prevalence of $\sim$1:500 and is characterised by abnormal hypertrophy in the left ventricle, particularly in the interventricular septum. More than 1000 mutations in over a dozen sarcomeric proteins have been linked to HCM. In particular, the β-myosin heavy chain encoding MYH7 and the myosin binding protein C encoding MYBPC3 are the two most common genes involved which account for around half of the HCM patients, followed by mutations in cTn.[38]

DCM is characterised by a thinning of the heart walls and an increased volume in the left ventricular chamber which leads to a reduced ejection fraction from the heart. The disease is more rare than HCM with an estimated prevalence of $\sim$1:5000 and is associated with a wider array of mutations, both sarcomeric and non-sarcomeric.

Extensive research into the disease pathway and functional effects of disease-causing mutations have afforded some understandings of the cellular mechanisms of inherited cardiomyopathies. Many adaptive responses in the heart exist to compensate genetic defects, in order to maintain the normal contractile functionality; however, sustained activation of these signaling pathways lead to cardiac dysfunction.[39]

Through the past two decade's biochemical experiments, it has become evident that many mutations associated with cardiomyopathies can lead to disruptions of the phosphorylation-dependent control of the Ca^{2+}-regulatory mechanism. It is well-established that the major

effect of HCM-causing mutations is to increase myofibrillar Ca^{2+}-sensitivity. Studies on the molecular mechanism of DCM have resulted in conflicting evidence over whether it causes an increase or a decrease of the Ca^{2+}-sensitivity in cardiomyocytes. Subsequent investigations showed that different mutants may increase, decrease or have no effect on the mutant's Ca^{2+}-sensitivity compared to the wild-type.[40,41] By studying isolated thin filaments with the quantitative *in vitro* motility assay (IVMA) it was found that the myofilament Ca^{2+}-sensitivity is not reduced upon TnI phosphorylation by PKA in many DCM-causing mutants. This relationship is termed the *uncoupling* phenomenon. It was further proposed that this uncoupling phenomenon blunts the lusitropic response which is necessary and sufficient to cause the DCM phenotype.[40–42] The mutation study of this project focuses on the TNNC1 G159D mutation which is located at the C-terminal of cTnC that results in genetic DCM.[43]

The recent focus of the research community has been using modern experimental and computational techniques to elucidate the stuctural effects of mutations, in order to link to and rationalise the mutants' functional anomalies. Troponin, as the Ca^{2+}-sensing apparatus, is at the center of these studies.[44] Liu *et al.* demonstrated that mutations on cTn alters thin filament's Ca^{2+} association and dissociation rate.[45] Using small-angle X-ray scattering, Matsuo *et al.* studied the structural changes introduced by the HCM-causing mutation TNNT2 E244D. It was found that the structural changes caused by the mutation in the troponin core domain are propagated throughout the molecule and modulating the intermolecular interactions between Tn/Tm and Tn/actin. In addition, the findings suggested that the mutation disrupts the normal movement of Tn and Tm upon Ca^{2+}-activation.[46]

1.5.1 The TNNC1 G159D Mutation

Genetic investigations of cardiac troponin from clinical data have discovered the TNNC1 Gly159 to Asp mutation as a DCM gene.[47] This single amino acid mutation is of particular interest because of its location on the C-terminal domain of cTnC which is generally considered to serve a structural function rather than regulatory.

Using recombinant mutant human troponin subunits, it was first determined that the G159D mutation causes no change in the Ca^{2+}-sensitivity or the cooperativity of force generation

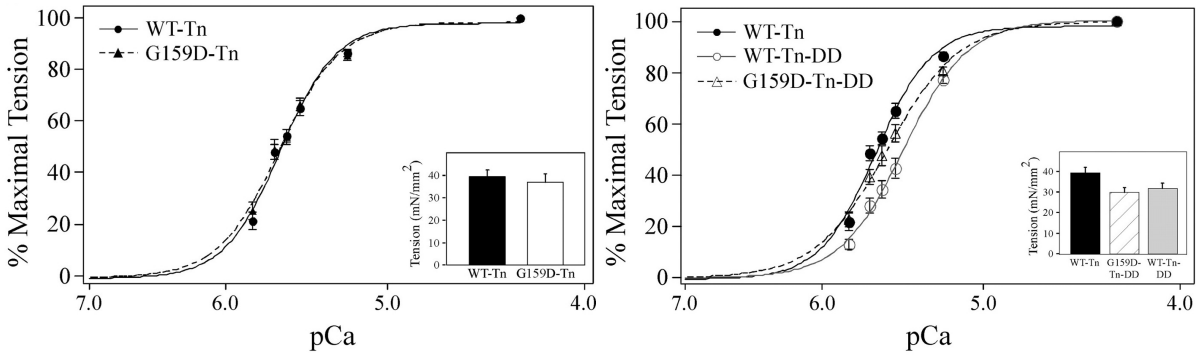


Figure 8: Measurements of Ca^{2+}-activated development of steady-state tension. Left graph shows that Tn containing the TnC G159D mutation does not alter trabeculae function without phosphorylation. Right graph compares pseudo-phosphorylated WT and mutant which shows that exchange of Tn containing the cTnC G159D and Ser22/23 pseudo-phosphorylated cTnI into trabeculae blunts the response of cTnI phosphorylation. Reproduced with permission.[48]

when compared to the wild-type.[49] However, the mutant reduces the muscle activation rate by 50%. Measurements of isometric tension in relationship to free Ca^{2+} concentration (Figure 8) have demonstrated that the G159D mutation does not alter the myofilament response to Ca^{2+} without phosphorylation modulation.[48,50] The dissociation of calcium from troponin is on the millisecond timescale which is beyond the achievable computational simulation length. In the presence of pseudo-phosphorylated cTnI at Ser22/23, which reduces Ca^{2+} sensitivity and enhances cross-bridge cycling in the wild type, the G159D mutant blunted the phosphorylation induced decrease in Ca^{2+}-dependent tension development without altering cross-bridge cycling. Uncoupling results in blunting of the lusitropic response to β_1 receptor stimulation and this has been proposed to be causative of the DCM phenotype.[51] Measurements in purified troponin confirmed that this G159D blunting of myofilament desensitization results from altered Ca^{2+}-binding to cTnC.[48]

Memo *et al.* used quantitative *in vitro* motility assay to quantify filament force generation with respect to free Ca^{2+} concentration. For G159D mutant, the EC_{50} of Ca^{2+} sensitivity of unphosphorylated troponin is $0.095\,\mu M$ which is a slight increase from the value of wild-type ($0.059\,\mu M$). In the phosphorylated state, the EC_{50} of Ca^{2+} sensitivity decreases from $0.14\,\mu M$ in wild-type to $0.092\,\mu M$ in the mutant.[41] This is an uncoupling phenomenon as no change in Ca^{2+} sensitivity was observed upon phosphorylation because of the mutation.

Further solution NMR studies with cTnI phosphomimetic mutants S22D/S23D, S41D/S43D

have shown that the DCM-associated mutations (cTnI K35Q, cTnC D75Y, and cTnC G159D) disrupt the active orientation of NcTnC. The active orientation of NcTnC is stabilized by the flexible tails of cTnI, $cTnI_{1-37}$ and $cTnI_{135-209}$. And when the modified cTnC without either tails is incorporated into the troponin complex, the NcTnC domain is still immobilized, revealing a new interaction between NcTnC and the IT-arm that stabilises a "dormant" orientation. Mahmud *et al.* proposed that the calcium sensitivity of the cardiac troponin complex is regulated by an equilibrium between active and dormant orientations, which can be shifted through post-translational modifications or DCM-associated mutations.[52]

Stevens *et al.* obtained the Ca^{2+} affinity of NcTnC with several clinically relevant mutations (A8V, L29Q, A31S, L48Q, Q50R, and C84Y) via isothermal titration calorimetry. It was found that Ca^{2+} affinity of the N-terminal domain of cTnC in isolation is insufficient to explain the pathogenicity of these mutations.[53] This disproved the hypothesis that mutations directly modulate TnC Ca^{2+} affinity. Further investigation with molecular dynamics simulations revealed that mutations modify the structural dynamics of TnC rather than the regulatory Ca^{2+} binding site which could help explain the disruptive nature of mutations located on the C-terminal end of TnC.

Overall, the TNNC1 G159D mutation is an interesting system to study. As one of the first DCM-linked mutation, it has well characterised biochemical data and experimental results. Its location on the troponin core domain is set for a simulation system well-balanced between model size and timescale. The studies presented in this thesis aim to reveal further structural and molecular insights in understanding how the mutation causes cardiac dysfunction.

1.6 Therapeutic Small Molecules for Cardiomyopathies

Small molecule studies, coined as the fourth wave of muscle research, aim to find therapeutic small molecules that have potential for reversing the abnormal contractile functions introduced by mutations in the myocytes.[54] To date, there are limited compounds for the treatments or prevention of sarcomeric cardiomyopathies. Current pharmaceutic compounds are directed at alleviating symptoms and preventing disease progression rather than treating the root cause of the diseases. There are three general directions for combating the functional

effects caused by mutations: Ca^{2+}-sensitisers and -desensitisers for correcting the mutants' Ca^{2+}-sensitivity to the wild-type level, recouplers for restoring the coupled relationship between phosphorylation state and Ca^{2+}-sensitivity in mutants, and myosin activators and inhibitors for modulating the heart's work output.[55] A number of modulators that aim to correct the abnormal calcium sensitivity in diseased mutants are targeting the sarcomere. Of those compounds, some are in clinical and preclinical stages and some are for research only.[56]

Some of the earliest quests towards a therapeutic molecule started with epigallocatechin-3 gallate (EGCG) which is an abundant polyphenol compound found in tea.[57,58] It was originally demonstrated that EGCG is a Ca^{2+}-desensitiser and later found that it is also able to restore the coupled relationship between TnI phosphorylation and Ca^{2+}-sensitivity in cTn mutants. The compounds bearing this recoupling property are termed *recouplers*.[51,57] Other compounds possess the desensitisation property include Nebivolol, W7, epicatechin gallate (ECG), and Silybin A; while Silybin B, resveratrol, pterostilbene, and novobiocin were found to be recouplers.

Bepridil, Levosimedan, and EMD57033 are examples of Ca^{2+}-sensitisers and have been extensively studied for their ability to increase Ca^{2+}-affinity for cardiac troponin C which in turn increases the sarcomeric Ca^{2+}-sensitivity. However, these compounds are unsuitable as general treatment as they are not specific enough to target cardiac troponin and have significant phosphodiesterase inhibiting acitivity. Moreover, quantitative IVMA have showed that both EMD57033 and Bepridil uncoupled Ca^{2+}-sensitivity from TnI phosphorylation, mimicking the effect of HCM mutations.[57]

NMR studies have revealed the binding interactions between cTn and several small molecules that can modulate the sarcomeric Ca^{2+}-sensitivity.[26,59–64] Most of the binding sites are localised on cTnC and have been found in both the regulatory and the structural domain. Small molecules targeting the NcTnC-cTnI interface to modulate the Ca^{2+}-sensitivity has also been a prominent area of research for its potential to treat heart diseases. This mode of action was used to rationalise the functional effects of compounds that were shown to interact with the hydrophobic pocket of NcTnC. Bowman *et al.* used an umbrella sampling method to elucidate the impact on the free energy of cNTnC hydrophobic patch opening due to

small molecule Ca^{2+}-sensitivity modulators.[65] It was found that the studied compounds (W7, bepridil, dfbp-o, 3-methyldiphenylamine) all lowered the free energy of hydrophobic patch opening without cTnI. With cTnI switch peptide present, dfbp-o and 3-methyldiphenylamine stabilise the open configuration while W7 and bepridil destablise the open state.

The combination of desensitisation with recoupling activity, as shown by EGCG is in principle an ideal functional profile for the treatment of HCM. However, EGCG is known to be inotropic and is a promiscuous molecule which make it unsuitable as a drug candidate. Nevertheless, EGCG was previously studied as a Hsp90 inhibitor and a large variety of structural analogues of EGCG were synthesised and tested for their activities towards diseased sarcomere.[66] Sheehan *et al.* assayed the Ca^{2+}-sensitivity of TPM1 E180G mutation of troponin *in vitro* when treated with 40 different small molecules that are structurally related to EGCG or silybin.[67] It was found that the recoupling property of small molecule is not correlated with activity towards Hsp90 inhibition. Twenty-three of the assayed compounds have shown to be recouplers and were able to restore the phosphorylation-mediated Ca^{2+}-sensitivity shift to the wild-type level. Interestingly, one structurally similar compound was found to cause an increase in Ca^{2+}-sensitivity upon phosphorylation in TPM1 E180G mutant, which is termed the *reverse-recoupler*. This study also revealed that the Ca^{2+}-desensitising and recoupling activities are separated, and it was proposed that there are four functional states for troponin (Figure 9). In normal cardiac tissue, phosphorylation of troponin can shift the contractile apparatus from a high Ca^{2+} affinity state to a low Ca^{2+} affinity state. However, this transition is blocked in disease-causing mutants and cTn is stuck in the high Ca^{2+} affinity state. It was further hypothesised that by selectively binding, hence stabilising, the phosphorylated state, recouplers can restore the normal change in Ca^{2+} affinity upon phosphorylation.

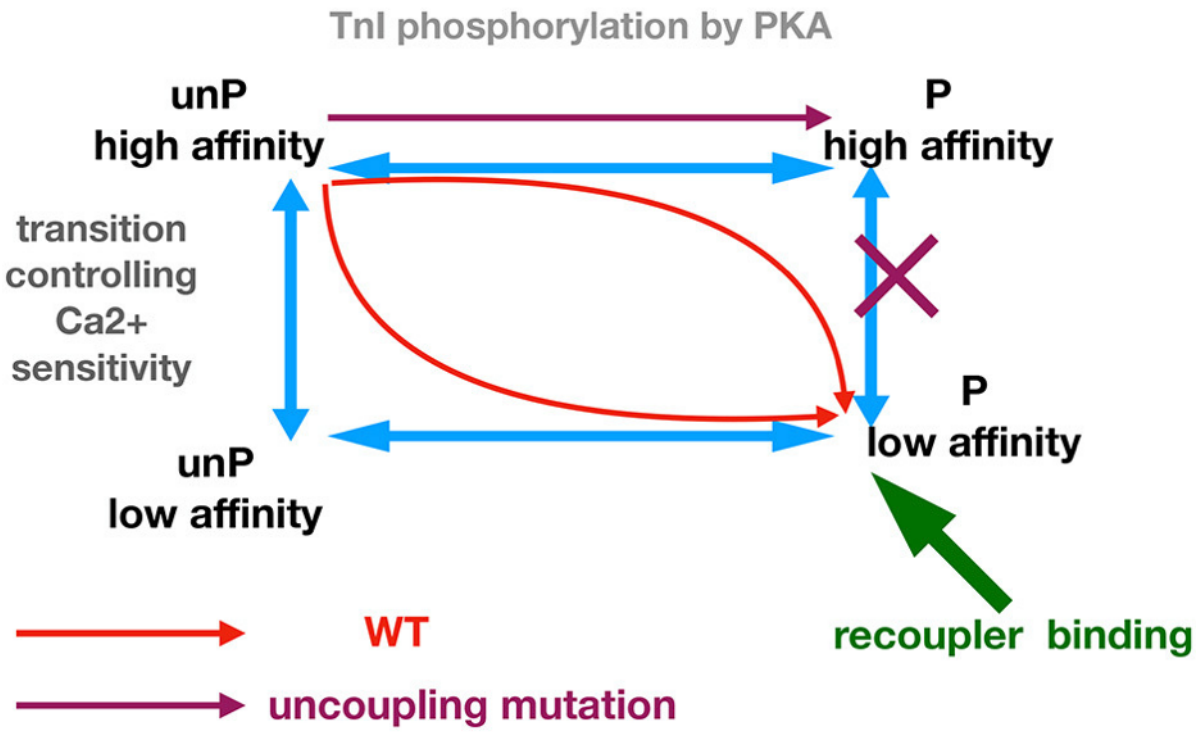


Figure 9: The four proposed functional states of troponin.[67] The wild-type troponin transitions from the high affinity state to the low affinity state upon phosphorylation while the pathway is blocked with mutantions. By stabilising the phosphorylated low affinity state, small molecules can recouple troponin mutants.
©2018 Sheehan, Messer, Papadaki, Choudhry, Kren, Biedermann, Blagg, Khandelwal and Marston. Licensed under the Creative Commons Attribution License (CC BY).

In this thesis, I report the structure-function study on a group of well-characterised, diverse small molecules (EGCG, resveratrol, SilybinA, SilybinB) in Section 4 and a group of derived compounds with interesting properties (ECG, DHSA, DHSB, r07) in Section 5.

1.7 Computational Simulations of Cardiac Muscle

The rapid development of the semiconductor industry following Moore's Law has provided computational biologists exponentially more powerful computer clusters, and a much wider landscape to apply cutting-edge software techniques for simulations. Application Specific Integrated Circuits (ASIC) specially designed for molecular dynamics (MD) have also been developed to conduct molecular simulations with unprecedented speed. It is now possible to simulate systems at the cellular-level down to atomistic detail or simulate small protein sys-

tems over physiologically relevant, millisecond timescales.[68,69] Recent commoditisation of high-performance Graphics Processing Unit (GPU), thanks to the gaming industry and machine learning fields, compounded with the advancements of general-purpose GPU computation toolkits, has allowed researchers access to routine long-timescale all-atom simulations.

Simulations in the cardiac muscle research field have been applied to a wide range of scales, from individual subunits (e.g. NcTnC, actin) to the full thin filament.[70] In the early days, separated domains of cTnC were simulated which showed the crucial hydrophobic patch open/closed dynamics.[71–73] Minimal systems were constructed to study the interaction of individual peptide segments, such as NcTnC and the TnI switch peptide. Subsequently, the troponin core domain was constructed to better represent the contractile regulator and study the interactions among its subunits. In the last decade, full cardiac thin filament systems were built to provide molecular insight into disease mechanism and facilitate future rational drug design.[74–77] In addition to troponin, MD has been applied to study other constituents of the sarcomere, including myosin,[78] tropomyosin-actins/myosin binding,[79] and titin.[80]

Molecular dynamics derived free energy methods were used to computationally estimate the binding affinity. Hanz *et al.* used steered MD to get the potential of mean force along a pulling path to estimate the catalytic calcium binding affinity; they observed the correct trends for the N-terminal TnC mutation effects on the catalytic calcium binding affinity with calcium sensitising and desensitising mutants.[81] The ability to rapidly delineate the mutation effect on troponin calcium sensitivity will help guide the design of therapeutic mutants and gene therapy that can reverse the mutation effects.[77,82]

The recent experimental studies on troponin have been focusing on understanding the disease-causing mutants' functional mechanism as well as finding drug-like molecules that can reverse the effects of mutantions. Multi-scale computational methods, with a variety of accuracy-speed trade-offs, are well-suited for the study of small molecules and can complement experimental techniques. Mechanistic insights gained from simulations can be further applied for discovering novel therapeutic molecules via high-throughput virtual screening or validation of rationally designed molecules.

The hydrophobic patch has been a target for high-throughput virtual screening and several

calcium sensitizers were discovered through this process, including NSC147866 from the NCI II diversity set, NSC600285 and NSC611817 from the entire NCI database.[83,84] Combined with experimental intuition, analogue compounds to diphenylamine were docked into NcTnC and allowed for the identification of the calcium sensitizer, 3-methyldiphenylamine.[25] Small molecules bound to NcTnC have also been studied with an umbrella sampling scheme to show their influence on the free energy landscape.[65]

1.8 Thesis Outline

In this chapter, I have introduced the general background on cardiac muscle and troponin as well as the current research focus on these topics. Details specific to each chapter are provided at the beginning of the chapter where relevant. Section 2 introduces the computational methods utilised in this thesis. Section 3 describes and discusses the phosphorylation effects on wild-type troponin and its G159D mutant. Sections 4 and 5 present small molecule studies that connect experimental observations to simulated structural effects. Finally, **??** provides a general conclusion on this work along with areas for future research.

The work in this thesis is built on past explorations and knowledge. Combining the latest experimental insights with well-balanced simulation size and timescale trade-off, I aim to uncover the structural effects of the TNNC1 G159D mutation and a variety of small molecules with potential therapeutic applications.

2 Computational Methods

Computational simulations have accompanied many major scientific discoveries over the past few decades. In the pharmaceutical industry, it is now routine to use various modelling methods in all parts of the drug discovery pipeline. Here we use simulations to further our understanding of troponin's regulatory mechanism, in particular the phosphorylation modulation and mutation induced defects. In this work, I utilised all-atom classical molecular dynamics (MD) to study the troponin complex. The unphosphorylated (uP) and doubly-negatively charged phosphorylated (SEP) wild-type troponin complex in the Ca^{2+} activated state, as well as its mutant, TNNC1 G159D were simulated. Additionally, simulations were done with the presence of small molecules to uncover the protein-ligand interactions and their impact on troponin dynamics. The model used in this study contains all three regions of cTn with the full troponin C subunit, and partial structure of troponin I and T that form the core domain of the cTn complex. To date, it is one of the most complete models simulated on the microsecond timescale.

In this chapter, I will give a brief overview on the theoretical background of MD, simulation protocols, and the post-simulation analyses I employed in this work.

2.1 Molecular Dynamics

Molecular dynamics (MD) is a method of studying molecular systems, ranging from small molecules, proteins, to solid-state materials. MD uses classical mechanics to model the physical interaction between pairs of atoms. The overall force acting on an atom is determined with molecular mechanics (MM), otherwise known as force-field equations. The net force is derived from the force-field equation which describes the total potential energy. Here, I used the AMBER software suite and its accompanied force-field for simulations.[85] The force acting on each atom can be calculated by taking the derivative relative to position. The aim of MD is to study the time-dependent dynamics of a system, and with numerical integration method we can simulate atom motions by displacing the particles for each timestep to obtain a movement trajectory over time.

2.1.1 AMBER force-field

The AMBER (Assisted Model Building with Energy Refinement) force-field describes the total potential energy of a molecular system.[86] The equation has the following form:

$$
\begin{aligned}
E_{total} = & \sum_{i \in \text{bonds}} k_{bi}(l_i - l_i^0)^2 \\
& + \sum_{i \in \text{angles}} k_{ai}(\theta_i - \theta_i^0)^2 \\
& + \sum_{i \in \text{dihedrals}} V_i[1 + \cos(n\phi_i - \gamma_i)] \\
& + \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} \left[\frac{A_{ij}}{r_{ij}^{12}} - \frac{B_{ij}}{r_{ij}^{6}} + \frac{q_i q_j}{\epsilon r_{ij}} \right]
\end{aligned}
$$

The first term sums over the directly bonded atom pairs where the potential energy is modelled as a harmonic oscillator. The second term sums over the angles of atom pairs that are two bonds apart. This energy arises when the bond angle deviates from the typical valence bonding angle. The third term sums over all dihedral pairs three covalent bonds apart and describes the torsional strain imposed due to the orientation of the bonds. The fourth terms sums over all atom pairs of the molecular system and describes the total non-covalent interaction energies. It comprises the van der Waals interaction modelled by the 12-6 Lennard-Jones potential and the electrostatic interaction between charged particles.

2.1.2 Velocity Verlet Algorithm

To simulation the molecular system over time, a time-reversible algorithm is needed to integrate Newton's equation of motion. The velocity Verlet algorithm, a variation on the Verlet algorithm, is commonly employed to obtain the position and velocity update for the next timestep from the current position and velocity.

For $\mathbf{x}(t)$ which represent the coordinates of a molecular system at time t, its Taylor expansion at $(t + \Delta t)$ to the third term is:

$$\mathbf{x}(t+\Delta t) = \mathbf{x}(t) + \mathbf{v}(t)\Delta t + \frac{1}{2}\mathbf{a}(t)\Delta t^2 \tag{1}$$

The velocity at each timestep is then updated as the following:

$$\mathbf{v}(t+\Delta t) = \mathbf{v}(t) + \frac{1}{2}[\mathbf{a}(t) + \mathbf{a}(t+\Delta t)]\Delta t \tag{2}$$

It should be noted that the velocity at $(t+\Delta t)$ can only be computed when the position has already been updated, as the velocity update depends on the updated acceleration which is only available with the updated postion. This assumes that the acceleration only depends on the particle position and not on the velocity. It can be shown that the velocity Verlet algorithm is equivalent to the original coordinate based Verlet algorithm hence the magnitude of the error term is the same.

2.1.3 Molecular Dynamics Simulation Protocol

The cTn model used for simulations in this work was constructed and extensively studied in the group.[87,88] It was built by combining the original crystal structure from Takeda *et al.* with missing disordered regions taken from NMR studies or modelled in.[87] The mutated residues in the crystal structures were restored to the native human sequence. Simulations started from a similar conformational structure in all systems which was derived from a G159D SEP trajectory from a previous study in the group. All simulations were performed using the AMBER software suite[85] and the CUDA accelerated PMEMD[89] code in the SPFP precision mode.[90] All parameter and topology files were prepared using LEaP, which is part of AmberTools, with AMBER ff14SB force-field library for amino acids and phosaa10 parameter library for phosphoserines (SEP).[91,92]

For the apo simulations, the troponin complex was placed at the center of a $140\,\text{Å}$ cubic waterbox. For the simulations of troponin with ligands the dimension of the cubic waterbox was increased to $150\,\text{Å}$ to accommodate a reasonable separation between the protein and the ligand in the starting structure. Work from the earlier part of the project, including sim-

ulations with ligand EGCG, RVL, SilybinA, and SilybinB, were conducted in a variable sized periodic boundary box where the box was chosen to be as least 20, 20 and 30 Å from the complex's x, y, and z sides, as noted in the previous study.[87] These systems also started with different troponin conformations which resulted in different solvent box sizes for WT uP, WT SEP, G159D uP, and G159D SEP. The respective box sizes are WT uP: 133 Å, 133 Å, 130 Å; WT SEP: 133 Å, 133 Å, 130 Å; G159D uP: 125 Å, 135 Å, 148 Å; G159D SEP: 140 Å, 137 Å, 128 Å. The structure was first neutralised by adding Cl^- or Na^+ as counter ions then solvated by adding water molecules, modelled with TIP3P water model.[93] To simulate physiological conditions, water molecules were replaced by ions at random to achieve a concentration of 0.15 M NaCl. Each system underwent a 5-stage minimisation, heating, and equilibration process before production runs.

1. An initial minimisation of the solvent by restraining the protein complex atoms with force constant of $10\,000\,\mathrm{kcal\,mol^{-1}\,Å^{-2}}$ for a maximum of 10000 cycles. The minimisation algorithm was switched from steepest descent to conjugate gradient after 500 steps.

2. A second minimisation of all atoms in the system for a maximum of 2500 steps. The system is first minimised for 1000 steps with steepest descent then switched algorithm to conjugate gradient.

3. A first-stage heating of the system went from 0 K to 100 K with an NVE ensemble and weak restraints on the protein complex ($10\,\mathrm{kcal\,mol^{-1}\,Å^{-2}}$), except the N-terminal of TnI (1-32). The temperature was increased linearly over 50 ps.

4. A second-stage heating of the system brought the temperature to 320 K over 60 ps then cooled down to 300 K over 40 ps with an NPT ensemble.

5. Equilibration of the system at 300 K for 10 ns (early simulations used 1 ns equilibration time) with an NPT ensemble.

The minimisation steps were done under constant volume with a non-bonded interactions cutoff of 12 Å. The NPT ensemble was controlled through Langevin dynamics with a collision frequency of $1\,\mathrm{ps^{-1}}$. A Monte Carlo barostat was used to maintain a 1 bar pressure with a relaxation time of 2.0 ps. For the heating, equilibration, and production runs, a non-bonded

interactions cutoff of $8\,\text{Å}$ was employed and the SHAKE algorithm was used to constrain hydrogen bonds to allow for a larger timestep.[94] Hydrogen mass repartition was employed for the equilibration and production steps to allow a $4\,\text{fs}$ timestep.[95] Trajectories were saved at a five frames per nanosecond frequency. Each production simulation was run for $1500\,\text{ns}$ and repeated 5 times which generated $7.5\,\mu\text{s}$ of aggregated simulation time for each system (37500 frames).

2.1.4 Ligand Parameterisation

The force field parameters for the ligands in this study were based on a general AMBER force field (GAFF) which is suitable for small organic molecules.[96] Ligand partial charges were derived from the ground state structures which were found by following conformational search, geometry optimisation, and thermal correction procedures, as detailed below:

1. An ensemble of conformers were generated for each ligand using the ETKDG algorithm as implemented by RDKit.[97]
2. Each conformer was optimised to a local minimum.
3. Each optimised structure was subjected to thermal corrections to calculate the Gibbs free energy.[98–100] The structure with the lowest ground state energy was picked as the ground state structure

From the ground state structures, restrained electrostatic potential (RESP) calculations were used to obtain the charge distribution on the molecule.[101] Gaussian16[102] was used for both geometry optimisation and RESP calculation. Geometry optimisation was done with B3LYP functional with cc-pVDZ basis set and the polarisable continuum model (PCM) in water as implicit solvent.[103–106] RESP calculation was done with Hartree-Fock method and 6-31G* basis set in vacuum.[107] AmberTools was then used to generate the charge parameters for each atom.

2.1.5 Data validation

Extensive care was put into validating the simulation data integrity to ensure all data corruption issues in the model and simulations have been fixed. These include: missing time values, erroneous coordinates, missing frames, etc. The minimum distance between the protein and its periodic neighbours were calculated post-simulation to check and ensure minimal self-interactions.

The simulations in this work all started with a well-equilibrated starting structure and underwent a long equilibration period before the production runs. By assessing pairwise root-mean-square deviation (RMSD) over the course of trajectories and time series data for system properties (Figure S1), large scale strucutral changes that were time-dependent were not observed. Hence, no cut-off was done to the initial segment of the simulations to account for a warm up period.

2.1.6 Representative Structures

Molecular dynamics sample a range of diverse structures therefore no single structure would best represent the entire system. I employed a simple procedure to find the most representative structure for each system that is then used for illustrative purpose throughout this thesis. For each system, the pairwise RMSD was calculated between all frames from the five production run trajectories, based on all protein heavy atoms. The structure that has the lowest combined distance to all other conformers in each system was picked as the representative structure.

2.1.7 Binding Region Clustering

Lacking a native binding site, small molecules interact with a plethora of residues during the simulation. This procedure aims to narrow down the binding regions to a few stable and long lasting ones that are likely to be the active binding pocket. To identify clusters of ligand positions, the simulated conformers were first superposed to its respective representa-

tive structure. The ensures that the ligands were in the relative position to the same protein conformer. The pairwise distance between the center of mass of each ligand position was then calculated. Density-based spatial clustering of applications with noise (DBSCAN) procedure as part of scikit-learn was used to find the clusters.[108,109] Eps, the maximum distance between two ligands' center of mass to be considered neighbours, was set to 10 Å and it required minimum 1% of the total number of frames to form a cluster.

2.2 Binding Energy Estimation

The molecular mechanics Poisson Boltzmann and surface area method (MMPBSA) utilises molecular mechanics and implicit continuum solvent model to estimate the free energies of ligand-receptor binding.[110] MMPBSA.py, part of the AmberTools, was used to carry out the calculations.[111] This methodology was used to calculate both the ligand binding energy and the peptide binding energy from segments of troponin. The entropic term is omitted in all MMPBSA based results shown in this work due to the computational intractability to accurately assess entropy for all simulations. The basic, albeit rather crude, assumption is that the entropy for similar interacting systems would be reasonably close hence it is acceptable to compare the MMPBSA energies without the entropic calculations.

2.3 Residue Flexibility

Root mean squared fluctuation (RMSF) was used to quantify the range of motion observed for each residue. It was calculated by aligning structures from one trajectory to a reference, in which case the first frame of the trajectory was used. An average C-alpha position for each residue was computed and the root-mean-squared deviation of C-alpha atom from its mean position was calculated. The 95% confidence interval for the fluctuation was calculated from the five repeats for each system.

2.4 Interhelical Angle Analysis

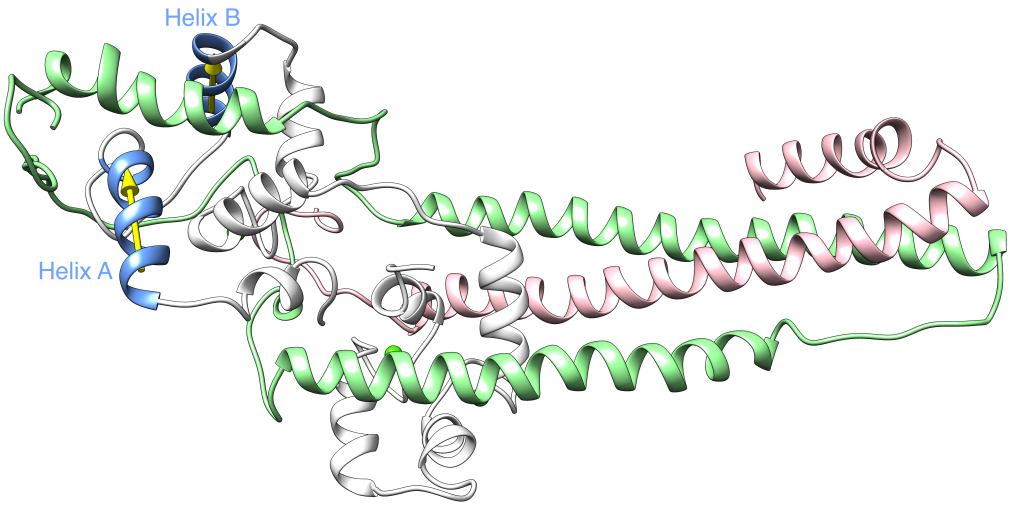


Figure 10: An illustration of the interhelical A/B angle on troponin. The three subunits of troponin are shown in white (TnC), light green (TnI), and light pink (TnT). Helix A and B of TnC are coloured in ligh blue. The two yellow arrows represent the vectors used to calculate the angle.

Computation of the angle between helix A and B of cTnC was implemented in Python with NumPy.[112] The coordinates of 10 backbone atoms at each end of each helix were extracted from the trajectory with Pytraj.[113,114] The coordinates at each end were averaged and used to calculate the vector that describes the direction of the helix (Figure 10). The interhelical angle was calculated from the two vectors. Note that the vector direction was from N-terminal to C-terminal for each helix, hence the reported angle is the supplement of the interhelical angle *between* the helices, which follows the conventions reported in previous publications.

2.5 Hinge Angle Analysis

Hinge angle is defined as the angle between the regulatory domain (vector from hinge to head) and the IT arm (vector from hinge to end of IT arm). The coordinates of the three points were defined as following:

- Head: the average of C-alpha coordinates of TnC_{3-85};

- Hinge: the average of C-alpha coordinates of TnC_{94-157};
- End of IT arm: the average of C-alpha coordinates of $TnT_{241-251}$, TnI_{69-76}

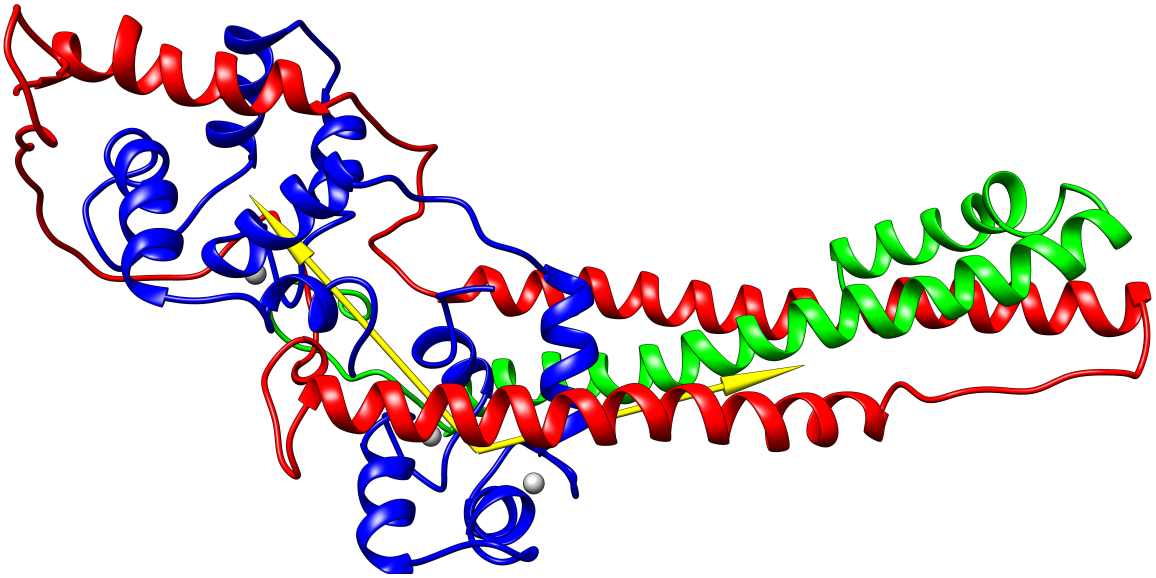


Figure 11: An illustration of the hinge angle on troponin. The two arrows represent the vectors used to calculate the angle.

2.6 Effect Size

MD simulations produce a large quantity of correlated data. Traditional statistical tests tend to deal with experimental data with a low number of observations. This means that most of the comparisons stemming from this study are likely to produce significant test statistics when the null hypothesis is that the mean of the two distributions are different although the actual difference could be negligible. Effect size is a more descriptive metric in that it qualitatively describes how big the difference is between the two samples.[115]

I used Cohen's D to assess the effect size between the means of the samples in this work which is defined as the following:

$$d = \frac{\bar{x}_1 - \bar{x}_2}{s} \tag{3}$$

where $\bar{x}$ is the sample mean and s is the pooled standard deviation. Note that d could be

negative depends on the order of comparison. Cohen defined s as:

$$s = \sqrt{\frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}} \tag{4}$$

where n is the sample size. The variance, s^2, is defined as:

$$s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2 \tag{5}$$

Colloquially, when interpreting the value of Cohen's D, some empirical descriptions are used to describe the absolute value of the quantity. With $|d| \sim 0.2$ called small, $|d| \sim 0.5$ medium, $|d| \sim 0.8$ large, and $|d| \sim 1.2$ very large. I adopted the same language when referring to effect size in this thesis.

2.7 Atomic Interaction Profiling

A major incentive for using all-atom simulation is that any observed large-scale structural changes can potentially be rationalised by changes in non-covalent residue interactions. I have employed two packages to quantify the interactions within the simulated systems, protein-ligand interaction profiler (PLIP) and Arpeggio.[116–118] PLIP uses geometry-based features such as bond length and bond angle to detect hydrogen bonds, hydrophobic contacts, π-stacking, π-cation interactions, salt bridges, water bridges, metal complexes and halogen bonds between ligands and targets. This affords us a systematic description of a ligand's interaction pattern with surrounding residues, particularly useful when looking at the results from MD simulations which contains tens of thousands of ligand binding poses.

Arpeggio employs a similar methodology but not limited to ligand interactions. It instead quantifies all interactions present in a given structure down to atomic level. This results in a much larger dataset than PLIP produces and necessitates additional considerations in subsequent analyses. The ability to profile the atomic interactions among residues over the course of a MD simulation provides additional insights into protein dynamics due to mutations or perturbations. And combined with time-series analysis methods, one can study the struc-

tural changes in a systematical way. One of the time-series analysis methods I explored was Granger causality which is detailed in the following section.

2.8 Granger Causality Analysis

Granger causality is a statistical hypothesis testing framework for asserting causal effects in time-series data and has been extensively used in economics and finance to genomics and neuroscience since its discovery.[119,120] It should be noted that causality is used in a restricted sense here where statistical causal effect observed in data does not necessarily mean the true causation of two events. Whether Granger causality truely uncovers causal effects has been debated since its inception. Granger causality boils down to the fundamental question: does the inclusion of past values of time-series $\mathbf{A}$ in addition to the past values of $\mathbf{B}$ improve the variance of the prediciton of future values of time-series $\mathbf{B}$?

When using the past values of a dataset to forecast itself, we can build a simple linear regression model based on n past values, at $t-1, t-2, ..., t-n$. This is known as an autoregressive model.

$$x_t = a_{t-1}x_{t-1} + a_{t-2}x_{t-2} + ... + a_{t-n}x_{t-n} + c \tag{6}$$

where x_t and a_t denote the value of the time-series and regression coefficient at time t, respectively. c is an added constant.

The original model proposed by Granger uses vector formulation hence utilises vector autoregression (VAR) which is a multivariate generalisation on the univariate autoregressive model. However, bivariate cases are more commonly considered in practice. For this project, bivariate autoregression is suitable as we know that the additional interactions introduced by perturbations is the independent variable.

We then construct two models where one depends on the past values of x only, Equation (7),

and one depends on the past values of both y and x, Equation (8).

$$\text{Model 1: } x_t = \sum_{i=1}^{n} a_{t-i} x_{t-i} + c \tag{7}$$

$$\text{Model 2: } x_t = \sum_{i=1}^{n} a_{t-i} x_{t-i} + \sum_{i=1}^{n} b_{t-i} y_{t-i} + c \tag{8}$$

Effectively, **Model 1** is a special case of **Model 2** where the coefficients for y are zero. When one model is a simplified version of another, they are known as nested models and likelihood ratio test is appropriate for comparing these two models.[121]

Granger causality is not widely adopted to analyse MD data yet. Previously, Sobieraj *et al.* used Granger causality on inter-residue distance to rationalise protein folding and unfolding dynamics in chignolin.[122] Other similar works used transfer entropy to assert causal relationships in residue pair motions.[123,124] To my knowledge, this is the first case where Granger causality is used to analyse atomic interaction data from MD at atomic level.

Another consideration in using Granger causality is the selection of a sensible lag time, i.e. how many historical values to use when forecasting future values. Unfortunately, there is no universal rule for lag time selection and it depends on the system under study. For the purpose of this work, I am interested in the propogation of direct molecular interactions hence limiting the maximum lag time to $20\,\mathrm{ns}$ (100 timesteps). Bayesian information criterion (BIC) is then calculated for all lag times and the model with the minimum BIC is selected as the optimal model.[125]

Within the context of this project, the core question answered with Granger causality is "which molecular interactions does the perturbation create or disrupt?" and there are a multitude of resolutions we can apply the methodology to. For example, we can test the interaction data on the atomic level or coarse grain the data to the residue level, and we can use either distance metric or binary on/off data for fitting the model. To maintain simplicity and aid interpretability, the residue level binary data was used in this work for Granger causality analysis.

3 Effects of Phosphorylation and Mutation on Troponin

In the healthy population, cardiac troponin possesses a phosphorylation-controlled calcium sensitivity which is crucial in response to a higher level of demanding physical activity or stress. Experiments have shown that many mutations within the thin filament would disrupt the coupled relationship between troponin's phosphorylation state and its calcium sensitivity.[48,126] Much effort has been put into understanding the molecular mechanism of the regulatory behaviour of phosphorylation and the disruptive consequences of mutations, both experimentally and computationally, while only studying a portion of the troponin trimer system. The rapid advance in recent years in both computational power and simulation software has allowed us to study the full core of the troponin complex over a microsecond timescale. We hope that in using a more complete structure of the troponin complex in this work would provide us the most comprehensive picture in how troponin's normal functionality is modulated by phosphorylation and disrupted with mutations.

To study the changes in structure and dynamics introduced by phosphorylation and mutation, molecular dynamics was used to simulate the wild type (WT) troponin and its TNNC1 G159D (G159D) mutant in both unphosphorylated (uP) and bis-phosphorylated and doubly-negatively-charged (SEP) states. Each system was repeated 5 times and each run lasted $1.5\,\mu s$. The full simulation protocol used in this work was given in Section 2.1.3. This work is built on past research in our group where troponin was studied with adaptive simulations coupled with Markov state models which attempted to illuminate the dynamics of the trimer system.[87,88] In comparison, I have fixed a couple of modelling errors in the previous work, namely residue protonation state and inadequate solvent box size, and employed novel analysis methods. In this chapter, I will first describe the global structure features observed in the four systems (WT uP, WT SEP, G159D uP, G159D SEP), measured changes in regional structural rearrangements, differences in peptide binding energy in regions of interest, and discuss the underlying molecular mechanisms for the observed changes based on residue interaction profiling on the simulated structures as well as through Granger causality analysis.

3.1 Disordered Regions Have Higher Flexibility

Root mean square fluctuation (RMSF) measures residue's range of movement within a simulation trajectory. A detailed description of the methodology was given in Section 2.3. In line with previous findings, the simulation with the newer ff14SB force-field showed a systematically lower fluctuations compared to older studies.[87] In general, the RMSF result shows that disordered regions of the troponin complex have a higher degree of flexibility. This means molecular dynamics simulations can reproduce experimentally observed secondary structure features of troponin which establishes more confidence in later analyses based on the simulated structures.

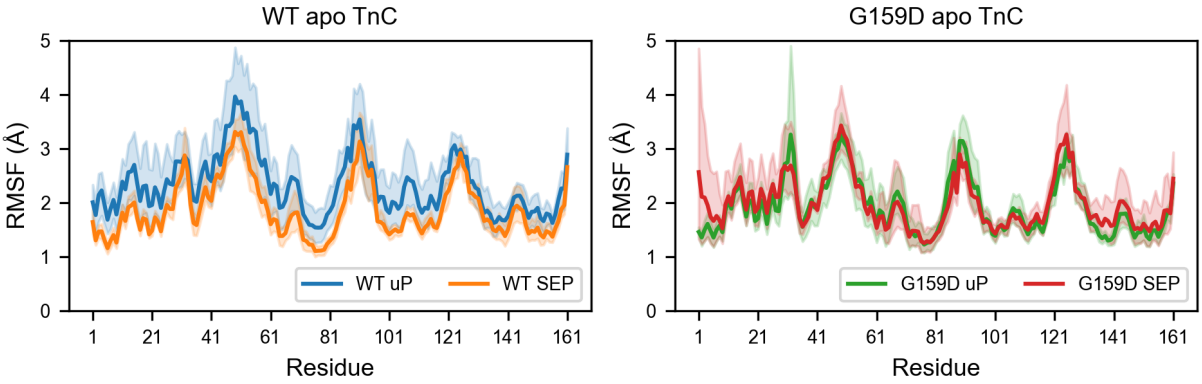


Figure 12: Root mean squared fluctuation for TnC. Shaded area represents the 95% confidence interval.

For wild-type TnC (Figure 12), the uP state showed a slighly higher flexibility in general while the 95% confidence intervals (shaded region) of RMSF of the uP and the SEP state have some overlap. For the G159D mutant, the mean RMSF for both states were very similar for most residues, except the first 10 residues of the N-terminal region which showed higher flexibility in the SEP state. Spatially, the N-terminal was close to the mutation site and this G159D SEP specific behaviour could be linked to the abnormal calcium sensitivity of G159D SEP (Figure 8). None of the residues in TnC had a large range of motion since the helical structures of TnC were ordered and only the loop regions between the EF-hands and the linker (residue 81-96) connecting the two domains had a higher degree of flexibility. The loop region between helix A and B (25-38) showed a more elevated RMSF than other ion binding sites, likely because this first ion binding site of TnC was unoccupied and related to the A-B angle open

and closed dynamics.

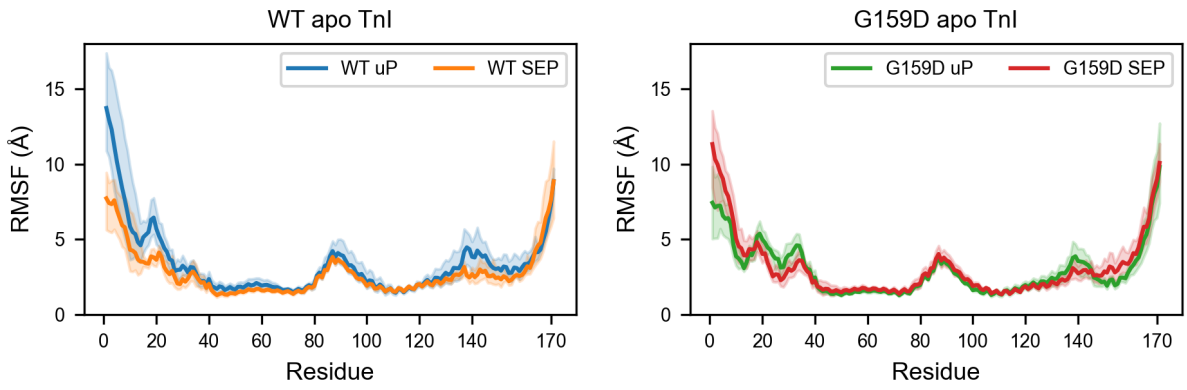


Figure 13: Root mean squared fluctuation for TnI. Shaded area represents the 95% confidence interval.

The most mobile parts of TnI were the two terminal domains where they were disoredered (Figure 13). In the N-terminal domain of TnI, residues around Arg20 showed a higher flexibility in WT uP while the RMSF for this region was reduced upon phosphorylation. The positively charged arginines in this region were found to interact with the neighbouring phosphoserines upon phosphorylation hence a reduced flexibility can be expected. On the other hand, similar RMSF values, around 5 Å, were found for this region in both phosphorylation states in the mutant. This reduced difference in calcium sensitivity upon phosphorylation in the mutant could be related to the reduced effect of phosphorylation on this region's flexibility. Region between residue 30 and 40 showed an increased flexibility in both states of the mutant but not in the wild type. This region connects the N-terminal domain to the helix H1 of the IT arm segment of TnI. Therefore, the increased flexibility could affect how the N-terminal segment interacts with TnC and lead to mutation-induced defects in phosphorylation-controlled calcium sensitivity. This is also the region that CcTnC starts to interact with TnI hence would also affect the anchoring of TnC to TnI. Residues 80-90 form an unordered U-turn connecting helix H1 and H2 of TnI hence a higher mean RMSF value was observed in all systems for this region. Elevated flexibility was also seen in the inhibitory peptide region (residues 137-148) where the uP state had a higher mean RMSF value than the SEP state in both wild type and the mutant. The switch peptide region (residues 149-163) showed a marginally higher flexibility in the uP state in the wild type while the opposite was observed in the mutant.

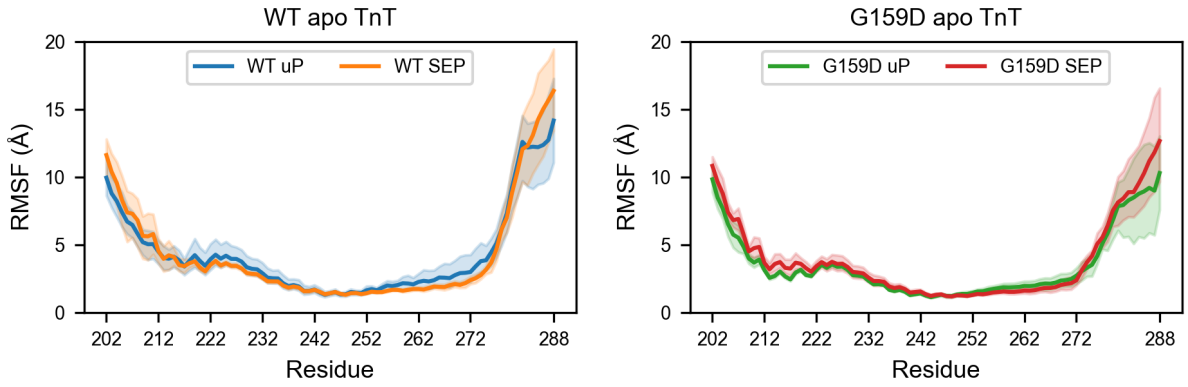


Figure 14: Root mean squared fluctuation for TnT. Shaded area represents the 95% confidence interval.

The TnT model included in this work mainly consists of two helical regions hence fewer flexible regions were observed (Figure 14). The two termini of the model indicated higher mobility however the N-terminal of the model starts from residue 202 therefore a higher RMSF can be expected from the truncated ends. The C-terminal end of TnT is thought to interact with actin or tropomyosin and anchor the troponin complex on the thin filament. The indicated higher flexibility of the C-terminal domain of TnT is likely to be a simulation artifact too as the TnT model in the simulations exists in isolation without interacting with the rest of the thin filament. The loop region between helix H1 and H2 of TnT showed a slightly higher RMSF value in all systems which is expected for unstructured regions.

3.2 Global Structure Changes from Principal Component Analysis

Principal component analysis (PCA) is a widely adopted dimensionality reduction method in the machine learning community. It is a linear dimensionality reduction method and is still used today owing to its simplicity and computational efficiency. PCA has increasingly been applied to analyse MD simulations in assessing large scale conformational changes and the principal motion of a trajectory.[127,128]

PCA was used to project frames of trajectories onto a lower dimensional space using Pytraj with their backbone heavy atom coordinates.[113] The first two principal components (PCs) only accounted for 30% of the observed variance in the backbone heavy atoms motion. The relative low percentage of the captured variance means that troponin undergoes complex

dynamical motions in all parts of the protein and the motions can not be reduced to a couple of coordinates.

Projection of explored conformations of the four states onto the first two PCs shows that majority of the conformational space overlaps and no state occupies a distinct cluster (Figure 15). The points in the projections represent frames from the trajectories. The projection shows a diagonal streaked pattern for all systems. At the extreme ends of the diagonals, the WT uP state is more prominent in the lower left corner while G159D SEP has a higher presence in the top right corner. It seems that structurally, all systems were in a similar projected space while phosphorylation and mutation biased troponin's structure to extreme ends of the space.

Inspecting the troponin strucures at the either end of the axes (Figure 16), we can see structural features that the PCs relates to. Broadly, the biggest difference is the change in the postition of the N-terminal domain of TnC in relationship to the IT arm. Measurement of this feature is discussed in detail in the following section.

3.3 Changes in Regional Structural Features

3.3.1 Interdomain Hinge Angle

The PCA analysis identified that the principal motion during troponin's dynamical movement was a large scale structural change between the regulatory head (N-terminal domain of TnC) and the IT arm (TnI, TnT, and C-terminal domain of TnC). The angle between the vectors of C-terminal TnC to N-terminal TnC and N-terminal TnC to the end of the IT arm (see Section 2.5 for a detailed methodology description) is termed the *interdomain hinge angle* and used to quantify the broad scale conformation of troponin.

Figure 17 illustrates the distribution of hinge angles for the four systems as kernel density estimation (KDE) plots, and the key measurements are tabulated in Table 4. KDE is a non-parametric method to estimate the underlying probability density function by modelling each observed data point as a Gaussian kernel. The wild type in both states showed a wider

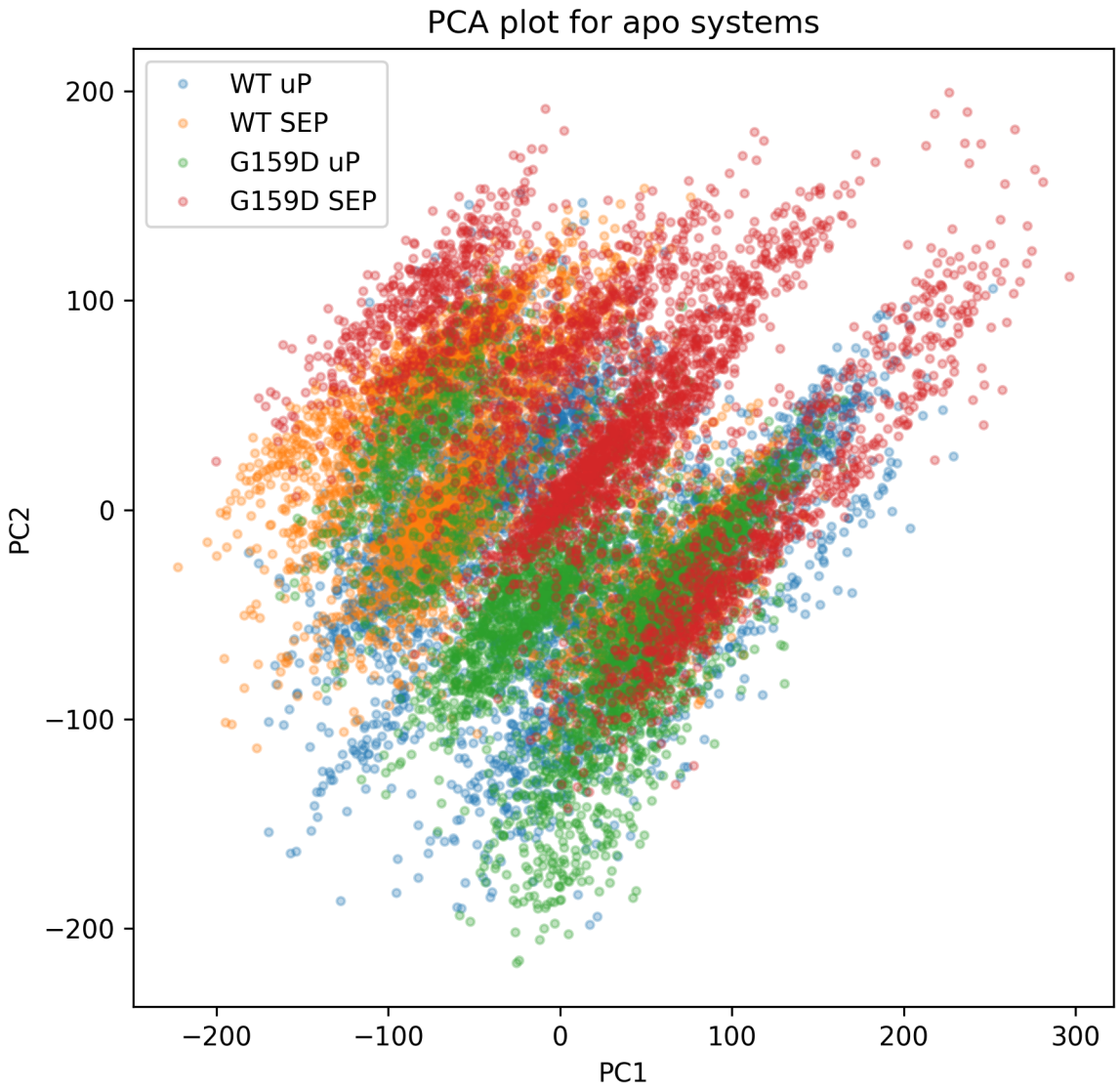


Figure 15: Scatter plot of troponin conformations in the first two principal components.

distribution than the G159D mutant. Upon phosphorylation, the WT peak got narrower and the mean was shifted to the higher angle direction (mean: WT uP 121.6°; WT SEP 123.3°). The opposite trend was observed in the mutant where the mean of the peak shifted from 121.5° to 116.2° with phosphorylation and the distribution was widened based on both standard deviation and the full-width at half-max (FWHM) metric upon phosphorylation. The G159D uP peak was centered at the same angle as the wild-type one, 121.5° and 121.6°, respectively, while the G159D uP peak was significantly narrower. Overall, the main mutation effect seems to be a reduction in the range of explored hinge angles in both states. And with phosphorylation, the wild type leans towards the more open flat conformation while the mutant tends to

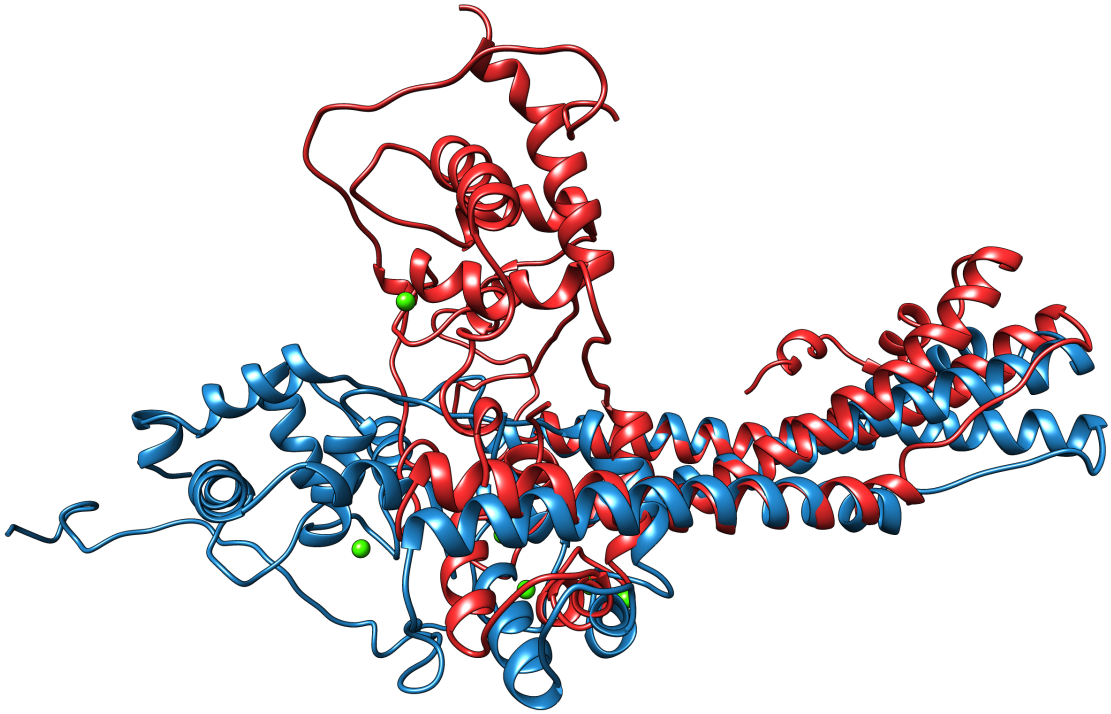


Figure 16: Depiction of troponin structures at the two diagonal ends (bottom left and top right) of the PCA plot. The structure coloured in blue is from WT uP simulations and represents the low PC1 and low PC2 state. The structure coloured in red is from G159D SEP simulations and represents the high PC1 and high PC2 state.

be more L-shaped and shifts to a more compact and rigid state.

	WT uP	WT SEP	G159D uP	G159D SEP
Peak	122.9°	122.5°	121.7°	117.1°
Mean	121.6°	123.3°	121.5°	116.2°
Std. Dev.	7.2°	6.0°	4.8°	6.0°
FWHM	22.0°	16.4°	11.5°	12.3°

Table 4: The hinge angle metrics for the four apo systems.

3.3.2 Contraction of The Regulatory Domain

Another notable feature from the PCA analysis is the contraction of the two lobes of TnC which are connected by a flexible linker. The separation between the N- and C-terminal domains were reduced in the high PC direction. I quantified the contraction by measuring the

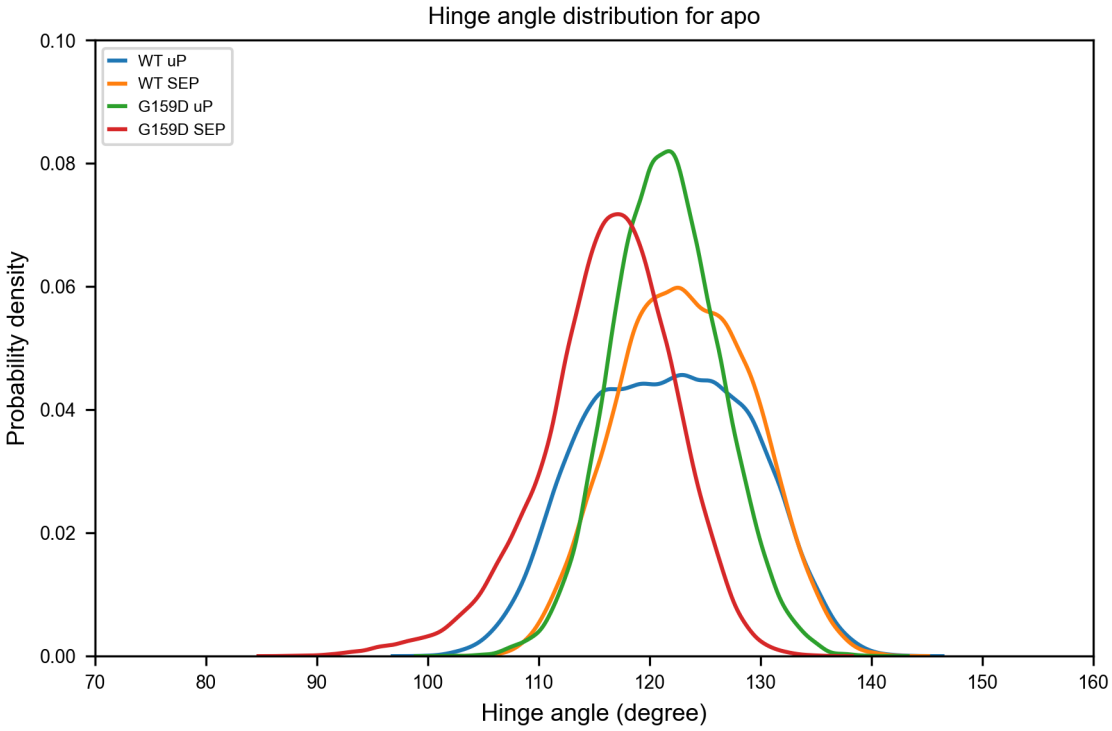


Figure 17: The hinge angle distributions as KDE plots.

distance between the centers of the two lobes.

	WT uP	WT SEP	G159D uP	G159D SEP
Peak	31.1 Å	29.6 Å	29.0 Å	29.2 Å
Mean	30.2 Å	29.9 Å	29.2 Å	29.6 Å
Std. Dev.	1.3 Å	0.9 Å	1.0 Å	1.2 Å
FWHM	3.0 Å	2.4 Å	1.6 Å	1.5 Å

Table 5: The TnC domain distance metrics for the four apo systems.

Here we see that the domain distance in TnC was reduced upon phosphorylation in the wild type, despite only a small change was observed for both the mean and the major peak. More pronounced difference was seen in the standard deviation and FWHM metrics where we see a reduction in the distribution of distances. In the mutant, the mean and the peak position were lower than WT SEP in both uP and SEP and the FWHM metric was significantly reduced from their respective WT state. Phosphorylation also caused a very slight increase in the mean and peak values in the mutant which was the opposite of the wild-type behaviour.

These metrics show that the G159D mutation caused compaction between the two lobes in TnC and reduced its range of motion.

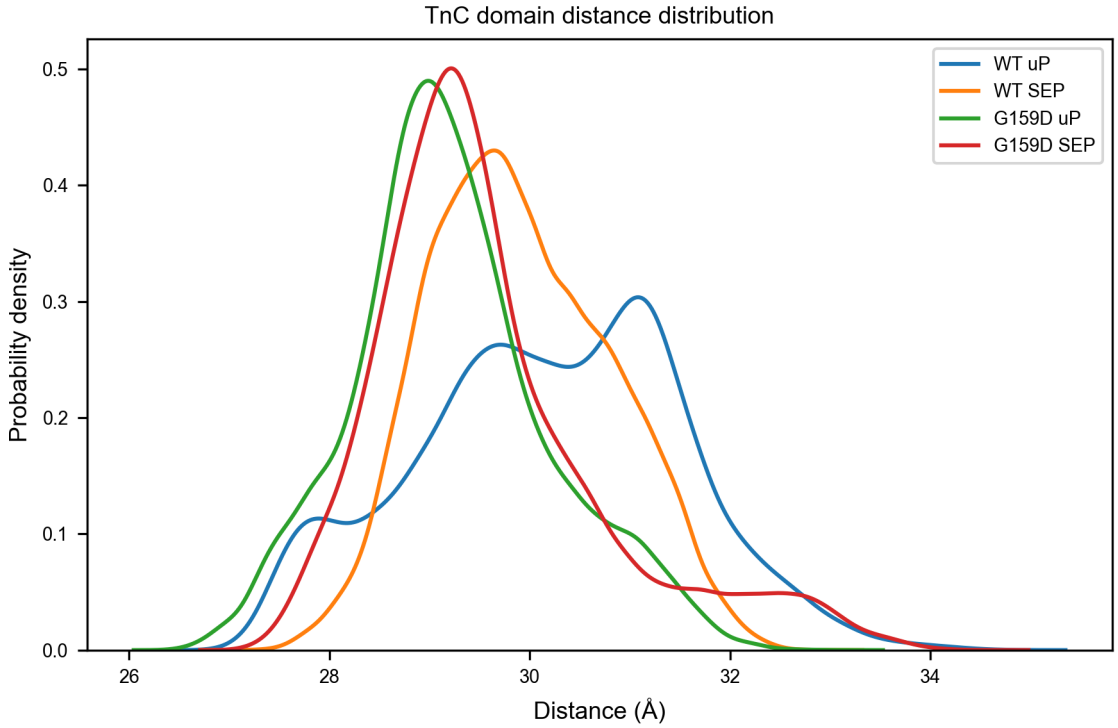


Figure 18: The TnC domain distance distributions as KDE plots.

The TnC domain distance distribution plot (Figure 18) was more evident of the compaction of TnC upon phosphorylation in the wild type as well as the effect of mutation. The WT uP state exhibited a much broader distance distribution, showing peaks centered around 27.5 Å, 29.5 Å, and 31.5 Å. Phosphorylation of the wild type reduced its width and the distribution was centered around 29.5 Å. On the other hand, the G159D mutation locked TnC into a closer packed conformation (around 29 Å) as the higher distanced density no longer existed and both peaks of G159D's TnC domain distance distributions were more alike than the wild type system.

3.3.3 Interhelical A-B Angle

The interhelical angle between helix A and B of TnC was previously used as a metric to quantify the hydrophobic patch's opening-closing dynamics. As suggested in the sarcomeric Ca^{2+}

mediated muscle contraction process, the opening of the hydrophobic patch is a crucial step in the switch peptide binding event, and it is of interest to study the observed interhelical angles for different systems (Figure 9). In this work, all simulations had the catalytic calcium bound and the majority of the interhelical angles were in the switch peptide-bound and A-B angle open state. The interhelical angle description was adopted from Kekenes-Huskey *et al.* where angles below 110° is open and angles above 130° is closed.[73]

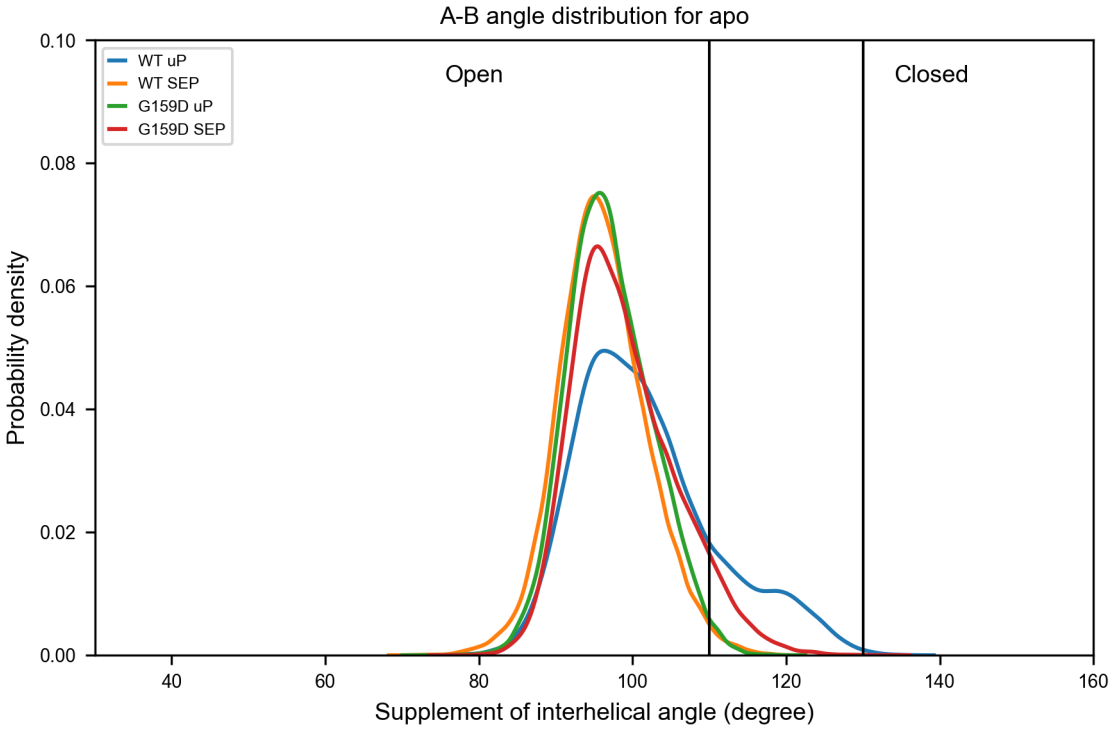


Figure 19: The interhelical A-B angle distributions for all apo systems.

In previous works, Cheng *et al.* used a similar model with TnI S22D/S23D phosphomimic systems to conclude that phosphorylation does not induce significant changes in the A-B angle.[129] The use of doubly charged phosphoserines and one order of magnitude more sampling seems to provide another picture of this metric. Figure 19 illustrates the distribution of the interhelical A-B angle for WT and G159D under both uP and SEP states. The distributions showed a predominent peak—the highest density point on KDE—which were centered close to each other (Table 6). Phosphorylation shifted the interhelical angle of the wild type to the more open direction. WT showed a pronounced reduction in peak angle upon phosphorylation, from 96.3° to 95.0°, where G159D's main peak saw a negligible change, from

95.7° to 95.4°. All distributions were right-skewed which is foreseeable as the systems are in the calcium-bound open state and the hydrophobic patch would naturally have more limited sampling of the higher-angled closed state. A bigger change in the data's mean upon phosphorylation was observed due to the skewness of the distributions, with a reduction from 101.8° to 96.2° for WT and a slight increase from 97.1° to 99.0° for G159D. This opposite phosphorylation-dependent interhelical angle dynamics was further reinforced when considering the percentage of the data points not in the open states. For unphosphorylated WT this was 18% and for G159D it was 1%; the corresponding values for the phosphorylated states were 1% (WT) and 7% (G159D). The range of angles sampled by each system or the spread of the distributions was measured with the standard deviation of the data and the full width at half maximum (FWHM) metric. Both metrics indicated that the WT apo showed the broadest distribution out of the four systems which was greatly reduced upon phosphorylation. Meanwhile, the G159D uP distribution closely resembled the WT SEP state, both in terms of peak height and spread, and more density was shifted to beyond 110° upon phosphorylation in the mutant.

	WT uP	WT SEP	G159D uP	G159D SEP
Peak	96.3°	95.0°	95.7°	95.4°
Mean	101.8°	96.2°	97.1°	99.0°
Std. Dev.	9.2°	5.7°	5.5°	6.7°
> 110°	18%	1%	1%	7%
FWHM	17.0°	12.2°	12.5°	13.6°

Table 6: The interhelical angle metrics for the four apo systems.

The simulations produced a large number of correlated conformations. Traditional statistical test would produce significant test statistics even though the difference between the compared datasets could be negligible. The effect size metric, Cohen's D, was used to quantify the magnitude of the difference between two datasets.[115] The methodology was detailed in Section 2.6. As shown in Table 7, phosphorylation on WT has a medium to large effect (0.74) on the interhelical angle while the effect is small to medium and the opposite direction (-0.31) in the mutant. In the same vein, the changes introduced by mutation in the uP state is medium

to large (0.63) and small to medium (-0.45) in the SEP state.

	WT SEP	G159D uP	G159D SEP
WT uP	0.74	0.63	0.35
WT SEP		-0.16	-0.45
G159D uP			-0.31

Table 7: The interhelical A-B angle pairwise effect sizes for different systems.

Overall, the data shows that for the A-B angle open/closed dynamics, WT apo naturally sampled a larger proportion of non-open states while phosphorylation shifted the troponin dynamics more towards the open state, as well as reduced the range of sampled angles. The opposite trend was observed in the mutant where SEP state showed a broader distribution and higher density above 110°. This further reinforces the observations laid out in Section 3.3.2 where phosphorylation in WT caused a rigidification of the NTnC lobe and the G159D mutation caused the opposite effects of phosphorylation. As the interhelical angle is a direct indicator of the hydrophobic patch-switch peptide binding/releasing event, the data suggests that phosphorylation can modulate the hydrophobic patch dynamics hence affects muscle relexation rates and the abnormal behaviour observed in the mutant disrupts this post-translational modification linked control. This shows a potential cause for how mutation can abolish the phosphorylation induced control in muscle relaxation time.

3.4 Differences in Peptide Binding Affinity

The binding energy of key segments of TnI has been calculated using the MMPBSA methodology. MMPBSA is an end-point binding free energy method that quantifies the Gibbs free energy based on molecular mechanics with polar solvation energy modelled by the Poisson-Boltzmann method and non-polar solvation energy estimated from the solvent accessible surface area.[130] Four regions of interest were identified to calculate their binding energy in contact with the rest of the troponin core complex. The following segments were chosen for their structural and functional significance. The positions of the segments on troponin are illustrated in Figure 20.

- the *switch peptide* is in close contact with the hydrophobic patch formed by EF-hand I of TnC. It plays a crucial role in muscle movement as its binding initiates the cross-bridge formation and starts the muscle contraction process.
- the *inhibitory peptide* inhibits the myosin binding sites on actin and is crucial in the muscle contraction and relaxation process.
- the TnI_{34-71} *peptide* anchors TnC to the IT arm and is the part of TnI that is next to the G159D mutation site. Changes in binding energy here could indicate changes in troponin's trimer intersubunit interactions due to mutation.
- the *NcTnI* (residue 1-32) region is important in regulating troponin's phosphorylation-dependent calcium affinity. Phosphorylation changes its interaction with NcTnC and mutations in troponin was shown to affect this region's binding interaction with NcTnC.

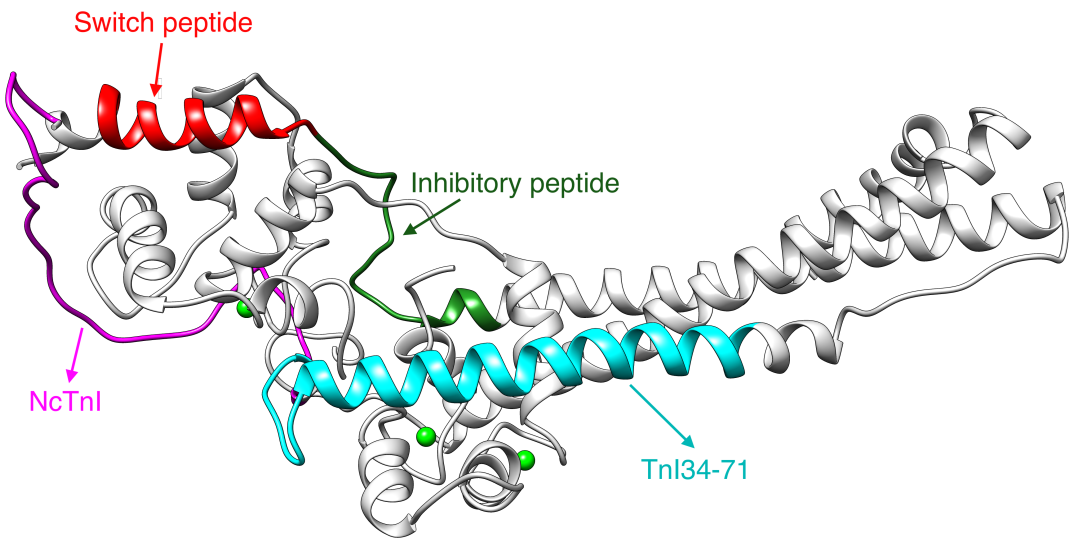


Figure 20: Illustration of the TnI peptide segments on troponin. The key segments are switch peptide (red), inhibitory peptide (green), TnI_{34-71} (cyan), and NcTnI (magenta).

3.4.1 Switch Peptide

Table 8 summarises the binding free energy of the switch peptide region to the rest of cTn. The mean and standard deviation for all four apo states are reported. The switch peptide bound stronger to the rest of the troponin complex upon phosphorylation in the wild type

	WT uP	WT SEP	G159D uP	G159D SEP
Mean	-108.4	-111.9	-113.5	-117.0
Std. Dev.	24.1	13.9	13.8	15.0

Table 8: Switch peptide binding energy to the rest of the troponin complex for the apo systems. Unit: $\mathrm{kcal\,mol^{-1}}$

($-108.4\,\mathrm{kcal\,mol^{-1}}$ to $-111.9\,\mathrm{kcal\,mol^{-1}}$). The same magnitude of increase was observed in the G159D mutant where the binding energy was $3.5\,\mathrm{kcal\,mol^{-1}}$ stronger in G159D SEP than G159D uP. However, the G159D mutant binding was $5\,\mathrm{kcal\,mol^{-1}}$ stronger than the wild type in both states. This means that the effect of the mutation at the C-terminal of TnC was causing changes in the N-terminal domain of TnC and led to a conformational change in the hydrophobic patch that made its binding to the switch peptide stronger. In a similar vein, phosphorylation of Ser22/23 on N-terminal TnI which are spatially close to the helix A and B of TnC caused the strengthened interaction between the hydrophobic patch and the switch peptide. Similar trend in the spread of the dataset in other metrics was seen in the switch peptide binding where the standard deviation of WT SEP was greatly reduced from WT uP and on par with G159D uP. The G159D SEP standard deviation was increased slightly from the unphosphorylated state but still much lower than WT uP. This trend reaffirms that the effect of phosphorylation and mutation is a rigidification of the regulatory head of troponin.

Baryshnikova *et al.* measured the dissociation constant of the switch peptide from NcTnC in the presence of $\mathrm{NcTnI_{1-29}}$ using NMR titration experiments.[131] The inclusion of the N-terminal segment of TnI in the experiment meant that the changes due to phosphorylation could be measured. This is most comparable to the WT results presented here as the NcTnC segment lacks the mutation site studied in this work. The measurement showed that the binding free energy was $-18.5\,\mathrm{kcal\,mol^{-1}}$ in the unphosphorylated case and slightly increased to $-19.7\,\mathrm{kcal\,mol^{-1}}$ upon phosphorylation. While the experimental binding free energy is not directly comparable to MMPBSA calculated binding energy, reassuringly, the data obtained from simulation is in the same trend as the experimental data.

	WT uP	WT SEP	G159D uP	G159D SEP
Mean	-153.7	-157.6	-168.5	-164.9
Std. Dev.	23.4	20.2	23.8	20.5

Table 9: Inhibitory peptide binding energy to the rest of the troponin complex for the apo systems. Unit: $\mathrm{kcal\,mol^{-1}}$

3.4.2 Inhibitory Peptide

Binding energy of the inhibitory peptide to the rest of the troponin complex increased by $3.9\,\mathrm{kcal\,mol^{-1}}$ with phosphorylation in the wild type (Table 9). The opposite trend was observed in G159D where the inhibitory peptide binding energy was weakened by phosphorylation ($-168.5\,\mathrm{kcal\,mol^{-1}}$ to $-164.9\,\mathrm{kcal\,mol^{-1}}$). A similar magnitude of decrease in the spread of the binding energy was seen with phosphorylation in both systems; in the wild type, the standard deviation reduced from $23.4\,\mathrm{kcal\,mol^{-1}}$ (uP) to $20.2\,\mathrm{kcal\,mol^{-1}}$ (SEP), and in the G159D mutant, the standard deviation reduced from $23.8\,\mathrm{kcal/mol}$ (uP) to $20.5\,\mathrm{kcal/mol}$ (SEP).

The sample mean shows that mutation increased the peptide binding energy by $14.8\,\mathrm{kcal\,mol^{-1}}$ in the unphosphorylated state and $7.3\,\mathrm{kcal\,mol^{-1}}$ in the phosphorylated state. And interestingly binding energy spread was not affected by mutation but only with phosphorylation in the inhibitory peptide case. Overall, the data shows that the G159D mutation led to a stronger inhibitory peptide binding to the rest of troponin by a significant amount. This could mean that returning to the myosin binding site inhibited state in the sarcomere will take longer and could lead to the phenotypes exhibited by dilated cardiomyopathies.

3.4.3 $\mathrm{TnI_{34-71}}$

	WT uP	WT SEP	G159D uP	G159D SEP
Mean	-215.7	-226.2	-212.4	-217.3
Std. Dev.	21.3	20.1	19.9	20.0

Table 10: $\mathrm{TnI_{34-71}}$ binding energy to the rest of the troponin complex for the apo systems. Unit: $\mathrm{kcal\,mol^{-1}}$

In the wild type, the TnI_{34-71} peptide binding energy (Table 10) was greatly increased with phosphorylation ($-215.7\,\mathrm{kcal\,mol^{-1}}$ to $-226.2\,\mathrm{kcal\,mol^{-1}}$) while the standard deviation reduced slightly ($21.3\,\mathrm{kcal\,mol^{-1}}$ to $20.1\,\mathrm{kcal\,mol^{-1}}$). In the G159D mutant, the mean binding energy increased by a small amount upon phosphorylation ($4.9\,\mathrm{kcal\,mol^{-1}}$) while the standard deviation remained at the same level. Comparing WT and G159D, it shows that the mutation reduced the peptide binding energy in both states and also reduced the increase in binding energy because of phosphorylation. NMR titrations were conducted by Baryshnikova *et al.* to assess the dissociation constant between the C-terminal domain of TnC and TnI_{34-71}.[132] It was found that the binding energy was stronger in the wild type than the G159D mutant in the unphosphorylated state (WT: $-34.5\,\mathrm{kcal\,mol^{-1}}$; $-31.7\,\mathrm{kcal\,mol^{-1}}$). The data obtained from simulations not only reproduced the trend in this mutation-induced effect but the difference is also comparable despite the lack of entropy calculations in the MMPBSA results. The weakened binding interaction could mean a weakened anchoring system of TnC to TnI.

3.4.4 NcTnI

	WT uP	WT SEP	G159D uP	G159D SEP
Mean	-129.2	-99.8	-111.3	-116.6
Std. Dev.	28.4	30.8	33.2	24.8

Table 11: Binding energy of NcTnI to the rest of troponin for the apo systems. Unit: $\mathrm{kcal\,mol^{-1}}$

In the wild type, the NcTnI binding energy was $-129.2\,\mathrm{kcal\,mol^{-1}}$ in the unphosphorylated state and greatly reduced to $-99.8\,\mathrm{kcal\,mol^{-1}}$ upon phosphorylation (Table 11). In contrast, the difference between NcTnI binding energy under the two phosphorylation states was much reduced in the mutant. G159D uP exihibits a $-111.3\,\mathrm{kcal\,mol^{-1}}$ binding energy in the uP state and increased slightly to $-116.6\,\mathrm{kcal\,mol^{-1}}$ in the SEP state. The standard deviations of data for the two phosphorylation states of WT were similar, $28.4\,\mathrm{kcal\,mol^{-1}}$ (uP) and $30.8\,\mathrm{kcal\,mol^{-1}}$ (SEP). The binding energy standard deviation saw a slight increase to $33.2\,\mathrm{kcal\,mol^{-1}}$ in G159D uP and was greatly reduced to the lowest of the group in G159D SEP, $24.8\,\mathrm{kcal\,mol^{-1}}$.

It is expected to see direct phosphorylation-dependent differences here as the phosphorylation sites are located on NcTnI. And indeed the results showed a significant reduction of binding energy upon phosphorylation in the wild type. The four negative charges introduced by phosphoserines interacted with their neighbouring postively-charged arginines therefore disrupted the interactions with NcTnC. This observation agrees with the experimental 2D NMR study where phosphorylation was found to change NcTnC domain positioning due to modified interactions with NcTnI.[23] The effect of the G159D mutation in the uP state seems to be a reduction of the binding affinity between NcTnI and surrounding residues. Interestingly, the opposite phosphorylation effect was observed in the mutant compared to the wild type, the G159D SEP had a stronger NcTnI binding interaction in relationship to the uP state. Overall, this dataset showed the effect of phosphorylation on NcTnI in the wild type; the G159D mutation completely disrupted the normal phosphorylation induced changes in binding energy of this crucial regulatory segment.

3.5 Residue Interaction Profiling

Molecular dynamics at the atomistic level is a powerful tool in simulating the atomic interactions among atoms which form the marcromolecules and, over time, we observe the local structural and global conformational changes. All changes mentioned in the previous analyses are downstream consequences of local changes in atomic interactions, more specifically, non-covalent bond forming and breaking. When investigating residue-residue contacts in previous studies, distance cutoff was a more commonly used method. However, this method requires setting a cutoff distance and does not provide interaction type level information. Arpeggio, originally developed by Jubbs *et al.*, uses distance and angle to determine the possible atomic interactions among neighbouring atoms.[118] Arpeggio directly looks for interacting atom pairs and characterises the potential interaction types based on geometry thereby provides richer details. Hence, to understand the origin of structural and energetics changes due to phosphorylation and mutation, I used Arpeggio to profile the protein's atomistic interaction pattern for each frame of the trajectories. Given the large quantity of atomic level interaction data, I further aggregated the data to the residue level—for each frame, if a particular type of interaction existed between any of the consituent atoms of one residue to another

then this particular type of interaction was counted between the two residues. The result was presented as the percentage of the total simulation where the interaction existed for each system.

The three subunits of troponin are well interlocked to each other through molecular interactions at the intersubunit interfaces. Figures S3 to S22 illustrate the ionic interaction pattern between troponin subunits as heatmaps. The major interaction points between the subunits and within the components themselves were first identified. Many of the observed changes in interactions occured on the peptides that existed as part of the original crystal structure which the MD models are based on. The modelling of these regions presents a higher confidence that simulations can represent troponin's native dynamics. The results presented in this section focus on ionic interactions and some prominent hydrogen bonds. Interactions between residue pairs 3 or 4 residues apart and as part of the same helix tend to be stable and serve a structural purpose. These interactions are not discussed in detail for brevity.

3.5.1 CcTnC and TnI Anchor Region

Between TnC and TnI (Table 12), the TnC Asp2 and TnI Lys45 ionic interaction firmly anchored NcTnC to helix H1 of TnI which was formed in more than 95% of the simulations except G159D SEP (89%). Another adjacent ionic interaction, between TnC Asp139 and TnI Lys45, also had a high propensity of formation with >80% in both WT phosphorylation states, 77% in G159D SEP, and a reduced 64% in G159D uP. Notably, Asp2 is located at the beginning of TnC and Asp139 is on helix G, the penultimate helix of the nine-helix dumbbell-shaped subunit. These two long lasting ionic interactions essentially formed the interface between the regulatory head and the IT-arm by joining the two ends of TnC with the starting point of TnI helix H1 which was right after the disordered N-terminal segment of TnI. The reduction in interaction count between TnC Asp139 and TnI Lys45 in the mutant was likely caused by the repositioning of the EF-hand IV (consists of helix G and H) due to the additional interactions formed with TnC Asp159 in the mutatant. This finding reinforces the results from the TnI_{34-71} binding energy measurements (Section 3.4.3) where the mutation was found to reduce the binding of the TnI segment to the rest of troponin which destabilises the anchoring

of CcTnC to TnI.

The next set of interactions are between the CcTnC and TnI interface. The aforementioned TnC Asp139 also formed ionic interaction with TnI Lys49 of TnI H1 for 35% and 31% of the time in WT uP and WT SEP, respectively. This interaction was increased to 39% in G159D uP and 43% in G159D SEP. TnI Arg44 was found to interact ionically with both TnC Glu126 and TnC Asp132 although minimal bridging was observed—TnI Arg44 interacted with both TnC residues simultaneously for less than 4% of the total simulation. Parallel to the TnC Asp139/TnI Lys45 pair was another ionic interaction connecting TnC helix G with TnI, TnC Glu135 and TnI Lys39, which was observed between 15% and 22% for all systems. TnC Glu135 also forms ionic interaction with TnI Lys45 with low percentage, 16% and 26% in WT uP and SEP, and 18% and 5% in G159D uP and SEP. Lastly, the last residue on the C-terminal of TnC, Glu161 formed ionic interaction with TnI Lys57 for 38% in WT uP, 35% in WT SEP and elevated to 49% in G159D uP, 60% in G159D SEP.

		WT uP	WT SEP	G159D uP	G159D SEP
TnC Asp2	TnI Lys45	97%	96%	98%	89%
TnC Asp139	TnI Lys45	82%	80%	64%	77%
TnC Glu135	TnI Lys45	16%	26%	18%	5%
TnC Asp139	TnI Lys49	35%	31%	39%	43%
TnC Glu135	TnI Lys39	16%	15%	17%	22%
TnC Glu126	TnI Arg44	51%	51%	45%	49%
TnC Asp132	TnI Arg44	28%	31%	34%	37%
TnC Glu161	TnI Lys57	38%	35%	49%	60%
TnC Arg102	TnI Glu63	15%	23%	19%	11%

Table 12: Ionic interactions at the CcTnC and TnI interface that anchored TnC to TnI.

3.5.2 Switch Peptide Region

Key ionic interactions formed close to the switch peptide region (Table 13). The TnC Glu19 and TnI Arg161 ionic interaction locked switch peptide with helix A which formed part of the

hydrophobic patch. This was the major ionic interaction that contributed to the switch peptide binding to TnC and appeared in more than 80% of the simulation in both states of the WT (uP 86%, SEP 81%) and G159D SEP (87%). In G159D uP, this interaction was reduced to 68%. Another ionic interaction was between TnI Glu164 and TnC Lys43 of helix B. This interaction existed in 18% of the simulation in WT uP but was reduced to 2% in WT SEP. G159D uP had a similar interaction percentage to WT SEP, at 2% and G159D SEP brought the frequency up to the WT uP level, at 19%. TnC Glu19 was involved in another prominent ionic interaction with TnI Arg9 which anchored the start of the TnC N-terminal region to the tip of the regulatory head. This interaction presented in around 70% of the trajectories in both WT states, 77% in G159D uP, and reduced to 44% in G159D SEP. TnI Arg9 interacted ionically with TnC Glu15 of helix A for around 50% of the time for WT uP, WT SEP, and G159D uP, and reduced to 39% in G159D SEP. TnI Arg161 also formed ionic interaction with TnI Glu10, essentially linking the start of NcTnI with the C-terminal end of the switch peptide. The overall effect was a clustered ionic interactions among two arginines and three glutamic acids that anchored the beginning of NcTnI to helix A of TnC, which was located at the tip of of the regulatory head, and bridged to the switch peptide portion of TnI.

		WT uP	WT SEP	G159D uP	G159D SEP
TnI Glu10	TnI Arg161	51%	65%	73%	40%
TnC Glu19	TnI Arg161	86%	81%	68%	87%
TnC Lys43	TnI Glu164	18%	2%	2%	19%
TnC Glu19	TnI Arg9	70%	71%	77%	44%
TnC Glu15	TnI Arg9	50%	48%	56%	39%

Table 13: Ionic interactions at the hydrophobic patch and switch peptide region.

3.5.3 Inhibitory Peptide Region

There are six positively charged residues (two lysines and four arginines) within the inhibitory peptide of TnI. Minimal self-interaction via ionic interactions was observed between inhibitory peptide and the rest of TnI. The ionic interactions between inhibitory peptide and TnT were scarce too. With TnC, majority of the interactions were localised on helix C which is part of

EF-hand II that binds to the catalytic Ca^{2+} (Table 14). TnI Arg140 formed ionic interactions with TnC Asp62 during majority of the simulations (WT uP 76%; WT SEP 87%; G159D uP 84%; G159D SEP 75%). For a smaller percentage of the time, TnI Arg140 briged between TnC Glu59 and TnC Asp62. The interaction frequency between TnI Arg140 and TnC Glu59 was between 22% and 33% for WT uP, WT SEP, and G159D SEP. The unphosphorylated G159D had a much elevated interaction count of 75% where the coexistence of both ionic interactions accounted for 63% of the simulations. More than half of the WT uP trajectories contained a TnI Arg144-TnC Glu56 ionic interaction (56%). This was reduced to 40% upon phosphorylation and again reduced to 27% in G159D uP with a slight increase to 35% in G159D SEP. For TnI Arg145, it bridged between TnC Glu59 and Glu63 and the frequency counts increased in the mutant for both states. TnI Arg145 also formed an ionic interaction with TnC Glu155 and the frequency counts increased from around 25% in the wild type to more than 50% in the mutant for both phosphorylation states. Upon inspecting the structure, it was noticable that the distance between TnI Arg145 and TnC Glu155 was prohibitive of any interactions when the two lobes of TnC were separated. It was only when the regulatory head was contracted and the distance between the two lobes reduced then this interaction could form. As detailed in Section 3.3.2, the G159D mutation locked TnC into a closer packed conformation hence this ionic interaction was more pronounced in the mutant. The last arginine of the inhibitory peptide bridged between TnC Glu63 and TnC Asp88. These two ionic interactions linked up helix C of TnC to the inhibitory peptide to the flexible linker between the two lobes of TnC. Both interactions presented between 52% and 71% of the simulations in all systems. Overall, the ionic interactions from the inhibitory peptide region firmly fixed the inhibitory peptide to TnC helix C and shared the connections to neighbouring TnC linker as well as the C-terminal end of TnC.

3.5.4 Catalytic Calcium Binding Site with TnI

TnI Lys35 was the main residue of TnI that ionically interacts with residues that formed part of the EF-hand II ion binding site (Table 15). The TnI Lys35/TnC Asp73 and TnI Lys35/TnC Glu76 ionic interactions presented in more than 50% of the simulations in both states of the wild type and reduced to 26-28% in G159D uP; both interactions slightly increased to 34%

		WT uP	WT SEP	G159D uP	G159D SEP
TnC Glu59	TnI Arg140	33%	22%	75%	27%
TnC Asp62	TnI Arg140	76%	87%	84%	75%
TnC Glu55	TnI Arg140	6%	3%	24%	1%
TnC Glu56	TnI Arg144	56%	40%	27%	35%
TnC Glu94	TnI Arg144	2%	2%	23%	25%
TnC Glu59	TnI Arg145	27%	21%	51%	41%
TnC Glu63	TnI Arg145	68%	72%	85%	83%
TnC Glu155	TnI Arg145	21%	28%	53%	50%
TnC Glu63	TnI Arg147	67%	67%	66%	71%
TnC Asp88	TnI Arg147	60%	57%	52%	57%
TnC Asp87	TnI Arg147	12%	23%	8%	5%
TnC Asp151	TnI Arg135	34%	41%	37%	53%

Table 14: Ionic interactions between inhibitory peptide and TnC.

in G159D SEP. The ionic interaction between TnC Asp75 and TnI Lys35 existed in 41% of the trajectories in WT uP and decreased to 27% upon phosphorylation. In the mutant, this interaction accounted for 25% of the simulations in both states. It should be noted that although TnC Asp75 is part of the loop of EF-hand II it did not contribute to calcium ion binding directly. The TnI Lys35/TnC Asp67 interaction accounted for 20% and 34% of the simulations for wild-type uP and SEP states, respectively. The percentages were reduced to 14% and 13% in the mutant. Additionally, a low frequency interaction was observed between TnC Asp75 and TnI Lys37 with 17% and 25% in WT uP and WT SEP, and reduced to 13% and 8% in G159D uP and G159D SEP. Overall, the interactions presented in this region influenced the positioning of NcTnI in relationship to NcTnC and the observed reduction in interaction frequency in the mutant would contribute to the reduced NcTnI binding energy reported in Section 3.4.4 (Figure 21).

		WT uP	WT SEP	G159D uP	G159D SEP
TnC Asp73	TnI Lys35	52%	52%	28%	34%
TnC Glu76	TnI Lys35	52%	54%	26%	34%
TnC Asp75	TnI Lys35	41%	27%	25%	25%
TnC Asp67	TnI Lys35	20%	34%	14%	13%
TnC Asp75	TnI Lys37	17%	25%	13%	8%

Table 15: Ionic interactions between TnI and the catalytic calcium binding site.

3.5.5 NcTnC and NcTnI

Some major ionic interactions between the arginines of NcTnI and the negatively charged residues on NcTnC attached the disordered TnI region to the N-terminal of TnC (Table 16). TnI Arg12 bridged between TnC Asp25 and TnC Glu32. Both TnC residues were located on the loop region between helix A and helix B and a change in the interaction pattern here likely indicates a modulation on the interhelical A-B angle. For both interactions, an increase in the interaction counts was observed upon phosphorylation of the wild type. The G159D uP state was similar in interaction percentage to the WT SEP state and decreased upon phosphorylation for the mutant. These phosphorylation-dependent interactions had opposite changes in the mutant when compared to the wild type and this trend seemed to coincide with the angle measurement observations as detailed in Section 3.3.3. The ionic interaction between TnC Asp33 and TnI Arg19 existed for 38% in WT uP and 25% in G159D uP and decreases to 8% in WT SEP and 12% in G159D SEP. A similar reduction in interaction percentage upon phosphorylation was seen in the ionic interaction between TnC Asp33/TnI Arg20 and TnC Asp33/TnI Arg21. Asp33 is located on the loop region between helix A and B and is one of the cardiac specific variant of TnC (glycine in skeletal isoform). This change influenced how NcTnI was positioned on NcTnC and in relationship to the hydrophobic patch which would affect the catalytic calcium binding/unbinding process. The ability of TnC Asp33 to form interactions with the cluster of TnI arginines signifies the importance this anchor point to the relative positioning of NcTnI to NcTnC. Cumulatively, the TnC Asp33/TnI Arg ionic interactions accounted for 77% and 74% of the simulations in WT uP and G159D uP. This was reduced to 21% and 31% upon phosphorylation in WT SEP and G159D SEP, respectively. TnI

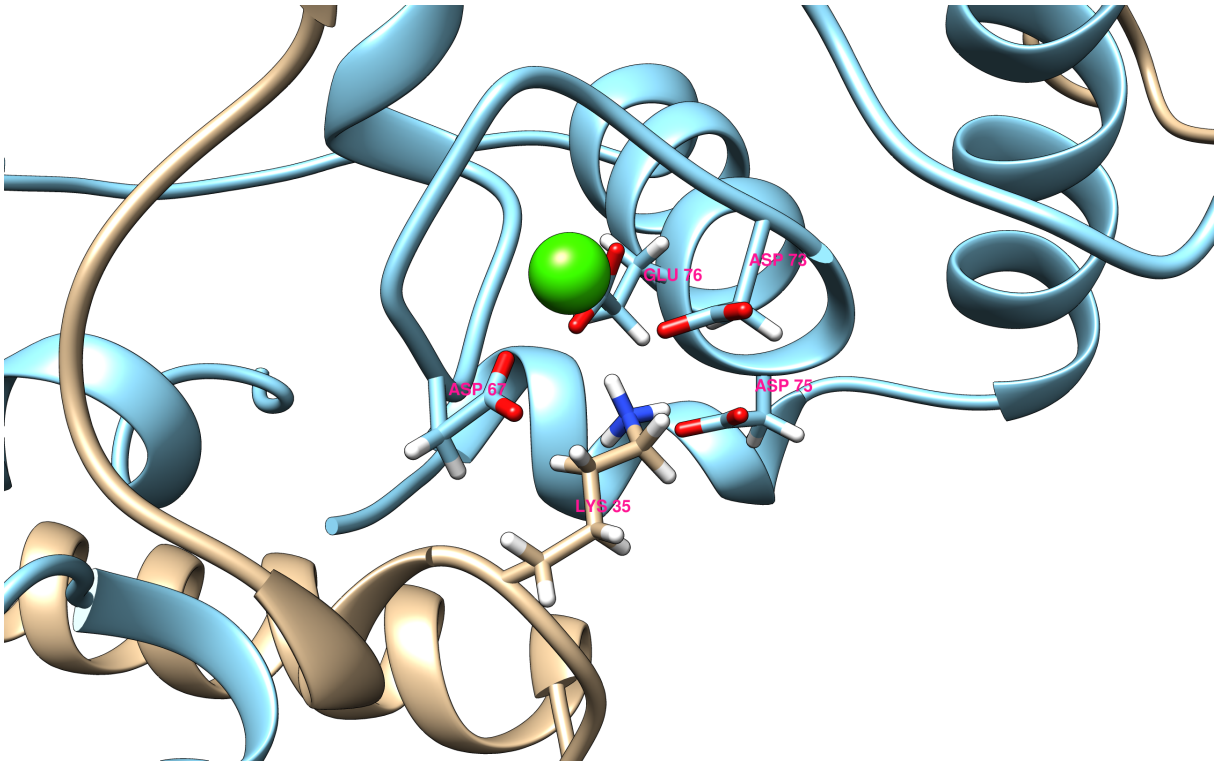


Figure 21: Depiction of the charged residues near the catalytic calcium binding site that interacted with TnI Lys35. TnI subunit is coloured in beige and TnC subunit is coloured in blue.

Arg21, adjacent to the phosphorylation site Ser22, formed ionic interaction with TnC Glu40, part of TnC helix B. This interaction was reduced upon phosphorylation in the wild type, 22% to 2%, but increased from 29% to 43% with phosphorylation in the mutant. The TnC Glu55 and TnI Arg26 ionic interaction fixed NcTnI to TnC helix C and existed on a similar level in both unphosphorylated wild type and the mutant (WT uP 65%, G159D uP 67%) and reduced to 29% in WT SEP and drastically reduced to 7% in G159D SEP. The reduction in interaction frequency observed with TnI Arg19, Arg20, Arg21, and Arg26 upon phosphorylation was due to NcTnI self-interactions. The extra negative charges introduced by phosphoserines formed ionic interactions with the arginines hence limited their ability to maintain other ionic interactions.

3.5.6 TnC and TnT Interface

At troponin's hinge region, the main anchor point of TnC is with TnI where the two EF-hands of CcTnC clamp helix H1 of TnI while the opposite face forms interactions with helix H2 of

		WT uP	WT SEP	G159D uP	G159D SEP
TnC Asp25	TnI Arg12	63%	75%	77%	60%
TnC Glu32	TnI Arg12	42%	57%	52%	32%
TnC Asp33	TnI Arg19	38%	8%	25%	12%
TnC Asp33	TnI Arg20	33%	13%	34%	24%
TnC Asp33	TnI Arg21	15%	1%	19%	0%
TnC Glu40	TnI Arg21	22%	2%	29%	43%
TnC Glu55	TnI Arg26	65%	29%	67%	7%

Table 16: Ionic interactions between NcTnI and NcTnC.

TnT. Minimal long-lasting ionic interactions were found between TnC and TnT except one between TnC Asp151 and TnT Arg267 (Table 17). This interaction was present in 90% of WT uP and 98% of WT SEP. The G159D mutation in the uP state reduced the interaction frequency to 77% which reverted back to 96% in G159D SEP. TnC Arg147 also formed ionic interaction with TnT Asp270 for 10% of the time in WT uP which increases to 37% upon phosphorylation. In the mutant, this interaction was a reduced 6% in G159D uP and increased to 23% in G159D SEP. TnT Arg267 formed another ionic interaction with TnC Glu94 which was located at the C-terminal end of the TnC flexible linker. This interaction was 24% and 10% in WT uP and SEP and mutation caused a reduction to 10% and 1% in G159D uP and SEP. The drop in interaction counts in the mutant could be related to the destablised TnC anchoring to the IT arm.

		WT uP	WT SEP	G159D uP	G159D SEP
TnC Asp151	TnT Arg267	90%	98%	77%	96%
TnC Arg147	TnT Asp270	10%	37%	6%	23%
TnC Glu94	TnT Arg267	24%	10%	10%	1%

Table 17: Ionic interactions at the TnC and TnT interface.

3.5.7 Phosphorylation Induced Changes

Cheng *et al.* studied the TnI S22D/S23D phosphomimic system and found that phosphorylation led to the formation of intrasubunit interactions between the N-terminal domain (NcTnI) and the inhibitory peptide (TnI_{IP}) of cTnI.[129] From arpeggio results, no ionic interactions between NcTnI and TnI_{IP} were observed but there were elevated hydrogen bond interactions between TnI Arg140 and the backbone oxygen of residues Arg26, Ala27, and Tyr28 upon phosphorylation (Table 18). Figure 22 illustrates the change in hydrogen bond interactions between NcTnI and $TnI_{136-170}$. An overall gain of interactions were observed for both WT and G159D. In the wild type, the guanidino group of TnI Arg140 formed hydrogen bond with the backbone of TnI Arg26 in 49% of the unphosphorylated simulations, and this was increased to 72% upon phosphorylation. The unphosphorylated mutant showed a greatly reduced interaction frequency compared to the wild type but elevated to a similar level upon phosphorylation (17% to 63%). A similar trend was also observed in the pairs of Arg140/Ala27 and Arg140/Tyr28 in the wild type where the interaction frequency increased with phosphorylation. However, a minimal increase was seen in the mutant's case.

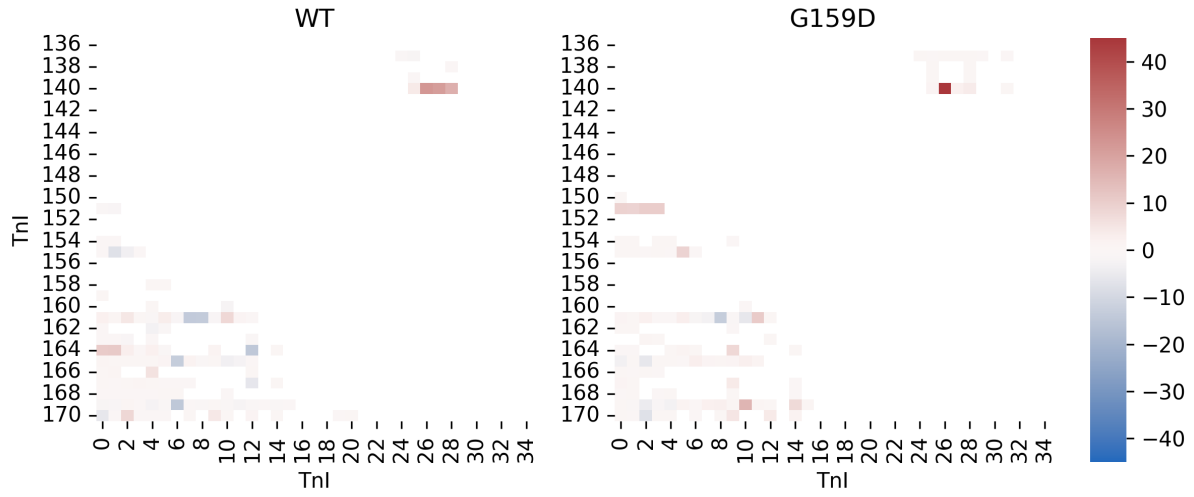


Figure 22: Heatmap showing the difference in hydrogen bond interaction percentage upon phosphorylation for WT and SEP. The graph focuses on the NcTnI and C-terminal end of TnI of the model used in this work.

The two phosphoserines introduce a range of new ionic interactions to the systems (Table 19). Between TnC and TnI, the only new interaction was with TnC Lys39. TnC Lys39 and TnI Sep23

		WT uP	WT SEP	G159D uP	G159D SEP
TnI Arg26	TnI Arg140	49	72	17	63
TnI Ala27	TnI Arg140	10	31	4	6
TnI Tyr28	TnI Arg140	7	24	6	10

Table 18: Hydrogen bonding between NcTnI and TnI_{IP}.

formed ionic interaction for 67% and 90% of the simulations in WT SEP and G159D SEP, respectively. A small percentage of interaction was also observed between TnC Lys39 and TnI Sep22 in WT SEP (12%) and none in G159D SEP. TnC Lys39 is located on helix B of the hydrophobic patch and interacting with phosphoserines represents a repositioning of NcTnI upon phosphorylation which in turn can influence the hydrophobic patch open/closed dynamics. Comparing WT SEP with G159D SEP, it can be seen that a higher percentage of TnC Lys39/TnI Sep23 interaction coincided with a higher percentage of TnC Glu40/TnI Arg21 interaction. It appears that the TnC Lys39/TnI Sep23 and TnC Glu40/TnI Arg21 ionic interaction pairs stabilised each other where the formation of one interaction is likely to lead to the formation of the other interaction.

		WT SEP	G159D SEP
TnC Lys39	TnI Sep23	67%	90%
TnC Lys39	TnI Sep22	12%	0%
TnI Arg19	TnI Sep22	24%	43%
TnI Arg20	TnI Sep22	10%	41%
TnI Sep22	TnI Lys35	3%	0%
TnI Arg19	TnI Sep23	68%	49%
TnI Arg21	TnI Sep23	72%	90%
TnI Sep23	TnI Arg26	95%	99%
TnT Lys280	TnI Sep22	20%	15%
TnT Lys280	TnI Sep23	10%	0%

Table 19: Ionic interactions from TnI phosphoserines 22 and 23.

There were some major TnI ionic self-interactions with the adjacent arginine residues in the

SEP state where in general TnI Sep23 formed more longer lasting interactions. Both TnI Sep22 and Sep23 interacted with TnI Arg19 in the WT SEP state with 24% and 68% interaction frequency, respectively. In the mutant, the Sep22 interaction was increased to 43% while the Sep23 interaction was reduced to 49%. This indicates a difference in the local NcTnI arrangement caused by the mutation where TnI Arg19 had higher propensity to form interaction with Sep23 in the wild type and the difference in propensity was much reduced due to mutation. TnI Arg20 exclusively formed ionic interaction with Sep22 with 10% in WT SEP and 41% in G159D SEP. Both Arg21 and Arg26 formed long lasting ionic interactions with Sep23. The interaction percentage in WT SEP was 72% for Arg21 and 95% for Arg26. In the mutant, both interaction frequencies were increased with Arg21 to 90% and Arg26 to 99%. Lastly, despite an almost negligible 3% interaction count it was interesting to see Sep22 get in a close proximity with TnI Lys35 which was found to directly interact with multiple residues in the catalytic calcium binding site.

A handful of low-count ionic interactions also formed between Sep22/23 and TnT. The only one with greater than 10% interaction frequency was with TnT Lys280 at the fulcrum region of troponin. Between TnI Sep22 and TnT Lys280, the interaction represented 20% of the simulation in the WT SEP state and droped to 15% in the mutant. With TnI Sep23, the interaction was 10% in the wild type and vanished in the G159D SEP state.

Overall, the findings here illustrate how phosphorylation can change the disposition of NcTnI in relationship to NcTnC and the rest of troponin. The altered interactions between NcTnI and TnC could play a crucial role in modulating troponin's Ca^{2+}-sensitivity and was shown previously that mono- and bisphosphorylation gradually reduces the TnI-TnC affinity which has been replicated computationally in this work (Section 3.4.4).[133] The increased interactions between NcTnI and TnI C-terminal part of the model also corroborated with past experimental studies on the TnI phosphorylation effects where the two terminmus were shown to be in closer proximity with less dynamic movement upon PKA phosphorylation.

3.5.8 Mutation Induced Changes

The G159D mutation introduces a negatively charged residue (aspartic acid) at the C-terminal of TnC to replace the neutral residue (glycine) in the wild type. The direct impact of this single nucleotide missense mutation is the additional negative charge density in the mutation site and the extra interactions that form from Asp159 to the surrounding residues (Table 20). Most frequently, Asp159 formed an ionic interaction with TnC Arg83 (uP: 80%, SEP: 76%) which is located on helix D of the N-terminal domain of TnC. This residue was close to the TnC flexible linker and situated at the TnC domain interface. A strong and frequent interaction between Asp159 and Arg83 would bring the two domains of TnC closer and lock them into a close proximity. This provides a direct rationale for the observations detailed in Section 3.3.2 where a contraction and loss of dynamic range was seen in the regulatory domain with mutation. A low frequency ionic interaction was seen in G159D SEP between TnC Lys90 and TnC Asp159 (6%) but not in the unphosphorylated state. Phosphorylation of G159D increased TnC Asp159 ionic interaction frequency with two residues in the inhibitory peptide region of TnI. For both TnI Arg145 and TnI Arg147, the interaction counts with TnC Asp159 in the unphosphorylated state was 1%. The interaction frequencies were increased to 14% and 4% in G159D SEP for Arg145 and Arg157, respectively. These two interactions represent a modulation of the mutated residue on the inhibitory peptide binding pattern. The additional ionic interactions directly contribute to the increased inhibitory peptide binding energy detailed in Section 3.4.2.

		G159D uP	G159D SEP
TnC Arg83	TnC Asp159	80%	76%
TnC Lys90	TnC Asp159	0%	6%
TnC Asp159	TnI Arg145	1%	14%
TnC Asp159	TnI Arg147	1%	4%

Table 20: Ionic interactions formed from TnC Asp159 in the mutant.

3.5.9 TnC self-interactions

TnC is a highly organised subunit of troponin that consists of nine helices folded into a dumbbell shape, connected by a flexible linker. As such, many prominent ionic self-interactions link up regions of TnC to maintain its overall geometry (Table 21).

The ionic interaction between TnC Glu63 and TnC Arg83 was present in 84% of the simulations for WT uP and reduced to 77% in WT SEP. The interaction frequency was similar in the mutant with 82% in G159D uP and 79% in G159D SEP. This interaction connected the C-terminal of hexlix C, Glu63, to the C-terminal of helix D, Arg83, which was also located near the interdomain flexible linker. TnC Glu63 also formed a low percentage interaction with TnC Lys158 in the wild type (WT uP 4%, WT SEP 12%) and the frequency was greatly increased to 31% in G159D uP and 53% in G159D SEP. Notably, TnC Arg83 formed long-lasting ionic interactions with Asp159 in the mutant which would bring Lys158 into closer proximity with Glu63 hence increase the propensity of ionic bonding. The increased ionic interactions between TnC Glu63 and residues of the inhibitory peptide and the increased inhibitory peptide binding energy in the mutant were described before. This ionic interaction network forms part of the hypothesis in rationalising how the G159D mutation can lead to an elevated inhibitory peptide binding energy. When discussing the ionic interactions from the inhibitory peptide in the previous sections, the inhibitory peptide arginines were found to bridge some negatively charged TnC residues including TnC Glu63 and Asp88. Hence, it was not surprising to see bridging from TnC Arg83 to Glu63 and Asp88 as well. The ionic interaction between TnC Arg83 and Asp88 existed for 51% of the simulations in WT uP and was reduced to 37% in the unphosphorylated mutant. In both phosphorylated systems, the interaction frequency was 32%. Lastly, an ionic interaction between TnC Asp3 and Arg83 connected the two ends of the N-terminal domain of TnC, forming the globular shape. This interaction accounted for 27% and 37% in WT uP and SEP, respectively, and increased to around 60% in both states of the mutant. This increased interaction frequency in the mutant is likely linked to increased structure rigidity because of the extra ionic interactions from TnC Asp159 in the mutant.

Additionally in the NcTnC side, Lys6 and Glu66 connected the beginning of NcTnC to the catalytic calcium loop region. This interaction accounted for more than 75% of the simula-

tion trajectories in all systems. The ionic interaction between Lys21 and Glu32 connected the middle of helix A to the middle of EF-hand I loop region. This interaction had similar frequencies in all systems, around 50%. On the CcTnC side, Lys92 and Glu161 joined the two ends of the C-terminal lobe together. This interaction accounted for 40% and 31% of the simulation in WT uP and WT SEP, and a slightly increased 45% and 38% in G159D uP and G159D SEP, respectively. Lys118 and Glu130 connected helix E and F which in turn joins the EF-hand III and IV. Both Asp109 and Asp113, from the ion binding loop of EF-hand III, formed ionic interaction (between 23-33% for all systems) with Arg147 at the calcium binding site of EF-hand IV. This provided furthur structural rigidity for the two structural calcium binding sites.

		WT uP	WT SEP	G159D uP	G159D SEP
TnC Glu63	TnC Arg83	84%	77%	82%	79%
TnC Glu63	TnC Lys158	4%	12%	31%	53%
TnC Arg83	TnC Asp88	51%	32%	37%	32%
TnC Asp3	TnC Arg83	27%	37%	61%	60%
TnC Lys6	TnC Glu66	78%	88%	77%	76%
TnC Lys21	TnC Glu32	48%	55%	52%	50%
TnC Lys92	TnC Glu161	40%	31%	45%	38%
TnC Lys118	TnC Glu130	36%	40%	42%	40%
TnC Asp109	TnC Arg147	33%	24%	28%	29%
TnC Asp113	TnC Arg147	32%	23%	27%	29%

Table 21: Ionic interactions formed within TnC.

3.5.10 IT Arm Interface

The IT arm domain of troponin consists of the intertwined helices of TnI and TnT. The ionic interactions present here serve a structural role in maintaining the rigid structure of the IT arm. Most of the interactions here are stable and no significant differences were observed across systems (Table S1).

3.6 Granger Causality Analysis

Using residue interaction profiling, the molecular interactions presented in the MD simulations were quantified and the significant changes in interactions were identified in the previous section. This allowed us to hypothesise how the changes in non-covalent bonding could influence troponin's dynamics that leads to the change in Ca^{2+} sensitivity and causes the dilated cardiomyopathy phenotype in the case of the G159D mutation. These insights were helpful in understanding the structural changes when looking at the cross-sectional data yet it does not provide reasoning in *how* those interactions were perturbed by the additional ionic interactions introduced by phosphorylation and mutation. I have developed a novel residue interaction analysis method based on Granger causality to unveil the causal effects between connected molecular interactions. The mathematical background on Granger causality is detailed in Section 2.8. In essence, Granger causality analysis aims to reveal if one molecular interaction is statistically linked to an adjacent molecular interaction, in which case a putative causal relationship can be asserted. This procedure provides a direct pathway for linking the allosteric effects from the perturbation site to the effect site and gives us further insights into troponin's regulatory mechanism.

3.6.1 Phosphorylation Induced Perturbations in the Wild Type

Figure 23 illustrates the propagation of the newly formed ionic interactions from the phosphoserines. The blue boxes represent residues and the edges between boxes represent interactions. As shown in the legend, hydrogen bond is coloured in blue; ionic interaction is coloured in orange; and in the case that the two boxes are connected by both ionic interaction and hydrogen bond, the edge is coloured in red. An edge between residues means that this interaction is causative towards the interactions between the connecting residue and residues in the next hop, from left to right. The direct effect of phosphorylation is the introduction of 4 negative charges which changes the interaction pattern between NcTnI and NcTnC. It is notable that only one of the phosphoserines is present in this graph and TnI Sep22 is absent from the interaction network. This means that the interactions formed from TnI Sep22 were stable and did not influence the formation or dissociation of other interactions. The

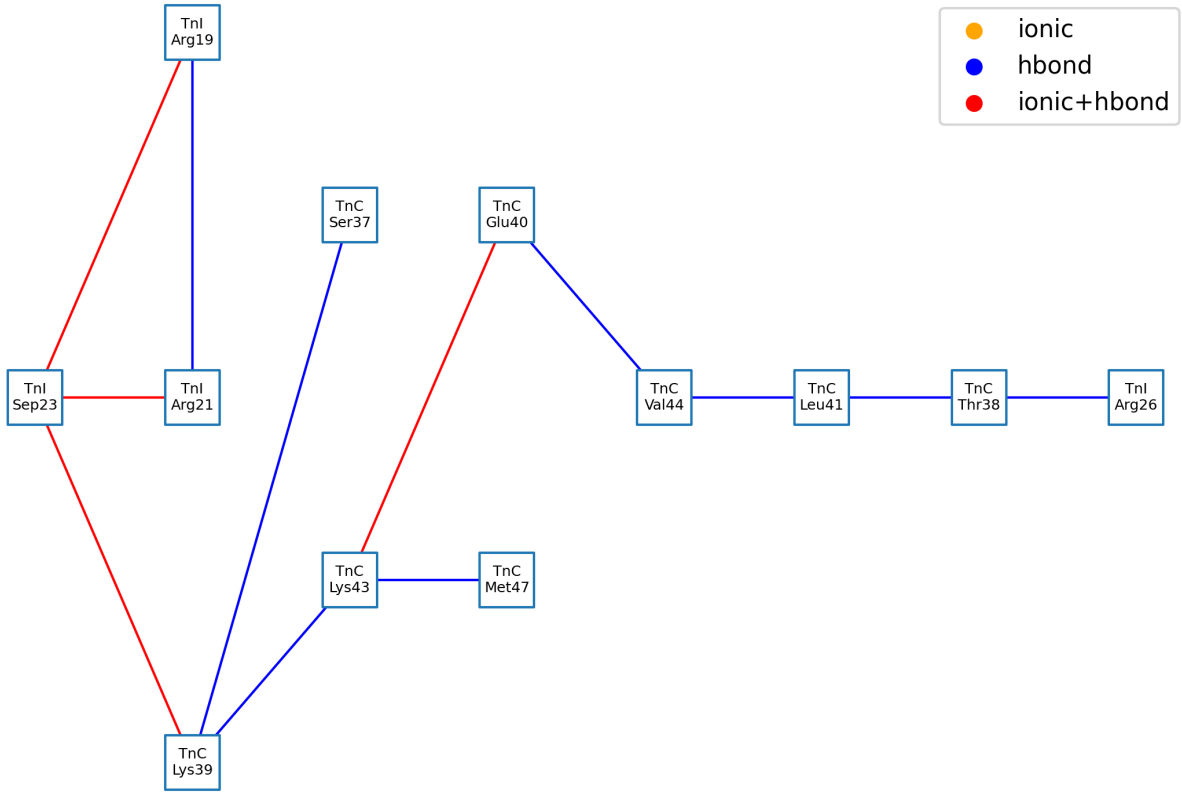


Figure 23: Granger causality inferred network for direct ionic and hydrogen bond interactions between residues for WT SEP. This network shows the perturbation effects of additional interactions introduced by the phosphoserines.

graph shows that TnI Sep23 formed both ionic interaction and hydrogen bonding with both TnI Arg19 and TnI Arg21. Both of these interactions then have had an effect on the hydrogen bond between TnI Arg19 and Arg21. This influence is expected since either interactions with TnI Sep23 will have an effect on the relative position of these two residues. TnI Sep23 also formed ionic and hydrogen bonding with TnC Lys39 and this interaction was shown to affect hydrogen bonding between TnC Lys39/TnC Ser37 and TnC Lys39/TnC Lys43. The effect is passed down to other interactions on helix B of TnC, mainly through the backbone hydrogen bonds that maintain the helical structure. Eventually, the perturbation was passed on to the hydrogen bond between TnC Thr38 and TnI Arg26.

The Granger causality analysis on WT SEP suggests that out of the two phosphoserines, TnI Sep22 predominently partakes in TnI self-interaction with surrounding positively charged residues in order to reduce the binding interactions between NcTnI and NcTnC. TnI Sep23 not only forms ionic interactions with neighbouring arginines but more importantly mod-

ulates the backbone hydrogen bonds on TnC helix B via ionic interaction with TnC Lys39. This pathway discovered by Granger causality analysis gave us a direct mechanism in how phosphorylation modulates the interhelical A-B angle which in turn would affect troponin's Ca^{2+}-sensitivity.

3.6.2 Mutation Induced Perturbations

Broadly, the effects of the G159D mutation spread more widely and influenced more molecular interactions than the phosphorylation effect in the wild type (Figure 24). This is likely because the mutated residue was deeply buried unlike the phosphorylation sites. The most direct influenced regions were the residues near the TnC domain interface. The first hop contained the direct interactions from TnC Asp159. These were the hydrogen bonds with the N-terminal of TnC and TnC Ser89 which was located on the flexible linker. The ionic interaction with TnC Arg83 was strong and long lasting which was the main causer for the reduced TnC domain distance and increased TnC rigidity in the mutant. The second hop saw the residue interactions spread to surrounding residues in the same proximity of NcTnC helix N and the region near the flexible linker. The only residue outside these areas was TnC Glu63 which was previously shown to participate in the TnC-inhibitory peptide interface. As expected, the third hop involved residues in the inhibitory peptide region (TnI Arg145 and Arg147). From the mutation site, the TnC Asp159/TnC Arg83 interaction was shown to influence the TnC Arg83/TnC Glu63 and the TnC Arg83/TnC Asp88 interactions which in turn influenced the TnC Glu63/TnI Arg145 and TnC Asp88/TnI Arg147 interactions. This represents a pathway that connects from the mutation site to the inhibitory peptide region hence provides a molecular explaination for the modulation of the inhibitory peptide binding due to the G159D mutation.

The third hop involved residues in the C-terminal end of TnC, such as Lys158 which served as a hub to pass down perturbations to other CcTnC residues in the next hop. The third hop also saw interactions further spread in the TnC helix C (Glu59, Met60, Ile61). The next few hops mainly contained two noteworthy observations. Firstly, the influence of the perturbation spreaded to the CcTnC region and the TnI region that CcTnC was anchored to. This

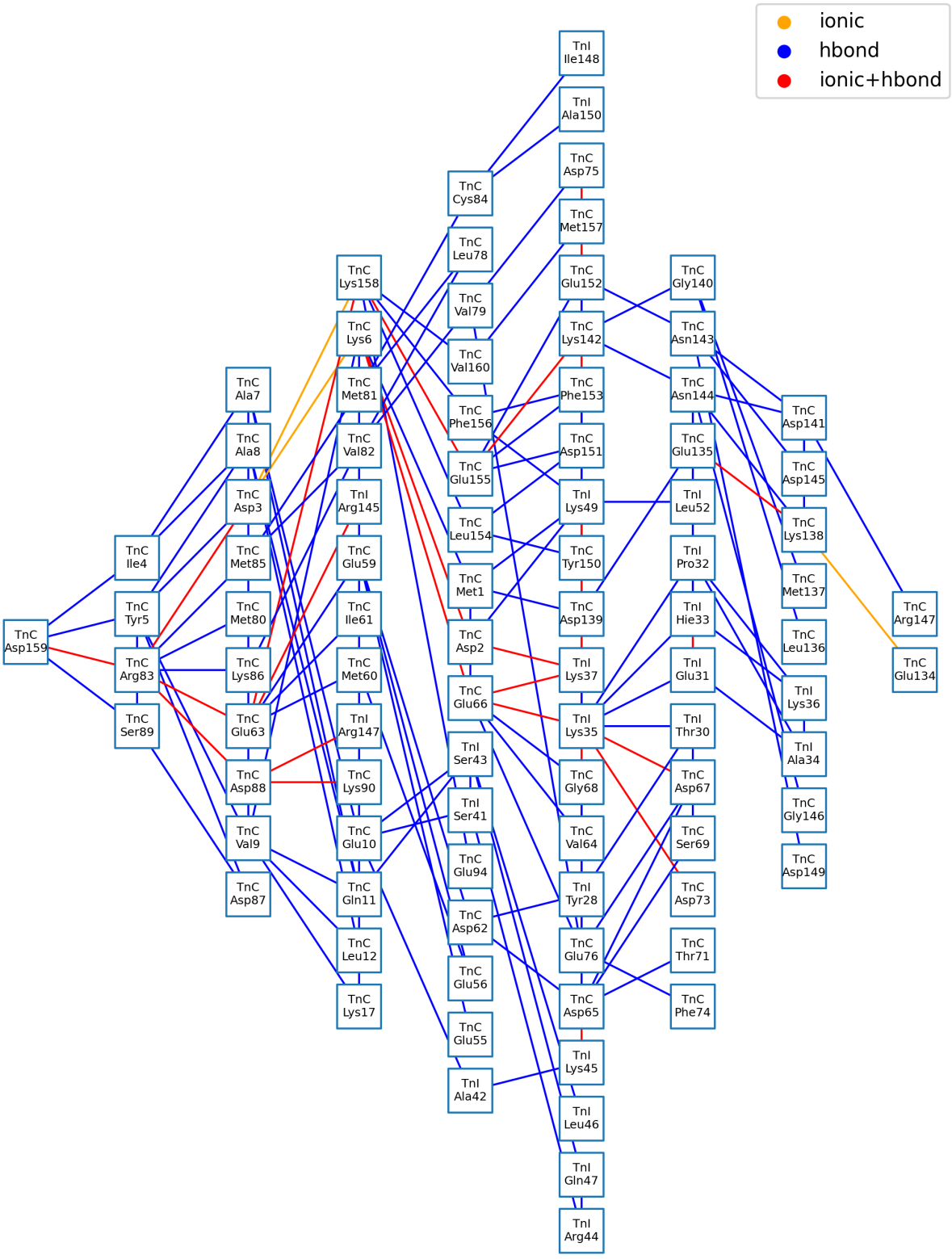


Figure 24: Granger causality inferred network for direct ionic and hydrogen bond interactions between residues for G159D uP. This network shows the perturbation effects of additional interactions introduced by the mutated residue.

pathway explains how mutation changed the CcTnC/TnI anchor's interaction pattern and influenced their binding energy. It was interesting that the changes in CcTnC/TnI anchoring due to mutation seemed to take an indirect pathway. The effect of the mutation was not directly passed on to the TnI anchoring region given the proximity but instead propagated through helix N of TnC and residues like Glu10 and Gln11. Secondly, the mutation effect was propagated further down EF-hand II and affected some residues directly in the catalytic calcium binding site. In the TnC helix N region, the ionic interaction between Asp3 and Lys6 influenced the interaction between Lys6 and TnC Glu66 which in turn perturbed the interaction between TnC Glu66 and TnI Lys35. TnI Lys35 was previous identified as a TnI residue that directly interacted with TnC residues that formed the ion coordination site and here it affected residues TnC Asp67 and Asp73. This pathway represents a direct effect on the catalytic calcium's coordination site because of the G159D mutation.

Overall, the Granger causality analysis on the G159D uP system was fruitful in shining lights on how mutation affected troponin's dynamics. The analysis was able to delineate three pathways that connected the mutation site to the inhibitory peptide region, CcTnC/TnI anchor region, and EF-hand II calcium binding site. These causal relationships provided plausible explanations in how the G159D mutation changes the structural and energetics features which eventually leads to the dilated cardiomyopathy phenotype.

3.6.3 Molecular Interactions in G159D SEP

In the case of G159D SEP, the causal graph is more extensive and largely encapsulated the mutation effects discovered in G159D uP but presented key differences in the phosphorylation effects(Figure 25). For the mutation effects introduced by TnC Asp159, a similar pathway to G159D uP was seen that connects the mutation site to the inhibitory peptide region via TnC Arg83, TnC Glu63, and TnC Asp88. Similarly, the effect of the newly formed interactions with the mutated residue was passed on to the CcTnC/TnI anchoring region via the N-terminal end of TnC. Some interactions involved TnC helix C were influenced by the mutation and passed on to the catalytic calcium coordination residues via TnI Lys35, similar to G159D uP. Interestingly, TnC Asp33, the cardiac-specific residue that previously noted as forming multiple ionic interactions with NcTnI arginines, did not appear in the causal graph. In the case

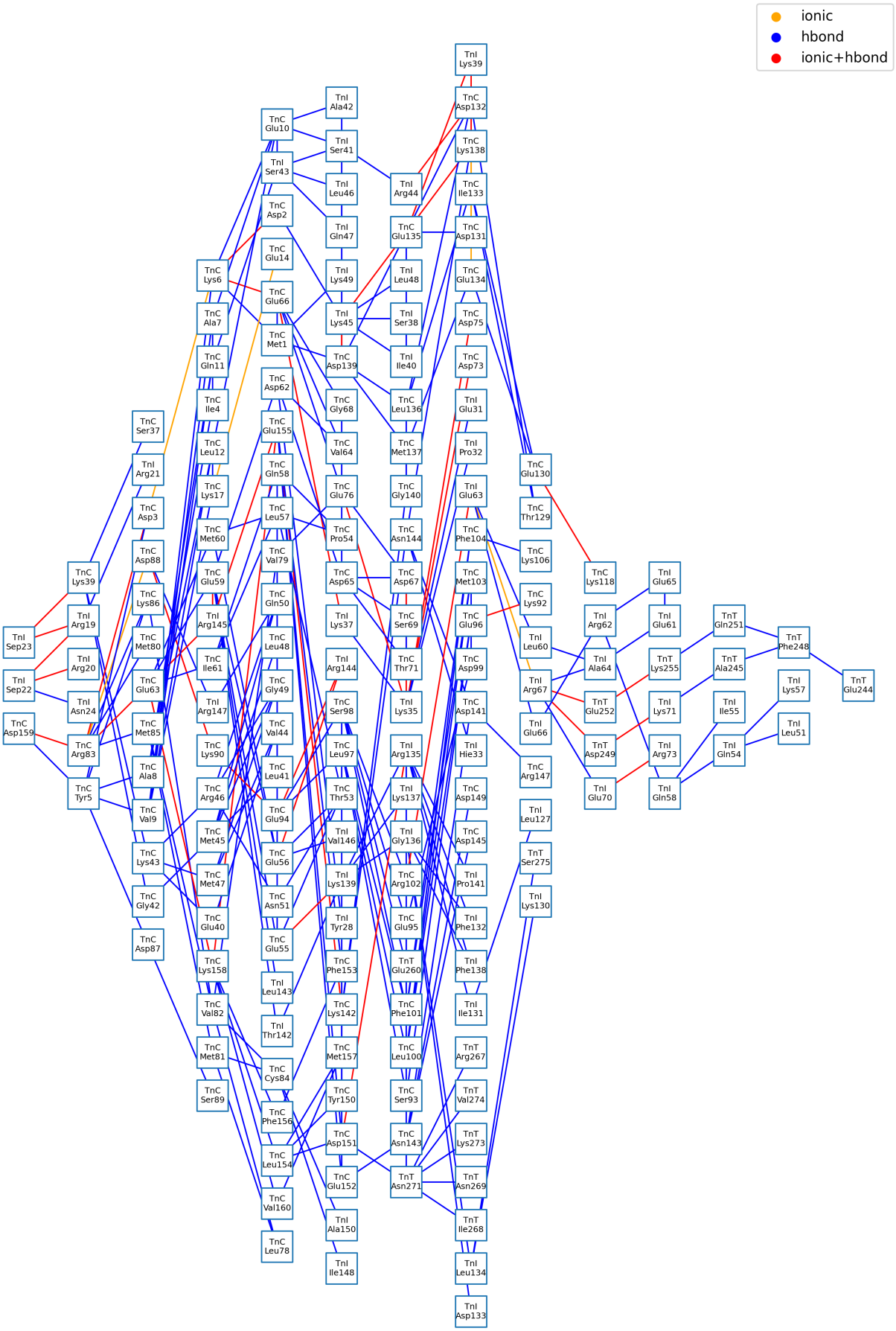


Figure 25: Granger causality inferred network for direct ionic and hydrogen bond interactions between residues for G159D SEP. This network shows the perturbation effects of additional interactions introduced by TnC Asp159 and TnI Sep22/23.

of G159D SEP, both phosphoserines appear in the network and connected to neighbouring residues however their effect was not propogated further. Akin to WT SEP, the main effect of phosphorylation was conducted via TnC Lys39 to the rest of TnC helix B. This represents a direct modulation on the A-B interhelical angle dynamics from the phosphorylation sites.

In summary, G159D SEP combines the effects of phosphorylation in WT SEP and mutation perturbations in G159D uP. Both phosphorylation and mutation affect NcTnC dynamics and together they may present a competitive effect. The interactions between TnI and TnC catalytic calcium binding site were perturbed by the mutation site which could be more pronounced than the phosphorylation effect on EF-hand I. This would then prohibit the change in calcium sensitivity upon phosphorylation in the mutant.

3.6.4 Limitations of the Granger Causality Analysis

One of the major limitations of the Granger causality method proposed here is that only directly connected interactions were considered. In reality, it is foreseeable that when one interaction forms it also brings adjacent residues into closer proximity hence likely to form interactions. Another caveat is that when an interaction between residues is stable, the time-series sees less variations and is unlikely to be deemed statistically significant in causing another interaction. This means that the Granger causality analysis should be used in conjunction with residue interaction profiling as some major changes in interaction percentage could be omitted from the results due to their highly stable nature.

Overall, I believe the Granger causality method represents a new direction in analysing MD trajectories where the causal effects of intricate molecular interactions could be inferred with statistical rigor. The results presented in this section largely agreed with our own intuition and past literatures yet provided a more detailed picture of the potential pathways. This methodology warrants further development to include molecular interactions that do not share a common residue and better visualisations to aid the researchers in interpreting the results.

3.7 Discussion

Troponin as the heart contraction regulator possesses a delicate phosphorylation-dependent mechanism. MD was used in this chapter to characterise the changes in troponin's structure and dynamics introduced by phosphorylation as well as the TNNC1 G159D mutation. Several key findings regarding troponin's structural features, energetics, and dynamics are discussed here.

3.7.1 Phosphorylation and Mutation Lead to Rigidification

Working as a molecular switch, the dynamical movement and structural rearrangement upon perturbation to satisfy the functional requirements under different states have fascinated biochemists over the years. Multiple NMR studies have shed light on the influence of TnI Ser22/23 phosphorylation and many concluded that phosphorylation reduces the flexibility of parts of troponin. This observation was reproduced in the MD simulations presented here. Firstly manifested as slightly higher RMSF values in wild-type phosphorylated TnC compared to the uP state. Many of the metrics discussed in this chapter, including interdomain hinge angle, TnC domain distance, interhelical angle, experienced a reduction in their standard deviation upon phosphorylation in the wild type. In the same vein, mutation also causes a reduction in the range of movement of troponin when comparing G159D uP to WT uP, and often times the phosphorylation of G159D uP led to a less rigid structure which is the opposite phosphorylation effect as observed in the wild type.

3.7.2 Mutation Increases Inhibitory Peptide Binding Energy

Measurements of the inhibitory peptide binding energy to the rest of troponin with the MMPBSA protocol have revealed that the G159D mutation tightens the binding in both phosphorylation states compared to the wild type. This observation can be explained by residue interaction profiling combined with causality analysis. Direct interactions from TnC Asp159 to TnI Arg145 and Arg147 were observed. These ionic interactions do not exist in the wild type and likely contribute to the increased binding energy in the mutant. Residue interaction profiling

showed increased ionic interactions between TnI Arg145 and TnC Glu59, Glu63, and Glu155 in the mutant. The Granger causality analysis was able to delineate pathways that led to the modulation of the interactions involve TnI Arg145 and Arg147. In the mutant, the additonal interaction between TnC Asp159/TnC Arg83 perturbed the TnC Arg83/TnC Glu63 and TnC Arg83/TnC Asp88 interactions which in turn influenced the TnC Glu63/TnI Arg145 and the TnC Asp88/TnI Arg147 interactions. These ionic interactions that were perturbed by the mutated residues can potentially explain the observed mutation effect in the inhibitory peptide binding energy measurements. The modulations on the inhibitory peptide binding energy parallels the results from Cool's work where the phosphorylation of an inhibitory peptide residue was found to increase the effective concentration of switch peptide to the hydrophobic patch.[21] In the same vein, the changes in inhibitory peptide binding interaction could directly alter the switch peptide and hydrophobic patch's binding dynamics.

3.7.3 Mutation and Phosphorylation Change TnC/TnI Anchoring Stability

Measurements of the TnI_{34-71} peptide binding energy gave an indication of the TnC/TnI anchoring stability. The results show that the G159D mutation reduced the anchoring stability in both phosphorylation states and phosphorylation increased the stability in both the mutant and the wild type. Crucially, this result agrees with NMR titration measured changes in TnI_{34-71} peptide binding energy with the G159D mutation. The simulation was able to reproduce the changes in the anchoring stability that was observed experimentally. RMSF metric showed the increased flexibility of TnI residues 30-40 in the mutant when compared to the wild type. Residue interaction profiling also demonstrated the reduced interaction percentage between CcTnC and TnI_{34-71} in the mutant, especially the ionic interaction between TnC Asp139 and TnI Lys45. Granger causality graph illustrated multiple pathways that propagated the mutation effect to the TnI_{34-71} region via helix N of TnC. The reduction in stability caused by mutation means that the relative position of TnC to TnI would be more variable and the switch peptide is less available to bind to the hydrophobic patch due to proximity effect which in turn causes a decrease in activation kinetics.

3.7.4 A-B Angle Dynamics are Linked to Interactions Between TnI and TnC

One crucial metric reported in this work and past literature is the interhelical angle between helix A and B of TnC. The simulation data showed that the A-B angle reduced upon phosphorylation in the wild type and increased with phosphorylation in the mutant. This observation coincides with the phosphorylation-dependent changes of interactions between TnI and TnC in both systems. Namely, the TnC Glu19/TnI Arg161 and the TnC Lys43/TnI Glu164 ionic interactions which involved residues from the hydrophobic patch and the switch peptide reduced in percentage upon phosphorylation in the wild type but the opposite was observed in the mutant. The same kind of opposite trends between WT and G159D was also seen in interactions between NcTnI and NcTnC, such as TnC Asp25/TnI Arg12, TnC Glu32/TnI Arg12, and TnC Glu40/TnI Arg21. Furthermore, in the Granger causality analyses, both the mutation and the phosphorylation were shown to affect interactions that involve residues within helix A and B of TnC. These interactions likely contribute to the modulation of the A-B angle as well as the switch peptide binding in different systems.

3.7.5 Phosphorylation and Mutation Cause Reposition of NcTnI

Experiments have shown that phosphorylation of both Ser22/23 on NcTnI leads to reduced interactions between NcTnI and NcTnC in the wild type.[23] The MMPBSA measurement (Section 3.4.4) has shown a drastic reduction in NcTnI binding energy of $30\,\mathrm{kcal\,mol^{-1}}$ upon phosphorylation in the wild-type troponin. The effect of mutation presented in this dataset was a reduction in NcTnI binding energy for the unphosphorylated state and the large differential between the two phosphorylation states was abolished in the G159D mutant. Effect size quantifies the WT changes as large (-0.99) and the G159D system possesses a small and opposite (0.18) effect due to phosphorylation.

The phosphorylation-dependent changes in NcTnI binding energy can be explained as the negative charges introduced by the phosphoserines neutralised the surrounding positively charged arginines hence limit the intersubunit interactions that presented in WT uP. Utilising novel residue interaction profiling tools, the residue interaction percentage data were able to

provide quantified evidence in describing the changes in molecular interactions. A systematic reduction in NcTnI ionic interactions with NcTnC was observed across TnI Arg19/20/21/26 (Section 3.5.5) with phosphorylation in the wild type. Similar trend followed in the mutant yet subtle differences was shown especially with interaction pairs not involving TnI Asp33. TnI Arg12 experienced increased ionic interactions in the wild type with phosphorylation while the opposite was observed in the mutant. And the absolute interaction percentages for Arg12 in G159D uP more closely resembled the WT SEP state. The cumulative interaction percentages between TnC Asp33 and NcTnI were similar in both WT uP and G159D uP yet the G159D SEP was 10% more frequent than WT SEP. These factors combined can provide a molecular rationale for the reduction in binding energy in the unphosphorylated state and the increase in binding energy in the phosphorylated state for the mutant.

3.8 Conclusion

In this chapter, molecular dynamics was used to simulate the core of cardiac troponin and studied the effects of phosphorylation and mutation. Based on structural and energetics measurements, the significant changes in troponin's dynamics were identified. Some of the changes in binding energy were inline with experimentally determined changes and other key observations could be corroborated with past experimental and computational studies. Through residue interaction profiling, crucial changes in molecular interactions were quantified and compared between systems to provide rational explanations in perturbation induced differences. Lastly, a novel Granger causality based analysis method that infers the causal relationship between molecular interactions was developed. This method has provided direct pathways into explaining the allosteric effects of phosphorylation and mutations.

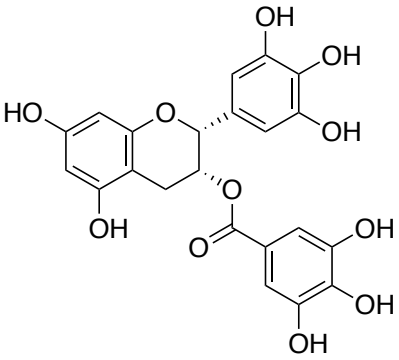
Overall, the results discovered in this section represent key metrics that describe the changes introduced by phosphorylation and mutation. These will serve as a fundation for studying the effects of small molecules by providing a set of comparable metrics.

4 Study of EGCG, SilybinA, SilybinB, and Resveratrol

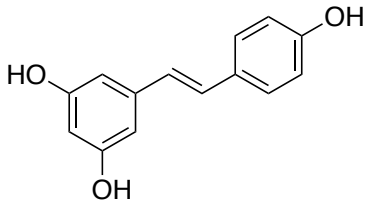
The widespread adoption of genetic testing in the clinical setting has afforded many phenotype-genotype correlation studies regarding cardiomyopathies.[134] Novel algorithms were also researched to predict the pathogenicity of variants of unknown significance in cardiac thin filament.[77] However, despite the prevalence of cardiomyopathies, there is currently limited drugs on the market that directly aim at treating sarcomeric cardiomyopathies by targeting the root cause of the disease.

Based on the apo simulations described in Section 3, I have observed key changes that correlated with the mutation and phosphorylation effects in cardiac troponin. These can serve as comparable metrics for quantifying the effects of introducing small molecules in troponin simulations presented in this chapter. The first group of natural products were selected based on their historical interest, structural features, and existing experimentally characterised data. The four molecules are epigallocatechin-3-gallate (EGCG), *trans*-resveratrol (RVL), silybin A, and silybin B (Figure 26).

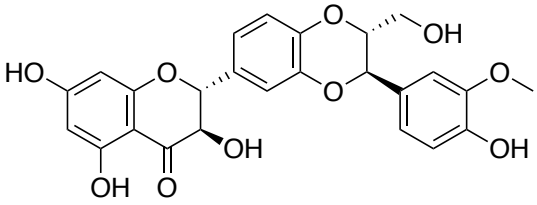
EGCG is a biologically active polyphenol found in green tea that possesses potent antiproliferative effect against human colorectal cancer cells.[135] It was one of the first natural products studied for its effect on cardiac muscle functions as catechins are known for their protective properties against cardiovascular diseases.[58] It was first shown to be a Ca^{2+} desensitiser by direct binding to TnC.[136] In the wild-type thin filament, EGCG decreased calcium sensitivity of phosphorylated and unphosphorylated states equally hence retaining the coupling property. Papadaki *et al.* demonstrated using IVMA that EGCG can restore the coupled relationship between Ca^{2+} sensitivity and TnI phosphorylation in a variety of thin filaments with both DCM and HCM phenotypes that are uncoupled.[57] The ability to recouple a range of mutations that present in different parts of the thin filament, including TnC, TnI, TnT, Actin, tropomyosin, and myosin binding protein C, suggests that EGCG's mode of action is likely to be universal and unrelated to the specific mutation. EGCG is a rather sizable molecule with four highly substituted aromatic rings. Sheehan *et al.* experimented with compounds structurally resembling EGCG but simpler, such as epicatechin (EC), epigallocatechin (EGC), gallocatechin (GC), and epicatechin gallate (ECG). It was found that all but ECG can recouple the TPM1



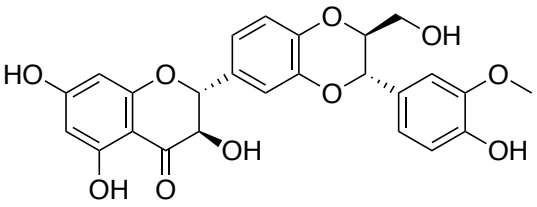
epigallocatechin-3-gallate



trans-resveratrol



silybin A



silybin B

Figure 26: Chemical structure of EGCG, RVL, silybin A, and silybin B.

E180G HCM mutation.[67] This study also expanded the compound scope to other putative Hsp90 inhibitors including silybin, its derivatives and stereoisomers.

RVL is a polyphenol with antioxidative properties that is abundant in wine. Moderate consumption of red wine has displayed cardioprotective effect based on epidemiological studies.[137] Experimental evidence suggests that RVL directly modulates cardiac myocyte shortening and was found to increase the myofilament Ca^{2+}-sensitivity of guinea pig myocytes.[138] RVL is mostly flat and does not possess the 3-dimensional structural complexity as EGCG yet was shown to be a prominent recoupler against the TPM1 E180G HCM mutation. Hence, it is of interest to understand its binding interaction and potential mode of action.

Silybin A and B are two major active flavonolignans found in milk thistle fruit extract.[139,140] Silybin A is a pure desensitiser and silybin B was shown to be a pure recoupler that decreased the mutant's Ca^{2+}-sensitivity in the phosphorylated state and left the unphosphorylated state unaffected. Differing only in stereochemistry, these two diastereomers pose interesting struc-

ture activity relationships to understand the modulation of troponin's calcium sensitivity by small molecules.

Past structural studies involved troponin in complex with ligands have identified several pharmacophores that contributed to the ligands binding motifs.[64] Multiple NMR studies have determined the structures of small molecule in complex with either the regulatory domain or the structural domain of TnC. Pineda-sanabria *et al.* showed that RVL binds to the hydrophobic pocket of CcTnC with a minor conformational change in the absence of TnI.[26] EGCG was also shown to interact with CcTnC hydrophobic pocket in the presence of TnI_{34-71}.[63] It was postulated that EGCG may compete with TnI_{34-71} binding and weaken the anchoring of TnC to the thin filament. Experimental evidence and the modelling results presented in Section 3 showed that the G159D mutation which is located at CcTnC weakens the affinity of TnI_{34-71} for the troponin complex. Robertson *et al.* suggested that EGCG may work in a similar manner to protect the heart from the development of hypertrophy.[63] Although this does not explain the apparent recoupling ability of EGCG towards DCM mutations. Levosimendan is a positive inotrope and together with its analogue were shown to bind to the regulatory domain of TnC.[64] Computational studies and virtual screenings also used the N-terminal domain of TnC in isolation to discover novel calcium sensitisers to some success.[83,141] Lacking a native binding site, it is generally accepted that active compounds would interact with TnC directly to modulate troponin's calcium sensitivity. The question, however, remains that if the EF-hands' hydrophobic pockets are accessible for the small molecules in a more intact system. For instance, NMR study of EGCG interaction with CcTnC has shown a drastically changed binding interaction in the presence of TnI_{34-71}. Therefore, an unbiased approach was taken towards the simulations of troponin in the presence of small molecules where the ligands were placed near the regulatory head of troponin but not directly inside the postulated binding sites.

In this chapter, MD simulations of troponin and its TNNC1 G159D mutant in both unphosphorylated (uP) state and phosphorylated and doubly-negatively-charged (SEP) state in the presence of individual small molecules, including EGCG, RVL, silybin A, and silybin B are presented. Similar to the apo studies, five repeats were conducted for each system and each production run lasted $1.5\,\mu s$. Structures were saved at $0.2\,ns$ intervals for 7500 frames per tra-

jectory and 37500 total frames per system. The ligand binding regions are described and the key metrics are measured to compare with the apo simulations.

4.1 Global Structure Changes from Principal Component Analysis

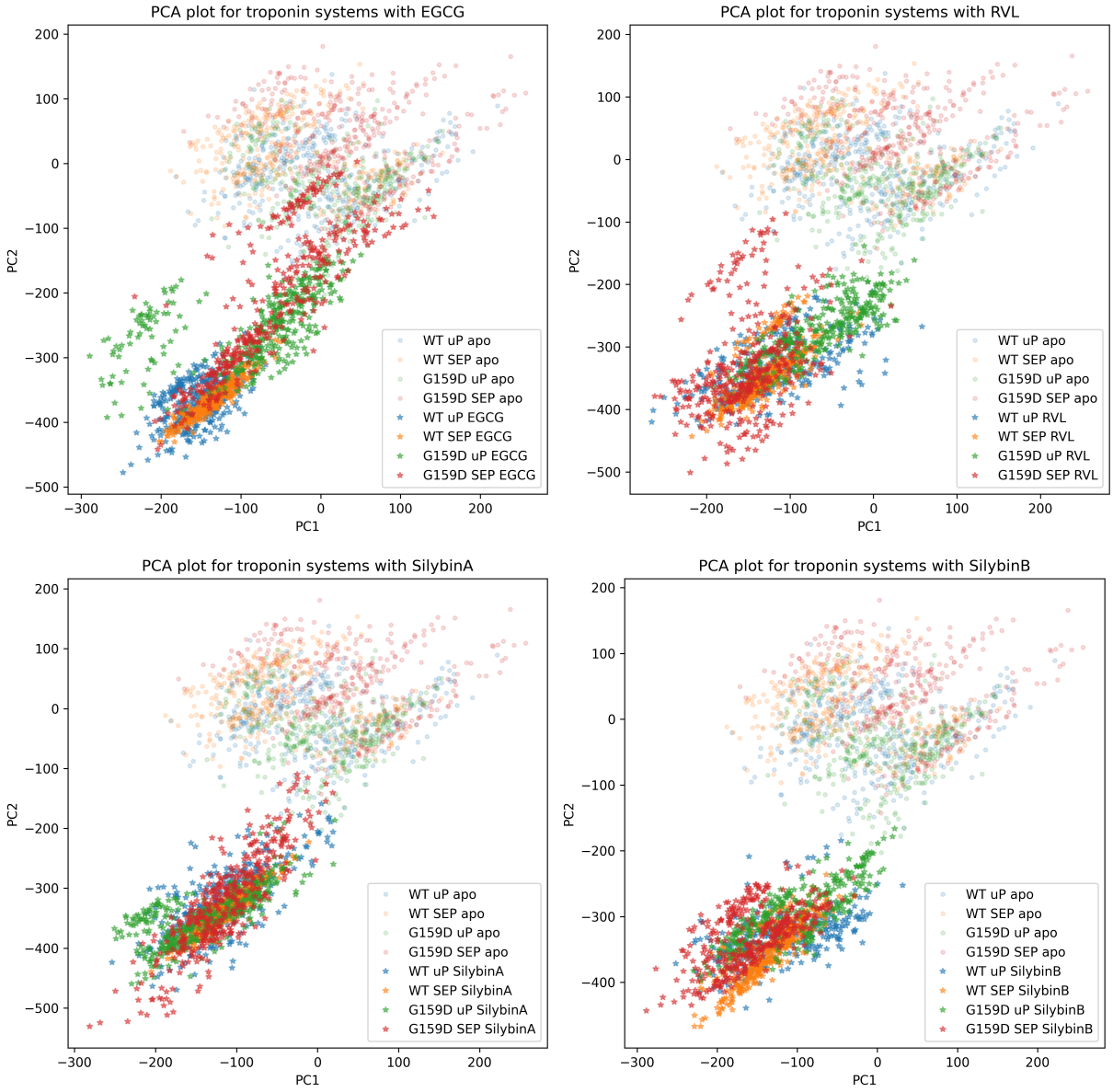


Figure 27: Scatter plot of troponin conformations in the apo systems' first two principal components. The faded circles represent conformations from the apo system. The coloured stars represent the ligand systems.

Figure 27 illustrates the PCA projections of the simulation trajectories containing EGCG, RVL, silybin A, and silybin B onto the same principal component space as the apo study. Broadly, all ligand systems are located towards the lower left corner of the PC1/PC2 plot in comparison to the apo systems. This shows the effects of small molecule perturbations have shifted troponin's global conformations into a different space in the first two principal components. It worth bear in mind that the first two principal components only represent 30% of the captured structural variance. In all the four ligand systems presented here, there is a great overlap among wild type and mutant and their phosphorylation states. EGCG showed a bigger overlap with the apo systems, G159D SEP in particular, than the rest of the ligand systems. Silybin A and B showed a more overlapping conformational space within the four systems. In all ligand systems, G159D SEP state showed a bigger deviation from the general ligand space and more leaning towards the apo space.

4.2 Changes in Regional Structural Features

4.2.1 Interhelical A-B Angle

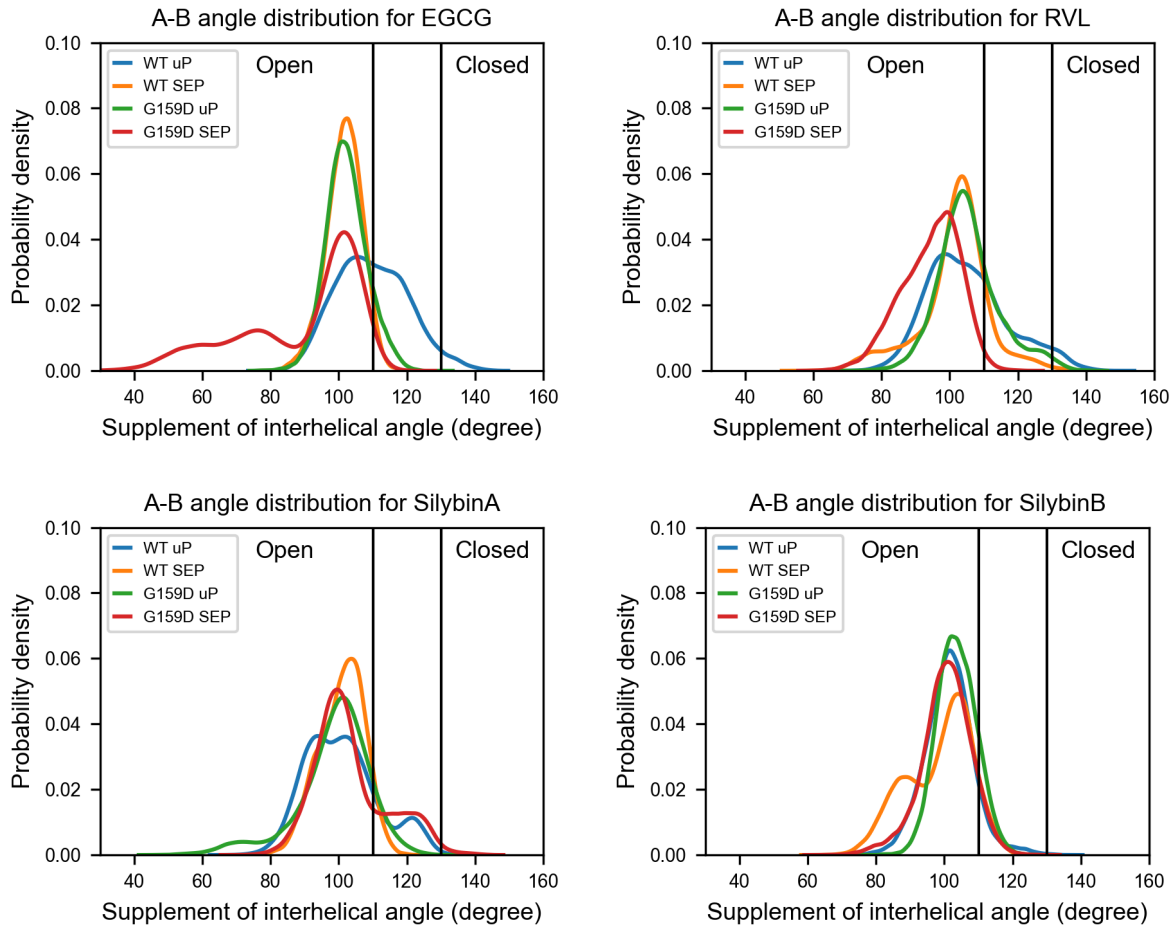


Figure 28: The interhelical angle distributions for EGCG, RVL, silybin A, and silybin B systems.

The percentage of density above 110° quantifies the proportion of the simulations that the interhelical A-B angle was beyond the open state. This metric showed that phosphorylation moved the wild type's hydrophobic patch into the more open conformation while the mutant was shifted to the higher angle direction and a more closed state upon phosphorylation (Figure 19). This was deemed as a significant effect of mutation on troponin's hydrophobic patch dynamics which would influence the switch peptide binding event thus affect troponin's calcium sensitivity.

	Peak	Mean	Std. Dev.	>110°	FWHM
WT uP apo	96.3°	101.8°	9.2°	18%	17.1°
WT uP EGCG	105.6°	109.8°	10.6°	47%	28.4°
WT uP RVL	98.7°	105.2°	12.0°	31%	24.3°
WT uP SilybinA	93.8°	100.8°	10.8°	18%	23.8°
WT uP SilybinB	101.6°	101.3°	7.2°	9%	14.3°
WT SEP apo	95.0°	96.2°	5.7°	1%	12.2°
WT SEP EGCG	102.4°	101.6°	5.4°	5%	12.1°
WT SEP RVL	103.6°	101.4°	10.4°	14%	12.6°
WT SEP SilybinA	103.7°	100.8°	6.7°	7%	16.8°
WT SEP SilybinB	104.0°	98.4°	9.7°	9%	13.7°
G159D uP apo	95.7°	97.1°	5.5°	1%	12.5°
G159D uP EGCG	101.1°	102.2°	6.1°	11%	12.8°
G159D uP RVL	103.8°	105.5°	9.2°	25%	14.8°
G159D uP SilybinA	101.2°	98.6°	11.0°	11%	17.4°
G159D uP SilybinB	102.2°	103.6°	5.8°	14%	14.2°
G159D SEP apo	95.4°	99.0°	6.7°	7%	13.6°
G159D SEP EGCG	101.6°	89.0°	18.1°	3%	14.1°
G159D SEP RVL	99.2°	94.3°	8.7°	1%	20.4°
G159D SEP SilybinA	99.6°	103.4°	10.9°	24%	14.1°
G159D SEP SilybinB	101.0°	100.5°	7.4°	9%	15.0°

Table 22: The interhelical angle metrics for the systems with EGCG, RVL, silybin A, and silybin B.

Figure 28 illustrates the distributions of the A-B angle in different systems as KDE plots. In WT uP, EGCG and RVL both caused a shift towards the more closed direction compared to the apo state (apo: 18%; EGCG: 47%; RVL: 31%) while silybin A did not change the >110° percentage and silybin B reduced the percentage to 9% (Table 22). Upon phosphorylation, all the ligand systems except silybin B increased the percentage to be above the apo level while maintained the differences between uP and SEP states to be a shift towards the open state. In the case of silybin B, the SEP percentage was elevated to be the same level as the uP state.

In G159D uP, all ligands caused shifts of A-B angle density towards to the higher angle direction which drastically increased the percentage to be above the apo level. In the phosphorylated mutant, EGCG and RVL brought the percentage to be below the mutant apo level. Silybin B caused the >110° percentage to go slightly over the apo level in G159D SEP state (apo: 7%; silybin B: 9%). Silybin A introduced a significant increase in this metric to 24%. The overall effect of EGCG, RVL, and silybin B to the phosphorylation induced changes in the mutant was a shift of the A-B angle towards the open direction. This fixed the abnormal behaviour of the mutant to be consistent with the wild-type apo systems. The desensitiser Silybin A caused the A-B angle to be more closed upon phosphorylation in the mutant and did not reverse the change in A-B angles to the same direction as WT apo.

In the apo study, it was noted that the sampled range of A-B angles as measured by standard deviation and FWHM showed a great reduction upon phosphorylation in the wild type and the opposite was observed in the mutant. In this set of data, almost all simulations with ligands showed an increased standard deviation compared to the respective apo systems. In the wild type, the reduction in range of sampled angles was largely maintained except for silybin B systems. In the mutant, RVL showed a reduced standard deviation in the SEP state, silybin A displayed the same standard deviation, and EGCG and silybin B exihibited an increased range of sampled angles.

Overall, the phosphorylation induced shift in A-B angle in the mutant was restored to the same directionn as seen in WT apo systems when recouplers EGCG, RVL, or silybin B were introduced to the simulations. Silybin A could not restore this mutational anomaly. This finding is consistent the experimental observation that EGCG, RVL, and silybin B could recouple

the mutant while silybin A could not.

4.2.2 Interdomain Hinge Angle

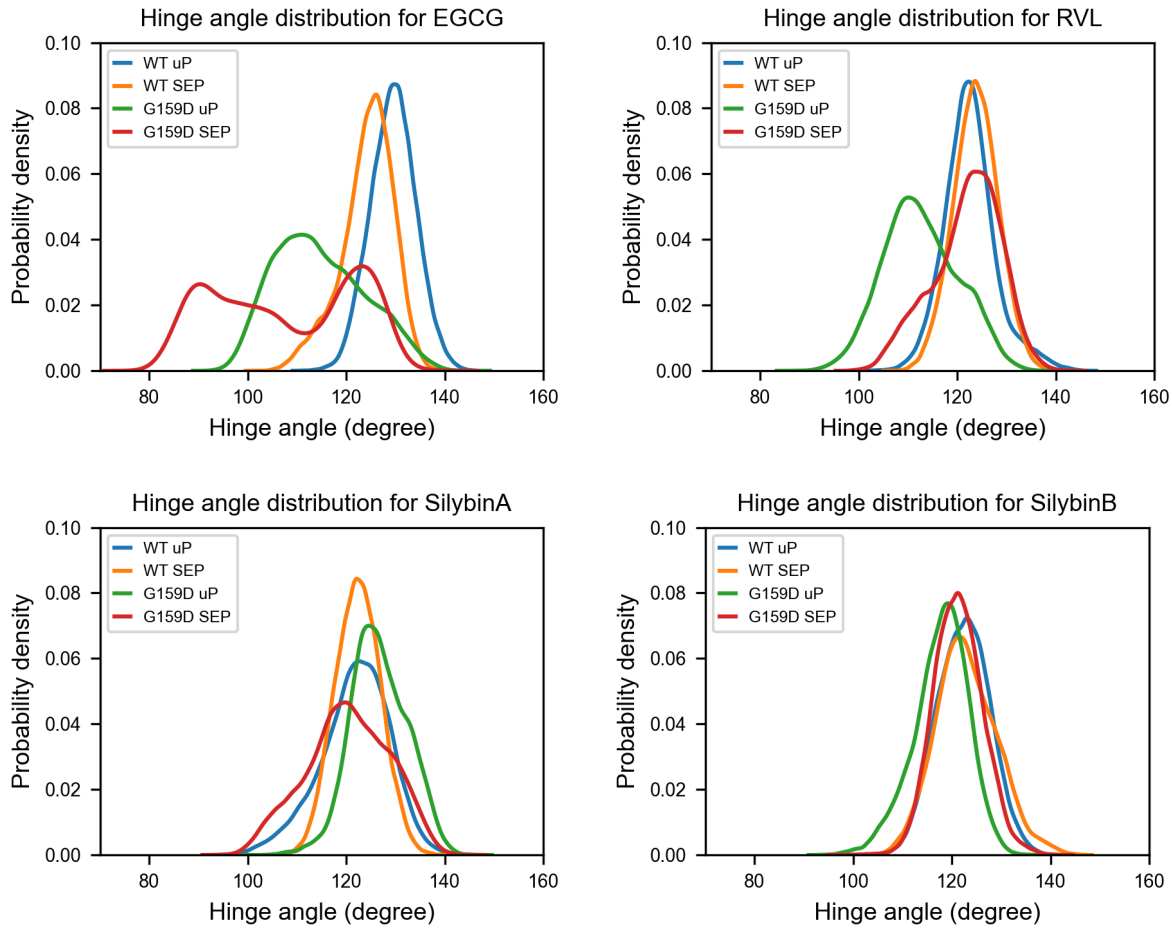


Figure 29: The interdomain angle distributions for EGCG, RVL, silybin A, and silybin B systems.

In the apo systems, it was found that both phosphorylation and mutation reduced the range of explored hinge angles (Figure 17). Phosphorylation shifted the hinge angle towards the higher angle direction in the wild type while the mutant was shifted to the lower angle direction. In the ligand systems (Figure 29 and Table 23), changes to the hinge angle distributions can be seen both compared to the apo systems as well as against each other. All systems were perturbed as can be expected from ligand insertion into parts of the protein system. For all wild-type simulations in both phosphorylation states with ligands, the range of explored

	Peak	Mean	Std. Dev.	FWHM
WT uP apo	122.9°	121.6°	7.2°	22.0°
WT uP EGCG	129.7°	129.5°	4.6°	10.7°
WT uP RVL	122.3°	122.6°	5.3°	9.8°
WT uP SilybinA	122.6°	121.8°	7.1°	15.0°
WT uP SilybinB	123.1°	122.3°	5.3°	13.3°
WT SEP apo	122.5°	123.3°	6.0°	16.4°
WT SEP EGCG	126.0°	124.2°	5.4°	10.5°
WT SEP RVL	123.6°	123.8°	4.6°	10.6°
WT SEP SilybinA	122.1°	122.4°	4.6°	11.3°
WT SEP SilybinB	121.7°	123.0°	6.2°	14.3°
G159D uP apo	121.7°	121.5°	4.8°	11.5°
G159D uP EGCG	110.9°	114.6°	9.2°	24.0°
G159D uP RVL	110.1°	112.6°	7.6°	18.3°
G159D uP SilybinA	124.6°	126.6°	5.8°	14.1°
G159D uP SilybinB	119.1°	117.7°	5.6°	11.5°
G159D SEP apo	117.1°	116.2°	6.0°	12.3°
G159D SEP EGCG	123.3°	107.6°	14.5°	22.5°
G159D SEP RVL	123.7°	121.9°	7.0°	13.9°
G159D SEP SilybinA	119.7°	120.3°	8.5°	20.8°
G159D SEP SilybinB	121.2°	121.3°	5.0°	11.6°

Table 23: The interdomain angle metrics for the systems with EGCG, RVL, silybin A, and silybin B.

hinge angles was reduced when compared to the apo simulations. In EGCG systems, both WT uP and WT SEP peaks were shifted to the higher angle direction and phosphorylation induced change in peak position was the opposite of the apo observations. In the mutant, however, the distributions were greatly broadened and further shifted to the lower angle conformation. Interestly, G159D SEP EGCG showed a clear bimodal distribution which was not seen in the other systems. In RVL systems, the narrowed wild type's distributions displayed a smilar phosphorylation-dependent shift in the mean (WT uP: 122.6°; WT SEP: 123.8°) as the apo systems. The unphosphorylated mutant with RVL was centered around 112.6° and moved to 121.9° upon phosphorylation which is closer to WT SEP apo than G159D SEP apo. The mutant with RVL also showed widened peaks on par with WT uP apo in terms of standard deviation. In both silybin A and B, the distributions of the four states were closer to each other compared to RVL and EGCG systems. With silybin A, phosphorylation shifted the peak to the left in the mutant, same as the apo mutants. However, both G159D uP and G159D SEP were moved to the higher angle end in presence of silybin A. In the wild type, silybin A introduced minimal changes in terms of peak position and spread. With silybin B, all four states have similar peak width and are centered close together. Upon phosphorylation, the wild type was shifted to the left and the mutant was shifted to the right.

Overall, the changes to the hinge angle distributions seemed varied. The recouplers, EGCG, RVL, and silybin B, all changed the mutant's phosphorylation-dependent shift in hinge angles back to the wild type apo direction. EGCG was an interesting case as its G159D SEP simulation presented a bimodal distribution where some of the density moved to the right and some of the density moved to the left from the G159D uP peak. Moreover, these compounds also perturbed the wild type's hinge angle dynamics significantly. Perhaps the measurement of troponin's overall conformation using hinge angle is not the most indicative metric to determine troponin's physiological behaviour especially when troponin is simulated in isolation from the thin filament.

4.2.3 TnC Domain Distance

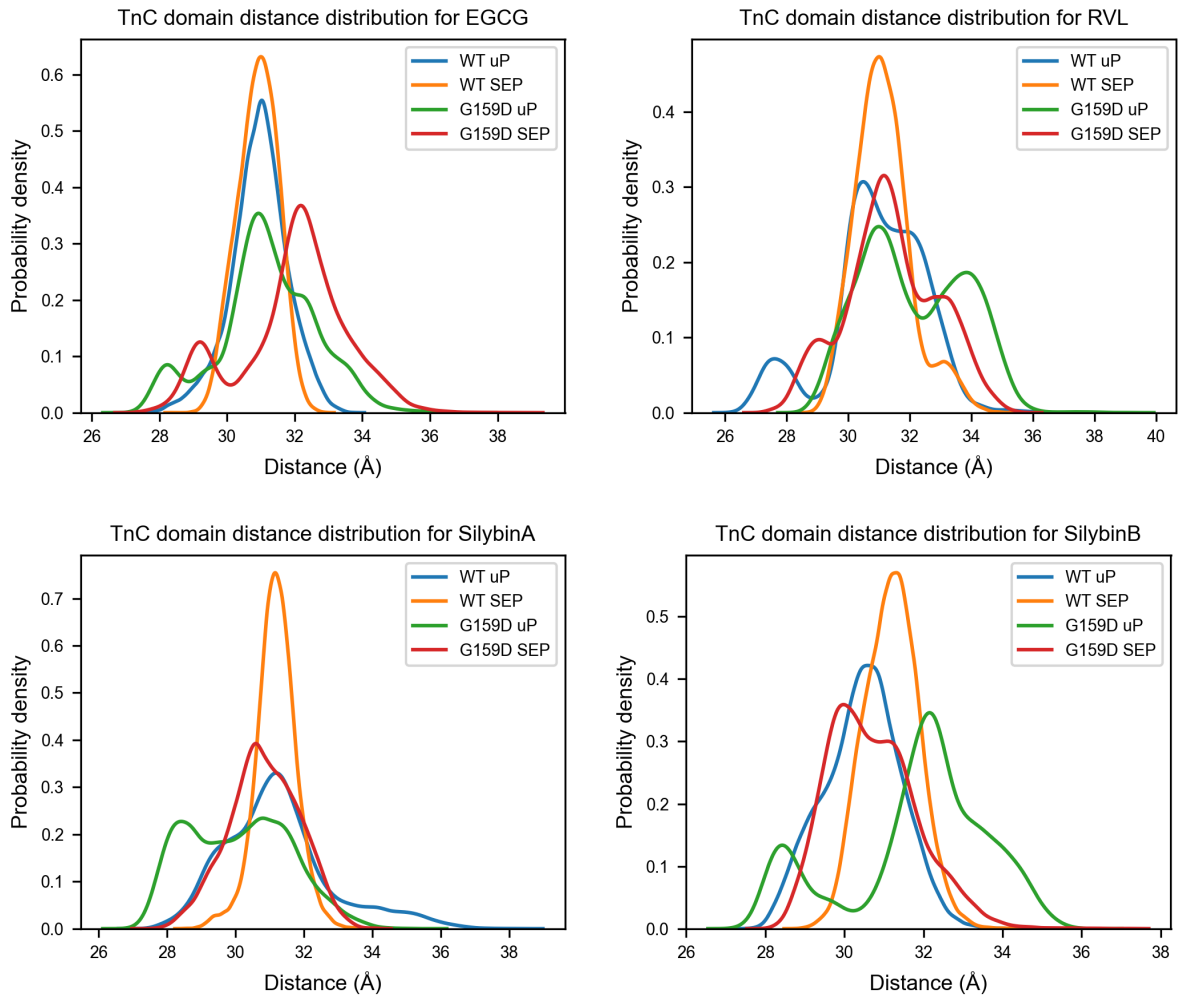


Figure 30: The TnC domain distance distributions for the systems with EGCG, RVL, silybin A, and silybin B.

The compaction of the regulatory domain, quantified by the distance between the N- and C-terminal domains of TnC, was found as a significant mutation effect that was casued directly by the G159D mutation site (Figure 18). By introducing ligand into the simulations (Figure 30 and Table 24), the mean TnC domain distance was increased in all systems and the range of explored angles was increased to be similar to the wild type apo level. In WT uP, the mean distance for EGCG, RVL, and silybin A systems increased from the apo level (30.2 Å) to between 31.0-31.2 Å. WT uP silybin B saw a very slight increase to 30.4 Å. Both EGCG and silybin B showed a reduction in peak spread while RVL and silybin A's standard deviation increased

	Peak	Mean	Std. Dev.	FWHM
WT uP apo	31.1 Å	30.2 Å	1.3 Å	3.0 Å
WT uP EGCG	31.0 Å	31.0 Å	0.8 Å	1.5 Å
WT uP RVL	30.5 Å	31.0 Å	1.6 Å	3.0 Å
WT uP SilybinA	31.2 Å	31.2 Å	1.6 Å	2.8 Å
WT uP SilybinB	30.6 Å	30.4 Å	1.0 Å	2.2 Å
WT SEP apo	29.6 Å	29.9 Å	0.9 Å	2.4 Å
WT SEP EGCG	31.0 Å	30.9 Å	0.6 Å	1.5 Å
WT SEP RVL	31.0 Å	31.1 Å	0.9 Å	1.9 Å
WT SEP SilybinA	31.2 Å	31.2 Å	0.6 Å	1.2 Å
WT SEP SilybinB	31.3 Å	31.2 Å	0.7 Å	1.7 Å
G159D uP apo	29.0 Å	29.2 Å	1.0 Å	1.6 Å
G159D uP EGCG	30.9 Å	31.2 Å	1.5 Å	2.4 Å
G159D uP RVL	31.0 Å	32.1 Å	1.7 Å	4.9 Å
G159D uP SilybinA	30.8 Å	30.1 Å	1.5 Å	4.4 Å
G159D uP SilybinB	32.2 Å	31.8 Å	1.8 Å	1.9 Å
G159D SEP apo	29.2 Å	29.6 Å	1.2 Å	1.5 Å
G159D SEP EGCG	32.2 Å	32.0 Å	1.6 Å	1.8 Å
G159D SEP RVL	31.2 Å	31.4 Å	1.5 Å	2.1 Å
G159D SEP SilybinA	30.6 Å	30.8 Å	1.0 Å	2.4 Å
G159D SEP SilybinB	30.0 Å	30.7 Å	1.1 Å	2.6 Å

Table 24: The TnC domain distance metrics for the systems with EGCG, RVL, silybin A, and silybin B.

to 1.6 Å from WT uP apo's 1.3 Å. In WT SEP, all ligand systems' mean distance increased to around 31 Å with similar or slightly reduced standard deviation when compared to the apo metrics. All ligand systems' mean TnC domain distances were comparable between WT uP and WT SEP state except silybin B, which saw a more pronounced increase. The reduction in dynamic range upon phosphorylation was also maintained in all ligand systems. In the G159D mutant, all systems saw an increase in the TnC domain separation in both phosphorylation states. The spread of distributions has increased for all G159D uP systems when ligands were introduced. EGCG and RVL introduced an increase in standard deviation in G159D SEP while silybin A and B introduced a slight decrease in the spread of data. With phosphorylation, the movement in the domain separation is more pronounced in the mutant than the wild type (Figure 30). EGCG and silybin A have seen a moderate shift towards the more distant direction while RVL and silybin B shifted to the more closely packed direction.

Overall, the addition of ligands increased the TnC domain distance in all systems and relieved troponin from the locked position as seen in G159D apo systems. This could be caused by the ligand binding to the CcTnC region as shown by previous NMR studies where it can disrupt the ionic interactions from TnC Asp159 in the mutant and restore TnC's dynamic movements.[26,63]

4.3 Changes in Peptide Binding Affinity

In the apo studies, the binding energy of four peptides of interest to the rest of troponin was measured. The effects of phosphorylation and mutation on these peptides' interactions presented an interesting landscape in how these perturbations could affect troponin's physiological functions. Where possible, the computationally determined changes in binding energy were compared to experiments and the direction of changes corroborated with experimentally determined binding affinity. Here, the changes in peptide binding energy were described when EGCG, RVL, silybin A, and silybin B were introduced to the simulations individually.

	Mean	Std. Dev.		Mean	Std. Dev.
WT uP apo	-108.4	24.0	G159D uP apo	-113.5	13.8
WT uP EGCG	-103.2	17.6	G159D uP EGCG	-109.3	13.7
WT uP RVL	-101.8	13.9	G159D uP RVL	-104.9	15.7
WT uP SilybinA	-108.3	17.2	G159D uP SilybinA	-113.6	15.0
WT uP SilybinB	-109.4	15.5	G159D uP SilybinB	-106.7	17.9
WT SEP apo	-111.9	13.9	G159D SEP apo	-117.0	15.0
WT SEP EGCG	-137.9	16.0	G159D SEP EGCG	-122.9	13.6
WT SEP RVL	-128.2	19.6	G159D SEP RVL	-106.9	15.2
WT SEP SilybinA	-129.4	16.2	G159D SEP SilybinA	-111.8	16.1
WT SEP SilybinB	-125.7	14.9	G159D SEP SilybinB	-106.8	13.5

Table 25: Switch peptide binding energy to the rest of the troponin complex for troponin systems with EGCG, RVL, silybin A, and silybin B. Unit: $\mathrm{kcal\,mol^{-1}}$

4.3.1 Switch Peptide

The switch peptide which binds to the hydrophobic patch formed by helix A and B has a significant role in the thin filament regulatory process. In the apo systems, an increase in binding energy was observed in both WT and G159D upon phosphorylation and the mutant had a stronger interaction over the wild type. In WT uP, the binding was weakened with the introduction of EGCG and RVL (Table 25). WT uP silybin B strengthened the interation slightly and silybin A did not introduce a significant change. In WT SEP, all ligand systems saw a big increase in the binding energy from $-111.9\,\mathrm{kcal\,mol^{-1}}$ in apo to more than $-125.7\,\mathrm{kcal\,mol^{-1}}$ with EGCG system increased the most to $-137.9\,\mathrm{kcal\,mol^{-1}}$. Similar changes were seen in the mutant. In G159D uP, EGCG, RVL, and silybin B all reduced the switch peptide binding energy to be more similar to the WT uP apo level. Interestingly, silybin A did not influence the average binding energy in G159D uP, same as in the unphosphorylated wild type. In G159D SEP systems, EGCG increased the binding energy while the other molecules weakened the interactions. From G159D SEP apo's $-117.0\,\mathrm{kcal\,mol^{-1}}$, EGCG increased the binding energy to $-122.9\,\mathrm{kcal\,mol^{-1}}$ and RVL and silybin B both reduced the binding energy to around $-106\,\mathrm{kcal\,mol^{-1}}$. Silybin A reduced the switch peptide binding energy to $-111.8\,\mathrm{kcal\,mol^{-1}}$

which was very similar to the WT SEP apo level. In the wild type, phosphorylation uniformly strengthened the switch peptide interaction in all ligand systems and the increase in binding energy was much increased compared to the apo systems. The phosphorylation effect was more varied in the mutant with EGCG and RVL still maintaining an increase, silybin B system was not changed upon phosphorylation and silybin A system was weakened from uP to SEP.

The significant increase in switch peptide binding energy in WT SEP was interesting and it may potentially be related to the desensitising effects of these molecules. Silybin A introduced minimal changes in the uP states in both the wild type and the mutant which could be linked to its inactive properties towords coupling.

4.3.2 Inhibitory Peptide

	Mean	Std. Dev.		Mean	Std. Dev.
WT uP apo	-153.7	23.4	G159D uP apo	-168.5	23.8
WT uP EGCG	-139.6	20.4	G159D uP EGCG	-149.3	19.4
WT uP RVL	-142.0	24.1	G159D uP RVL	-147.0	24.1
WT uP SilybinA	-137.0	18.8	G159D uP SilybinA	-138.3	18.1
WT uP SilybinB	-145.4	21.9	G159D uP SilybinB	-148.9	23.6
WT SEP apo	-157.6	20.2	G159D SEP apo	-164.9	20.5
WT SEP EGCG	-150.4	16.7	G159D SEP EGCG	-153.7	24.0
WT SEP RVL	-149.3	20.1	G159D SEP RVL	-157.0	21.9
WT SEP SilybinA	-145.4	14.3	G159D SEP SilybinA	-144.1	18.8
WT SEP SilybinB	-149.8	14.1	G159D SEP SilybinB	-152.2	24.4

Table 26: Inhibitory peptide binding energy to the rest of the troponin complex for troponin systems with EGCG, RVL, silybin A, and silybin B. Unit: $\mathrm{kcal\,mol^{-1}}$

The inhibitory peptide presents a direct effect on the inhibition of the actin binding sites on tropomyosin. In the apo study, it was observed that phosphorylation strengthened the inhibitory peptide binding to the rest of the troponin complex in the wild type. Mutation strengthened the inhibitory peptide interactions in both phosphorylation states and caused the phosphorylated mutant to have a weaker interaction which was the opposite change

compared to the wild type. In ligand simulations (Table 26), it was found that all inhibitory peptide binding energies were weakened from their respective apo state. In the wild type, the binding energy was weakened by around $10\,\mathrm{kcal\,mol^{-1}}$ in both uP and SEP systems with all ligands. The phosphorylation induced increase in binding energy was also maintained in ligand systems. In G159D uP, the ligand systems' binding energies were reduced to be on par with WT uP apo level except silybin A, where a $-30.2\,\mathrm{kcal\,mol^{-1}}$ reduction in binding energy brought it to be significantly weaker than WT uP apo. Similar changes in binding energies in the presence of ligands can be seen in G159D SEP where EGCG, RVL, silybin B reduced the binding energy to be more similar with WT SEP apo and silybin A reduced the energy to be well below the WT SEP apo level. When considering the phosphorylation induced changes in the mutant, all ligand systems had strengthened inhibitory peptide binding energy in the SEP state which is in line with the wild-type behaviour.

4.3.3 $\mathbf{TnI_{34-71}}$

	Mean	Std. Dev.		Mean	Std. Dev.
WT uP apo	-215.7	21.2	G159D uP apo	-212.4	19.9
WT uP EGCG	-230.1	24.0	G159D uP EGCG	-206.7	18.8
WT uP RVL	-213.9	22.1	G159D uP RVL	-200.8	20.3
WT uP SilybinA	-214.8	23.9	G159D uP SilybinA	-211.9	18.6
WT uP SilybinB	-216.9	22.9	G159D uP SilybinB	-209.9	19.6
WT SEP apo	-226.2	20.1	G159D SEP apo	-217.3	20.0
WT SEP EGCG	-233.2	20.3	G159D SEP EGCG	-212.5	23.6
WT SEP RVL	-227.5	18.3	G159D SEP RVL	-223.4	21.3
WT SEP SilybinA	-225.4	19.5	G159D SEP SilybinA	-217.1	20.3
WT SEP SilybinB	-230.2	19.8	G159D SEP SilybinB	-224.2	22.5

Table 27: TnI_{34-71} peptide binding energy to the rest of the troponin complex for troponin systems with EGCG, RVL, silybin A, and silybin B. Unit: $\mathrm{kcal\,mol^{-1}}$

The TnI_{34-71} region is critical in the anchoring of TnC to the IT arm and the positioning of the regulatory head in relationship to the rest of the troponin complex. It was proposed that the

G159D mutation which is located close to the TnI_{34-71} region causes cardiac defects partially due to the weakened TnC anchoring. In the apo study, both the wild type and the mutant showed an increase in the TnI_{34-71} peptide binding energy upon phosphorylation however the mutant's TnI_{34-71} interaction was weakened especially in the phosphorylated state when compared to the wild type. In WT uP, EGCG introduced a significant strengthening in the TnI_{34-71} binding energy, from $-215.7\,\mathrm{kcal\,mol^{-1}}$ in WT uP apo to $-230.1\,\mathrm{kcal\,mol^{-1}}$ (Table 27). This is likely caused by EGCG directly binding to the CcTnC/TnI region as shown by NMR and computational studies hence directly strengthening this interaction.[63,142] The other ligands have introduced minimal changes, less than $2\,\mathrm{kcal\,mol^{-1}}$ to the binding energy. A similar pattern was seen in the phosphorylated state where EGCG induced a more significant increase in binding energy than silybin B and RVL and silybin A system stayed largely the same. In terms of the phosphorylation effect in the wild type with ligands, the strengthening effect was maintained in all systems. The difference in binding energy was largely the same for RVL, silybin A, and silybin B. EGCG systems saw a much reduced increase due to the significant increased binding in the WT uP state. In the mutant, the ligands further weakened the TnI_{34-71} peptide binding in the uP state. RVL reduced the binding energy the most, from $-212.4\,\mathrm{kcal\,mol^{-1}}$ in G159D uP apo to $-200.8\,\mathrm{kcal\,mol^{-1}}$ while silybin A only marginally changed the binding energy ($-211.9\,\mathrm{kcal\,mol^{-1}}$). In G159D SEP, EGCG further reduced the binding energy from $-217.3\,\mathrm{kcal\,mol^{-1}}$ to $-212.5\,\mathrm{kcal\,mol^{-1}}$. RVL and silybin B strengthened the binding energy to $-223.4\,\mathrm{kcal\,mol^{-1}}$ and $-224.2\,\mathrm{kcal\,mol^{-1}}$, respectively, which were more akin to the WT SEP apo level. Silybin A did not change the binding energy much and stayed at $-217.1\,\mathrm{kcal\,mol^{-1}}$. In simulations of the CcTnC segment in the presence of TnI_{34-71}, Botten *et al.* found that the inter-protein interactions were reduced when EGCG was introduced into the system.[142] The energetics calculation presented here is in agreement with this study whether in the phosphorylated or unphosphorylated state of the mutant although the opposite was observed in the wild type. When comparing G159D SEP to G159D uP, all ligand systems increased the binding energy in the phosphorylated state. Overall, the inactive silybin A introduced minimal changes in TnI_{34-71} binding energies in all systems.

	Mean	Std. Dev.		Mean	Std. Dev.
WT uP apo	-129.2	28.4	G159D uP apo	-111.3	33.2
WT uP EGCG	-124.7	26.1	G159D uP EGCG	-111.0	27.2
WT uP RVL	-119.0	25.8	G159D uP RVL	-108.1	24.4
WT uP SilybinA	-102.2	32.3	G159D uP SilybinA	-116.8	27.1
WT uP SilybinB	-112.8	21.0	G159D uP SilybinB	-116.1	28.0
WT SEP apo	-99.8	30.8	G159D SEP apo	-116.6	24.8
WT SEP EGCG	-103.8	20.4	G159D SEP EGCG	-103.2	26.7
WT SEP RVL	-99.0	23.4	G159D SEP RVL	-87.5	39.2
WT SEP SilybinA	-112.6	22.8	G159D SEP SilybinA	-82.4	28.1
WT SEP SilybinB	-93.7	24.1	G159D SEP SilybinB	-102.2	44.3

Table 28: Binding energy of NcTnI to the rest of troponin for troponin systems with EGCG, RVL, silybin A, and silybin B. Unit: $\mathrm{kcal\,mol^{-1}}$

4.3.4 NcTnI

NcTnI which contains the phosphorylation sites wraps around NcTnC and phosphorylation has shown to influence the NcTnI/NcTnC interactions significantly. In the apo systems, phosphorylation was found to reduce the NcTnI binding energy in the wild type and the opposite change was found in the mutant. In WT uP, all ligand simulations reduced the NcTnI binding energy (Table 28), likely caused by the ligand binding to the regulatory head and modulating the interactions between NcTnI and NcTnC. In WT SEP, EGCG and silybin A increased the binding interaction from the apo level of $-99.8\,\mathrm{kcal\,mol^{-1}}$ to $-103.8\,\mathrm{kcal\,mol^{-1}}$ and $-112.6\,\mathrm{kcal\,mol^{-1}}$, respectively. Silybin B reduced the NcTnI binding energy to $-93.7\,\mathrm{kcal\,mol^{-1}}$ and RVL introduced a negligible change in the binding energy. Comparing WT SEP to WT uP, EGCG, RVL, and silybin B all maintained the reduction in NcTnI binding energy upon phosphorylation while silybin A caused a strengthening in this interaction. In the mutant, silybin A and B increased the binding energy in the unphosphorylated state and RVL reduced the interaction while EGCG did not change this interaction energy. In the phosphorylated state, all ligand systems introduced significant reductions to the binding energy from the G159D SEP apo level of $-116.6\,\mathrm{kcal\,mol^{-1}}$. EGCG and silybin B reduced the binding energy to a sim-

ilar level of $-103.2\,\mathrm{kcal\,mol^{-1}}$ and $-102.2\,\mathrm{kcal\,mol^{-1}}$, respectively. RVL and silybin A further reduced the binding energy to $-87.5\,\mathrm{kcal\,mol^{-1}}$ and $-82.4\,\mathrm{kcal\,mol^{-1}}$, respectively. These significant weakening in NcTnI interactions meant that the normal phosphorylation induced change in NcTnI binding energy was restored in all the ligand simulations in the mutant. RVL and silybin A, in particular, restored this change to a similar magnitude as in the WT apo system.

4.4 Change in Residue Interactions

The direct effect of the missense G159D mutation is the additional ionic interactions formed from the negatively charged aspartic acid. In the apo studies, using residue interaction profiling method, it was identified that the majority of ionic interactions formed from TnC Asp159 was with TnC Arg83 (Section 3.5.8). This long lasting interaction was found in 80% of the simulations in G159D uP and 76% in G159D SEP along with low percentage interactions with TnC Lys90 and inhibitory peptide residues TnI Arg145 and TnI Arg147. The additional ionic interactions were postulated to directly contribute to the defects in the troponin mutant, such as the more rigid dynamics, the closer TnC domain distance, and the increased inhibitory peptide binding energy. Table 29 summarises the interaction percentage between TnC Asp159 and three TnC residues. The major ionic interaction between TnC Asp159 and TnC Arg83 was greatly subsidised in all ligand systems to be around or below 50%. In G159D uP, the largest reduction was from the apo level of 80% to 8% in silybin A simulations. RVL system saw the least reduction (51%) while EGCG and silybin B had 22% and 17% of interactions, respectively. G159D SEP saw smaller changes compared to the uP state. EGCG system had the least change, from 76% in apo to 53%, and silybin B produced the largest change, to 23%. RVL and silybin A reduced the interaction percentage to 46% and 40%, respectively. Part of the reduced percentage was shifted to interactions between TnC Asp159 and TnC residues Lys86 and Lys90. The ionic interaction betwen TnC Asp159 and TnC Lys86 was not seen in the apo systems while it existed for 4-26% of the simulations in the ligand systems. TnC Asp159 and TnC Lys90 ionic interactions existed in G159D SEP apo for 6% but not in the uP system. In the ligand systems, this interaction was seen for 2-9% of the time.

	TnC Arg83		TnC Lys86		TnC Lys90	
	uP	SEP	uP	SEP	uP	SEP
apo	80%	76%	0%	0%	0%	6%
EGCG	22%	53%	12%	6%	8%	2%
RVL	51%	46%	8%	4%	2%	3%
SilybinA	8%	40%	11%	17%	9%	2%
SilybinB	17%	23%	19%	26%	3%	6%

Table 29: Ionic interaction percentages between TnC Asp159 and TnC residues in simulations with EGCG, RVL, silybin A, and silybin B.

In addition to the changes in interaction pattern from TnC Asp159 to TnC residues, the mutant residue also formed more interactions with arginines that are part of the TnI inhibitory peptide. As tabulated in Table 30, major increase in interaction percentage was seen in the existing ionic interactions with TnI Arg145 and TnI Arg147 alongside a new interaction with TnI Arg144 that was not seen in the apo systems. For these three ionic interactions, the uP state saw bigger increases where the change in the SEP state was more gradual. With TnI Arg144, the interaction existed for 12% and 23% of the simulations with EGCG and RVL in the uP state where silybin B saw a small increase to 2% and RVL system was not changes. In the SEP state, only EGCG system increased the interaction percentage very slightly. Similar pattern was seen in the interactions with TnI Arg145 where EGCG and RVL increased the percentage in the uP state more than the other two ligands. In the SEP state, the interaction existed for 14% of the simulations in the apo system. EGCG and silybin B reduced the interaction to 9% and 6%, respectively. Silybin A maintained a similar percentage where RVL annihilated this interaction. With TnI Arg147, bigger increses was seen in EGCG and RVL systems in the uP state. Silybin B also increased the interaction percentage drastically to 37%. Less than 5% change was seen in all ligand systems in the SEP state.

The change in the TnC Asp159 interaction pattern in the mutant when ligands were introduced represent both a shift in interactions with TnC and towards inhibitory peptide residues. The reduced interactions with TnC Arg83 and increased interactions with TnC Lys86 and Lys90 means the CcTnC mutant residue was less locked together with the N-terminal domain residue and more moved towards the C-terminal end of the flexible linker. This would relieve

	TnI Arg144		TnI Arg145		TnI Arg147	
	uP	SEP	uP	SEP	uP	SEP
apo	0%	0%	1%	14%	1%	4%
EGCG	12%	1%	10%	9%	25%	0%
RVL	23%	0%	15%	0%	13%	2%
SilybinA	0%	0%	2%	16%	3%	4%
SilybinB	2%	0%	0%	6%	37%	9%

Table 30: Ionic interaction percentages between TnC Asp159 and TnI residues in simulations with EGCG, RVL, silybin A, and silybin B.

TnC from the compact locked conformation as seen in the apo mutant systems hence allow the TnC domain separation to increase to the wild-type level as described in Section 4.2.3. Although the ligands increased the interactions between the inhibitory peptide and TnC Asp159, particularly in the uP state, the binding energy of the inhibitory peptide to the rest of the troponin system greatly reduced based on MMPBSA measurements (Section 4.3.2). The relative bigger reduction in the G159D uP state compared to the SEP state means that the phosphorylation induced change in the inhibitory peptide was restored to the wild-type direction. Hence the change in the interaction percentages might represent a repositioning of the inhibitory peptide due to ligand binding which reduced its binding interactions with the rest of the troponin complex.

4.5 Binding Region Clustering

This section presents the clusters of binding regions explored by the ligands during the simulations. The illustrations represent the ligand cluster centers with their respective representative protein structure of the system. The ligands are coloured based on the cluster size, from large to small the order follows pink, orange, yellow, and cyan, which are also given the numerical numbering of cluster 1 to 4. The results quoted in this section represent the cluster size as a percentage of the total simulation trajectories for each system.

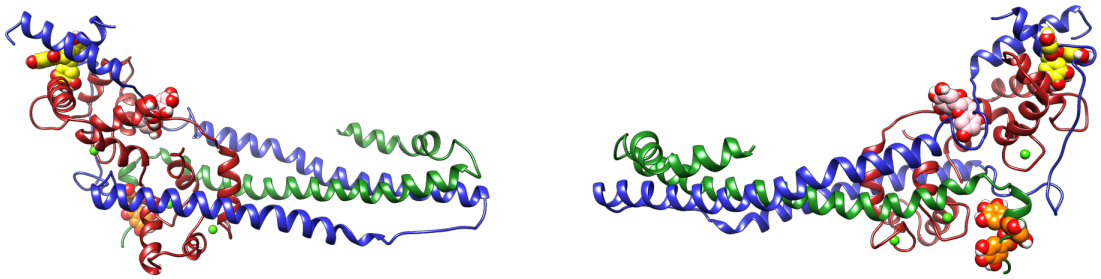
4.5.1 EGCG

	Pink	Orange	Yellow	Cyan
WT uP	42%	36%	20%	
WT SEP	36%	34%	18%	5%
G159D uP	58%	39%		
G159D SEP	83%	14%	1%	

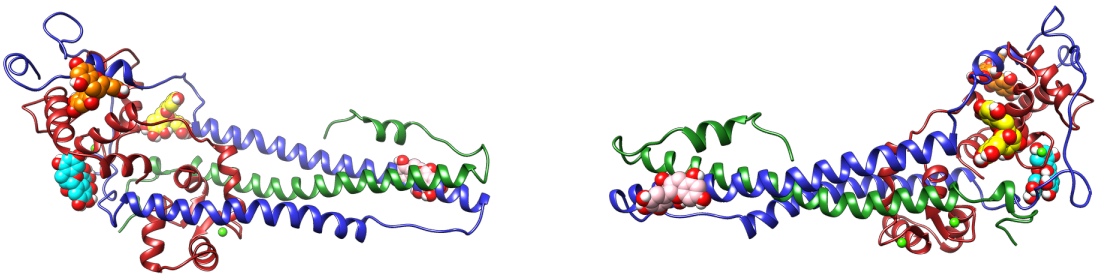
Table 31: Binding region cluster percentage for EGCG systems.

Figure 31 shows the troponin systems with EGCG at cluster centers. For each system, the depiction on the right is the left image rotated 180° to show the other side of the protein-ligand binding interactions. In WT uP, the most populated cluster accounted for 42% of the ligand positions (Table 31) and was located around EF-hand II hydrophobic cleft. In this position, EGCG also interacted with the TnI inhibitory peptide and the TnC flexible linker. The second largest cluster with 36% population was located at the hinge region where the ligand interacted with NcTnI, EF-hand IV, and TnT. The least populated cluster (20%) was located at the tip of the regulatory head. In this position, EGCG was inserted into the hydrophobic cleft of EF-hand I and interacted with TnI switch peptide and N-terminal end of TnI. In the phosphorylated wild type, the most populated cluster (36%) was around the end of the IT arm region where EGCG interacted with the helices. These interactions tend to be weaker than when EGCG was inserted deep into the binding pocket. The second cluster (34%) was near the NcTnC hydrophobic cleft where EGCG also formed interactions with the switch peptide. The third cluster (18%) was located near the hinge region and included interactions with the inhibitory peptide, TnT, and NcTnI. The last cluster which accounted for 5% of the simulations was located at the back of the regulatory head. In this region, EGCG interacted with NcTnI and EF-hand II close to the ion binding loop.

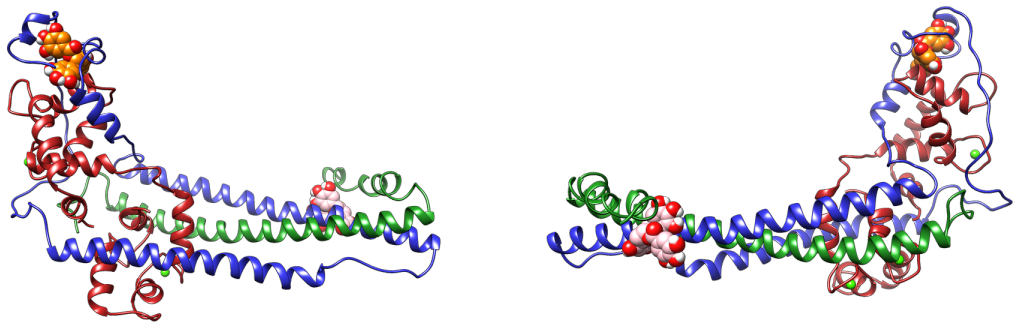
In the mutant unphosphorylated state, EGCG stayed near the tail of the IT arm for 58% of the simulations. The second cluster which occupied 39% of the simulations was located at the tip of the regulatory head and EGCG interacted with EF-hand I and N- and C-terminal end of TnI. In the phosphorylated case, EGCG stayed at the back of the regulatory head for majority



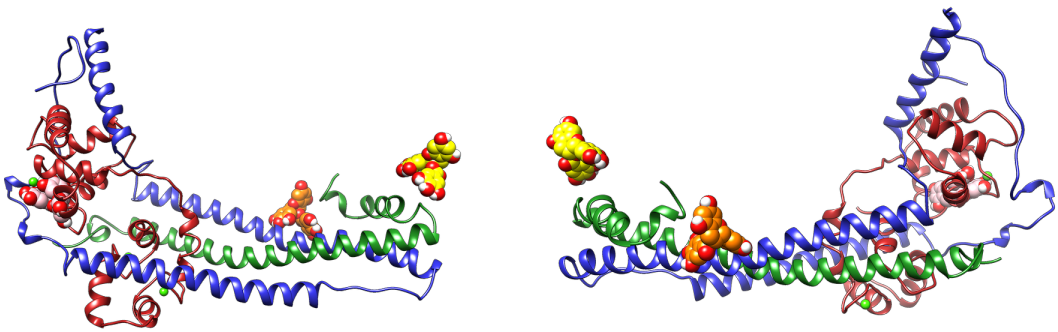
(a) WT uP



(b) WT SEP



(c) G159D uP



(d) G159D SEP

Figure 31: Binding site clusters for systems with EGCG.

of the simulations (83%). In this position, EGCG interacted with NcTnI and EF-hand I around the loop region. For 14% and 1% of the simulations, the ligand stayed near the middle and the end of the IT arm, respectively.

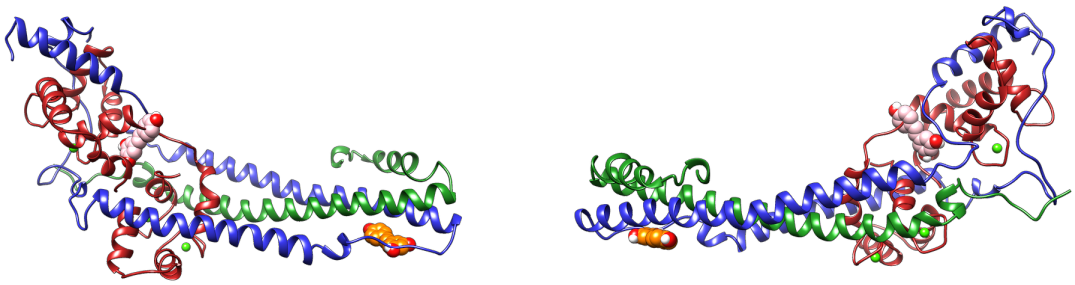
Overall, the EGCG binding position clusters were mainly centered around the regulatory head and the end of the IT arm. Several prominent binding sites were centered around the hydrophobic pocket of NcTnC, TnC domain interface, and the back of the hinge next to the EF-hand II ion binding loop. EGCG interacted less frequently with the regulatory head upon phosphorylation in the wild type and the opposite was observed in the mutant where majority of the binding interactions happened at the regulatory head in the phosphorylated state.

4.5.2 RVL

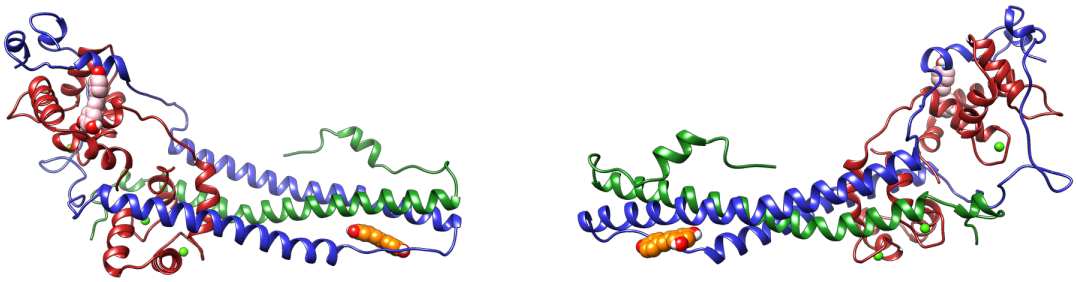
	Pink	Orange	Yellow	Cyan
WT uP	84%	2%		
WT SEP	66%	25%		
G159D uP	73%	7%	3%	
G159D SEP	85%			

Table 32: Binding region cluster percentages for RVL systems.

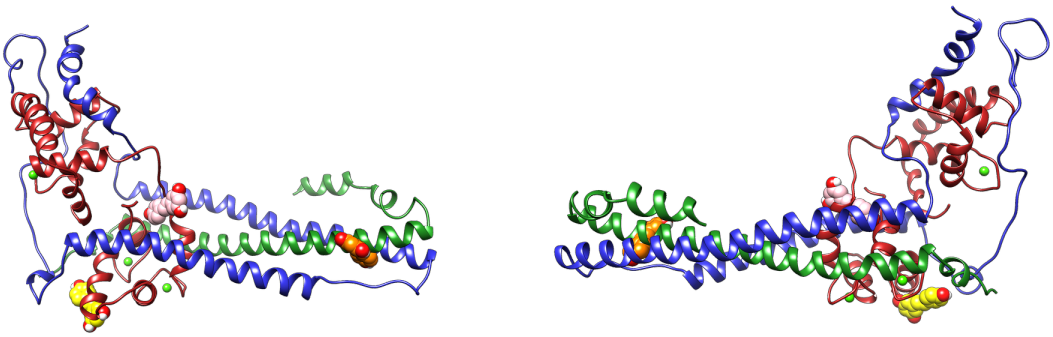
For the RVL systems, the main clusters were all located at the regulatory head (Section 4.5.2 and Table 32). In WT uP, the main cluster accounted for 84% of the simulations and the cluster center was at the hinge region of troponin. The minor cluster was at the tail of the IT arm and accounted for 2% of the simulations. The phosphorylated wild type presented a similar binding pattern to the unphosphorylated one with the main cluster centered around NcTnC hydrophobic patch region (66%). In this state, the percentage of the minor cluster centered around the tail of the IT arm increased to 25%. In the unphosphorylated mutant, the largest cluster (73%) shifted more towards the CcTnC direction. RVL was found between the TnC and TnI anchoring region and next to the mutation site. The cluster at the tail of the IT arm accounted for 7% of the simulations. In 3% of the simulations, the ligand was found at the back of the hinge and interacted with CcTnC. In G159D SEP, there was a single cluster cen-



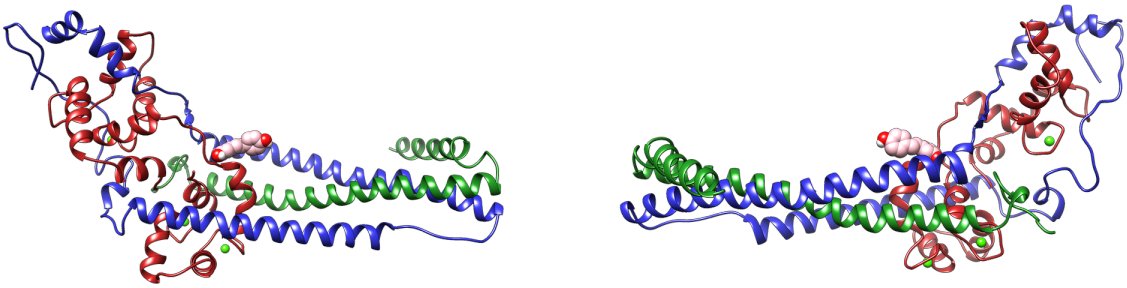
(a) WT uP



(b) WT SEP



(c) G159D uP



(d) G159D SEP

Figure 32: Binding site clusters for systems with RVL.

tered around the hinge region next to CcTnC and the inhibitory peptide. This single cluster accounted for 85% of the simulations.

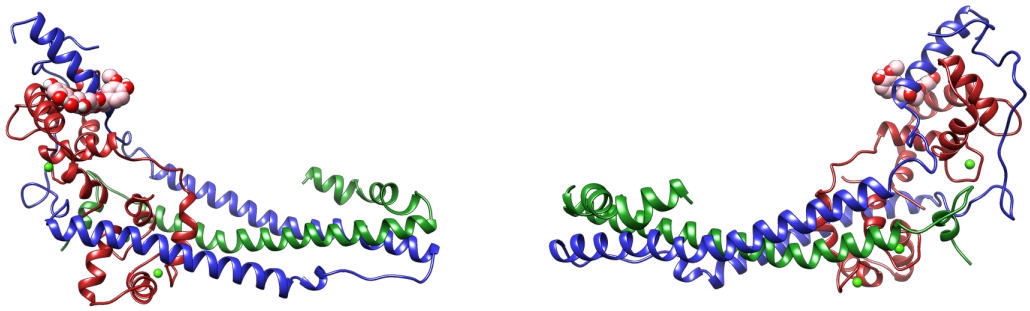
Overall, majority of the binding interactions of RVL happened at the regulatory head. A shift in binding area was observed from the wild type to the mutant. In WT, the major clusters tend to centered around NcTnC area while they were shifted to the CcTnC area in the mutant. This could be caused by the extra negative charge introduced by the aspartic acid which could attract the phenyl groups of RVL.

4.5.3 Silybin A

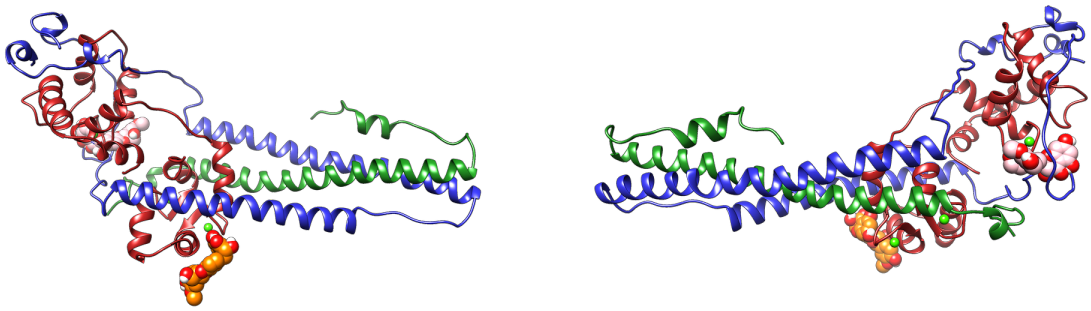
	Pink	Orange	Yellow	Cyan
WT uP	98%			
WT SEP	83%	8%		
G159D uP	67%	20%	10%	
G159D SEP	63%	29%		

Table 33: Binding region cluster percentages for Silybin A systems.

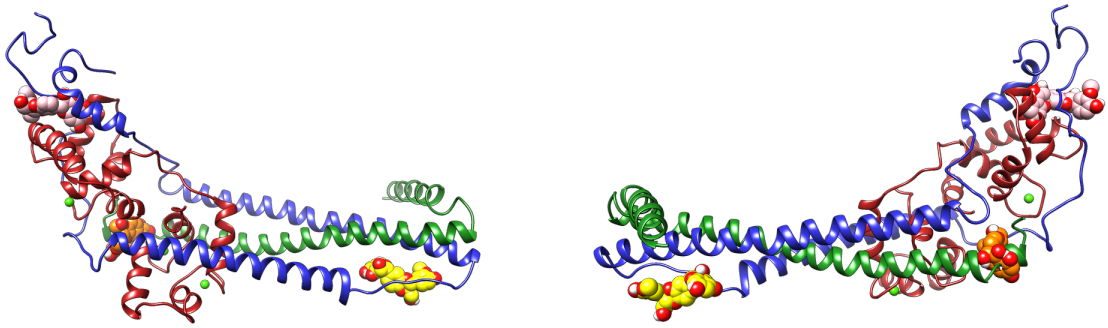
In silybin A systems, majority of the binding interactions happened at the regulatory head (Figure 33 and Table 33). In WT uP, silybin A almost exclusively interacted with NcTnC (98% of the simulations). At the cluster center, silybin A was inserted into the hydrophobic patch of EF-hand I and formed extensive interactions with the switch peptide. In the phosphorylated wild type, the major cluster (83%) was centered around the back of the hinge region. Silybin A inserted inbetween NcTnI and TnC and interacted with EF-hand II's loop region. The minor cluster accounted for 8% of the simulations and was located at the tail of the IT arm. In the unphosphorylated mutant, the largest cluster of 67% was centered around the tip of the regulatory head and interacted with the hydrophobic pocket of EF-hand I, and N- and C-terminal region of TnI. The second largest cluster of 20% was centerd around the back of the hinge region. In this area, silybin A was at the domain interface of TnC and interacted with NcTnI. The last cluster was located at the tail of the IT arm and accounted for 10% of the simulations. For the phosphorylated mutant, the major cluster (63%) was located at the hinge



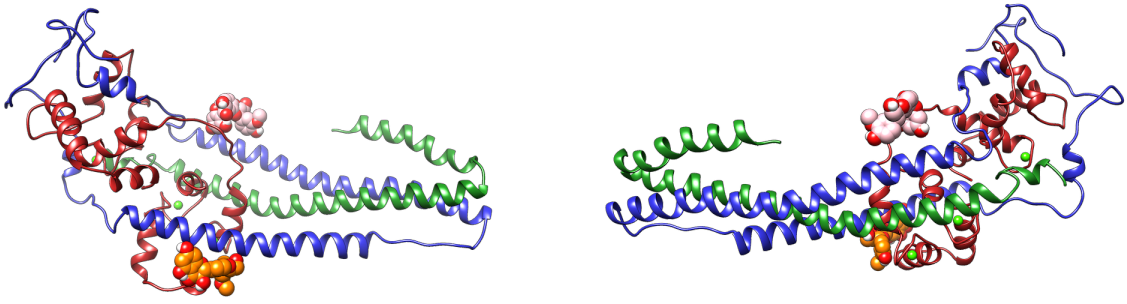
(a) WT uP



(b) WT SEP



(c) G159D uP



(d) G159D SEP

Figure 33: Binding site clusters for systems with Silybin A

region where the cluster was centerd around the TnC flexible linker and the TnI inhibitory peptide. The secondary cluster of 29% was located near CcTnC and interacting with nearby TnI helix as well.

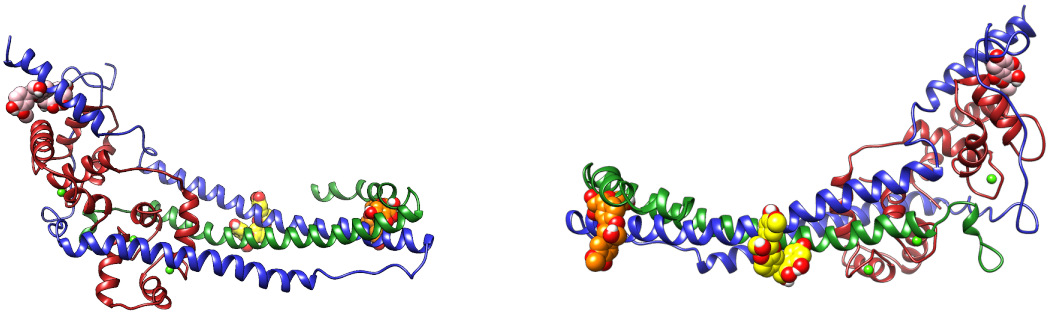
Overall, silybin A has a high tendency to interact with NcTnC in WT uP. This behaviour was seen in G159D uP as well with a secondary binding cluster at the back of the hinge region. The interactions at the back of the hinge region fixed NcTnI to TnC subunit and was prominent in WT SEP. Silybin A also formed interactions with CcTnC and TnI in the phosphorylated state of both WT and G159D. Finally, silybin A in G159D SEP possessed a unique cluster at the hinge region and interacted with both the inhibitory peptide and TnC flexible linker.

4.5.4 Silybin B

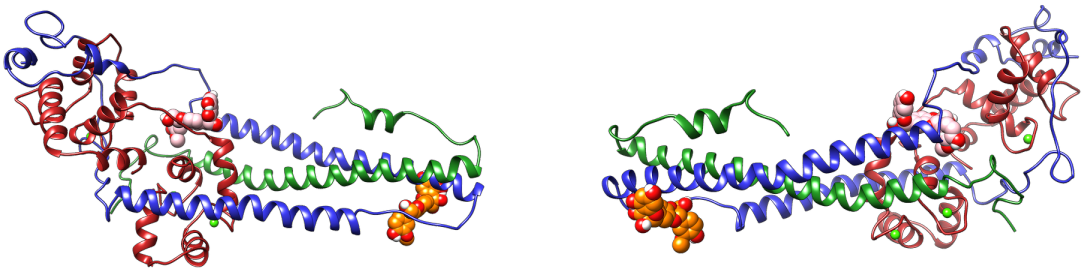
	Pink	Orange	Yellow	Cyan
WT uP	83%	10%	2%	
WT SEP	93%	1%		
G159D uP	97%			
G159D SEP	58%	36%	3%	

Table 34: Binding region cluster percentages for silybin B systems.

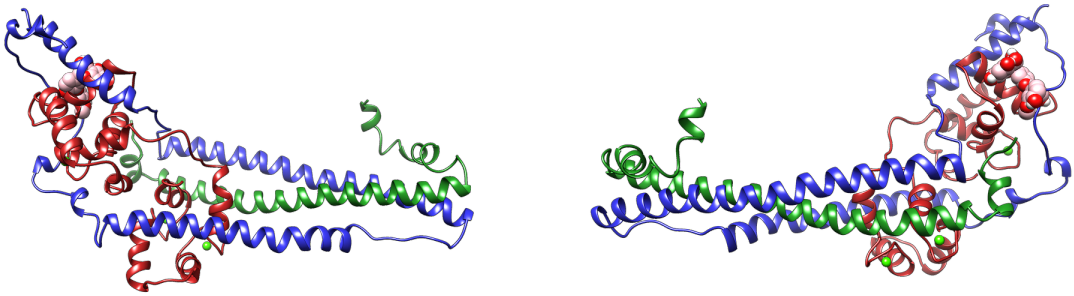
Majority of the silybin B interactions took place at the regulatory head (Figure 34 and Table 33). In WT uP, silybin B interacted around EF-hand I and switch peptide for 83% of the simulations. Two minor clusters were centered around the middle and the end of the IT arm for 10% and 2% of the simulations, respectively. In the phosphorylated state, the major cluster was shifted more towards CcTnC direction. This binding position was inserted into the hinge region and between the inhibitory peptide and TnC flexible linker. The minor cluster of 1% was at the tail of the IT arm. In the mutant, the unphosphorylated state presented a single cluster centered at the tip of the regulatory domain, similar to the WT uP system. This cluster accounted for 97% of the simulations. In G159D SEP, the major cluster (58%) was at the troponin domain interface where the cluster center was inserted in between TnC and NcTnI and formed interactions with the inhibitory peptide. The second largest cluster (36%)



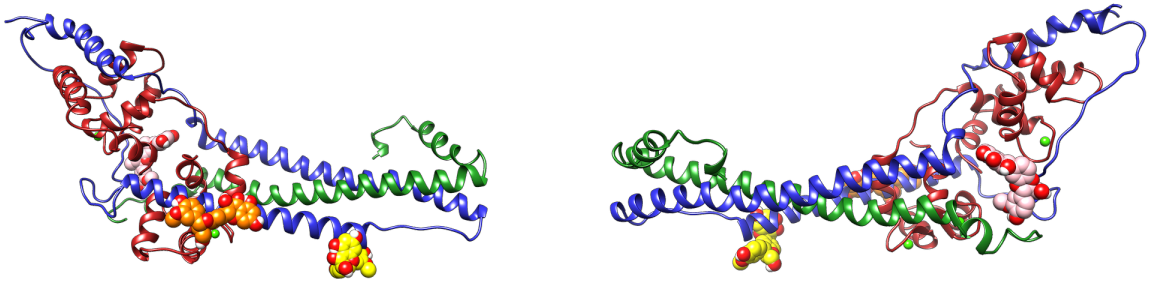
(a) WT uP



(b) WT SEP



(c) G159D uP



(d) G159D SEP

Figure 34: Binding site clusters for systems with Silybin B

was located at the region where CcTnC anchors with TnI helix and 3% of the simulations had silybin B around the middle of the IT arm. Overall, the interaction clusters shifted towards the CcTnC region upon phosphorylation in both wild type and the mutant.

4.6 Ligand binding pattern

In the previous section, the general regions that ligands interacted with were discussed and several common binding areas on the regulatory head and the IT arm were identified. In this section, the ligand interactions were further delineated by presenting the quantitative per residue interaction counts. Protein-Ligand Interaction Profiler (PLIP) was used to determine the residues interacting with the ligand for each frame and it covers several interaction types, including hydrogen bonds, hydrophobic contacts, pi interactions, salt bridges, etc.[116] PLIP functions by detecting interactions on single atom level by following a set of geometric rules, similar to Arpeggio used in Section 3.5. The results are presented as aggregated counts across the whole 37500 frames for each system.

4.6.1 EGCG

In the wild type TnC, EGCG interacted across the N-terminal, linker, and C-terminal domain in the uP state (Figure 35). The interaction counts were lower in the phosphorylated state and mostly localised around NcTnC helix A. This finding is reinforced by the TnI interaction pattern where EGCG interacted with both terminal segments of TnI as they joined at the tip of the regulatory head. The interaction pattern in the TnI unit is largely separated by phosphorylation states. In uP, EGCG interacted with NcTnI, TnI_{51-71} in conjunction with CcTnC, and the inhibitory peptide. In particular, the interactions with TnI_{51-71} directly caused the increase in TnI_{34-71} binding energy. In SEP, the interactions were mainly focused on the IT arm region near the U-turn and just before the inhibitory peptide. The patterns in TnT reflected the EGCG interactions with the IT arm where it was seen near the back of the hinge region in uP and the tail of the IT arm in SEP.

In the mutant, EGCG presented more interactions with TnC in the phosphorylated state than

the unphosphorylated state (Figure 36). The interactions in the uP state were mainly located around EF-hand I. In the SEP state, the interactions were similarly positioned around helix B but with doubled counts with additional interactions around helix N and the flexible linker region. In both phosphorylation states, EGCG interacted with both ends of TnI. This represents a repositioning of TnI in relationship to NcTnC upon phosphorylation as the change in the TnI terminal interactions was not as drastic as the change in NcTnC interactions. Another interesting observation is the interactions with NcTnI near the phosphorylation sites (residues 20-40). In the wild type, phosphorylation greatly reduced these interactions while they are exclusively seen in the phosphorylated state of the mutant.

Overall, EGCG presented a varied picture where the mutant binding interactions were more concentrated near NcTnC while the wild type saw a bigger spread across TnC.

4.6.2 RVL

In the RVL systems, the interaction counts were lower in comparison to other ligand systems. This is due to the flat geometry and the relatively smaller size of RVL which can be expected to form less complex interactions. In WT TnC, RVL formed interactions across the board in uP while phosphorylation greatly subsidised the interactions at CcTnC (Figure 37). In WT TnI, interactions can be seen at both ends of TnI which correspond to ligand binding with NcTnC. The uP state had more interactions between the ligand and TnI_{41-51} while SEP interacted more with the IT arm U-turn region. In the mutant, majority of the residues RVL interacted with were part of helix C, E, and H in the unphosphorylated state (Figure 38). The TnC interactions were much reduced in the SEP state. A similar trend can be seen in the TnI interactions where the interactions with TnI_{51-61} and the inhibitory peptide in uP greatly subsidised and in turn RVL interacted more with the IT arm.

4.6.3 Silybin A

In the wild type, silybin A interacted with NcTnC and the linker region in uP and these interactions were greatly subsidised in SEP (Figure 39). Silybin A also interacted with CcTnC

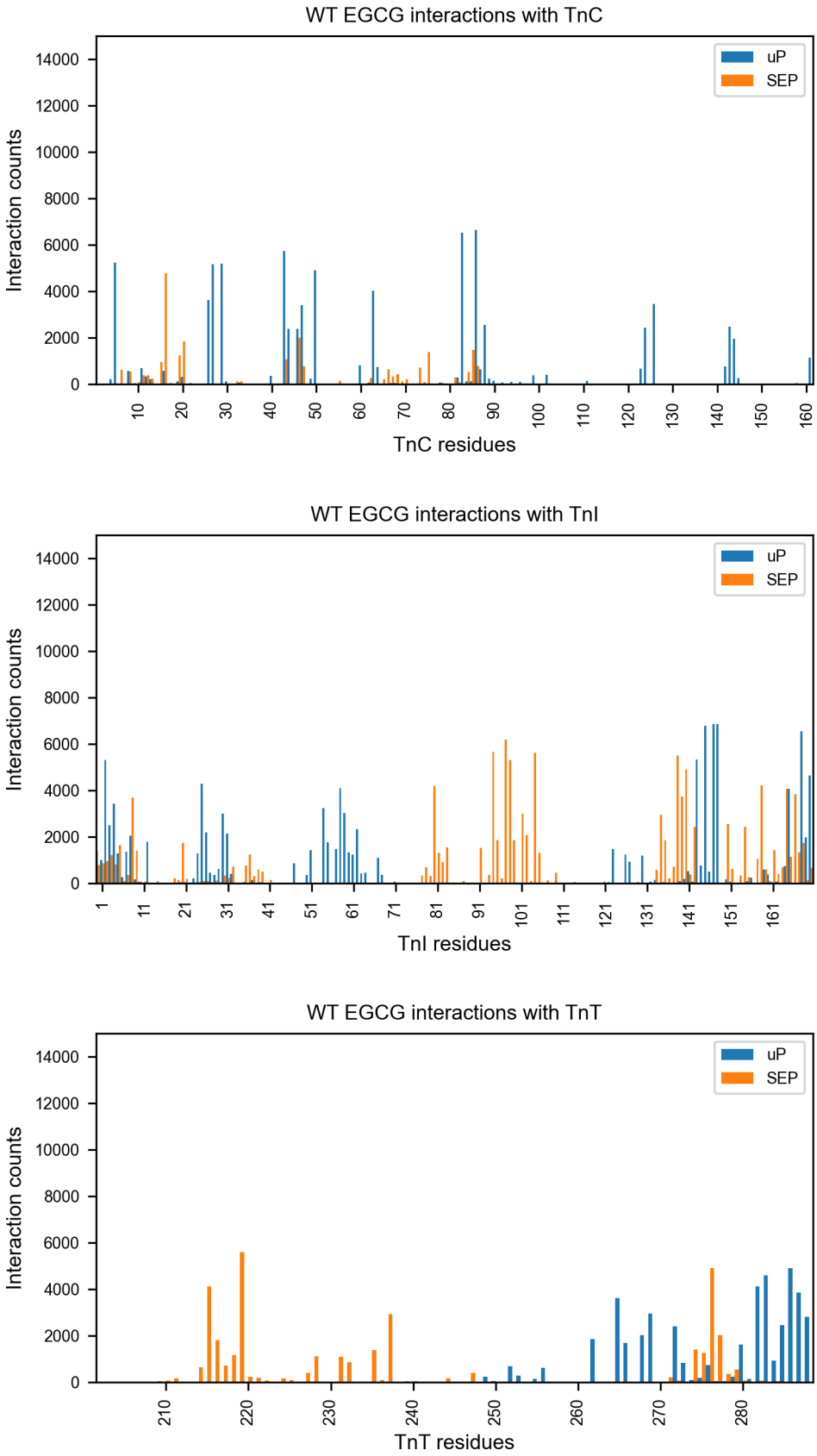


Figure 35: Interaction counts of EGCG with residues in WT systems.

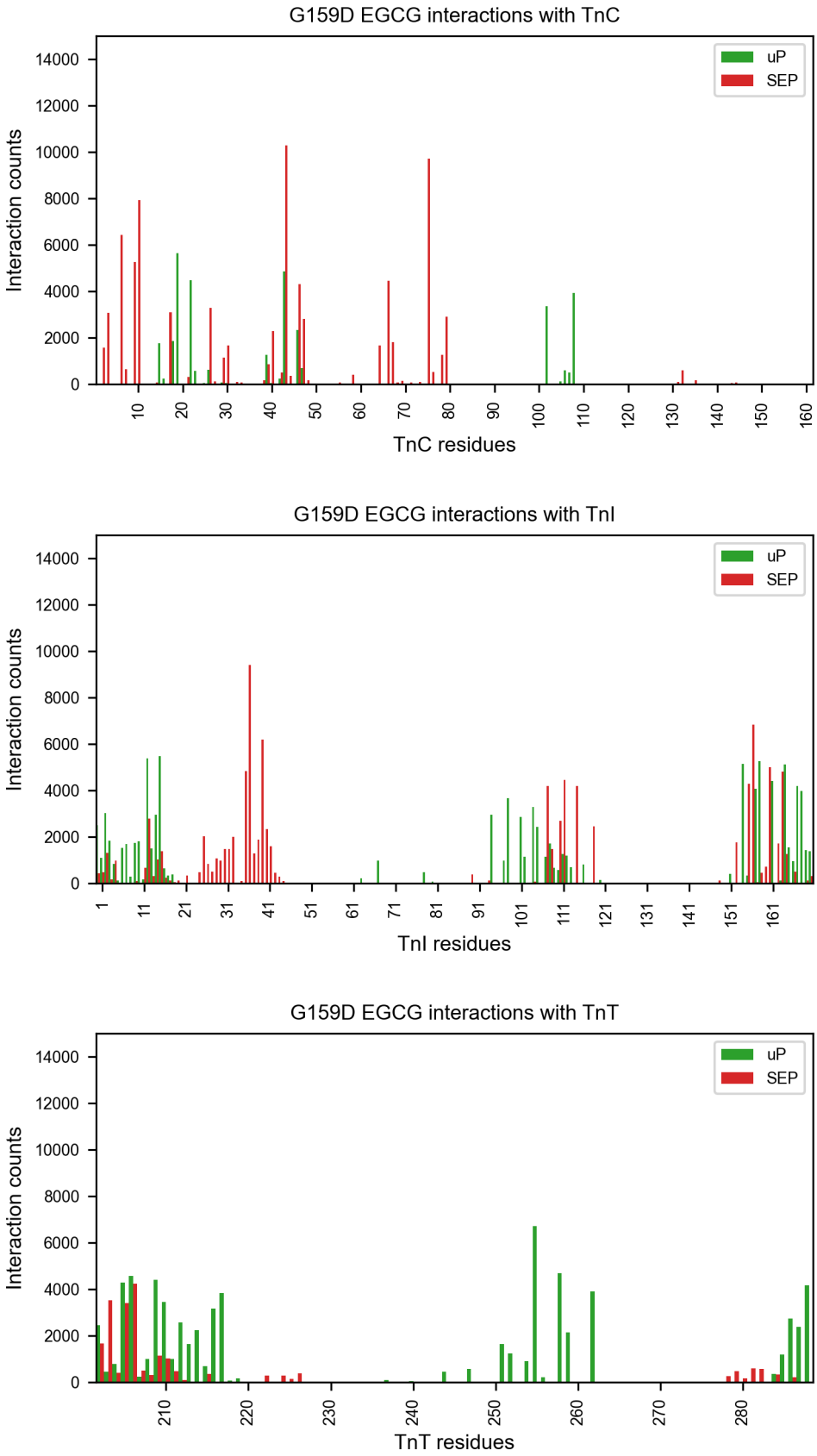


Figure 36: Interaction counts of EGCG with residues in G159D systems.

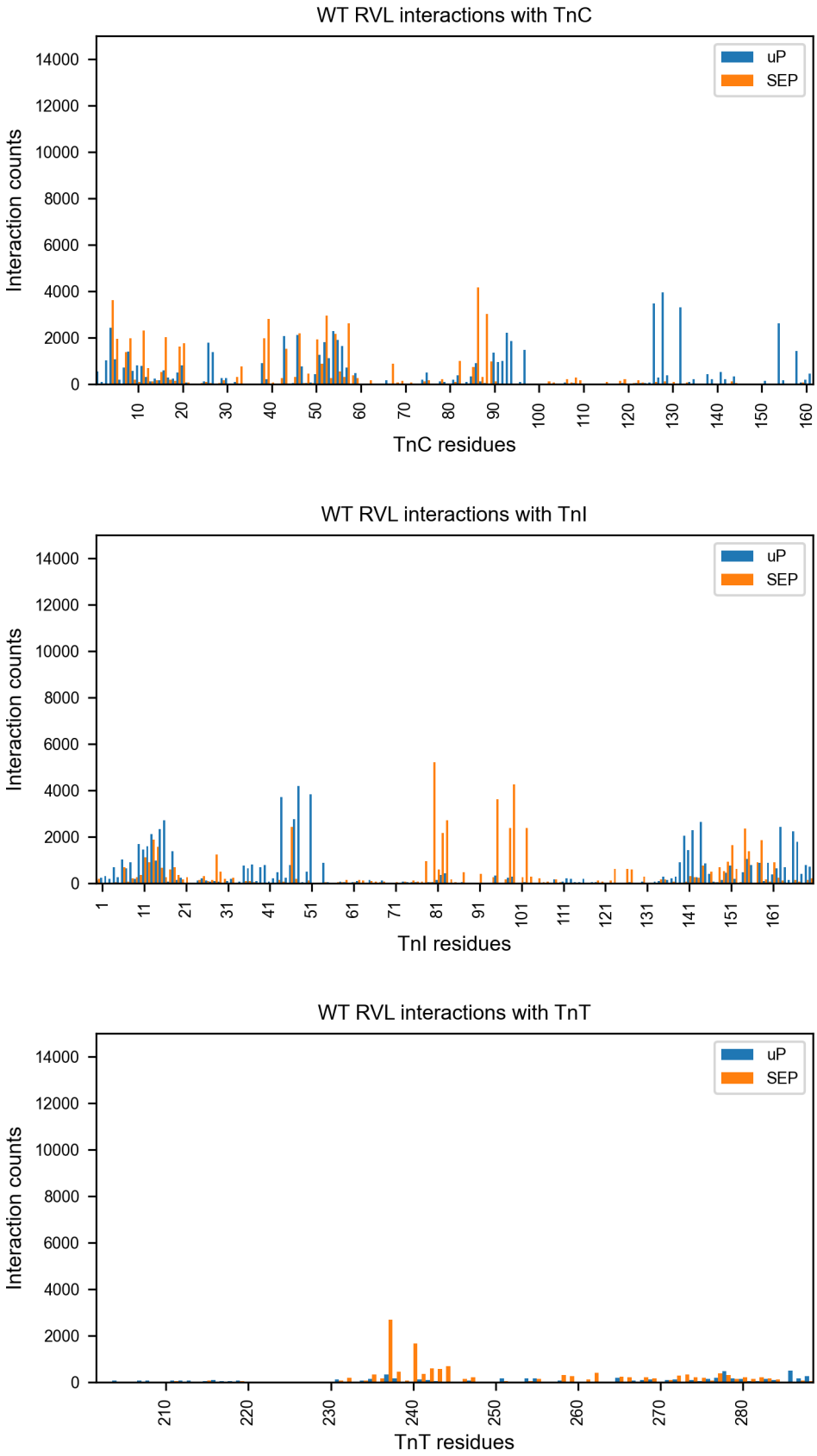


Figure 37: Interaction counts of RVL with residues in WT systems.

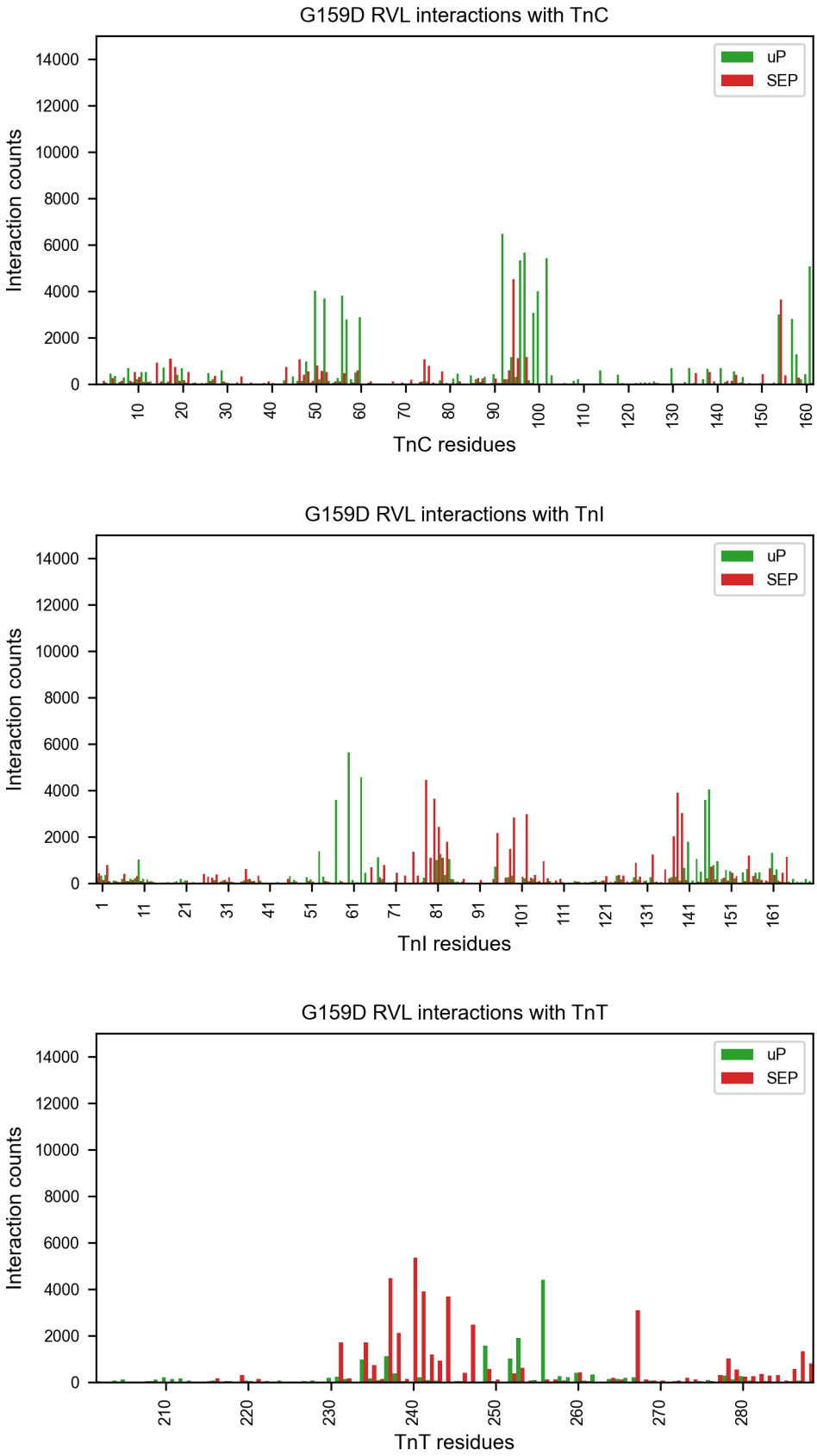


Figure 38: Interaction counts of RVL with residues in G159D systems.

with low counts in SEP. In TnI, a polarising distinction can be seen between the two phosphorylation states. Silybin A interacted more with the inhibitory and switch peptide regions in conjunction with NcTnC in the uP state. In SEP, silybin A more interacted with NcTnI from the back of the hinge as evident from the TnT interactions as well. This pose likely caused the drastic increase in the NcTnI binding energy. In the mutant, silybin A interacted with both EF-hand I and IV of TnC in uP (Figure 40). In G159D SEP, silybin A interactions shifted towards regions between EF-hand III and IV which also involved TnI_{41-61} and $TnT_{202-220}$. In both phosphorylation states silybin A interacted with NcTnI while the uP state saw more interactions in the switch peptide region. Silybin A also interacted with the TnI U-turn and CcTnT in the uP state.

4.6.4 Silybin B

In the wild type TnC, majority of the silybin B interactions were around EF-hand I for the uP state and around helix C and E for the SEP state (Figure 41). These are correlated with extensive interactions with the switch peptide in uP and the inhibitory peptide in SEP. Silybin B had more spread out interactions in the NcTnI region in uP while more interactions were seen around TnI Pro11 and Arg12 in SEP. In WT SEP TnI, the extensive interactions with TnI_{51-71} was in connection with the CcTnC interaction. In the mutant, NcTnC formed frequent interactions with silybin B in uP, larged from helix A, B, and C (Figure 42). Upon phosphorylation, the interactions shifted towards helix C and the region between EF-hand III and IV. The interaction pattern in uP TnC was seen along the interactions with NcTnI and the inhibitory and switch peptide. The NcTnI interactions with silybin B were largely formed when the ligand was at the back of the hinge region.

4.7 Ligand Binding Energy

The ligand binding was assessed with binding energy calculations. The MMPBSA methodology was used to estimate the binding affinity of ligands with the troponin complex over the course of the whole simulation. Multiple binding and unbinding events were observed in each trajectory. The ligand was also seen changing binding regions within a single simula-

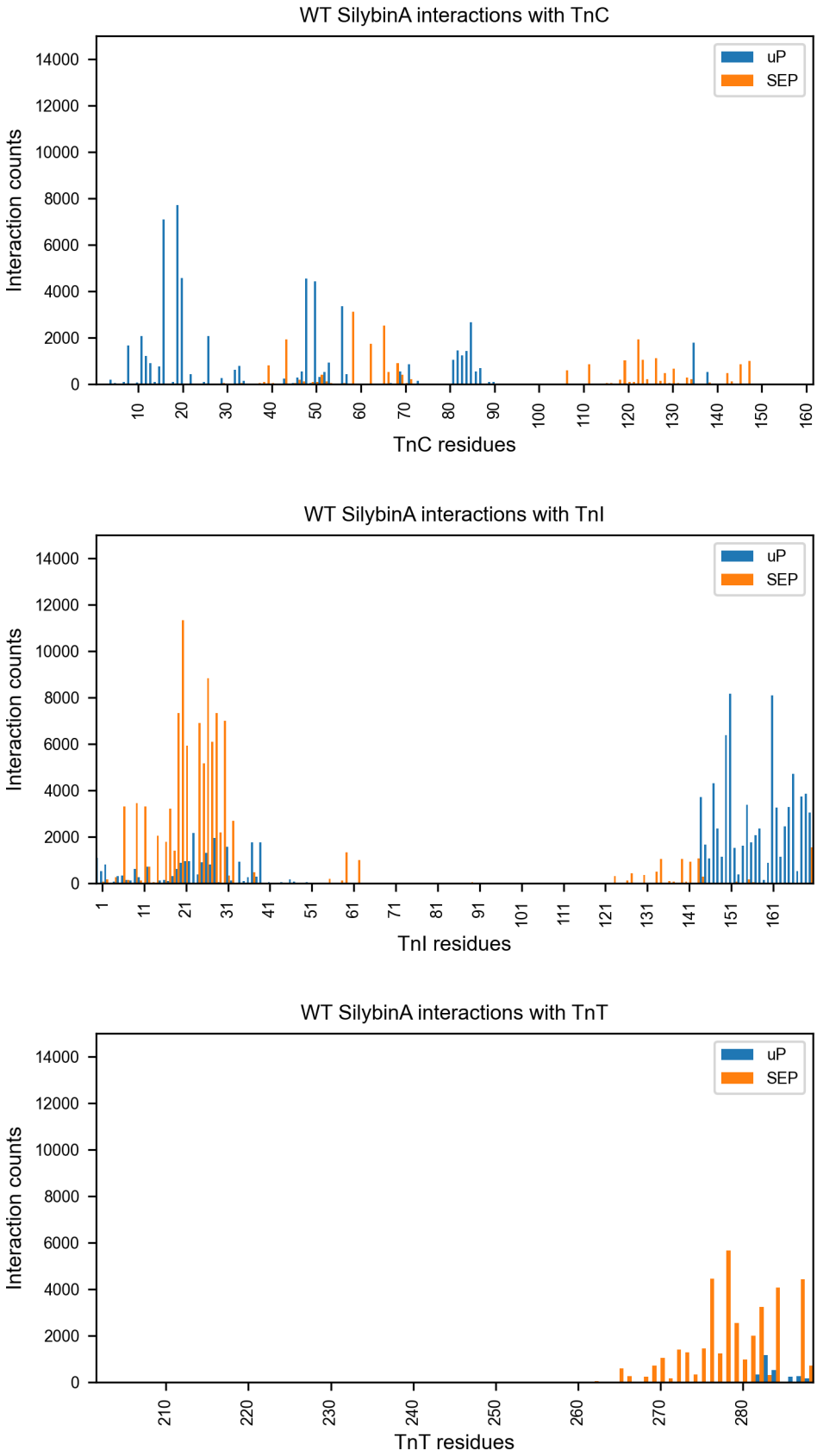


Figure 39: Interaction counts of silybin A with residues in WT systems.

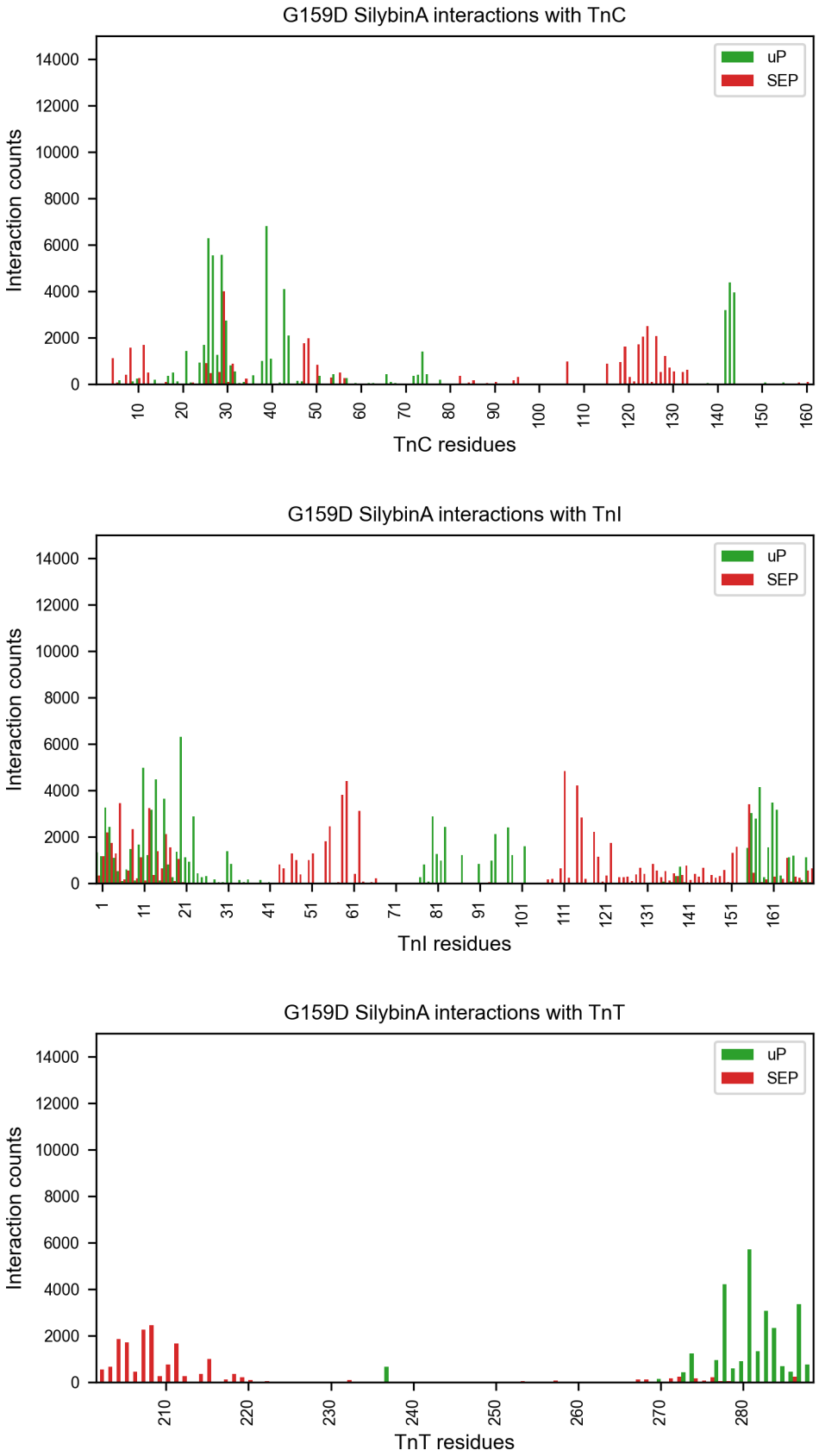


Figure 40: Interaction counts of silybin A with residues in G159D systems.

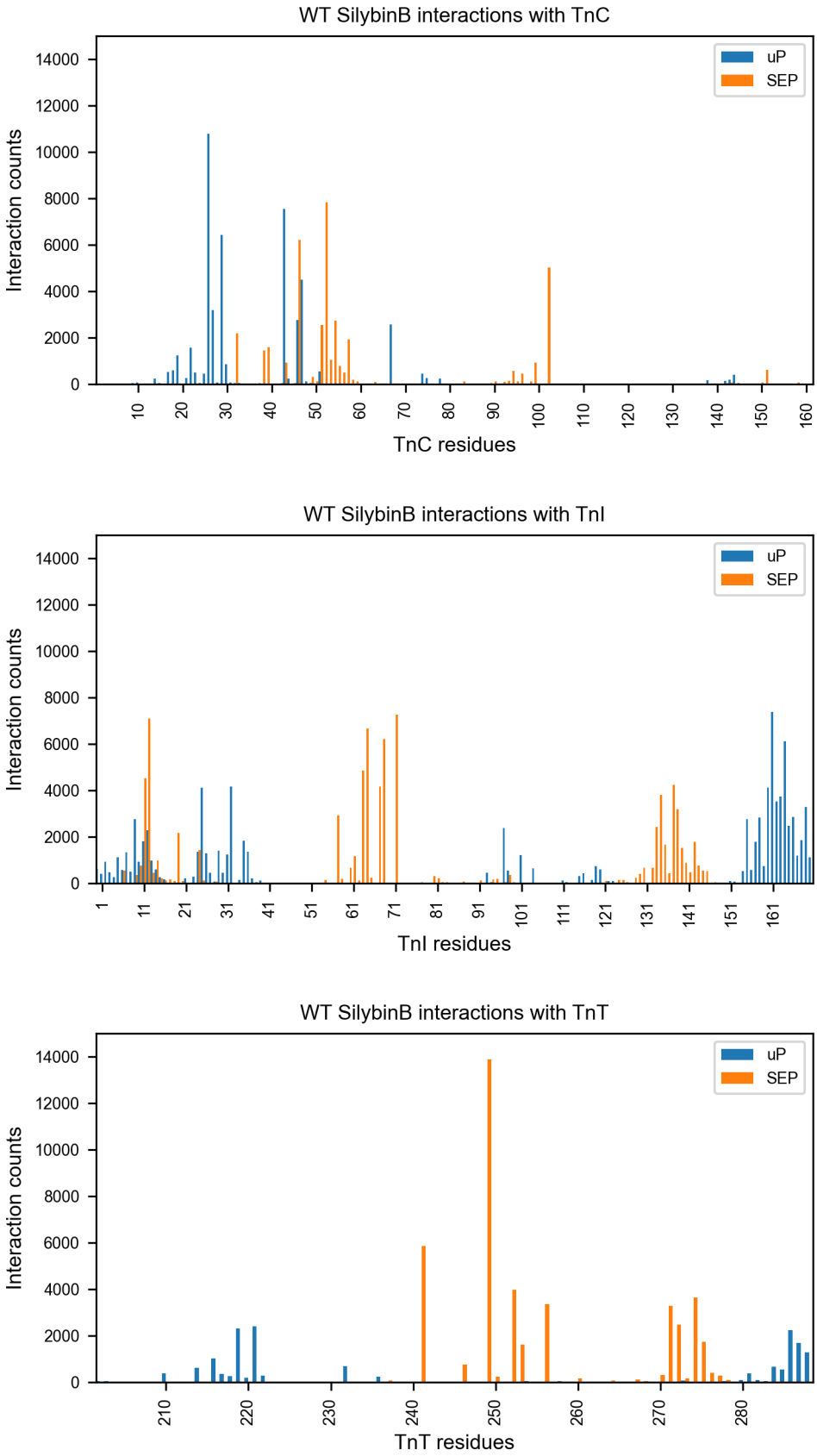


Figure 41: Interaction counts of silybin B with residues in WT systems.

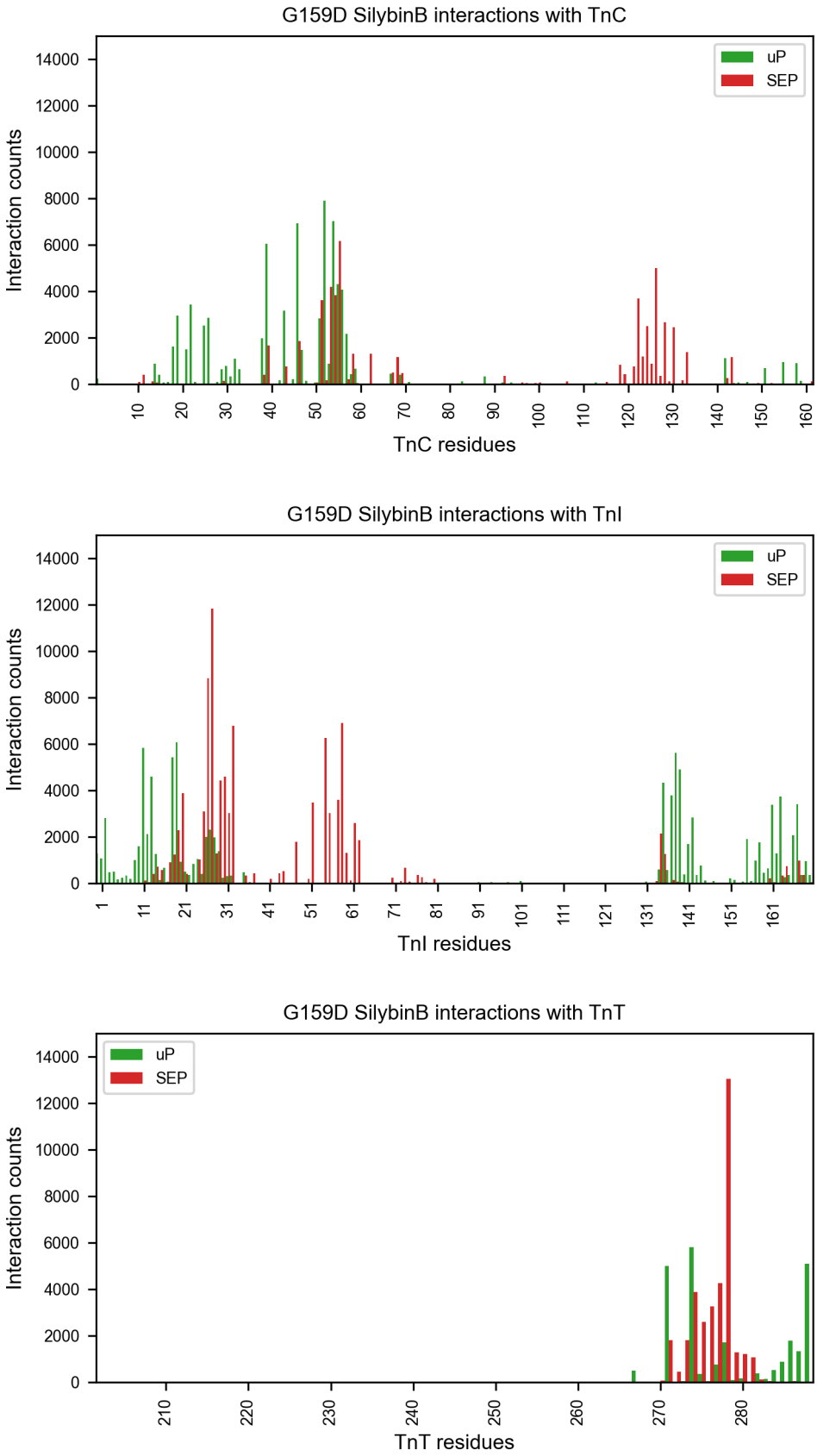


Figure 42: Interaction counts of silybin B with residues in G159D systems.

tion. The mean binding energy is summarised in Table 35 and aims to provide a broad view into how strongly the ligands interacted with the protein.

	WT uP	WT SEP	G159D uP	G159D SEP
EGCG	-30.4	-22.7	-26.1	-25.4
RVL	-13.8	-12.8	-13.7	-12.4
Silybin A	-23.2	-22.4	-25.5	-19.1
Silybin B	-21.7	-23.0	-26.2	-26.6

Table 35: Mean ligand binding energy in troponin systems with EGCG, RVL, silybin A, and silybin B. Unit: $\mathrm{kcal\,mol^{-1}}$

Over all systems, the binding energy is influenced by the ligand's size. RVL had the lowest binding energy while EGCG, silybin A, and silyin B largely have similar binding energies. RVL has the smallest size and a flat geometry which would not help in forming complex 3D binding interactions. EGCG binds to WT SEP significantly weaker than WT uP by almost $8\,\mathrm{kcal\,mol^{-1}}$ while the binding energy in G159D SEP is on par with G159D uP. Another ligand with significant phosphorylation modulated binding energy was DHSA in G159D where the ligand binding energy was reduced by $6.4\,\mathrm{kcal\,mol^{-1}}$ upon phosphorylation. RVL presented similar binding energies in all systems. Comparing silybin A with silybin B, the binding energies were similar in all systems except in G159D SEP where silybin A interacted weaker than silybin B with a $7.5\,\mathrm{kcal\,mol^{-1}}$ difference.

The time averaged binding energy also included the poses where the ligands were dissociated from the protein. To further compare the interaction energy of ligands at similar binding regions, the mean binding energy for each ligand cluster, as described in Section 4.5, was calculated (Table 36).

Across different ligand systems, the strongest binding cluster tends to be located on the regulatory head. In EGCG systems, the ligand cluster with the strongest binding energy was located at EF-hand I where the ligand also interacted with both ends of TnI. In this pose, EGCG was fully enclosed by TnC and TnI and deeply inserted into the hydrophobic patch of the EF-hand. The second strongest binding cluster was located at the same place too. The other

	Pink	Orange	Yellow	Cyan
WT uP EGCG	-32.6	-21.6	-44.8	
WT uP RVL	-15.6	-11.7		
WT uP SilybinA	-23.8			
WT uP SilybinB	-23.3	-20.6	-17.1	
WT SEP EGCG	-26.5	-22.3	-24.3	-19.6
WT SEP RVL	-14.7	-12.7		
WT SEP SilybinA	-25.3	-19.1		
WT SEP SilybinB	-24.4	-17.0		
G159D uP EGCG	-22.5	-33.8		
G159D uP RVL	-16.4	-13.4	-16.8	
G159D uP SilybinA	-24.6	-33.0	-23.1	
G159D uP SilybinB	-26.8			
G159D SEP EGCG	-25.1	-31.7	-19.7	
G159D SEP RVL	-14.3			
G159D SEP SilybinA	-20.7	-21.3		
G159D SEP SilybinB	-31.8	-21.5	-14.4	

Table 36: Ligand binding energy in troponin systems with EGCG, RVL, silybin A, and silybin B separated by clusters. Unit: $\mathrm{kcal\,mol^{-1}}$

strong binding clusters were mostly located near NcTnC as well. For RVL systems, multiple similar strong binding clusters were located at the TnC domain interface. Due to RVL's flat geometry and small size, it can insert into the hinge region hence be buried in the neighbouring inhibitory peptide and the TnC linker more easily. In silybin A, the strongest binding cluster was cluster 2 of G159D uP system. Silybin A was in a similar position as the strongest binding pose of EGCG where it interacted with both TnC's EF-hand I and ends of TnI. Notably, however, this was not the most populated cluster of the system likely due to the mobility of the terminal regions of TnI which meant the pose was not the most stable. In silybin B systems, the strongest binding cluster (G159D SEP cluster 1) had the ligand positioned at the back of the hinge region where the ligand interacted with TnC helix C, NcTnI, as well as CcTnT. This represents a departure from the observation in the EGCG and silybin A systems where the

cluster at the tip of the regulatory head had the strongest binding interactions. The presence of this binding site is also the reason silybin B binds to G159D SEP stronger than silybin A. Cluster 1 of G159D uP silybin B was situated at the tip of the regulatory head and represented the second strongest binding cluster in silyin B systems. Cluster 1 of WT SEP silybin B also saw the ligand half inserted into the hinge region similar to the RVL system, which was another strong binding cluster of the ligand. In all ligand systems, bindings towards the IT arm tend to be relatively weak except when the ligand was enclosed by the N-terminal end of TnT, such as in G159D SEP EGCG cluster 2.

4.8 Discussion

4.8.1 Recouplers Restored Key Metrics

Both the structural measurements (A-B angle, hinge angle, TnC domain distance) and the energetics calculations (switch peptide, inhibitory peptide, TnI_{34-71}, NcTnI) showed a varied picture in the ligand effect. Broadly, the recouplers restored some mutant metrics to the corresponding wild-type levels and the inactive silybin A left some metrics unchanged. However, more often than not, the mutant metrics were not simply restored to the wild-type values and the wild type can be influenced by ligand binding as well. One hypothesis is that as long as the relative phosphorylation induced changes are maintained in the same direction as the healthy troponin and a differential in the calcium affinity exists, thin filament and cardiac muscles are able to adapt to the changed physiology and maintain the heart's normal functionality. Figure 43 summarises all the measured metrics and illustrates the difference in key metrics between the phosphorylated and unphosphorylated states as bar plots.

EGCG restored the changes in A-B angle metric, inhibitory peptide and NcTnI binding energy to the wild type direction in the mutant while hinge angle and TnC domain distance were not fixed. Notably, EGCG also disrupted the normal behaviour of hinge angle in the wild type. This might be related to EGCG's dual function as recoupler and desensitiser. RVL restored all erroneous metrics in the mutant. However, it also shifted the change in TnC domain distance into the opposite direction in the wild type. Silybin B restored all erroneous metrics in mutant

but also eliminated the differences in the switch peptide binding energy in G159D and the A-B angle metric in WT. Lastly, silybin A restored the inhibitory peptide and NcTnI binding energy in the mutant but disrupted NcTnI binding energy in the wild type and switch peptide binding energy in the mutant.

Overall, the pattern between recouplers and inactive compound is clear. Despite some disruption to the wild type metrics the recouplers restored majority of the mutant metrics. The pure desensitiser compound silybin A not only was unable to restore the structural metrics but also introduced new defects into both the wild type and the mutant systems. Although it is unclear the relative importance of these metrics or the consequence of disrupting metrics displaying the correct changes, the result has provided evidence for the suitability in using these metrics as preliminary screening measurements.

4.8.2 Common Binding Regions Present in All Simulated Systems

Based on the ligand binding region clusters, common binding regions were found across different ligand systems and ligand interactions happened at the regulatory head for majority of the time. The key binding regions were EF-hand I hydrophobic patch, hinge region with inhibitory peptide and TnC flexible linker, back of the hinge region with NcTnI and EF-hand II, and CcTnC anchoring region with TnI. While some binding regions were found at the IT arm side, they were generally weaker and unlikely when the full TnT subunit is in place.

For EGCG and RVL, the experimental findings were not reproduced.[26,63] Instead of interactions with CcTnC, majority interactions were found in the N-terminal domain. In the NMR experiments, TnC were used in isolation of the N- and C-terminal regions of TnI which were found to be crucial in the ligands binding motif. For many interaction poses, EGCG interactions with NcTnC in concert with N- and C-terminal ends of TnI. This stable interaction would be greatly weakened in absence of TnI which could partially explain the deviation from NMR data. The presence of the full TnI core would also make the full insertion of the ligand into the CcTnC domain much harder. Comparing the diastereomers silybin A and B, differentials in binding regions were seen across the systems. This illustrates the major impact of stereoconfiguration on ligand interaction patterns. Silybin B also showed a stronger selective

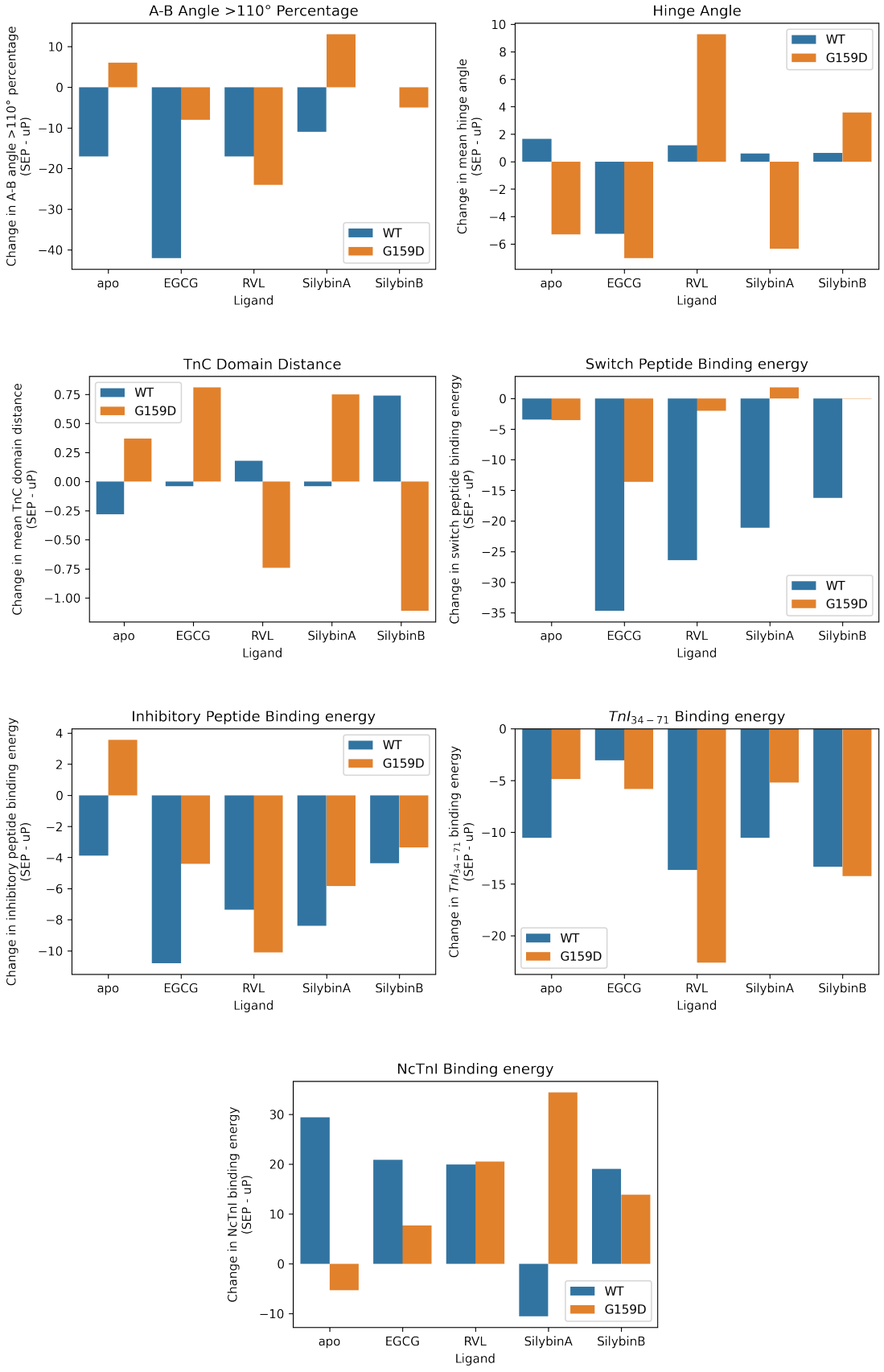


Figure 43: Comparison of changes in key metrics upon phosphorylation in ligand systems with apo systems.

binding towards G159D SEP compared to silybin A which fits the previously proposed functional states where recouplers would selectively binds to the phosphorylated mutant (Figure 9).

Lacking a native binding site, it was difficult to discern exactly which binding region would lead to the desired physiological outcome. No apparent pattern was seen to differentiate the recouplers from the inactive compounds or which binding region would contribute to the desensitisation function. It is also possible that ligand effect can not be explained by a single binding pose. Perhaps the combination of bindings to different regions, such as EF-hand I, hinge region, TnC anchoring region, hence fixing multiple defects explains the ligand's physiological functions.

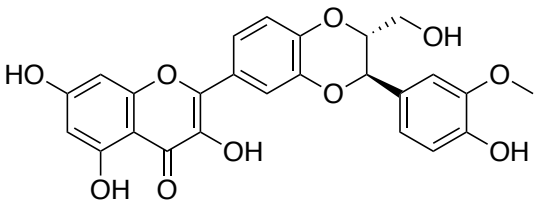
4.8.3 Ligand Binding Disrupts Ionic Interactions from the mutation site

The residue interaction profiling data showed that the significant ionic interaction formed between the mutant residue and TnC Arg83 was disrupted with ligand interactions. This was accompanied with increased interactions with TnC Lys86 and Lys90 and changed interaction patterns with inhibitory peptide arginines. Two ligand binding regions could modulate the ionic interactions directly. When the ligand was inserted into the hinge region which is also the TnC domain interface, it interacts with both the inhibitory peptide and the TnC flexible linker. This would increase the domain separation and disrupt the ionic interactions from the mutant residue. Another potential mode of action is when the ligand interacted with the CcTnC/TnI anchoring region. This would change the dynamics of the C-terminal end of TnC which is where the mutation site is located.

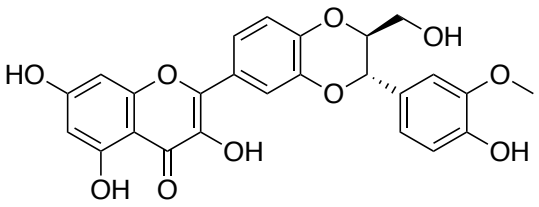
These changes in interactions provide direct explanations for the observed effects of ligands on TnC domain distance and inhibitory peptide binding energy. When NcTnC and CcTnC are relieved from the locked position due to ionic interactions from TnC Asp159, the hinge angle dynamics is influenced as well. Therefore, the modulation on the mutation residue interactions could explain many of the observed changes in the measured metrics.

5 Study of ECG, DHSA, DHSB, and r07

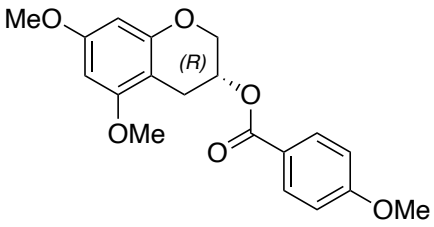
This chapter builds on the previous discussions on the apo and small molecule studies. I further expanded the scope of ligand simulations to include five additional molecules that possess interesting *in vitro* assay results. These compounds are analogues of EGCG, silybin A, and silybin B yet represent diverse physiological responses. This set of compounds (Figure 44) include 2,3-dehydrosilybin A and B (DHSA and DHSB), r07, and epicatechin gallate (ECG).



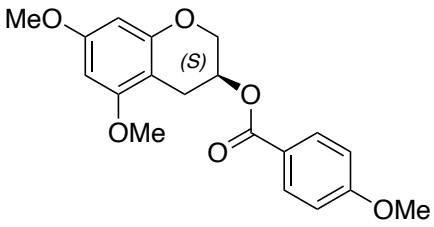
2,3-dehydrosilybin A



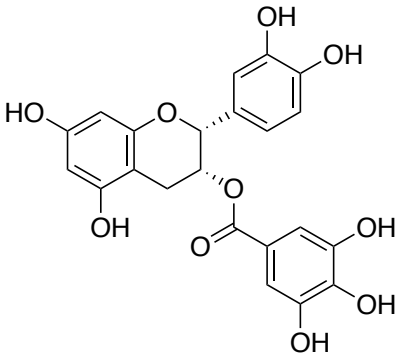
2,3-dehydrosilybin B



r07R



r07S



epicatechin gallate

Figure 44: Chemical structure of DHSA, DHSB, r07R, r07S, and ECG.

2,3-dehydrosilybin A and B (DHSA and DHSB) are the oxidative derivatives of silybin A and

B. Both enantiomers were known for their antioxidant acitivities.[143] Single Ca^{2+} concentration screening has shown that DHSB is a pure recoupler while DHSA is inactive.[67] Both compounds were also tested against the G159D mutant and DHSB recoupled the mutant while DHSA introduced minimal difference in the calcium sensitivity. It is of interest to see if DHSB produces similar structural and energetics changes as silybin B since both are recouplers with minimal structural differences. In the case of comparing DHSA to silybin A, we want to see if the difference between an inactive compound and desensitiser can be discerned.

As part of the compounds screened by Sheehan *et al.*, r07 was found as the only reverse-recoupler in that it caused the calcium sensitivity in the phosphorylated state to be higher than the unphosphorylated state.[67] This unique behaviour was fascinating since the other structurally similar Hsp90 inhibitor analogues were either recouplers or inactive. The synthetic route of r07 resulted in a racemic mixture (named r07R and r07S) of the compound and both enantiomers were present when r07 was tested with troponin mutants.[66] The isomers could have different binding patttern hence affect the troponin's dynamics differently, as evidented by the silybin A/B and DHSA/DHSB pairs. For this reason, both isomers were simulated in this work.

ECG is a pure desensitiser. It was the only analogue of EGCG that did not show a recoupling property. When tested against the uncoupling TPM1 E180G mutant that exihibits a HCM phenotype, ECG reduced the Ca^{2+} sensitivity of both phosphorylated and unphosphorylated thin filaments equally so the system remained uncoupled.[67] ECG only differs from EGCG by missing one hydroxyl group. We want to see if this difference can be captured by ligand binding regions or structural features.

This chapter follows the same layout as Section 4. The same MD simulation protcol as the previous chapter was employed here. The key structural metrics, peptide binding energies, ligand binding regions, and residue interaction changes are briefly described without reiterating their significance. The potential connections between ligand binding and changes in protein dynamics are then discussed in the context of the small molecules' physiological effects.

5.1 Global Structure Changes from Principal Component Analysis

Figure 45 illustrates the PCA projections of the simulation trajectories containing DHSA, DHSB, r07R, r07S, and ECG onto the same principal component space as the apo studies. Similar to the previously described ligand systems, these five small molecules shifted the troponin conformations towards the low PC1 and low PC2 direction. The projections of the simulations with DHSA/B, r07R/S, and ECG all showed clear separation from the apo systems. Within each ligand systems, different phosphorylation states of WT and G159D are mostly mixed together and occupy the same space. The G159D SEP state in each ligand systems displayed higher tendency to wander into the apo space which was seen in the previous group of small molecules. The tendency for G159D SEP system to lean towards high PC1 and high PC2 direction is also the trend observed in the apo systems which means some of the global conformational characteristics are preserved in the ligand simulations.

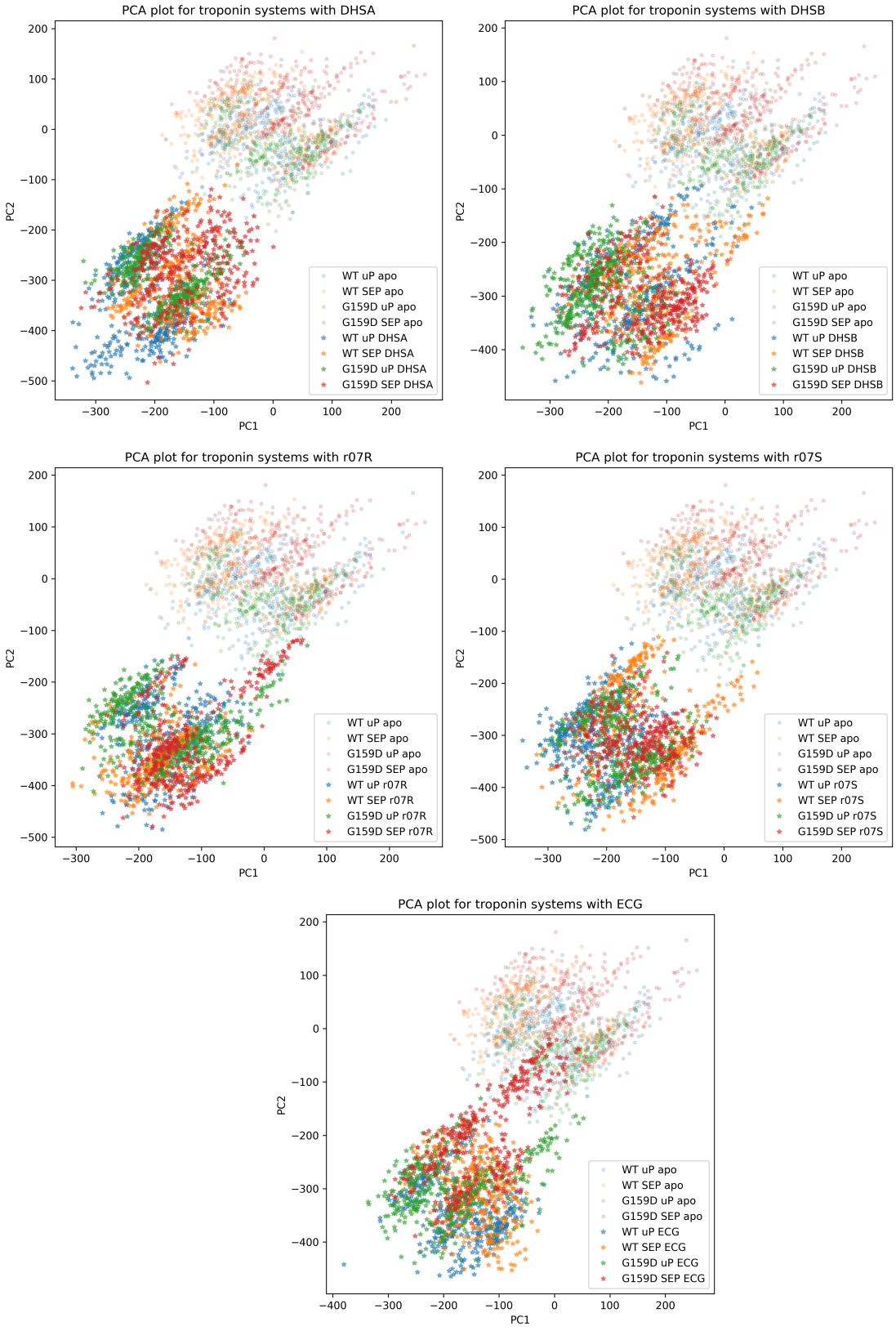


Figure 45: Scatter plot of troponin conformations in the apo systems' first two principal components. The faded circles represent conformations from the apo system. The coloured stars represent the ligand systems.

5.2 Changes in Regional Structural Features

5.2.1 Interhelical A-B Angle

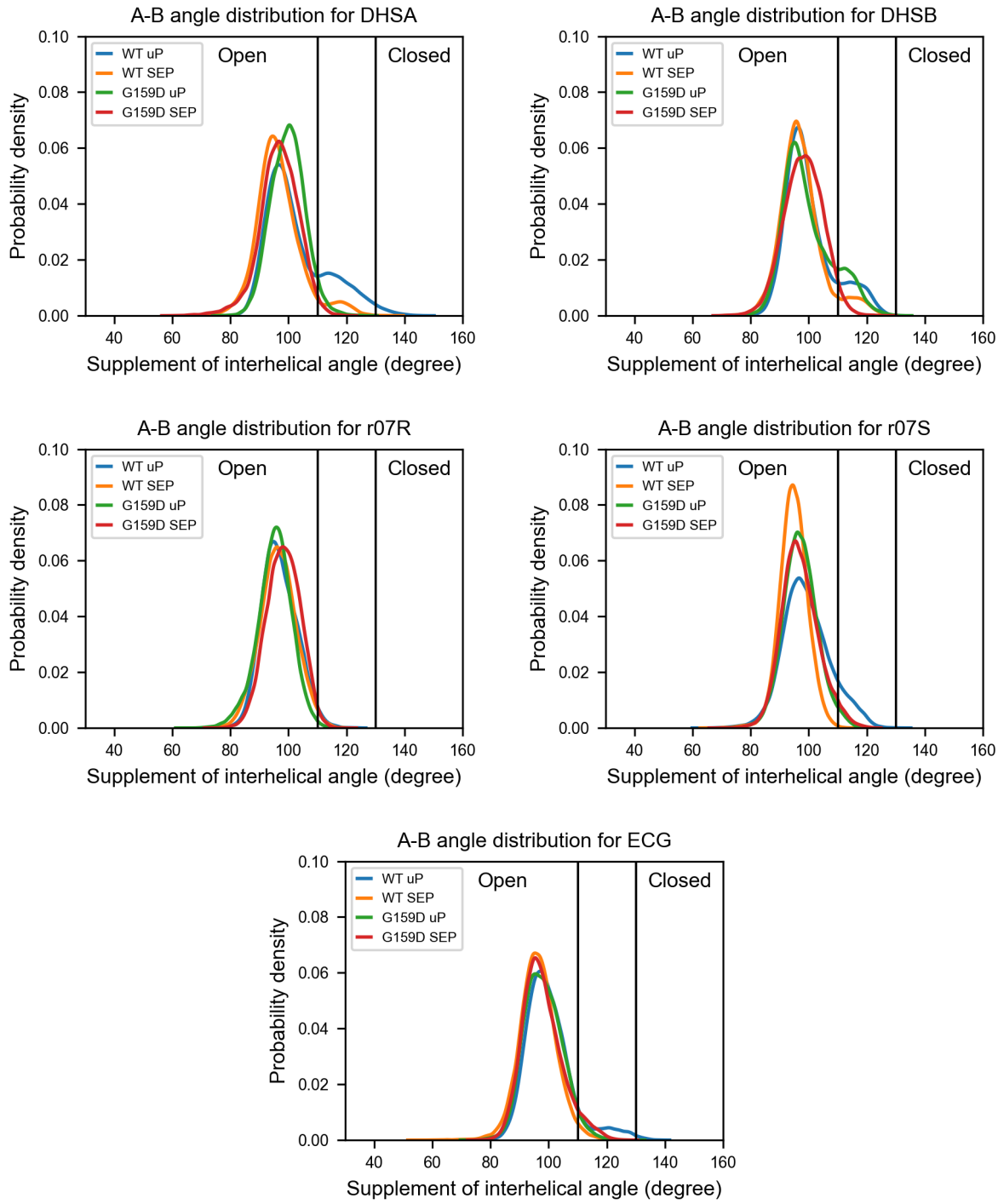


Figure 46: The interhelical angle distributions for DHSA, DHSB, r07R, r07S, and ECG systems. Compare to apo figures in Figure 19.

In WT uP, DHSA shifted the interhelical angle distribution towards the more closed direction and increased the >110° percentage from the apo level, 18%, to 24% (Figure 46 and Table 37). DHSB, r07S, and ECG reduced the percentage to 15%, 11% and 9%, respectively while r07R drastically reduced the percentage to 2%. In the phosphorylated wild type, DHSA and DHSB increased the A-B angle >110° percentage to 6% and 8%, respectively. Both r07R and ECG reduced the percentage to 2% while r07S shifted all the density into the open region. Comparing the two phosphorylation states, the ordinary reduction of the A-B angle >110° percentage in WT apo systems were maintained in DHSA, DHSB, r07S, and ECG. No change in the percentage was seen in r07R systems upon phosphorylation. In G159D uP, DHSA, r07R, r07S, and ECG, all changed the >110° percentage slightly from 1% in G159D uP apo to be within 0-4%. DHSB, on the other hand, greatly increased the percentage to be 15% which is more akin to the WT uP apo level. In G159D SEP, all ligand systems reduced the percentage to be between 1-6% from G159D SEP apo's 7%. DHSA changed the phosphorylation induced change to be a reduction in the angle percentage with a small difference of 2%. r07R, r07S, and ECG increased the percentage slightly upon phosphorylation. DHSB changed the difference in the angle >110° percentage between G159D SEP and G159D uP to resemble the WT apo systems the most which is also the only recoupler in this group of small molecules.

	Peak	Mean	Std. Dev.	>110°(%)	FWHM
WT uP apo	96.3°	101.8°	9.2°	18.0°	17.1°
WT uP DHSA	96.8°	103.0°	10.9°	24.0°	13.1°
WT uP DHSB	96.3°	100.1°	8.5°	15.0°	11.4°
WT uP r07R	94.9°	97.0°	6.0°	2.0°	14.6°
WT uP r07S	96.6°	99.3°	8.2°	11.0°	16.5°
WT uP ECG	97.2°	100.0°	8.2°	9.0°	15.0°
WT SEP apo	95.0°	96.2°	5.7°	1.0°	12.2°
WT SEP DHSA	94.6°	96.7°	7.9°	6.0°	13.2°
WT SEP DHSB	95.7°	97.6°	7.6°	8.0°	12.0°
WT SEP r07R	96.0°	96.4°	6.2°	2.0°	13.9°
WT SEP r07S	94.5°	94.9°	4.8°	0.0°	10.4°
WT SEP ECG	95.3°	96.4°	6.4°	2.0°	13.3°
G159D uP apo	95.7°	97.1°	5.5°	1.0°	12.5°
G159D uP DHSA	100.4°	99.8°	5.9°	4.0°	14.0°
G159D uP DHSB	95.1°	99.7°	8.5°	15.0°	12.0°
G159D uP r07R	95.9°	95.2°	6.0°	0.0°	12.6°
G159D uP r07S	96.1°	97.1°	6.0°	2.0°	13.0°
G159D uP ECG	95.4°	98.1°	6.4°	4.0°	16.1°
G159D SEP apo	95.4°	99.0°	6.7°	7.0°	13.6°
G159D SEP DHSA	96.7°	96.8°	6.8°	2.0°	14.8°
G159D SEP DHSB	98.9°	98.2°	6.7°	3.0°	16.6°
G159D SEP r07R	98.1°	98.2°	5.8°	1.0°	15.0°
G159D SEP r07S	95.4°	97.1°	6.4°	4.0°	13.6°
G159D SEP ECG	95.3°	97.9°	6.9°	6.0°	13.5°

Table 37: The interhelical angle metrics for the systems with DHSA, DHSB, r07R, r07S, and ECG.

5.2.2 Interdomain Hinge Angle

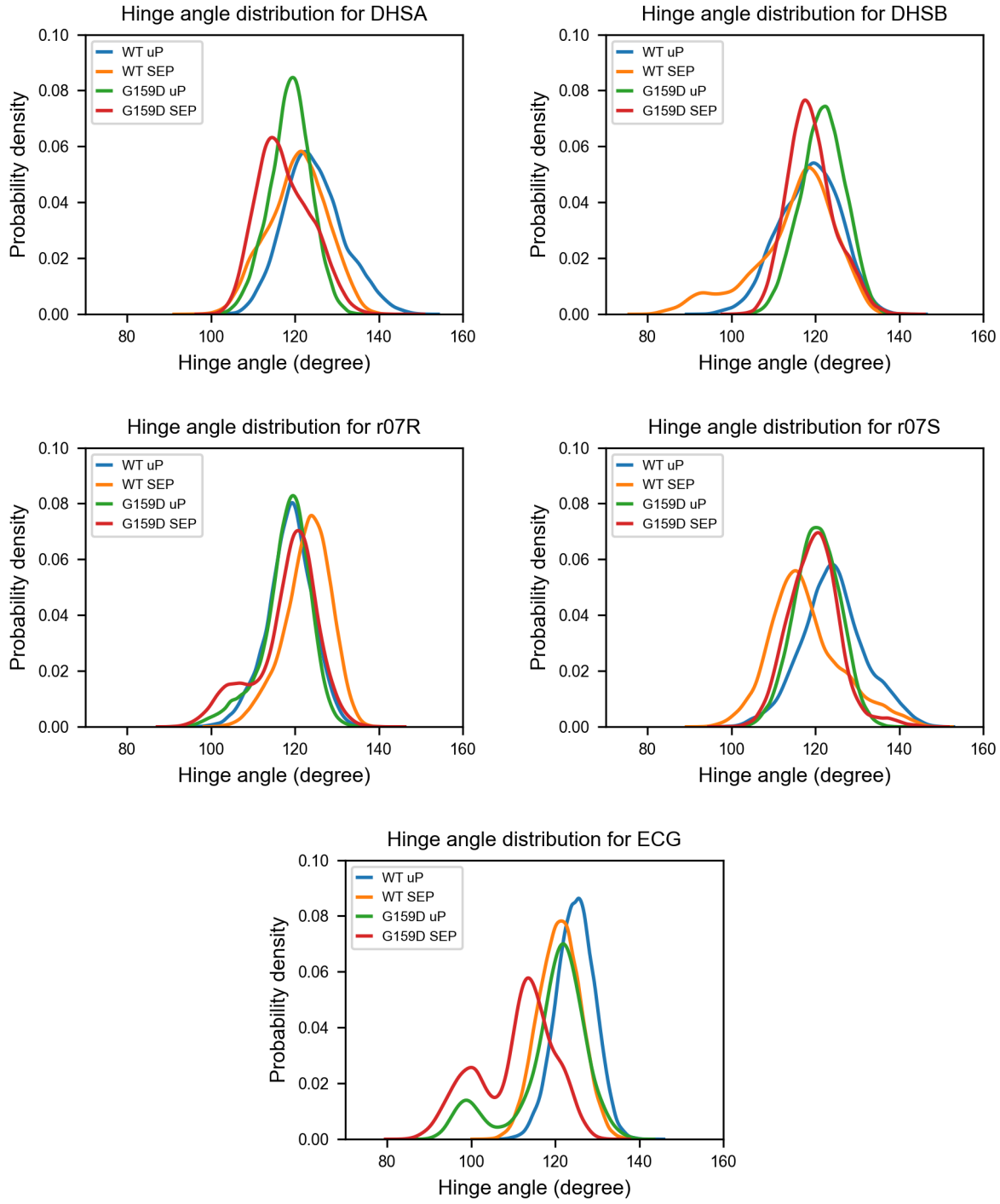


Figure 47: The interdomain angle distributions for DHSA, DHSB, r07R, r07S, and ECG systems. Compare to apo figures in Figure 17.

In WT uP, DHSA, r07S, and ECG increased the mean hinge angle to around 124° from the apo level of 121.6° (Figure 47 and Table 38). Both DHSB and r07R reduced the mean to around 118°. In WT SEP, all ligands caused reductions in the mean hinge angle from the apo level which means a shift towards the more L shaped conformation. Comparing WT SEP with WT uP, all ligands seem to have disrupted the normal shift in the mean hinge angle except r07R, which further increased the increase of the mean angle upon phosphorylation. In G159D uP, DHSA, r07R, r07S, and ECG all reduced the mean hinge angle from the apo level while DHSB did not move the mean significantly. In G159D SEP, DHSA, DHSB, r07R, r07S all increased the mean hinge angle slightly, from 116.2° in apo to be between 117.5° and 119.7°. ECG, in contrast, greatly reduced the hinge angle mean to 110.6°. Phosphorylation in the mutant apo system introduced a decrease in the hinge angle which none of the ligands were able to correct. DHSA, DHSB, and r07S were able to reduce the difference while ECG exacerbated the decrease. r07R, the only ligand that maintained the apo changes in the WT systems, moved the mean hinge angle of both phosphorylation states to be the same.

	Peak	Mean	Std. Dev.	FWHM
WT uP apo	122.9°	121.6°	7.2°	22.0°
WT uP DHSA	122.4°	124.5°	7.2°	15.3°
WT uP DHSB	119.6°	118.6°	7.2°	18.5°
WT uP r07R	119.3°	118.7°	5.5°	11.2°
WT uP r07S	124.0°	124.3°	7.7°	14.6°
WT uP ECG	125.6°	124.7°	4.5°	11.0°
WT SEP apo	122.5°	123.3°	6.0°	16.4°
WT SEP DHSA	121.4°	120.5°	7.0°	16.2°
WT SEP DHSB	118.4°	115.0°	9.8°	15.1°
WT SEP r07R	123.9°	122.9°	5.6°	11.9°
WT SEP r07S	115.2°	117.8°	8.7°	14.1°
WT SEP ECG	121.4°	121.0°	4.9°	12.3°
G159D uP apo	121.7°	121.5°	4.8°	11.5°
G159D uP DHSA	119.5°	118.8°	4.9°	10.7°
G159D uP DHSB	122.3°	121.7°	5.3°	12.8°
G159D uP r07R	119.5°	117.9°	5.9°	10.2°
G159D uP r07S	120.1°	120.4°	5.4°	13.4°
G159D uP ECG	121.9°	118.9°	9.0°	11.1°
G159D SEP apo	117.1°	116.2°	6.0°	12.3°
G159D SEP DHSA	114.5°	117.5°	6.4°	16.3°
G159D SEP DHSB	117.5°	119.0°	5.6°	11.3°
G159D SEP r07R	120.7°	118.0°	7.6°	10.7°
G159D SEP r07S	120.6°	119.7°	6.1°	13.6°
G159D SEP ECG	113.6°	110.6°	9.0°	11.7°

Table 38: The interdomain angle metrics for the systems with DHSA, DHSB, r07R, r07S, and ECG.

5.2.3 TnC Domain Distance

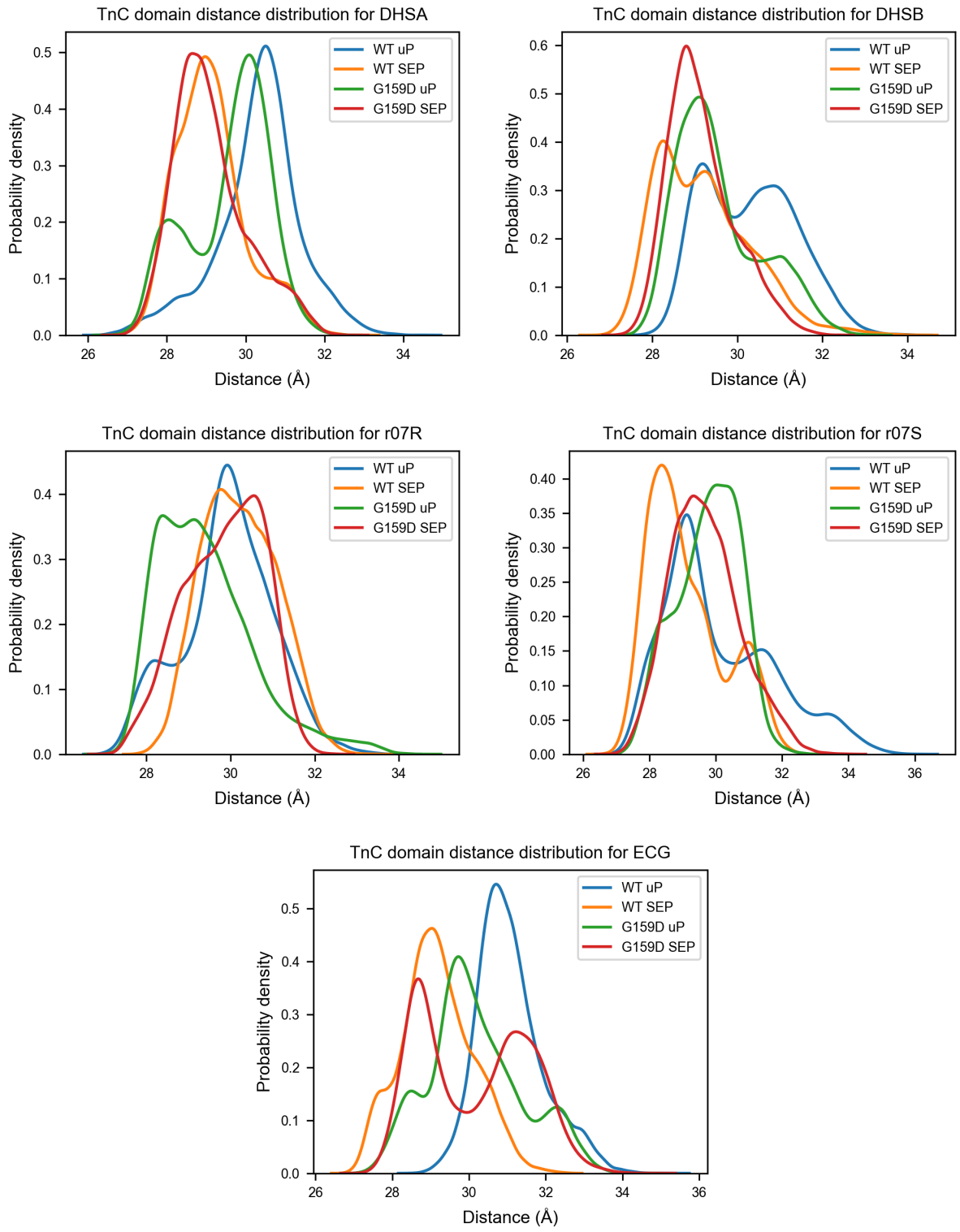


Figure 48: The TnC domain distance distributions for the systems with DHSA, DHSB, r07R, r07S, and ECG. Compare to apo figures in Figure 18.

	Peak	Mean	Std. Dev.	FWHM
WT uP apo	31.1 Å	30.2 Å	1.3 Å	3.0 Å
WT uP DHSA	30.5 Å	30.3 Å	1.1 Å	1.4 Å
WT uP DHSB	29.2 Å	30.3 Å	1.1 Å	3.0 Å
WT uP r07R	29.9 Å	29.9 Å	1.1 Å	1.7 Å
WT uP r07S	29.1 Å	30.1 Å	1.7 Å	1.7 Å
WT uP ECG	30.7 Å	31.1 Å	0.9 Å	1.5 Å
WT SEP apo	29.6 Å	29.9 Å	0.9 Å	2.4 Å
WT SEP DHSA	29.0 Å	29.1 Å	0.9 Å	1.8 Å
WT SEP DHSB	28.3 Å	29.2 Å	1.1 Å	2.4 Å
WT SEP r07R	29.8 Å	30.2 Å	0.9 Å	2.5 Å
WT SEP r07S	28.4 Å	29.1 Å	1.2 Å	2.0 Å
WT SEP ECG	29.0 Å	29.2 Å	0.9 Å	1.8 Å
G159D uP apo	29.0 Å	29.2 Å	1.0 Å	1.6 Å
G159D uP DHSA	30.1 Å	29.6 Å	1.0 Å	1.4 Å
G159D uP DHSB	29.1 Å	29.6 Å	1.0 Å	1.5 Å
G159D uP r07R	28.4 Å	29.4 Å	1.2 Å	2.6 Å
G159D uP r07S	30.0 Å	29.8 Å	1.0 Å	2.6 Å
G159D uP ECG	29.7 Å	30.2 Å	1.2 Å	1.8 Å
G159D SEP apo	29.2 Å	29.6 Å	1.2 Å	1.5 Å
G159D SEP DHSA	28.6 Å	29.1 Å	0.9 Å	1.6 Å
G159D SEP DHSB	28.8 Å	29.2 Å	0.8 Å	1.4 Å
G159D SEP r07R	30.6 Å	29.8 Å	0.9 Å	2.6 Å
G159D SEP r07S	29.3 Å	29.7 Å	1.1 Å	2.4 Å
G159D SEP ECG	28.7 Å	30.2 Å	1.4 Å	2.5 Å

Table 39: The TnC domain distance metrics for the systems with DHSA, DHSB, r07R, r07S, and ECG.

Unlike the previous set of the small molecules which increased the TnC domain distance in most systems, this set of ligands only introduced small changes to the TnC domain separation metric. In WT uP, DHSA, DHSB, r07R, and r07S only perturbed the system minimally and ECG introduced a large increase to the mean TnC domain distance (Figure 48 and Table 39). In WT SEP, DHSA, DHSB, r07S, and ECG all reduced the mean distance from the apo level of $29.9\,\text{Å}$ to between 29.1-$29.2\,\text{Å}$. r07R introduced a small increase to $30.2\,\text{Å}$. The decrease in TnC domain separation upon phosphorylation in the wild type apo system was conserved in WT systems with DHSA, DHSB, r07S, and ECG. r07R increased the mean distance upon phosphorylation in WT. In G159D uP, all ligand systems had increased mean distance when compared to the apo level with ECG caused the largest increase, from $29.2\,\text{Å}$ to $30.2\,\text{Å}$. In the phosphorylated mutant, DHSA and DHSB reduced the mean distance from $29.6\,\text{Å}$ to $29.1\,\text{Å}$ and $29.2\,\text{Å}$, respectively. r07R and r07S caused small increases to TnC domain separation while ECG increased the distance more significantly. In terms of the phosphorylation effect in the mutant, DHSA and DHSB corrected the change in TnC domain distance to a comparable level as the WT apo systems. r07S and ECG systems produced minimal differences between the phosphorylation states. r07R, however, did not reverse the change observed in the mutant apo system.

5.3 Peptide Binding Energy

5.3.1 Switch Peptide

In WT uP, most ligands increased the switch peptide binding energy while reduced the standard deviation (Table 40). DHSA, DHSB, r07S, and ECG changed the mean of the binding affinity from the apo level of $-108.4\,\text{kcal}\,\text{mol}^{-1}$ to around -110-$112\,\text{kcal}\,\text{mol}^{-1}$. r07R weakened the binding to $-105.5\,\text{kcal}\,\text{mol}^{-1}$. In the phosphorylated wild type, DHSA reduced the binding energy from $-111.9\,\text{kcal}\,\text{mol}^{-1}$ in the apo system to $-108.0\,\text{kcal}\,\text{mol}^{-1}$. r07R and ECG strengthened the binding by a small amount to around $-113\,\text{kcal}\,\text{mol}^{-1}$ while DHSB did not changed the switch peptide binding energy significantly. r07S introduced a big increase in the binding affinity by more than $10\,\text{kcal}\,\text{mol}^{-1}$. The spread of distributions as quantified by standard deviation all stayed similar to the apo level. These ligand effects mean that phosphorylation

	Mean	Std. Dev.		Mean	Std. Dev.
WT uP apo	-108.4	24.0	G159D uP apo	-113.5	13.8
WT uP DHSA	-109.8	14.4	G159D uP DHSA	-108.0	14.5
WT uP DHSB	-111.7	19.2	G159D uP DHSB	-107.4	16.2
WT uP r07R	-105.5	14.9	G159D uP r07R	-108.6	13.0
WT uP r07S	-110.3	14.8	G159D uP r07S	-113.0	17.7
WT uP ECG	-112.0	15.2	G159D uP ECG	-108.1	15.2
WT SEP apo	-111.9	13.9	G159D SEP apo	-117.0	15.0
WT SEP DHSA	-108.0	13.8	G159D SEP DHSA	-112.1	16.2
WT SEP DHSB	-111.5	14.6	G159D SEP DHSB	-114.1	20.2
WT SEP r07R	-112.6	13.5	G159D SEP r07R	-112.3	13.6
WT SEP r07S	-122.1	13.1	G159D SEP r07S	-115.8	11.3
WT SEP ECG	-113.5	13.0	G159D SEP ECG	-113.0	16.7

Table 40: Switch peptide binding energy to the rest of the troponin complex for troponin systems with DHSA, DHSB, r07R, r07S, and ECG. Unit: $\mathrm{kcal\,mol^{-1}}$

weakens DHSA and DHSB WT systems which is the opposite of the apo observation. In the r07R, r07S, and ECG systems, the switch peptide bindings were still strengthened upon phosphorylation except by a much bigger magnitude with r07S.

In G159D uP, ligands weakened the mean switch peptide binding energy in all systems. DHSA, DHSB, r07R, and ECG all reduced the affinity from the apo level of $-113.5\,\mathrm{kcal\,mol^{-1}}$ to around $-108\,\mathrm{kcal\,mol^{-1}}$ which is similar to the WT uP apo value while r07S only introduced a small reduction in the binding energy. In the phosphorylated mutant, all ligands reduced the switch peptide binding energy. Both DHSA and r07R reduced the affinity to around $-112\,\mathrm{kcal\,mol^{-1}}$ which is similar to the WT SEP apo value. DHSA, r07S, and ECG all caused smaller reductions. The similar change in binding energy in both phosphorylation states means that all ligands still followed the phosphorylation induced change in the WT systems. And the binding energies in the presence of ligands were more similar to the WT level.

	Mean	Std. Dev.		Mean	Std. Dev.
WT uP apo	-153.7	23.4	G159D uP apo	-168.5	23.8
WT uP DHSA	-156.5	20.3	G159D uP DHSA	-148.4	20.8
WT uP DHSB	-157.0	26.7	G159D uP DHSB	-164.3	20.4
WT uP r07R	-161.0	21.3	G159D uP r07R	-155.1	14.9
WT uP r07S	-161.3	22.3	G159D uP r07S	-159.9	21.8
WT uP ECG	-155.6	20.6	G159D uP ECG	-158.4	19.7
WT SEP apo	-157.6	20.2	G159D SEP apo	-164.9	20.5
WT SEP DHSA	-164.6	21.3	G159D SEP DHSA	-166.5	21.5
WT SEP DHSB	-164.5	19.9	G159D SEP DHSB	-159.2	17.7
WT SEP r07R	-158.6	16.8	G159D SEP r07R	-159.8	19.4
WT SEP r07S	-171.3	20.5	G159D SEP r07S	-168.2	17.9
WT SEP ECG	-161.9	21.8	G159D SEP ECG	-175.5	22.6

Table 41: Inhibitory peptide binding energy to the rest of the troponin complex for troponin systems with DHSA, DHSB, r07R, r07S, and ECG. Unit: $\mathrm{kcal\,mol^{-1}}$

5.3.2 Inhibitory Peptide

In the wild type systems, all ligands increased the inhibitory peptide binding in both the uP and SEP states (Table 41). For the unphosphorylated systems, DHSA, DHSB, and ECG strengthened the binding energy by around $3\,\mathrm{kcal\,mol^{-1}}$ from the apo level of $-153.7\,\mathrm{kcal\,mol^{-1}}$ while r07R and r07S increased the binding energy more substantially to $-161\,\mathrm{kcal\,mol^{-1}}$. In WT SEP, r07R increased the affinity by $1\,\mathrm{kcal\,mol^{-1}}$ from apo's $-157.6\,\mathrm{kcal\,mol^{-1}}$ while r07S introduced the largest increase to $-171.3\,\mathrm{kcal\,mol^{-1}}$. DHSA and DHSB both strengthened the binding to $-164\,\mathrm{kcal\,mol^{-1}}$ and ECG increased the value to $-161.9\,\mathrm{kcal\,mol^{-1}}$. Overall, the effect of phosphorylation in the wild type is consistent with the apo systems in simulations with DHSA, DHSB, r07S, and ECG. r07R produced the opposite change mainly due to the bigger increase in the WT uP state compared to the SEP state.

Contrasting to the across the board increase in the wild type ligand systems, more reductions were seen in the mutant systems. In G159D uP, all ligands weakened the inhibitory peptide binding energy with DHSA produced the largest decrease by $20\,\mathrm{kcal\,mol^{-1}}$. DHSB introduced

the least amount of decrease, from apo's $-168.5\,\mathrm{kcal\,mol^{-1}}$ to $-164.3\,\mathrm{kcal\,mol^{-1}}$. In G159D SEP, DHSB and r07R both reduced the binding energy to around $-159\,\mathrm{kcal\,mol^{-1}}$. DHSA and r07S strengthened the interactions moderately to around -166-$168\,\mathrm{kcal\,mol^{-1}}$. ECG introduced a large increase in binding affinity to $-175.5\,\mathrm{kcal\,mol^{-1}}$. Mostly of the ligands, except DHSB, reverted the change in binding energy in the mutant to be an increase upon phosphorylation. DHSB did not change the trend due to the sizable reduction in binding energy in the SEP state but not in the uP state.

5.3.3 $\mathbf{TnI_{34-71}}$

	Mean	Std. Dev.		Mean	Std. Dev.
WT uP apo	-215.7	21.2	G159D uP apo	-212.4	19.9
WT uP DHSA	-210.1	19.3	G159D uP DHSA	-212.0	20.5
WT uP DHSB	-216.3	20.4	G159D uP DHSB	-215.0	20.4
WT uP r07R	-218.2	20.0	G159D uP r07R	-213.4	20.6
WT uP r07S	-213.1	20.5	G159D uP r07S	-216.2	21.8
WT uP ECG	-215.7	18.8	G159D uP ECG	-213.5	22.2
WT SEP apo	-226.2	20.1	G159D SEP apo	-217.3	20.0
WT SEP DHSA	-220.8	22.5	G159D SEP DHSA	-221.1	19.9
WT SEP DHSB	-225.5	21.9	G159D SEP DHSB	-214.6	20.5
WT SEP r07R	-216.5	20.3	G159D SEP r07R	-213.3	19.4
WT SEP r07S	-217.8	20.8	G159D SEP r07S	-216.8	20.0
WT SEP ECG	-223.3	20.1	G159D SEP ECG	-217.0	18.6

Table 42: $\mathrm{TnI_{34-71}}$ peptide binding energy to the rest of the troponin complex for troponin systems with DHSA, DHSB, r07R, r07S, and ECG. Unit: $\mathrm{kcal\,mol^{-1}}$

In WT uP, DHSA and r07S reduced the $\mathrm{TnI_{34-71}}$ peptide binding energy (Table 42). DHSB and r07R strengthened the interaction slightly while ECG did not cause a change. In the phosphorylated state, all ligands weakened the $\mathrm{TnI_{34-71}}$ peptide binding interaction. From apo's value of $-226.2\,\mathrm{kcal\,mol^{-1}}$, r07R and r07S produced the largest decrease to around $217\,\mathrm{kcal\,mol^{-1}}$. DHSA and ECG reduced the binding energy moderately while DHSB introduced a small de-

crease. Considering the phosphorylation effect in the wild type, DHSA and DHSB maintained the stronger binding interaction in the SEP state by a similar margin compared to the apo systems. r07S and ECG kept the direction of change but with smaller magnitude. r07R, however, caused the SEP state to be weaker than the uP state which is the opposite of the apo trend.

In G159D uP, DHSA did not change the TnI_{34-71} peptide binding energy meaningfully while the other ligands increased the binding affinity between 1-4 $\mathrm{kcal\,mol^{-1}}$. In the phosphorylated state, DHSA increased the binding energy while the other ligands reduced the binding affinity by less than $4\,\mathrm{kcal\,mol^{-1}}$. Overall, DHSA, r07S, and ECG systems all had stronger binding interactions in the SEP state. DHSB and r07R reversed the apo trend in both WT and G159D to cause a very slightly weaker interaction upon phosphorylation.

5.3.4 NcTnI

	Mean	Std. Dev.		Mean	Std. Dev.
WT uP apo	-129.2	28.4	G159D uP apo	-111.3	33.2
WT uP DHSA	-112.1	38.5	G159D uP DHSA	-121.8	29.5
WT uP DHSB	-121.5	29.4	G159D uP DHSB	-105.2	28.2
WT uP r07R	-123.8	29.7	G159D uP r07R	-120.6	50.0
WT uP r07S	-106.7	40.2	G159D uP r07S	-113.0	28.1
WT uP ECG	-127.7	28.0	G159D uP ECG	-116.0	29.6
WT SEP apo	-99.8	30.8	G159D SEP apo	-116.6	24.8
WT SEP DHSA	-117.7	37.1	G159D SEP DHSA	-99.0	36.0
WT SEP DHSB	-106.4	34.0	G159D SEP DHSB	-109.2	31.3
WT SEP r07R	-88.2	30.8	G159D SEP r07R	-118.1	35.9
WT SEP r07S	-106.8	35.1	G159D SEP r07S	-103.1	42.6
WT SEP ECG	-123.8	28.4	G159D SEP ECG	-110.4	32.8

Table 43: Binding energy of NcTnI to the rest of troponin for troponin systems with DHSA, DHSB, r07R, r07S, and ECG. Unit: $\mathrm{kcal\,mol^{-1}}$

In WT uP, all ligand systems weakened the NcTnI binding energy to the rest of the troponin complex (Table 43). r07S caused the largest decrease, from apo's $-129.2\,\mathrm{kcal\,mol^{-1}}$ to $-106.7\,\mathrm{kcal\,mol^{-1}}$

while ECG produced the smallest decrease of $-1.5\,\mathrm{kcal\,mol^{-1}}$. In the phosphorylated state, DHSA and ECG increased the binding energy significantly to $-117.7\,\mathrm{kcal\,mol^{-1}}$ and $-123.8\,\mathrm{kcal\,mol^{-1}}$, respectively, from apo's value of $-99.8\,\mathrm{kcal\,mol^{-1}}$. Both DHSB and r07S increased the binding affinity to around $-106\,\mathrm{kcal\,mol^{-1}}$ while r07R reduced the value to $-88.2\,\mathrm{kcal\,mol^{-1}}$. These shifts led to drastically different phosphorylation induced changes in the ligand systems. DHSA caused an increase in the binding energy upon phosphorylation in the wild type and r07S abolished the change. DHSB, r07R, and ECG kept the same trend as the apo systems but ECG produced a much reduced difference.

In G159D uP, DHSA and r07R strengthened the binding energy to around $-120\,\mathrm{kcal\,mol^{-1}}$ from apo's $-111.3\,\mathrm{kcal\,mol^{-1}}$. DHSB reduced the value to $-105.2\,\mathrm{kcal\,mol^{-1}}$ while r07S and ECG introduced small increases. In G159D SEP, DHSA caused the largest reduction in the binding energy, from apo's $-116.6\,\mathrm{kcal\,mol^{-1}}$ to $-99.0\,\mathrm{kcal\,mol^{-1}}$ which is very close to the WT SEP apo level. DHSB, r07S, and ECG also reduced the binding energy but to a lesser extend. r07R introduced a very slight increase in the binding interaction. Comparing uP and SEP, DHSA, r07R, r07S, and ECG all caused a weaker interaction in the phosphorylated state which is the trend in the WT apo system. Notably, however, recoupler DHSB was the only one did not reverse this trend in the mutant system.

5.4 Change in Residue Interactions

In Section 4.4, the ionic interaction between TnC Asp159 and TnC Arg83 in the mutant was disrupted by ligand binding. The subsequent shift in the interactions from TnC Asp159 to TnC linker and TnI inhibitory peptide was likely to be one of the mechanisms that ligands can modulate calcium sensitivity in the troponin mutant. The resulting changes in the G159D simulations with ligands in this chapter were less uniform (Table 44). DHSA, r07R, and ECG increased the TnC Arg83 and TnC Asp159 ionic interaction percentage by 1-5% in the uP state. DHSB reduced the interaction from apo's 80% to 62% and r07S caused a more significant reduction to 48%. In the phosphorylated state, DHSA, DHSB, and ECG all increased the interaction percentage by 7-10% from apo's 76%. r07R and r07S reduced the percentage to 72% and 70%, respectively. These data showed that the inactive compounds DHSA and ECG both

increased the interaction percentages in uP and SEP. The recoupler DHSB reduced the interaction in uP but increased in SEP. The enantiomers of r07 displayed opposite behaviour in the uP state while both reduced the percentages in SEP. As the key ionic interaction was not significantly subsidised and increased in some cases, the shift towards interactions with TnC Lys86 and Lys90 was not as pronounced. The main increase has been observed in G159D uP r07S where the ionic interaction between TnC Asp159 and TnC Lys86 existed for 22% of the simulations. This interaction was seen for 5% and 6% in G159D uP DHSB and G159D SEP r07R. 1-2% interaction percentages were seen in other systems.

	TnC Arg83		TnC Lys86		TnC Lys90	
	uP	SEP	uP	SEP	uP	SEP
apo	80%	76%	0%	0%	0%	6%
DHSA	81%	83%	0%	0%	0%	0%
DHSB	62%	84%	5%	0%	0%	0%
r07R	85%	72%	0%	6%	2%	1%
r07S	48%	70%	22%	1%	1%	2%
ECG	82%	86%	2%	1%	0%	0%

Table 44: Ionic interaction percentages between TnC Asp159 and TnC residues in simulations with DHSA, DHSB, r07R, r07S, and ECG.

In terms of interactions from the mutant residues to the inhibitory peptide, subtle changes were seen in the ionic interactions (Table 45). No significant interactions were observed between TnC Asp159 and TnI Arg144 in all systems. All ligands increased the ionic interaction between TnC Asp159 and TnI Arg145 in G159D uP where r07R caused the biggest increase, from 1% in apo to 20%, and r07S caused the least increase, to 4%. For this interaction in the SEP state, DHSA, DHSB, and r07R all eliminated the interation while r07S and ECG increased the interation to 16% and 32%, respectively. With TnI Arg147, no meaningful change was seen in the uP state for the ionic interactions from the mutant residue. In G159D SEP, DHSB increased the interation percentage to 14% and other ligands caused less than 2% changes.

	TnI Arg144		TnI Arg145		TnI Arg147	
	uP	SEP	uP	SEP	uP	SEP
apo	0%	0%	1%	14%	1%	4%
DHSA	0%	0%	9%	1%	0%	0%
DHSB	0%	0%	16%	1%	3%	14%
r07R	1%	0%	20%	0%	0%	2%
r07S	0%	0%	4%	16%	0%	1%
ECG	1%	0%	16%	32%	0%	6%

Table 45: Ionic interaction percentages between TnC Asp159 and TnI residues in simulations with DHSA, DHSB, r07R, r07S, and ECG.

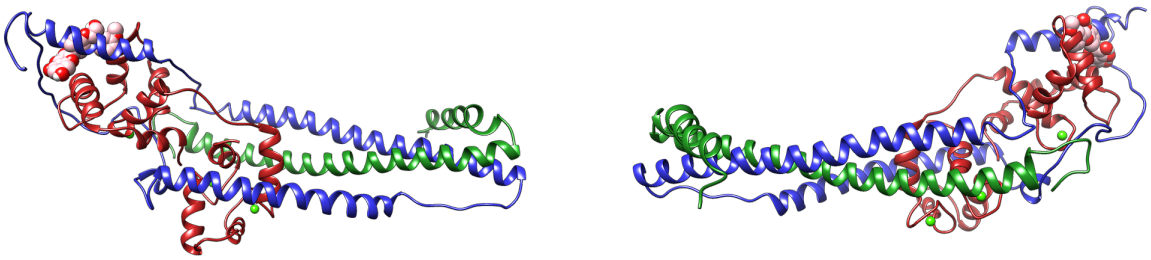
5.5 Binding Region Clustering

Several common binding regions were identified in Section 4 and most of them were located on the regulatory head of troponin. Majority of the binding regions of compounds described in this chapter followed the same pattern and are described in details in the following sections.

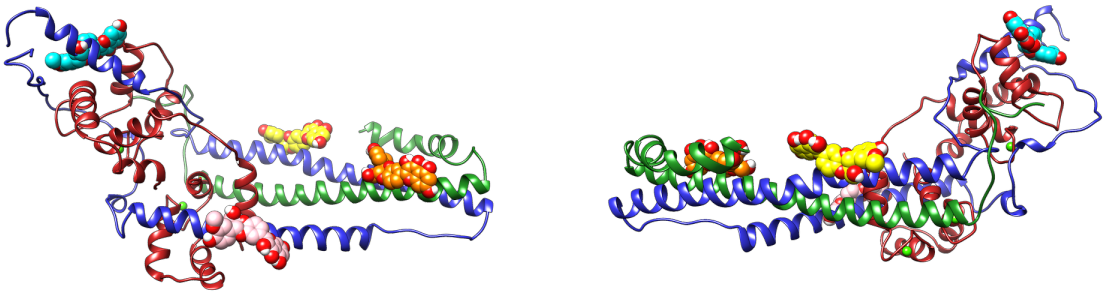
5.5.1 DHSA

In WT uP, DHSA exclusively bound to the tip of the regulatory head where the cluster center was located at the EF-hand I hydrophobic patch and interacted with both ends of TnI Figure 49 and Table 46. Upon phosphorylation, DHSA was seen at a diverse range of binding regions. The major cluster (32%) was centered near EF-hand III and interacted with TnI H1 too. The second cluster (24%) was located at the end of the IT arm and the third cluster (17%) was seen at the middle of the IT arm and interacted with TnI H2. The last binding region of 13% was at the tip of the regulatory head where the major cluster was in the WT uP system.

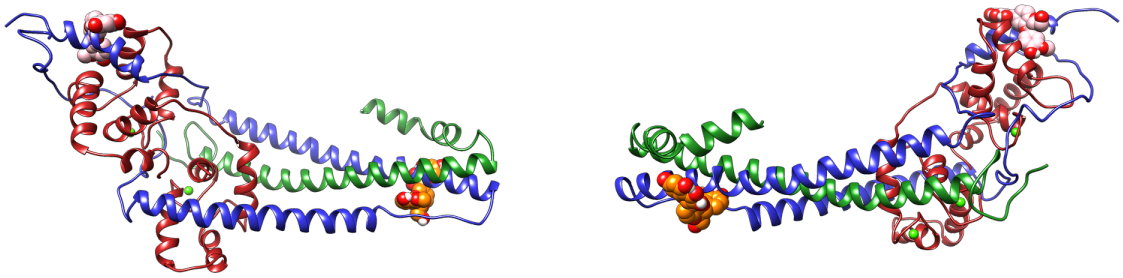
G159D presented a similar binding pattern in both phosphorylation state. The major cluster was around EF-hand I and interacted with switch peptide in both systems (G159D uP: 58%; G159D SEP: 57%). The secondary cluster was seen at the end of the IT arm and was similar in size in both systems as well (G159D uP: 39%; G159D SEP: 37%).



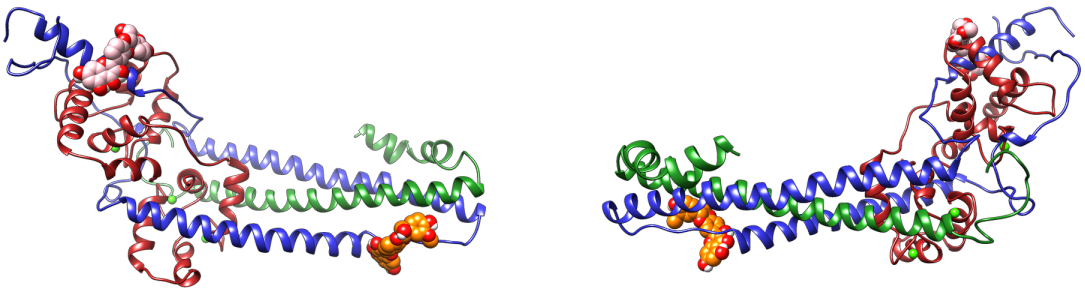
(a) WT uP



(b) WT SEP



(c) G159D uP



(d) G159D SEP

Figure 49: Binding site clusters for systems with DHSA.

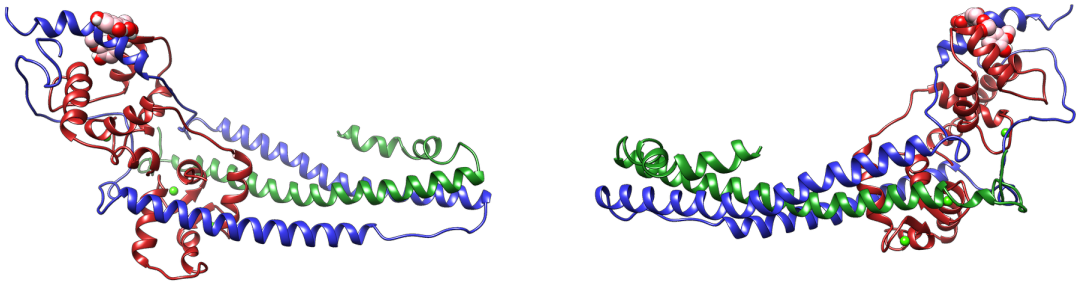
	Pink	Orange	Yellow	Cyan
WT uP	100%			
WT SEP	32%	24%	17%	13%
G159D uP	58%	39%		
G159D SEP	57%	37%		

Table 46: Binding region cluster percentage for DHSA systems.

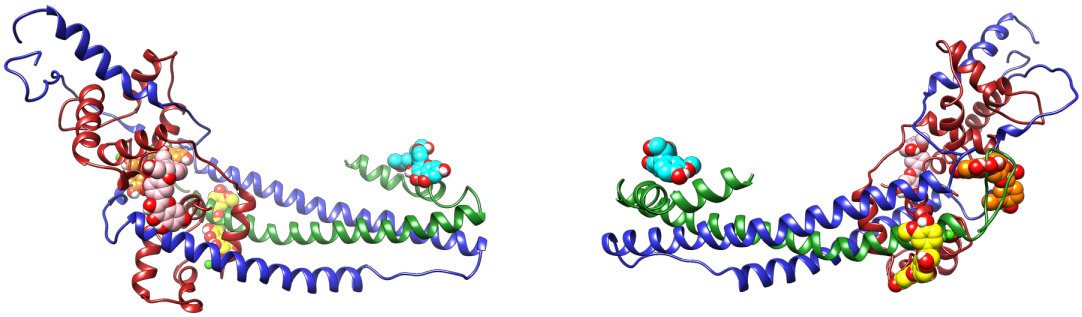
5.5.2 DHSB

Similar to its enantiomer, DHSB also exclusively interacted with TnC EF-hand I and the ends of TnI including the switch peptide in the unphosphorylated wild type (Figure 50 and Table 47). In WT SEP, a more diverse range of binding regions existed, similar to the DHSA system. The major cluster of 59% was centered around the TnC flexible linker and close to the TnI helix region that was bound to CcTnC. The secondary cluster (20%) was located at the back of the hinge region and connected NcTnI with C-terminal end of TnT. The third cluster (14%) was close to the second cluster but shifted further down the IT arm and interacted with part of CcTnC as well. The last minor cluster of 4% was at the tail of the IT arm.

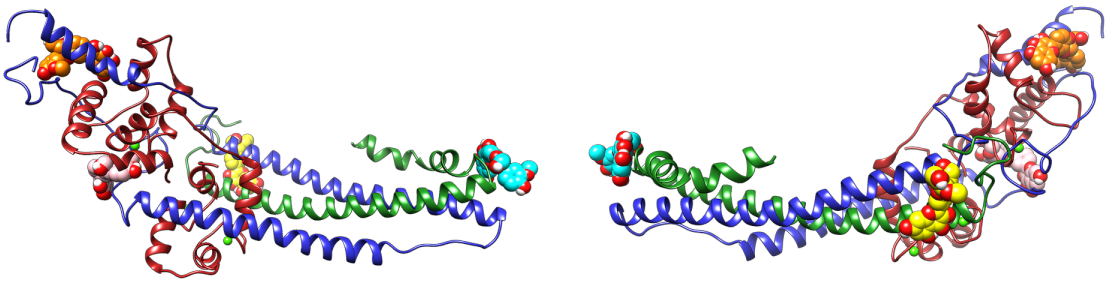
In contrast to the G159D DHSA systems, the binding regions of DHSB varied when phosphorylation was introduced to the mutant. In G159D uP, the major cluster (41%) was located at the back of the hinge region and centered between NcTnI and EF-hand II. The second cluster of 24% was in a similar position as the major cluster in WT uP, located at the tip of the regulatory head. The third cluster (19%) interacted with both TnI H1 and TnT near the hinge region. This position is similar to the third ligand cluster of the WT SEP system. The last cluster of 3% was at the end of the IT arm. The DHSB binding regions mirrored the DHSA system in G159D SEP where the major cluster (73%) was at the tip of the regulatory head and the secondary cluster (22%) at the end of the IT arm.



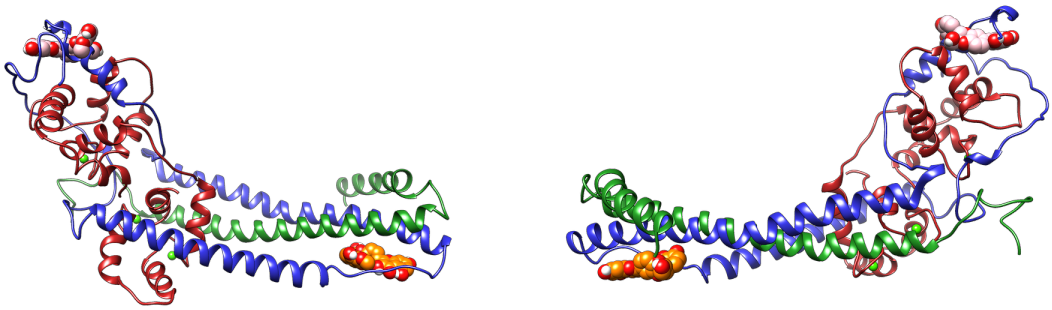
(a) WT uP



(b) WT SEP

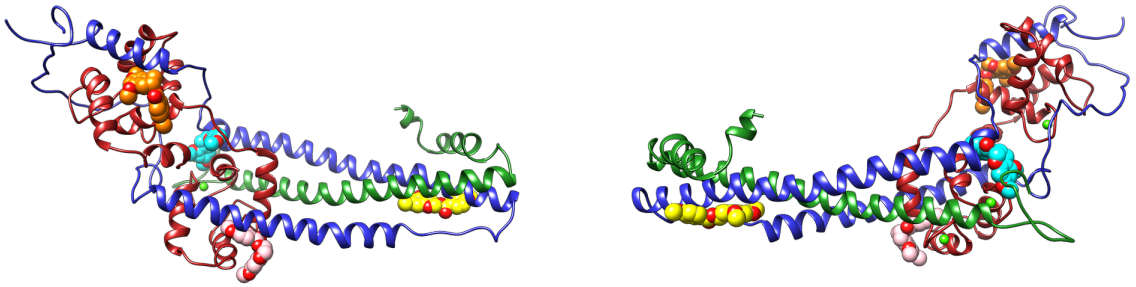


(c) G159D uP

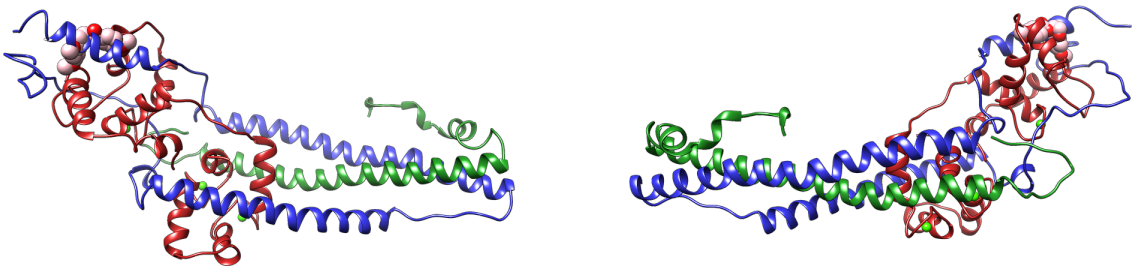


(d) G159D SEP

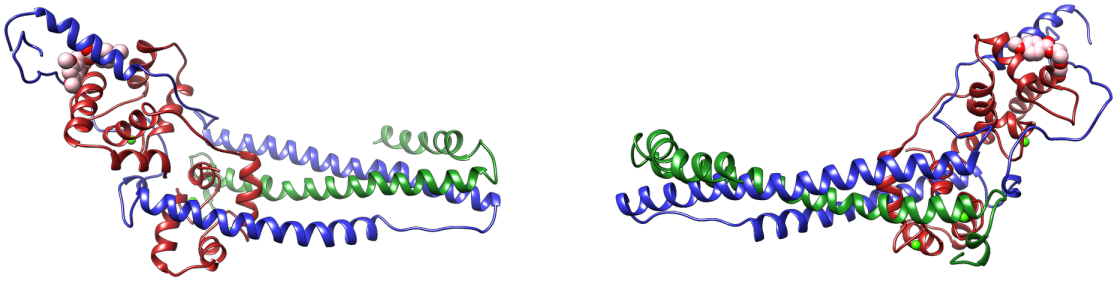
Figure 50: Binding site clusters for systems with DHSB.



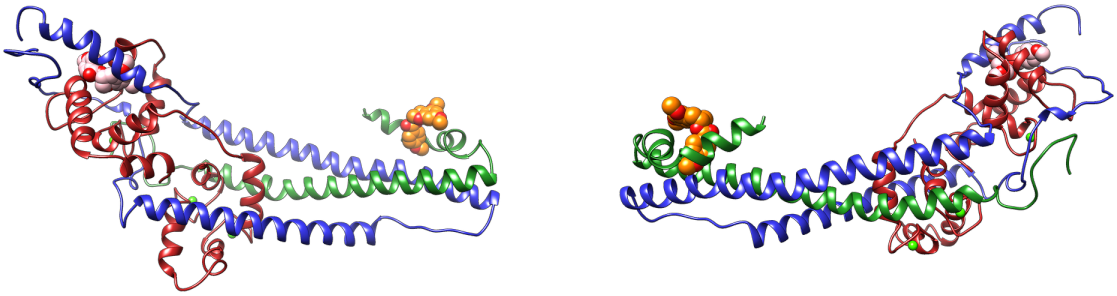
(a) WT uP



(b) WT SEP



(c) G159D uP



(d) G159D SEP

Figure 51: Binding site clusters for systems with r07R.

	Pink	Orange	Yellow	Cyan
WT uP	96%			
WT SEP	59%	20%	14%	4%
G159D uP	41%	24%	19%	3%
G159D SEP	73%	22%		

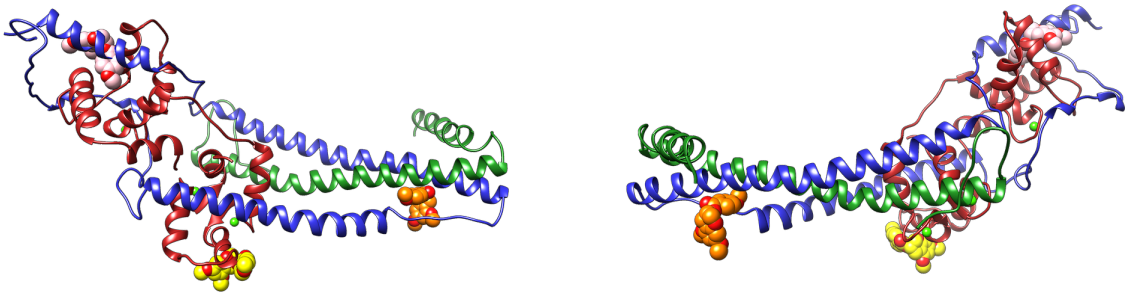
Table 47: Binding region cluster percentage for DHSB systems.

5.5.3 r07R

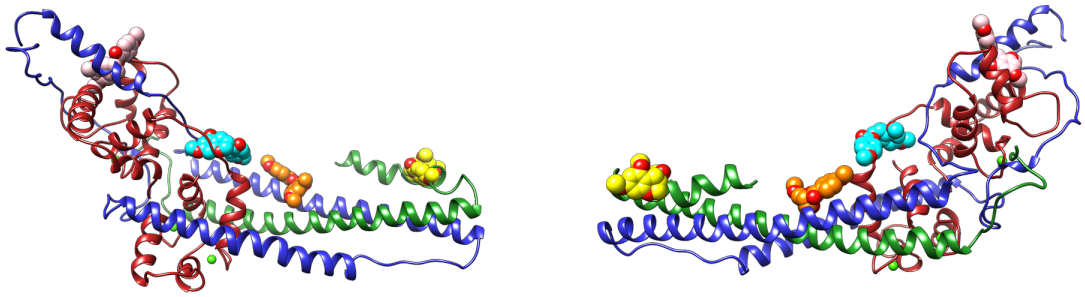
In WT uP, r07R formed binding regions in four similarly occupied clusters (Figure 51 and Table 48). The major cluster (26%) was centered at EF-hand III. In this pose, the ligand was partially inserted into the hydrophobic patch. This was the binding site reported by NMR and computational studies with EGCG yet not observed in the simulations of this work. The much reduced size of r07R likely enabled this to happen. The second cluster of 25% was seen at NcTnC and interacted with TnC helix A, D, and the switch peptide. The third cluster (21%) was at the tail of the IT arm and the last cluster (20%) was at the back of the hinge region and interacted with NcTnI, the inhibitory peptide, and CcTnC. Both WT SEP and G159D uP presented a very similar binding pose. Only a single cluster existed in both systems, 98% in WT SEP and 90% in G159D uP, which was located at the tip of the regulatory head. In G159D SEP, the major cluster of 78% had the ligand inserted into EF-hand I and between the hydrophobic patch and the switch peptide. The minor cluster (8%) was seen at the end of the IT arm.

	Pink	Orange	Yellow	Cyan
WT uP	26%	25%	21%	20%
WT SEP	98%			
G159D uP	90%			
G159D SEP	78%	8%		

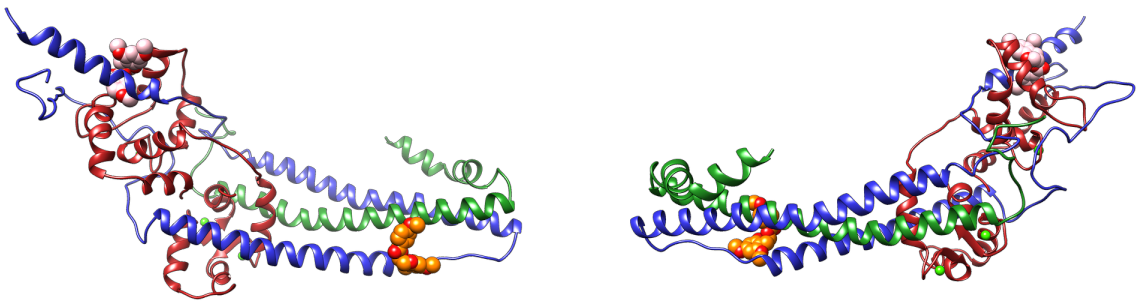
Table 48: Binding region cluster percentage for r07R systems.



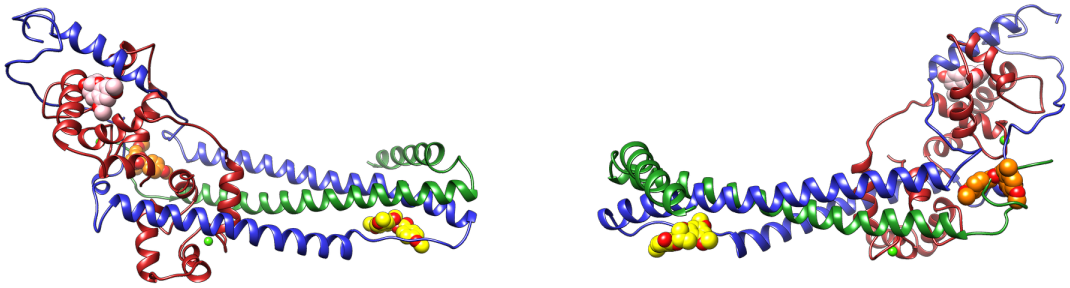
(a) WT uP



(b) WT SEP



(c) G159D uP



(d) G159D SEP

Figure 52: Binding site clusters for systems with r07S.

5.5.4 r07S

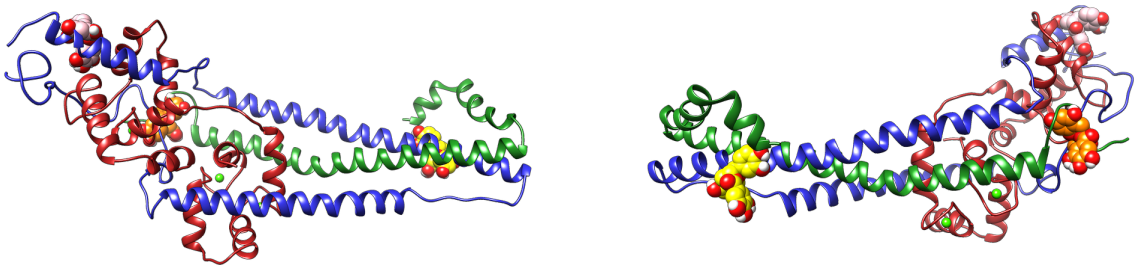
With ligand r07S, the major cluster in all systems were similarly located at EF-hand I of TnC and interacted with the switch peptide (Figure 52 and Table 49). In WT uP, a secondary cluster of 26% was found at the tail end of the IT arm and a minor cluster of 1% was near CcTnC. In WT SEP, both the second and the third clusters occupied 5% of the simulations and were located at the mid and the end of the IT arm, respectively. In G159D uP, the secondary cluster of 12% was on the IT arm and near the U-turn of TnI. In the phosphorylated mutant, the secondary cluster (19%) was at the back of the hinge region and interacted with NcTnI, TnI inhibitory peptide, and TnT.

	Pink	Orange	Yellow	Cyan
WT uP	60%	26%	1%	
WT SEP	76%	5%	5%	1%
G159D uP	77%	12%		
G159D SEP	58%	19%	13%	

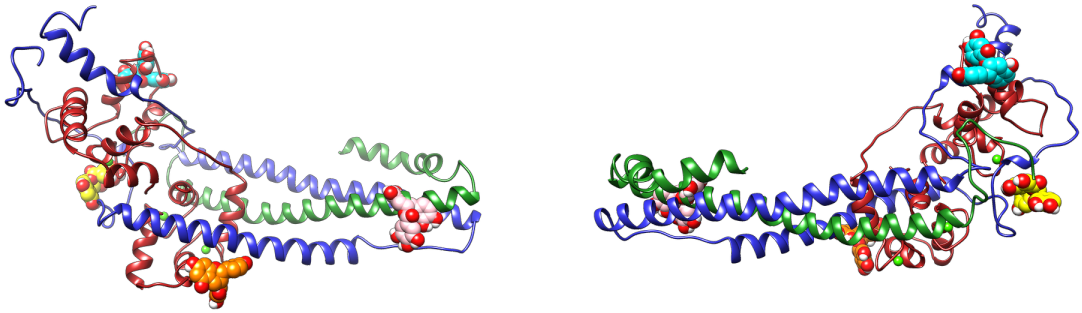
Table 49: Binding region cluster percentage for r07S systems.

5.5.5 ECG

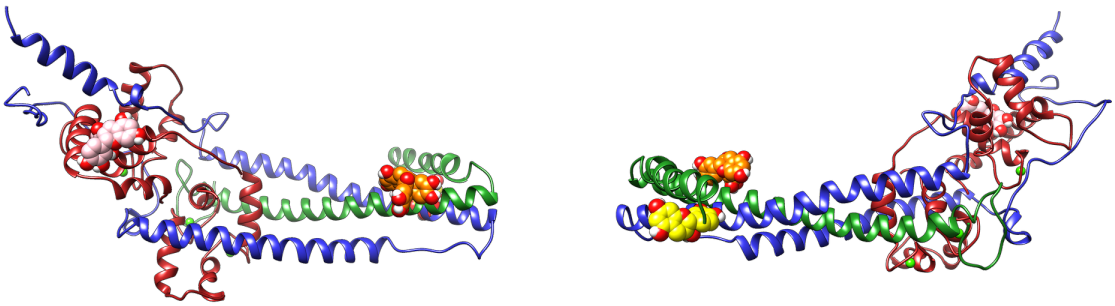
In WT uP, the major cluster of ECG was located at TnC EF-hand I and formed interactions with the C-terminal end of TnI (Figure 53 and Table 50). This cluster accounted for 69% of the simulations. The secondary cluster (19%) was at the back of the hinge region and the third cluster (3%) was at the end of the IT arm. In WT SEP, the largest cluster (51%) moved to TnI U-turn region. The second cluster (16%) was near EF-hand III and the third cluster (10%) was at the back of the hinge region. The smallest cluster of 5% was near the tip of the regulatory head. In the unphosphorylated mutant, the major cluster of 57% was located near NcTnC and interacted with TnC helix A, B, and D. The second and the third cluster were both at the end of the IT arm and accounted for 33% and 3% of the simulations, respectively. In G159D SEP, the major cluster (46%) was at the tail end of the IT arm, similar to WT SEP. The secondary cluster was at the tip of the regulatory head and accounted for 37% of the simulations. The third (7%)



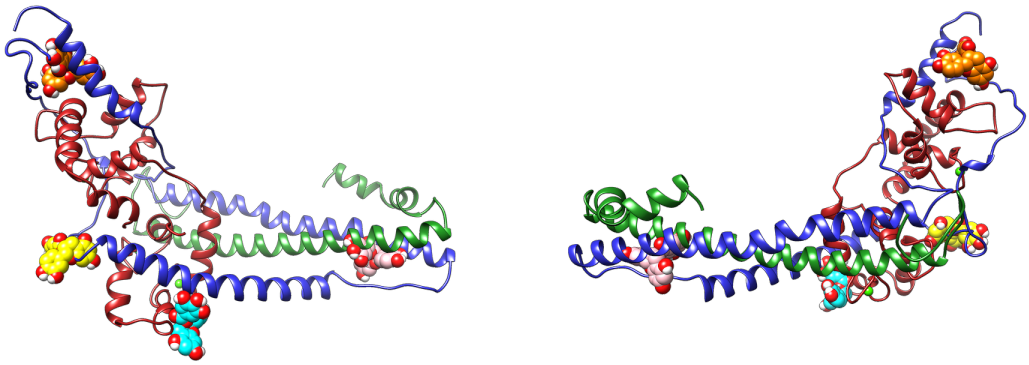
(a) WT uP



(b) WT SEP



(c) G159D uP



(d) G159D SEP

Figure 53: Binding site clusters for systems with ECG.

and the last (3%) cluster were at the back of the hinge region and near CcTnC, respectively.

	Pink	Orange	Yellow	Cyan
WT uP	69%	19%	3%	
WT SEP	51%	16%	10%	5%
G159D uP	57%	33%	3%	
G159D SEP	46%	37%	7%	3%

Table 50: Binding region cluster percentage for ECG systems.

5.6 Ligand binding pattern

This section further extend the general ligand binding regions from the previous section and discuss the detailed ligand interaction patterns with troponin residues. The ligand interacting residues were found for each frame and the results are presented as aggregated counts across the whole 37500 frames for each system.

5.6.1 DHSA

In WT DHSA systems, majority of the ligand interactions with TnC was at the EF-hand I region and the flexible linker in the wild type (Figure 54). This was evident in the TnI interactions as DHSA interacted exclusively with the two ends of TnI. In WT uP TnC, DHSA saw a much subsidised interaction counts with NcTnC and the linker region with more interactions located around EF-hand III loop region. These bindings were reflected in the TnI interaction pattern as well. Major interactions were seen at TnI_{41-71} and TnI H2 part of the IT arm. The TnT interactions were at the tail end of the IT arm.

In the mutant, the similarity in binding regions between uP and SEP was again reflected here (Figure 57). DHSA formed interactions with NcTnC and TnC linker in both systems and interacted with both termini of TnI. Subtle difference between the ligand positions can also be seen from the interaction counts. In G159D SEP, DHSA interacted with a wider range of residues of NcTnI. Binding at the end of the IT arm can be seen in both TnI and TnT's inter-

action counts.

5.6.2 DHSB

In WT uP, residue interactions with DHSA were concentrated at NcTnC, particularly EF-hand I (Figure 56). This was accompanied by interactions with TnC flexible linker, NcTnI, and the switch peptide/CcTnI. In the phosphorylated wild type, ligand interactions were more spread out on TnC than WT uP. The ligand formed interactions with TnC helix N, helix C, flexible linker, and regions between EF-hand III and IV. These interactions were reflected in the NcTnI and inhibitory peptide bindings as well as the CcTnT interactions.

In G159D uP, DHSB presented a similar interaction pattern to WT SEP where it interacted with multiple regions of TnC, including helix N, A, B, D, F, and G (Figure 57). The binding with the back of the hinge region was reflected in interaction counts in the NcTnI region as well as CcTnT. In the phosphorylated mutant, the interactions were more concentrated around EF-hand I, TnI termini, and the TnI U-turn.

5.6.3 r07R

In WT uP, r07R interacted with TnC helix N, A, flexible linker, and EF-hand IV (Figure 58). The interactions at the N-terminal end of TnC were reflected in TnI as the interactions with inhibitory peptide and the lower interaction counts with the switch peptide. The TnI interactions between residues 41-61 were near EF-hand IV. The bindings with the IT arm were shown in both the TnI U-turn region and TnT. The binding of r07R with WT SEP was more concentrated at the tip of the regulatory head where it also interacted with both termini of TnI. r07R also interacted with the flexible linker of TnC.

In both phosphorylation states of the mutant, the ligand interacted with NcTnC together with the N- and C-terminal ends of TnI (Figure 59). Substle difference was seen in the ligand positioning at NcTnC where it interacted more with helix A in the uP state and more with helix B in the SEP state. Minor interactions with TnT was shown in the SEP state.

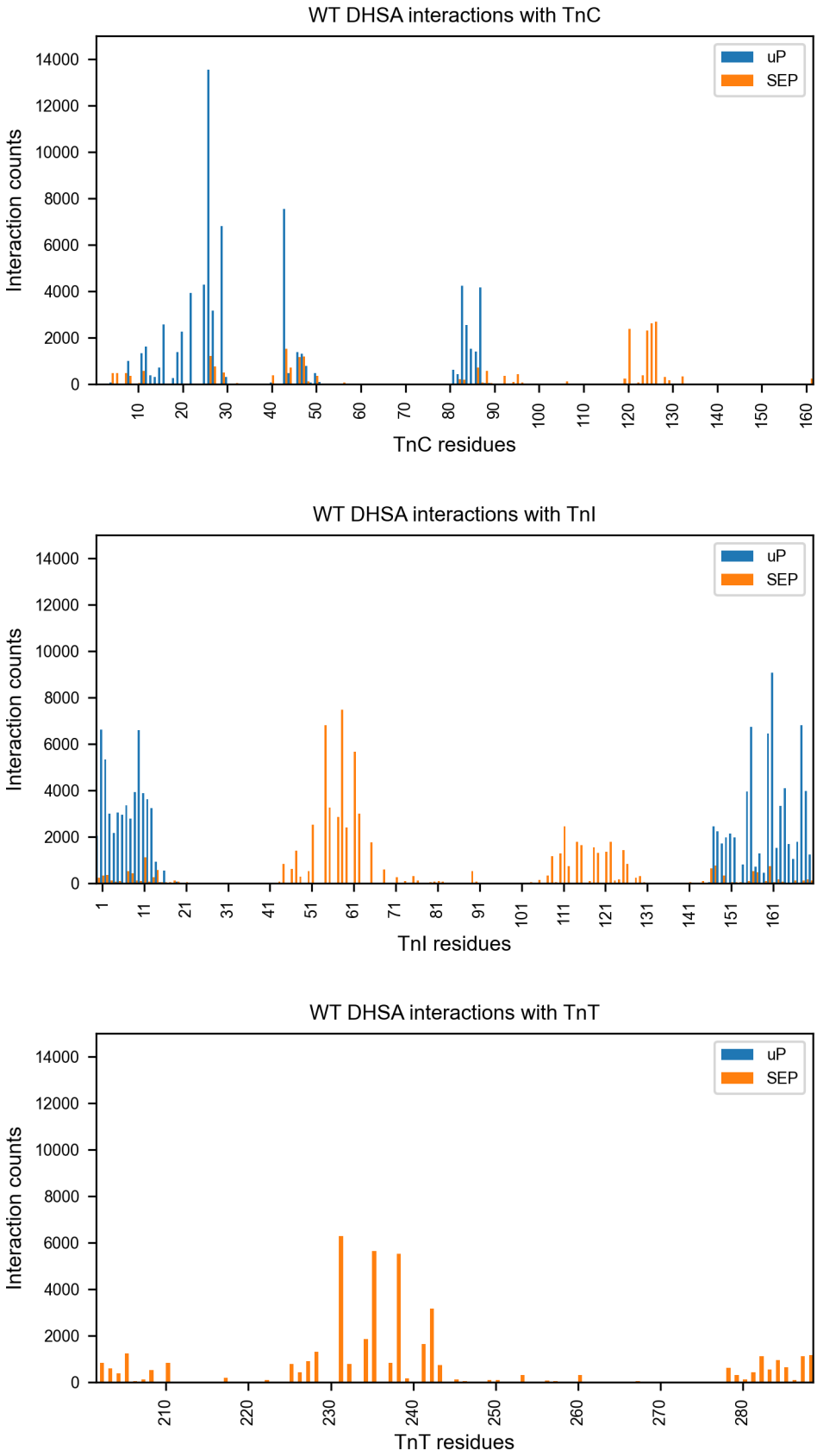


Figure 54: Interaction counts of DHSA with residues in WT systems.

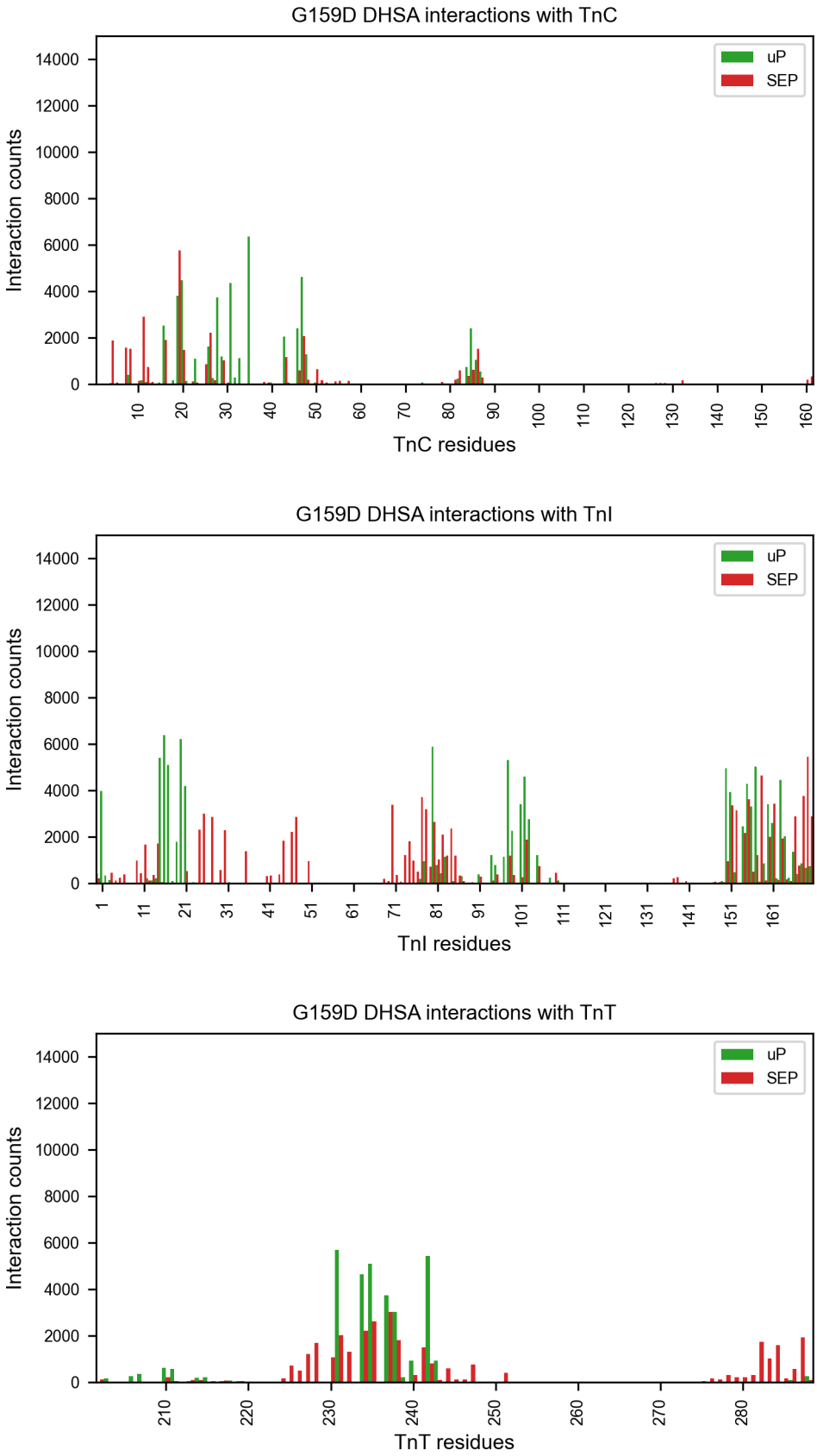


Figure 55: Interaction counts of DHSA with residues in G159D systems.

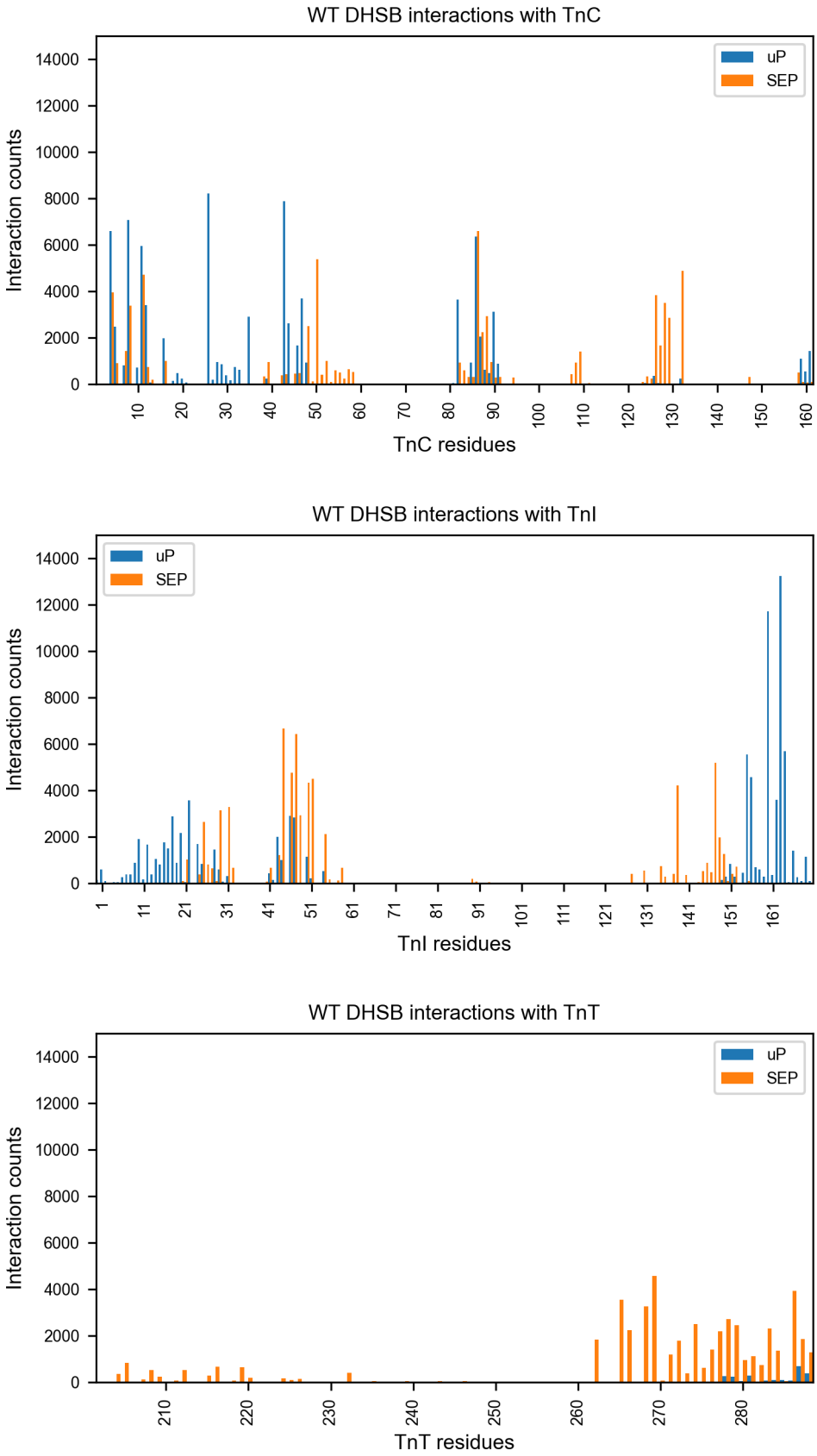


Figure 56: Interaction counts of DHSB with residues in WT systems.

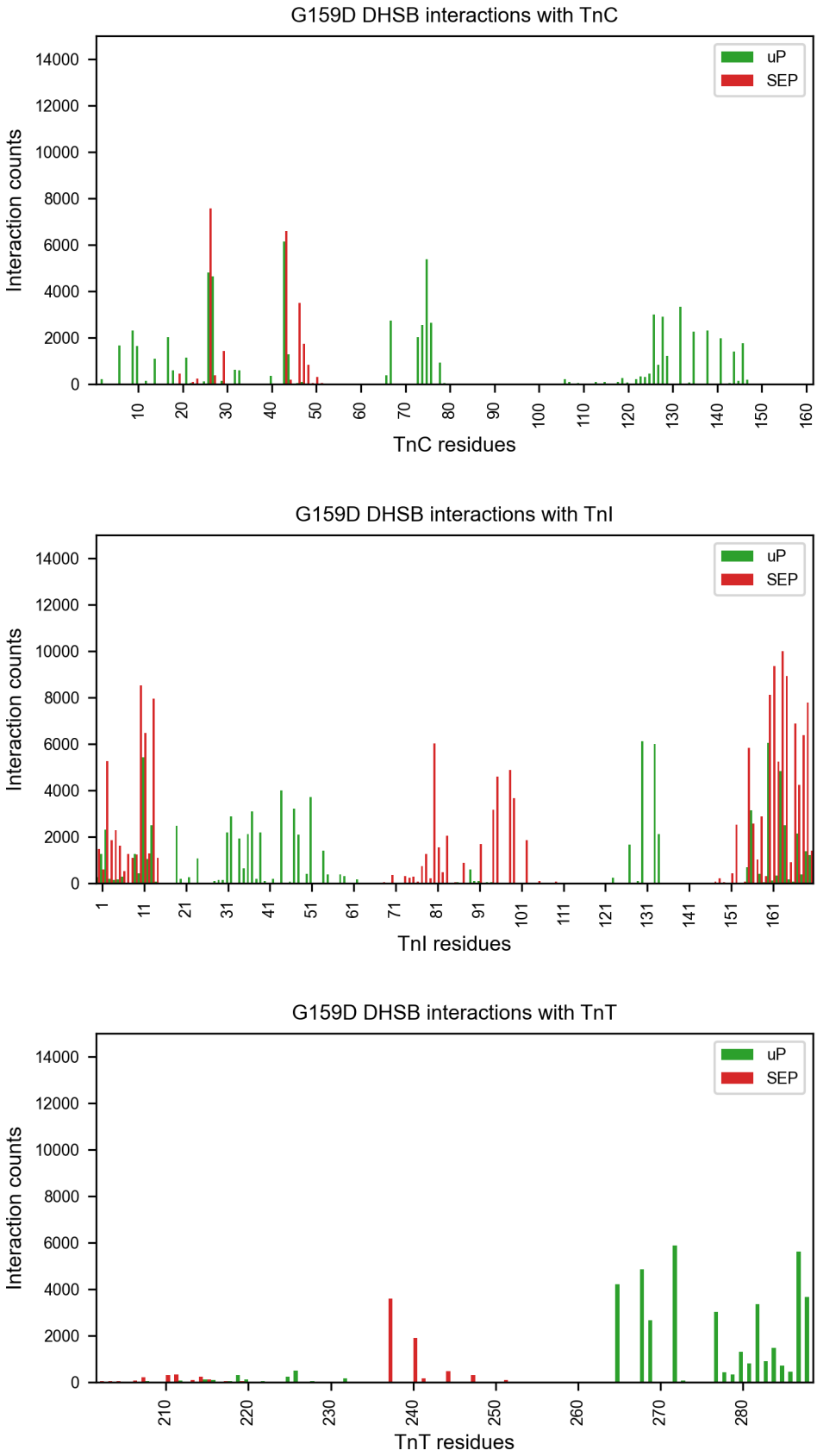


Figure 57: Interaction counts of DHSB with residues in G159D systems.

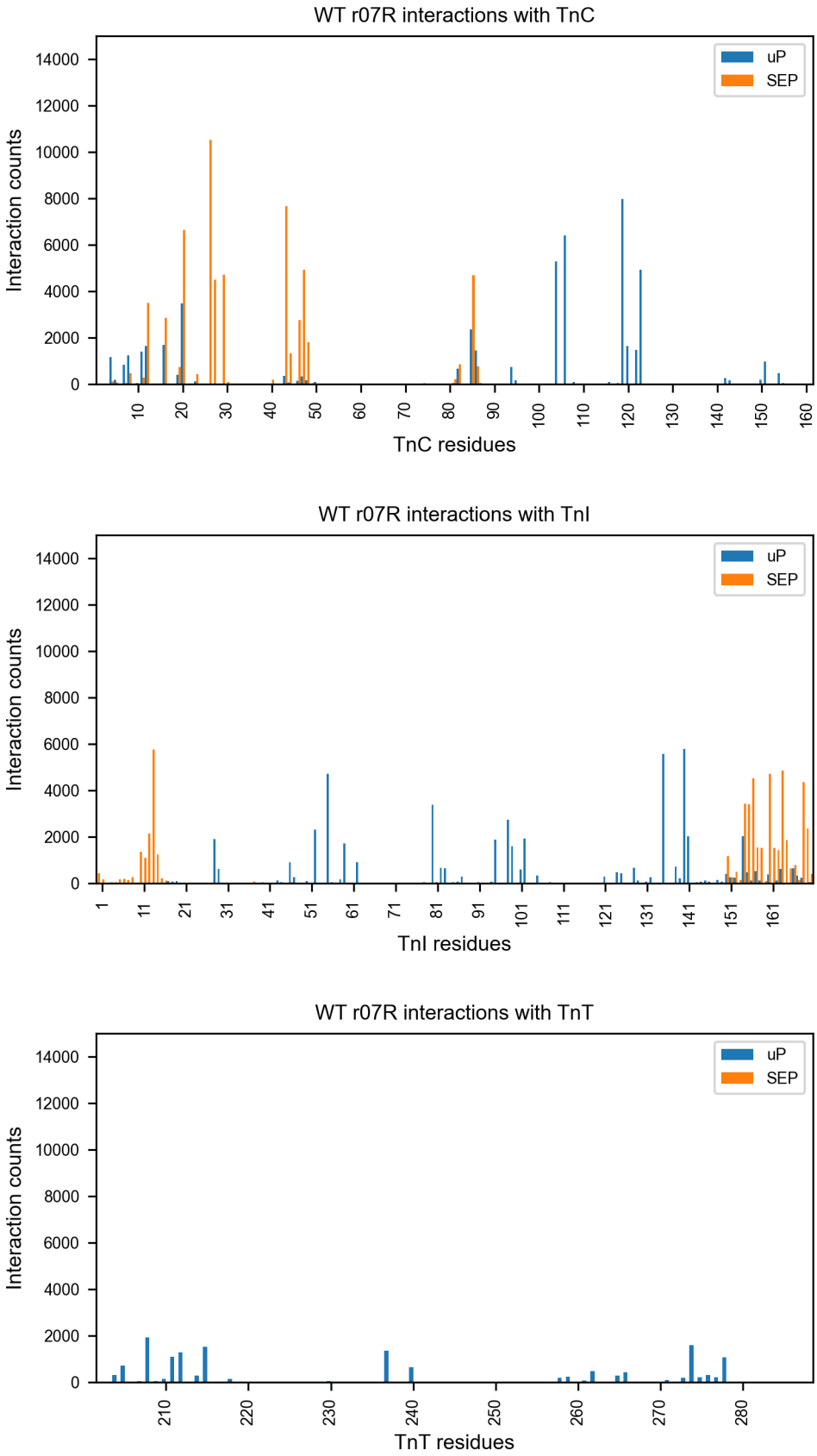


Figure 58: Interaction counts of r07R with residues in WT systems.

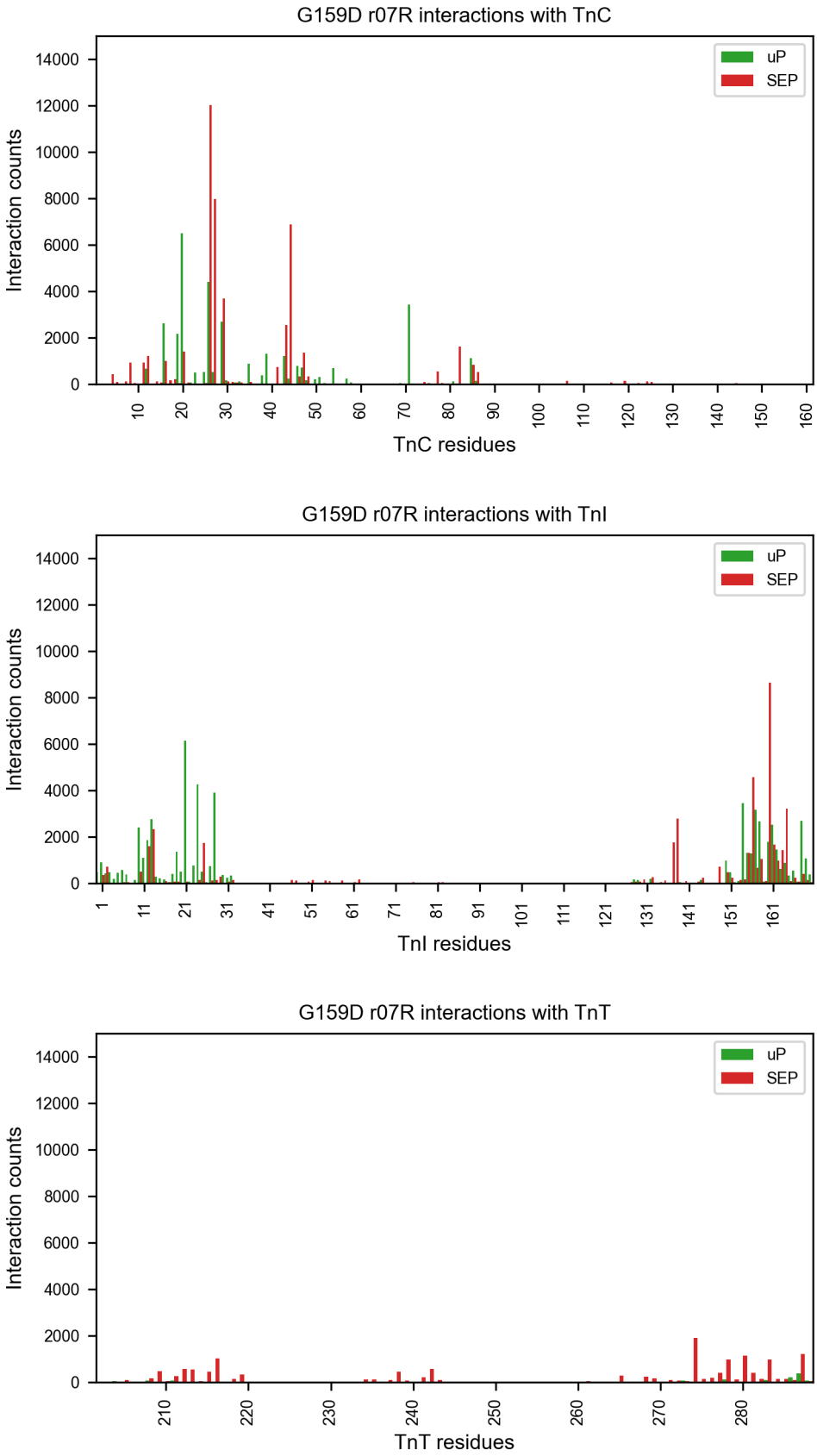


Figure 59: Interaction counts of r07R with residues in G159D systems.

5.6.4 r07S

In the wild type, both uP and SEP presented similar interaction patterns with r07S (Figure 60). TnC interactions were concentrated on NcTnC and SEP state showed higher interaction counts with residues between 40 and 50. Minor interations with TnC flexible linker were observed. The NcTnC interactions were reflected in the TnI subunit as major interactions in the switch peptide region along with some interactions in NcTnI. WT uP showed some interactions near the end of the IT arm. The mutant systems presented very similar patterns to the wild type systems except G159D uP showed higher interaction counts in the EF-hand I region than G159D SEP (Figure 61). Low counts interactions were seen spread across TnI and TnT in both phosphorylation states.

5.6.5 ECG

In WT uP, ECG formed interactions with helix A, B, and C in TnC (Figure 62). These interactions were accompanied by interactions with NcTnI and the switch peptide. The C-terminal end of TnT was also involved in ECG interactions. In WT SEP, the NcTnC interactions with ECG was much weakened and some additional interactions were found between EF-hand III and IV. No interactions with the switch peptide was seen and the NcTnI interactions occured when the ligand was located at the back of the hinge region, together with some CcTnT interactions. Extensive binding between ECG and the IT arm was also seen in the phosphorylated state.

In the mutant systems, ECG was found to form high counts interactions with the tail of the IT arm in both uP and SEP states (Figure 63). The ligand was seen inserted inbetween TnI and TnT hence forming stable interactions. ECG also formed extensive interactions with both termini of TnI in both uP and SEP. In the TnC subunit, ECG interacted with helix N, A, and the flexible linker in uP while mainly interacted with helix A in the phosphorylated state.

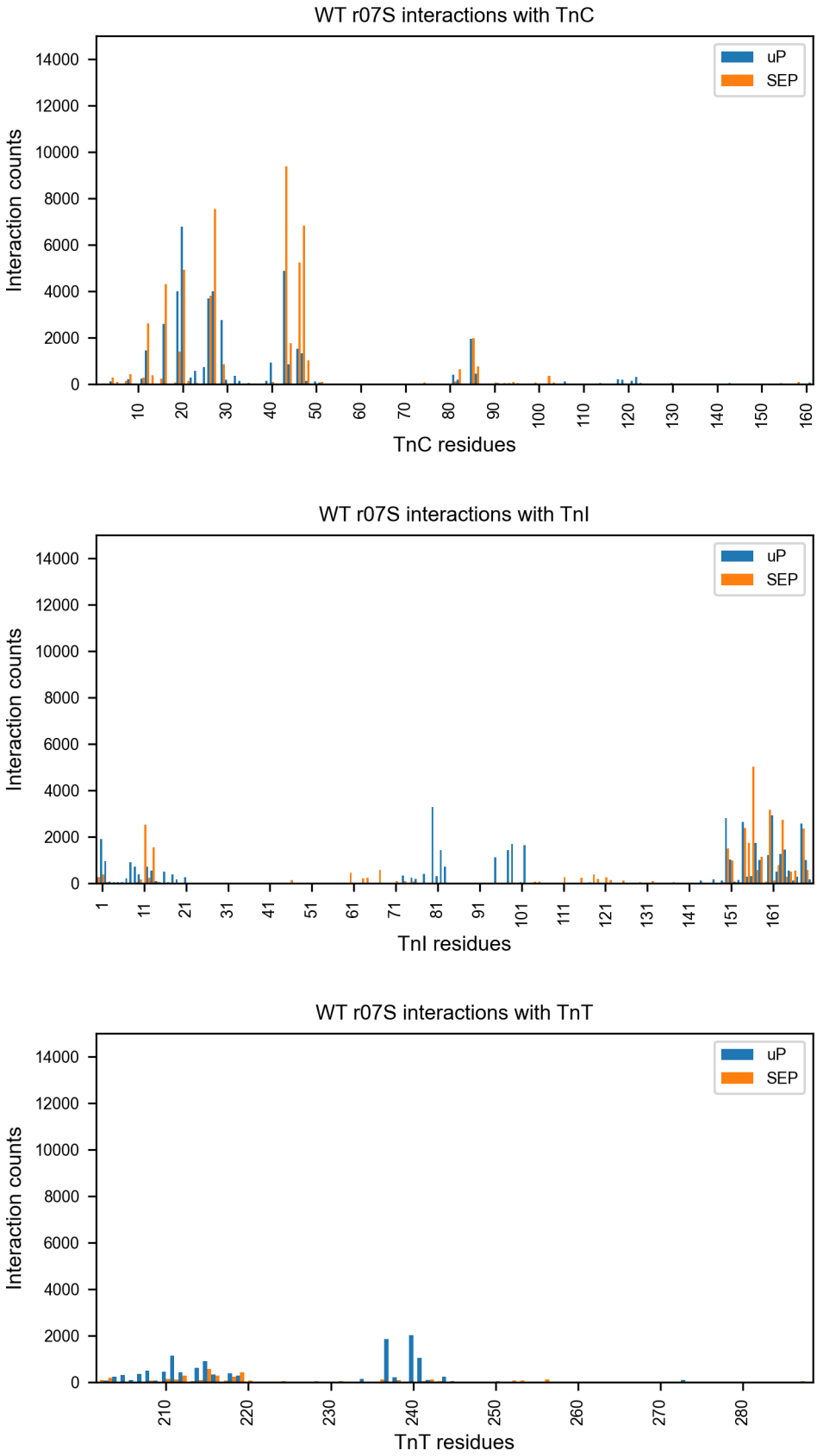


Figure 60: Interaction counts of r07S with residues in WT systems.

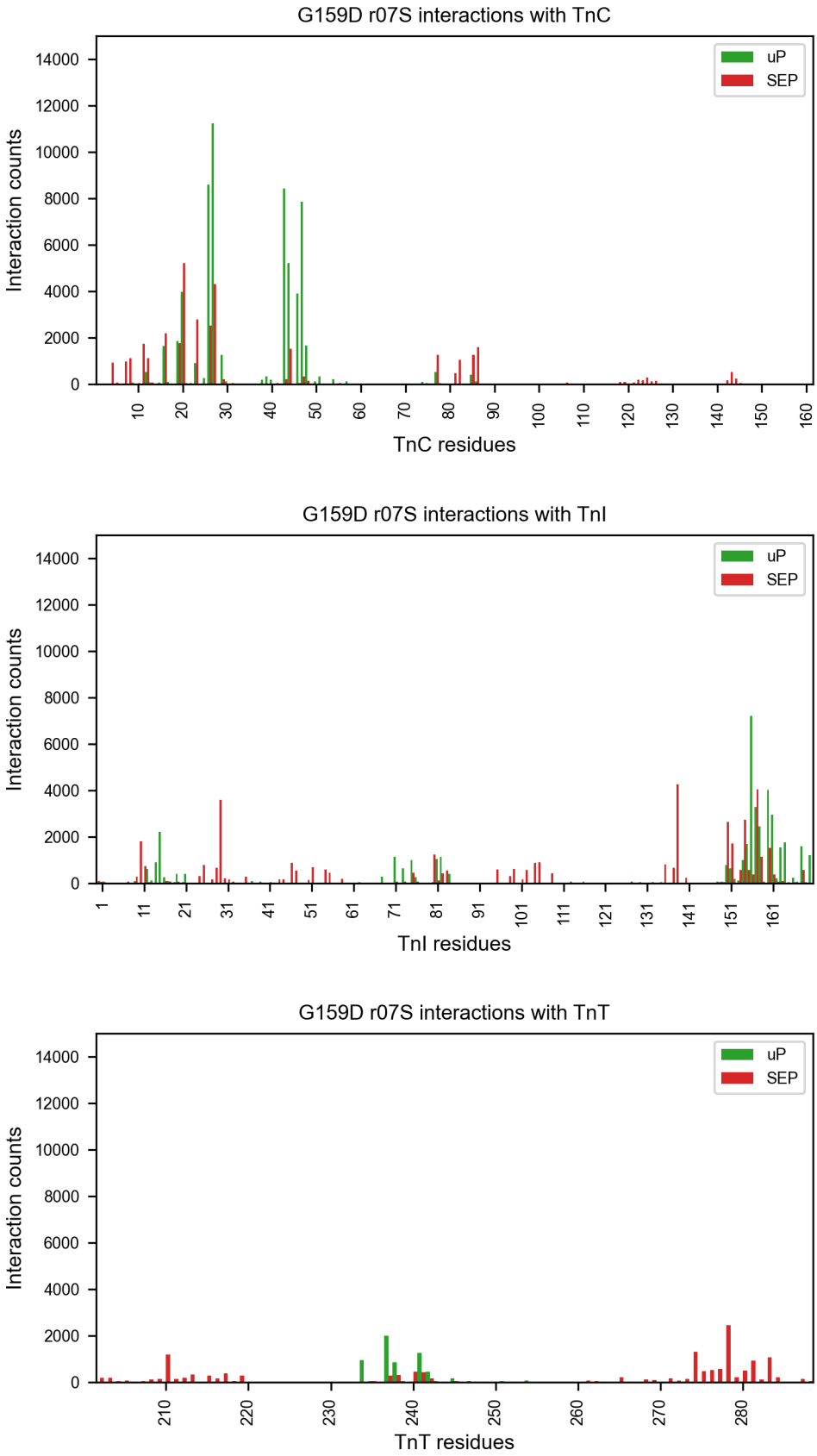


Figure 61: Interaction counts of r07S with residues in G159D systems.

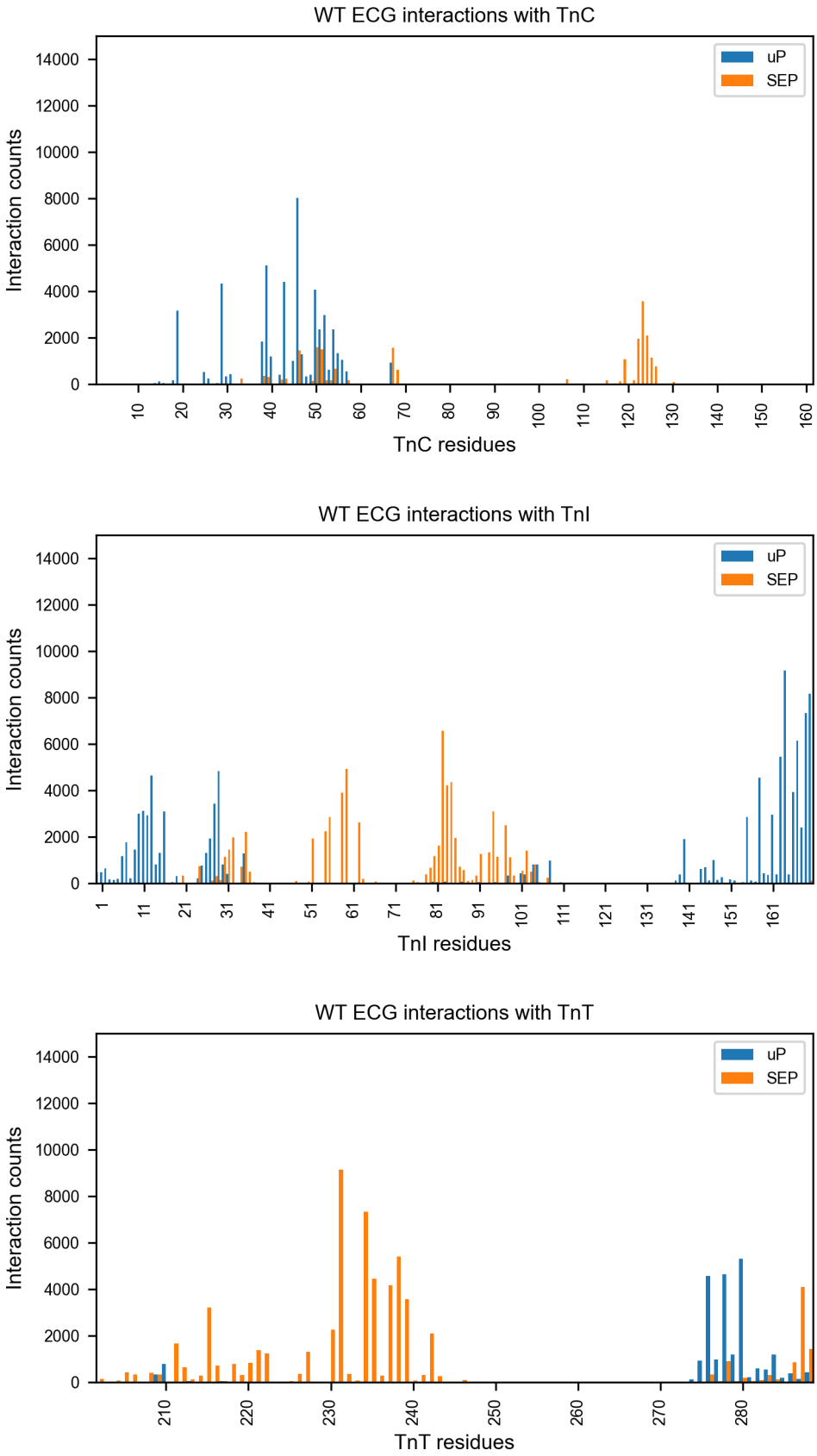


Figure 62: Interaction counts of ECG with residues in WT systems.

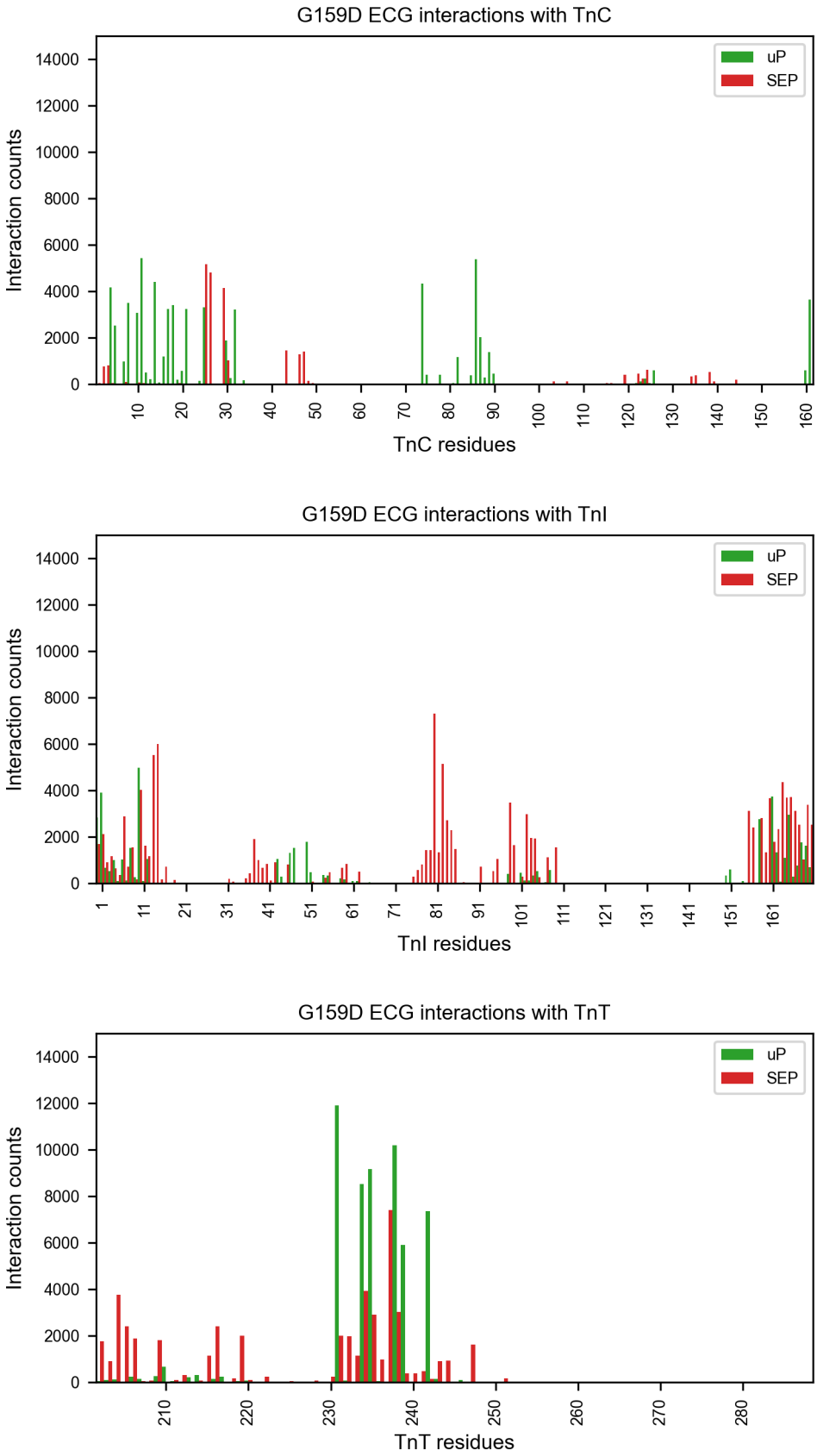


Figure 63: Interaction counts of ECG with residues in G159D systems.

	WT uP	WT SEP	G159D uP	G159D SEP
DHSA	-24.6	-18.4	-24.2	-22.4
DHSB	-21.7	-23.8	-23.7	-27.2
r07R	-18.6	-23.0	-19.3	-20.2
r07S	-17.6	-17.6	-20.3	-17.9
ECG	-24.7	-20.3	-25.0	-25.9

Table 51: Mean ligand binding energy in troponin systems with DHSA, DHSB, r07R, r07S, and ECG. Unit: $\mathrm{kcal\,mol^{-1}}$

5.7 Ligand Binding Energy

Table 51 summarised the averaged binding energies across the whole simulations for each system. Ligand binding energies were mostly comparable yet some outliers exist. DHSA had a lower binding energy in the WT SEP state compared to the other systems. DHSB had a stronger binding energy in G159D SEP even though the major binding region was similar to the WT uP system. Recoupler DHSB was also the strongest G159D SEP binding ligand out of this group of small molecules. This coincides with the hypothesis that recouplers selectively bind to hence stabilise the phosphorylated state of the mutant. r07R had a stronger interaction with WT SEP while r07S had a stronger binding with G159D uP than other troponin systems of the same ligand. Of ECG systems, WT SEP had the lowest binding energy. Comparing the enantiomer pairs, both DHSA/DHSB and r07R/r07S showed significantly different binding energies for the same system. DHSA had a stronger binding energy in WT uP while DHSB formed stronger interactions in WT SEP and G159D SEP. Comparing r07R and r07S, r07R bounded to troponin stronger in WT SEP and G159D SEP than r07S and other systems were more comparable.

To compare the ligand binding strength when bound to different regions, Table 52 contains the averaged binding energy for each cluster of the ligand systems. Across the board, r07R and r07S tend to bind weaker due to their smaller size and bindings near the tail of the IT arm tend to be weaker. The strongest binding region ($-35.6\,\mathrm{kcal\,mol^{-1}}$) was when DHSB bound to the CcTnT in G159D uP. However, this pose had the TnT wrapped around the ligand and enclosed it together with TnI and TnT helical portion of the IT arm which might be unrealistic

	Pink	Orange	Yellow	Cyan
WT uP DHSA	-24.6			
WT uP DHSB	-22.6			
WT uP r07R	-20.7	-20.6	-18.5	-20.2
WT uP r07S	-21.3	-17.0	-13.0	
WT uP ECG	-28.6	-22.8	-24.1	
WT SEP DHSA	-22.7	-21.4	-19.0	-21.3
WT SEP DHSB	-23.7	-28.7	-23.3	-16.2
WT SEP r07R	-23.3			
WT SEP r07S	-21.2	-11.3	-14.8	-10.4
WT SEP ECG	-24.4	-27.5	-22.5	-23.0
G159D uP DHSA	-27.8	-20.6		
G159D uP DHSB	-25.1	-25.7	-35.6	-13.6
G159D uP r07R	-21.3			
G159D uP r07S	-23.5	-16.7		
G159D uP ECG	-26.5	-27.6	-20.7	
G159D SEP DHSA	-26.9	-18.9		
G159D SEP DHSB	-29.5	-24.4		
G159D SEP r07R	-24.0	-14.4		
G159D SEP r07S	-21.5	-18.7	-14.4	
G159D SEP ECG	-24.9	-33.3	-19.4	-22.1

Table 52: Ligand binding energy in troponin systems with DHSA, DHSB, r07R, r07S, and ECG separated by clusters. Unit: $\mathrm{kcal\,mol^{-1}}$

as this segment of TnT was suggested to be involved in actin interactions. The next strongest binding region was ECG in G159D SEP, with a binding energy of $-33.3\,\mathrm{kcal\,mol^{-1}}$. Similar to the strongest binding pose of EGCG, ECG was encapsulated by TnI EF-hand I, switch peptide, and NcTnI at the tip of the regulatory head. In majority systems, the strongest binding clusters of each system were all centered around EF-hand I. In WT SEP, the strongest binding sites with DHSB and ECG ($-28.7\,\mathrm{kcal\,mol^{-1}}$ and $-27.5\,\mathrm{kcal\,mol^{-1}}$) were located at the back of the hinge region with interactions formed with NcTnI, TnT, and NcTnC.

5.8 Discussion

5.8.1 Changes in Key Metrics with Phosphorylation

Figure 64 summarised the differences in the measured metrics when comparing uP and SEP for each system. The recoupler DHSB was able to restore the changes in A-B angle and TnC domain distance in G159D back to the wild-type trend. It could not correct the hinge angle anomaly and inhibitory and NcTnI binding energy changes. Moreover, DHSB appeared to have introduced defects into hinge angle and switch peptide binding in WT and TnI_{34-71} binding in G159D. This represents a big deviation from silybin B's results where it was able to correct all erroneous metrics in G159D. The inactive enantiomer, DHSA, was able to fix A-B angle, TnC domain distance, inhibitory peptide binding energy, and NcTnI binding enerngy although it also introduced defects in hinge angle, switch peptide binding, and NcTnI binding in the wild type. Comparing the enantiomers r07R and r07S, they largely caused similar changes in metrics. r07R reverted inhibitory peptide and NcTnI binding energy to the wild type trend. It also abolished the differences in hinge angle and TnI_{34-71} binding energy in G159D and A-B angle in WT. r07S removed the differences in NcTnI binding energy in WT and greatly reduced TnI_{34-71} binding energy difference in G159D. This lack of differentials between phosphorylation states maybe connected to r07's reverse-recoupling properties. The pure desensitiser ECG did not make a sufficient difference in the structural metrics but corrected the inhibitory peptide and NcTnI binding energy in G159D. In WT, ECG also reversed the change in hinge angle metric. Overall, the feature observations in the inactive compounds were less consistent than the recouplers presented in the previous chapter. The stark differ-

ence between DHSB and silybin B shows the impact of minor changes in ligand which can lead to different behaviours in the measured metrics.

5.8.2 Ligand Binding Regions

Across the board, the ligand binding regions discovered in this chapter are similar to the previous group of compounds. The reverse-recoupler r07 did not interact with a unique binding pocket. Similar to silybin A and B, the stereoisomers in this chapter presented differentials in binding regions in all systems as well. Same as the previous findings, the strongest binding region in all systems is at the tip of the regulatory head where the ligand form interactions with EF-hand I, NcTnI, and the switch peptide. The fully enclosed nature of the binding site ensures strong interactions while the flexibility of the two termini of TnI let the ligands insert the pocket with ease. In many cases, the ligand also explored multiple binding regions in one system and no apparent patterns can be seen for the inactive compounds or desensitisers. This fits with the previously proposed hypothesis that ligand interacts with multiple regions of troponin collectively to produce the observed physiological response.

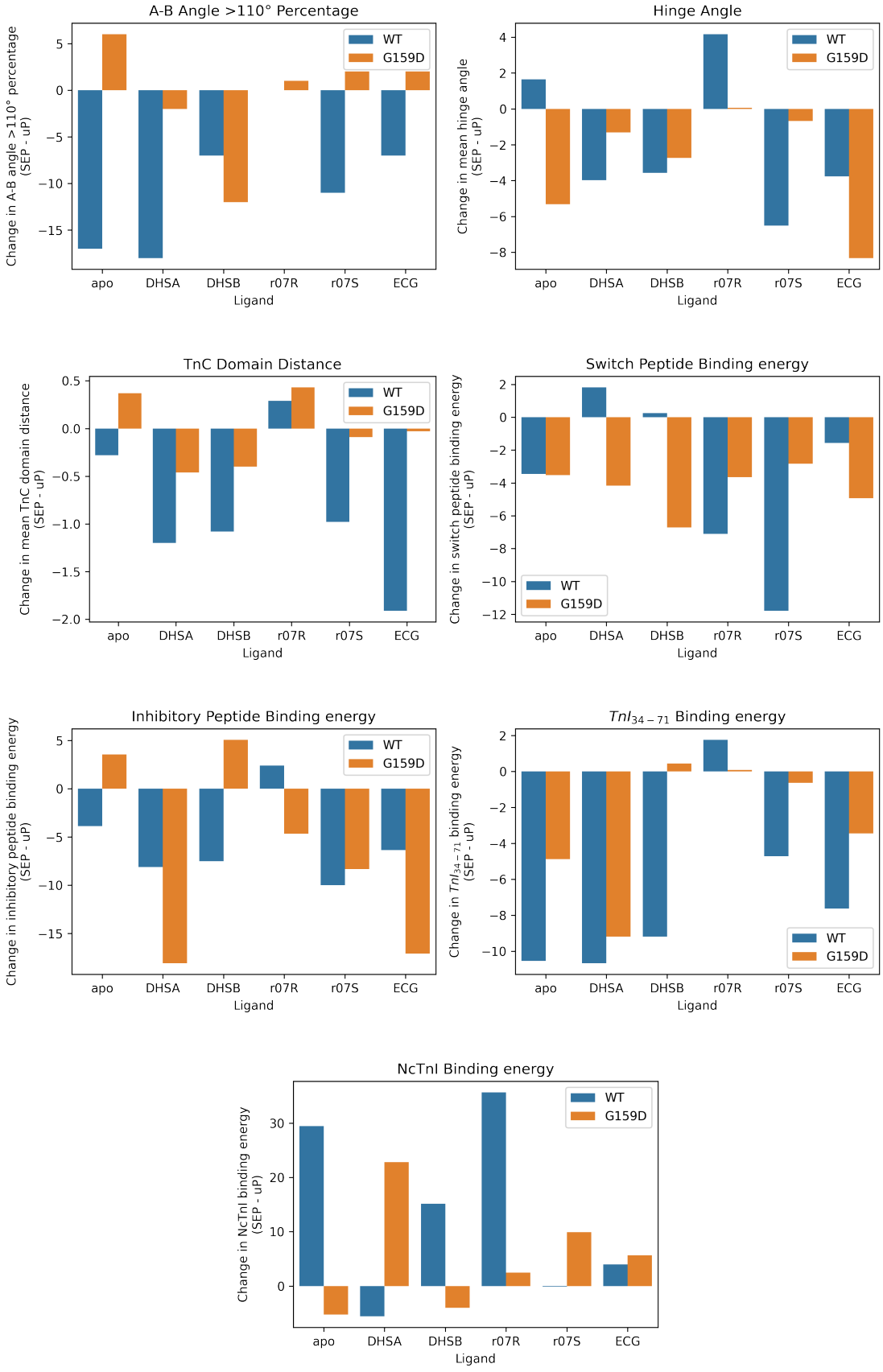


Figure 64: Comparison of changes in key metrics upon phosphorylation in ligand systems with apo systems.

6 Conclusion

All-atom MD simulation was used to study the core of human cardiac troponin and its TNNC1 G159D mutant in the Ca^{2+} activated, unphosphorylated and phosphorylated state. In the apo study, key metrics (A-B angle, hinge angle, TnC domain distance; switch pepide, inhibitory peptide, TnI_{34-71}, and NcTnI binding energy) were measured which provided a clear picture in understanding the phosphorylation and mutation induced changes in troponin's structural and energetics features. Where comparable to experimentally measured binding energies, the computed results produced the same trend of changes due to phosphorylation and mutation. Extensive residue interaction profiling was used to discover how mutation and phosphorylation changed the protein's residue interactions which led to observed changes in stucture and peptide binding. This provided quantitative results in assessing troponin's structural dynamics. Lastly, I have employed a novel Granger causality method to analyse the protein's residue interaction data and uncovered allosteric pathways that transmit the phosphorylation and mutation effect to the rest of the troponin complex. This methodology represents a first case of causal inference using residue interaction data from MD simulations which gave convincing results.

Based on the apo study's result, ligand studies were conducted to include nine ligands into the simulations. The simulation trajectories demonstrated a good sampling coverage of the protein complex's surface for potential binding sites. The recouplers (EGCG, RVL, silybin B, DHSB) were able to correct many of the erroneous metrics caused by mutation while the results for inactive compounds and desensitisers were more varied. This parallels *in vitro* results where recouplers were also able to fix experimentally measured parameters of contraction magnitude and speed. From the ligand binding region clustering and the ligand residue interaction counts, common binding regions were discovered. Although these ligand interactions can be linked to the changes in structural and energetics features, the complex landscape of multiple binding regions made it impossible to assert correlations between ligand bindings and *in vitro* assay outcomes. It is also possible that the small molecule effects can not be explained by the occupancy of a single binding site. Instead, I propose that perhaps the ligand has to interact with multiple regions thus correcting several structural and

energetics defects to produce the desired outcome.

6.1 Future Work

6.1.1 A More Complete Model

With the rapid advance of computer hardware, the field of cardiac muscle simulation has gradually moved to studying interactions within the full thin filament systems. The cutting edge GPU is able to simulation multiple nanoseconds of the thin filaments every day. Recent seminal work in thin filament cryo-EM provided high quality starting structures for MD simulations. Szatkowski *et al.* presented a computational study using docking and the fully atomistic model of the cardiac thin filament to uncover candidate drug molecules to restore the normal cardiac functions in the mutants which was experimentally validated later.[144] A more complete model generally reduces modelling artifacts introduced by unrealistic interactions as seen in this work and allows us to study the detailed interactions on a bigger scale. Modelling the full thin filament also allows the study of many DCM- and HCM-causing mutations that are located outside the troponin core.

6.1.2 Further Development on Causal Inference Methods

In the apo studies, I have presented a novel method based on Granger causality that uncovered molecular interaction pathways which brought explainable causal effects to MD simulations. These pathways uncovered by causal inference agreed with our intuition and allowed a more autonomous unbiased analysis of molecular interactions from MD trajectories. I believe this limited case of using time-series causal inference method in modelling molecular interactions has demonstrated its potential and I think it warrants further development. It would be possible to apply the Granger causality method to residue distance as continuous values which might increase the inference power. Using atomic level interactions would further increase the granularity although also introduce more noise. Finally, applying the method to ligand systems might uncover how ligand binding changes residue interaction pathways but this requires careful treatment of ligand binding positions.

7 Appendix

7.1 Chapter 2

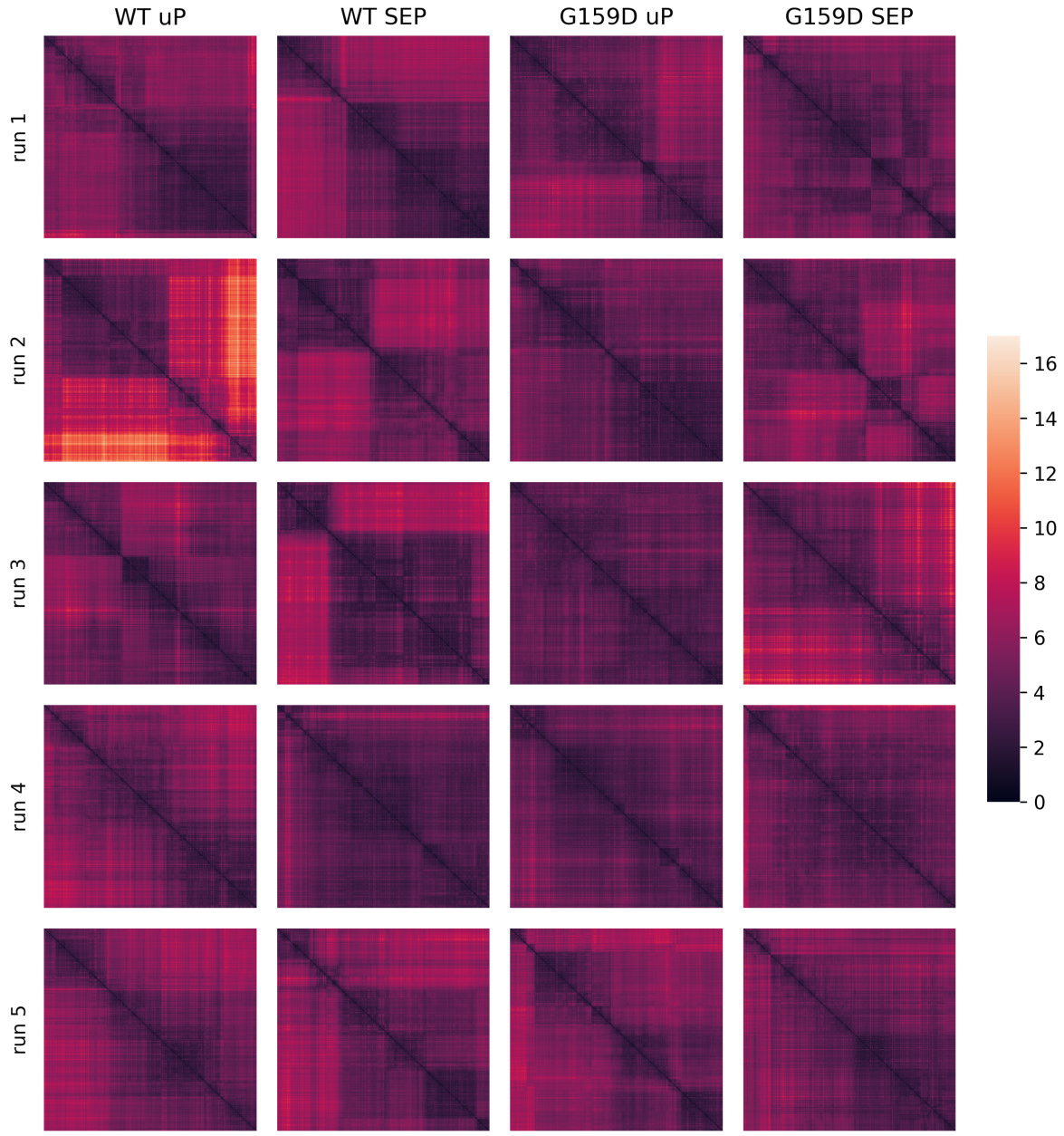


Figure S1: Pairwise protein backbone root-mean-square deviation (RMSD) between all frames of each trajectory for all apo systems. No obvious burn-in period was observed.

7.2 Chapter 3

		WT uP	WT SEP	G159D uP	G159D SEP
TnT Arg216	TnI Asp104	77%	85%	97%	93%
TnT Lys247	TnI Glu109	96%	95%	94%	94%
TnT Asp249	TnI Lys71	88%	93%	92%	93%
TnT Glu214	TnI Arg97	63%	68%	87%	69%
TnT Asp222	TnI Arg97	69%	63%	81%	62%
TnT Glu244	TnI Arg102	70%	65%	49%	60%
TnT Glu252	TnI Arg67	62%	60%	64%	62%
TnT Glu244	TnI Arg78	41%	39%	57%	43%
TnT Arg205	TnI Glu108	13%	36%	17%	34%
TnT Arg265	TnI Glu123	37%	32%	38%	32%
TnT Arg205	TnI Asp107	10%	31%	15%	25%
TnT Glu246	TnI Arg110	28%	19%	23%	29%
TnT Lys210	TnI Glu108	9%	21%	28%	17%
TnT Lys209	TnI Asp107	11%	16%	27%	22%
TnT Arg278	TnI Asp133	17%	11%	25%	10%
TnT Lys209	TnI Glu108	18%	18%	10%	24%

Table S1: Ionic interactions between TnT and TnI.

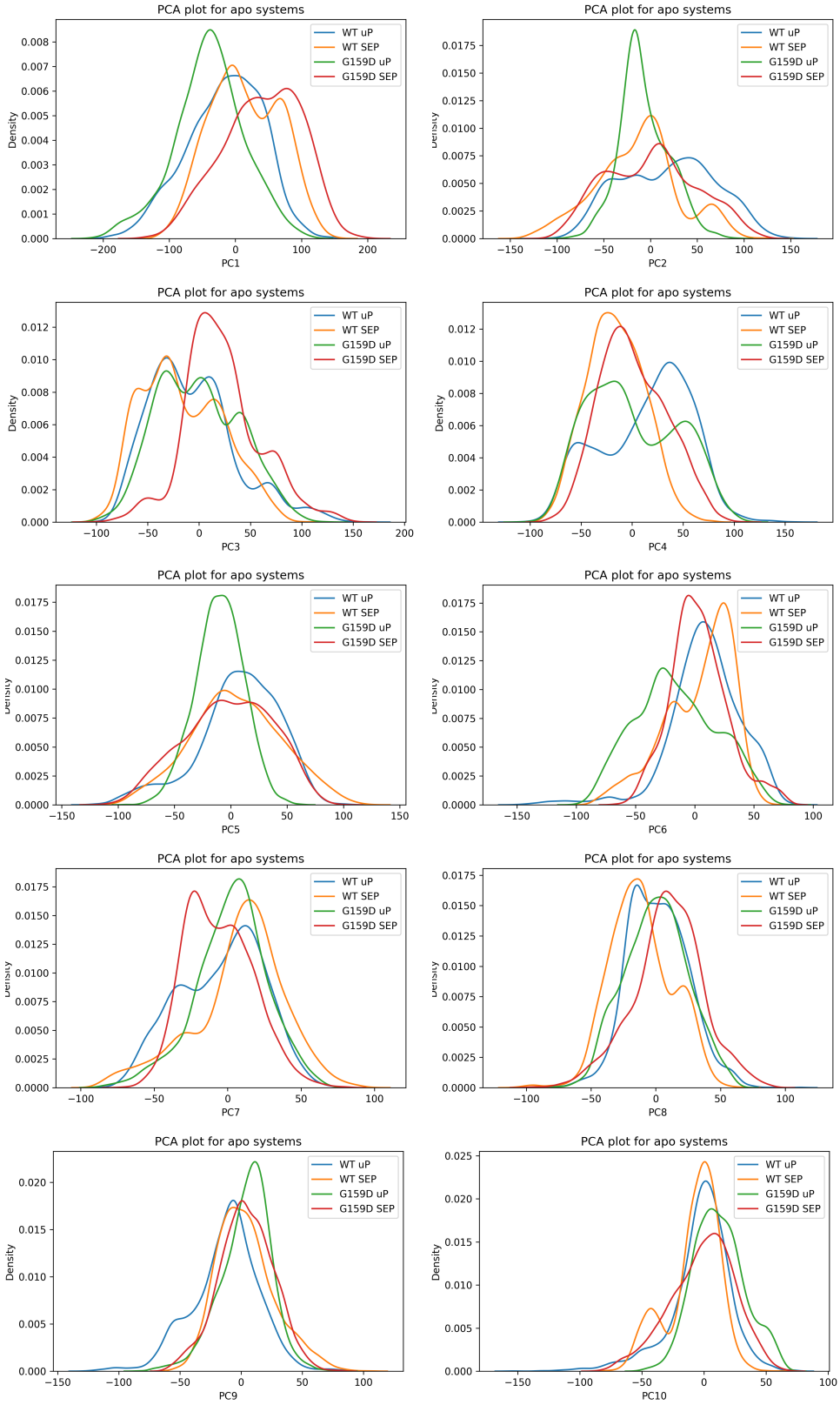


Figure S2: Distributions of simulated conformations for the apo systems along the first ten principal components as KDE plots.

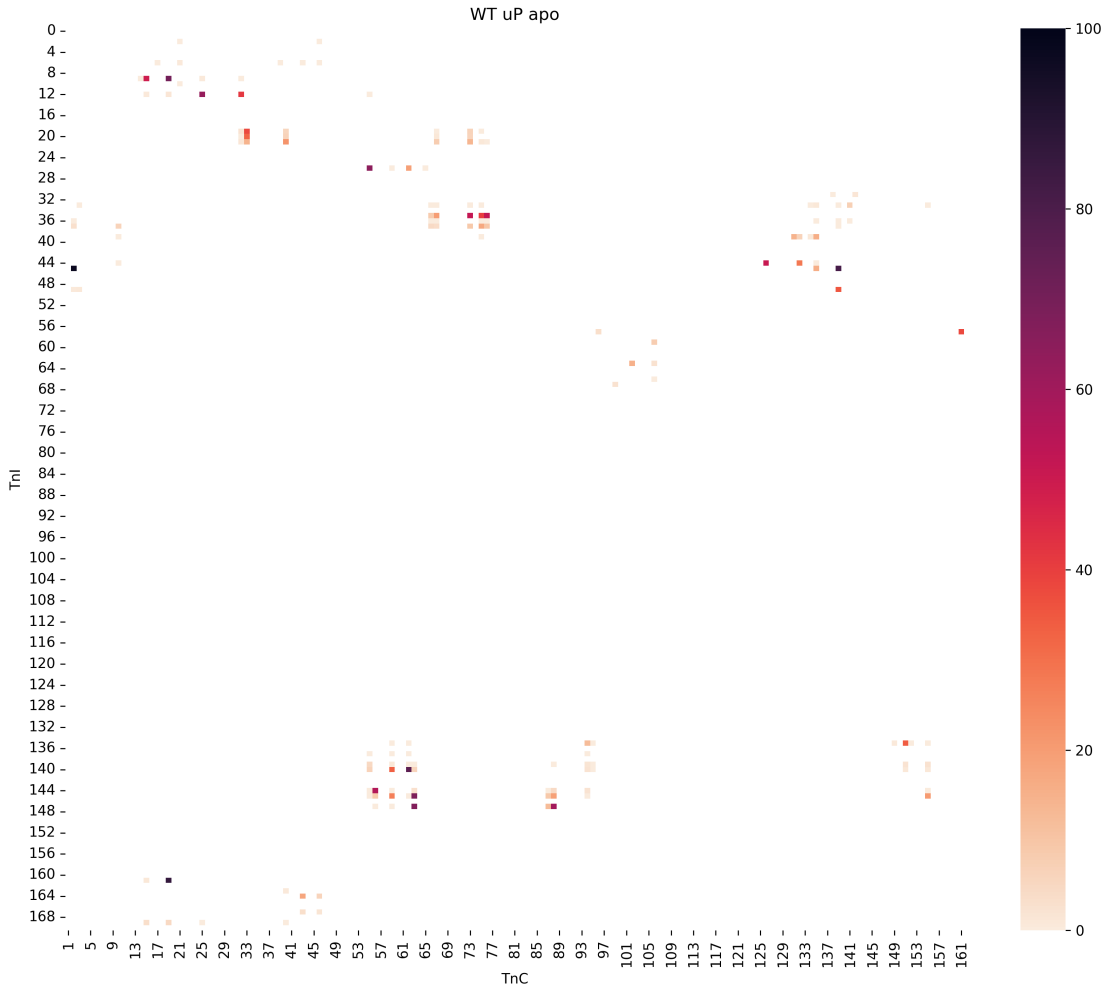


Figure S3: Heatmap showing the ionic interaction pattern between TnC and TnI for WT uP apo system. The interaction count is expressed as a percentage of the total number of frames.

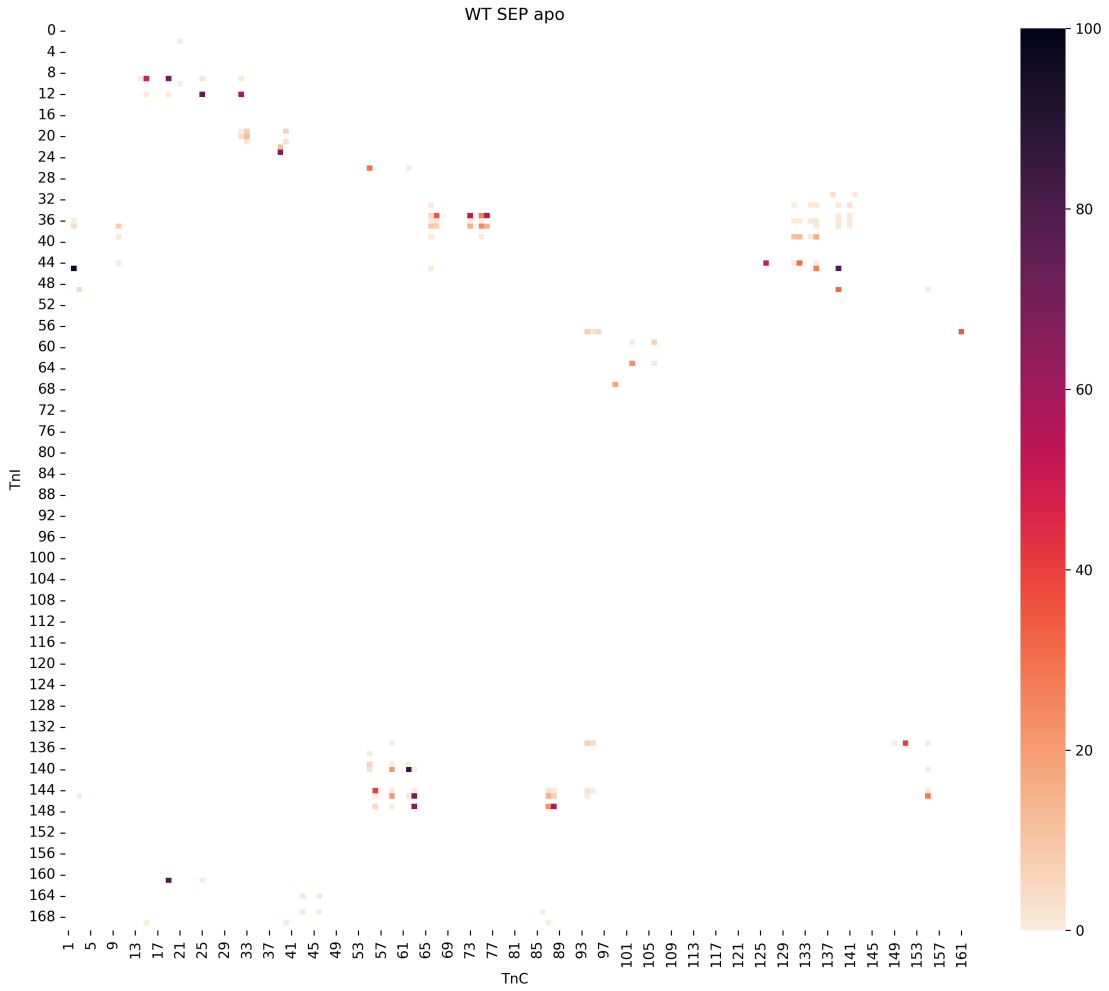


Figure S4: Heatmap showing the ionic interaction pattern between TnC and TnI for WT SEP apo system. The interaction count is expressed as a percentage of the total number of frames.

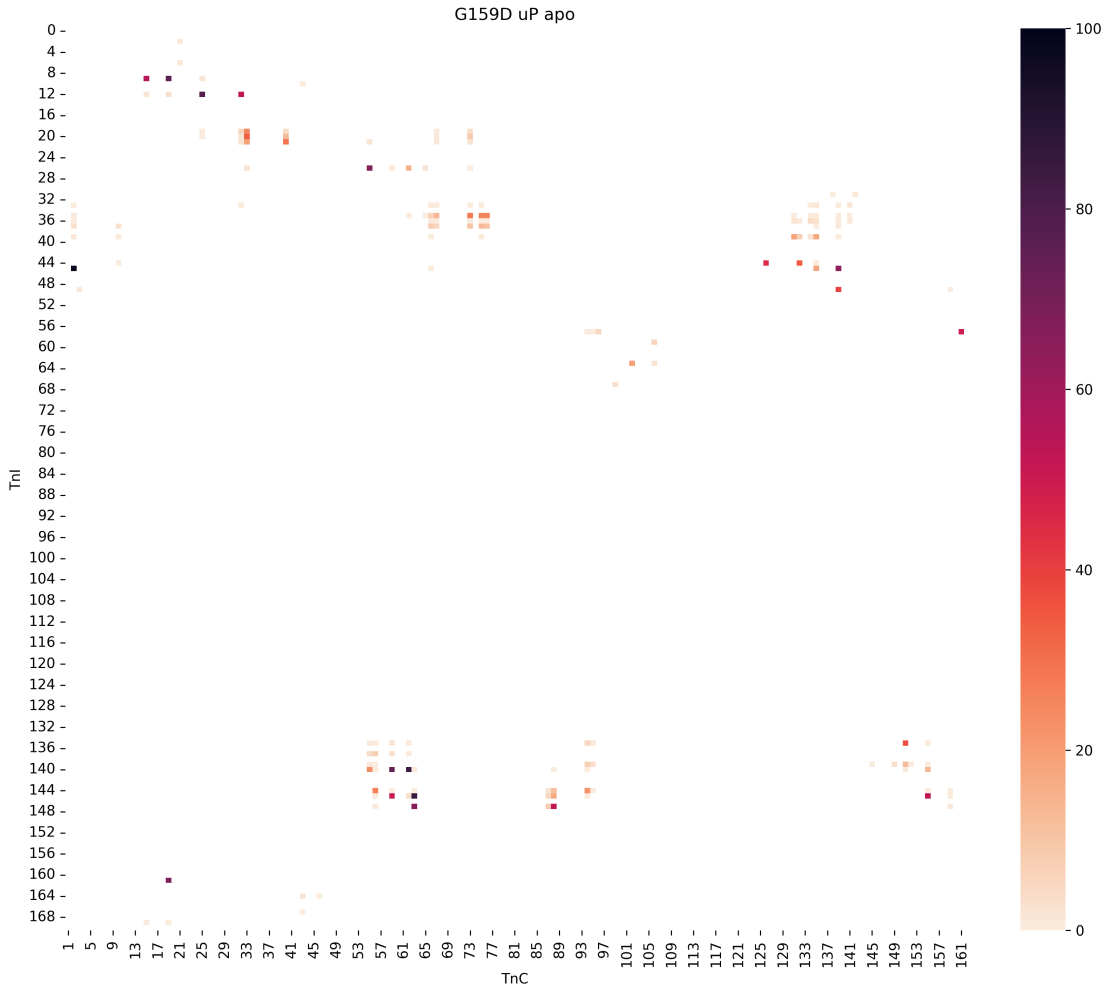


Figure S5: Heatmap showing the ionic interaction pattern between TnC and TnI for G159D uP apo system. The interaction count is expressed as a percentage of the total number of frames.

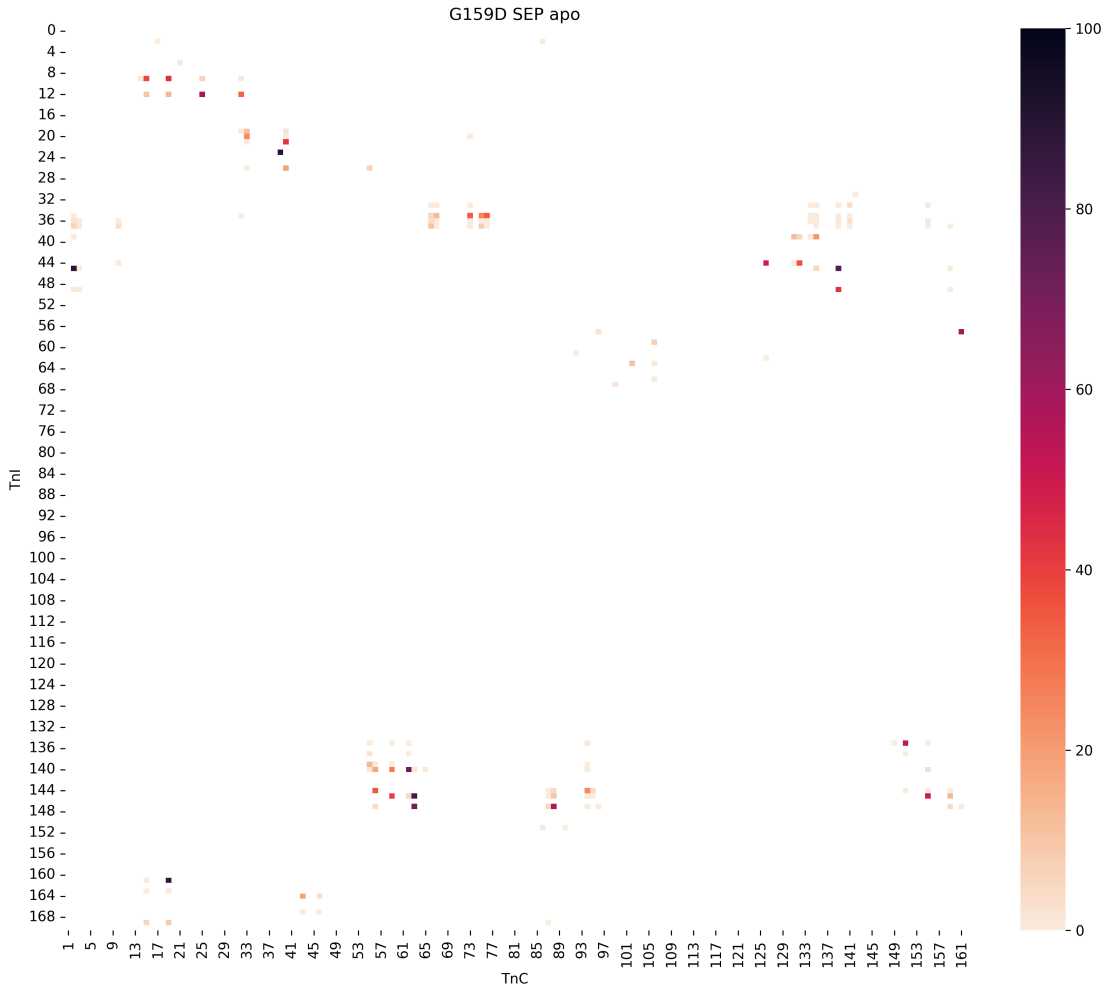


Figure S6: Heatmap showing the ionic interaction pattern between TnC and TnI for G159D SEP apo system. The interaction count is expressed as a percentage of the total number of frames.

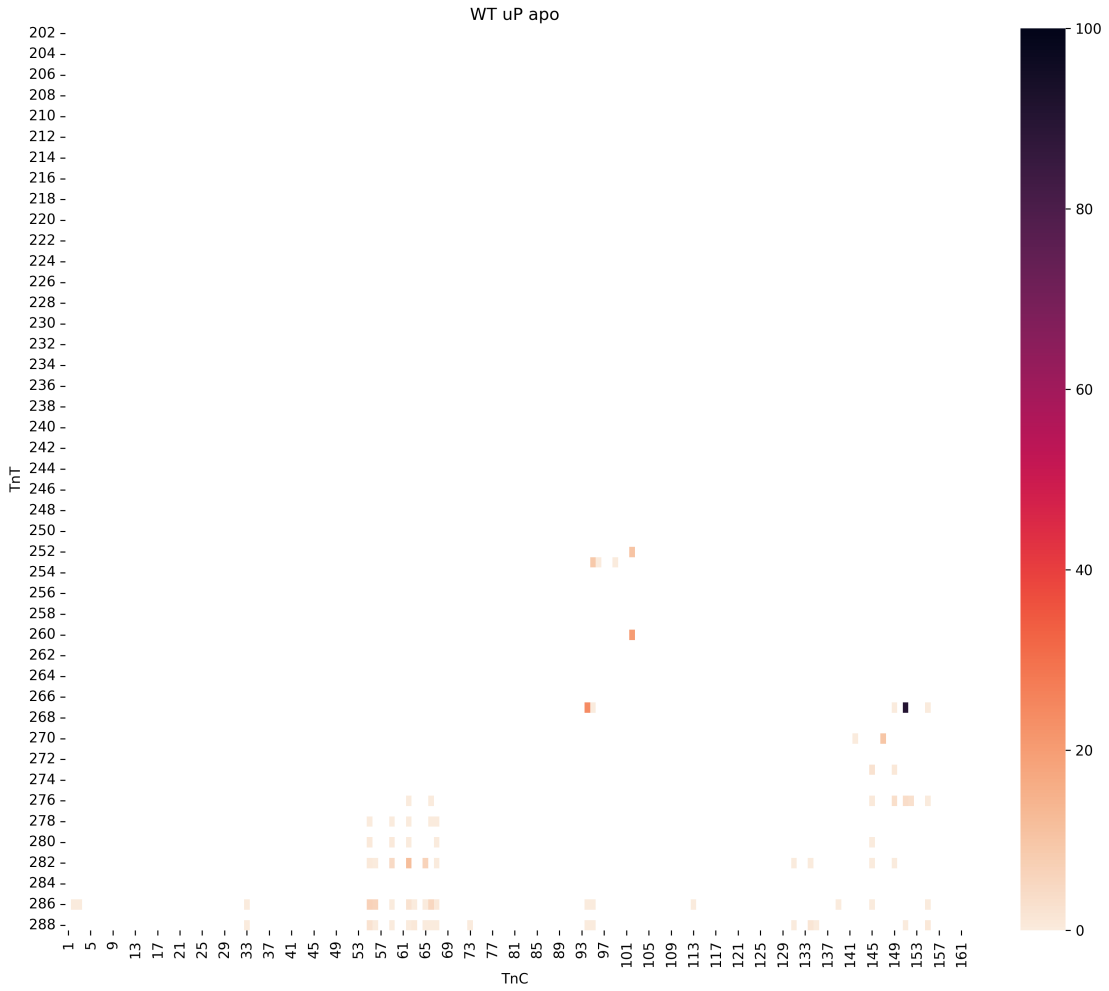


Figure S7: Heatmap showing the ionic interaction pattern between TnC and TnT for WT uP apo system. The interaction count is expressed as a percentage of the total number of frames.

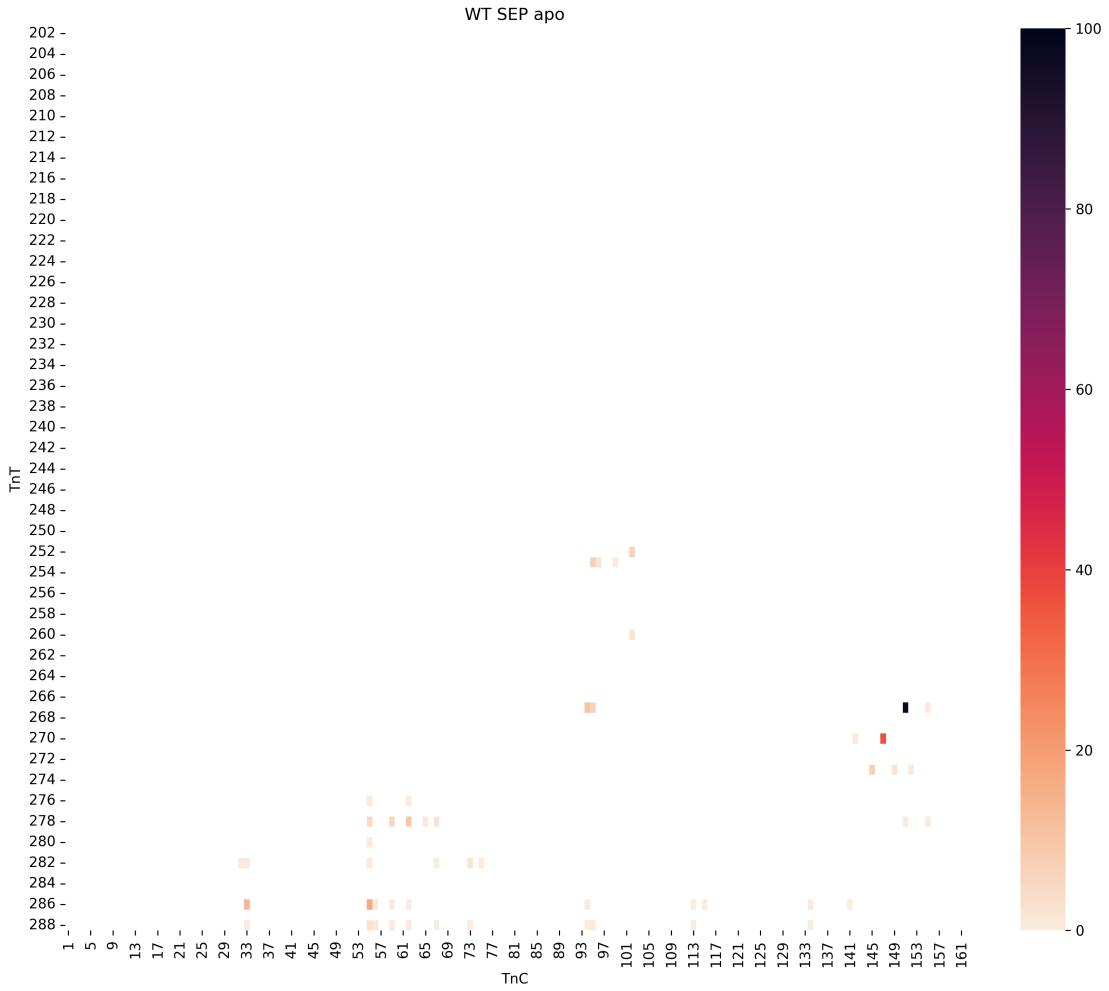


Figure S8: Heatmap showing the ionic interaction pattern between TnC and TnT for WT SEP apo system. The interaction count is expressed as a percentage of the total number of frames.

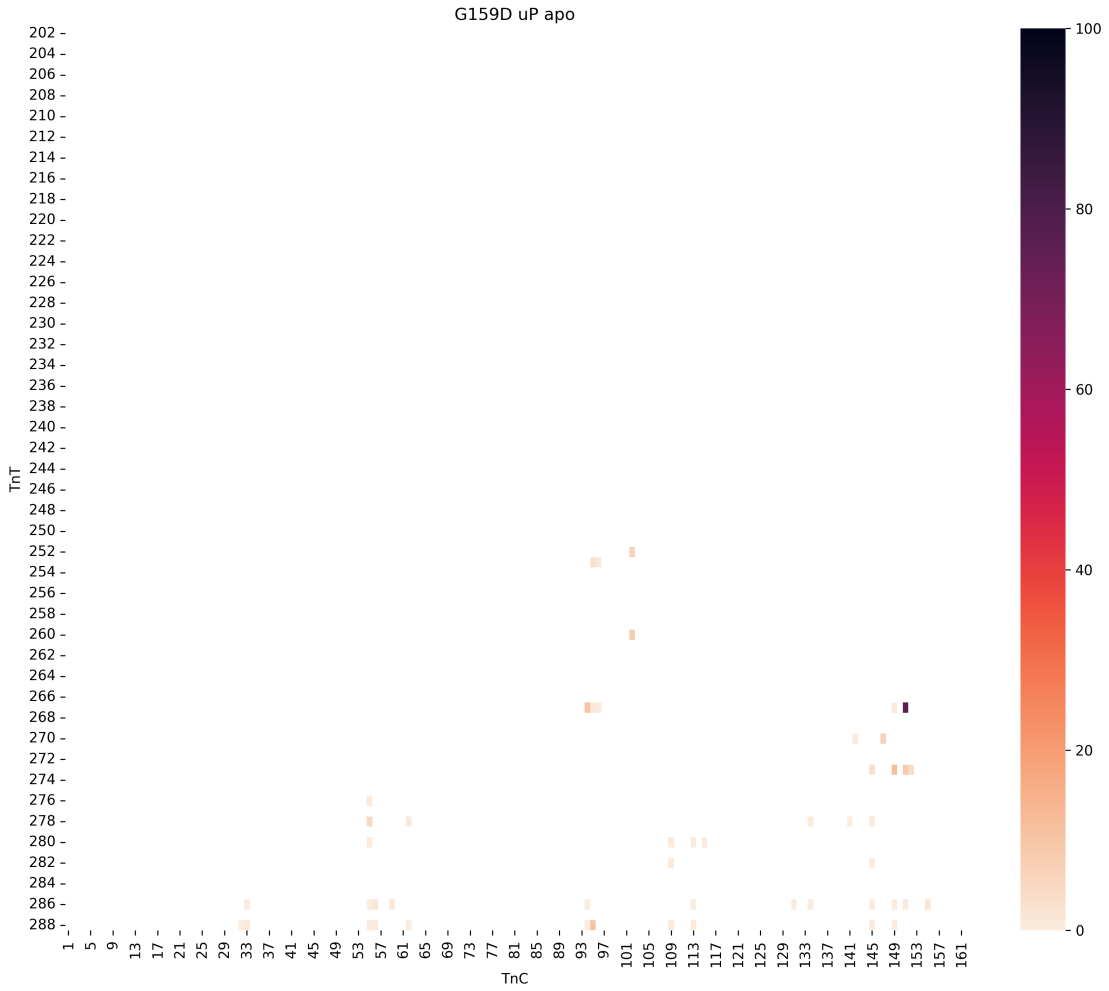


Figure S9: Heatmap showing the ionic interaction pattern between TnC and TnT for G159D uP apo system. The interaction count is expressed as a percentage of the total number of frames.

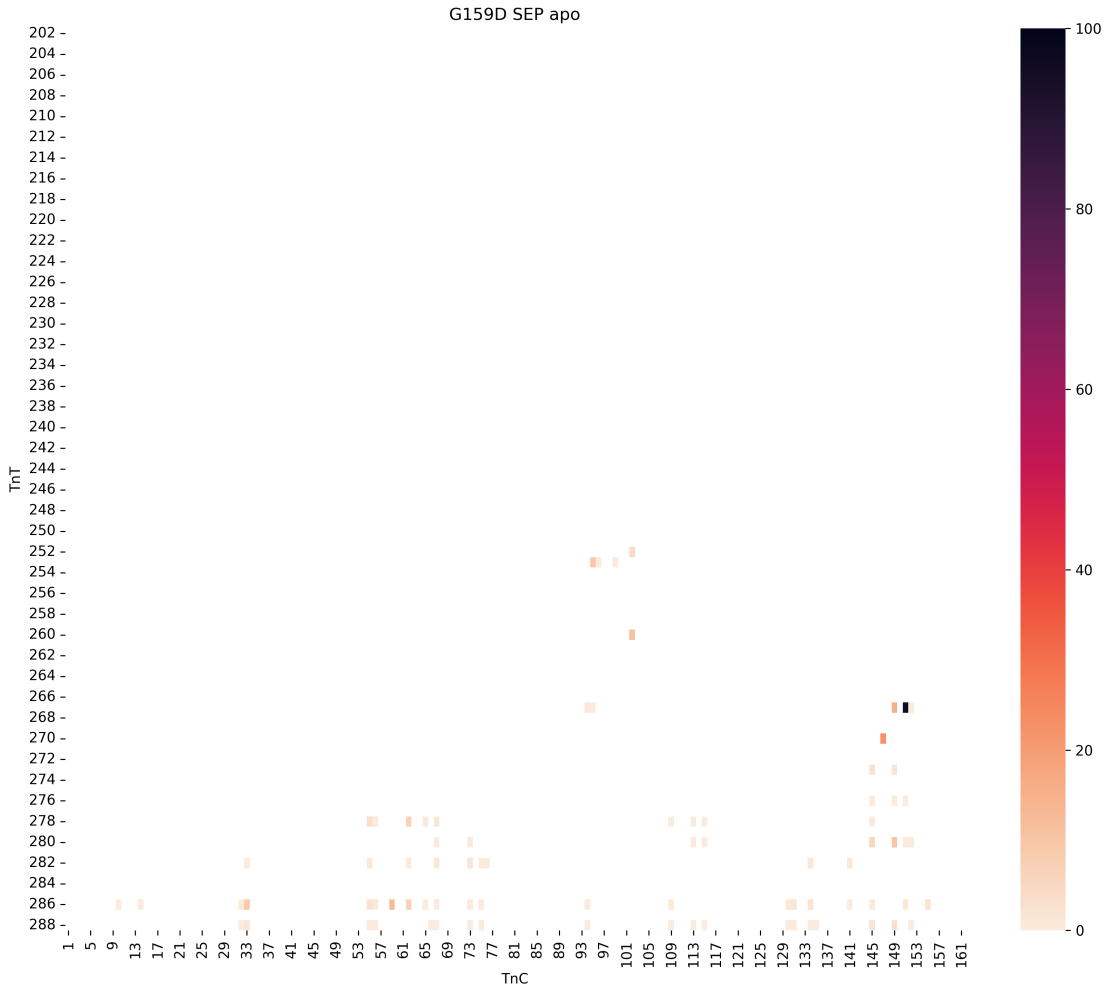


Figure S10: Heatmap showing the ionic interaction pattern between TnC and TnT for G159D SEP apo system. The interaction count is expressed as a percentage of the total number of frames.

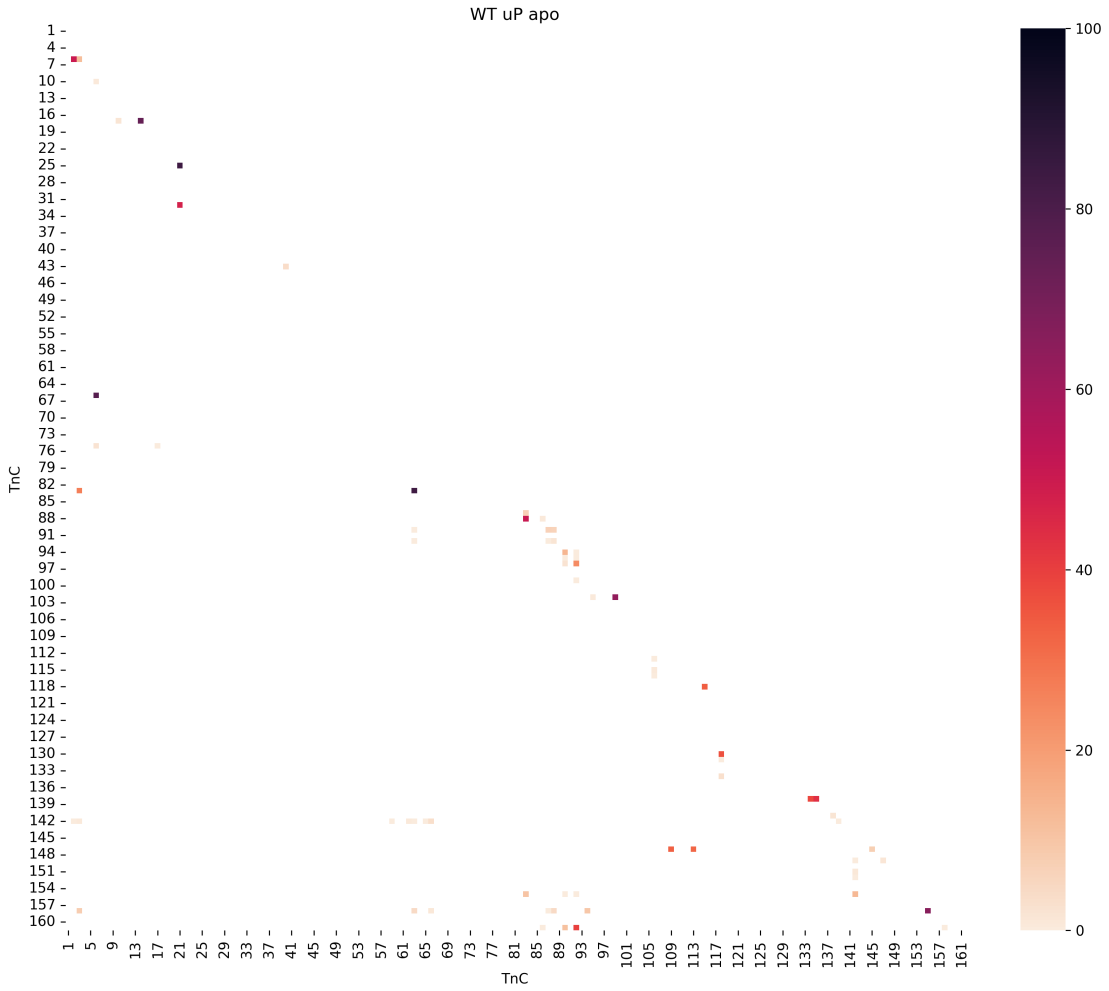


Figure S11: Heatmap showing the ionic interaction pattern between TnC and TnC for WT uP apo system. The interaction count is expressed as a percentage of the total number of frames.

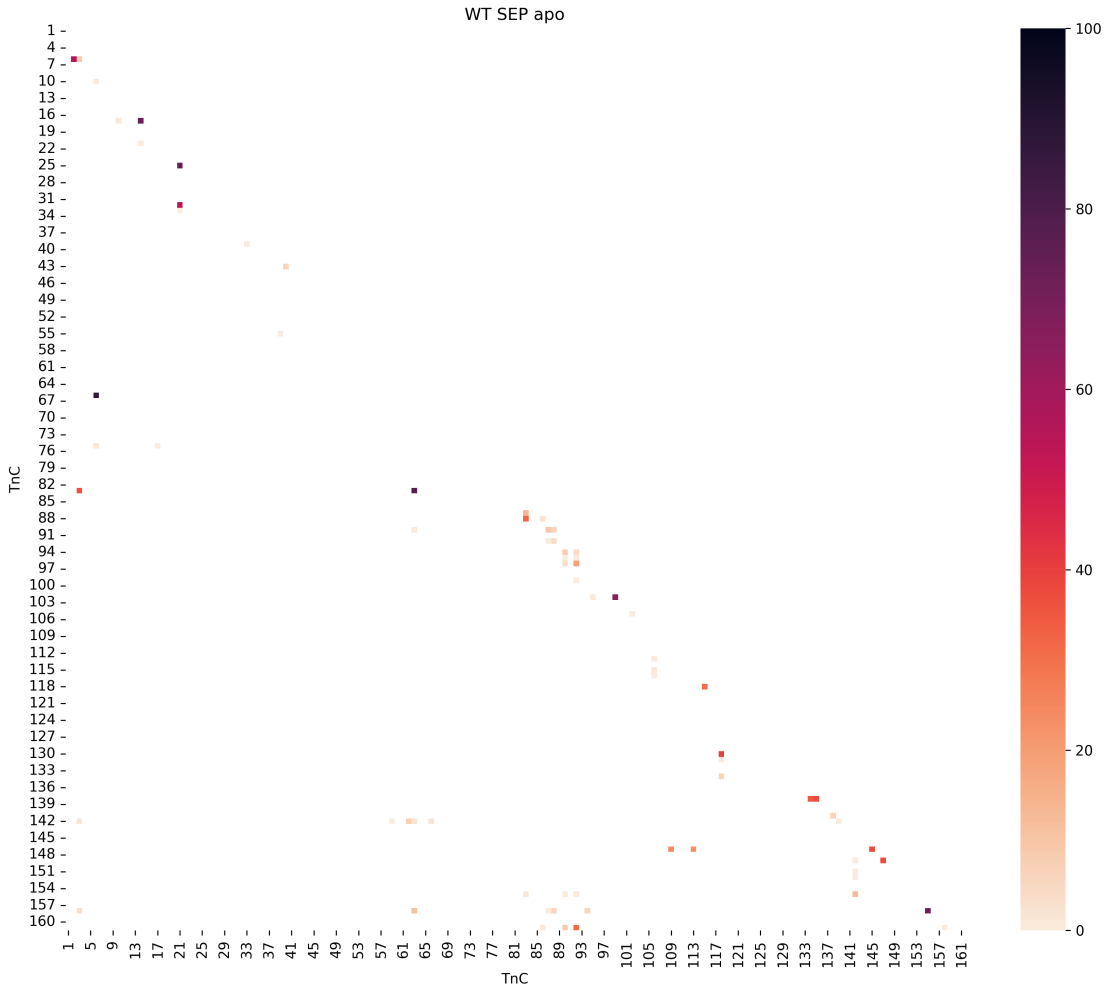


Figure S12: Heatmap showing the ionic interaction pattern between TnC and TnC for WT SEP apo system. The interaction count is expressed as a percentage of the total number of frames.

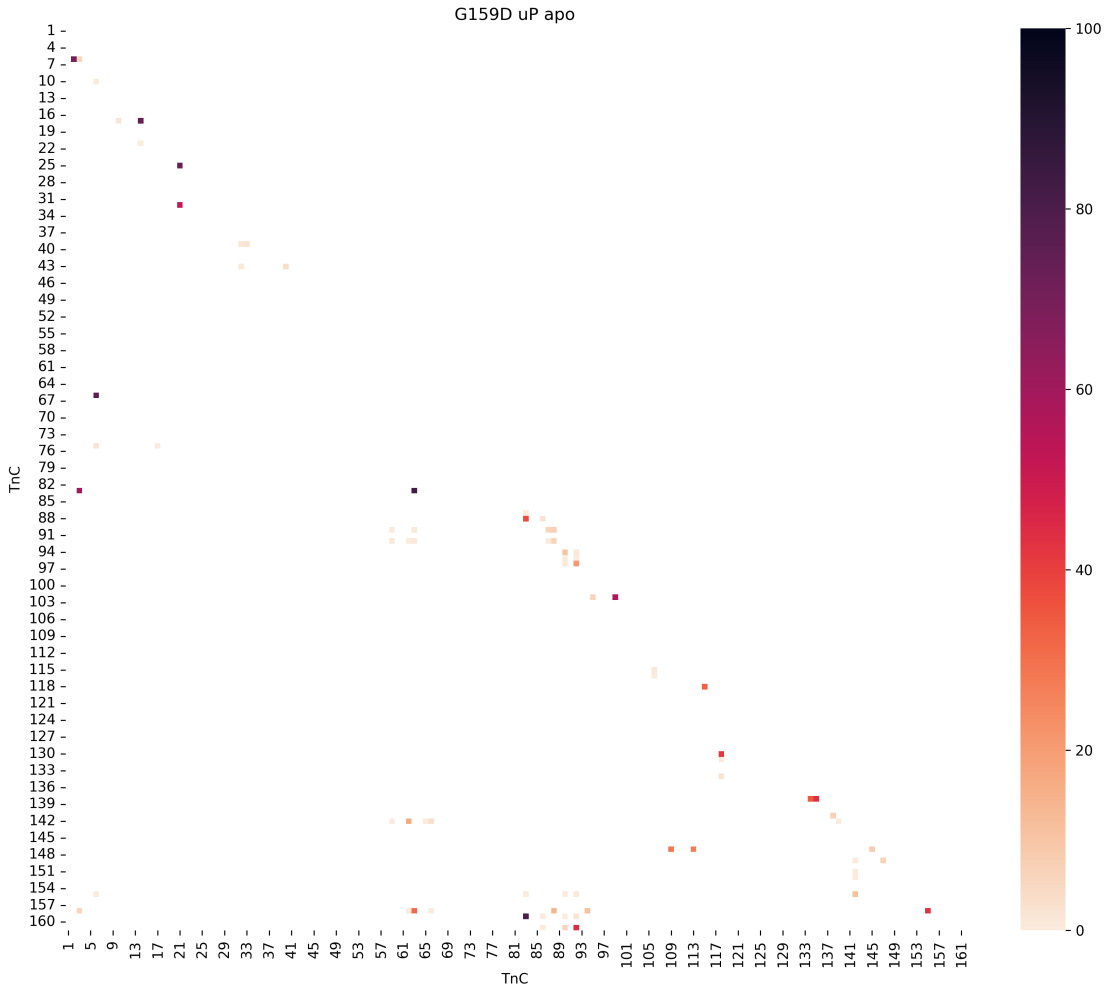


Figure S13: Heatmap showing the ionic interaction pattern between TnC and TnC for G159D uP apo system. The interaction count is expressed as a percentage of the total number of frames.

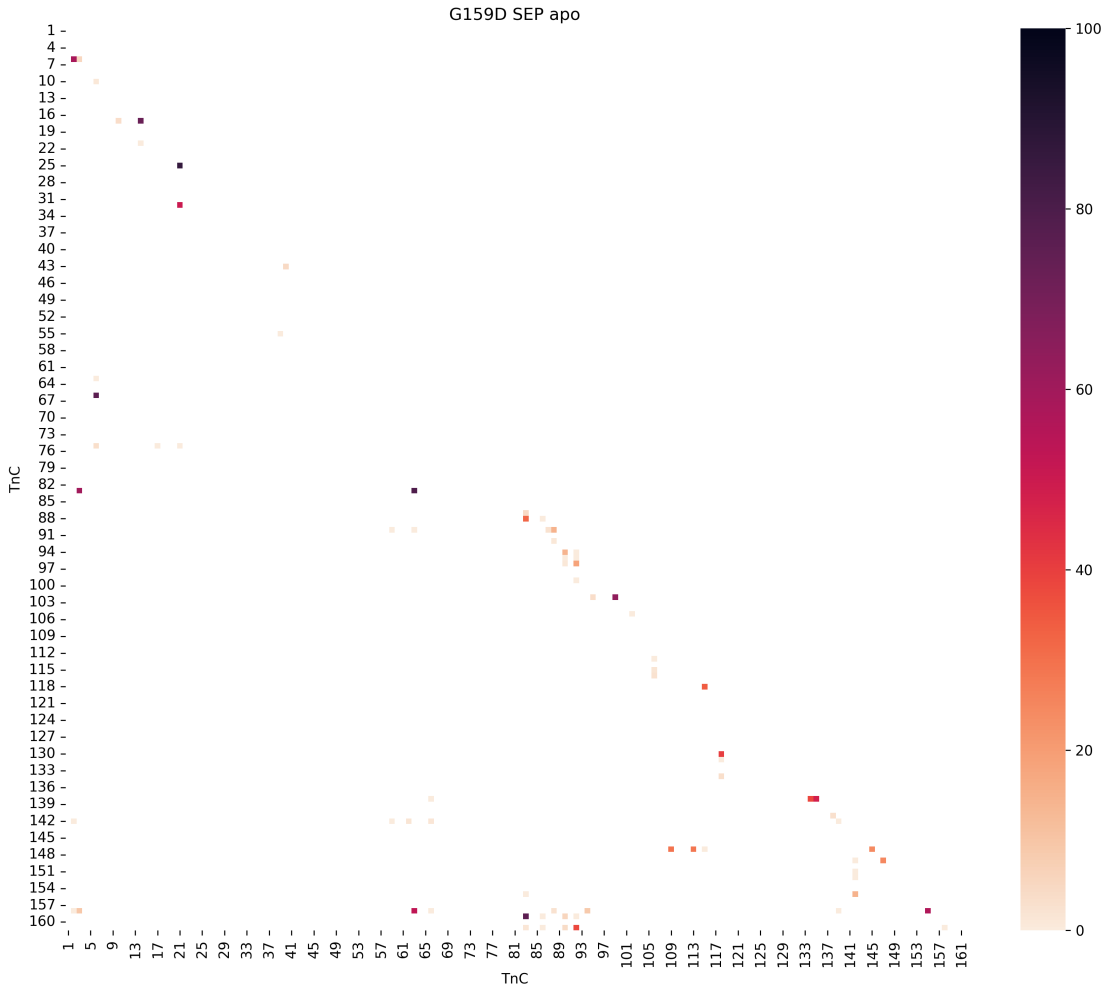


Figure S14: Heatmap showing the ionic interaction pattern between TnC and TnC for G159D SEP apo system. The interaction count is expressed as a percentage of the total number of frames.

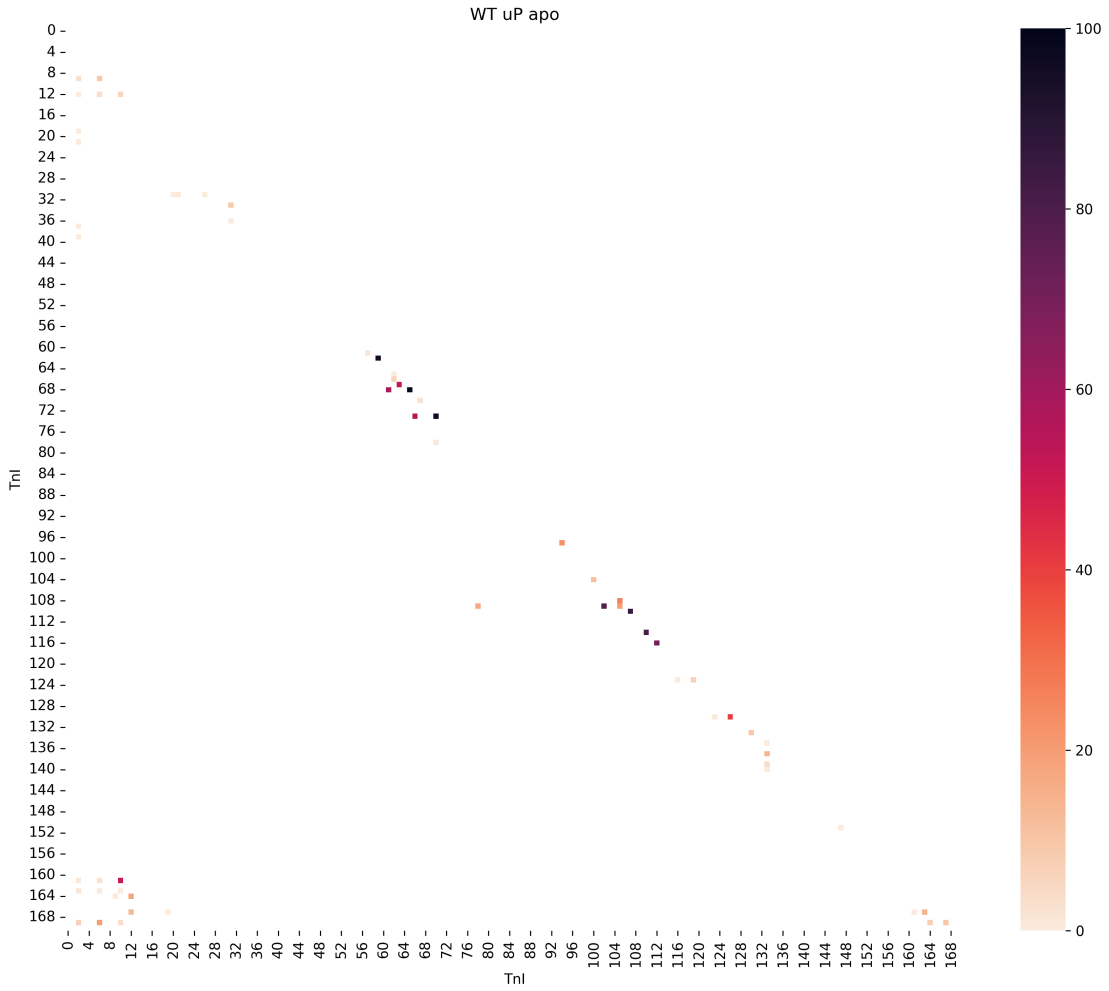


Figure S15: Heatmap showing the ionic interaction pattern between TnI and TnI for WT uP apo system. The interaction count is expressed as a percentage of the total number of frames.

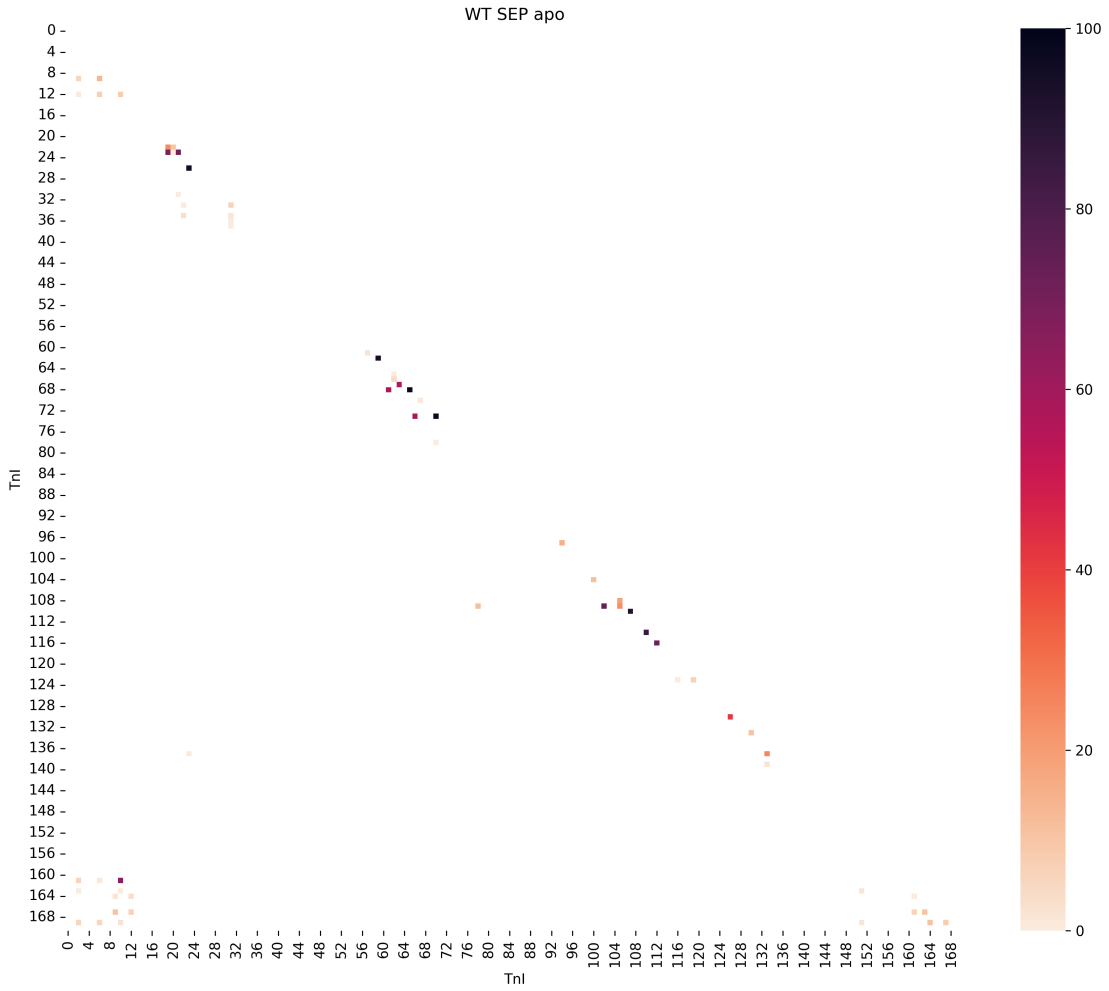


Figure S16: Heatmap showing the ionic interaction pattern between TnI and TnI for WT SEP apo system. The interaction count is expressed as a percentage of the total number of frames.

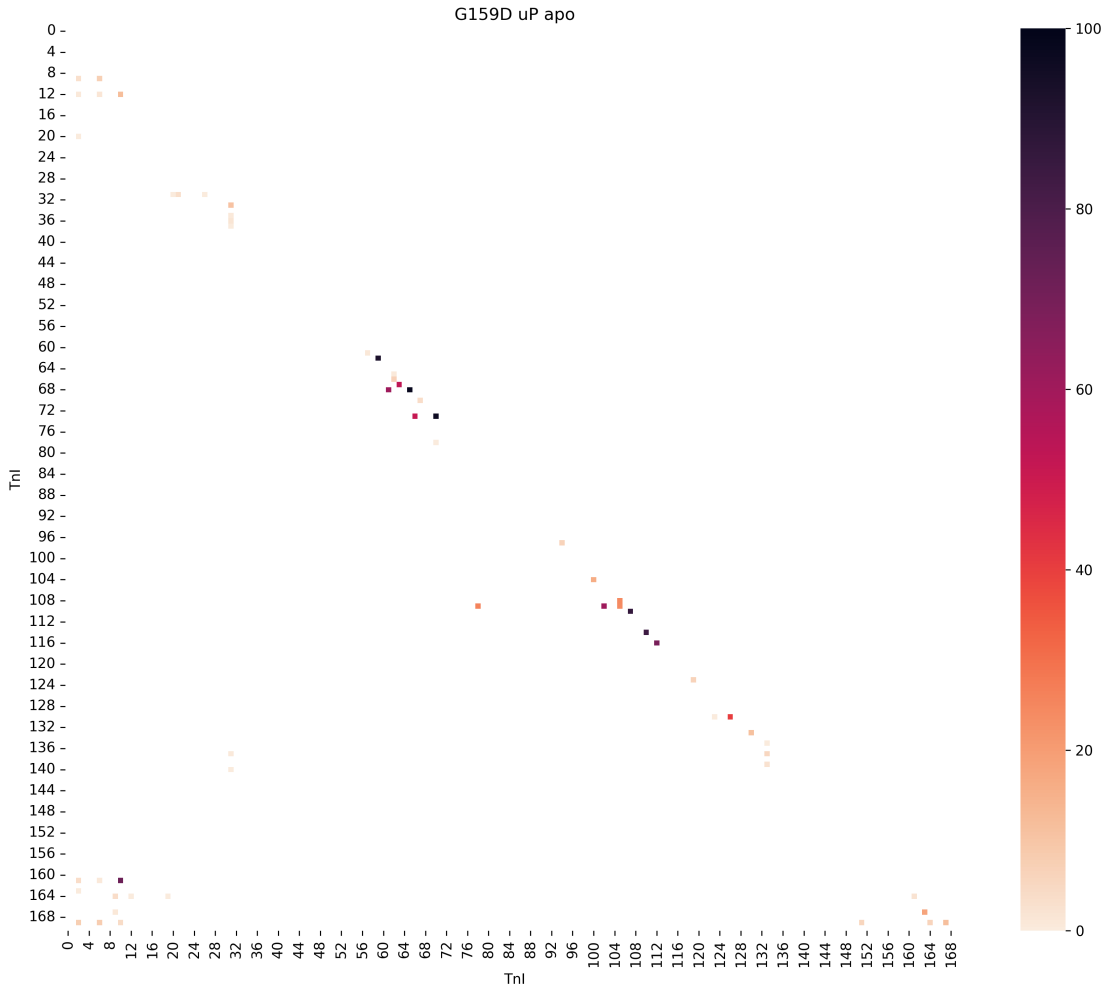


Figure S17: Heatmap showing the ionic interaction pattern between TnI and TnI for G159D uP apo system. The interaction count is expressed as a percentage of the total number of frames.

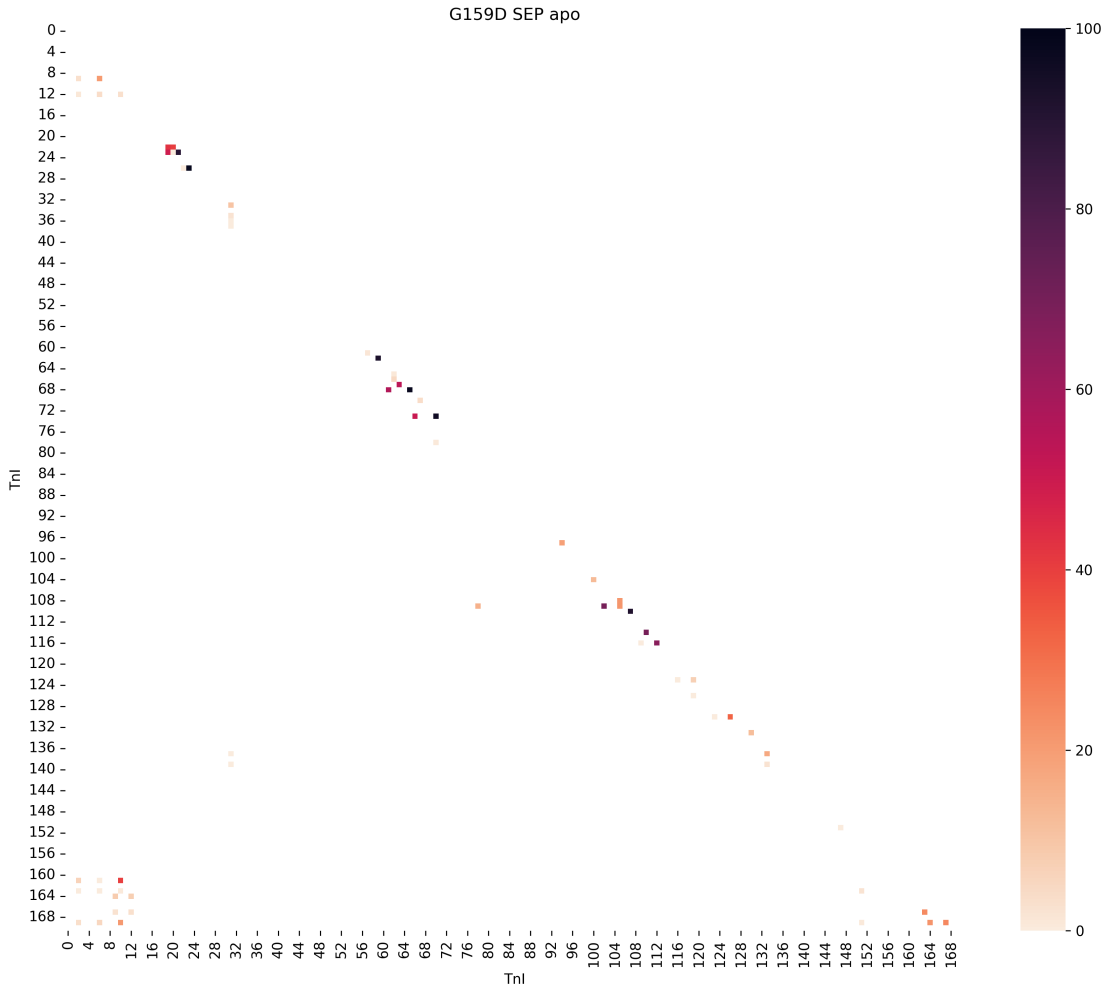


Figure S18: Heatmap showing the ionic interaction pattern between TnI and TnI for G159D SEP apo system. The interaction count is expressed as a percentage of the total number of frames.

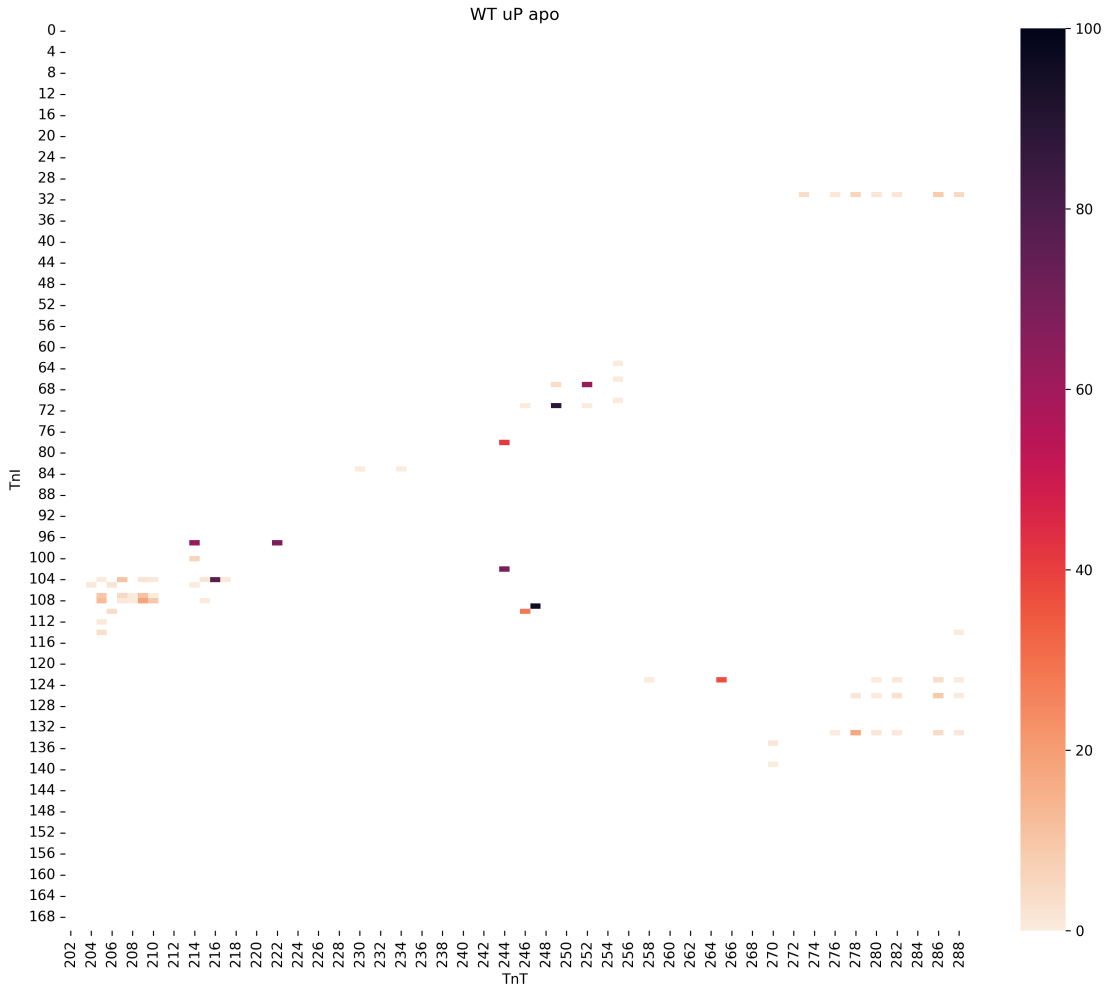


Figure S19: Heatmap showing the ionic interaction pattern between TnT and TnI for WT uP apo system. The interaction count is expressed as a percentage of the total number of frames.

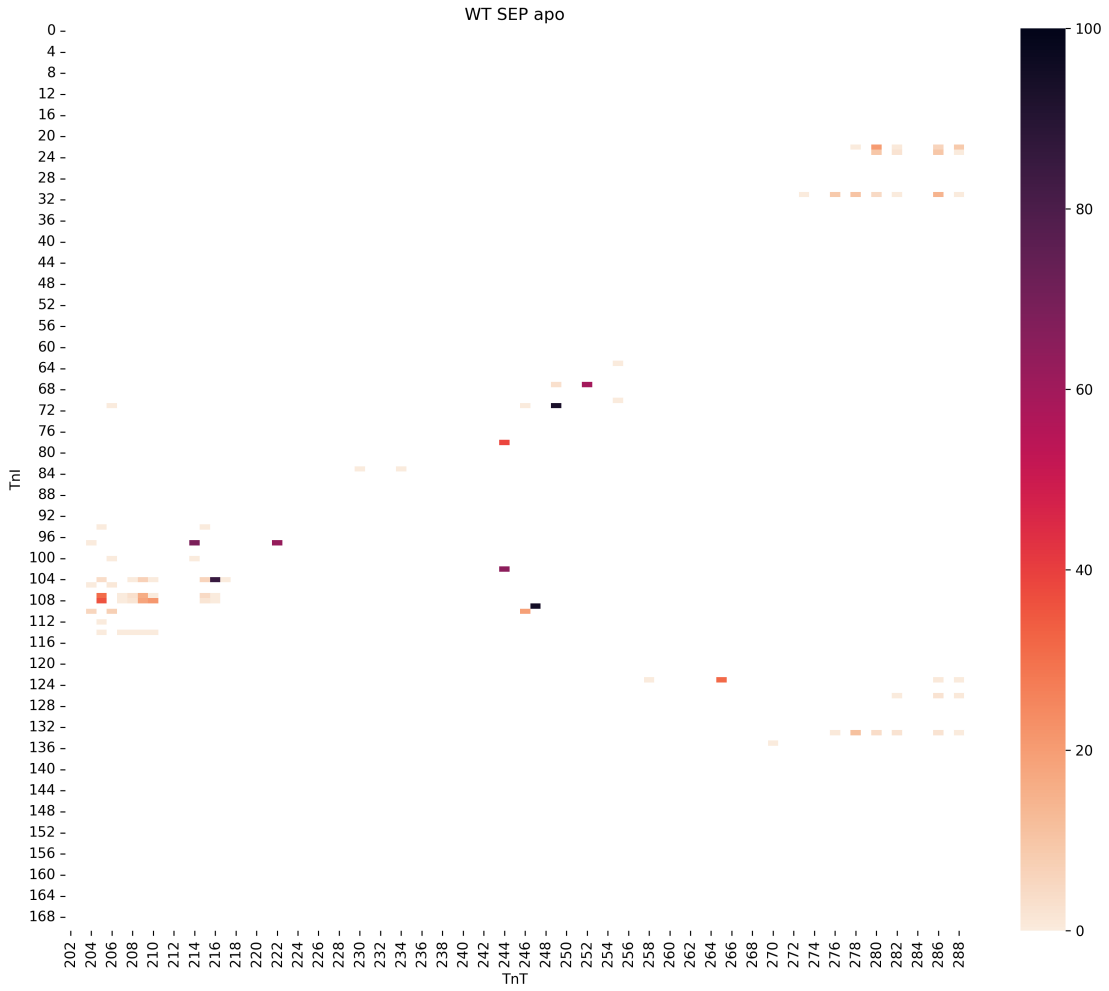


Figure S20: Heatmap showing the ionic interaction pattern between TnT and TnI for WT SEP apo system. The interaction count is expressed as a percentage of the total number of frames.

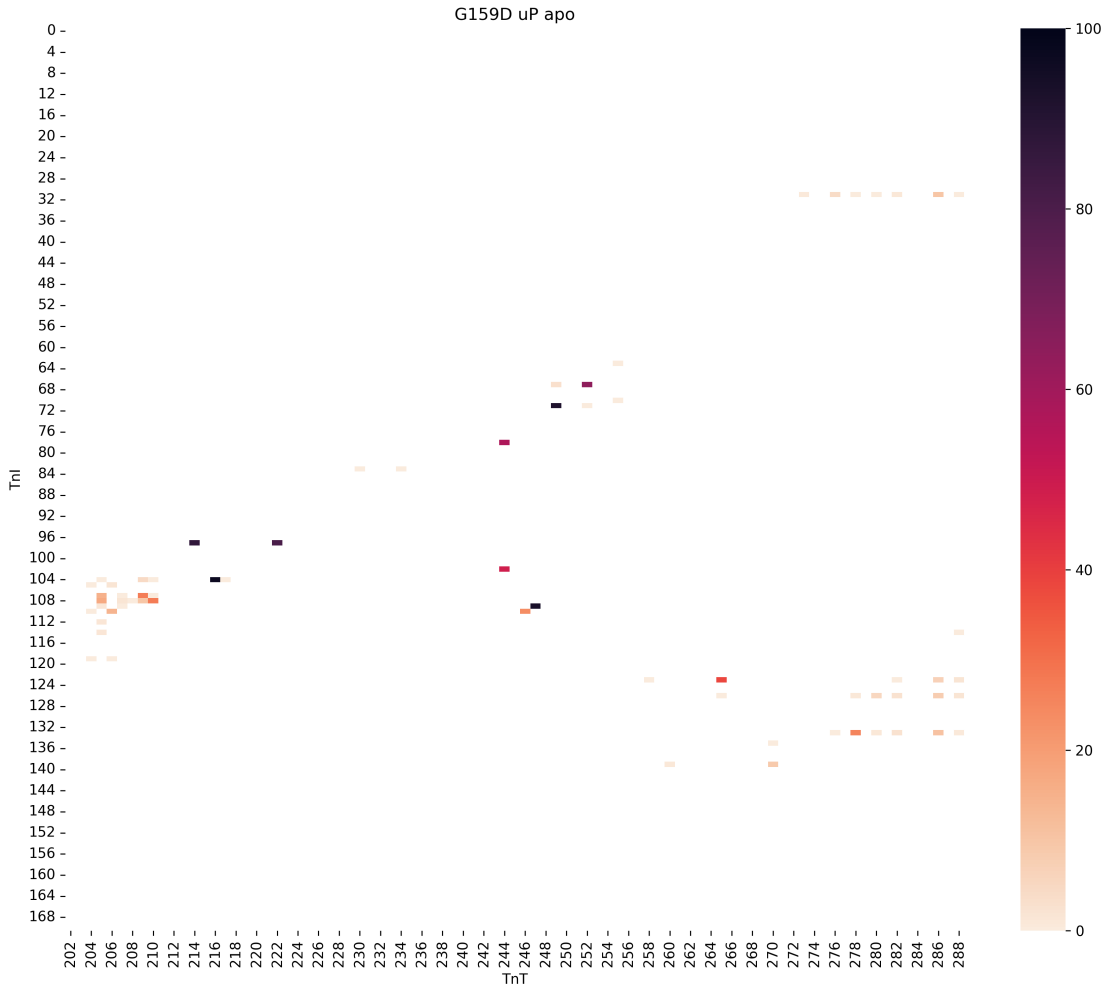


Figure S21: Heatmap showing the ionic interaction pattern between TnT and TnI for G159D uP apo system. The interaction count is expressed as a percentage of the total number of frames.

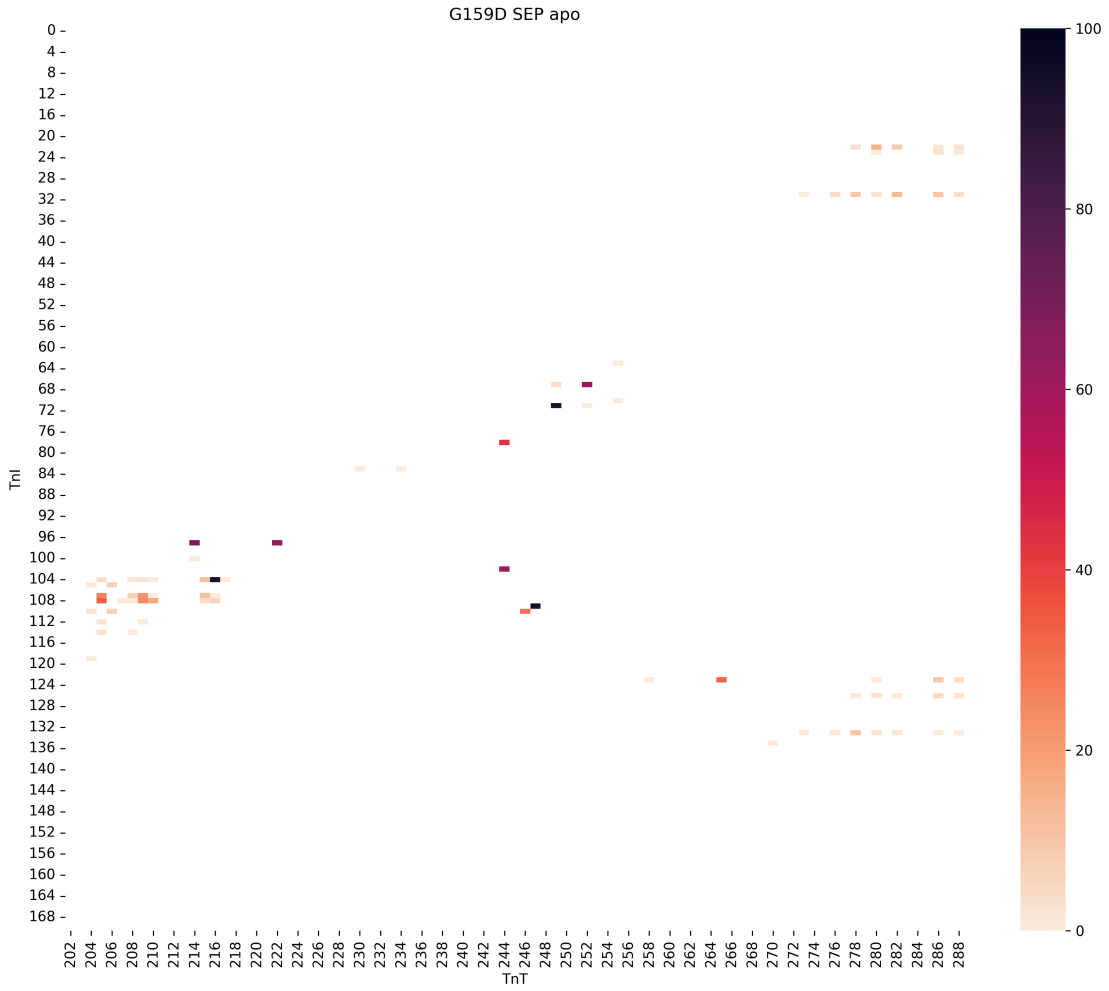


Figure S22: Heatmap showing the ionic interaction pattern between TnT and TnI for G159D SEP apo system. The interaction count is expressed as a percentage of the total number of frames.

References

(1) J. Sarko and C. V. Pollack, *The Journal of Emergency Medicine*, 2002, **23**, 57–65, DOI: 10.1016/S0736-4679(02)00463-8.

(2) Y. Cheng and M. Regnier, *Archives of Biochemistry and Biophysics*, 2016, **601**, 11–21, DOI: 10.1016/j.abb.2016.02.004.

(3) C. C. Glembotski, *Physiology*, 2012, **27**, 343–350, DOI: 10.1152/physiol.00034.2012.

(4) M. Michalak and M. Opas, *Trends in Cell Biology*, 2009, **19**, 253–259, DOI: 10.1016/j.tcb.2009.03.006.

(5) A. Fabiato, *American Journal of Physiology-Cell Physiology*, 1983, **245**, C1–C14, DOI: 10.1152/ajpcell.1983.245.1.C1.

(6) M. Watanabe, Y. Yoshino and S. Morimoto, *Biochemical and Biophysical Research Communications*, 2003, **307**, 236–240, DOI: 10.1016/S0006-291X(03)01170-7.

(7) V. B. Patchell, C. E. Gallon, M. A. Hodgkin, A. Fattoum, S. V. Perry and B. A. Levine, *European Journal of Biochemistry*, 2002, **269**, 5088–5100, DOI: 10.1046/j.1432-1033.2002.03227.x.

(8) D. M. Bers, *Circulation Research*, 2000, **87**, 275–281, DOI: 10.1161/01.RES.87.4.275.

(9) H. L. Kirschenlohr, A. A. Grace, J. I. Vandenberg, J. C. Metcalfe and G. A. Smith, *Biochemical Journal*, 2000, **346**, 385–391.

(10) B. D. Sykes, *Nature Structural & Molecular Biology*, 2003, **10**, 588–589, DOI: 10.1038/nsb0803-588.

(11) R. Dominguez, *Biophysical Journal*, 2011, **100**, 797–798, DOI: 10.1016/j.bpj.2011.01.018.

(12) S. Marston and J. E. Zamora, *Journal of Muscle Research and Cell Motility*, 2020, **41**, 71–89, DOI: 10.1007/s10974-019-09513-1.

(13) D. G. Vassylyev, S. Takeda, S. Wakatsuki, K. Maeda and Y. Maéda, *Proceedings of the National Academy of Sciences*, 1998, **95**, 4847–4852, DOI: 10.1073/pnas.95.9.4847.

(14) A. Y. Li, J. Lee, D. Borek, Z. Otwinowski, G. F. Tibbits and M. Paetzel, *Journal of Molecular Biology*, 2011, **413**, 699–711, DOI: 10.1016/j.jmb.2011.08.049.

(15) X. L. Zhang, G. F. Tibbits and M. Paetzel, *Acta Crystallographica Section D Biological Crystallography*, 2013, **69**, 722–734, DOI: 10.1107/S0907444913001182.

(16) S. Takeda, A. Yamashita, K. Maeda and Y. Maéda, *Nature*, 2003, **424**, 35–41, DOI: 10.1038/nature01780.

(17) Y. Yamada, K. Namba and T. Fujii, *Nature Communications*, 2020, **11**, 153, DOI: 10.1038/s41467-019-14008-1.

(18) T. Oda, H. Yanagisawa and T. Wakabayashi, *Journal of Structural Biology*, 2020, **209**, 107450, DOI: 10.1016/j.jsb.2020.107450.

(19) W. Lehman, E. Pavadai and M. J. Rynkiewicz, *Biochemical and Biophysical Research Communications*, 2021, **551**, 27–32, DOI: 10.1016/j.bbrc.2021.03.010.

(20) E. Pavadai, M. J. Rynkiewicz, Z. Yang, I. R. Gould, S. B. Marston and W. Lehman, *Archives of Biochemistry and Biophysics*, 2022, **725**, 109282, DOI: 10.1016/j.abb.2022.109282.

(21) A. M. Cool and S. Lindert, *The Journal of Physical Chemistry B*, 2021, **125**, 7388–7396, DOI: 10.1021/acs.jpcb.1c03844.

(22) J. W. Howarth, J. Meller, R. J. Solaro, J. Trewhella and P. R. Rosevear, *Journal of Molecular Biology*, 2007, **373**, 706–722, DOI: 10.1016/j.jmb.2007.08.035.

(23) P. M. Hwang, F. Cai, S. E. Pineda-Sanabria, D. C. Corson and B. D. Sykes, *Proceedings of the National Academy of Sciences of the United States of America*, 2014, **111**, 14412–14417, DOI: 10.1073/pnas.1410775111.

(24) S. E. Pineda-Sanabria, I. M. Robertson and B. D. Sykes, *Biochemistry*, 2015, **54**, 3583–3593, DOI: 10.1021/acs.biochem.5b00178.

(25) F. Cai, M. X. Li, S. E. Pineda-Sanabria, S. Gelozia, S. Lindert, F. West, B. D. Sykes and P. M. Hwang, *Journal of Molecular and Cellular Cardiology*, 2016, **101**, 134–144, DOI: 10.1016/j.yjmcc.2016.10.016.

(26) S. E. Pineda-Sanabria, I. M. Robertson and B. D. Sykes, *Biochemistry*, 2011, **50**, 1309–1320, DOI: 10.1021/bi101985j.

(27) M. X. Li and P. M. Hwang, *Gene*, 2015, **571**, 153–166, DOI: 10.1016/j.gene.2015.07.074.

(28) V. Zabrouskov, Y. Ge, J. Schwartz and J. W. Walker, *Molecular & Cellular Proteomics*, 2008, **7**, 1838–1849, DOI: 10.1074/mcp.M700524-MCP200.

(29) T. Kobayashi, S. E. Patrick and M. Kobayashi, *The Journal of Biological Chemistry*, 2009, **284**, 20052–20060, DOI: 10.1074/jbc.M109.001396.

(30) J.-P. Jin and S. M. Chong, *Archives of biochemistry and biophysics*, 2010, **500**, 144–150, DOI: 10.1016/j.abb.2010.06.001.

(31) J. C. Kentish, D. T. McCloskey, J. Layland, S. Palmer, J. M. Leiden, A. F. Martin and R. J. Solaro, *Circulation Research*, 2001, **88**, 1059–1065, DOI: 10.1161/hh1001.091640.

(32) R. J. Solaro, P. Rosevear and T. Kobayashi, *Biochemical and biophysical research communications*, 2008, **369**, 82–87, DOI: 10.1016/j.bbrc.2007.12.114.

(33) H. E. Salhi, S. D. Walton, N. C. Hassel, E. A. Brundage, P. P. de Tombe, P. M. L. Janssen, J. P. Davis and B. J. Biesiadecki, *Journal of Molecular and Cellular Cardiology*, 2014, **76**, 257–264, DOI: 10.1016/j.yjmcc.2014.09.013.

(34) J. G. Seidman and C. Seidman, *Cell*, 2001, **104**, 557–567, DOI: 10.1016/S0092-8674(01)00242-2.

(35) A. V. Gomes, J. D. Potter and D. Szczesna-Cordary, *IUBMB Life*, 2002, **54**, 323–333, DOI: 10.1080/15216540216037.

(36) B. J. Maron, J. J. Doerer, T. S. Haas, D. M. Tierney and F. O. Mueller, *Circulation*, 2009, **119**, 1085–1092, DOI: 10.1161/CIRCULATIONAHA.108.804617.

(37) A. Malhotra and S. Sharma, *European Cardiology Review*, 2017, **12**, 80–82, DOI: 10.15420/ecr.2017:12:1.

(38) A. J. Marian and E. Braunwald, *Circulation research*, 2017, **121**, 749–770, DOI: 10.1161/CIRCRESAHA.117.311059.

(39) P. A. Harvey and L. A. Leinwand, *The Journal of Cell Biology*, 2011, **194**, 355–365, DOI: 10.1083/jcb.201101100.

(40) S. B. Marston, *Journal of Cardiovascular Translational Research*, 2011, **4**, 245–255, DOI: 10.1007/s12265-011-9266-2.

(41) M. Memo, M.-C. Leung, D. G. Ward, C. dos Remedios, S. Morimoto, L. Zhang, G. Ravenscroft, E. McNamara, K. J. Nowak, S. B. Marston and A. E. Messer, *Cardiovascular Research*, 2013, **99**, 65–73, DOI: 10.1093/cvr/cvt071.

(42) S. Marston, *Journal of Muscle Research and Cell Motility*, 2022, DOI: 10.1007/s10974-022-09630-4.

(43) J. P. Kaski, M. Burch and P. M. Elliott, *Cardiology in the Young*, 2007, **17**, 675–677, DOI: 10.1017/S1047951107001291.

(44) Q.-W. Lu, X.-Y. Wu and S. Morimoto, *Journal of Geriatric Cardiology : JGC*, 2013, **10**, 91–101, DOI: 10.3969/j.issn.1671-5411.2013.01.014.

(45) B. Liu, S. B. Tikunova, K. P. Kline, J. K. Siddiqui and J. P. Davis, *PLoS ONE*, 2012, **7**, e38259, DOI: 10.1371/journal.pone.0038259.

(46) T. Matsuo, F. Kono and S. Fujiwara, *Journal of Structural Biology*, 2019, **205**, 196–205, DOI: 10.1016/j.jsb.2018.12.005.

(47) J. Mogensen, R. T. Murphy, T. Shaw, A. Bahl, C. Redwood, H. Watkins, M. Burke, P. M. Elliott and W. J. McKenna, *Journal of the American College of Cardiology*, 2004, **44**, 2033–2040, DOI: 10.1016/j.jacc.2004.08.027.

(48) B. J. Biesiadecki, T. Kobayashi, J. S. Walker, R. John Solaro and P. P. de Tombe, *Circulation Research*, 2007, **100**, 1486–1493, DOI: 10.1161/01.RES.0000267744.92677.7f.

(49) L. C. Preston, S. Lipscomb, P. Robinson, J. Mogensen, W. J. McKenna, H. Watkins, C. C. Ashley and C. S. Redwood, *Pflügers Archiv - European Journal of Physiology*, 2007, **453**, 771–776, DOI: 10.1007/s00424-006-0161-7.

(50) E. C. Dyer, A. M. Jacques, A. C. Hoskins, D. G. Ward, C. E. Gallon, A. E. Messer, J. P. Kaski, M. Burch, J. C. Kentish and S. B. Marston, *Circulation: Heart Failure*, 2009, **2**, 456–464, DOI: 10.1161/CIRCHEARTFAILURE.108.818237.

(51) A. E. Messer and S. B. Marston, *Frontiers in Physiology*, 2014, **5**, DOI: 10.3389/fphys.2014.00315.

(52) Z. Mahmud, P. S. Dhami, C. Rans, P. B. Liu and P. M. Hwang, *Journal of Molecular Biology*, 2021, **433**, 167010, DOI: 10.1016/j.jmb.2021.167010.

(53) C. M. Stevens, K. Rayani, G. Singh, B. Lotfalisalmasi, D. P. Tieleman and G. F. Tibbits, *Journal of Biological Chemistry*, 2017, **292**, 11915–11926, DOI: 10.1074/jbc.M116.770776.

(54) S. Marston, *Journal of Muscle Research and Cell Motility*, 2019, **40**, 69–76, DOI: 10.1007/s10974-019-09526-w.

(55) K. Alsulami and S. Marston, *International Journal of Molecular Sciences*, 2020, **21**, 9599, DOI: 10.3390/ijms21249599.

(56) P. M. Hwang and B. D. Sykes, *Nature Reviews Drug Discovery*, 2015, **14**, 313–328, DOI: 10.1038/nrd4554.

(57) M. Papadaki, P. G. Vikhorev, S. B. Marston and A. E. Messer, *Cardiovascular research*, 2015, **108**, 99–110, DOI: 10.1093/cvr/cvv181.

(58) N. Tadano, C.-K. Du, F. Yumoto, S. Morimoto, M. Ohta, M.-F. Xie, K. Nagata, D.-Y. Zhan, Q.-W. Lu, Y. Miwa, F. Takahashi-Yanaga, M. Tanokura, I. Ohtsuki and T. Sasaguri, *British Journal of Pharmacology*, 2010, **161**, 1034–1043, DOI: 10.1111/j.1476-5381.2010.00942.x.

(59) R. M. B. Hoffman and B. D. Sykes, *Biochemistry*, 2009, **48**, 5541–5552, DOI: 10.1021/bi9001826.

(60) M. Oleszczuk, I. M. Robertson, M. X. Li and B. D. Sykes, *Journal of Molecular and Cellular Cardiology*, 2010, **48**, 925–933, DOI: 10.1016/j.yjmcc.2010.01.016.

(61) F. Cai, P. M. Hwang and B. D. Sykes, *Biochemistry*, 2018, **57**, 6461–6469, DOI: 10.1021/acs.biochem.8b00882.

(62) X. Wang, M. X. Li and B. D. Sykes, *Journal of Biological Chemistry*, 2002, **277**, 31124–31133, DOI: 10.1074/jbc.M203896200.

(63) I. M. Robertson, M. X. Li and B. D. Sykes, *Journal of Biological Chemistry*, 2009, **284**, 23012–23023, DOI: 10.1074/jbc.M109.021352.

(64) I. M. Robertson, Y.-B. Sun, M. X. Li and B. D. Sykes, *Journal of Molecular and Cellular Cardiology*, 2010, **49**, 1031–1041, DOI: 10.1016/j.yjmcc.2010.08.019.

(65) J. D. Bowman, W. H. Coldren and S. Lindert, *Journal of Chemical Information and Modeling*, 2019, **59**, 2964–2972, DOI: 10.1021/acs.jcim.9b00256.

(66) A. Khandelwal, J. A. Hall and B. S. J. Blagg, *The Journal of Organic Chemistry*, 2013, **78**, 7859–7884, DOI: 10.1021/jo401027r.

(67) A. Sheehan, A. E. Messer, M. Papadaki, A. Choudhry, V. Kren, D. Biedermann, B. Blagg, A. Khandelwal and S. B. Marston, *Frontiers in Physiology*, 2018, **9**, 243, DOI: 10.3389/fphys.2018.00243.

(68) M. Feig and Y. Sugita, *Annual Review of Cell and Developmental Biology*, 2019, **35**, 191–211, DOI: 10.1146/annurev-cellbio-100617-062542.

(69) D. E. Shaw, *Journal of Biomolecular Structure and Dynamics*, 2013, **31**, 108–108, DOI: 10.1080/07391102.2013.786408.

(70) M. J. Rynkiewicz, E. Pavadai and W. Lehman, *Frontiers in Physiology*, 2022, **13**.

(71) E. L. Mehler, J.-L. Pascual-Ahuir and H. Weinstein, *Protein Engineering, Design and Selection*, 1991, **4**, 625–637, DOI: 10.1093/protein/4.6.625.

(72) S. Lindert, P. M. Kekenes-Huskey and J. A. McCammon, *Biophysical Journal*, 2012, **103**, 1784–1789, DOI: 10.1016/j.bpj.2012.08.058.

(73) P. M. Kekenes-Huskey, S. Lindert and J. A. McCammon, *PLOS Computational Biology*, 2012, **8**, e1002777, DOI: 10.1371/journal.pcbi.1002777.

(74) E. P. Manning, J. C. Tardiff and S. D. Schwartz, *Biochemistry*, 2011, **50**, 7405–7413, DOI: 10.1021/bi200506k.

(75) M. R. Williams, S. J. Lehman, J. C. Tardiff and S. D. Schwartz, *Proceedings of the National Academy of Sciences*, 2016, **113**, 3257–3262, DOI: 10.1073/pnas.1519541113.

(76) M. R. Williams, J. C. Tardiff and S. D. Schwartz, *The journal of physical chemistry letters*, 2018, **9**, 3301–3306, DOI: 10.1021/acs.jpclett.8b00958.

(77) A. B. Mason, M. L. Lynn, A. P. Baldo, A. E. Deranek, J. C. Tardiff and S. D. Schwartz, *JCI Insight*, 2021, **6**, DOI: 10.1172/jci.insight.154350.

(78) T. Masuda, *Biosystems*, 2015, **132–133**, 1–5, DOI: 10.1016/j.biosystems.2015.04.001.

(79) W. Zheng and H. Wen, *Proteins: Structure, Function, and Bioinformatics*, 2019, **87**, 805–814, DOI: 10.1002/prot.25707.

(80) M. Abidi and R. Soheilifard, *Journal of Molecular Graphics and Modelling*, 2020, **96**, 107534, DOI: 10.1016/j.jmgm.2020.107534.

(81) E. R. Hantz and S. Lindert, *Journal of Chemical Information and Modeling*, 2021, **61**, 3052–3057, DOI: 10.1021/acs.jcim.1c00419.

(82) J. D. Bowman and S. Lindert, *The Journal of Physical Chemistry B*, 2018, **122**, 7874–7883, DOI: 10.1021/acs.jpcb.8b05435.

(83) S. Lindert, M. X. Li, B. D. Sykes and J. A. McCammon, *Chemical Biology & Drug Design*, 2015, **85**, 99–106, DOI: 10.1111/cbdd.12381.

(84) M. L. Aprahamian, S. B. Tikunova, M. V. Price, A. F. Cuesta, J. P. Davis and S. Lindert, *Journal of Chemical Information and Modeling*, 2017, **57**, 3056–3069, DOI: 10.1021/acs.jcim.7b00536.

(85) D. A. Case, H.M. Aktulga, K. Belfon, I.Y. Ben-Shalom, J.T. Berryman, S.R. Brozell, D.S. Cerutti, T.E. Cheatham, III, G.A. Cisneros, V.W.D. Cruzeiro, T.A. Darden, N. Forouzesh, G. Giambaşu, T. Giese, M.K. Gilson, H. Gohlke, A.W. Goetz, J. Harris, S. Izadi, S.A. Izmailov, K. Kasavajhala, M.C. Kaymak, E. King, A. Kovalenko, T. Kurtzman, T.S. Lee, P. Li, C. Lin, J. Liu, T. Luchko, R. Luo, M. Machado, V. Man, M. Manathunga, K.M. Merz, Y. Miao, O. Mikhailovskii, G. Monard, H. Nguyen, K.A. O'Hearn, A. Onufriev, F. Pan, S. Pantano, R. Qi, A. Rahnamoun, D.R. Roe, A. Roitberg, C. Sagui, S. Schott-Verdugo, A. Shajan, J. Shen, C.L. Simmerling, N.R. Skrynnikov, J. Smith, J. Swails, R.C. Walker, J. Wang, H. Wei, X. Wu, Y. Wu, Y. Xiong, Y. Xue, D.M. York, S. Zhao, Q. Zhu and P.A. Kollman, *University of California, San Francisco*, 2023.

(86) W. D. Cornell, P. Cieplak, C. I. Bayly, I. R. Gould, K. M. Merz, D. M. Ferguson, D. C. Spellmeyer, T. Fox, J. W. Caldwell and P. A. Kollman, *Journal of the American Chemical Society*, 1995, **117**, 5179–5197, DOI: 10.1021/ja00124a002.

(87) J. E. Zamora, M. Papadaki, A. E. Messer, S. B. Marston and I. R. Gould, *Physical Chemistry Chemical Physics*, 2016, **18**, 20691–20707, DOI: 10.1039/C6CP02610A.

(88) J. Eiros Zamora, 2019, DOI: 10.25560/78681.

(89) R. Salomon-Ferrer, A. W. Götz, D. Poole, S. Le Grand and R. C. Walker, *Journal of Chemical Theory and Computation*, 2013, **9**, 3878–3888, DOI: 10.1021/ct400314y.

(90) S. Le Grand, A. W. Götz and R. C. Walker, *Computer Physics Communications*, 2013, **184**, 374–380, DOI: 10.1016/j.cpc.2012.09.022.

(91) J. A. Maier, C. Martinez, K. Kasavajhala, L. Wickstrom, K. E. Hauser and C. Simmerling, *Journal of Chemical Theory and Computation*, 2015, **11**, 3696–3713, DOI: 10.1021/acs.jctc.5b00255.

(92) T. Steinbrecher, J. Latzer and D. A. Case, *Journal of Chemical Theory and Computation*, 2012, **8**, 4405–4412, DOI: 10.1021/ct300613v.

(93) D. J. Price and C. L. Brooks, *The Journal of Chemical Physics*, 2004, **121**, 10096–10103, DOI: 10.1063/1.1808117.

(94) J.-P. Ryckaert, G. Ciccotti and H. J. C. Berendsen, *Journal of Computational Physics*, 1977, **23**, 327–341, DOI: 10.1016/0021-9991(77)90098-5.

(95) C. W. Hopkins, S. Le Grand, R. C. Walker and A. E. Roitberg, *Journal of Chemical Theory and Computation*, 2015, **11**, 1864–1874, DOI: 10.1021/ct5010406.

(96) J. Wang, R. M. Wolf, J. W. Caldwell, P. A. Kollman and D. A. Case, *Journal of Computational Chemistry*, 2004, **25**, 1157–1174, DOI: 10.1002/jcc.20035.

(97) S. Riniker and G. A. Landrum, *Journal of Chemical Information and Modeling*, 2015, **55**, 2562–2574, DOI: 10.1021/acs.jcim.5b00654.

(98) G. Luchini, J. V. Alegre-Requena, I. Funes-Ardoiz and R. S. Paton, *GoodVibes: Automated Thermochemistry for Heterogeneous Computational Chemistry Data*, 2020, DOI: 10.12688/f1000research.22758.1.

(99) Y.-P. Li, J. Gomes, S. Mallikarjun Sharada, A. T. Bell and M. Head-Gordon, *The Journal of Physical Chemistry C*, 2015, **119**, 1840–1850, DOI: 10.1021/jp509921r.

(100) S. Grimme, *Chemistry – A European Journal*, 2012, **18**, 9955–9964, DOI: 10.1002/chem.201200497.

(101) C. I. Bayly, P. Cieplak, W. Cornell and P. A. Kollman, *The Journal of Physical Chemistry*, 1993, **97**, 10269–10280, DOI: 10.1021/j100142a004.

(102) M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, G. A. Petersson, H. Nakatsuji, X. Li, M. Caricato, A. V. Marenich, J. Bloino, B. G. Janesko, R. Gomperts, B. Mennucci, H. P. Hratchian, J. V. Ortiz, A. F. Izmaylov, J. L. Sonnenberg, D. Williams-Young, F. Ding, F. Lipparini, F. Egidi, J. Goings, B. Peng, A. Petrone, T. Henderson, D. Ranasinghe, V. G. Zakrzewski, J. Gao, N. Rega, G. Zheng, W. Liang, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, K. Throssell, J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. J. Bearpark, J. J. Heyd, E. N. Brothers, K. N. Kudin, V. N. Staroverov, T. A. Keith, R. Kobayashi, J. Normand, K. Raghavachari, A. P. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, J. M. Millam, M. Klene, C. Adamo, R. Cammi, J. W. Ochterski, R. L. Martin, K. Morokuma, O. Farkas, J. B. Foresman and D. J. Fox, *Gaussian16*, 2016.

(103) A. D. Becke, *The Journal of Chemical Physics*, 1993, **98**, 5648–5652, DOI: 10.1063/1.464913.

(104) C. Lee, W. Yang and R. G. Parr, *Physical Review B*, 1988, **37**, 785–789, DOI: 10.1103/PhysRevB.37.785.

(105) T. H. Dunning, *The Journal of Chemical Physics*, 1989, **90**, 1007–1023, DOI: 10.1063/1.456153.

(106) S. Miertuš, E. Scrocco and J. Tomasi, *Chemical Physics*, 1981, **55**, 117–129, DOI: 10.1016/0301-0104(81)85090-2.

(107) R. Ditchfield, W. J. Hehre and J. A. Pople, *The Journal of Chemical Physics*, 1971, **54**, 724–728, DOI: 10.1063/1.1674902.

(108) M. Ester, H.-P. Kriegel, J. Sander and X. Xu, Proceedings of the Second International Conference on Knowledge Discovery and Data Mining, AAAI Press, Portland, Oregon, 1996, pp. 226–231.

(109) F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot and É. Duchesnay, *The Journal of Machine Learning Research*, 2011, **12**, 2825–2830.

(110) P. A. Kollman, I. Massova, C. Reyes, B. Kuhn, S. Huo, L. Chong, M. Lee, T. Lee, Y. Duan, W. Wang, O. Donini, P. Cieplak, J. Srinivasan, D. A. Case and T. E. Cheatham, *Accounts of Chemical Research*, 2000, **33**, 889–897, DOI: 10.1021/ar000033j.

(111) B. R. Miller, T. D. McGee, J. M. Swails, N. Homeyer, H. Gohlke and A. E. Roitberg, *Journal of Chemical Theory and Computation*, 2012, **8**, 3314–3321, DOI: 10.1021/ct300418h.

(112) S. van der Walt, S. C. Colbert and G. Varoquaux, *Computing in Science & Engineering*, 2011, **13**, 22–30, DOI: 10.1109/MCSE.2011.37.

(113) H. Nguyen, J. Swails, D. R. Roe and Cody, *Amber-MD/Pytraj: V2.0.0*, Zenodo, 2017, DOI: 10.5281/zenodo.439134.

(114) D. R. Roe and T. E. I. Cheatham, *Journal of Chemical Theory and Computation*, 2013, **9**, 3084–3095, DOI: 10.1021/ct400341p.

(115) D. Lakens, *Frontiers in Psychology*, 2013, **4**, 863, DOI: 10.3389/fpsyg.2013.00863.

(116) S. Salentin, S. Schreiber, V. J. Haupt, M. F. Adasme and M. Schroeder, *Nucleic Acids Research*, 2015, **43**, W443–W447, DOI: 10.1093/nar/gkv315.

(117) M. F. Adasme, K. L. Linnemann, S. N. Bolz, F. Kaiser, S. Salentin, V. J. Haupt and M. Schroeder, *Nucleic Acids Research*, 2021, **49**, W530–W534, DOI: 10.1093/nar/gkab294.

(118) H. C. Jubb, A. P. Higueruelo, B. Ochoa-Montaño, W. R. Pitt, D. B. Ascher and T. L. Blundell, *Journal of Molecular Biology*, 2017, **429**, 365–371, DOI: 10.1016/j.jmb.2016.12.004.

(119) C. W. J. Granger, *Econometrica*, 1969, **37**, 424–438, DOI: 10.2307/1912791.

(120) A. Shojaie and E. B. Fox, *Annual Review of Statistics and Its Application*, 2022, **9**, 289–319, DOI: 10.1146/annurev-statistics-040120-010930.

(121) F. Lewis, A. Butler and L. Gilbert, *Methods in Ecology and Evolution*, 2011, **2**, 155–162, DOI: 10.1111/j.2041-210X.2010.00063.x.

(122) M. Sobieraj and P. Setny, *Journal of Chemical Theory and Computation*, 2022, **18**, 1936–1944, DOI: 10.1021/acs.jctc.1c00945.

(123) H. Kamberaj and A. van der Vaart, *Biophysical Journal*, 2009, **97**, 1747–1755, DOI: 10.1016/j.bpj.2009.07.019.

(124) S. Vatansever, Z. H. Gümüş and B. Erman, *Scientific Reports*, 2016, **6**, 37012, DOI: 10.1038/srep37012.

(125) G. Schwarz, *Annals of Statistics*, 1978, **6**, 461–464, DOI: 10.1214/aos/1176344136.

(126) A. E. Messer, A. M. Jacques and S. B. Marston, *Journal of Molecular and Cellular Cardiology*, 2007, **42**, 247–259, DOI: 10.1016/j.yjmcc.2006.08.017.

(127) A. Amadei, A. B. M. Linssen and H. J. C. Berendsen, *Proteins: Structure, Function, and Bioinformatics*, 1993, **17**, 412–425, DOI: 10.1002/prot.340170408.

(128) O. Beckstein, E. J. Denning, J. R. Perilla and T. B. Woolf, *Journal of Molecular Biology*, 2009, **394**, 160–176, DOI: 10.1016/j.jmb.2009.09.009.

(129) Y. Cheng, S. Lindert, P. Kekenes-Huskey, V. S. Rao, R. J. Solaro, P. R. Rosevear, R. Amaro, A. D. McCulloch, J. A. McCammon and M. Regnier, *Biophysical Journal*, 2014, **107**, 1675–1685, DOI: 10.1016/j.bpj.2014.08.008.

(130) R. M. Levy, L. Y. Zhang, E. Gallicchio and A. K. Felts, *Journal of the American Chemical Society*, 2003, **125**, 9523–9530, DOI: 10.1021/ja029833a.

(131) O. K. Baryshnikova, M. X. Li and B. D. Sykes, *Journal of Molecular Biology*, 2008, **375**, 735–751, DOI: 10.1016/j.jmb.2007.10.062.

(132) O. K. Baryshnikova, I. M. Robertson, P. Mercier and B. D. Sykes, *Biochemistry*, 2008, **47**, 10950–10960, DOI: 10.1021/bi801165c.

(133) S. U. Reiffert, K. Jaquet, L. M. G. Heilmeyer and F. W. Herberg, *Biochemistry*, 1998, **37**, 13516–13525, DOI: 10.1021/bi980280j.

(134) A. Tower-Rader and M. Y. Desai, *Circulation: Cardiovascular Imaging*, 2017, **10**, e006066, DOI: 10.1161/CIRCIMAGING.117.006066.

(135) G.-J. Du, Z. Zhang, X.-D. Wen, C. Yu, T. Calway, C.-S. Yuan and C.-Z. Wang, *Nutrients*, 2012, **4**, 1679–1691, DOI: 10.3390/nu4111679.

(136) F. W. Friedrich, F. Flenner, M. Nasib, T. Eschenhagen and L. Carrier, *Frontiers in Physiology*, 2016, **7**, 607, DOI: 10.3389/fphys.2016.00607.

(137) M. Law and N. Wald, *BMJ : British Medical Journal*, 1999, **318**, 1471–1480.

(138) R. Liew, M. A. Stagg, K. T. MacLeod and P. Collins, *European Journal of Pharmacology*, 2005, **519**, 1–8, DOI: 10.1016/j.ejphar.2005.06.017.

(139) M. Bijak, *Molecules : A Journal of Synthetic Chemistry and Natural Product Chemistry*, 2017, **22**, 1942, DOI: 10.3390/molecules22111942.

(140) V. Křen and K. Valentová, *Natural Product Reports*, 2022, **39**, 1264–1281, DOI: 10.1039/D2NP00013J.

(141) W. H. Coldren, S. B. Tikunova, J. P. Davis and S. Lindert, *Journal of Chemical Information and Modeling*, 2020, **60**, 3648–3661, DOI: 10.1021/acs.jcim.0c00452.

(142) D. Botten, G. Fugallo, F. Fraternali and C. Molteni, *PLOS ONE*, 2013, **8**, e70556, DOI: 10.1371/journal.pone.0070556.

(143) J. Vrba, B. Papoušková, K. Lněničková, P. Kosina, V. Křen and J. Ulrichová, *Antioxidants*, 2021, **10**, 954, DOI: 10.3390/antiox10060954.

(144) L. Szatkowski, M. L. Lynn, T. Holeman, M. R. Williams, A. P. Baldo, J. C. Tardiff and S. D. Schwartz, *ACS Omega*, 2019, **4**, 6492–6501, DOI: 10.1021/acsomega.8b03340.

WOLTERS KLUWER HEALTH, INC. LICENSE
TERMS AND CONDITIONS

Jul 09, 2023

This Agreement between Zeyu Yang ("You") and Wolters Kluwer Health, Inc. ("Wolters Kluwer Health, Inc.") consists of your license details and the terms and conditions provided by Wolters Kluwer Health, Inc. and Copyright Clearance Center.

License Number	5571070013362
License date	Jun 16, 2023
Licensed Content Publisher	Wolters Kluwer Health, Inc.
Licensed Content Publication	Circulation Research
Licensed Content Title	The Troponin C G159D Mutation Blunts Myofilament Desensitization Induced by Troponin I Ser23/24 Phosphorylation
Licensed Content Author	Brandon J. Biesiadecki, Tomoyoshi Kobayashi, John S. Walker, et al
Licensed Content Date	Apr 19, 2007
Licensed Content Volume	100
Licensed Content Issue	10
Type of Use	Dissertation/Thesis
Requestor type	University/College
Sponsorship	No Sponsorship
Format	Electronic

Will this be posted online?	Yes, on an unrestricted website
Portion	Figures/tables/illustrations
Number of figures/tables/illustrations	2
Author of this Wolters Kluwer article	No
Will you be translating?	No
Intend to modify/change the content	No
Title	Atomistic Simulation of Cardiac Troponin with Phosphorylation and Mutation Effects
Institution name	Imperial College London
Expected presentation date	Jul 2023
Portions	Figure 2, Figure 3
Requestor Location	Zeyu Yang 20 Beaufort Gardens London, NW4 3QP United Kingdom Attn: Zeyu Yang
Publisher Tax ID	13-2932696
Total	0.00 GBP
Terms and Conditions	

Wolters Kluwer Health Inc. Terms and Conditions

1. **Duration of License:** Permission is granted for a one time use only. Rights herein do not apply to future reproductions, editions, revisions, or other derivative works. This permission shall be effective as of the date of execution by the parties for the maximum period of 12 months and should be renewed after the term expires.
    i. When content is to be republished in a book or journal the validity of this agreement should be the life of the book edition or journal issue.
    ii. When content is licensed for use on a website, internet, intranet, or any publicly accessible site (not including a journal or book), you agree to remove the material from such site after 12 months, or request to renew your permission license
2. **Credit Line:** A credit line must be prominently placed and include: For book content: the author(s), title of book, edition, copyright holder, year of publication; For journal content: the author(s), titles of article, title of journal, volume number, issue number, inclusive pages and website URL to the journal page; If a journal is published by a learned society the credit line must include the details of that society.
3. **Warranties:** The requestor warrants that the material shall not be used in any manner which may be considered derogatory to the title, content, authors of the material, or to Wolters Kluwer Health, Inc.
4. **Indemnity:** You hereby indemnify and hold harmless Wolters Kluwer Health, Inc. and its respective officers, directors, employees and agents, from and against any and all claims, costs, proceeding or demands arising out of your unauthorized use of the Licensed Material
5. **Geographical Scope:** Permission granted is non-exclusive and is valid throughout the world in the English language and the languages specified in the license.
6. **Copy of Content:** Wolters Kluwer Health, Inc. cannot supply the requestor with the original artwork, high-resolution images, electronic files or a clean copy of content.
7. **Validity:** Permission is valid if the borrowed material is original to a Wolters Kluwer Health, Inc. imprint (J.B Lippincott, Lippincott-Raven Publishers, Williams & Wilkins, Lea & Febiger, Harwal, Rapid Science, Little Brown & Company, Harper & Row Medical, American Journal of Nursing Co, and Urban & Schwarzenberg - English Language, Raven Press, Paul Hoeber, Springhouse, Ovid), and the Anatomical Chart Company
8. **Third Party Material:** This permission does not apply to content that is credited to publications other than Wolters Kluwer Health, Inc. or its Societies. For images credited to non-Wolters Kluwer Health, Inc. books or journals, you must obtain permission from the source referenced in the figure or table legend or credit line before making any use of the image(s), table(s) or other content.
9. **Adaptations:** Adaptations are protected by copyright. For images that have been adapted, permission must be sought from the rightsholder of the original material and the rightsholder of the adapted material.
10. **Modifications:** Wolters Kluwer Health, Inc. material is not permitted to be modified or adapted without written approval from Wolters Kluwer Health, Inc. with the exception of text size or color. The adaptation should be credited as follows: Adapted with permission from Wolters Kluwer Health, Inc.: [the author(s), title of book, edition, copyright holder, year of publication] or [the author(s), titles of article, title of journal, volume number, issue number, inclusive pages and website URL to the journal page].
11. **Full Text Articles:** Republication of full articles in English is prohibited.
12. **Branding and Marketing:** No drug name, trade name, drug logo, or trade logo can be included on the same page as material borrowed from *Diseases of the Colon & Rectum, Plastic Reconstructive Surgery, Obstetrics & Gynecology (The Green Journal), Critical Care Medicine, Pediatric Critical Care Medicine, the American Heart Association publications and the American Academy of Neurology publications*.
13. **Open Access:** Unless you are publishing content under the same Creative Commons license, the following statement must be added when reprinting material in Open Access journals: "The Creative Commons license does not apply to this content. Use of the material in any format is prohibited without written permission from the

publisher, Wolters Kluwer Health, Inc. Please contact permissions@lww.com for further information."

14. **Translations:** The following disclaimer must appear on all translated copies: Wolters Kluwer Health, Inc. and its Societies take no responsibility for the accuracy of the translation from the published English original and are not liable for any errors which may occur.
15. **Published Ahead of Print (PAP):** Articles in the PAP stage of publication can be cited using the online publication date and the unique DOI number.
    i. Disclaimer: Articles appearing in the PAP section have been peer-reviewed and accepted for publication in the relevant journal and posted online before print publication. Articles appearing as PAP may contain statements, opinions, and information that have errors in facts, figures, or interpretation. Any final changes in manuscripts will be made at the time of print publication and will be reflected in the final electronic version of the issue. Accordingly, Wolters Kluwer Health, Inc., the editors, authors and their respective employees are not responsible or liable for the use of any such inaccurate or misleading data, opinion or information contained in the articles in this section.
16. **Termination of Contract:** Wolters Kluwer Health, Inc. must be notified within 90 days of the original license date if you opt not to use the requested material.
17. **Waived Permission Fee:** Permission fees that have been waived are not subject to future waivers, including similar requests or renewing a license.
18. **Contingent on payment:** You may exercise these rights licensed immediately upon issuance of the license, however until full payment is received either by the publisher or our authorized vendor, this license is not valid. If full payment is not received on a timely basis, then any license preliminarily granted shall be deemed automatically revoked and shall be void as if never granted. Further, in the event that you breach any of these terms and conditions or any of Wolters Kluwer Health, Inc.'s other billing and payment terms and conditions, the license is automatically revoked and shall be void as if never granted. Use of materials as described in a revoked license, as well as any use of the materials beyond the scope of an unrevoked license, may constitute copyright infringement and publisher reserves the right to take any and all action to protect its copyright in the materials.
19. **STM Signatories Only:** Any permission granted for a particular edition will apply to subsequent editions and for editions in other languages, provided such editions are for the work as a whole in situ and do not involve the separate exploitation of the permitted illustrations or excerpts. Please view: STM Permissions Guidelines
20. **Warranties and Obligations:** LICENSOR further represents and warrants that, to the best of its knowledge and belief, LICENSEE's contemplated use of the Content as represented to LICENSOR does not infringe any valid rights to any third party.
21. **Breach:** If LICENSEE fails to comply with any provisions of this agreement, LICENSOR may serve written notice of breach of LICENSEE and, unless such breach is fully cured within fifteen (15) days from the receipt of notice by LICENSEE, LICENSOR may thereupon, at its option, serve notice of cancellation on LICENSEE, whereupon this Agreement shall immediately terminate.
22. **Assignment:** License conveyed hereunder by the LICENSOR shall not be assigned or granted in any manner conveyed to any third party by the LICENSEE without the consent in writing to the LICENSOR.
23. **Governing Law:** The laws of The State of New York shall govern interpretation of this Agreement and all rights and liabilities arising hereunder.
24. **Unlawful:** If any provision of this Agreement shall be found unlawful or otherwise legally unenforceable, all other conditions and provisions of this Agreement shall remain in full force and effect.

For Copyright Clearance Center / RightsLink Only:

1. **Service Description for Content Services:** Subject to these terms of use, any terms set forth on the particular order, and payment of the applicable fee, you may make the

following uses of the ordered materials:

i. **Content Rental:** You may access and view a single electronic copy of the materials ordered for the time period designated at the time the order is placed. Access to the materials will be provided through a dedicated content viewer or other portal, and access will be discontinued upon expiration of the designated time period. An order for Content Rental does not include any rights to print, download, save, create additional copies, to distribute or to reuse in any way the full text or parts of the materials.

ii. **Content Purchase:** You may access and download a single electronic copy of the materials ordered. Copies will be provided by email or by such other means as publisher may make available from time to time. An order for Content Purchase does not include any rights to create additional copies or to distribute copies of the materials

Other Terms and Conditions:

v1.18

Questions? customercare@copyright.com.