## Supplementary Information

## Troponin structure: its modulation by Ca<sup>2+</sup>and phosphorylation studied by molecular dynamics simulations

Juan Eiros Zamora<sup>a</sup>, Maria Papadaki<sup>b</sup>, Andrew E. Messer<sup>b</sup>, Steven B. Marston<sup>b</sup> and Ian R.

Gould<sup>a</sup>

<sup>a</sup>Department of Chemistry, Institute of Chemical Biology, Imperial College London, SW7 2AZ London, United Kingom

<sup>b</sup>National Heart & Lung Institute, Myocardial Function Section, Imperial College London, W12 ONN London, United Kingdom

Corresponding Author:

Ian R. Gould, PhD. Department of Chemistry, Imperial College London, SW7 2AZ London, United Kingdom. Email: <u>i.gould@imperial.ac.uk</u>

#### Table of Contents

MODEL PRIMARY SEQUENCE	3
ROOT MEAN SQUARE DEVIATIONS	3
WILD TYPE CARDIAC TROPONIN	3
PHOSPHORYLATED CARDIAC TROPONIN	6
FIRST PRINCIPAL COMPONENT MOTION REPRESENTATION	8
CLUSTERING PARAMETERS	8
ROOT MEAN SQUARE FLUCTUATIONS	11
CATALYTIC CA <sup>2+</sup> COORDINATION SPHERE DISTANCES	12
TIME SERIES	12
DENSITY COMPARISONS	16
CONTACT MAPS	18
DIFFERENCE CONTACT MAPS	19

## Model primary sequence

sp/P63316/TNNC1_HUMAN/1-161	1 MDD I YKAAV EQLTE EQKNEFKAAFD I FVLGAEDGC I STKELGKVMRMLGQNPTPEELQEM I DEVDEDGSGTVDFDEFLVM	M81
<pre>critc_moder.pub/1-101</pre>		101
cTnC_model.pdb/1-161	82 VRCMKDDSKGKSEEELSDLFRMFDKNADGYTDLDELKTMLQATGETTTEDDTEELMKDDKNNDGRTDYDEFLEFMKGVE 82 VRCMKDDSKGKSEEELSDLFRMFDKNADGYTDLDELKTMLQATGETTTEDDTEELMKDGDKNNDGRTDYDEFLEFMKGVE	161
sp P45379 TNNT2_HUMAN/1-298 cTnT_model.pdb/1-87	1 MSD I E EVVE E Y E E E Q E E AAVE E E E DWR E DE DE Q E E AAEEDAEAEAETE E T R AEEDE E E E E AKEAEDGPME E	72
sp P45379 TNNT2_HUMAN/1-298 cTnT_model.pdb/1-87	73 SKPKPRSFMPNLVPPK IPDGERVDFDDI HRKRMEKDLNELQAL I EAHFENRKKE E E E LVSLKDR I ERRRA ER	144
sp P45379 TNNT2_HUMAN/1-298 cTnT_model.pdb/1-87	145 A EQQR I R N ER EK ERQNR LA E ERAR R E E E E N R K A E DE AR K K K A L S NMH H G G Y I Q K Q A Q T E R K S G K R Q T E R E 1	216 5
sp P45379 TNNT2_HUMAN/1-298 cTnT_model.pdb/1-87	217 K K K K I LA ERRK V LA I DHLNEDQLR EKAK E LWQ S I YNLEA EK FDLQ EK FKQQK Y E I NVLRNR I NDNQK V SK TR 6 K K K K I LA ERRK V LA I DHLNEDQLR EKAK E LWQ S I YNLEA EK FDLQ EK FKQQK Y E I NVLRNR I NDNQK V SK TR	288 77
sp P45379 TNNT2_HUMAN/1-298 cTnT_model.pdb/1-87	289 <mark>GKAKVTGRWK</mark> 78 <mark>GKAKVTGRWK</mark>	298 87
sp P19429 TNNI3_HUMAN/1-210 cTnl_model.pdb/1-172	1 MADG S SDAAR E P R P A P A P I R R R S S N Y R A Y A T E P HAKKK SK I SA S R K LQ L K T L L LQ I A KQ E L E R E A E E R R G E K C 1 MADG S SDAAR E P R P A P A P I R R R S S N Y R A Y A T E P HAKKK SK I SA S R K LQ L K T L L LQ I A KQ E L E R E A E E R R G E K C	73 73
sp P19429 TNNI3_HUMAN/1-210 cTnl_model.pdb/1-172	74 RALSTRCQP LE LAGLGFA E LQDLCRQLHAR VDK VDE ERYDI EAK VTKNITE I ADLTQK I FDLRGK FKRPTLRF 74 RALSTRCQP LE LAGLGFA E LQDLCRQLHAR VDK VDE ERYDI EAK VTKNITE I ADLTQK I FDLRGK FKRPTLRF	146 146
sp P19429 TNNI3_HUMAN/1-210 cTnl_model.pdb/1-172	147 <mark>VR I SADAMMQA LLGARAK E SLDLRA</mark> HLKQVKK EDT EK ENR E VGDWRKN I DALSGM EGRKKK FE S 147 <mark>VR I SADAMMQA LLGARAK E SLDLRA</mark>	210 171
Figure S1. Sequence a	alignment of our cardiac Troponin model to the canonical human sequences of cTnC (t	top),

cTnT (middle) and cTnI (bottom).

## Root Mean Square Deviations

#### Wild Type cardiac Troponin

HMR refers to the MD trajectories obtained using the hydrogen mass repartition protocol, allowing for a time step of 4 fs. cMD refers to the trajectories obtained with a 2 fs timestep.



**Figure S2.** RMSD evolution through time for each individual WT trajectory. The RMSD is calculated for the backbone atoms against the first frame of each run.



Figure S3. Summary of the RMSD results shown in Figure S2. Average RMSD values are plotted as dots and the standard deviation as error bars.

MD run	Average RMSD (Å)	Std. Dev. (Å)
1 (HMR)	9.14	0.72
2 (HMR)	9.41	1.38
3 (HMR)	9.91	1.32
4 (HMR)	10.66	1.55
5 (HMR)	9.45	1.44
6 (HMR)	13.65	2.40
7 (cMD)	12.18	2.89
8 (cMD)	9.44	1.13
9 (cMD)	14.43	2.27
10 (cMD)	7.04	0.69

Table S1. Values used to generate Figure S3.

RMSD analysis of the common regions of the model to the Takeda crystal structure, discarding the first 50 ns of simulation.

The residues in the model that match the crystal structure are presented in Table S2. **Table S2.** Correspondence between the residues present both in our 419-residue cTn model and the PDB 1J1D (chains A to D) crystal structure.

Subunit	PDB 1J1D	cTn model
	numbering	numbering
cTnC	1-89 (chain A)	1-89
cTnC	92-161 (chain A)	92-161
cTnT	202-271 (chain B)	162-231
cTnl	35-136 (chain C)	283-384
cTnl	145-160 (chain C)	393-408



**Figure S4.** RMSD deviation of each run to the Takeda crystal structure (PDB 1J1D). The RMSD is calculated for the backbone atoms of the residues that are present in both structures. The first 50 ns of each simulation are discarded.



Figure S5. Summary of the RMSD results shown in Figure S4. Average RMSD values are plotted as dots and the standard deviations as error bars.

MD run	Average RMSD (Å)	Std. Dev. (Å)
1 (HMR)	6.46	0.51
2 (HMR)	6.85	1.27
3 (HMR)	6.51	1.28
4 (HMR)	7.12	1.18
5 (HMR)	6.50	0.96
6 (HMR)	9.91	1.78
7 (cMD)	9.61	2.35
8 (cMD)	5.70	0.89
9 (cMD)	8.65	1.02
10 (cMD)	4.83	0.63

<b>Table 33.</b> Values used to generate Figure 33
--

#### Phosphorylated cardiac Troponin



S1P refers to the bis-phosphorylated systems that have each phosphoserine with a -1 chage, and SEP to each phosphoserine with a -2 charge.

**Figure S6.** RMSD evolution through time for each individual SP23/SP24 trajectory to the Takeda crystal structure(PDB 1J1D). The RMSD is calculated for the backbone atoms of the residues that are present in both structures. The first 50 ns of each simulation are discarded.



Figure S7. Summary of the RMSD results shown in Figure S6. Average RMSD values are plotted as dots and the standard deviations as error bars.

MD run	Average	Std. Dev.
	RIVISD (A)	(A)
1	6.61	0.7
2	7.72	0.63
3	11.2	1.46
4	7.85	1.57
5	6.87	1.38
6	6.77	1.27
7	7.96	1.65
8	10.46	1.12
9	6.34	0.99
10	7.2	1.13
11	7.31	1.22
12	5.67	0.62

Table S4. Values used to generate Figure S7.

## First Principal Component motion representation



**Figure S8.** Hinge representation due to the contribution of each residue to the 1<sup>st</sup> Principal Component of motion. The backbone is colored from red to blue based on the specific contribution (blue being low and red high).



**Figure S9.** DBSCAN clustering parameters (left: sorted K-dist plot; right: pSF and DBI metrics) for the CcTnI region (cTnI<sub>135-171</sub>) of the HMR WT runs. In this case, a cluster count of 4 is the best option.

#### Clustering parameters





**Figure S10.** DBSCAN clustering parameters for the following regions and type of MD simulation, in order: CcTnIcMD, CcTnT-cMD, CcTnT-HMR, NcTnI-cMD and NcTnI-HMR; where cMD corresponds to the WT simulations performed with a 2fs timestep and HMR to the ones using a 4fs timestep. The left graphs are the sorted K-dist plots, used to narrow down the area of  $\varepsilon$  and MinPnts parameters, and the right graphs are the pSF (top) and DBI (bottom) clustering metrics as a function of the cluster count, used to determine the optimal combination of  $\varepsilon$  and MinPnts.

### Root Mean Square Fluctuations



**Figure S11.** RSMF comparison between the two different types of SP23/SP24 systems that were simulated. S1P corresponds to both phosphoserines carrying a -1 net charge, and SEP to both carrying a -2 net charge (-2 and -4 total net charge on each system, respectively). As can be appreciated by direct comparison and from the T-test analysis, no significant difference in fluctuations can be distinguished between the two types of phosphorylation.



**Figure S12.** Results for the T-test between the RMSF values of the WT and SP23/SP24 systems. The Welch variant assuming varying variance and unequal sample size was used to calculate the 95% CI. Only the C-terminal cTnT region presents a significantly lower RMSF value after phosphorylation. As discussed in the main text, this could be an issue caused by the fact that the phosphorylated runs were started from the conformations obtained after the WT cluster analysis.

# Catalytic Ca<sup>2+</sup> coordination sphere distances $\underline{\text{Time series}}$



D67 OD1 and OD2





D75 OD1 and OD2

15 14 12	15 14 13	15 14 12 13	15 14 12 13	15 14 MD run 1	13 14 MD run 2	13 14 MD run 3	15 14	15 14 MD run 5
			ali ka shu ni ni falimili					
2	2	2	2	S . I. balling an Hills	S Steam and the sta	<sup>3</sup> <sup>1</sup>	Soll Hilling, Managered	§ Sectional times of a section of the section of th
100 200 300 400 500 600 700 Time (ng)	200 200 300 400 500 600 700 Time (na)	100 200 300 400 500 600 700 Time (ng)	100 200 300 400 500 600 700 Time (na)	5	5	5	5	5
15 075001 5725 5724 14 - MD run 5	15 14 14 15 14 15 14 15 15 15 15 15 15 15 15 15 15 15 15 15	15 075001 5725 5724 14 - MD run 7	13 14 14 11 11	2	3	,	,	>
and a start of the second s	Anticology and two entered initiality of the of	interferent i beneam vie antik beiten die else als entstelligt	and an increase the state of the second second	2	2	2-	2-	2
	alle de la contrata d		A REAL PROFESSION OF THE REAL PROFESSION OF	0 100 200 300 400 500 600 700 Time (m)	0 100 200 300 400 500 600 700 Time (m)	0 100 200 300 400 500 600 700 Time (ral)	0 100 200 300 400 500 600 300 Time (m)	0 100 200 300 400 500 600 700 Time (rol)
2 1 4 C 4 C 10 10 10	an an the second to the second second	A distance of the second states of the second state	bear an in the second second	15 D750D1 WT	15 075001 WT	15 D75001 WT	15 D750D1 WT	15 075001 WT
4	4 3- 2	4 3 2	4 3 2	14 · · · · · · · · · · · · · · · · · · ·	11	11	η.	14
1 0 300 200 300 400 500 600 700	10 100 200 300 400 500 600 700	10 200 300 400 500 600 700	10 200 300 400 500 600 700	in the later . Mittanti into it	- Link milita site al	" Li sa the siles hade	والال والبيد بروا	······································
13 D750D1 \$P23-\$P24	15 D75001 5923-5924	13 D75001 5923-5924	13 D75001 5P23-5P24					• AND A BUILDING
14 13 12	14 MD run 10 13 12	14	10			1 Albertan August		Sector and the sector of
	ر المالية من الشكار فالشكاليات.	A CONTRACTOR OF A CONTRACTOR O		Strict March Ma	8 S. Shan, Edite & March	S S MARINE MILLING	a lan ann ann ann an	8
a finde a life of the state of the second	In Albert of Invalors		a a contair dhi indina a contair i s	4.	5	3	5	5
5		5	5	3	3	3	3	3
3	2	3	2	1	2	1	1	1
0 200 200 300 600 500 600 700 Time (rol)	0 200 200 300 400 500 400 700 Time (rol)	0 100 200 300 400 500 600 700 Time (rst)	0 100 200 300 400 500 400 700 Time (nd)	0 100 200 300 400 500 400 700 Time (rst)	9 50 100 150 200 250 300 350 400 450 5 Time Ind	0 30 100 150 200 250 300 350 400 450 50 Time (na)	0 50 100 150 200 250 300 350 400 450 500 Time (rst)	50 100 150 200 250 300 350 400 458 50 Time (vol
15 D750D2 5P23-5P24	15 075002 5923-5924	13 075002 5923-5924	13 D75002 5923 5924	13 D75002 WT	15 D75002 WT	13 D75002 WT	15 D75002 WT	15 D75002 WT
10 12 12 delle à thirthe de contes, sel some det , dit binard auss.		10- 10- 12-		14 MO NOR 1	14 10	14 MD TOR 3	14 MD Fight 4	14
a frankriger af de la state					12 In the state base for the state	and the second states	in a statistic	". The second contraction of the second
a statistication and a statistical statistical statistication and the statistication of the statistical st		ומימידיה חדי ייוסיין	lad on aldalli, Addin ut					"When the follows.
		5	5	s Lillian a Harda	: 1000 UK M . M	s (Libitit, a) can	s a da la fa di si	
2	2	2	2	<sup>8</sup> Lut. (telefablics) all, all	S The color and data	<sup>3</sup> A. R. J. LAMMARKARK	<sup>3</sup> <sup>1</sup>	<sup>8</sup> Show and the state of the second seco
0 300 200 300 400 500 600 700 Time (raj)	0 <sup>1</sup> 300 200 300 400 500 600 700 Time (raj)	0 100 200 300 400 500 600 700 Time (na)	100 200 300 400 500 400 700 Time (ns)	5	5	5	5	5
15 D75002 \$P23-\$P24 14	15 D75002 5P23-5P24	15 075002 5723-5724 14	15 16 16 16 175002 5725 5724 1725 5724 16 175002 5725 5724	3-	3	3	3	3
in and the second secon	a selecte adoressa i contra l'azi titali (il denit	finning said files a shalls	hilitet) and in the dented in the set	1	2	2	2	2
	Negli Julio Personali S		a des de la Constillo - All	0 100 200 300 400 500 600 700 Tinut (m)	0 100 200 300 400 500 600 700 Time (m)	0 100 200 300 400 500 600 700 Time (rel)	0 100 200 300 400 500 600 300	0 100 200 300 400 500 600 700
	and the second states of a second states of a second states of a second state of a s	the state of the second state of the second	transition of a provider out that a	13 D75002 WT	13 075002 WT	15 D75002 WT	15 D75002 WT	15 D75002 WT
4	4	3	3	14	14	11	14	14
1 0 200 200 300 600 500 600 200	100 200 100 400 500 500 700	100 200 100 400 500 400 700		no. No. and a consider the set	and the distance of all the second	n a chadrain a da b	Class of Mirror and Mali	12 12 million fills at a Long of
Time Insl 075002 5P23-5P24	Time (m) D75002 5P23-5P24	Time (m) 075002 5P23-5P24	Time (n) 075002 5P23-5P24				• <b>•</b> • • • • • • • • • • • • • • • • •	
14	14 MD run 10 13	14	14	5 - CIT (	s 1. (Yu (1. ), 4. (9)			
A A A A A A A A A A A A A A A A A A A	A ALL PRIMA ALL ADDIVISION	AND A PROPERTY OF A PARTY OF A PA	A RATÉRICAN MANAGEMENTAL IN	State of the state of the		Statistics, and read	i si si si mana na si	State of Milling
in the second se		A MARY . MANAGE	ALTER AND A DESCRIPTION OF THE	5	5	3	5	5
5	5	5	5	3	3	3	3	3
3	2	3	3	2	2	2	2	5
0 100 200 300 400 500 600 700 Time (ns)	0 100 200 100 400 500 600 700 Time (nd)	0 100 200 100 400 500 600 700 Time (ns)	0 100 200 100 400 500 600 700 Time (nd)	0 100 200 300 600 500 600 100 Time (vs)	0 50 100 150 200 250 300 353 400 450 5 Time 3%3	0 0 100 150 200 250 300 350 400 450 50 Time (na)	0 50 100 150 200 250 300 350 400 450 500 Time Inti	e 50 100 150 200 253 300 350 400 458 50 Time (vol

E66 OE1 and OE2

	E660E1 5P23-5P24	E660E1 5P23-5P24			E66OE1 WT		E66OE1 WT	E66OE1 WT	
14 13 12 11	- MD run 1	MD run 2	- M0 run 3	14 - MD run 4 12	13 14 13	15 14 13	14 14 13	10 14 13	10 14 10
Distance (A)	Based in the second sec	A DESCRIPTION OF TAXABLE PARTY			an address in	n Marine and the first section.	and the second		Mar Ballow hat
			5 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4	2 4 2 2 2 4 2 2 2 2 2 2 2 2 2 2 2 2 2 2	and a special plate		11 Part March 1916		Based of the second sec
15 14	200 200 300 400 500 600 700 Trime incl 6660E1 5P23-5P24 15 	200 200 300 400 500 600 700 Trine ful E660E1 \$P23-\$P24	100 200 300 400 500 600 300 10 6660[1 5P23-5P24 	100 200 300 400 500 600 300 11 E660E1 5P23-5P24 14	5	5- 4- 3-	5- 4- 3-	5	e blade a
e (h)		-			2 1 0 100 200 300 400 500 600 700	2 0 1 00 200 300 400 500 600 700	2 1 0 100 200 300 400 500 600 700	2 1 0 100 200 300 400 500 600 300	2 3 6 100 200 300 400 500 600 700
2 0 5 4 1	Base	Deve	Director	a statistic production of the state	15 E660(3 WT 14 MD nun 6	Time (n) 13 14 14 14	15 E6660(1 WT 14	Time (n) 6660(3 WT 14 	15 (6660(1 WT 16 M0 run 10)
210	2 2 300 200 300 400 500 600 200 Time (m)	300 200 300 400 500 600 700 Tirre (m)	2 3 0 100 200 300 400 500 600 700 Time (no)	2 0 100 200 300 400 500 400 300 Time (%)		n n n du hidh	Later and march		n Mahlu Manu, Mu
15 14 13 12	0660(1 \$923-\$924	E660(1 923-5924	5 E660[1.923-5924	11 [660011 5923-5924 14 [97] [97] [97] [97] [97] [97] [97] [97]	les stands, elses s	the during and	distribution and the		The Martin Law
nancelli n v n e 81			WW.U.		a a ta angerera	s fulls days of Modell down		8 6 5	
	5		5 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	5	3- 2- 1-	3	3	3	3
-	200 200 300 400 500 600 700 Vina (m)	200 200 300 400 500 600 700 York (set)	0 200 200 800 500 600 700	0 100 200 300 400 500 400 300	0 100 200 300 400 500 400 300 Yona (un)	0 30 100 150 200 200 300 350 400 450 500 Yuna (un)	9 50 200 250 200 250 300 350 400 450 500	50 100 150 200 250 300 350 400 450 50	0 \$10 100 100 200 200 300 400 450 50
								1000 000	1004-010
15 14 13 12	E66012 5923-5924	666012 5923-5924	666002 5923-5924	15 E66002 5923 5924 14	15 E66062 WT 14 MD run 1	13 E66062 WT 14	13 14 13	13 E66622 WT 14 MD nun 4 13	15 E66012 WT 14 MD num 5 13
15 14 12 12 13 14 15 17 15 15 15 15 15 15 15 15 15 15 15 15 15	15660(25925924 MD run 1) 15 15 15 15 15 15 15 15 15 15 15 15 15	194012 1923-1924 	CECCI 2:923-924	1 E6002 5925-9924	14 [66012 WT 14 [10 num1] 10 11 13 11 14 11 15 11 16	15 E66012 WT 10 M M M M M M M M M M M M M M M M M M M	E66002 WT     E66002 WT     MD run 3     J	10 00000000000000000000000000000000000	10 CEGODUWT
10 + 2 + 2 + 7 + 10 + 11 12 12 14 13 - 2 + 3 - 2	EKECC 2923-5924	(conc) 50/3 40/4 (	10000 973 974 	56002 923 994 - 100 mm 1 - 1	1 Au COO3		CECCO2 WT	EXCELOS IN EXCELOS IN MO run () 10 10 10 10 10 10 10 10 10 10	The Close of
15 14 13 12 13 10 9 8 7 8 5 4 3 2 1 0 17 10 004160		EMOCI 973 974           Image: State Stat					A Constant of the second secon		
11 11 11 11 10 9 8 7 8 5 8 1 2 1 0 11 11 11 11 11 11 11 11 11 11 11 11							Concernent of the second secon	Torrest and the second	
Dodance (A) 5 + 2 + 4 + 5 (1) (1) (1) (1) (2 + 5 + 5 + 5 + 6 (1) (1) (1) (1) (1) (1) (1) (1) (1) (1)									How we have a set of the set of t
Distance (A) Dista							A CONTROL OF A CON	MOTOR WE MOTOR AND	
Distance (A) 0 = 2 = 2 = 2 = 2 = 2 = 2 = 2 = 2 = 2 =					A DESCRIPTION OF THE OWNER OF THE OWNER OF THE OWNER				
00M Distance (A Di					Here and the second sec				
Desarce (A) Desarc					Here is a set of the s				



E76 OE1 and OE2





**Figure S13.** Time evolution of the distances between the catalytic Ca<sup>2+</sup> and cTnC site II amino acids for each run. The WT systems are plotted in blue and the phosphorylated systems in orange.







**Figure S14.** Comparison of the distance distributions between the WT and the SP23/SP24 systems for the aggregated time series of Figure S13. The distribution is smoothened with a Gaussian kernel density estimator.



**Figure S15.** Average contact maps for the N-terminal cTnC (1-88) region against the N terminal cTnI region (1-41) in the unphosphorylated (WT) and phosphorylated (SP23/SP24 states).



**Figure S16.** Average contact maps for the unphosphorylated (left), phosphorylated (right) systems between the complete cTnC molecule and the cTnI inhibitory peptide.



**Figure S17.** Average contact maps for the unphosphorylated (left) and phosphorylated (right) systems between the C terminal region of cTnT and the N terminal region of cTnI.



**Figure S18.** Average contact maps for the unphosphorylated (left) and phosphorylated (right) systems between the C terminal region of cTnT and the cTnI inhibitory peptide.

#### Difference contact maps

The following six difference contact maps highlight the changes in fraction of contact between the studied regions introduced by phosphorylation. The win/loss of contact values are mapped from the color bar, and the associated standard deviations are annotated as text in each cell. As can be observed, the standard deviation is always equal (or larger) than the absolute value of the difference.







**Figure S20.** Difference contact map between the N- terminal region of cTnC and the N – terminal region of cTnI.









Figure S23. Difference contact map between the C terminal region of cTnT and the inhibitory peptide of cTnI.



