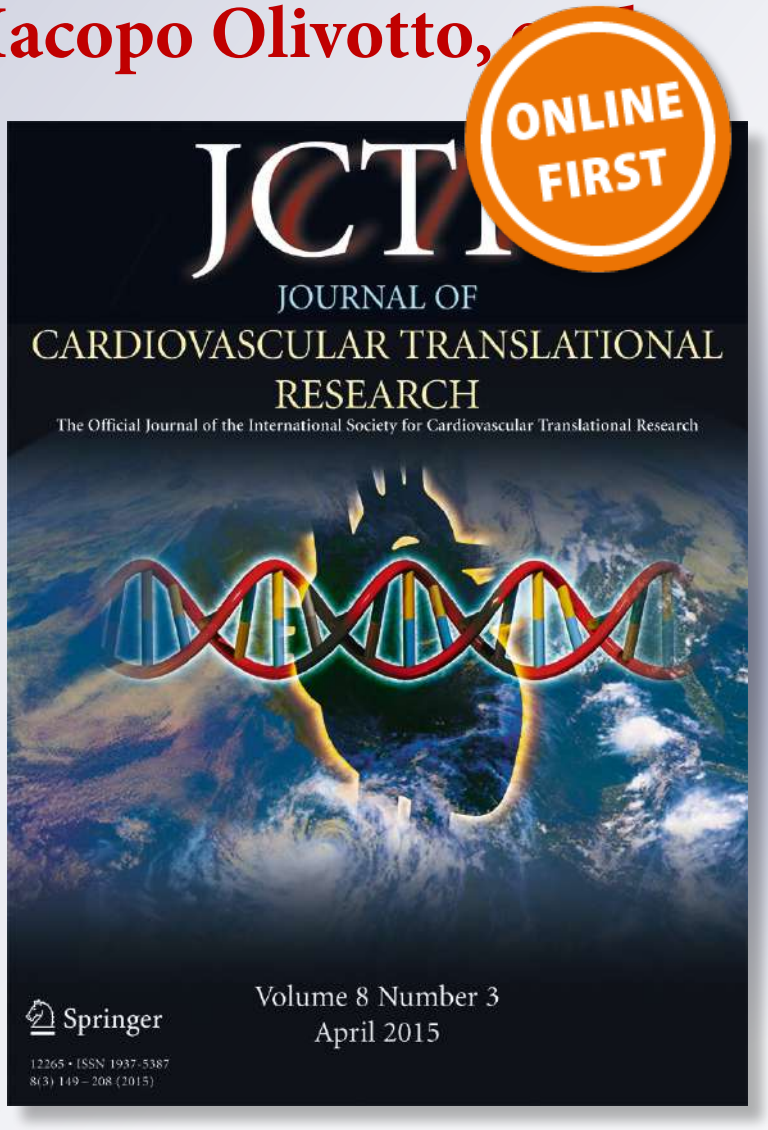
An Investigation of the Molecular Mechanism of Double cMyBP-C Mutation in a Patient with End-Stage Hypertrophic Cardiomyopathy

Poornima Gajendrarao, Navaneethakrishnan Krishnamoorthy, Senthil Selvaraj, Francesca Girolami, Franco Cecchi, Iacopo Olivotto,

Journal of Cardiovascular Translational Research

ISSN 1937-5387

J. of Cardiovasc. Trans. Res.
DOI 10.1007/s12265-015-9624-6



Springer

J. of Cardiovasc. Trans. Res.
DOI 10.1007/s12265-015-9624-6

An Investigation of the Molecular Mechanism of Double cMyBP-C Mutation in a Patient with End-Stage Hypertrophic Cardiomyopathy

Poornima Gajendrarao[1,4] · Navaneethakrishnan Krishnamoorthy[1,4] · Senthil Selvaraj[1] · Francesca Girolami[2,3] · Franco Cecchi[2] · Iacopo Olivotto[2] · Magdi Yacoub[1,4]

Received: 10 December 2014 / Accepted: 7 April 2015

Abstract Mutations in the gene coding for cardiac myosin binding protein-C (cMyBP-C), a multi-domain (C0-C10) protein, are a major causative factor for inherited hypertrophic cardiomyopathy. Patients carrying mutations in this gene have an extremely heterogeneous clinical course, with some progressing to end-stage heart failure. The cause of this variability is unknown. We here describe molecular modeling of a double mutation in domains C1 (E258K) and C2 (E441K) in a patient with severe HCM phenotype. The three-dimensional structure for the C1-motif-C2 complex was constructed with double and single mutations being introduced. Molecular dynamic simulations were performed for 10 ns under physiological conditions. The results showed that both E258K and E441K in isolation can predominantly affect the native domain as well as the nearby motif via conformational changes and result in an additive effect when they coexist. These changes involve important regions of the motif such as phosphorylation and potential actin-binding sites. Moreover, the charge reversal mutations altered the surface electrostatic properties of the complex. In addition, we studied protein expression, which showed that the mutant proteins were expressed and we can suppose that the severe phenotype was not due to haploinsufficiency. However, additional studies on human gene expression will need to confirm this hypothesis. The double mutation affecting the regulatory N-terminal of cMyBP-C have the potential of synergistically interfering with the binding to neighbouring domains and other sarcomeric proteins. These effects may account for the severe phenotype and clinical course observed in the complex cMyBP-C genotypes.

Keywords Cardiac myosin binding protein-C · Double mutation · Hypertrophic cardiomyopathy · Protein expression · Molecular dynamics simulation · ?Structure–function relationship

Introduction

Hypertrophic cardiomyopathy (HCM) is a common inherited heart disease caused by mutations in sarcomeric proteins and most commonly in cardiac myosin-binding protein-C (cMyBP-C) [1]. HCM is a prevalent cause of sudden cardiac death and heart failure-related disability in the young [2–4]. Clinical presentation is extremely variable with an important minority of patients progressing to severe left ventricular dysfunction and heart failure, occasionally requiring heart transplant [5, 6]. The causes of disease progression are largely unresolved, although several causative features have been identified [7]. Complex genotypes characterized by double mutations are over-represented in HCM patients with early symptomatic onset, increased arrhythmic risk and adverse clinical course, presumably reflecting a gene dosage effect in this disease [8]. However, it is not clear that the missense proteins are actually expressed due to the complexity in the

Associate Editor Daniel P. Judge oversaw the review of this article

Poornima Gajendrarao and Navaneethakrishnan Krishnamoorthy contributed equally to this work.

Electronic supplementary material The online version of this article (doi:10.1007/s12265-015-9624-6) contains supplementary material, which is available to authorized users.

✉ Magdi Yacoub
m.yacoub@imperial.ac.uk

[1] Qatar Cardiovascular Research Centre, Qatar Foundation, Doha, Qatar

[2] Referral Centre for Myocardial Diseases, Careggi University Hospital, Florence, Italy

[3] Genetic Unit, Careggi University Hospital, Florence, Italy

[4] Heart Science Centre, National Heart and Lung Institute, Imperial College London, Harefield UB9 6JH, UK

Published online: 14 May 2015

genotypes of cMyBP-C. Also, the molecular mechanisms responsible for adverse prognosis in HCM patients with double mutations have not been addressed. Therefore, we used molecular modeling to start to define how two point mutations in the gene *MYBPC3* may lead to alterations in the structural conformation. Structural changes in cMyBP-C may alter the ability of this protein to bind and signal to its targets downstream.

Patient and Methods

Clinical Features of the Study Proband

The patient, carrying the E258K and E441K mutations, was diagnosed with non-obstructive HCM at 21 years of age due to dyspnea on effort. His echocardiogram at the time showed asymmetric left ventricular (LV) hypertrophy with maximum septal thickness of 22 mm, normal systolic function (ejection fraction 67 %) and moderate diastolic dysfunction (pseudonormalized pattern). In the following years, he developed progressive heart failure with repeated hospitalizations due to acute pulmonary edema and paroxysmal atrial fibrillation. At age 40, his LV ejection fraction had dropped to 33 %, with restrictive LV filling pattern; septal thinning occurred due to extensive fibrous substitution of the myocardium (Fig. 1). Functional mitral regurgitation developed due to annulus dilatation, which required surgical correction by valvuloplasty. Due to further worsening of clinical conditions to NYHA class IV, he was evaluated for cardiac transplant, but was not eligible for listing due to irreversible pulmonary hypertension. A Jarvick ventricular assist device was implanted, with significant clinical improvement. However, the patient died after 4 months, at age 42, of cerebral haemorrhage.

The E258K mutation is a pathogenic G>A transition on the last nucleotide of exon 6 which results in a substitution of the amino acid lysine for glutamic acid at position 258 in cMyBP-C. This is the most studied and prevalent *MYBPC3* mutation, with 39 probands identified in 11 independent studies [9], and occurs in 13 % of all HCM patients in Tuscany [10]. This variant has an allele frequency of 0.004 % in ExAC database (http://exac.broadinstitute.org/variant/). It is usually associated with a severe phenotype, a poor prognosis and is highly penetrant [1]. Recently, Mearini provided the first evidence of successful correction of the phenotype by 5′ trans splicing in vivo in a mouse model [11], thus demonstrating the importance of the expressed mutant protein in causing the disease. This manuscript applies to one of the three mutant transcripts predicted to occur with the c.772G>A mutation: The mutant-3 that causes deletion/insertion is due to skipping of exon 6. The E441K is a G>A transition which results in a substitution of the amino acid lysine for glutamic acid at position 441 in cMyBP-C. The E441K variant has an allele frequency of 0.016 % in ExAC database (http://exac.broadinstitute.org/variant/). This mutation was first described in 2005 by Seidman et al. (Cardiogenomics) [12] and in 2009 by Marsiglia et al. (JDC 2009) [13] in compound heterozygosity with E258K. Marsiglia suggested that the combination of the two mutations might be responsible for severe phenotype, whereas, in isolation, both are capable of causing the disease but in its milder form.

The two *MYBPC3* mutations were identified following screening by automatic sequencing of the eight sarcomeric

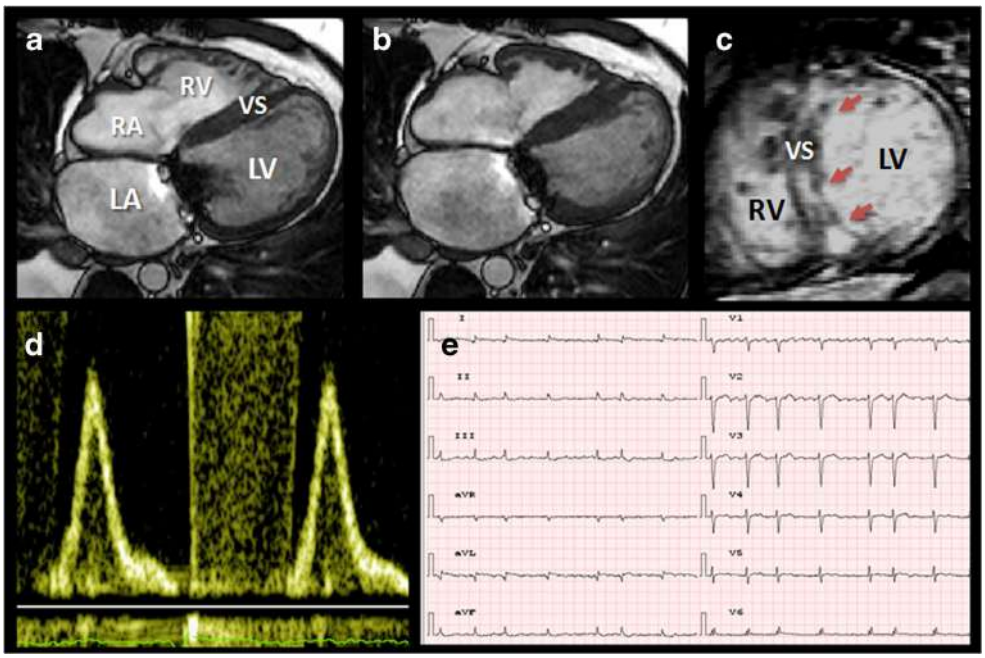


Fig. 1 Features of the HCM patient with double mutation in the gene coding for cMyPB-C at age 41. **a, b** End-diastolic and end-systolic MRI frames in four chamber views showing severe systolic dysfunction, LV dilatation with mild residual septal hypertrophy, prior mitral plasty and bi-atrial dilation. **c** MRI short axis view showing diffuse septal late gadolinium enhancement (*arrows*) indicating fibrous substitution of the myocardium. **d** Echocardiographic pulsed wave Doppler trans-mitral flow showing severe diastolic dysfunction with restrictive pattern. **e** 12-Lead ECG in atrial fibrillation

genes including *MYH7*, *MYBPC3*, *TNNT2*, *TNNI3*, *ACTC1*, *TPM1*, *MYL2* and *MYL3* by standard Sanger method [10].

Cell Culture and Transfection

The rat embryonic, heart-derived cardiomyoblast cell line, H9C2 was obtained from American Type Culture Collection (ATCC, Manassas, VA). Cells were cultured in Dulbecco modified Eagle medium (DMEM; InvitrogenGibco, Carlsbad, CA, USA) supplemented with heat-inactivated 10 % foetal bovine serum (FBS; Gibco, MD) and antibiotics (100 units/ml penicillin and 100 μg/ml streptomycin) in a water-saturated atmosphere of 5 % CO_2 at 37 C. For transient transfection of GFP-tagged WT and mutant *MYBPC3* (E258K, E441K & E258K-E441K; GeneArt Gene Synthesis, Life Technologies, CA, USA), cells were transfected in OptiMEM with Lipofectamine reagent 2000 (Invitrogen, Carlsbad, CA, USA) as per the manufacturer's instructions.

Western Blotting Analysis

To perform Western blot analyses, cells were lysed in RIPA buffer after 36 h of transfection. Protein concentrations were determined using the Bradford reagent (Bio-Rad Laboratories Inc., CA, USA) and resolved on NuPAGE 4–12 % Bis-Tris gel or NuPAGE 3–8 % Tris-acetate gels (Life Technologies, CA, USA), transferred to polyvinylidene difluoride (PVDF) membranes, and probed with respective antibodies. A 1:500 dilution of the primary antibody was used to probe for MYBPC-3 and 1:2000 for anti-GFP (Sigma-Aldrich, St. Louis, MO, USA). Anti-β-actin (1:5000; Sigma-Aldrich, St. Louis, MO, USA) was used to normalize for equal amounts of proteins and calculate the relative induction ratio. Peroxidase-conjugated respective secondary antibodies were used to label the proteins and detected using ECL reagent and Bio-Rad Quantity One (CA) system.

Statistical Analysis

Data analysis was performed using Origin 7.0 (OriginLab, Northampton, MA). Statistical comparisons were made using Student's *t* test. Experimental values are expressed as means± SEM.

Molecular Modeling

Building Structure for the Motif

The structure of cMyBP-C contains 11 globular domains including eight immunoglobulin (Ig)-like domains and three fibronectin (Fn)-like domains termed C0–C10 [14]. The two mutations in our patient were located in domains C1 (E258K) and C2 (E441K), for which 3D structures are available. These domains are connected by a linker region also known as motif (m), for which the 3D structure is unknown. However, a partial 3D structure of mouse cMyPB-C motif is available [15]. The complex C1-m-C2 is reported to be a region responsible for phosphorylation and for the regulation of the molecular mechanisms of muscle contraction via its interaction with actin and myosin [16].

The 3D structural data of the domains C1 (pdb ID: 2V6H) [17] and C2 (pdb ID: 1PD6) [18] were obtained from the protein data bank (www.rcsb.org). There are no structures available for human cMyBP-C motif. Thus, the amino acid sequence of the motif was taken from Uniprot for model building, and the iterative threading assembly refinement (I-TASSER) [19] was utilized to generate the models of the motif using the multiple threading method. The top 5 models with reliable structural properties were selected for quality assessment via Ramachandran analysis. The structure with the most favourable regions (98.8 % in the favourable region and 1.2 % (1 residue) in the disallowed region) was selected for the complex construction after energy minimization.

Construction of the Complex C1-m-C2

The complex of C1-m-C2 was generated by molecular docking using ClusPro [20]. A set of 40 structures were generated for the complex, and the model with lowest energy, highest affinity score and most reliable electrostatic properties was selected for further study. This complex was energy minimized to remove constraints using the parameters as stated in the molecular dynamics (MD) simulation section.

Four systems were developed from the complex C1-m-C2 for the MD simulations. These include the complex (i) without mutations (WT), (ii) double mutation, (iii) E258K in domain C1 (E258K) and (iv) E441K in domain C2 (E441K). Discovery studio (DS) [21] v3.5 was used to introduce mutations to the complexes as performed by Krishnamoorthy et al. [22].

MD Simulations of the Complexes

The GROMACS simulation package (v4.5.4) [23, 24] was used to perform the MD simulations, with explicit water and physiological conditions. The systems were solvated using the SPC3 [25] water model in a 0.8 nm cubic box, with applied periodic boundary conditions in all directions. The required counter ions were added to neutralize the systems. The resulting systems contained ~50,000 atoms. The systems were energy-minimized using the steepest descent algorithm with a tolerance of 2000 kJ/mol/nm, and the resulting structures were used as the starting structures for MD simulations by applying

GROMOS96 force field [26]. Van der Waals and electrostatic interactions were observed with a twin range cutoff of 0.8 nm and 1.4 nm, respectively. The LINCS algorithm [27] was employed to constraint the lengths of all the bonds and the SETTLE [28] to constraint the geometry of the water molecules. After energy minimization, the structures were pre-equilibrated for 100 ps. This was followed by 10 ns of production MD simulations with a time step of 2 fs, at constant temperature (300 K), pressure (1 atm) and number of particles, without any position restraints [29]. The trajectories of the simulations were collected at every 5 ps for various quantitative analyses using GROMACS tools. The interactions at the interface of the complexes were analysed using Ligplot [30] for the representative structures that are obtained by utilizing cluster analysis in GROMACS. In addition, structural analyses were carried out using DS and PyMOL (www.pymol.org).

Electrostatic Surface Calculation

Studying electrostatic properties could reveal key intra- and inter-molecular interactions, which provide information for structure–function relationships [31, 32]. Thus, here, we used the Delphi module of the DS to calculate the surface electrostatic potential for the WT and the mutants. The charges for the structures were applied using the Delphi force field, and the mapping of surface electrostatic potential was done by solving the Poisson-Boltzmann equation.

Results

Stability of Expressed Mutant Proteins in Mammalian Cells

It is important to examine the expression and stability of the single and double point mutations as a recent study showed complexity in the expression of point mutations [33]. To determine the isolated effect of E258K, E441K and their combination (E258K-E441K) on protein stability, we transfected H9C2 cells with WT or single or double mutation of *MYBPC3* constructs tagged with GFP. Single mutations and double mutation had no effect on MyBP-C protein level when compared with WT. Similarly, no significant changes were observed in β-actin levels (Fig. 2a and Supplement Fig. 1). The quantification showed that there was no significant difference in *MYBPC3* protein level between WT and mutant transcripts (Fig. 2b). Collectively, these results suggested that both single and double mutation did not affect the *MYBPC3* protein level. However, to understand the structural consequences of the expressed mutant proteins, here, we have used molecular modeling.

Structural Features of the Modelled Motif

We used the available partial 3D structure of the cMyBP-C motif (pdb ID: 2LHU) as a primary template to model the structure of human cMyBP-C motif. The structure of the

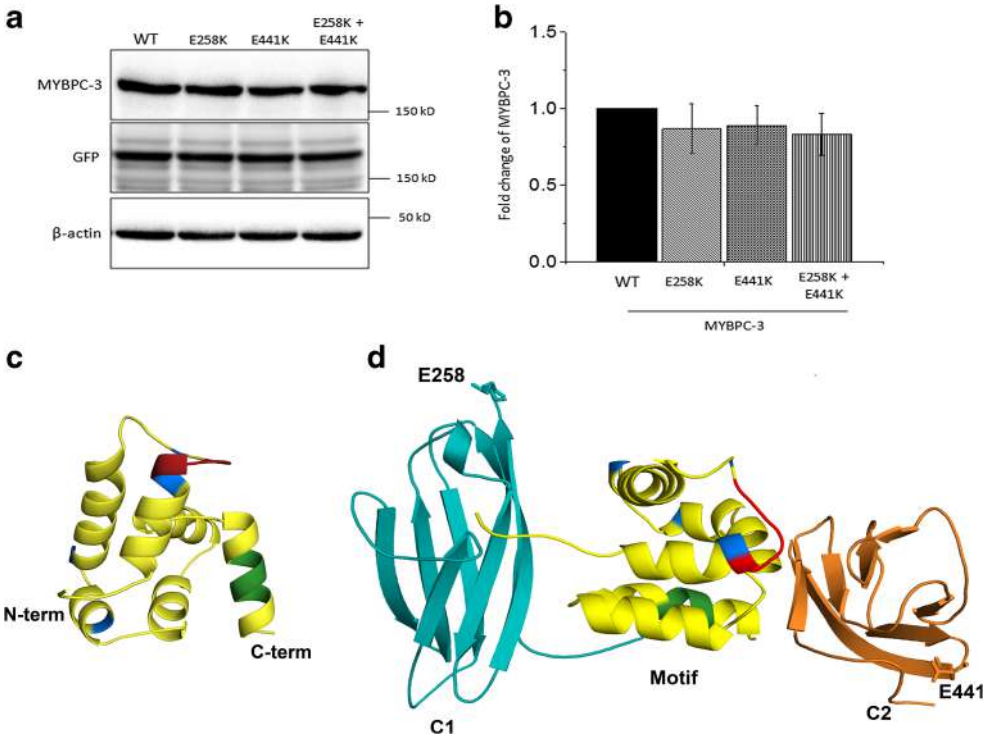


Fig. 2 C1-m-C2 complex with the constructed motif. **a** Effects of mutations on the protein stability in mammalian cells. H9C2 cells were transfected with GFP tagged wild-type (*WT*) or *MYBPC3* variants. Cell lysates were used to detect the protein level by immunoblotting. **b** Histogram represents the relative intensity of the protein normalized to β-actin. Values represent mean±SEM (n=3 experiments). **c** The modelled structure of the motif. **d** The complex C1-m-C2 including the constructed motif, where the native residues at position 258 and 441 are shown in sticks. The cardiac-specific loop, sites of phosphorylation and actin-binding site in the motif are represented in *red*, *blue* and *green*, respectively

C-subdomain of the motif was resolved recently from mouse with three helices. The remaining two helices at the N-subdomain have not been resolved [19]. In Fig. 2c, our model shows a plausible complete structure of a human cMyBP-C motif composed of five helical structures, as suggested previously, and a small helix that connects the subdomains N and C. The model also shows suitable positioning of the LAGGG RRIS loop at the N-terminal in the vicinity of the sites of phosphorylation. This signature loop is specific to cardiac isoforms, and its arrangement close to the sites might be required to regulate the site-specific phosphorylation [34]. Furthermore, the motif provides four accessible phosphorylation sites that are known: Ser275, Ser284, Ser304 and Ser311, and a well-exposed potential binding site LK(R/K)XK for actin [22] at the vicinity of the C-terminal. The observed structural features of the motif that correlates with the literature suggest that the constructed model is biologically reasonable.

The C1-m-C2 complex was constructed using molecular docking method (Fig. 2d). This WT complex was further used to produce the mutated complexes. Both the mutations, E258K from domain C1 and E441K from domain C2, are located at the C-terminal of the relevant domains, where the former mutation is adjacent to the functionally important motif.

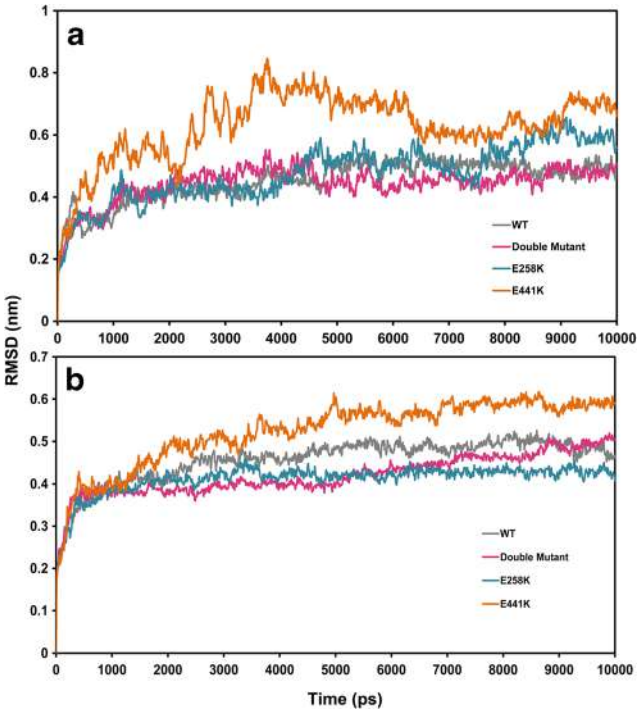


Fig. 3 Trajectory-based analyses for the complexes. Structural deviation (root mean square deviation (RMSD) of **a** the complexes and **b** the motif during the MD simulations

Structural Deviation of Protein Complexes in Dynamics

The trajectory-based analyses were applied to understand the overall structural deviation of the systems during dynamics. The root mean square deviation (RMSD) was calculated for the Cα atoms with respect to the initial structures throughout the MD simulations (Fig. 3a). The result showed that the WT and double mutation retained around 0.45 nm, though each followed a slightly distinct pattern. The behaviour of the systems with single mutation was entirely different. E258K slightly deviated after 8 ns by reaching 0.6 nm, whereas E441K reached 0.8 nm approximately at 4 ns. The deviation of individual domains was also calculated via RMSD with reference to the initial structure, which indicated that domains C1 and C2 were stable with average RMSD of 0.3 nm. However, the motif deviated more than 0.4 nm in all the systems (Fig. 3b). These results showed that the increase in the overall structural deviation might be due to both flexibility of the motif and effect of the mutations. Although this result provides overall stability of the complexes, further analyses required to understand the structural properties of this complex.

Conformational Changes of the Domains

To examine the conformational changes due to the mutations, a representative snapshot was taken for each complex using the cluster analysis within the GROMACS package. In this analysis, 2000 structures from 10 ns of each simulation were used to make different clusters, from which a representative structure was in turn selected from a top cluster with the highest number of recurrent structures and subjected to further structural analyses. Each domain was analysed individually to understand the structural consequences of the mutations.

Conformational changes were observed in all three mutated systems when they were compared with WT (Fig. 4). In the double mutation, near mutational spots, the C-terminal of domain C1 and domain C2 showed significant conformational changes (Fig. 4a), which are greater than that observed in E258K and E441K (Fig. 4b, c). However, the motif in the double mutation displayed minimal conformational shift compared to its shift in the single mutant systems. The systems with single mutation, E258K and E441K, showed that they could largely affect their native domain, C1 and C2, respectively (Fig. 4b, c). These results indicate that the single mutations studied here predominantly affect their native domains as well as the nearby motif. When they coexist, they are likely to have an additive effect on the complex.

Structural Deviations at the Functionally Important Regions of the Motif

Interestingly, conformational changes were noticed in the mutated systems on the key regions of the motif including phosphorylation sites (Ser275, Ser284, Ser304 and Ser311), the cardiac-specific loop LAGGGRRIS (reported as a regulator of phosphorylation) (Fig. 5) and the potential actin-binding

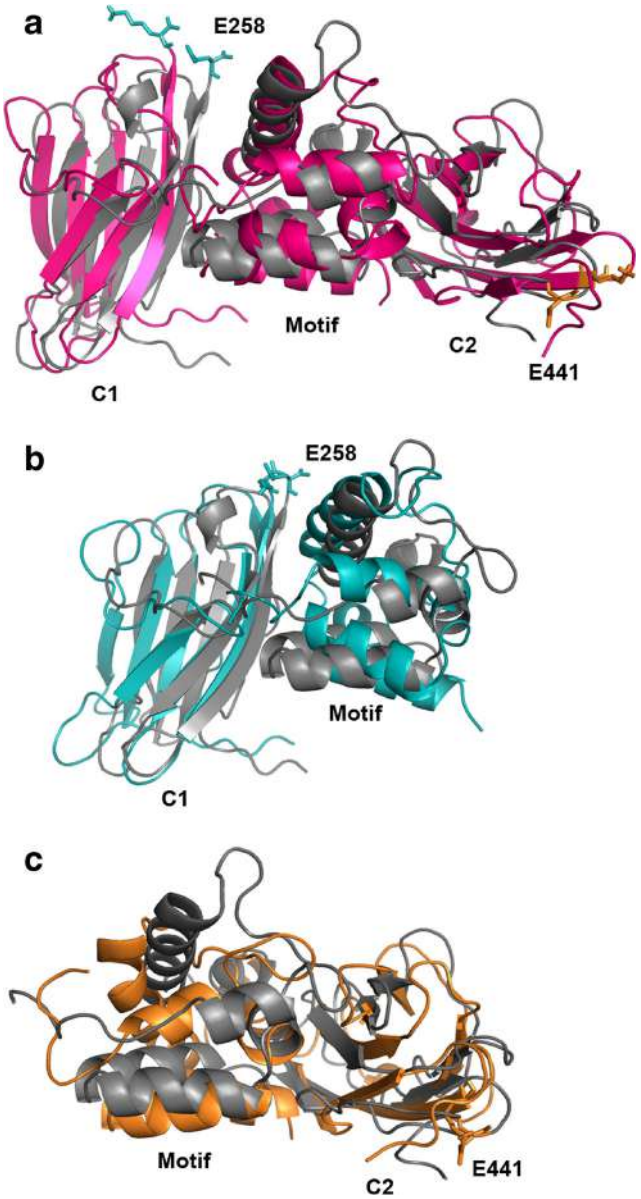


Fig. 4 Structural deviations of the domains. The representative snapshots of the WT and the mutants are superimposed with WT. **a** WT with double mutation, **b** WT with E258K and **c** WT with E441K. Here, WT, double mutation, E258K and E441K are shown in *grey*, *pink*, *aqua green* and *orange*, respectively. The residues 258 and 441 are represented in sticks. For the purpose of clarity, only the native domain and the motif are displayed in panels **b** and **c**

site (Fig. 6) LKRLK. The systems with double mutation and E441K largely affected the sites of phosphorylation and the cardiac-specific loop by shifting the coordinates of the relevant residues. It is important to note that the exposed Ser284 of WT was buried in the double mutation and in the E441K (Fig. 5b and d), whose phosphorylation is reported to be a prerequisite for phosphorylation of the rest of the sites. Surprisingly, in E258K, minimal effects were observed at these sites (Fig. 5c). The linker between the subdomains N and C of the motif displayed flexible behaviour in all the mutant systems. The mutations also induced prominent shifts on the backbone of the residues of the actin-binding sites at the C-terminal of the motif (Fig. 6). In which, both the double mutation and E441K in isolation induce maximal shift. Altogether, the structural shifts on the key regions of the motif indicate that both the double mutation and E441K induce severe changes on these sites compared to E258K. These changes might affect the binding of the C1-m-C2 to myosin S2 and actin, which could also impact phosphorylation.

Interactions at the Protein–Protein Interface

Studying the interactions at the interface of the C1-motif and motif-C2 is essential to understand their complex structural organization for the function. The inter-domain interactions of C1-motif and motif-C2 are listed in Table 1.

The interactions at the interfaces of WT show that C1-motif has 27 interactions (10 hydrogen bonds and 17 hydrophobic interactions), while, in contrast, motif-C2 has 29 inter-domain contacts (7 hydrogen bonds and 22 hydrophobic interactions). This suggests that in the C1-m-C2 complex, the motif might have slightly stronger affinity towards C2 than C1. Double mutation decreased both hydrogen bonds and hydrophobic interactions between C1 and motif compared to WT. Strikingly, at the interface of motif and C2, double mutation increased the number of hydrogen bonds from 7 to 14, while there was a reduction of hydrophobic interactions from 22 to 12. In the single mutation systems, hydrogen bonds are decreased at the C1-motif interface compared to WT but hydrophobic interactions are maintained. At the motif–C2 interface of the single mutants, we also observed a dramatic increase of hydrogen bonds (from 7 to 15) and a decrease of hydrophobic interactions (from 22 to 14). These variations in the interaction patterns indicate that the conformational changes induced by mutations in the C1-m-C2 could disturb the formation of native interface between the domains.

Changes on the Surface Electrostatic Properties

Surface electrostatic properties of the representative structures showed changes in the surface charge at the mutated spots, as the native residues were negatively charged glutamic acids which were mutated with positively charged lysines (Fig. 7). This change in the mutational spots affects their neighbouring regions. In addition, mutation-induced conformational shifts cause changes to the electrostatic properties at the front and rear side of the mutated complexes. It is worth noting that in the WT complex, the roof of the cMyBP-C motif creates an elevated surface with a distribution of negatively charged residues. By contrast, in the double mutation and in E258K, the C1 domain shows an additional elevation that is positively charged on its surface due to the mutation. As suggested by

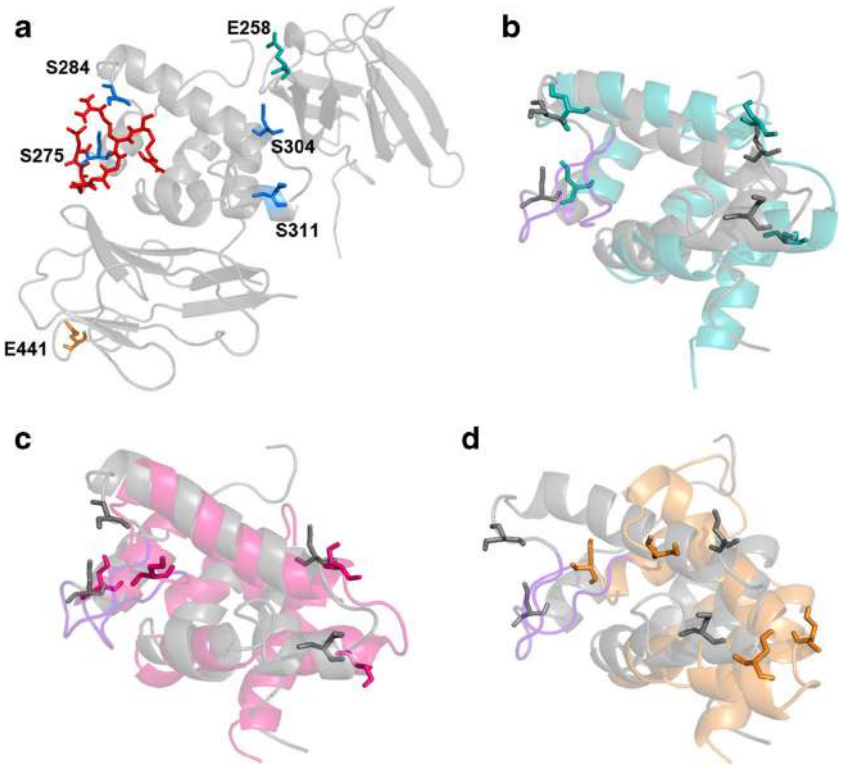


Fig. 5 Conformational changes at the site of phosphorylation. **a** The complex C1-m-C2 is shown with the phosphorylation site (*blue sticks*) and the cardiac specific loop (*red sticks*). The residues E258 and E441 are in *aqua green* and *orange sticks*, respectively. **b**–**d** superimposition of WT motif with **b** double mutation, **c** E258K and **d** E441K. Here, the colours *grey*, *pink*, *aqua green* and *orange* represent WT, double mutation, E258K and E441K, respectively. The cardiac-specific LAGGGRRIS loop is coloured in *violet* from **b** to **d**

earlier studies [21, 22], this might interfere with the native affinity of C1 and the motif towards their binding partners, myosin or/and actin. In E441K, due to severe conformational changes, the elevated shape of the motif on the top of the complex is flattened. The observed changes to the surface charge and shape of the complex reveal that the reverse charge mutations can impact profoundly on the electrostatic properties of C1-m-C2.

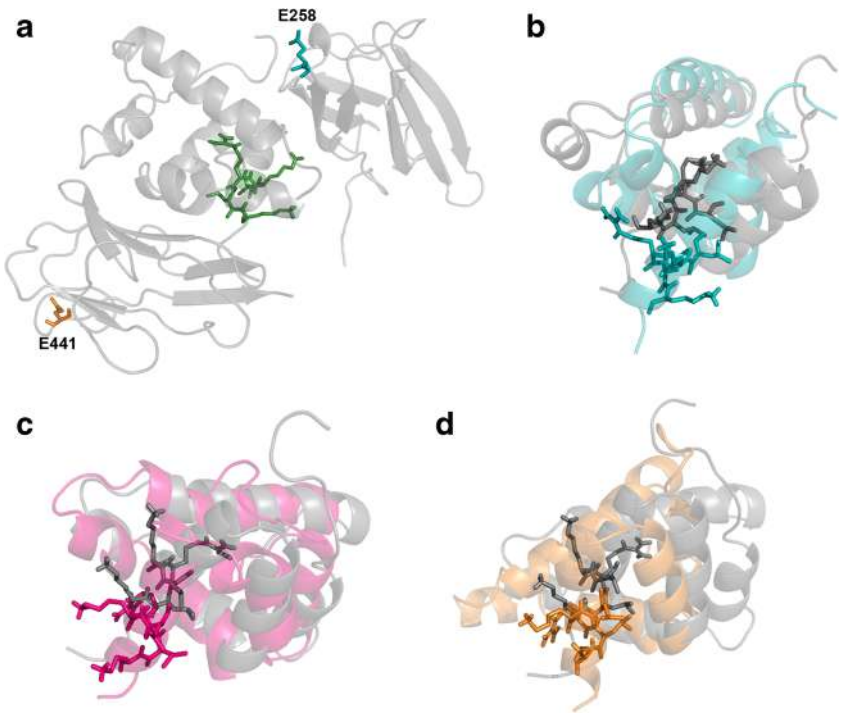


Fig. 6 Conformational changes at the potential actin-binding site. **a** Complex C1-m-C2 is shown with the location of mutated residues along with the actin binding site (*green sticks*). **b**–**d** Superimposition of WT motif with **b** double mutation, **c** E258K and **d** E441K. Here, the colour codes are used as in Fig. 5

Table 1 Inter-domain interactions of the complex C1-m-C2

Systems	C1 and motif		C2 and motif	
	Hydrogen bond	Hydrophobic interaction	Hydrogen bond	Hydrophobic interaction
WT	Asp 163- Lys350, Asn251-Lys350 Asn251-Arg346 Lys195-Gly261 Glu165=Thr307, Glu165-Arg306 Gly164-Arg306, Thr255=Arg306	**C1:** Tyr237, Ser236, Arg238, Gly194, Lys193, Trp196, Asn253, His257 **Motif:** Val342, Thr343, Thr262, Tyr340, Met260, Ala259, Asp303, Phe305, Ser304	Glu319-Lys380 Ala277-Lys418 Arg326-Leu369, Arg326-Glu370; Gly354-Gln374, Arg356-Gln374 Arg357-Val449	**Motif:** Gly279, Gly278, Glu323, Trp322, Glu314, Asp310, Leu313, Gln327, Pro329, Pro330, Met355 **C2:** Ile415, Gly416, Ala417, Arg382, Ile381, His379, Thr384, Tyr373, Ala372, Phe448, Asp431
Double mutation	Glu165-Ser304, Arg160-Arg309, Asn253-Asp303, Ser236= Met260, Lys195-Gly263, Arg238-Gly261, Asn251-Tyr340, Pro153-Arg346	**C1:** Asp163, Gln162, Asp248, Ser250, Cys249, Gly235, Gly194, **Motif:** Lys312, Arg306, Val342, Ala259, Thr262, Thr343	Gly278-Gly416, Gln327-Thr420, Gln327-Arg382, Arg357-Gln374, Arg326=Ala372, Arg326-Glu370; Gly354-Gln374, Met355-Gln374, Trp322-Lys380, Glu323-Arg382, Glu319-Arg382, Glu319-Lys380, Glu318-Lys380	**Motif:** Leu276, Ala277,Gly279, Arg356, Ala315 **C2:** Ile415, Lys418, Thr384, Phe448, Tyr373, Ile381, His379
E258K	Glu165=Thr307, Arg306; Ser236-Lys312, Asn253-Lys312, Thr255-Arg306, Lys195-Thr262, Trp196-Gly261	**C1:** Asn251, His257, Glu164, Asp163, Gly235, Leu254, Thr234, Lys193, Gly194 **Motif:** Thr343, Ser304, Phe305, Arg309, Asp303, Asp264, Ala259, Met260	Glu318=Ser376, Glu319-Lys380, Arg357-Gln374, Met355-Gln374, Arg326=Ala372, Arg326-Glu370, Arg326-Thr384, Arg356-Glu370, Gln327-Thr420, Gln327(3HB)Lys418, Leu276-Lys418	**Motif:** Val375, His379, Ile381, Arg382, Phe448, Tyr373, Leu369, Ile415 **C2:** Trp322, Glu323, Gly354, Ala328, Ala277, Gly278
E441K	Arg328-Asp264, Gly194-Asp264, Ser236-Gly261, Asn253-Ser296, Thr255-Asp294, His257-Asp294, Asp163-Lys300, Asp163-Lys301	**C1:** Ser250, Lys195, Lys193, Leu254, Gly235, Thr234, Val256, Arg160, Asn251 **Motif:** Thr343, Gly263, Thr262, Leu267, Leu293, Met260, Thr290, Val343, Ser297, Leu265	Arg357-Gln374, Arg326-Ala372, Arg326-Glu370, Arg326-Thr384, Gly354-Gln374, Met355, Gln374, Glu319-Lys380, Glu318-Ser376, Lys312-His379, Gln327-Thr420	**Motif:** Arg356, Trp322, Glu323, Gly278, Leu313 **C2:** Lys450, Phe448, Tyr373, Ile381, Val375, Arg382, Lys418, Ile415, Gly416, Gly378

The bold letters indicate the structural domains (C1, Motif and C2) of the complex

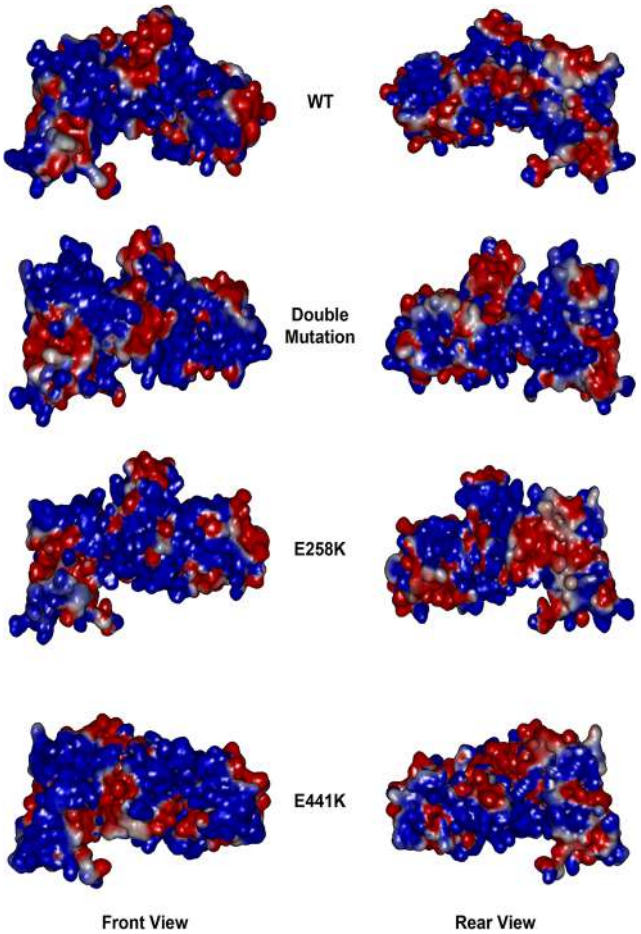


Fig. 7 Surface electrostatic potential. Front and rear view of the map of the surface electrostatic potential for the representative snapshots. Here, *blue*, *red* and *white* represent positive, negative and hydrophobicity, respectively

Discussion

This study on the complex C1-m-C2 suggests possible structural consequences of double mutation in the gene coding for cMyBP-C, accounting for the severe HCM phenotype observed in our patient. In MD simulations, both mutations produced distinctive and considerable changes on the sites of phosphorylation and potential binding regions of actin and myosin. These sites are affected through modifications of the complex conformation, interface and charge reversal on electrostatic properties at surface.

Previous studies have shown that mutation in *MYBPC3* alters the protein stability and produce truncated protein. The truncated proteins were rapidly removed from the cardiomyocytes by ubiquitin-proteasome system (UPS), and impairment in UPS function contributes to cardiac dysfunction [35]. Several studies have shown that truncation mutations in cMyBP-C diminish the total protein level and cause hypertrophic cardiomyopathy through haploinsufficiency [36–38]. However, a recent study suggests that consequences of the mutations cannot be predicted solely based on the type or location of the mutation [33]. Our in vitro results showed that full length cMyBP-C protein level was not significantly altered between wild-type, single mutations and double mutation, indicating that both mutant mRNA and protein are stable and UPS is unlikely to be involved in the regulation of mutant protein clearance in H9C2 cells. However, further studies are needed to validate these results in in vivo by using an animal model or cardiac tissues of HCM patients carrying the E258K or E441K or double cMyBP-C mutation.

Multiple phosphorylation sites on cMyBP-C suggest its role in normal cardiac function. Mutations in these sites reduce contractile function [34, 39, 40]. In addition, it has been suggested that phosphorylation of Ser284 in humans (Ser282 in mouse) is a prerequisite for phosphorylation of Ser275, Ser304 and Ser311 [34, 41]. This study provides structural insight into the mutational effect on the phosphorylation sites. The drastic conformational shift on the sites, in particular the burial of Ser284 and the deviation of the cardiac specific loop due to double mutation (as well as E441K in isolation), might affect normal cardiac function. The mutation E258K is adjacent to Ser304. However, it did not appear to affect any site of phosphorylation in a major way. This suggests that the E258K mutation might affect the binding of C1 to S2 of myosin directly by changing the surface electrostatic properties [13, 21, 22]. This is in agreement with our previous study on E258K [10]. However, the effects of mutations on the sites of phosphorylation have to be validated.

The biochemical or biophysical analyses and yeast two hybrid experiments [42–44] reported that the domain C1, motif and/or C1-m-C2 complex interacts with actin mainly via lysine residues [45]. The mutated systems appear to severely affect the backbone of the residues of potential actin-binding site LKRLK [19] in the motif. This indicates that the interaction between residues of the motif and actin might be based on electrostatic properties. Thus, the charge reversal mutations (single or double) studied here can induce modifications to surface electrostatic properties via conformational changes on the binding site and consequently impact its interaction with actin [46, 47].

Although this study suggests the structural abnormalities caused by the mutations in the domain C1-m-C2 using in silco models along with the expression data of the mutant proteins, it has a few limitations. In particular, it is unable to address the structural consequences of the complete structure of the protein and the complex interaction with other sarcomeric proteins. However, these limitations have also been a challenge for NMR and X-ray studies since the protein was identified, and the complete structure with its binding partners is yet to be resolved. Experimental evidences for the changes in the sites of phosphorylation and protein–protein interactions will be reported in the future.

Conclusions

In the contractile apparatus, cMyBP-C binds to both the thick and thin filament systems involved in organizing sarcomeric structure and cross-bridge regulation, in which the N-terminal of cMyBP-C, C1-m-C2, also plays a key role. This modeling study explains the importance of the conservation of conformational specificity, interface and electrostatic properties of the native structural organization of C1-m-C2 in order to retain its key functions. This type of structural conservation is affected considerably when E258K and E441K coexist, potentially impacting normal phosphorylation and binding with both myosin and actin. The studied mutational effects might explain the severe consequences of double mutation affecting these domains. This modeling may be helpful in designing structural hypotheses (phosphorylation and protein–protein interactions) for mechanism of the disease, and they need to be validated experimentally using small angle X-ray scattering, electron microscopic methods, NMR and X-ray crystallography. It is hoped that these observations will help to identify future therapeutic targets for HCM.

Acknowledgments We would like to thank Dr. Othmane Bouhali, Director of the Research Computing and Mr. Faisal Chaudhry, Senior Lead Systems Engineer, Texas A&M University in Qatar for providing supercomputing facility. We also extend our thanks to Dr. Brian P. Mitchelson, QCRC and Mr. Mark Radford for their suggestions. No human or animal studies were carried out by the authors of this article.

Funding This work was supported by Qatar Foundation through Qatar Cardiovascular Research Center, Doha, Qatar, and it was also supported by Magdi Yacoub Research Network, London, UK. IO and FG are supported by the Italian Ministry of Health (RF 2010 – 2313451 "Hypertrophic cardiomyopathy: new insights from deep sequencing and psychosocial evaluation") and NET-2011-02347173 (Mechanisms and treatment of coronary microvascular dysfunction in patients with genetic or secondary left ventricular hypertrophy). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Conflict of Interest None declared.

References

1. Richard, P., Charron, P., Carrier, L., Ledeuil, C., Cheav, T., Pichereau, C., Benaiche, A., Isnard, R., Dubourg, O., Burban, M., Gueffet, J. P., Millaire, A., Desnos, M., Schwartz, K., Hainque, B., & Komajda, M. (2003). EUROGENE Heart Failure Project. Hypertrophic cardiomyopathy: distribution of disease genes, spectrum of mutations, and implications for a molecular diagnosis strategy. *Circulation, 107*, 2227–2232. doi:10.1161/01.CIR.0000066323.15244.54.
2. Maron, B. J., Gardin, J. M., Flack, J. M., Gidding, S. S., Kurosaki, T. T., & Bild, D. E. (1995). Prevalence of hypertrophic cardiomyopathy in a general population of young adults. Echocardiographic analysis of 4111 subjects in the CARDIA study. Coronary artery risk development in (Young) adults. *Circulation, 92*, 785–789. doi:10.1161/01.CIR.92.4.785.
3. Maron, B. J., Shirani, J., Poliac, L. C., Mathenge, R., Roberts, W. C., & Mueller, F. O. (1996). Sudden death in young competitive athletes. Clinical, demographic, and pathological profiles. *JAMA, 276*, 199–204. doi:10.1001/jama.1996.03540030033028.
4. Cecchi, F., Yacoub, M. H., & Olivotto, I. (2005). Hypertrophic cardiomyopathy in the community: why we should care. *Nature Clinical Practice Cardiovascular Medicine, 2*, 324–325. doi:10.1038/ncpcardio0248.
5. Olivotto, I., Cecchi, F., Poggesi, C., & Yacoub, M. H. (2012). Patterns of disease progression in hypertrophic cardiomyopathy: an individualized approach to clinical staging. *Circulation Heart Failure, 5*, 535–546. doi:10.1161/CIRCHEARTFAILURE.112.967026.
6. Yacoub, M. H., Olivotto, I., & Cecchi, F. (2007). 'End–stage' hypertrophic cardiomyopathy: from mystery to model. *Nature Clinical Practice Cardiovascular Medicine, 4*, 232–233. doi:10.1038/ncpcardio0859.
7. Frey, N., Luedde, M., & Katus, H. A. (2011). Mechanisms of disease: hypertrophic cardiomyopathy. *Nature Reviews Cardiology, 9*, 91–100. doi:10.1038/nrcardio.2011.159.
8. Kelly, M., & Semsarian, C. (2009). Multiple mutations in genetic cardiovascular disease: a marker of disease severity? *Circulation Cardiovascular Genetics, 2*, 182–190. doi:10.1161/CIRCGENETICS.108.836478.
9. De Lange, W. J., Grimes, A. C., Hegge, L. F., Spring, A. M., Brost, T. M., & Ralphe, J. C. (2013). E258K HCM-causing mutation in cardiac MyBP-C reduces contractile force and accelerates twitch kinetics by disrupting the cMyBP-C and myosin S2 interaction. *Journal of General Physiology, 142*, 241–255. doi:10.1085/jgp.201311018.
10. Olivotto, I., Girolami, F., Ackerman, M. J., Nistri, S., Bos, J. M., Zachara, E., Ommen, S. R., Theis, J. L., Vaubel, R. A., Re, F., Armentano, C., Poggesi, C., Torricelli, F., & Cecchi, F. (2008). Myofilament protein gene mutation screening and outcome of patients with hypertrophic cardiomyopathy. *Mayo Clinic Proceedings, 83*, 630–638. doi:10.4065/83.6.630.
11. Mearini, G., Stimpel, D., Krämer, E., Geertz, B., Braren, I., Gedicke-Hornung, C., Précigout, G., Müller, O. J., Katus, H. A., Eschenhagen, T., Voit, T., Garcia, L., Lorain, S., & Carrier, L. (2013). Repair of Mybpc3 mRNA by 5′-trans-splicing in a mouse model of hypertrophic cardiomyopathy. *Molecular Therapy Nucleic Acids, 2*, e102. doi:10.1038/mtna.2013.31.
12. Merk, S. (2005). *Disease-causing mutations in the human cardiac myosin binding protein C gene*. Genomics of cardiovascular development, adaptation, and remodeling. NHLBI Program for Genomic Applications, Harvard Medical School. Available at: http://www.cardiogenomics.org [June 2010].
13. Marsiglia, J. D., Batitucci Mdo, C., Fd, P., Barbirato, C., Arteaga, E., & Araújo, A. Q. (2010). Study of mutations causing hypertrophic cardiomyopathy in a group of patients from Espirito Santo, Brazil. *Arquivos Brasileiros de Cardiologia, 94*, 10–17.
14. Einheber, S., & Fischman, D. A. (1990). Isolation and characterization of a cDNA clone encoding avian skeletal muscle C-protein: an intracellular member of the immunoglobulin superfamily. *Proceedings of the National Academy of Sciences of the United States of America, 87*, 2157–2161. doi:10.1073/pnas.87.6.2157.
15. Howarth, J. W., Ramisetti, S., Nolan, K., Sadayappan, S., & Rosevear, P. R. (2012). Structural insight into unique cardiac myosin-binding protein-C motif: a partially folded domain. *Journal of Biological Chemistry, 287*, 8254–8262. doi:10.1074/jbc.M111.309591.
16. Oakley, C. E., Chamoun, J., Brown, L. J., & Hambly, B. D. (2007). Myosin binding protein-C: enigmatic regulator of cardiac contraction. *International Journal of Biochemistry and Cell Biology, 39*, 2161–2166. doi:10.1016/j.biocel.2006.12.008.

17. Govada, L., Carpenter, L., da Fonseca, P. C., Helliwell, J. R., Rizkallah, P., Flashman, E., Chayen, N. E., Redwood, C., & Squire, J. M. (2008). Crystal structure of the C1 domain of cardiac myosin binding protein-C: implications for hypertrophic cardiomyopathy. *Journal of Molecular Biology, 378*, 387–397. doi:10.1016/j.jmb.2008.02.044.
18. Ababou, A., Gautel, M., & Pfuhl, M. (2007). Dissecting the N-terminal myosin binding site of human cardiac myosin-binding protein C. Structure and myosin binding of domain C2. *Journal of Biological Chemistry, 282*, 9204–9215. doi:10.1074/jbc.M610899200.
19. Roy, A., Kucukural, A., & Zhang, Y. (2010). I-TASSER: a unified platform for automated protein structure and function prediction. *Nature Protocols, 5*, 725–738. doi:10.1038/nprot.2010.5.
20. Comeau, S. R., Gatchell, D. W., Vajda, S., & Camacho, C. J. (2004). ClusPro: an automated docking and discrimination method for the prediction of protein complexes. *Bioinformatics, 20*, 45–50. doi:10.1093/bioinformatics/btg371.
21. Discovery Studio. (2012). *3.5 User Guide*. Accelrys Inc., San Diego, CA, USA.
22. Krishnamoorthy, N., Yacoub, M. H., & Yaliraki, S. N. (2011). A computational modeling approach for enhancing self-assembly and biofunctionalisation of collagen biomimetic peptides. *Biomaterials, 32*, 7275–7285. doi:10.1016/j.biomaterials.2011.06.074.
23. Van Der Spoel, D., Lindahl, E., Hess, B., Groenhof, G., Mark, A. E., & Berendsen, H. J. (2005). GROMACS: fast, flexible, and free. *Journal of Computational Chemistry, 26*, 1701–1718. doi:10.1002/jcc.20291.
24. Hess, B., Kutzner, C., van der Spoel, D., & Lindahl, E. (2008). GROMACS 4: algorithms for highly efficient, load-balanced, and scalable molecular simulation. *Journal of Chemical Theory and Computation, 4*, 435–447. doi:10.1021/ct700301q.
25. Berendsen, H. J. C, Postma J. P. M, van Gunsteren W. F, Hermans J. (1981). *Interaction models for water in relation to protein hydration*. In: *Intermolecular forces*, edited by Pullman B, D. Reidel. Publishing Company. pp 331–342.
26. van Gunsteren, W. F., Billeter, S. R., Eising, A. A., Hünenberger, P. H., Krüger, P., Mark, A. E., Scott, W. R. P., & Tironi, I. G. (1996). *Biomolecular simulation: the GROMOS96 manual and user guide* (pp. 1–1042). Zürich: Vdf Hochschulverlag AG an der ETH Zürich.
27. Hess, B., Bekker, H., Berendsen, H. J. C., & Fraaije, J. G. E. M. (1997). LINCS: a linear constraint solver for molecular simulations. *Journal of Computational Chemistry, 18*, 1463–1472. doi:10.1002/(SICI)1096-987X(199709)18.
28. Miyamoto, S., & Kollman, P. A. (1992). SETTLE: an analytical version of the SHAKE and RATTLE algorithm for rigid water models. *Journal of Computational Chemistry, 13*, 952–962. doi:10.1002/jcc.540130805.
29. Berendsen, H. J. C., Postma, J. P. M., van Gunsteren, W. F., DiNola, A., DiNola, A., & Haak, J. R. (1984). Molecular dynamics with coupling to an external bath. *Journal of Chemical Physics, 81*, 3684–3691.
30. Wallace, A. C., Laskowski, R. A., & Thornton, J. M. (1995). LIGPLOT: a program to generate schematic diagrams of protein-ligand interactions. *Protein Engineering, 8*, 127–134. doi:10.1093/protein/8.2.127.
31. Sinha, N., & Smith-Gill, S. J. (2002). Electrostatics in protein binding and function. *Current Protein and Peptide Science, 3*, 601–614. doi:10.2174/1389203023380431.
32. Krishnamoorthy, N., Gajendrarao, P., Eom, S. H., Kwon, Y. J., Cheong, G. W., & Lee, K. W. (2008). Molecular modeling study of CodX reveals importance of N-terminal and C-terminal domain in the CodWX complex structure of Bacillus subtilis. *Journal of Molecular Graphics and Modelling, 27*, 1–12. doi:10.1016/j.jmgm.2008.01.009.
33. Helms, A. S., Davis, F. M., Coleman, D., Bartolone, S. N., Glazier, A. A., Pagani, F., Yob, J. M., Sadayappan, S., Pedersen, E., Lyons, R., Westfall, M. V., Jones, R., Russell, M. W., & Day, S. M. (2014). Sarcomere mutation-specific expression patterns in human hypertrophic cardiomyopathy. *Circulation Cardiovascular Genetics, 7*(4), 434–443.
34. Gautel, M., Zuffardi, O., Freiburg, A., & Labeit, S. (1995). Phosphorylation switches specific for the cardiac isoform of myosin binding protein-C: a modulator of cardiac contraction? *EMBO Journal, 14*, 1952–1960.
35. Bahrudin, U., Morisaki, H., Morisaki, T., Ninomiya, H., Higaki, K., Nanba, E., Igawa, O., Takashima, S., Mizuta, E., Miake, J., Yamamoto, Y., Shirayoshi, Y., Kitakaze, M., Carrier, L., & Hisatome, I. (2008). Ubiquitin-proteasome system impairment caused by a missense cardiac myosin-binding protein C mutation and associated with cardiac dysfunction in hypertrophic cardiomyopathy. *Journal of Molecular Biology, 384*(4), 896–907. doi:10.1016/j.jmb.2008.09.070.
36. Marston, S., Copeland, O., Jacques, A., Livesey, K., Tsang, V., McKenna, W. J., Jalilzadeh, S., Carballo, S., Redwood, C., & Watkins, H. (2009). Evidence from human myectomy samples that MYBPC3 mutations cause hypertrophic cardiomyopathy through haploinsufficiency. *Circulation Research, 105*, 219–222. doi:10.1161/CIRCRESAHA.109.202440.
37. Marston, S., Copeland, O., Gehmlich, K., Schlossarek, S., & Carrier, L. (2012). How do MYBPC3 mutations cause hypertrophic cardiomyopathy? *Journal of Muscle Research and Cell Motility, 33*, 75–80. doi:10.1007/s10974-011-9268-3.
38. van Dijk, S. J., Dooijes, D., dos Remedios, C., Michels, M., Lamers, J. M., Winegrad, S., Schlossarek, S., Carrier, L., ten Cate, F. J., Stienen, G. J., & van der Velden, J. (2009). Cardiac myosin-binding protein C mutations and hypertrophic cardiomyopathy: haploinsufficiency, deranged phosphorylation, and cardiomyocyte dysfunction. *Circulation, 119*, 1473–1483. doi:10.1161/CIRCULATIONAHA.108.838672.
39. Sadayappan, S., Gulick, J., Osinska, H., Martin, L. A., Hahn, H. S., Dorn, G. W., 2nd, Klevitsky, R., Seidman, C. E., Seidman, J. G., & Robbins, J. (2005). Cardiac myosin binding protein-C phosphorylation and cardiac function. *Circulation Research, 97*, 1156–1163. doi:10.1161/01.RES.0000190605.79013.4d.
40. Tong, C. W., Stelzer, J. E., Greaser, M. L., Powers, P. A., & Moss, R. L. (2008). Acceleration of crossbridge kinetics by protein kinase A phosphorylation of cardiac myosin binding protein C modulates cardiac function. *Circulation Research, 103*, 974–982. doi:10.1161/CIRCRESAHA.108.177683.
41. Sadayappan, S., Gulick, J., Osinska, H., Barefield, D., Cuello, F., Avkiran, M., Lasko, V. M., Lorenz, J. N., Maillet, M., Martin, J. L., Brown, J. H., Bers, D. M., Molkentin, J. D., James, J., & Robbins, J. (2011). A critical function for Ser-282 in cardiac Myosin binding protein-C phosphorylation and cardiac function. *Circulation Research, 109*, 141–150. doi:10.1161/CIRCRESAHA.111.242560.
42. Moos, C., Mason, C. M., Besterman, J. M., Feng, I. N., & Dubin, J. H. (1978). The binding of skeletal muscle C-protein to F-actin, and its relation to the interaction of actin with myosin subfragment-1. *Journal of Molecular Biology, 124*, 571–586.
43. Squire, J. M., Luther, P. K., & Knupp, C. (2003). Structural evidence for the interaction of C-protein (MyBP-C) with actin and sequence identification of a possible actin-binding domain. *Journal of Molecular Biology, 331*, 713–724.
44. Razumova, M. V., Shaffer, J. F., Tu, A. Y., Flint, G. V., Regnier, M., & Harris, S. P. (2006). Effects of the N-terminal domains of myosin binding protein-C in an in vitro motility assay: evidence for long-lived cross-bridges. *Journal of Biological Chemistry, 281*, 35846–35854. doi:10.1074/jbc.M606949200.

45. Bhuiyan, M. S., Gulick, J., Osinska, H., Gupta, M., & Robbins, J. (2012). Determination of the critical residues responsible for cardiac myosin binding protein C's interactions. *Journal of Molecular and Cellular Cardiology, 53*, 838–847. doi:10.1016/j.yjmcc.2012.08.028.

46. Craig, R., Lee, K. H., Mun, J. Y., Torre, I., & Luther, P. K. (2014). Structure, sarcomeric organization, and thin filament binding of cardiac myosin-binding protein-C. *Pflügers Archiv, 466*, 425–431. doi:10.1007/s00424-013-1426-6.

47. Luther, P. K., Winkler, H., Taylor, K., Zoghbi, M. E., Craig, R., Padrón, R., Squire, J. M., & Liu, J. (2011). Direct visualization of myosin-binding protein C bridging myosin and actin filaments in intact muscle. *Proceedings of the National Academy of Sciences of the United States of America, 108*, 11423–11428. doi:10.1073/pnas.1103216108.