

Molecular Modeling of Double cMyBP-C Mutation Resulting in End-Stage Hypertrophic Cardiomyopathy

Poornima Gajendrarao^{1,3†}, Navaneethakrishnan Krishnamoorthy^{1,3†}, Francesca Girolami², Franco Cecchi², Iacopo Olivotto² and Magdi Yacoub^{1,3}

¹Qatar Cardiovascular Research Centre, Qatar Foundation, Doha, Qatar, ²Referral Centre for Myocardial Diseases, Careggi University Hospital, Florence, Italy, ³Heart Science Centre, National Heart and Lung Institute, Imperial College London, United Kingdom.

Name of the corresponding author: Magdi Yacoub

Email: m.yacoub@imperial.ac.uk

Address: Imperial College London, National Heart and Lung Institute, Heart Science Center, Harefield, UB9 6JH, UK

Telephone: +44 (0)1895 828 893, Fax: +44 (0) 1895 828900

† These authors contributed equally to this work

Summary

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 phenotype. The 3D structure for the C1-motif-C2 complex was constructed, with both double and single mutations being introduced and performed molecular dynamics simulations 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 resulted in an additive effect when they coexist. These changes involve important regions of the motif such as phosphorylation sites and potential actin binding region. Moreover, the charge reversal mutations altered the surface electrostatic properties. Double mutation affecting the regulatory N-terminal of cMyBP-C has the potential of synergistically interfering with the binding to neighbouring domains and to other sarcomeric proteins. These effects may account for the severe phenotype and clinical course observed in our patient.

Key Words: cardiac Myosin Binding Protein-C; Double Mutation; Hypertrophic Cardiomyopathy; Molecular Dynamics Simulation; Structure-function Relationship.

Introduction

Hypertrophic cardiomyopathy (HCM) is an inherited heart disease affecting all age groups, 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, 3, 4]. Clinical presentation and progression is extremely variable with some patients rapidly progressing to severe left ventricular dysfunction and heart failure, occasionally requiring heart transplant [5, 6]. The causes of disease progression are largely unresolved, with several factors thought to be instrumental [7]. Complex genotypes including double mutations have been shown to be 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, the molecular mechanisms responsible for adverse prognosis in HCM patients with double mutations have not been addressed. Molecular modeling is a powerful tool for studying the relationship between genotype and phenotype [9, 10, 11, 12]. Therefore, we here describe modeling of a double mutation in cMyBP-C that is encoded by the gene *MYBPC3*, which most often involved in HCM worldwide.

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 four 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 [13], and occurs in 13% of all HCM patients in Tuscany [14]. This variant is usually associated with a severe phenotype and a poor prognosis and it is highly penetrant [1]. Recently Mearini provided for this mutation the first evidence of successful 5' trans splicing *in vivo* [15] .

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. This mutation was first described in 2005 by Seidman *et al.*, (Cardiogenomics) [16] and in 2009 by Marsiglia *et al.*, (JDC 2009) [17] 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 structure of cMyBP-C contains 11 globular domains including eight immunoglobulin (Ig)-like domains and three fibronectin (Fn)-like domains, termed C0–C10 [18]. The double mutation in our patient was located in domains C1 (E258K) and C2 (E441K), for which 3D structures are available. The domains, C1 and C2 are connected through 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 [19]. 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 [20]. This modeling study on C1-m-C2 provides a plausible molecular explanation of the double mutation in the patient with severe phenotype.

Methods

Cell culture and transfection

The rat embryonic, heart derived cardiomyoblast cell line, H9c2 were obtained from American Type Culture Collection (ATCC, Manassas, VA). Cells were cultured in Dulbecco modified Eagle medium (DMEM, GIBCO, Carlsbad, CA, USA) supplemented with heat-inactivated 10% fetal bovine serum (FBS; Gibco, MD) and antibiotics (100 units/ml penicillin, 100 µg/ml streptomycin) in a water-saturated atmosphere of 5% CO₂ at 37 °C. For transient transfection of GFP tagged WT and mutant MYBPC-3 (E258K, E441K & E258K-E441K; GeneArt Gene Synthesis, Life Technologies, CA), cells were transfected in OptiMEM with Lipofectamine reagent 2000 (Invitrogen, Carlsbad, CA) 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), and resolved on NuPAGE 4-12% Bis-Tris gel (Life Technologies, CA), 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). Anti-β-actin (1:5000; Sigma-Aldrich) 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 \pm SEM.

Building Structure for the Motif

The 3D structural data of the domains C1 (pdb ID: 2V6H) [21] and C2 (pdb ID: 1PD6) [22] 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 I-TASSER (iterative threading assembly refinement) [23] 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, and 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 [24]. 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) [25] v3.5 was

used to introduce mutations to the complexes as performed by Krishnamoorthy *et al.*, 2011 [26].

MD simulations of the Complexes

The GROMACS simulation package (v4.5.4) [27,28] was used to perform the MD simulations, with explicit water and physiological conditions. The systems were solvated using the SPC3 [29] 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 [30]. Van der Waals and electrostatic interactions were observed with a twin range cut-off of 0.8 nm (long-range) and 1.4 nm, respectively. The LINCS algorithm [31] was employed to constrain the lengths of all the bonds, and the SETTLE [32] to constrain 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 [33]. 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 [34] 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 [35, 36]. 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 [37]. To determine the effect of E258K, E441K and double mutation (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). 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 mutations 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 C-subdomain of the motif was resolved recently from mouse, with 3 helices. The remaining 2 helices at the N-subdomain have not been resolved [19]. In Fig. 2a, our model shows a plausible complete structure of a human cMyBP-C motif, composed of 5 helical structures as suggested previously, and a small helix that connects the subdomains N and C. The model also shows suitable positioning of the LAGGGRRIS 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 [38]. Furthermore, the motif provides four accessible phosphorylation sites that are known;

Ser275, Ser284, Ser304 and Ser311 [38], 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. 2*b*). 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.

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 C α atoms with respect to the initial structures throughout the MD simulations (Fig. 3*a*). 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 systems (Fig. 3*b*). These results show 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 are observed in E258K and E441K (Fig. 4b and c). However, the motif in the double mutation displayed minimal conformational shift compared to its shift in the single mutant systems. Systems with single mutation, E258K and E441K, showed that they could largely affect their native domain, C1 and C2 respectively (Fig. 4b and 4c). These results indicate that the single mutations studied here predominantly affect their native domains as well as the nearby motif. When they co-exist, 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 LAGGRRIS (reported as a regulator of phosphorylation) (Fig. 5) and the potential actin binding 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. 5*b* and *d*), whose phosphorylation is reported to be a prerequisite for phosphorylation of rest of the sites. Surprisingly, in E258K, minimal effects were observed at these sites (Fig. 5*c*). 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). Again, both the double mutation and E441K in isolation induce maximal shift. Altogether, the structural shifts on the key regions of the motif indicate that the both double mutation and E441K induce severe changes on these sites compare to E258K. These changes might affect the binding of the C1-m-C2 to myosin S2 and actin also could 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 shown 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 mutations increased the number of hydrogen bonds from 7 to 14, whilst there was a reduction of

hydrophobic interactions from 22 to 12. In the single mutation systems, hydrogen bonds are also 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 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.

Discussion

This modeling study on the complex C1-m-C2 suggests possible structural consequences of the severe HCM phenotype associated with double mutation in the gene coding for cMyBP-C. In MD simulations, the mutations have been shown to produce distinctive and considerable changes on the sites of phosphorylation, 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.

Several studies have shown that truncation mutations in *MYBPC3* diminish the cMyBP-C protein level and cause hypertrophic cardiomyopathy through haploinsufficiency [39, 40, 41]. However, *MYBPC3* missense mutations, including E258K, cause a more malignant form of HCM through disruption of sarcomeric function than *MYBPC3* haploinsufficiency [37, 42, 13]. Marston et al., showed that E258K mutation marginally reduced the protein level in myocardium [39]. Our data showed that both single and double mutations unlikely affected the protein level and may supports the concept that haploinsufficiency may not be the primary mechanism of disease progression in HCM.

Multiple phosphorylation sites on cMyBP-C suggest its role in normal cardiac function. Mutations in these sites reduce contractile function [38, 43, 44, 45]. In addition, it has been suggested that phosphorylation of Ser284 in humans (Ser282 in mouse) is a prerequisite for phosphorylation of Ser275, Ser 304 and Ser311 [38, 45]. This study provides a structural insight into the mutational effect on the phosphorylation sites. The drastic conformational shift on the sites, and in particular the burial of Ser284, and the deviation of the cardiac specific loop due to double mutations (as well as E441K in isolation), might affect normal cardiac function. The mutation E258K is located near to Ser304 however it was not affecting any site of phosphorylation severely. This suggests that the E258K mutation might affect the

binding of C1 to S2 of myosin directly by changes on surface electrostatic properties [13, 21, 22]. This is in agreement with the observations of our previous study on E258K [10]. However, the effects of mutations on the sites of phosphorylation have to be validated.

Ca²⁺ signalling-based energy depletion is one of the possible mechanisms that causes HCM [7, 46], which found to be regulated by the normal phosphorylation in the motif [47]. The mutations induced changes revealed here on the sites of phosphorylation appear that it could influence Ca²⁺ signalling and energy supply for contractile function.

In vitro studies such as Immunofluorescence and yeast-two-hybrid, suggest that C1-m-C2 interacts with S2 of myosin to regulate phosphorylation and cross-bridge kinetics [21, 22, 48, 49]. Here, the interaction pattern at the interfaces revealed that the native C1-m and m-C2 is predominately built with hydrophobic interactions and a few hydrogen bonds. This native pattern was altered by the mutations studied, which could impact the normal functionality of the complex.

The biochemical or biophysical analyses and yeast-two-hybrid experiments [50, 51, 52] reported that the domain C1, motif and/or C1-m-C2 complex interacts with actin mainly via lysine residues [53]. 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 could be based on electrostatic properties. Thus, the charge reversal mutations (single or double) studied here can induce modifications to the surface electrostatic properties via conformational changes on the binding site, and consequently interfere its interaction with actin [54, 55].

Conclusions

In the contractile apparatus, cMyBP-C has to bind with both thick and thin filament systems that are 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 conservation of conformational specificity, interface and electrostatic properties to the native structural organization of C1-m-C2 for 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 lead to severe phenotype observed in our patient. It is hoped that these observations will help to identify future therapeutic targets for HCM.

Acknowledgements

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.

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, United Kingdom. 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.

References

- [1] Richard P, Charron P, Carrier L, Ledeuil C, Cheav T, Pichereau C, Benaiche A, Isnard R, Dubourg O, Burban M, Gueffet JP, 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 BJ, Gardin JM, Flack JM, Gidding SS, Kurosaki TT, Bild DE. 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 BJ, Shirani J, Poliac LC, Mathenge R, Roberts WC, Mueller FO. 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 MH, Olivotto I. 2005 Hypertrophic cardiomyopathy in the community: why we should care. *Nat. Clin. Pract. Cardiovasc. Med.* **2**, 324-325. (doi:10.1038/ncpcardio0248)

- [5] Olivotto I, Cecchi F, Poggesi C, Yacoub MH. 2012 Patterns of disease progression in hypertrophic cardiomyopathy: an individualized approach to clinical staging. *Circ. Heart. Fail.* **5**, 535-546. (doi: 10.1161/CIRCHEARTFAILURE.112.967026)

- [6] Yacoub MH, Olivotto I, Cecchi F. 2007 'End-stage' hypertrophic cardiomyopathy: from mystery to model. *Nat. Clin. Pract. Cardiovasc. Med.* **4**, 232-233. (doi:10.1038/ncpcardio0859)

- [7] Frey N, Luedde M, Katus HA. 2011 Mechanisms of disease: hypertrophic cardiomyopathy. *Nat. Rev. Cardiol.* **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? *Circ. Cardiovasc. Genet.* **2**, 182-190. (doi: 10.1161/CIRCGENETICS.108.836478)

- [9] Kang YS, Park SY, Yim CH, Kwak HS, Gajendrarao P, Krishnamoorthy N, Yun SC, Lee KW, Han KO. 2009 The CYP3A4*18 genotype in the cytochrome P450 3A4 gene, a rapid metabolizer of sex steroids, is associated with low bone mineral density. *Clin. Pharmacol. Ther.* **85**, 312-318. (doi: 10.1038/clpt.2008.215)
- [10] Gajendrarao P, Krishnamoorthy N, KassemHSh, Moharem-Elgamal S, Cecchi F, Olivotto I, Yacoub MH. 2013 Molecular modeling of disease causing mutations in domain C1 of cMyBP-C. *PLoS One* **8**, e59206. (doi: 10.1371/journal.pone.0059206)
- [11] Lower SK, Lamlertthong S, Casillas-Ituarte NN, Lins RD, Yongsunthorn R, Taylor ES, DiBartola AC, Edmonson C, McIntyre LM, Reller LB, Que YA, Ros R, Lower BH, Fowler VG Jr. 2011 Polymorphisms in fibronectin binding protein A of *Staphylococcus aureus* are associated with infection of cardiovascular devices. *Proc. Natl. Acad. Sci. USA*. **108**, 18372-18377. (doi: 10.1073/pnas.1109071108)
- [12] Kumar A, Purohit R. 2014 Use of long term molecular dynamics simulation in predicting cancer associated SNPs. *PLoS Comput. Biol.* **10**, e1003318. (doi: 10.1371/journal.pcbi.1003318)
- [13] De Lange WJ, Grimes AC, Hegge LF, Spring AM, Brost TM, Ralphe JC. 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. *J. Gen. Physiol.* **142**, 241-255. (doi: 10.1085/jgp.201311018)
- [14] Olivotto I, Girolami F, Ackerman MJ, Nistri S, Bos JM, Zachara E, Ommen SR, Theis JL, Vaubel RA, Re F, Armentano C, Poggesi C, Torricelli F, Cecchi F. 2008 Myofilament protein gene mutation screening and outcome of patients with hypertrophic cardiomyopathy. *Mayo. Clin. Proc.* **83**, 630-638. (doi: 10.4065/83.6.630)
- [15] Mearini G, Stimpel D, Krämer E, Geertz B, Braren I, Gedicke-Hornung C, Précigout G, Müller OJ, Katus HA, 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. *Mol. Ther. Nucleic Acids*. **2**, e102. (doi: 10.1038/mtna.2013.31)
- [16] Merk, Seidman. 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].

[17] Marsiglia JD, Batitucci Mdo C, Paula Fd, Barbirato C, Arteaga E, Araújo AQ. 2010 [Study of mutations causing hypertrophic cardiomyopathy in a group of patients from Espirito Santo, Brazil]. *Arq. Bras. Cardiol.* **94**, 10-17.

[18] S Einheber and D A Fischman. 1990 Isolation and characterization of a cDNA clone encoding avian skeletal muscle C-protein: an intracellular member of the immunoglobulin superfamily. *Proc. Natl. Acad. Sci. USA.* **87**, 2157–2161. (DOI:10.1073/pnas.87.6.2157)

[19] Howarth, J.W, Ramiseti, S, Nolan, K, Sadayappan, S, Rosevear, P.R. 2012 Structural insight into unique cardiac myosin-binding protein-C motif: a partially folded domain. *J. Biol. Chem.* **287**, 8254-8262. (doi: 10.1074/jbc.M111.309591)

[20] Oakley CE, Chamoun J, Brown LJ, Hambly BD. 2007 Myosin binding protein-C: enigmatic regulator of cardiac contraction. *Int. J. Biochem. Cell Biol.* **39**, 2161-2166. (DOI: 10.1016/j.biocel.2006.12.008)

[21] Govada L, Carpenter L, da Fonseca PC, Helliwell JR, Rizkallah P, Flashman E, Chayen NE, Redwood C, Squire JM. 2008 Crystal structure of the C1 domain of cardiac myosin binding protein-C: implications for hypertrophic cardiomyopathy. *J. Mol. Biol.* **378**, 387-397. (doi: 10.1016/j.jmb.2008.02.044)

[22] 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. *J. Biol. Chem.* **282**, 9204–9215. (doi: 10.1074/jbc.M610899200)

[23] Roy A, Kucukural A, Zhang Y. 2010 I-TASSER: a unified platform for automated protein structure and function prediction. *Nat. Protoc.* **5**, 725-738. (doi: 10.1038/nprot.2010.5)

[24] Comeau SR, Gatchell DW, Vajda S, Camacho CJ. 2004 ClusPro: an automated docking and discrimination method for the prediction of protein complexes. *Bioinformatics* **20**, 45-50. (doi: 10.1093/bioinformatics/btg371)

[25] Discovery Studio, 2.0 User Guide, 2005, Accelrys Inc., San Diego, CA, USA.

- [26] Krishnamoorthy N, Yacoub MH, Yaliraki SN. 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)
- [27] Van Der Spoel D, Lindahl E, Hess B, Groenhof G, Mark AE, Berendsen HJ. 2005 GROMACS: fast, flexible, and free. *J. Comput. Chem.* **26**, 1701-1718. (DOI: 10.1002/jcc.20291)
- [28] Hess B, Kutzner C, van der Spoel D, Lindahl E. 2008 GROMACS 4: algorithms for highly efficient, load-balanced, and scalable molecular simulation. *J. Chem. Theory. Comput.* **4**, 435–447. (DOI: 10.1021/ct700301q)
- [29] Berendsen HJC, Postma JPM, van Gunsteren WF, 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.
- [30] van Gunsteren WF, Billeter SR, Eising AA, Hünenberger PH, Krüger P, Mark AE, Scott WRP, Tironi IG. 1996 *Biomolecular Simulation: The GROMOS96 Manual and User Guide*, Vdf Hochschulverlag AG an der ETH Zürich, Zürich, Switzerland, pp. 1-1042.
- [31] Hess B, Bekker H, Berendsen HJC, Fraaije JGEM. 1997 LINCS: a linear constraint solver for molecular simulations. *J. Comput. Chem.* **18**, 1463–1472. (DOI: 10.1002/(SICI)1096-987X(199709)18)
- [32] Miyamoto S, Kollman PA. 1992 SETTLE: an analytical version of the SHAKE and RATTLE algorithm for rigid water models. *J. Comput. Chem.* **13**, 952–962. (DOI: 10.1002/jcc.540130805)
- [33] Berendsen HJC, Postma JPM, van Gunsteren WF, DiNola A, Haak JR. 1984 Molecular dynamics with coupling to an external bath. *J. Chem. Phys.* **81**, 3684–3691. (doi.org/10.1063/1.448118)
- [34] Wallace AC, Laskowski RA, Thornton JM. 1995 LIGPLOT: a program to generate schematic diagrams of protein-ligand interactions. *Protein Eng.* **8**, 127-134. (doi: 10.1093/protein/8.2.127)

- [35] Sinha N, Smith-Gill SJ. 2002 Electrostatics in protein binding and function. *Curr. Protein Pept. Sci.* **3**, 601-614. (DOI: 10.2174/1389203023380431)
- [36] Krishnamoorthy N, Gajendrarao P, Eom SH, Kwon YJ, Cheong GW, Lee KW. 2008 Molecular modeling study of CodX reveals importance of N-terminal and C-terminal domain in the CodWX complex structure of *Bacillus subtilis*. *J. Mol. Graph. Model.* **27**, 1-12. (doi: 10.1016/j.jmgm.2008.01.009)
- [37] Helms AS, Davis FM, Coleman D, Bartolone SN, Glazier AA, Pagani F, Yob JM, Sadayappan S, Pedersen E, Lyons R, Westfall MV, Jones R, Russell MW, Day SM. Sarcomere mutation-specific expression patterns in human hypertrophic cardiomyopathy. *Circ Cardiovasc Genet.* 2014 Aug;7(4):434-43
- [38] 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 J.* **14**, 1952–1960.
- [39] Marston, S., O. Copeland, A. Jacques, K. Livesey, V. Tsang, W.J. McKenna, S. Jalilzadeh, S. Carballo, C. Redwood, and H. Watkins. 2009. Evidence from human myectomy samples that MYBPC3 mutations cause hypertrophic cardiomyopathy through haploinsufficiency. *Circ. Res.* 105:219–222.
- [40] Marston, S., O. Copeland, K. Gehmlich, S. Schlossarek, and L. Carrier. 2012. How do MYBPC3 mutations cause hypertrophic cardiomyopathy? *J. Muscle Res. Cell Motil.* 33:75–80.
- [41] van Dijk, S.J., D. Dooijes, C. dos Remedios, M. Michels, J.M. Lamers, S. Winegrad, S. Schlossarek, L. Carrier, F.J. ten Cate, G.J. Stienen, and J. van der Velden. 2009. Cardiac myosin-binding protein C mutations and hypertrophic cardiomyopathy: haploinsufficiency, deranged phosphorylation, and cardiomyocyte dysfunction. *Circulation.* 119:1473–1483.
- [42] Niimura, H., L.L. Bachinski, S. Sangwatanaroj, H. Watkins, A.E. Chudley, W. McKenna, A. Kristinsson, R. Roberts, M. Sole, B.J. Maron, et al. 1998. Mutations in the gene for cardiac myosin-binding protein C and late-onset familial hypertrophic cardiomyopathy. *N. Engl. J. Med.* 338:1248–1257.
- [43] Sadayappan S, Gulick J, Osinska H, Martin LA, Hahn HS, Dorn GW 2nd, Klevitsky R, Seidman CE, Seidman JG, Robbins J. 2005 Cardiac myosin binding protein-C

phosphorylation and cardiac function. *Circ. Res.* **97**, 1156–1163. (doi: 10.1161/01.RES.0000190605.79013.4d)

[44] Tong CW, Stelzer JE, Greaser ML, Powers PA, Moss RL. 2008 Acceleration of crossbridge kinetics by protein kinase A phosphorylation of cardiac myosin binding protein C modulates cardiac function. *Circ. Res.* **103**, 974–982. (doi: 10.1161/CIRCRESAHA.108.177683)

[45] Sadayappan S, Gulick J, Osinska H, Barefield D, Cuello F, Avkiran M, Lasko VM, Lorenz JN, Maillet M, Martin JL, Brown JH, Bers DM, Molkentin JD, James J, Robbins J. 2011 A critical function for Ser-282 in cardiac Myosin binding protein-C phosphorylation and cardiac function. *Circ. Res.* **109**, 141–150. (doi: 10.1161/CIRCRESAHA.111.242560)

[46] Ashrafian H, Redwood C, Blair E, Watkins H. 2003 Hypertrophic cardiomyopathy: a paradigm for myocardial energy depletion. *Trends Genet.* **19**, 263–638. (DOI: 10.1016/S0168-9525(03)00081-7)

[47] Lu Y, Kwan AH, Jeffries CM, Guss JM, Trewhella J. 2012 The motif of human cardiac myosin-binding protein C is required for its Ca²⁺-dependent interaction with calmodulin. *J. Biol. Chem.* **287**, 31596–31607. (doi: 10.1074/jbc.M112.383299)

[48] Gruen M, Gautel M. 1999 Mutations in beta-myosin S2 that cause familial hypertrophic cardiomyopathy (FHC) abolish the interaction with the regulatory domain of myosin-binding protein-C. *J. Mol. Biol.* **286**, 933–949. (DOI: 10.1006/jmbi.1998.2522)

[49] Sadayappan S, Osinska H, Klevitsky R, Lorenz JN, Sargent M, Molkentin JD, Seidman CE, Seidman JG, Robbins J. 2006 Cardiac myosin binding protein C phosphorylation is cardioprotective. *Proc. Natl. Acad. Sci. USA.* **103**, 16918–16923. (doi: 10.1073/pnas.0607069103)

[50] Moos C, Mason CM, Besterman JM, Feng IN, Dubin JH. 1978 The binding of skeletal muscle C-protein to F-actin, and its relation to the interaction of actin with myosin subfragment-1. *J. Mol. Biol.* **124**, 571–586. (DOI: 10.1016/0022-2836(78)90172-9)

[51] Squire JM, Luther PK, Knupp C. 2003 Structural evidence for the interaction of C-protein (MyBP-C) with actin and sequence identification of a possible actin-binding domain. *J. Mol. Biol.* **331**, 713–724. (DOI: 10.1016/S0022-2836(03)00781-2)

- [52] Razumova MV, Shaffer JF, Tu AY, Flint GV, Regnier M, Harris SP. 2006 Effects of the N-terminal domains of myosin binding protein-C in an in vitro motility assay: evidence for long-lived cross-bridges. *J. Biol. Chem.* **281**, 35846-35854. (doi: 10.1074/jbc.M606949200)
- [53] Bhuiyan MS, Gulick J, Osinska H, Gupta M, Robbins J. 2012 Determination of the critical residues responsible for cardiac myosin binding protein C's interactions. *J. Mol. Cell. Cardiol.* **53**, 838-847. (doi: 10.1016/j.yjmcc.2012.08.028)
- [54] Craig R, Lee KH, Mun JY, Torre I, Luther PK. 2014 Structure, sarcomeric organization, and thin filament binding of cardiac myosin-binding protein-C. *Pflugers. Arch.* **466**, 425-31. (doi: 10.1007/s00424-013-1426-6)
- [55] Luther PK, Winkler H, Taylor K, Zoghbi ME, Craig R, Padrón R, Squire JM, Liu J. 2011 Direct visualization of myosin-binding protein C bridging myosin and actin filaments in intact muscle. *Proc. Natl. Acad. Sci. USA.* **108**, 11423-8. (doi: 10.1073/pnas.1103216108)