STRUCTURAL ANALISYS OF MUTATIONS ASSOSIATED WITH IDIOPATIC RESTRICTIVE CARDIOMYOPATHY IN CYTOSKELETAL AND SARCOMERIC PROTEINS

Svetlana Tarnovskaya[1], Artem Kiselev[2,3], Anna Kostareva[2], Dmitrij Frishman[1,4,5]

[1]*Department of Bioinformatics, St. Petersburg Polytechnic University, St. Petersburg, Russian Federation;* svetlanatarnovskaya@gmail.com

[2]*Almazov Federal Medical Research Centre, St. Petersburg, Russian Federation;*

[3]*ITMO University, Institute of translational Medicine, St. Petersburg, Russia;*

[4]*Department of Bioinformatics, Technische Universität München, Wissenschaftszentrum Weihenstephan, Freising, Germany;*

[5]*Institute for Bioinformatics and Systems Biology, HMGU German Research Center for Environmental Health, Neuherberg, Germany.*

Idiopathic restrictive cardiomyopathy (RCMP) is the least common type of cardiomyopathies often of genetic origin. Recently 77 non-synonymous single nucleotide polymorphisms (nsSNPs) in cardiomyopathy-associated genes were classified as potential disease-modifying variants leading to RCMP development (Kostareva et al., submitted). The aim of the present investigation was to analyze the structural these amino acid substitutions and to predict their effect on protein function.

Molecular modeling was conducted using PyMol software. 3D structures of human cytoskeletal and sarcomeric proteins were extracted from the PDB database (Berman, 2000). nsSNPs were classified as damaging or neutral by SNPs&GO, a machine learning method based on multiple amino acid features, evolutionary conservation, and functional annotation (Calabrese, Capriotti, Fariselli, Martelli, & Casadio, 2009). I-Mutant 2.0, CUPSAT and mCSM methods were used to predict the possible impact of amino acid substitutions on the stability and function of human proteins using structural evidence. Furthermore, domain composition of proteins was characterized using Pfam (Finn et al., 2014). Related protein sequence from other organisms were identified and multiply aligned by BLAST (Altschul, Gish, Miller, Myers, & Lipman, 1990) using E-value threshold 0.001. Sequence redundancy in protein families was reduced to 80% using CD-hit (Fu, Niu, Zhu, Wu, & Li, 2012).

We obtained 16 disease-assosiated nsSNPs, with 8 of them located in proteins with known three-dimensional structure (see Table 1). Three radical amino acids substitutions (R->W, G->E and L->P) are located in highly conserved regions and are predicted to destabilize protein structure and probably affect function (Table). Moreover, all structurally known

amino acids mutations lead to unfavorable torsion angles.

As an example we present here our analysis of the variant found in myomesin-1 (Myom1), a major component of the vertebrate myofibrillar M-band. In cardiac muscles, elastic M-band motions correlate with the heart beat rate. M-bridges in the M-bands connect thick filaments with each other and with titin (Ttn) and myosin (Myh7) filaments. The C-terminal part of the myomesin filament consists of an array of repetitive Ig domains followed by exposed α-helical linkers. The complete structure and extent of longitudinal elasticity of the entire C-terminal tail-to-tail myomesin filament My9–My10–My11–My12–(My13)2–My12'–My11'–My10'–My9' (My9–My13) is described in (Pinotsis et al., 2012) (PDB: 2Y25). They form structurally highly conserved Ig domain/helix interfaces (Lange et al., 2005; Musa et al., 2006; Schoenauer et al., 2008).

Table 1. Prediction of protein stability changes upon single point mutation

Protein (Uniprot id)	Mutation	PDB code	Secondary structure	I-Mutant 2.0	CUPSAT	mCSM
Ttn (Q8WZ42)	R2083I	1G1C	Loop	Destabilizing	Stabilizing	Stabilizing
Tnni3 (P19429)	R170W	1J1E	α-helix	Destabilizing	Destabilizing	Destabilizing
Myom1 (P52179)	G1424E	3RBS	β-sheet	Destabilizing	Destabilizing	Destabilizing
Ilk (Q13418)	R211C	3KMW	β-sheet	Destabilizing	Destabilizing	Destabilizing
Myh7 (P12883)	G768R	4P7H	α-helix	Stabilizing	Stabilizing	Destabilizing
Cacnb2 (Q08289)	D131G	4DEX	Loop	Stabilizing	Destabilizing	Destabilizing
Actn2 (P35609)	N175Y	4D1E	Loop	Stabilizing	Stabilizing	Destabilizing
Scn4b (Q8NF91)	L51P	4MZ2	β-sheet	Destabilizing	Destabilizing	Destabilizing

Ttn – titin isoform IC [Homo sapiens]; Tnni3 - troponin I, cardiac muscle [Homo sapiens]; Myom1 – myomesin-1 isoform a [Homo sapiens]; Ilk - integrin-linked protein kinase isoform 1 [Homo sapiens]; Myh7 - myosin-7 [Homo sapiens]; Cacnb2 - voltage-dependent L-type calcium channel subunit beta-2 isoform 4 [Homo sapiens]; Actn2 - alpha-actinin-2 isoform 2 [Homo sapiens]; Scn4b - sodium channel subunit beta-4 isoform 1 precursor [Homo sapiens].

According to all prediction methods the mutation in position 1424 affects has a strong destabilizing effect on the protein (see Table 1). A Gly to Glu substitution is located in a β-strand, which interacts with an α-helix and forms the Ig domain-helix interface. Main-chain atoms of both Glu in the mutant and Gly in the wild type form a hydrogen bond with the main chain of Leu 1443 located in the C-terminal part of the My domain, which directly

interacts with the α -helix. We therefore project that destabilization in this region can lead to changes in the My11 domain-helix interface and have a negative impact on myomesin stretching (Fig. 1).

A B

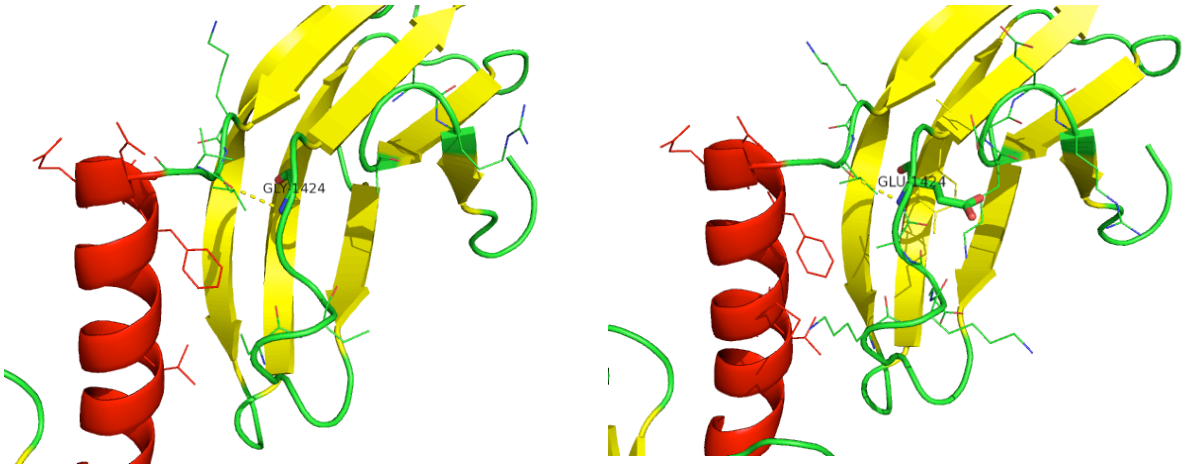

Fig. 1. Crystal structure of the IgH segment of the My11 domain (PDB: 2y25, Chain A). A: wild protein; B: mutated protein.

The main function of myomesin is to balance mechanical forces between molecules during the sarcomeric force-generating cycles. Disease-causing mutations are more likely to occur at positions that are conserved throughout evolution (Ng & Henikoff, 2006). Most probably they reduce stability, affect protein function, and lead to elasticity changes in the myofibril.

Detailed structure-based analysis of the impact of mutations on protein stability can shed light on their causal role in the idiopathic restrictive cardiomyopathy.

References

1. Altschul, S. F., Gish, W., Miller, W., Myers, E. W., & Lipman, D. J. (1990). Basic local alignment search tool. *Journal of Molecular Biology*, *215*(3), 403–10. doi:10.1016/S0022-2836(05)80360-2
2. Berman, H. M. (2000). The Protein Data Bank. *Nucleic Acids Research*, *28*(1), 235–242. doi:10.1093/nar/28.1.235
3. Calabrese, R., Capriotti, E., Fariselli, P., Martelli, P. L., & Casadio, R. (2009). Functional annotations improve the predictive score of human disease-related mutations in proteins. *Human Mutation*, *30*(8), 1237–44. doi:10.1002/humu.21047
4. Finn, R. D., Bateman, A., Clements, J., Coggill, P., Eberhardt, R. Y., Eddy, S. R., Punta, M. (2014). Pfam: the protein families database. *Nucleic Acids Research*, *42*(Database issue), D222–30. doi:10.1093/nar/gkt1223

5. Fu, L., Niu, B., Zhu, Z., Wu, S., & Li, W. (2012). CD-HIT: accelerated for clustering the next-generation sequencing data. *Bioinformatics (Oxford, England)*, *28*(23), 3150–2. doi:10.1093/bioinformatics/bts565
6. Lange, S., Himmel, M., Auerbach, D., Agarkova, I., Hayess, K., Fürst, D. O., Ehler, E. (2005). Dimerisation of myomesin: implications for the structure of the sarcomeric M-band. *Journal of Molecular Biology*, *345*(2), 289–98. doi:10.1016/j.jmb.2004.10.040
7. Musa, H., Meek, S., Gautel, M., Peddie, D., Smith, A. J. H., & Peckham, M. (2006). Targeted homozygous deletion of M-band titin in cardiomyocytes prevents sarcomere formation. *Journal of Cell Science*, *119*(Pt 20), 4322–31. doi:10.1242/jcs.03198
8. Ng, P. C., & Henikoff, S. (2006). Predicting the effects of amino acid substitutions on protein function. *Annual Review of Genomics and Human Genetics*, *7*, 61–80. doi:10.1146/annurev.genom.7.080505.115630
9. Pinotsis, N., Chatziefthimiou, S. D., Berkemeier, F., Beuron, F., Mavridis, I. M., Konarev, P. V., Wilmanns, M. (2012). Superhelical architecture of the myosin filament-linking protein myomesin with unusual elastic properties. *PLoS Biology*, *10*(2), e1001261. doi:10.1371/journal.pbio.1001261
10. Schoenauer, R., Lange, S., Hirschy, A., Ehler, E., Perriard, J.-C., & Agarkova, I. (2008). Myomesin 3, a novel structural component of the M-band in striated muscle. *Journal of Molecular Biology*, *376*(2), 338–51. doi:10.1016/j.jmb.2007.11.048