Identification of Calpain 10 Isoforms (*b, d, e, f & h*) Conserved Regions and Possible Functional Prophecy through Bioinformatics

Kausar Saboohi[1,2], Samina Bano[2] and Bushra Chaudhry[1,*]

[1]*Department of Biological & Biomedical Sciences, The Aga Khan University Medical College, P.O Box 3500, Stadium Road, Karachi, Pakistan*

[2]*Department of Biochemistry, University of Karachi, University Road, Karachi, 75270, Pakistan*

Abstract: Calpain 10 is an atypical calpain ubiquitously exist in all human tissues. It exhibits eight protein isoforms designated as "*a-h*" which play a vital role in glucose homeostasis but actual mechanism of action is yet to be ascertained. We have predicted the partial roles of Isoform *a, c* and *g* previously. They were envisaged to act partially as mu and m-calpain cysteine proteases. Here we predict the function of minor isoforms *b, d, e, f* and *h.* We have applied NCBI Blast and Conserved domain tool for nucleotide and protein alignments. Blast query indicated 87%, 84%, 87%, 94% and 34% identity of isoform *b, d, e, f* and *h* with canonical sequence of calpain 10 *a* isoform. Conserved domain analyses of protein sequences revealed significant structural similarities of their N-terminal domain I and II with catalytic domain of cysteine protease superfamily PC1 (e-value:CAPN10*b, d, e* = 2.41e-76, CAPN10*f* = 1.07e-43 and CAPN10*h* = 1.13e-17). Isoform *b, d* and *e* have one consecutive domain similar with C2 like subdomain III (e-value=2.92-32, 1.03e-35, 1.88e-14 respectively) and was classified in CAPN10 group of Palb subfamily. Isoform *f* and *h* were lacking this domain and had shorter sequences. Although structural similarities are not guaranteed for similar actions but domain homology predicted the existence of similar functions as of calpain I and II.

Keywords: Cysteine proteases, calpain 10 isoforms, conserved domain homology, structure function prediction, action mechanism, Type 2 diabetes, insulin secretion.

INTRODUCTION

The calpain 10 is an atypical cysteine protease. It plays an active role in the development of type 2 diabetes predicted after complete genome scan analyses [1]. Calpain 10 as a member of a superfamily of proteases, is thought to be calcium activated, non-lysosomal, neutral, cysteine protease which partially proteolysis its substrates. The partial proteolysis results in the modulation of its substrate functions by either activation or inhibition of the molecules. It is ubiquitously found in all tissues of human body. It is found in cytosol as well as in mitochondria [2]. The overexpression of calpain 10 proteins results in mitochondrial dysfunction [3, 4], as they cleave the electron transport chain proteins and a decrease in fatty acid oxidation for production of ATP, which increases the intracellular fatty acyl CoA and diacylglycerol (DAG) levels in mitochondria and disrupts insulin signaling. These molecules activate protein kinase C, which is also a substrate of calpain10. Increased amount of protein kinase C in turn activates a serine kinase cascade leading to increased serine phosphorylation of insulin receptor substrate 1 (IRS-1). This serine phosphorylation at critical sites inhibits the IRS-1 tyrosine phosphorylation by insulin receptor and results in inhibition of insulin-stimulated glucose transport and causing insulin resistance in skeletal muscle and liver [5]. Splicing of calpain 10 mRNA results in the formation of eight different isoforms (a-h). We hypothesized that the isoforms which we are analyzing for individual functional prediction may behave in a similar mode as described above. These isoforms are originating from calpain 10 largest isoform. The Genetic variation in calpain 10 gene has been studied in several ethnic populations of the world and resulted in low expression of calpain 10 proteins. The mechanism by which calpain 10 involved with disease(s) is not yet exactly known. It was previously reported that calpain 10 protein expression is induced by physical training [6]. A band of 60kDa of calpain 10 protein (represents isoform *b, c* and *d*) observed after Immunoblotting in skeletal muscles of endurance trained control subjects and not in diabetic patients. These results indicated that physical activity rather than insulin resistance may influence the regulation of calpain 10 protein levels in human skeletal muscle [6]. Bioinformatics tools are actively used to foretell the functions of molecules for their *in vivo* behavior and mechanisms of action on the basis of their nucleotide and amino acid sequences, three dimensional structures and conserve domain analyses. Three major and ubiquitously found isoforms (*a, c* and *g*) were previously been analyzed by our group and were found homologous to cysteine protease superfamily but 3D-structure analysis revealed that they have their own unique 3-dimentional structures.

*Address correspondence to this author at the Department of Biological & Biomedical Sciences, The Aga Khan University, Stadium Road, Karachi, Pakistan; Tel: 34930051, Ext.4326; E-mail: bushra.chaudhry@aku.edu

ISSN: 1814-8085 / E-ISSN: 1927-5129/13

The aim of our study was to predict the mechanism of action of calpain 10 minor isoforms *b, d, e, f* and *h* and their involvement in the development of type 2 diabetes and other diseases using bioinformatics approaches mainly on the basis of their domain homology, providing some information about the function.

MATERIALS AND METHODS

Protein sequences of Calpain 10 isoforms *b, d, e* and *h* were obtained from UniProt KB (www.uniprot.org). These minor isoforms were examined for nucleotide and protein sequence similarity using multiple alignment tools on NCBI blast Align tool (http://blast.ncbi.nlm.nih.gov/Blast.cgi). The blast nucleotide algorithm (blastn) first breaks the query sequence into the short subsequences (words) and then identifies the exact matches. These hits were used to generate final alignment. In contrast, standard protein-protein blast (blastp) identified a query sequence as well as other similar sequences in protein databases. Calpain 10 isoform *a was* used as canonical sequence to determine the alignment analysis of calpain 10 *b, d, e* and *h*. Domain homology of calpain 10 minor isoforms with other Homo sapiens proteins were analyzed by Conserved Domain tool (http://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb). It used templates from SMART, COG, Protein Clusters and Pfam databases for alignments [7-9]. It was used to predict the domain homology of a query protein with other proteins on the basis of structural similarity and used to predict the function of that particular domain.

RESULTS

Results of this study showed that minor isoforms *b, d* and *e* have similar structural (Figure **1**) and domain homology with isoforms *a*, whereas isoform *f* and *h* are different. The nucleotides and amino acids sequences were showed the blast query coverage of 87%, 84%, 87%, 94% and 34% identity with canonical sequence of calpain 10 *a* for isoform *b, d, e, f* and *h* respectively

Score = 875 bits (2260), Expect = 0.0, Method: Compositional matrix adjust.
Identities = 444/487 (91%), Positives = 452/487 (93%), Gaps = 2/487 (0%)

```
Query 1    MRAGRGATPARELFRDAAFPAADSSLFCDLSTPLAQFREDITWRRPQEICATPRLFPDDP   60
           MRAGRGATPARELFRDAAFPAADSSLFCDLSTPLAQFREDITWRRPQEICATPRLFPDDP
Sbjct  1   MRAGRGATPARELFRDAAFPAADSSLFCDLSTPLAQFREDITWRRPQEICATPRLFPDDP   60

Query 61   REGQVKQGLLGDCWFLCACAALQKSRHLLDQVIPPGQPSWADQEYRGSFTCRIWQFGRWV  120
           REGQVKQGLLGDCWFLCACAALQKSRHLLDQVIPPGQPSWADQEYRGSFTCRIWQFGRWV
Sbjct  61  REGQVKQGLLGDCWFLCACAALQKSRHLLDQVIPPGQPSWADQEYRGSFTCRIWQFGRWV  120

Query 361  FPSNPKFWLRVSEPSEVYIAVLQRSRLHAADWAGRARALVGDSHTSWSPASIPGKHYQAV  420
           FPSNPKFWLRVSEPSEVYIAVLQRSRLHAADWAGRARALVGDSHTSWSPASIPGKHYQAV
Sbjct  361 FPSNPKFWLRVSEPSEVYIAVLQRSRLHAADWAGRARALVGDSHTSWSPASIPGKHYQAV  420

Query 421  GLHLWKVEKRRVNLPRVLSMPPVAGTAC--HAYDREVHLRCELSPGYYLAVPSTFLKDAP  478
           GLHLWKV +      +    + C  H  Y + EV   C  L  G  Y   VPST+L D
Sbjct  421 GLHLWKVPEGGRSQDAPPLLLQEPLLSCVPHRYAQEVSRLCLLPAGTYKVVPSTYLPDTE  480
```

Figure 1: NCBI Align Blastp analysis of calpain 10 isoforms: Multiple alignments of Calpain 10 isoform *b, d, e* and *h* using NCBI Align Blastp tool. Different colored boxes show the pairwise similarity of protein sequences (PSPS). Black color indicates PSPS <40 amino acids, Blue color indicates PSPS 40-50 amino acids, Green color indicates PSPS 50-80 amino acids, Purple color indicates 80-200 and Red color indicates >200 amino acids pairwise similarity.

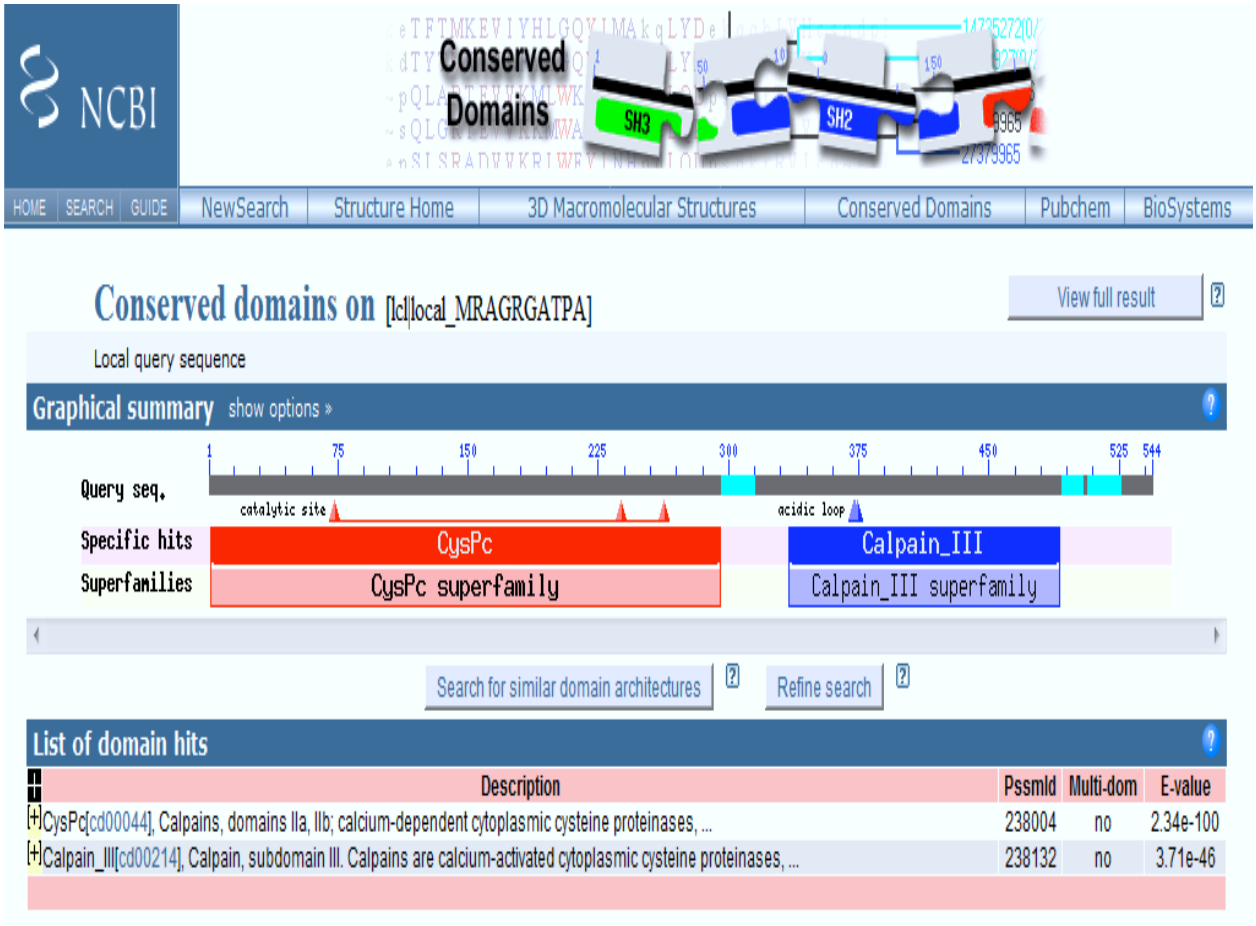

Figure 2: Conserve Domain Analysis of calpain 10 isoforms: web page results of conserve domain analysis of calpain 10 isoform *b* by NCBI Conserved Domain Tool. Graphical summary shows the hits to NCBI-curated domain models highlighted in balloons indicating non-specific hit of CysPc (Cysteine Protease) super family and specific hit of CysPc (Cysteine Protease) for domain IIa & IIb whereas Calpain_III (C2 like subdomain III) for domain III of Calpain 10 isoform *b*.

with an e-value = 0.00. Conserved domain analyses of protein sequences of isoforms (Figure **2** & Table **1**) revealed significant domain similarities of their N-terminal domain I and II with catalytic domain of cysteine protease superfamily PC1 and PC2 with an e-value of CAPN10*b* = 2.41e-76, CAPN10*d* = 2.47e-76, CAPN10*e* = 6.45e-75, CAPN10*f* = 1.07e-43 and CAPN10*h* = 1.13e-17. Calpain 10 isoform *b*, *d* and *e* have one consecutive domain (Figure **3**) which was found to be as similar with C2 like subdomain III of mu-like calpain (e-value=2.92-32, 1.03e-35, 1.88e-14) and was classified in CAPN10 group of Palb subfamily. Isoforms *f* and *h* were found to be shorter proteins and only have catalytic cysteine protease domain (Figure **4**).

DISCUSSION

Calpain 10 on the basis of its structure homology studies belongs to the cysteine protease super family. Generally, calpains function in the cytoskeletal remodeling process, cell differentiation, apoptosis and signal transduction as reported previously [10]. Calpains being proteases, serve as bio-modulators for other peptides. Our results predicted the possible activity and mechanism of action of studied protein isoforms of calpain 10 which have originated from post translational processing. Structural homology indicated that Isoform *b, d, e, f* and *h* have sequence similarity with canonical sequence of isoform *a* with shorter sequences. They do share common domains which may be involved in similar type of function as the largest isoform *a ensures* in the development of type 2 diabetes. Further all these isoforms have membrane bound sequences and actually are from mitochondrial origin as described for some members of superfamily [11]. They may play a critical role in mitochondrial dysfunction if there expression will decrease.

Conserved domain analysis of isoforms showed two types of proteolytic domains. Isoform *b, d* and *e* have

Table 1: Domain Homology Analysis of Calpain 10 Isoforms from NCBI Conserved Domain Tool

Calpain 10 Isoforms	Specific Domain Hit (Superfamily)	Conserved Domain Length (CD Length)	Bit Score	Conserve Domain (CD)	E-Value
CAPN10*b*	Cysteine proteinase	315	281.44	Cd00044	2.41e-76
	Calpain_III	150	145.42	Cd00214	2.28e-35
CAPN10*d*	Cysteine proteinase	315	281.44	Cd00044	2.47e-76
	Calpain_III	150	146.57	Cd00214	1.03e-35
CAPN10*e*	Cysteine proteinase	315	276.44	Cd00044	6.45e-75
	Calpain_III	150	75.70	Cd00214	1.88e-14
CAPN10*f*	Cysteine proteinase	315	171.66	Cd00044	1.07e-43
	Calpain_III	Absent	-	-	-
CAPN10*h*	Cysteine proteinase	315	48.40	Cd00044	5.54e-07
	Calpain_III	150	83.79	Cd00214	1.13e-17

Cd 00044(cl00051) CysPc (Cysteine Protease).
Cd00214 (Cl00165) Calpain_III (C2 like subdomain III).

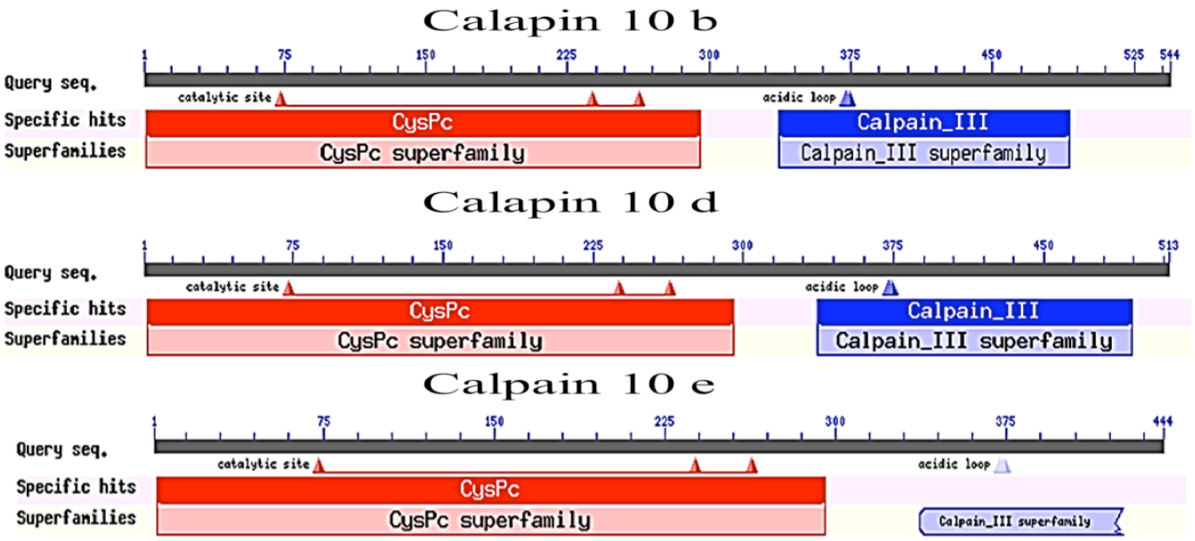

Figure 3: Conserved domain analysis of calpain 10 *b, d* and *e*: Red boxes showed the domain homologous specifically to cysteine protease superfamily and blue boxes showed domain homologous to C2 like subdomain III.

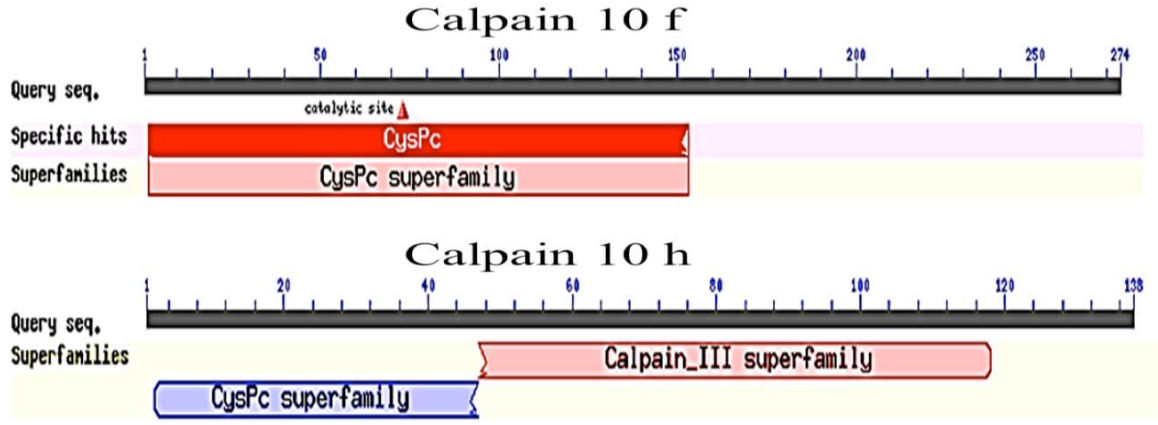

Figure 4: Conserved domain analysis of calpain 10 *f* and *h:* Red boxes showed the domain homologous specifically to cysteine protease superfamily for calpain 10 *f* and blue box of cysteine protease superfamily whereas pink box showed partially homologous domain of Calpain III superfamily in calpain 10 *h*.

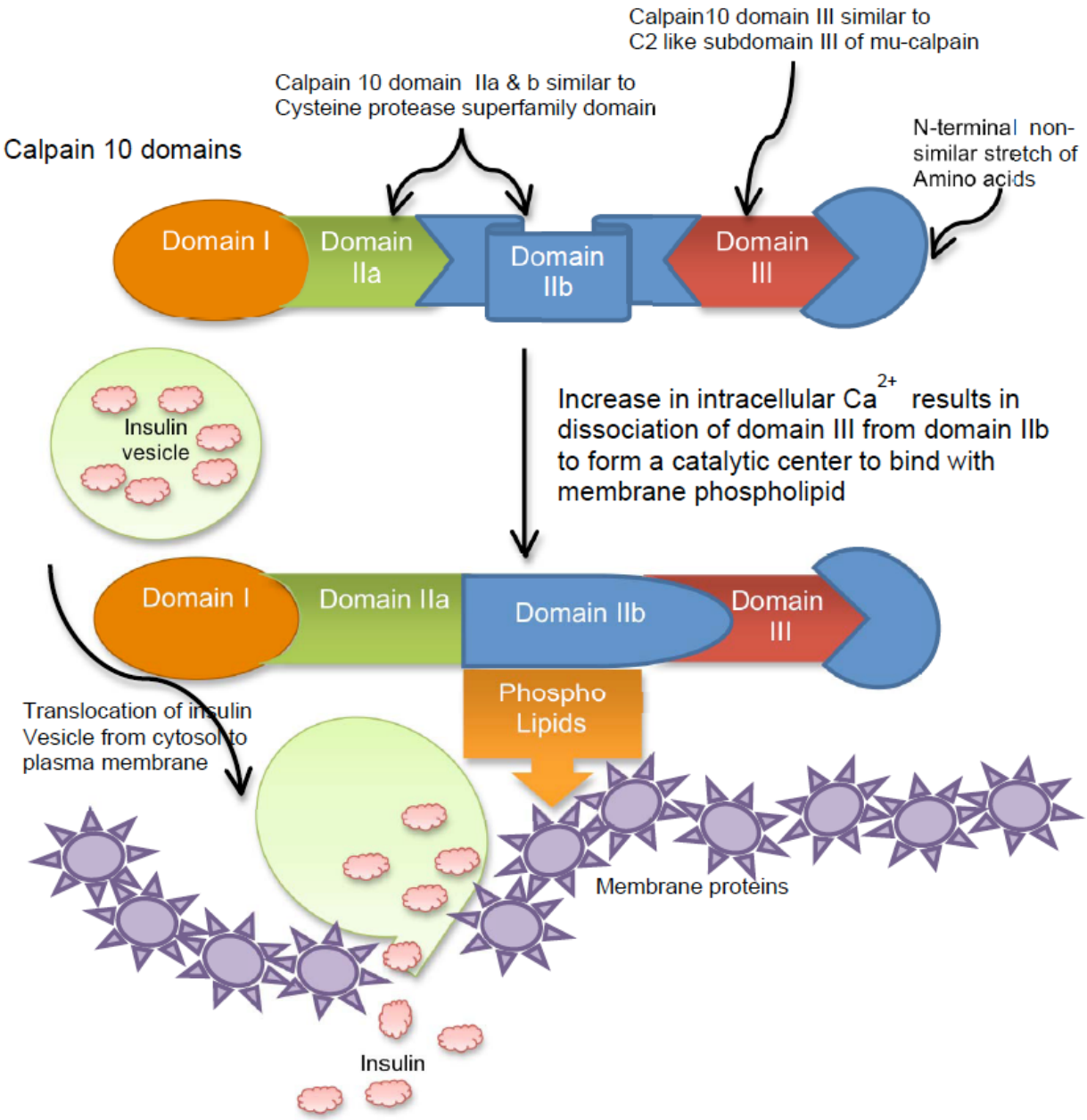

Figure 5: Model representing calpain 10 protein activation and binding: Calpain protein after activation is ready to bind with phospholipid and membrane proteins making away for the exocytosis of the insulin for secretion.

two N-terminal domains as similar to the cysteine protease superfamily domain PC1 and PC2. This domain may perform three functions 1) accepts calcium ions and induces activation of their own or other proteins, 2) it may involve in mediating interaction with phospholipids and their translocation to cytoplasmic/nuclear membranes, 3) it maintain the stability of cysteine proteases. These functions are predicted on the bases of observed similarity. This domain is followed by C2 like domain subdomain III (CD Accession: cd00214). There are two C2 like domain subdomain III in isoform *b* and *d* and one in isoform *e*. This observation predicts that isoform *b* and *d* acts like an electrostatic switch which to form a functional catalytic center which will be stronger in case of isform *b* and *d* as compared to isoform *e*. The presence of cysteine protease superfamily proteolytic domain in calpain 10 isoforms may play a vital role in calcium induced activation of their own or other proteins directly or indirectly by mediating interaction with phospholipids and their translocation to cytoplasmic/nuclear membranes. According to the previous reports, domain I and domain II exhibit protease activity because it contains catalytic active

site at locations Cys105, His262 and Asn286 [12]. Our results are lacking the application of tool to find the active site residues in isoforms. The function of subdomain III of mu-like calpain is known to have capability to bind with calcium and phospholipid as described previously [13]. We interpret that similarity with this domain in isoforms *b, d* and *e* might form electrostatic interaction with domain II to facilitate calpain interaction as shown in model (Figure **5**) with phospholipids and cell membranes (cytoplasmic/ nuclear) to release insulin which is similar as described previously for calpain 2 [12, 14]. Further we agree with the previous reports indicating that this interaction may assists to mediate the maintenance of the catalytic core in an inactive form, membrane binding and also for stabilization of the active enzyme [13, 15-17]. Isoform *f* and *h* have two domains similar to the cysteine protease superfamily; therefore it might lack the capability to bind with phospholipids and cell membrane and only possesses week protease activity, which may be random.

It is predicted from the observation that calpain 10 minor isoforms *b, d* and *e* isoforms may have a role of protein modulator as cysteine protease because of their domain I & II similarity with the protease domain of cysteine protease superfamily having domain III homology with subdomain III of mu-like calpain (calpain 2). Since these have found *b, d* and *e* isoforms contain membrane binding sequences and are originally as mitochondrial proteins there up and down regulation will be highly involved in the causation of insulin secretion, insulin resistance and hence diabetes type 2. Protein kinase C is regulated by calpain 10 isoforms being a substrate, decreased expression levels of calpains isoforms *b, d* and *e* will increase the amount of protein kinase C directly which in turn will activate the serine kinase cascade leading to increased serine phosphorylation of insulin receptor substrate 1 (IRS-1). This serine phosphorylation at critical sites will inhibits the IRS-1 tyrosine phosphorylation by insulin receptor and will results in inhibition of insulin-stimulated glucose transport and causing insulin resistance.

Further, secretion of insulin from pancreatic beta cells is a complex process involving the integration and interaction of multiple external and internal stimuli. It involves nutrients, hormones, neurotransmitters, and drugs all activate or inhibit insulin release. The calpain isoforms may also play a role because of the presence of domains, involved in binding with phospholipids make them candidates for the secretion of insulin from pancreatic beta cells.

In case of isoforms *b, d* and *e* the increased intracellular Ca^{2+} influx after elevation in the ATP/ADP ratio and cell membrane depolarization triggers the exocytosis of insulin. This may be a result of dissociation of domain III from domain II *b* to form a catalytic center to bind with membrane phospholipid and translocation of insulin vesicles from cytosol to membrane. In case if amount of calpain proteins are scanty (down regulated) there will be low level of calpain to bind with calcium and to release the catalytic centers for attachment to phospholipids and hence for release of insulin. Calpain 10 isoform *f* and *h* might have papain like activity as they only got cysteine protease superfamily domain, making them week proteases to act randomly to chop off proteins under certain physiological conditions. We include them in mitochondrial sub population as reported for some other proteins [18, 19].

REFERENCES

[1] Horikawa Y, Oda N, Cox NJ, Li X, Orho-Melander M, Hara M, *et al*. Genetic variation in the gene encoding calpain-10 is associated with type 2 diabetes mellitus. Nat Genet 2000; 26: 163-75. http://dx.doi.org/10.1038/79876

[2] Arrington DD, Van Vleet TR, Schnellmann RG. Calpain 10: a mitochondrial calpain and its role in calcium-induced mitochondrial dysfunction. Am J Physiol Cell Physiol 2006; 291: C1159-71. http://dx.doi.org/10.1152/ajpcell.00207.2006

[3] Smith MA, Covington MD, Schnellmann RG. Loss of calpain 10 causes mitochondrial dysfunction during chronic hyperglycemia. Arch Biochem Biophys 2012; 523: 161-8. http://dx.doi.org/10.1016/j.abb.2012.04.020

[4] Akopova OV, Sagach VF. Effect of Ca2+ on induction of the mitochondrial pore opening in the rat myocardium. Ukr Biokhim Zh 2004; 76: 48-55.

[5] Lowell BB, Shulman GI. Mitochondrial dysfunction and type 2 diabetes. Science 2005; 307: 384-7. http://dx.doi.org/10.1126/science.1104343

[6] Norton L, Parr T, Chokkalingam K, *et al*. Calpain-10 gene and protein expression in human skeletal muscle: effect of acute lipid-induced insulin resistance and type 2 diabetes. J Clin Endocrinol Metab 2008; 93: 992-8. http://dx.doi.org/10.1210/jc.2007-1981

[7] Aron Marchler-Bauer JBA, Farideh Chitsaz, Myra K. *et al*. "CDD: specific functional annotation with the Conserved Domain Database." Nucleic Acids Res 2009: 205-10. http://dx.doi.org/10.1093/nar/gkn845

[8] Marchler-Bauer A, Lu S, Anderson JB, *et al*. CDD: a Conserved Domain Database for the functional annotation of proteins. Nucleic Acids Res 2011; 39(Suppl 1): D225-D9. http://dx.doi.org/10.1093/nar/gkq1189

[9] Marchler-Bauer A, Bryant SH. CD-Search: protein domain annotations on the fly. Nucleic Acids Res 2004; 32(Suppl 2): W327-W31. http://dx.doi.org/10.1093/nar/gkh454

[10] Turner MD, Cassell PG, Hitman GA. Calpain-10: from genome search to function. Diabetes Metab Res Rev 2005; 21: 505-14. http://dx.doi.org/10.1002/dmrr.578

[11] Covington MD, Schnellmann RG. Chronic high glucose down regulates mitochondrial calpain 10 and contributes to renal cell death and diabetes-induced renal injury. Kidney Intl 2012; 81: 391-400. http://dx.doi.org/10.1038/ki.2011.356

[12] Franco SJ, Huttenlocher A. Regulating cell migration: calpains make the cut. J Cell Sci 2005; 118(Pt 17): 3829-38. http://dx.doi.org/10.1242/jcs.02562

[13] Croall DE, Ersfeld K. The calpains: modular designs and functional diversity. Genome Biol 2007; 8: 218. http://dx.doi.org/10.1186/gb-2007-8-6-218

[14] Tompa P, Emori Y, Sorimachi H, Suzuki K, Friedrich P. Domain III of calpain is a Ca^{2+}-regulated phospholipid-binding domain. Biochem Biophys Res Commun 2001; 280: 1333-9. http://dx.doi.org/10.1006/bbrc.2001.4279

[15] Reverter D, Strobl S, Fernandez-Catalan C, Sorimachi H, Suzuki K, Bode W. Structural basis for possible calcium-induced activation mechanisms of calpains. Biol Chem 2001; 382: 753-66. http://dx.doi.org/10.1515/BC.2001.091

[16] Moldoveanu T, Hosfield CM, Lim D, Elce JS, Jia Z, Davies PL. A Ca(2+) switch aligns the active site of calpain. Cell 2002; 108: 649-60. http://dx.doi.org/10.1016/S0092-8674(02)00659-1

[17] Strobl S, Fernandez-Catalan C, Braun M, *et al*. The crystal structure of calcium-free human m-calpain suggests an electrostatic switch mechanism for activation by calcium. Proceed Nat Acad Sci 2000; 97: 588-92. http://dx.doi.org/10.1073/pnas.97.2.588

[18] Saunders JE, Beeson CC, Schnellmann RG. Characterization of functionally distinct mitochondrial subpopulations. J Bioenerg Biomembr 2013; 45: 87-99. http://dx.doi.org/10.1007/s10863-012-9478-4

[19] Smith MA, Schnellmann RG. Calpains, mitochondria, and apoptosis. Cardiovascr Res 2012. http://dx.doi.org/10.1093/cvr/cvs163

Received on 18-06-2013 Accepted on 29-07-2013 Published on 14-08-2013

http://dx.doi.org/10.6000/1927-5129.2013.09.63