IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

Impact Factor 7.105 ⁜ Vol. 9, Issue 2, February 2022

DOI: 10.17148/IARJSET.2022.9277

Ubiquitin Associated Domain is involved in highly divergent functions: A structural and functional analysis

Niveditha M[1], Mounika R[2], Meghana C B[3], Divya Darshika A[4], Monisha M[5], Manohar G M[6], Rama T[7]

[1-7]Department of PG studies in Biotechnology, Nrupatunga University, Bangalore, Karnataka, India.

[6]halliLabs, Ragihalli, Anekal Tq, Bangalore, Karnataka

Abstract : Domains are the functional units of protein that folds and functions independently from the rest of the protein. Ubiquitin associated domains (UBA) are approximately 45 amino acid residues that help in cellular processes like nucleotide excision repair, spindle pole body duplication and cell growth. In order to see the extent of sequential and functional diversity in this domain 13 different protein-containing UBA domains in humans were randomly selected and their respective domain was retrieved through the SMART database. Multiple Sequence Alignment shows except for two amino acid positions the rest of the positions were not conserved. Phylogenetic tree showed an evolutionary relationship between these domains and showed the SQSTM1 domain as an outlier. GO molecular functions analysis from UNIPROT, showed that though there were common functions shared by some of the domains they had unique functions associated with each of them. The same methods were applied for UBA domains in different organisms. The RAD23 protein domain was taken for further analysis through Jackhammer. 2 isoforms UBA1 and UBA2 were identified in RAD23. UBA2 was more conserved among reptiles, aves, mammalians, fishes and amphibians. Though functions are spread across the different organisms they seem to be not completely unrelated. Mutational analysis on mutations in the SQSTM1 gene domain, causing Paget disease of bone (PDB), was done in model organisms. MSA and CONSURF analysis showed that I424S (UK), G425R, and A427D associated with PDB were found to be highly conserved.

Keywords: UBA domain , Multiple sequence alignment, Conserved position, Sequestasome, functional diversity, Rad23, TNRC6C, UBE2K, UBAC1, UBL7, USP5, UBXN1, UBQLN3, MARK1, CBLB, and SQSTM1, Phylogenetic tree

1. INTRODUCTION

Protein commonly consists of one or more sub molecule parts, which are termed as domains. Domain is a structural or functional module of protein, and it is usually an evolutionarily conserved unit. Differential association of domains provides a way to create new functions for organisms[17]. Ubiquitin associated Domain (UBA) interacts non-covalently with Ubiquitin in many proteins. This family of domains is involved in various cellular processes like nucleotide excision repair, spindle pole body duplication and cell growth, kinase binding, ion binding, ATP binding, and Ubiquitin ligase binding were obtained from UniProt. Structure of UBA has three helices[5]. All these proteins have a common sequence motif of approximately 45 amino acid residues.

Rad23A is a nucleotide excision repair protein which is encoded by RAD23 gene and contains UBA domain.The RAD23 homologue has a modular structure which includes N terminal ubiquitin-like (Ubl) domain and 2 UBA domains - UBA1 and UBA2. This protein binds to the ubiquitinated proteins through its two domains, that is, UBA1 and UBA2 domains[9].

When structurally compared UBA1 and UBA2, both will form similar folds with a large conserved hydrophobic surface patch which may be a common protein-interacting surface which can be seen in diverse UBA domains[6]. In many enzymes (example: Hen Lysozyme and Papain) the active sites are present at the domain interfaces between the domains[7].

Other UBA containing proteins were manually selected in humans, namely RAD23, TNRC6C, UBE2K, UBAC1, UBL7, USP5, UBXN1, UBQLN3, MARK1, CBLB, and SQSTM1. Multiple sequence alignment of the UBA domain of those selected proteins were done to find conserved residues in the Domain and found that their sequences were less conserved. Generally, proteins consist of multiple domains, the HMM model provides the constituting or corresponding domains and their locations in the amino acid sequences[10]. They are statistical probabilistic methods to obtain optimal sequence alignment, where matches are maximised and gaps are minimised, used in multiple protein sequences. HMM search of 12 UBA domains in humans revealed that there is a maximum match in between UBA domains and the HMM model sequences and HMM model for CBLB is not obtained. So, structural analysis is done by pymol.

Phylogenetic tree analysis of the above-mentioned genes which encode the UBA domains showed that SQSTM1 has evolved distinctly and functionally.
Mutations in the UBA domain have been reported in Paget's disease of bone (PDB) in humans. PDB disease is a chronic disorder which leads to increased activity of osteoblast, which results in increase of bone turnover at different sites throughout the skeleton. Common symptoms include bone pain, susceptibility to fracture and deformity[8].
Three dimensional structures were visualised through Pymol and structural analysis revealed that all the superimposed structures of the domains (UBE2K, USP5, TNRC6C, UBL7, SQSTM1, MARK1) have <3°A RMSD values, indicating that they are structurally similar[12], but sequentially they are less conserved. As RAD23 is involved in UV Excision repair mechanisms, we further analyzed the conservation and evolution of RAD23 UBA1 and 2 domains in mammals, amphibians, fishes and reptiles. GO molecular functions (uniprot) of RAD23 domain are compared in different organisms, and found that UBA2 is more conserved compared to UBA1.

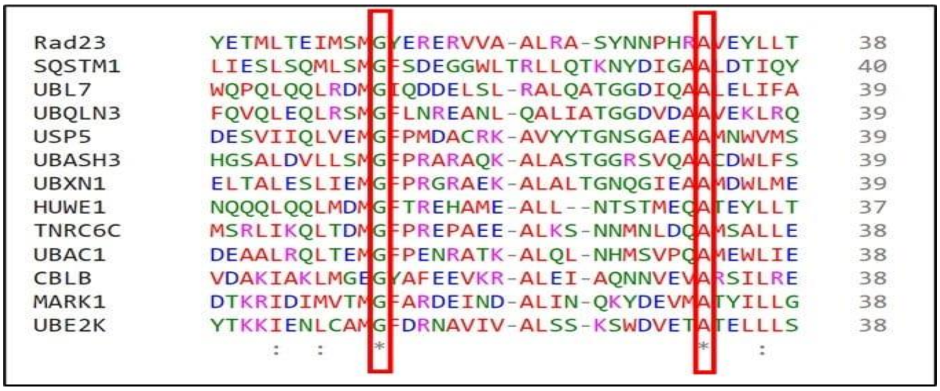
2. MATERIALS AND METHODS

Protein sequences as listed in table S1 were retrieved in FASTA format from the NCBI GenBank database (https://www.ncbi.nlm.nih.gov/genbank/). The UBA domain was identified in these sequences by using Pfam [1] (http://pfam.xfam.org/). Rad23 protein sequence(NP_001231653) retrieved from GenBank database from NCBI (https://www.ncbi.nlm.nih.gov/) was used as a query for BLASTp(database - non redundant protein sequences, Expect threshold: 0.05, Blosum62, Gap cost:- 11, extension:1) and delta blast was used (Database -non redundant protein sequences, Expected threshold - 0.05, psi blast threshold - 0.005, delta blast threshold - 0.05, matrix - Blosum 62 andGap cost:- 11, extension:1).
Through delta blast we obtained similar sequences along with isoforms, even in psi blast we got more similar sequences. By using these two database searches we selected the 13 sequences. SMART database [2] (http://smart.embl-heidelberg.de/) was used to retrieve UBA domain sequences. Multiple sequence alignment was done by CLUSTAL OMEGA(1.2.4) (https://www.ebi.ac.uk/Tools/msa/clustalo/). HMM models for the UBA domain sequences were obtained by using HMMER Tool (3.3.2). Hmmscan was used with HMMER Options: --cut_ga --hmmdb pfam. By using Jackhmmer, we selected the sequences for different organisms. GO Molecular functions are obtained from UNIPROT (https://www.uniprot.org/).
Phylogenetic tree was constructed using Phylogeny.fr (https://www.phylogeny.fr/)in the one click mode, and MUSCLE alignment.
Structural analysis was done by using PYMOL(2.4.1), RMSD values were compared from CBL B domain with other UBA domain sequences (if RMSD value is >3°A superimposed domains are structurally different, if <3° the superimposed structures are similar), and PHYRE2(2.0) [3] (http://www.sbg.bio.ic.ac.uk/phyre2/html/page.cgi?id=index) was used to analyze secondary structure of CBLB domain and compared with other UBA domains. Consurf [4] (https://consurfdb.tau.ac.il/) was used to find the conservation of residues in model organisms.

3.RESULTS

Analysis of Human UBA domains

Using a UBA domain as a key word in a SMART tool we manually retrieved the UBA domain containing proteins in humans**(S1)**. These proteins are RAD23, TNRC6C, UBASH3B, UBAC1, UBE2K, CBLB, MARK1, USP5, UBL7, SQSTM1, UBQLN3 and HUWE1, UBXN1 were obtained from SMART and delta blast. UBA Domain sequences of each of these proteins were retrieved from Smart database and used as a query in Blastp to check whether any other proteins have UBA domain. From the result list of sequences consisting of sequences from other animals, we filtered only human sequences (Homosapiens taxid:9606) and Blastp results of RAD23 in humans showed 28 sequences**(S2)** and SQSTM1 showed 10 sequences**(S3)**. Further we proceeded with Delta blast with the same parameters used in Blastp. The Delta blast result of RAD23 has shown 470 sequences**(S4)** and SQSTM1 281 sequences**(S5)** with UBA domain isoforms. As we got only similar sequences in Blastp and delta blast, we selected UBA domains in SMART. A total of 13 sequences (11 UBA domain containing protein sequences were chosen from SMART and HUWE1 and UBXN1 UBA domain containing proteins were obtained from delta BLAST) were retrieved from SMART for further analysis in humans. Doing multiple sequence alignment of those 13 sequences **(Fig 1)** we observed that only glycine (at 12) and alanine (at 34) residues are conserved in all the sequences and in the 21st position only SQSTM1 has threonine(T) residue, rest of the sequences do not have any residue in that position. Further we analyzed the GO molecular functions(UNIPROT) and evolutionary relationship of each protein by constructing a phylogenetic tree **(Fig 2)** and we observed that all the functions (listed in **S6**) are not found in all proteins. Only a few proteins has the same function- like polyubiquitin modification function was observed in RAD23,UBXN1,UBQLN3,SQSTM1 and kinase binding function was observed in RAD23, MARK1, CBLB, and SQSTM1. SQSTM1(sequestosome1) has evolved more distinctly compared to other UBA domain proteins and has more functions.



```
Rad23       YETMLTEIMSMGYERERVVA-ALRA-SYNNPHRAVEYLLT    38
SQSTM1      LIESLSQMLSMGFSDEGGWLTRLLQTKNYDIGAALDTIQY    40
UBL7        WQPQLQQLRDMGIQDDELSL-RALQATGGDIQAALELIFA    39
UBQLN3      FQVQLEQLRSMGFLNREANL-QALIATGGDVDAAVEKLRQ    39
USP5        DESVIIQLVEMGFPMDACRK-AVYYTGNSGAEAAMNWVMS    39
UBASH3      HGSALDVLLSMGFPRARAQK-ALASTGGRSVQAACDWLFS    39
UBXN1       ELTALESLIEMGFPRGRAEK-ALALTGNQGIEAAMDWLME    39
HUWE1       NQQQLQQLMDMGFTREHAME-ALL--NTSTMEQATEYLLT    37
TNRC6C      MSRLIKQLTDMGFPREPAEE-ALKS-NNMNLDQAMSALLE    38
UBAC1       DEAALRQLTEMGFPENRATK-ALQL-NHMSVPQAMEWLIE    38
CBLB        VDAKIAKLMGEGYAFEEVKR-ALEI-AQNNVEVARSILRE    38
MARK1       DTKRIDIMVTMGFARDEIND-ALIN-QKYDEVMATYILLG    38
UBE2K       YTKKIENLCAMGFDRNAVIV-ALSS-KSWDVETATELLLS    38
                :   :    *                 *    :
```

Fig 1 : Multiple Sequence alignment of Different UBA domains inHuman(CLUSTAL OMEGA 1. 2.4)

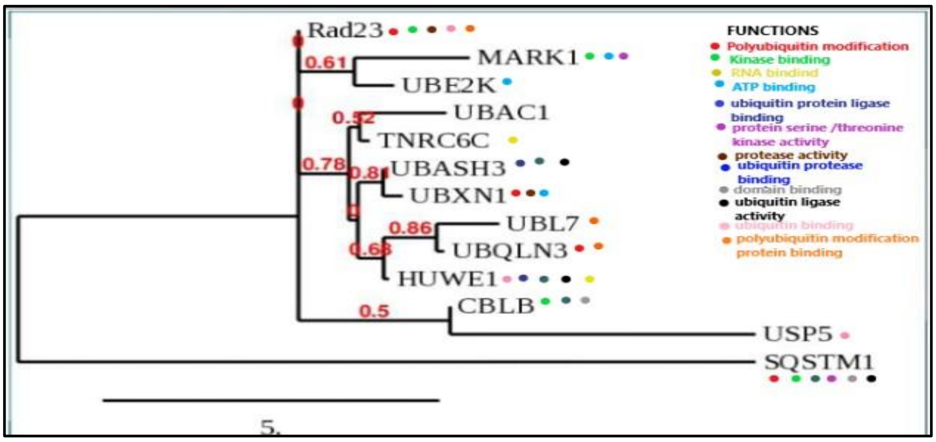

Rad23
0.61
MARK1
UBE2K
0.52
UBAC1
TNRC6C
0.78
0.8
UBASH3
UBXN1
0.86
UBL7
0.63
UBQLN3
HUWE1
0.5
CBLB
USP5
SQSTM1
5.
FUNCTIONS
● Polyubiquitin modification
● Kinase binding
● RNA bindind
● ATP binding
● ubiquitin protein ligase binding
● protein serine /threonine kinase activity
● protease activity
● ubiquitin protease binding
● domain binding
● ubiquitin ligase activity
● ubiquitin binding
● polyubiquitin modification protein binding

Fig 2 : Phylogenetic tree of 13 UBA domains of human (molecular functions are highlighted)

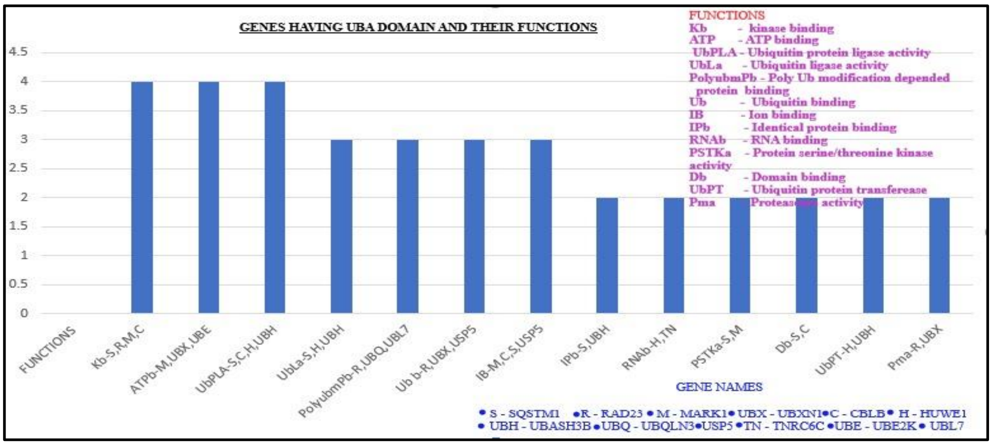


Fig 3 : Bar graph showing the functions and respective genes in x - axis and the number of genes possessing the same functions in y – axis

DOI: 10.17148/IARJSET.2022.9277

Bar graph was plotted for functions possessed by each UBA domain containing proteins(**Fig 3**). We found that SQSTM1 gene has 10 functions and RAD23, MARK1, UBASH3B, CBLB has 5 functions each and HUWE1, UBXN1 has 4 functions and UBQLN3, and USP5 has 2 functions. All the genes do not possess all the functions.

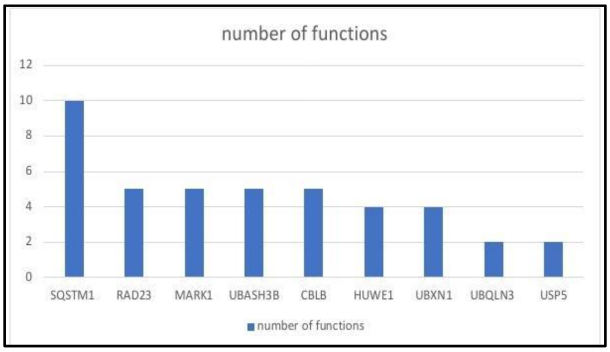


Fig 4 : BAR graph showing number of functions in y - axis and genes in x-axis

HMM model was constructed for UBA domain containing proteins to find the maximum matches from the databases. There was a maximum match between each of the domain sequences and the HMM model, but HMM of the CBLB UBA domain was not found. To get more information about it, further we proceed with the phyre2 tool and pymol for structural analysis. Secondary structure of CBLB and RAD23 UBA domains obtained from phyre2 we found that RAD23 has 3 helices**(S7)** and CBLB had shown 2 helices**(S8)**, but the confidence of the alpha helix(in between second and third helix)was low, which means that there is some possibility that it might not be a helix and there is a break in the long helix. So there is a possibility that it could be 3 helices. Then we proceeded with structural comparison of CBLB with other UBA domain proteins PYMOL(2.4.1) and 3D structures **(S9)** are observed.

Here is the list of PDB ID of UBA domains and RMSD value

3k9o (UBE2K)	0.950	2dkl (TNRC6C)	2.015
2dai (UBAC1)	1.805	2cwb (UBL7)	2.265
2dah (UBQLN3)	1.852	1qo2 (SQSTM1)	1.757
2dag (USP5)	1.987	2hak (MARK1)	1.757

The RMSD values are found to be less than 3A°,so these domain sequences are structurally similar to CBLB.

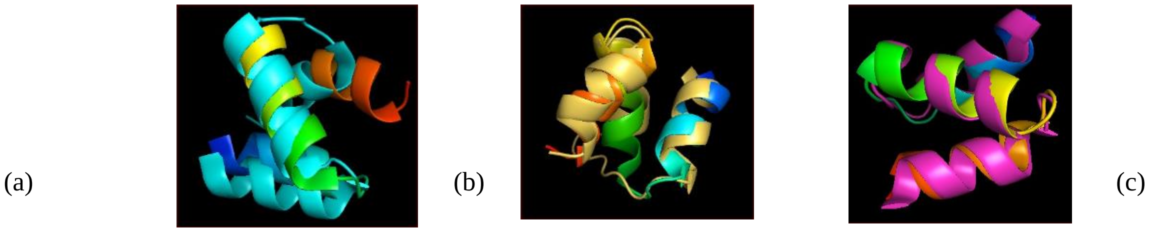
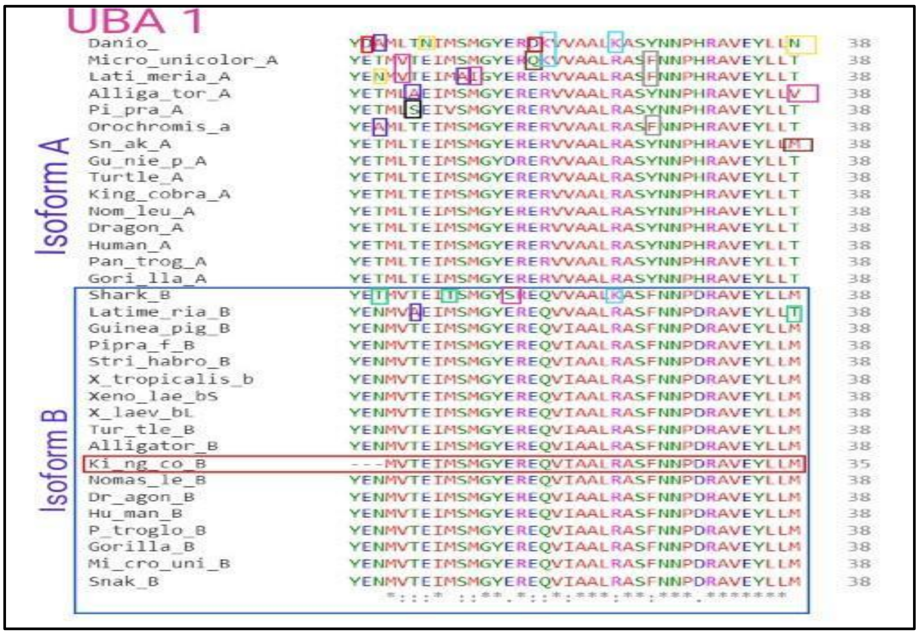


Fig 5: Superimposed structures of CBLB with (a)MARK1 (b) SQSTM1 (c)UBE2K

UBA Domain Analysis In Different Organisms

RAD23 UBA domains are analyzed in different organisms from the JACKHAMMER and we studied their evolution and functions. Rad23 has both the isoforms of the UBA domain. In different organisms RAD23 UBA 1 and 2 domain sequences are retrieved from SMART and were further analyzed. Multiple sequence alignment and Phylogenetic tree of RAD23 UBA1 and 2 domain isoforms were done separately. The UBA 1 **(Fig 5)** domain of RAD23 isoform A has few sequences showing change in amino acid residues compared to RAD23 B isoform(only fishes were showing change in few residues). In isoform B King cobra -first 3 residues are deleted. Both isoforms of RAD23 UBA2 **(S10)** domain are highly conserved in all organisms with a change in few residues in isoform B (in tiger snake, and bony fishes-Danio rerio, Oreochromis). In isoform A -Latimeria (vertebrate fish) there is asparagine (N)amino acid at the 16th position, but the rest of the sequence are having serine(S) residue. Hence the multiple sequence alignment showed that the UBA2 of RAD23 was highly conserved when compared to UBA1 of RAD23 domain.

Result obtained from the phylogenetic tree**(S11)** have revealed that branch length of Danio rerio(isoform A) is more compared to other organisms, isoform A and B evolved differently and has common ancestors and though 3 residues in King cobra was deleted (from multiple sequence alignment), it is evolved as same as other organisms. GO molecular functions of the UBA1 RAD23 domain were obtained from UNIPROT **(S12)** and evolution of functions were analyzed. Damaged DNA binding, kinase binding and ubiquitin binding functions were found in both isoform A and B of all of the organisms. Some functions are the same in some organisms which means not all genes possess all functions. Human RAD23 domains have shown more functions compared to domains in other organisms. Ubiquitin specific protease binding, DNA binding and proteasome binding these functions are found only in Nomascus(monkey), Gorilla, Human and Pantroglodytes isoform A and in isoform B this function was found only in humans. Though the Danio rerio sequence has longer branch length, it also has the same functions compared to other organisms.



```
Danio_              YDAMLTNIMSMGYERDKVVAALKASYNNPHRAVEYLLN   38
Micro_unicolor_A    YETMVTEIMSMGYEHQKVVAALRASFNNPHRAVEYLLT   38
Lati_meria_A        YENMVTEIMAEGYERERVVAALRASFNNPHRAVEYLLT   38
Alliga_tor_A        YETMLAEIMSMGYERERVVAALRASYNNPHRAVEYLLV   38
Pi_pra_A            YETMLSEIVSMGYERERVVAALRASYNNPHRAVEYLLT   38
Orochromis_a        YEAMLTEIMSMGYERERVVAALRASFNNPHRAVEYLLT   38
Sn_ak_A             YETMLTEIMSMGYERERVVAALRASYNNPHRAVEYLLM   38
Gu_nie_p_A          YETMLTEIMSMGYDRERVVAALRASYNNPHRAVEYLLT   38
Turtle_A            YETMLTEIMSMGYERERVVAALRASYNNPHRAVEYLLT   38
King_cobra_A        YETMLTEIMSMGYERERVVAALRASYNNPHRAVEYLLT   38
Nom_leu_A           YETMLTEIMSMGYERERVVAALRASYNNPHRAVEYLLT   38
Dragon_A            YETMLTEIMSMGYERERVVAALRASYNNPHRAVEYLLT   38
Human_A             YETMLTEIMSMGYERERVVAALRASYNNPHRAVEYLLT   38
Pan_trog_A          YETMLTEIMSMGYERERVVAALRASYNNPHRAVEYLLT   38
Gori_lla_A          YETMLTEIMSMGYERERVVAALRASYNNPHRAVEYLLT   38
Shark_B             YETMVTEITSMGYSREQVVAALKASFNNPDRAVEYLLM   38
Latime_ria_B        YENMVAEIMSMGYEREQVVAALRASFNNPDRAVEYLLT   38
Guinea_pig_B        YENMVTEIMSMGYEREQVIAALRASFNNPDRAVEYLLM   38
Pipra_f_B           YENMVTEIMSMGYEREQVIAALRASFNNPDRAVEYLLM   38
Stri_habro_B        YENMVTEIMSMGYEREQVIAALRASFNNPDRAVEYLLM   38
X_tropicalis_b      YENMVTEIMSMGYEREQVIAALRASFNNPDRAVEYLLM   38
Xeno_lae_bS         YENMVTEIMSMGYEREQVIAALRASFNNPDRAVEYLLM   38
X_laev_bL           YENMVTEIMSMGYEREQVIAALRASFNNPDRAVEYLLM   38
Tur_tle_B           YENMVTEIMSMGYEREQVIAALRASFNNPDRAVEYLLM   38
Alligator_B         YENMVTEIMSMGYEREQVIAALRASFNNPDRAVEYLLM   38
Ki_ng_co_B          ---MVTEIMSMGYEREQVIAALRASFNNPDRAVEYLLM   35
Nomas_le_B          YENMVTEIMSMGYEREQVIAALRASFNNPDRAVEYLLM   38
Dr_agon_B           YENMVTEIMSMGYEREQVIAALRASFNNPDRAVEYLLM   38
Hu_man_B            YENMVTEIMSMGYEREQVIAALRASFNNPDRAVEYLLM   38
P_troglo_B          YENMVTEIMSMGYEREQVIAALRASFNNPDRAVEYLLM   38
Gorilla_B           YENMVTEIMSMGYEREQVIAALRASFNNPDRAVEYLLM   38
Mi_cro_uni_B        YENMVTEIMSMGYEREQVIAALRASFNNPDRAVEYLLM   38
Snak_B              YENMVTEIMSMGYEREQVIAALRASFNNPDRAVEYLLM   38
                       *:::*  ::**.*::*:***:**:***.*******
```

Fig 6 : Isoform A and B of UBA1 (RAD23) in different organisms

SQSTM1 mutation

Paget disease of bone(PDB) associated UBA domain mutation of SQSTM1 exert distinct effects on protein structure and function . Mutation sites are relatively common in UK PDB patients, mutational sites are I424S (I-Isoleucine, S- serine), G425R(G-glycine, R-arginine), A427D (A-alanine,D-aspartate) and these are associated with a severe phenotypic changes in southern Italian patients[8]. To check whether these mutations were there in other organisms, multiple sequence alignment (S13) and phylogenetic tree(S14) is constructed by taking a SQSTM1 domain sequence in model organisms(mammals,fishes,reptiles and aves). Mutational sites A427D and G425R are conserved in all model organisms (S16) and CONSURF results(S15) of SQSTM1 UBA domain (PDB ID : 2k0b)showed that at position 40 Isoleucine is conserved, at 41 Glycine is variable, and at 43 Alanine is partially conserved. Phylogenetic tree showed the evolutionary characteristics between 17 model organisms. Xenopus tropicalis and pinecone soldierfish are showing the different branches compared to other 15 Model organisms in the phylogenetic tree, upon evolution they acquired unique functions too (UNIPROT)(16).

4.DISCUSSION

The structure of the UBA domains was conserved across different genes, though the sequence were radically diverged, indicates that there is an evolutionary constraint on the structure for binding and other functions. The UBA domain is an extensive example as to however though the sequence varies, the structure remains the same. Since we found the same UBA domain in multiple genes sometimes each with multiple functions we think this may shed light on how new genes evolve. Our results give credence to the theory that new genes could evolve through mix and match of already existing domains.

5.CONCLUSION

In humans UBA domain in RAD23, TNRC6C, UBE2K, UBAC1, UBL7, USP5, UBXN1, UBQLN3, MARK1, CBLB, and SQSTM1 domains are sequentially less conserved but similar in their structures. They all share a common origin but have diverged sequentially. They have different functions but still belong to the UBA family of domains. This shows that a single domain may be involved in multiple and unrelated functions both in Humans and in other organisms. Mixing a domain with other domains may result in novel functions. In different organisms RAD23 UBA2 is more conserved than UBA1. Though functions are spread across the different organisms they seem to be not completely unrelated.

ACKNOWLEDGEMENT :

We deeply appreciate the help and mentoring from Dr. Meenakshi, Ph D, NCBS, for regular inputs for this research.

Supplementary material :

https://docs.google.com/document/d/1QkYlrIdnwM1SqLGgzqOslZAuC3XtcO-0SaCmCkfusPQ/edit

REFERENCE

1. Pfam: The protein families database in 2021: J. Mistry, SChugurensky, L. Williams, M. Qureshi, G.A.Salazar, E.L.L. Sonnhammer, S.C.E. Tosatto, L. Paladin, S. Raj, L.J. Richardson, R.D. Finn, A. Bateman , Nucleic Acids Research (2020) doi: 10.1093/nar/gkaa913
2. Ivica Letunic, Peer Bork, 20 years of the SMART protein domain annotation resource, Nucleic Acids Research, Volume 46, Issue D1, 4 January 2018, Pages D493–D496,
3. Kelley, L., Mezulis, S., Yates, C. et al.The Phyre2 web portal for protein modelling, prediction and analysis. Nat Protoc 10, 845–858 (2015).
4. Ben Chorin A., Masrati G., Kessel A., Narunsky A., Sprinzak J., Lahav S., Ashkenazy H. and Ben-Tal N. (2020). ConSurf-DB: An accessible repository for the evolutionary conservation patterns of the majority of PDB proteins. Protein Science 29:258–267. Goldenberg O., Erez E., Nimrod G. and Ben-Tal N. (2009). The ConSurf-DB: Pre-calculated evolutionary conservation profiles of protein structures. Nucleic Acids Research (Database issue), 37:D323-D327; PMID: 18971256.
5. Dieckmann, T et al. “Structure of a human DNA repair protein UBA domain that interacts with HIV-1 Vpr.” Nature structural biology vol. 5,12 (1998): 1042-7. doi:10.1038/4220Dieckmann, T et al. “Structure of a human DNA repair protein UBA domain that interacts with HIV-1 Vpr.” Nature structural biology vol. 5,12 (1998): 1042-7. doi:10.1038/4220Dieckmann, T et al. “Structure of a human DNA repair protein UBA domain that interacts with HIV-1 Vpr.” Nature structural biology vol. 5,12 (1998): 1042-7. doi:10.1038/4220
6. Mueller, Thomas D, and Juli Feigon.“Solution structures of UBA domains reveal a conserved hydrophobic surface for protein-protein interactions.” Journal of molecular biology vol.319,5(2002):1243-55.doi:10.1016/S0022-2836(02)00302-9
7. Janin, J, and S J Wodak. “Structural domains in proteins and their role in the dynamics of protein function.” Progress in biophysics and molecular biology vol. 42,1 (1983): 21-78. doi:10.1016/0079-6107(83)90003-2
8. Layfield, R & Ciani, Barbara & Ralston, Stuart & Hocking, Lynne & Sheppard, P & Searle, M & Cavey, J. (2004). Structural and functional studies of mutations affecting the UBA domain of SQSTM1 (p62) which cause Paget's disease of bone: Figure 1. Biochemical Society transactions. 32. 728-30. 10.1042/BST0320728.
9. Goh, Amanda M et al. “Components of the ubiquitin-proteasome pathway compete for surfaces on Rad23 family proteins.” BMC biochemistry vol. 9 4. 30 Jan. 2008, doi:10.1186/1471-2091-9-4
10. Yoon, Byung-Jun. “Hidden Markov Models and their Applications in Biological Sequence Analysis.” Current genomics vol. 10,6 (2009): 402-15. doi:10.2174/138920209789177575 (HMM,which%20are%20not%20directly%20observable.&text=The%20hidden%20states%20form%20a,depends%20on%20the%20underlying%20state.
11. https://www.cell.com/current-biology/comments/S0960-9822(97)70070-8
12. Ramírez, David, and Julio Caballero. “Is It Reliable to Take the Molecular Docking Top Scoring Position as the Best Solution without Considering Available Structural Data?” Molecules, vol. 23, no. 5, Apr. 2018, p. 1038. Crossref,
13. Ye, Yu et al. “Dissection of USP catalytic domains reveals five common insertion points.” Molecular bioSystems vol. 5,12 (2009): 1797-808. doi:10.1039/b907669g
14. Andersen, Katrine M et al. “Ubiquitin-binding proteins: similar, but different.” Essays in biochemistry vol. 41 (2005): 49-67. doi:10.1042/EB0410049
15. Swanson, Kurt A et al. “Structural basis for monoubiquitin recognition by the Ede1 UBA domain.” Journal of molecular biology vol. 358,3 (2006): 713-24. doi:10.1016/j.jmb.2006.02.059

DOI: 10.17148/IARJSET.2022.9277

16. Zhou, Zi-Ren et al. “Differential ubiquitin binding of the UBA domains from human c-Cbl and Cbl-b: NMR structural and biochemical insights.” Protein science : a publication of the Protein Society vol. 17,10 (2008): 1805-14. doi:10.1110/ps.036384.108
17. Xue Li, Lifeng Yang, Xiaopan Zhang, Xiong Jiao, "Prediction of Protein-Protein Interactions Based on Domain", Computational and Mathematical Methods in Medicine, vol. 2019, Article ID 5238406, 7 pages, 2019.