Computational Structural Analysis of C-Terminal Residues of Proteins Containing Transmembrane Regions

M.C. Rishyakulya[1], K.M. Saravanan[1*]

[1] Centre of Excellence in Bioinformatics, School of Biotechnology, Madurai Kamaraj University, Madurai 625 021

Article Info	ABSTRACT
	For the past few years, the numbers of transmembrane protein structures in Protein Data Bank have been increased substantially. It is of interest to analyze the terminal residues of transmembrane proteins by using computational approaches. Also, up to our knowledge, no analysis was reported in the literature on the study of terminal residues in transmembrane proteins. While the N-terminal position of alpha and beta transmembrane proteins are composed of signal peptides, in the present work, a careful, in-depth, computational analysis such as residue preference, stability upon mutation, solvent accessibility, hydrogen bonding and carboxy terminal pentapeptide pattern search respectively has been done on C-terminal residues. Alanine in alpha transmembrane proteins and phenylalanine in beta transmembrane proteins are highly preferred. Glutamic acid and glycine residues can be substituted at the terminal sites of alpha and beta transmembrane proteins without affecting the protein's overall stability. Hydrogen bonding of terminal residues is studied in detail. Pattern search of carboxy pentapeptides shows that identical pentapeptides with reference to the position can adopt a different secondary structure. The results discussed in this paper may help to understand the role of carboxy terminal residues in alpha and beta transmembrane proteins. From our analysis, we insist that the preferences and structural analysis of carboxy terminal residues in alpha and beta transmembrane proteins, can help to model and design novel transmembrane proteins.

Corresponding Author:

KM Saravanan,
Centre of Excellence in Bionformatics,
Madurai Kamaraj University,
Madurai, india.
Email: saravananbioinform@gmail.com

How to Cite:

Rishyakulya *et. al.* Computational Structural Analysis of C-Terminal Residues of Proteins Containing Transmembrane Regions. IJCB. 2015; Volume 4 (Issue 1): Page 44-54.

1.   INTRODUCTION

Proteins are linear amino acid polymers flanked by amino (N-) and carboxy (C-) termini. The amino and carboxy termini of proteins play a vital role in the folding and functioning of proteins [1]. They are also important for the conformational stability of proteins [2]. These terminal residues serve as a binding site to interact with other proteins to form protein-protein interactions to perform a specific function [3-7]. Another well known function of terminal residues is their ability to serve as a cellular targeting signal [8]. Also, there are several hypotheses that the terminal residues and last few amino acid residues may interact with various release factors involved in translational termination [9]. Considering the above facts, protein science groups have made several detailed theoretical and experimental analysis on amino and carboxy terminal of folded globular proteins.

One such experiment suggests that the terminal residues have comparatively more accessible surface area than other residues [10]. This is because, if the terminal residues are exposed, it can easily interact with other proteins. Another analysis on protein termini reveals that the proteins having terminal residues in the core (not on the surface) are difficult to fold and hence the tendency of the terminal residues to be on the surface is increased [11]. Pal and Chakrabarti have made a computational analysis on the known protein structures to find out if there is any pattern in the type of residues used and their conformation at the two terminal positions of the polypeptide chains [12]. Likewise, several analyses on protein termini is reported on globular proteins [13-17]. By looking the literature carefully, it is found that the studies and reports on terminal residues of membrane proteins are limited. It is of interest to analyze the terminal residues of membrane proteins using various computational procedures, and hence we have made a detailed analysis on carboxy terminal residues of proteins containing transmembrane regions.

Transmembrane proteins are major classes of membrane proteins which is involved in diverse functions in the cell like transport, ion transmission and catalysis. These transmembrane proteins are further classified into two groups based on its structure such as alpha transmembrane proteins (made up of alpha helices) and beta transmembrane proteins (made up of beta barrels). Although, transmembrane proteins are targets of many pharmaceutical developments, they are not paid much attention like globular proteins. This is due to the hardness of structure elucidation of transmembrane proteins. The crystal structures of several membrane proteins are of low resolution [18]. Currently, the number of membrane protein structures is increased in the Protein Data Bank and now it has more than 300 membrane-protein structures with 30-40 unique folds [19]. Considering the above facts, in the present work, a careful computational analysis has been done on the carboxy terminal residues of the proteins containing transmembrane regions. The carboxy terminal residues are subjected to mutation analysis to find favorable and unfavorable residues at that position. The five residues at the terminal is considered and subjected to pattern search in the dataset of proteins containing regular secondary structures. Further, hydrogen bonding patterns of the terminal residues are studied.

2.   RESEARCH METHOD

2.1 Dataset
We have obtained non-redundant alpha and beta transmembrane protein sequences and structures from PDBTM database [18]. As of December 2013, there are 176 alpha transmembrane proteins and 65 beta transmembrane proteins. It is noted that the significant fraction of the proteins in the dataset is composed of several chains or domains. Further, in order to make a pattern search of terminal carboxy pentapeptide residues with globular proteins, we used a pre-compiled dataset of globular proteins derived from the latest release of the PDB [19]. The dataset contains 5877 nonredundant protein chains and each residue in the dataset is assigned with regular secondary structure (Helix, Sheet or Coil) and solved experimentally at high resolution (less than 3 Å). The proteins in the dataset share less than 30% sequence identity [20].

2.2 Extraction of terminal residues and terminal carboxy pentapeptides
Since, the N-terminal region of the membrane proteins consist of signal peptides, the C-terminal residues were considered in our analysis. Terminal residues were extracted from the PDB files by using in-house Perl scripts. For example, the protein with PDB ID: 1YC9, the C-terminal of chain A is considered. The last residue in the chain is not a phenylalanine as shown in PDB file. A visual inspection of the structure reveals that the C-terminal of chain A is actually far from the membrane, but there is a phenylalanine in the C-terminal segment of

the transmembrane domain. Hence, we have not omitted above proteins from the dataset and our analysis is with respect to the domain organization of proteins. Also, there are some missing residues in the C-terminal region of the structure. Five carboxy terminal residues (Carboxy pentapeptides) were also extracted and kept in appropriate format for further computations. We have computed the percentage of number of occurrences of twenty amino acid residues at carboxy terminal position of alpha and beta transmembrane proteins.

2.3 Phylognetc analysis and gene ontology terms of transmembrane proteins

We have performed multiple sequence alignment on the alpha and beta transmembrane protein sequences to understand the evolutionary relationship (65 alpha and 63 beta transmembrane proteins) in the context of C-terminal residues using clustalw. Further, we have extracted gene ontology terms from the annotations page of the Protein Data Bank to understand the functional consequence of each proteins in the dataset [19].

2.4 Computation of protein stability upon point mutations

The carboxy terminal residues of alpha and beta transmembrane proteins were subjected to point mutation with all other 19 amino acids and the ΔΔG is calculated by using CUPSAT (Cologne University Protein Stability Analysis Tool) server [21]. This program uses structural environment specific atom potentials and torsion angle potentials to predict ΔΔG. ΔΔG is the difference in free energy of unfolding between wild-type and mutant proteins. The ΔΔG values obtained from the CUPSAT server were considered for further analysis. Also, this server calculates the solvent accessibility of the mutated residues.

2.5 Computation of hydrogen bonding patterns

Hydrogen bonding patterns of carboxy terminal residues were examined using Struct tools server at high performance computing at National Institute of Health (NIH). The server is freely accessible at http://helixweb.nih.gov/structbio/. This server computes hydrogen bond distance and the angle between the atoms of donor (CNH_O) and acceptor (NH_OC). The results are tabulated in appropraite format to perform further analysis.

2.6 Searching identical carboxy pentapeptide patterns in folded proteins

Five residue carboxy terminal segments (carboxy pentapeptide) of each of the transmembrane proteins were searched against the ccPDB sequences (5877 sequences) [20]. Secondary structures of identically matched segments were obtained. The results thus obtained were classified into two major groups such as identical pentapeptides with identical secondary structures and identical pentapeptides with different secondary structures. The whole searching and secondary structure assignment process was performed using an in-house Perl program.

3.   RESULTS AND ANALYSIS

3.1 Percentage of occurrences of amino acid residues at C-terminal position

In protein sequence analysis, looking at the number of occurrences of a particular amino acid type at a specific position is important. Hence, the percentage of occurrences of twenty amino acids at carboxy terminal sites of transmembrane (alpha and beta) and globular proteins is computed and shown in figure 1. From the figure, it is found that the alanine occurs in higher percentage (13.84%) in terminal sites of alpha transmembrane proteins whereas, phenylalanine (57.14%) occurs in higher percentage in terminal sites of beta transmembrane proteins. In alpha transmembrane proteins, charged amino acid residues such as arginine and glutamic acid occur relatively at higher percentage in carboxy terminal sites. The amino acid 'leucine' occurs relatively in higher percentage along with alanine in terminal sites of both alpha and beta transmembrane proteins.
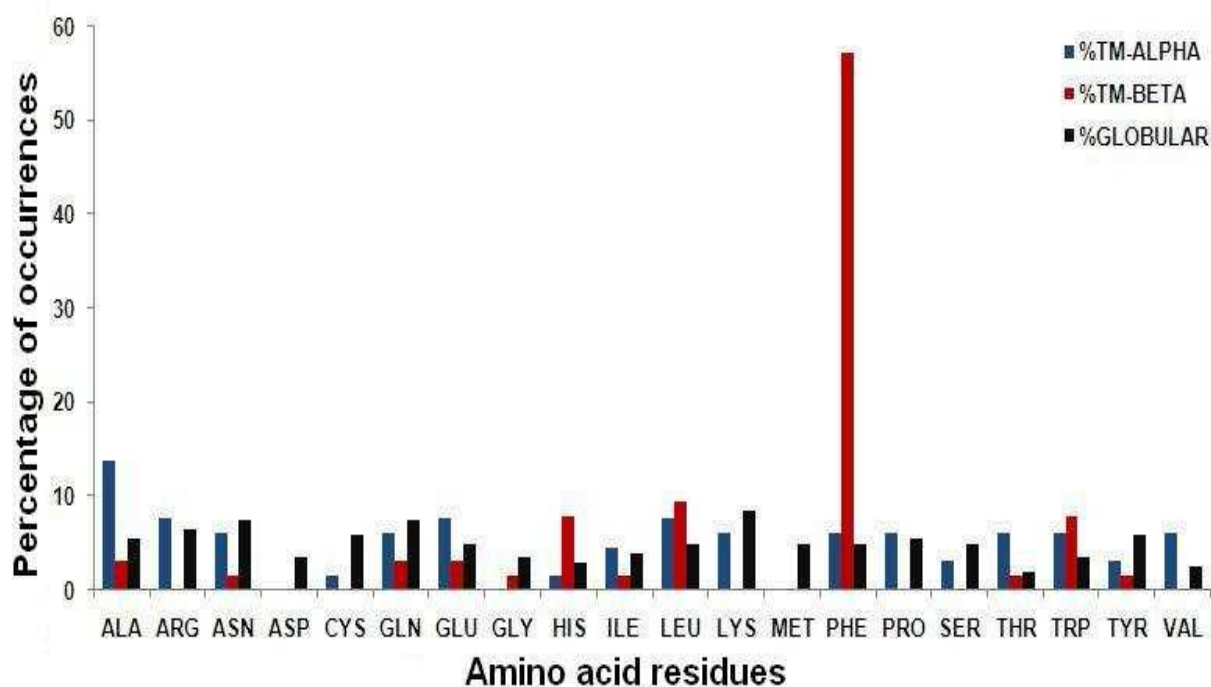
Figure 1. Percentage of occurrences of amino acid residues at C-terminal positions

The observation that the occurrences of higher percentage of alanine residues in terminal sites of alpha transmembrane proteins agree with the observation of occurrences of higher percentage of alanine at terminal sites of globular proteins [12]. This implies that the frequent distribution of alanine at the terminal site of alpha transmembrane proteins. Occurrence of phenylalanine at the carboxy terminal in beta transmembrane proteins is important for efficient and correct assembly of the protein and it is also observed that the residues with hydrophobic property occurs relatively in higher percentage which suggests that the presence of hydrophobic residues make the chain termini less mobile [22]. In the case of alpha transmembrane proteins, residues such as aspartic acid, glycine and methionine no longer occurs at carboxy terminal whereas in the case of beta transmembrane proteins, residues such as arginine, cysteine, lysine, methionine, proline and serine occurs no longer at the carboxy terminal. In both classes of transmembrane proteins, methionine and aspartic acid have no occurrence.

3.2. Phylogeny and gene ontology terms of transmembrane proteins

The phylogenetic tree of alpha and beta transmembrane proteins is shown in figure 2. By looking the figures carefully, it is found that the multiple sequence alignment of transmembrane proteins cannot form clusters or groups based on the terminal residues. This is due to the low sequence identity between the proteins in the dataset. The dataset thus used is non-redundant and the sequences in the dataset share less than 30% sequence identity. The gene ontology terms of each of the proteins in the dataset is shown in Supplementary Material S1. Most of the alpha transmembrane proteins in the dataset have the gene ontology terms as electron carrier and metal ion binding. There are other terms such as ion channel activity, transporter, ammonium binding and proteolysis repectively. In the case of beta transmembrane ptoeins, most of the proteins have gene ontology terms as transporter and porin activity. Eleven beta transmembrane proteins have no annotations.
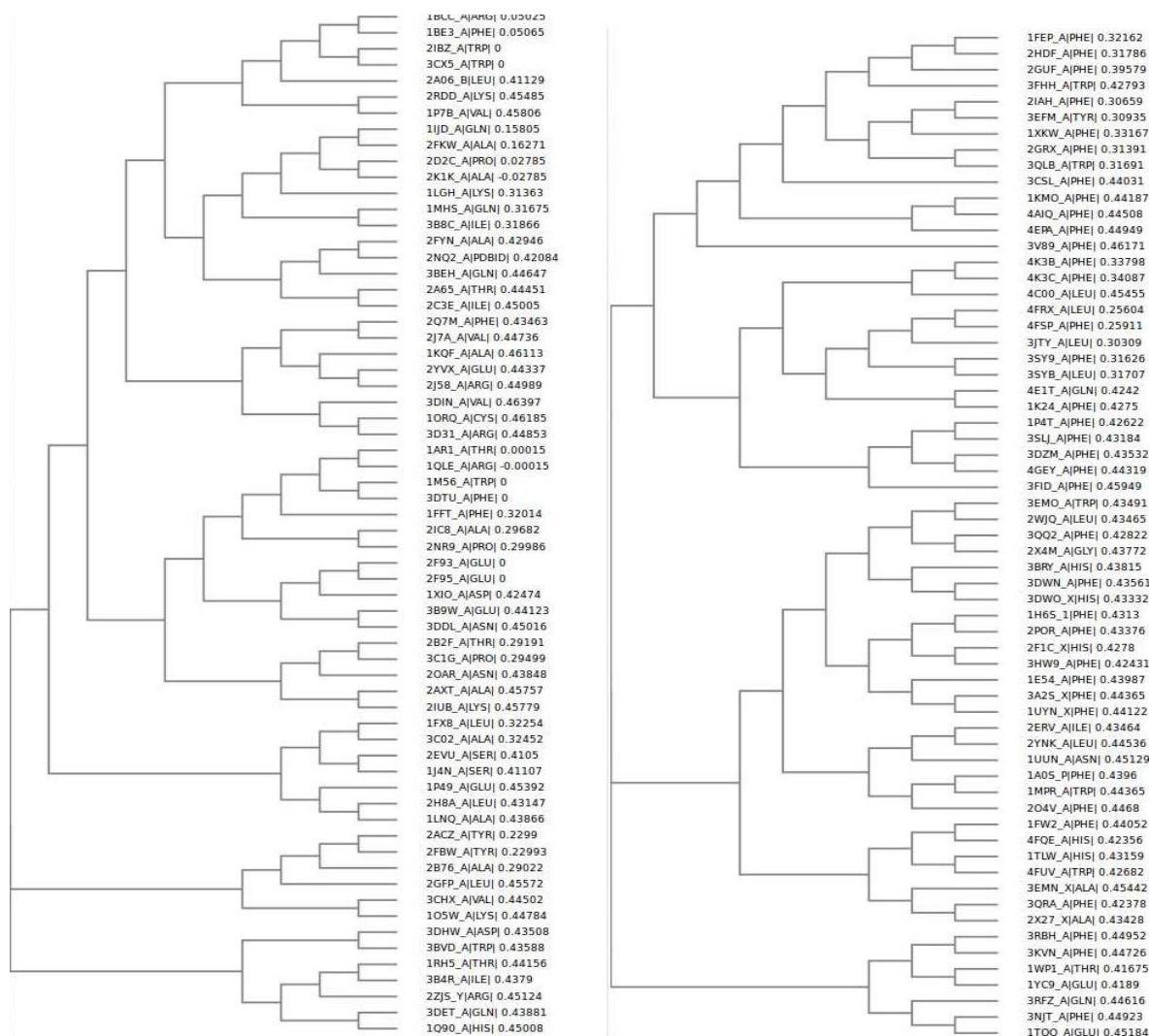
Figure 2. Phylogenetic tree of alpha (left) and beta (right) transmembrane proteins

3.3 Mutational analysis and solvent accessibility of carboxy terminal sites of transmembrane protein

In order to find the mutable and immutable residues at the carboxy terminal site of the alpha and beta transmembrane proteins, mutational analysis is carried out. The terminal site is mutated with all nineteen amino acids (except wild type amino acid). The average $\Delta\Delta G$ values for each amino acid residue at carboxy terminal is computed and presented in figure 3. In the case of alpha transmembrane proteins, the terminal sites can be mutated by glutamic acid residue whereas in the case of all beta transmembrane proteins, small residue like glycine can be substituted on behalf of any residue at the terminal site. The residues like tryptophan, isoleucine, tyrosine and leucine cannot be substituted at terminal sites of alpha transmembrane proteins. In the case of beta transmembrane proteins, residues such as isoleucine, tyrosine, methionine, proline, phenylalanine and valine cannot be substituted at carboxy terminals. Interestingly, it is observed that the residue glycine can be substituted at the carboxy terminal site instead of any residue in alpha and beta transmembrane proteins. The role of glycine in folded proteins is extensively studied by various groups [23, 24, 25, 26]. Similarly, sulphur containing amino acid residue, cysteine can be substituted at the terminal site of alpha and beta transmembrane proteins. The results imply that the carboxy terminal sites cannot be replaced by hydrophobic amino acids like isoleucine and valine.
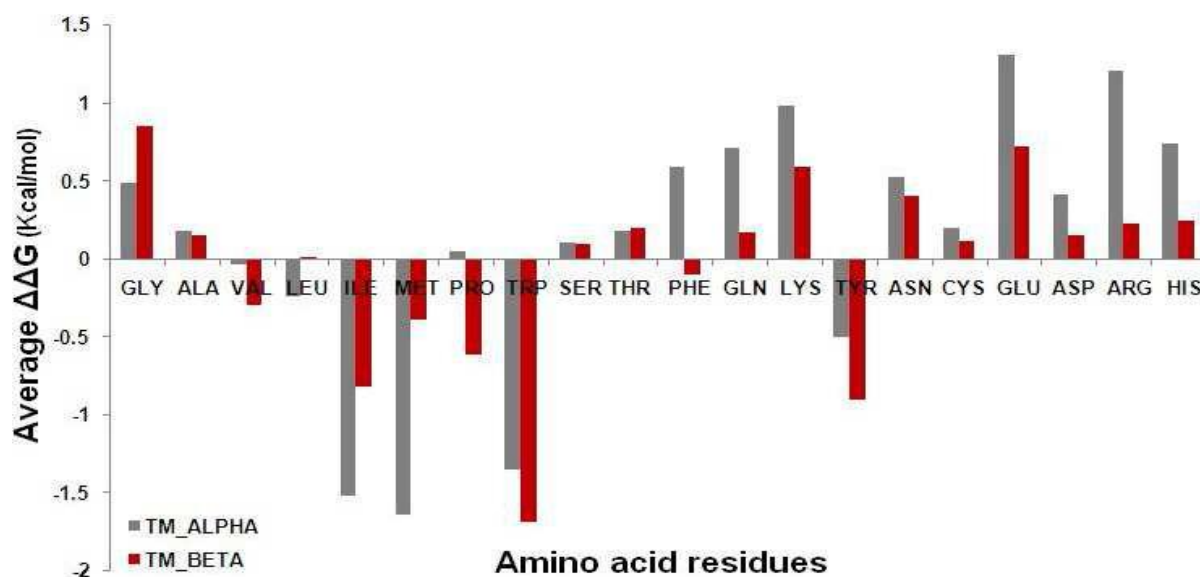
Figure 3. Average ΔΔG values for each amino acid residue at C-terminal position

The charged amino acid residues can be substituted at the carboxy terminal position without affecting the stability of the protein in alpha and beta transmembrane proteins. The substitution of residues such as threonine and serine at the carboxy terminal site also seems to be stabilizing in the alpha and beta transmembrane proteins. Accessible Surface Area is closely related to the molecular surface which plays a vital role in binding process [27]. The average solvent accessibility of carboxy terminal residues in alpha and beta transmembrane proteins is 87.12% and 67.64% respectively. The average solvent accessibility of terminal site residues of alpha transmembrane proteins is higher than the beta transmembrane proteins.

3.4 Carboxy terminal pentapeptide patterns

From several previous reports, it was proved that the identical pentapeptides at different positions in unrelated proteins can adopt different secondary structures [23, 28, 29, 30] . The five terminal residues (carboxy pentapeptides) are extracted from the dataset and these carboxy pentapeptides are subjected to pattern search against the dataset of proteins with regular secondary structures. The carboxy pentapeptides, hits obtained by pattern matching and their corresponding secondary structure, percentage of similarity between secondary structures of obtained hits are shown in table 1 (A and B). In other words, the above shows that the identical pentapeptides with different and similar secondary structure with respect to carboxy pentapeptide. From our observations, it is found that the identical pentapeptides adopt a different secondary structure. The transition of secondary structure is frequent from helix to coil and extended to coil and rare for helix to extended transition. Due to the positional context along sequence, carboxy pentapeptide and hits obtained do not possess a similar secondary structure. The terminal residues are mostly said to be unstructured in both alpha and beta transmembrane proteins, but it is found to have different secondary structure in globular proteins. While looking at the amino acid composition for C-terminal pentapeptides in alpha transmembrane proteins, leucine has occupied highest percentage of occurrences (11.40%). Lecuine is the most distributed amino acid among the proteins and so it has occupied highest percentage. But while considering only the terminal residue, alanine is found to occupy more. Arginine (10.20%) has a relatively higher

percentage of occurrence when compared with leucine. Cysteine and methionine have a minimum percentage of occurrences. In the case of beta transmembrane proteins phenylalanine is found to occupy highest percentage of occurrence in both cases [only terminal residue (57%) and terminal pentapeptide (18.40%)]. Cysteine is not found in any carboxy pentapeptides.

Carboxy pentapeptide PDB ID (Ref)	Hits obtained in 5877 sequences with regular secondary structures (Hit)	Terminal-carboxy penta-peptides	Ref Secondary Structure	Hit Secondary Structure	Secondary Structure Difference (%)
1KQF_A	4I17_A	NIEKA	EEEEC	CHHHH	100
1P7B_A	1VHS_A	DTEPV	CCCCC	CCCCC	0
1XIO_A	2WH7A,3F62B,3GRHA	RNLND	HHCCC	EECCE,ECCCC,CCCCC	60,40,40
2A06_B	3HMCA,1DC1A,1JBKA,1K92A,1SURA	FIDEL	CCCCC	HHHHH,HHHHH,EEECC, HHHHC,HHHHH	100, 100, 60, 80,100
2A65_A	3UPIA	ESAGT	HHHCC	ECCCH	80
2AXT_A	2QQ8A,3K6OA	PLDLA	CCCCC	CHHHH,HHHHH	80,100
2B76_A	1QSAA	TLPPA	CCCCC	CCHHH	60
2C3E_A	2PA8D,3U28A	LYDEI	HHCCC	CCEEE,CCCEE	100, 80
2F93_A	2ZGIA,1SA3A	TLRAE	HHHHC	EEEEE,HCCCC	100,80
2F95_A	2ZGIA,1SA3A	TLRAE	HHHHC	EEEEE,HCCCC	100, 80
2FBW_A	3IGFA	AIRSY	CCCCC	HHHHH	100
2FKW_A	2D73A	VKKAA	CCCCC	HHHHH	100
2GFP_A	4EIYA	CWLPL	CCCCC	HHHHH	100
2J7A_A	3Q9VA	QKRLV	CEECC	CCCEE	80
2K1K_A	3RD5A	RSRRA	HHCCC	CCCCC	40
2NR9_A	2RFAA,3PICA	LELVP	HCCCC	CCCCC,HHHCC	20, 40
2OAR_A	1XHNA	LAQTN	HHHCC	CCCCC	60
2RDD_A	1V74A	RFSRK	CCCCC	CCCHH	40
2YVX_A	1W1DA,2OO0A,3BC8A,3CJDA,3DSHA,3E2VA,3L6DA,3QXFA,3SL7A,1DUGA,1O1YA	RLLLE	HCCCC	HHHHH,HHHHH,HHHHH,HHHHH,HHHHH,CEEEC,HHHHH,HHHHH,HHHCC,HHHHH.HHHHH	80,80, 80, 80,80, 80, 80,80,80
2ZJS_Y	4AVAA	QLMLR	HHHHC	EEEEE	100
3B4R_A	3E8TA,3Q4OA	IFKNI	HHHCC	HHHHH,HHHHH	40, 40
3C02_A	3E96A,1I1QA,2KINA,3EL6A,3SJQC	LVDLA	HHHHC	HHHHC,HHHHH,EEECC, HHHHH,HHHHH	0,20,80,20,20
3DHW_A	3FFVA	QFAGD	HHHCC	EECCC	60
3DIN_A	2V76A,3C26A	YVLRV	HCCCC	EEEEE,EEEEE	100, 100

Table 1(A) Carboxy-terminal pentapeptide of alpha transmembrane proteins pattern search in 5877 sequences with regular secondary structure

While looking at the amino acid composition for C-terminal pentapeptides in alpha transmembrane proteins, leucine has occupied highest percentage of occurrences (11.40%). Lecuine is the most distributed amino acid among the proteins and so it has occupied highest percentage. But while considering only the terminal residue, alanine is found to occupy more. Arginine (10.20%) has a relatively higher percentage of occurrence when compared with leucine. Cysteine and methionine have a minimum percentage of occurrences. In the case of beta transmembrane proteins phenylalanine is found to occupy highest percentage of occurrence in both cases [only terminal residue (57%) and terminal pentapeptide (18.40%)]. Cysteine is not found in any carboxy pentapeptides.

Carboxy pentapeptide PDB ID (Ref)	Hits obtained in 5877 sequences with regular secondary structures (Hit)	Terminal-carboxy penta-peptides	Ref Secondary Structure	Hit Secondary Structure	Secondary Structure Difference (%)
1H6S_1	3E4W_A	VRFDF	EEEEC	EEEEE	20
1KMO_A	3BJE_A	GSLKF	EEEEC	HHCCE	100
1P4T_A	2NPT_B	VRVKF	EEEEC	EEEEE	20
1TLW_A	2P9W_A	NFHHH	EECCC	CCCCC	40
1YC9_A	3V7P_A,1BY2_A	ALGFE	HHCCC	HHCCC,HCCCC	0,20
2GUF_A	2X8T_A,4AHW_A, 1NQE_A(TERMINAL)	GSYTF	EEEEC	CCEEC,EEEEE,EEEEC	40,20,0
2HDF_A	4FNV_A	VDYRF	EEEEC	HHHHH	100
3KVN_A	3D1B_A	LSLDF	EEEEC	CCHHH	100
3RFZ_A	1ZE3_H(TERMINAL)	TFVYQ	EEEEC	EEECC	20
3SY9_A	3OCM_A	PIDVF	EEEEC	HHHHH	100
3V89_A	1D8H_A	LEMKF	EEEEC	EEEEE	20
4AIQ_A	2GAG_A	VNYKF	EEEEC	CCCCC	80
4C00_A	3NUI_A	LGPEL	CCCCC	ECHHH	80
4E1T_A	3HLZ_A	VLEYQ	CEEEC	EEEEC	20
4FRX_A	1XKP_B,2YJG_A,3GUZ_A	TLPLL	EECCC	CHHHH,HHHHH,CHHHH	100 ,100,100
4K3B_A	1ZUH_A,2CL3_A	LGTTF	EECCC	CEEEE,EECCE	60,20

Table 1(B) Carboxy-terminal pentapeptide of beta transmembrane proteins pattern search in 5877 sequences with regular secondary structure

3.5 Hydrogen bonding patterns of carboxy terminal residues

Proteins are made up of hundreds of amino acids folded to a well-defined structure that is stabilized by various types of interactions. Discretization of the 3D structure of proteins can be done in many ways. The formation of hydrogen bonds between amino acid

residues in a protein is one such way to understand protein folding and structure. In the present work, the hydrogen bonding of terminal residues with other residues in proteins is analyzed in detail. The terminal residue (donor) and other residue (acceptor) forming hydrogen bond distance and details of acceptor and donor is presented in Supplementary Material S2. In the case of alpha transmembrane proteins, the terminal residue of 20 proteins does not form hydrogen bond whereas, in the case of beta transmembrane proteins, terminal residues of ten proteins do not form hydrogen bonds. In the case of alpha and beta transmembrane proteins, the hydrogen bonding distance between terminal residue and other residue is approximately 3Å. 42% of terminal residues of transmembrane proteins form a long-range hydrogen bond with other residues. Long-range hydrogen bonds are nothing but the residues forming hydrogen bonds that are far in sequence and close in three dimensional structure (distance between two residues along the sequence greater than 4) [31, 32]. From our analysis, it is found there is no clear discrimination or difference in hydrogen bonding patterns of carboxy terminal residues between alpha and beta transmembrane proteins.

4. CONCLUSION

Terminal residues of the proteins help in binding to other proteins and for their correct folding and assembly. Although studies have been done on terminal residues of globular proteins, not much of the analysis has been done on terminal residues of transmembrane proteins. These proteins were found to be involved in diverse biological processes. Since, the N-terminal region of the proteins pertain the signal peptides, the significance of the C-terminal residues were subjected to computational analysis. These residues were subjected to mutation analysis to elucidate the favorable and unfavorable residues along their individual positions and it was found that residues like, tryptophan, isoleucine, tyrosine and leucine cannot be substituted at terminal sites of alpha transmembrane proteins and in the case of beta transmembrane proteins, residues such as isoleucine, tyrosine, methionine, proline, phenylalanine and valine cannot be substituted at carboxy terminals respectively. The terminal carboxy pentapeptide patterns of transmembrane proteins were matched with the globular protein sequences and the identical pentapeptides adopting different secondary structures were identified. The frequency of transition was more from helical to coil conformation and from extended to coil conformation. The solvent accessibility and hydrogen bonding patterns were studied and it was inferred that the average hydrogen bonding distance was found to be 3Å and about 42% of the terminal residues of transmembrane proteins form long-range hydrogen bonds with other residues. The results obtained in the present study will provide clues to understand protein folding and binding patterns in transmembrane proteins.

SUPPLEMENTARY FILES

1. Gene ontology terms of alpha and beta transmembrane proteins
2. Hydrogen bonding pattern of carboxyl terminal in alpha and beta transmembrane proteins.

REFERENCES

[1]   Thornton JM, Sibanda BL., Amino and carboxy-terminal regions in globular proteins, J Mol Biol, 1983; 167: 443-60.
[2]   Takano K, Tsuchimori K, Yamagata , Yutani K., Effect of foreign N-terminal residues on the conformational stability of human lysozyme, Eur J Biochem, 1999; 266: 675-82.
[3]   Kaufmann E, Klaus W, Norbert G., Intermediate filament forming ability of desmin derivatives lacking either the amino-terminal 67 or the carboxy-terminal 27 residues, J Mol Biol, 1985; 185: 733-42.
[4]   Krishna MMG, Englander SW., The N-terminal to C-terminal motif in protein folding and function, Proc Natl Acad Sci USA, 2005; 102: 1053-8.
[5]   Chung JJ, Shikano S, Hanyu Y, Li M., Functional diversity of protein C-termini: more than zipcoding? Trends Cell Biol, 2002; 12: 146-50.
[6]   Fitzpatrick P, Shattil SJ, Ablooglu AJ., Carboxy-terminal-COOH of integrin β1 is necessary for β1 association with the kindlin-2 adapter protein, J Biol Chem, 2014; 289: 11183-93.
[7]   Curtis FA, Reed P, Wilson LA, Bowers LY, Yeo RP, Sanderson JM, Walmsley AR, Sharples GJ., The C-terminus of the phage λ Orf recombinase is involved in DNA binding, J Mol Recog, 2010; 24: 333-340.
[8]   Wang HG, Rikitake Y, Carter MC, Yaciuk P, Abraham SE, Zerler B, Moran E., Identification of specific adenovirus E1A N-terminal residues critical to the binding of cellular proteins and to the control of cell growth, J Virol, 1993; 67: 476-88.
[9]   Chen J, Skehel JJ, Wiley DC., N-and C-terminal residues combine in the fusion-pH influenza hemagglutinin HA2 subunit to form an N cap that terminates the triple-stranded coiled coil, Proc Natl Acad Sci USA, 1999; 96: 8967-72.
[10]   Rost B, Sander C., Conservation and prediction of solvent accessibility in protein families. Proteins, 1994; 20: 216-26.
[11]   Friedman R, Nachliel E, Gutman M., Molecular dynamics of a protein surface: ion-residues interactions, Biophys J, 2005; 89: 768-81.
[12]   Pal D, Chakrabarti P., Terminal residues in protein chains: residue preference, conformation, and interaction, Biopolymers, 2000; 53: 467-475.
[13]   Rajasri B, Pal D, Chakrabarti P., Secondary structures at polypeptide-chain termini and their features, Acta Crystallogr D Biol Crystallogr, 2002; 58: 1793-802.
[14]   Austin RS, Provart NJ, Cutler SR., (2007) C-terminal motif prediction in eukaryotic proteomes using comparative genomics and statistical over-representation across protein families, BMC genomics, 2007; 8: 191.
[15]   Javier S, Serrano L, Fersht AR., Histidine residues at the N-and C-termini of. alpha.-helixes: perturbed pKas and protein stability, Biochemistry, 1992; 31: 2253-58.
[16]   Takahashi E, Kobayashi H, Yamanaka H, Nakanishi M, Tateishi A, Abet T, Arimoto S, Negishi T, Okamoto K., Analysis of carboxy terminal domain of metalloprotease of elastolytic Aeromonas hydrophila, Biol Pharma Bull, 2013; 36: 1174-82.
[17]   Nelson TK, Sorgen PL, Burt JM., Carboxy terminus and pore-forming domain properties specific to Cx37 are necessary for Cx37-mediated suppression of insulinoma cell proliferation, Am J Physiol-Cell, 2013; 305: C1246-C56.
[18]   Kozma D, Simon I, Tusnády GE., PDBTM: Protein Data Bank of transmembrane proteins after 8 years, Nucleic Acids Res, 2013; 41: D524-D9.
[19]   Rose PW, Bi C, Bluhm WF, Christie CH, Dimitropoulos D, et al., The RCSB Protein Data Bank: new resources for research and education, Nucleic Acids Res, 2013; 41: D475-D82.
[20]   Singh H, Chauhan JS, Gromiha MM, Consortium OSDD, Raghava GPS., ccPDB: compilation and creation of data sets from Protein Data Bank, Nucleic Acids Res, 2012; 40: D486-9.
[21]   Parthiban V, Gromiha MM, Schomburg D., CUPSAT: Prediction of protein stability upon point mutations, Nucleic Acids Res, 2006; 34, W239-42.
[22]   Cock H, Struyve M, Kleerebezem M, Krift TV, Tommassen J., Role of the Carboxy-terminal Phenylalanine in the Biogenesis of Outer Membrane Protein PhoE of Escherichia coli K-12, J Mol Biol, 1997; 269: 473-478.
[23]   Saravanan KM, Selvaraj S., Search for identical octapeptides in unrelated proteins: Structural plasticity revisited, Biopolymers, 2012; 98: 11-26.
[24]   Saravanan KM, Krishnaswamy S., Analysis of dihedral angle preferences for alanine and glycine residues in alpha and beta transmembrane regions, J Biomol Struct Dyn, 2014; DOI: 10.1080/07391102.2014.895678
[25]   Javadpour MM, Eilers M, Groesbeek M, Smith SO., Helix packing in polytopic membrane proteins: role of glycine in transmembrane helix association, Biophysical J, 1999; 77: 1609- 18.
[26]   Ho BK, Brasseur R., The Ramachandran plots of glycine and pre-proline, BMC Struct Biol, 2005; 5: 14.
[27]   Adamczak R, Porollo A, Meller J., Combining prediction of secondary structure and solvent accessibility in proteins, Proteins, 2005; 59: 467-75.
[28]   Saravanan KM, Balasubramanian H, Nallusamy S, Samuel S., Sequence and structural analysis of two designed proteins with 88% identity adopting different folds, Protein Eng Des Sel, 2010; 23: 911-918.

[29]   Saravanan, KM, Selvaraj S., Performance of secondary structure prediction methods on proteins containing conformationally ambivalent sequence fragments, Biopolymers, 2012; 100:148-53

[30]   Saravanan KM, Selvaraj S., Search and Analysis of Identical Reverse Octapeptides in Unrelated Proteins, Genomics, Proteomics Bioinformatics, 2013; 11: 114-21.

[31]   Gromiha MM., Influence of cation–π interactions in different folding types of membrane proteins, Biophy Chem, 2003; 103: 251-258.

[32]   Saravanan KM, Selvaraj S., Analysis and Visualization of Long-Range Contacts and Networks in Homologous Families of Proteins, Open Struct Biol J, 2009; 3: 104-125.