ISSN 0975-6299 Vol 2/Issue 1/ Jan-Mar 2011

International Journal of Pharma and Bio Sciences

RESEARCH ARTICLE

BIOINFORMATICS

HUNTINGTIN PROTEIN MODELING AND STRUCTURE ALIGNMENT STUDIES

Corresponding Author

A. KRISHNA CHAITANYA

Department of Biochemistry & Bioinformatics, GIS, GITAM University, Visakhapatnam-45, India

Co Authors

I. BHASKAR REDDY, DSVGK. KALADHAR AND GNV. SANTHOSH

Department of Biochemistry & Bioinformatics, GIS, GITAM University, Visakhapatnam-45, India

ABSTRACT

Huntington's disease results from genetically programmed degeneration of nerve cells, called neurons, in certain areas of the brain. This degeneration causes uncontrolled movements, loss of intellectual faculties, and emotional disturbance. The disease is caused due to the mutant huntingtin protein for which the function is unclear. This protein interrupts the other interacting proteins activity resulting in abnormal functioning of the nerve cells. To deduce the exact function of huntingtin protein, an attempt is made to determine the structure of Htt protein which is very large in length, through automated modelling server and phylogeny analysis as well structure alignment studies were performed using CLC Workbench & MVD software respectively for inferring the conservativity of the Htt protein as well its probable function. From this study, the CREB-binding protein & Multidrug transporter protein are found to be much similar in structural features & functional aspects of the Htt protein.

ISSN 0975-6299 Vol 2/Issue 1/ Jan-Mar 2011

KEYWORDS

Huntington, Htt Protein, Structural Bioinformatics, Structural Alignment.

INTRODUCTION

Structural Bioinformatics is related to the analysis and prediction of the three-dimensional structure of biological macromolecules. It deals with generalizations about macromolecular 3D structure such as comparisons of overall folds and local motifs, principles of molecular folding, evolution, and binding interactions, and structure/function relationships, working both from experimentally solved structures and from computational models.

The Huntingtin gene, also called HTT or HD (Huntington disease) gene, or the IT15 ("interesting transcript 15") gene codes for a protein called the huntingtin protein. The genetic defect responsible for Huntington's disease is a small sequence of DNA on chromosome 4 in which several base pairs are repeated many, many times. The normal gene has three DNA bases, composed of the sequence CAG. In people with Huntington's disease, the sequence abnormally repeats itself dozens of times.

The individuals having normal range of CAG repeats is ≤ 28 and ≥ 40 repeats of CAG repeats are at higher risk. As the length of the protein sequence is very larger (~3142 aminoacids), the structure determination is difficult. As well the function of the Htt protein is unclear.

The structure prediction can be done using modeling servers and the conservativity and similarity studies can be done through the structural alignment studies which are comparatively better to assign the function to the protein.

MATERIALS AND METHODS

(i) Retrieval of the Huntingtin & its Interacting Protein Sequences

Protein sequence of Huntingtin has been retrieved from UniProtKB/Swiss-Prot database. (http://expasy.org/sprot/) Details of the protein sequence: Accession No. **P42858** (HD_HUMAN). Over 100 interacting proteins have been found, in which only some of the sequences retrieved which are available in UniProtKB/Swiss-Prot database.

(ii) Huntingtin & its Interacting Proteins Phylogenetic Signal Analysis

The Huntingtin protein has no sequence homology with other proteins as well function is also unclear. Hence an attempt is made to find a useful phylogenetic signal, a term that is used generally to denote whether interacting proteins tend to resemble each other and with respect to Huntingtin protein. Alignment of sequences was done using the CLC Free Workbench.

(iii) Modeling of Huntingtin Protein

Normal huntingtin is generally to be ~3142 amino acids in size. As the protein sequence is very large, it is difficult to model or predict the structure of the protein using automated protein modelling servers. Hence a novel approach is followed to model the protein by disuniting the sequence in to fragments of length of 300 amino acid residues respectively.

Each 300-length fragmentised protein sequence is submitted to the ModWeb server for comparative protein structure modeling.

(iv) Structure Alignment Studies

The 3D structure of proteins are more conserved than the sequence. Therefore, structure-based function prediction approaches are considered to be more reliable, as function can be assigned to proteins. Hence, an array of methods such as structure alignment studies can be used to assign/predict the molecular function of proteins. The structure alignment was carried out using Molegro Virtual Docker (MVD). The academic version of MVD was used for the structure alignment studies.

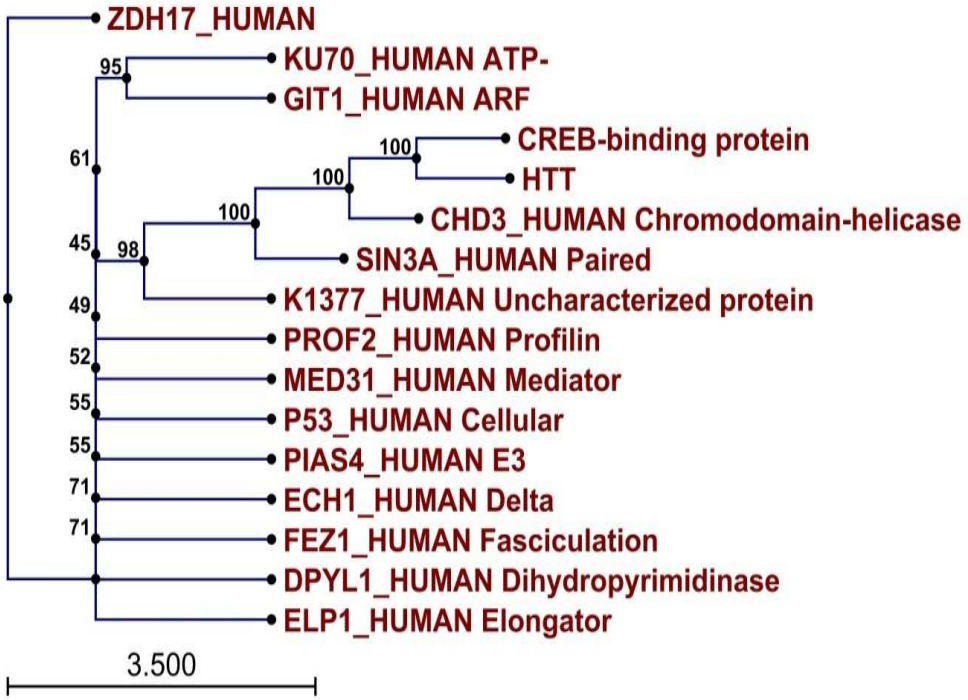
RESULTS & DISCUSSION

(i) Phylogeny of Huntingtin Interacting Proteins with Huntingtin Protein

Fifteen of the Huntingtin Interacting Proteins of Homosapiens along with the Huntingtin protein have been carried out through Phylogenetic analysis using Neighbour Joining (NJ) method which infers that GIT 1-Human (GTPase-activating protein) is evolutionarily divergent to the other interacting proteins. This analysis also infers about the convergence of the Huntington protein with the CREB-Binding protein (cAMP response element-binding). CREB has a well-documented role in neuronal plasticity and long-term memory formation in the brain.

Figure 1

Phylogeny of Huntingtin Interacting Proteins along with Htt Protein using NJ Method



(ii) Htt Protein Modeling Studies

Each 300-length fragmentized sequences are generated manually, and submitted to ModWeb server from which only seven segments are modeled. ModWeb builds the structure for the given sequence on the basis of the available structures in the Protein Data Bank (PDB). Hence the protein in the PDB can be easily identified from which we can easily derive its information as well its function. The details of the Htt protein fragmentized segments are shown in the following table 1.

Table 1

Htt Protein (Target) Fragmentized Segments & Modeled Regions of Htt with Respective (Template) Structures from PDB and Sequence Identity

S. No.	Modeled Fragments	Region of the Submitted Htt Sequence (Target)	Region of the Htt Protein Modeled	Total No. of Residues Modeled	PDB ID (Template)	Sequence Identity of the Alignment %
1	1.a.	1-300	122-269	147	1B3U(A)	26
2	1.b.		41-100	59	1NAY(A)	49
3	1.c.		1-30	30	3IOR(C)	100
4	1.d.		1-56	56	3IOR(B)	100
5	1.e.		139-281	142	1M5N(S)	26
6	2.	301-600	491-585	94	2ENY(A)	38
7	3.	601-900	802-894	92	2IBI(A)	28
8	4.a.	1801-2100	2011-2077	66	2DNO(A)	38
9	4.b.		1917-2052	135	1SZ9(A)	24
10	5.	2101-2400	2136-2288	152	2Q1Z(B)	28
11	6.	2401-2700	2773-2858	85	2DHH(A)	34
12	7.	2701-3000	2869-2977	108	2AXJ(A)	39

(iii) Htt Protein Structure Alignment Studies

The Modeled PDB Structures are structurally aligned with their respective Template Structures obtained from PDB. Structure alignment is carried out with MVD software by which Root Mean Square Deviation (RMSD) is taken into consideration. The lower the RMSD values, the higher the structural similarity.

From the structure alignment studies, we can infer that the Htt modeled protein at different segments may have distinct functionalities (Protein Classification) and is highly structurally similar to multidrug transporter protein because of less RMSD value among other protein structures.

Table 2

Htt Protein Modeling with Respective (Template) Structures from PDB and RMSD Values

S. No.	Region of the Submitted Htt Sequence (Target)	PDB ID (Template)	Protein Name	Protein Classification	RMSD Value
1	1-300	1B3U(A)	Protein Phosphatase PP2A	Scaffold	5.21185
2		1NAY(A)	Designed (GlyProPro)10foldon	Structural	--
3		3IOR(C)	Huntingtin Amino-Terminal Region With 17 Gln Residues	Signaling	--
4		3IOR(B)	Huntingtin Amino-Terminal Region With 17 Gln Residues	Signaling	4.76081
5		1M5N(S)	Heat Repeats (1-11) of Importin b	Protein Transport	7.59196
6	301-600	2ENY(A)	Ig-Like Domain of Human Obscurin	Contractile	--
7	601-900	2IBI(A)	Ubiquitin-carboxyl-terminal hydrolase 2	Hydrolase	--
8	1801-2100	2DNO(A)	RNA Binding Domain in Trinucleotide Repeat Containing 4 Variant	RNA Binding	--
9		1SZ9(A)	The Rna Polymerase Ii Ctd In Mrna Processing	Transcription	--
10	2101-2400	2Q1Z(B)	SigE in complex with the anti-sigma ChrR	Transcription	13.4831
11	2401-2700	2DHH(A)	Multidrug Transporter	Membrane	3.2001
12	2701-3000	2AXJ(A)	T Cell Receptor Beta Chains related to Rheumatoid Arithritis	Immune System	6.75036

CONCLUSION

Huntingtin protein is potential target for treating the Huntington disease. In order to understand structural features, automated modeling tools are used and by performing the Phylogeny studies, CREB - Binding Protein and Htt are found to be more conserved among the other interacting proteins. By the Huntingtin Structure Prediction & Modeling Studies, it is identified that the Huntingtin Protein shares the major structural similarity with the Multidrug Transporter, Protein Phosphatase PP2A, T-cell receptor β-Chain with RMSD Values of 3.2, 5.2, and 6.7 respectively. Hence, there is a chance of featuring the similar kind of role and biological function. This study also provides new in-depth inferring the structural aspects as well new approaches to model a protein of larger size.

REFERENCES

1. Bonilla E, Huntington disease: A review. Invest Clin, 41(2):117-41, (2000)
2. Roze E, Bonnet C, Betuing S, Caboche J., Huntington's disease. Adv Exp Med Biol., 685:45-63, (2010)
3. Borrell-Pagès M, Zala D, Humbert S, Saudou F., Huntington's disease: from huntingtin function and dysfunction to therapeutic strategies. Cell Mol Life Sci., 63(22):2642-60, (2006)
4. Humbert S, Saudou F., Huntington's disease: intracellular signaling pathways and neuronal death. J Soc Biol.;199(3):247-51, (2005)
5. Walling HW, Baldassare JJ, Westfall TC. Molecular aspects of Huntington's disease. J Neurosci Res. Nov:1;54(3):301-8, (1998)
6. Li SH, Li XJ., Huntingtin-protein interactions and the pathogenesis of Huntington's disease. Trends Genet, Mar;20(3):146-54, (2004)
7. Truant R, Atwal R, Burtnik A., Hypothesis: Huntingtin may function in membrane association and vesicular trafficking., Biochem Cell Biol.;84(6):912-7. (2006)
8. Borrell-Pagès M, Zala D, Humbert S, Saudou F., Huntington's disease: from huntingtin function and dysfunction to therapeutic strategies. Cell Mol Life Sci. Nov;63(22):2642-60. (2006)
9. http://expasy.org/sprot/
10. http://www.clcbio.com/index.php?id=28 http://modbase.compbio.ucsf.edu/ModWeb20-html/modweb.html