ISSN: 0975-8585

Research Journal of Pharmaceutical, Biological and Chemical Sciences

Computational 3D Modelling and *In Silico* Characterization of Human Interferon Induced Transmembrane Proteins.

Pialy Chatterjee[1], Lalit Samant[2*], Sandeepan Mukherjee[1], Sanket Bapat[1], and Abhay Chowdhary[1,2].

Department of Virology and Immunology[1], Systems Biomedicine Division[2], Haffkine Institute for Training, Research and Testing, Acharya Donde Marg, Parel, Mumbai -400012, India.

ABSTRACT

The interferon system protects the cell from viral infection through the induction of a group of genes collectively known as the Interferon Stimulated Genes (ISGs). In recent times, the Interferon Induced Transmembrane (IFITM) genes which belong to the family of small ISGs have gained importance owing to their anti-viral & anti-proliferative properties. In this study, we have developed the three dimensional structure of two of the protein products of these genes namely IFITM1 & IFITM3 using homology modeling approach. Both IFITM1 and IFITM3 sequences were characterized *in silico* using various bioinformatics tools such as the Expasy ProtParam, Cysteine Recognition server, CDD, SOPMA and TMHMM. Secondary structure analysis showed 44% and 42.11% alpha helix content in IFITM1 and IFITM3 respectively making it stable for modeling. Homology modeling of the IFITM1 AND IFITM3 protein structure was done using I-TASSER workspace and validated using RAMPAGE and ArgusLab. Ramachandran plot analysis of IFITM1 and IFITM3 proteins showed that conformations for 85.4 % and 84.7% of amino acid residues respectively are within the most favored regions and the C-score, expected TM score and RMSD results validated the final refined models as reliable. Energy minimization for IFITM1 and IFITM3 proteins was done by using ArgusLab and the results suggested the final model to be steadfast. The structures thus modeled may enable conceptualizing prophylactic & therapeutic strategies against many viral diseases.

Keywords: Interferon; IFITM 1, IFITM 3, CD225, Homology Modeling, In Silico, Transmembrane proteins.

Corresponding author

ISSN: 0975-8585

INTRODUCTION

The IFITM gene family was first identified in a cDNA screen from IFN-treated neuroblastoma cells back in 1984 [1]. The IFITM genes IFITM1, IFITM2 and IFITM3 (IFITM1-3) belong to the family of small ISGs. Its members are induced in response to viral infections [2]. Human IFITM 1, IFITM 2, IFITM 3 and IFITM 5 genes lie adjacently on chromosome 11 whereas mouse Ifitm1, Ifitm2, Ifitm3 and Ifitm5 genes are located on chromosome 7 and are orthologues of their human counterparts. The chromosomal arrangement of these genes is not conserved between human, mouse and rat both in length and in number of genes present in the chromosomal locus [3].

The IFITM proteins comprise a family of viral restriction factors that play critical roles in the interferon-mediated control of Influenza A virus (IAV). A recent study indicates that the IFITM3 protein is required both for basal levels of resistance, as well as for the heightened defenses elicited by IFN γ and α which are known to strongly decrease basal levels of IAV infection [4]. An unique attribute of these antiviral proteins is that they interfere with a step in viral replication preceding fusion of the viral & cellular membranes [5,6]. Several implications of this early restriction step have been indicated in a recent study. First, IFITM-mediated restriction precedes the induction of type I IFNs in infected cells thus providing basal resistance which is then enhanced by amplification of IFITM protein expression on IFN induction thus protecting uninfected cells in a paracrine manner. Second, viral escape from restriction by IFITM proteins could be more challenging than escape from inhibitory factors that function at later stages of the viral replication cycle [7].

In addition to influenza A virus, IFITM proteins restrict infection by several other enveloped viruses. These include retroviruses (HIV-1), flaviviruses (dengue virus and WNV)(4), filoviruses (Marburg virus and Ebola virus) and coronaviruses (such as severe acute respiratory syndrome (SARS) coronavirus)(5). Viruses that are restricted by IFITM proteins tend to fuse with host cell membranes in a late endosome or lysosome. Experiments using retroviruses coated with SARS virus entry protein have established that the site of viral fusion is crucial for the antiviral activity of IFITM proteins [5].

There seems to be specialization among the antiviral functions of IFITM proteins. Cell-culture experiments indicate that the IFITM3 protein is especially effective in controlling influenza A virus whereas the IFITM1 protein restricts filoviruses and SARS coronavirus more effectively than the IFITM3 protein does [5]. Pathogen infections have shown to result in long-lasting upregulation of IFITM proteins, suggesting a function of these proteins in host defense. This finding indicates towards the ability of these proteins to function as good biomarkers for past or chronic infections [8].

Taking into consideration the antiviral potential of the IFITM proteins, the present study aims at developing fully modeled structures of IFITM 1 and IFITM 3 proteins *in silico*. The fully developed structures shall help elucidate the functional mechanisms of these proteins. Information thus attained may enable docking studies thus paving new ways for further novel prophylactic or therapeutic strategies for a myriad of dreaded viral diseases.

ISSN: 0975-8585

METHODS

Sequences

The query sequence for IFITM1 (NP_003632.3) and IFITM3 (AFF60355.1) were retrieved from NCBI database.

Physiochemical and functional characterization

The physio-chemical properties of the query sequences were computed using Expasy's ProtParam server while the Cysteine Recognition server [9] yielded the count, position and pattern of all the cysteine residues in both the protein sequences. The conserved regions were detected using Conserved Domain Database (CDD) [10] tool to establish functional domains and regular expression pattern. Finger print analysis [11] was also done to build diagnostic signatures of the protein family membership and the fingerprints thus created were used to identify the distant relatives of the protein in PRINTS database. The secondary structures for the query sequences were predicted using SOPMA [12,13] based on the steric properties of amino acids and the numbers of proline and glycine residues as they exhibit reduced and complete torsional freedom respectively. Prediction of transmembrane regions and orientation of the glycoprotein query sequence was done using TMHMM [14,15]. BlastP [16] was used in the present study to find homologous sequences to the query sequences and were arranged in the ascending values of E.

Comparative modeling

The three dimensional structures of IFITM1 and IFITM3 proteins was modeled using I-TASSER model workspace [17-19]. The theoretical structures of the IFITM1 & IFITM3 proteins were generated using 4iaqA, 3h6pB, 4bwzA, 1h25B, 3vr8D, 2hr2B, 3vr9D, 3wguA, 1rsiD, 3ukmA templates and 4bwsA, 3iynQ, 4bw5A, 3iynq, 2koga, 3vvnA, 3iynQ, 3ayfA, 1ddbA, 3bw5A respectively as templates. The best among the resultant modeled structures was selected depending on confidence score. The model was validated by ArgusLab and RAMPAGE [20,21] by verifying the parameters like Ramachandran plot quality, peptide bond planarity, main chain hydrogen bond energy, alpha chirality and over-all G factor and the side chain parameters like standard deviations of chi gauche minus, trans and plus pooled standard deviations of chi with respect to the refined structure [20,22].

RESULTS AND DISCUSSION

Physiochemical and functional characterization

The IFITM1 and IFITM3 protein sequences (Table 1) having accession number NP_003632.3 and AFF60355.1 respectively were retrieved from NCBI database. The query sequences were subjected to compositional analysis using various bioinformatics tools. The physiochemical parameters viz., theoretical isoelectric point (pI), molecular weight, total number of positive and negative residues, extinction coefficient, half-life, instability index, aliphatic index (AI) and grand average hydropathy (GRAVY) were computed using the Expasy's ProtParam tool (Table 2).

ISSN: 0975-8585

	Query Sequence
IFITM1	>gi|150010589|ref|NP_003632.3| interferon-induced transmembrane protein 1 [Homo sapiens] MHKEEHEVAVLGAPPSTILPRSTVINIHSETSVPDHVVWSLFNTLFLNWCCLGFIAFAYSVKSRDRKMVG DVTGAQAYASTAKCLNIWALILGILMTIGFILLLVFGSVTVYHIMLQIIQEKRGY
IFITM3	>gi|381148134|gb|AFF60355.1| IFITM3 [Homo sapiens] MNHTVQTFFSPVNSGQPPNYEMLKEEHEVAVLGAPHNPAPPTSTVIHIRSETSVPDHVVWSLFNTLFMNP CCLGFIAFAYSVKSRDRKMVGDVTGAQAYASTAKCLNIWALILGILMTILLIVIPVLIFQAYG

Table 1: Query sequences of IFITM1 and IFITM3

	IFITM1	**IFITM3**
Length (amino acids)	125	133
Molecular weight	13938.4	14632.1
Theoretical pI	7.78	6.49
Negatively charged residues (-R)	8	8
Positively charged residues (+R)	9	7
Abs 0.1% (=1 g/l)	1.620/1.611	1.168/1.159
The N-terminal of the sequence	M (Met)	M (Met)
Estimated half-life (hours)	30 (mammalian reticulocytes *in vitro*)	30 (mammalian reticulocytes *in vitro*)
Extinction Coefficient ($M^{-1}cm^{-1}$)	22585	**17085**
Instability Index (II)	54.03	44.63
Aliphatic Index (AI)	116.16	101.13
GRAVY	0.555	0.381

Table 2: Expasy ProtPram result of IFITM1 and IFITM3 amino acid sequence

Both IFITM1 and IFITM3 proteins were found to possess acidic and basic amino acid residues with Instability Indices (II) of 54.03 and 44.63 respectively, and were therefore classified to be biologically non stable proteins since II is above 40. The aliphatic index (AI), a positive factor for the increase of thermal stability of globular proteins was found to be high indicating greater amount of aliphatic to aromatic residues. Extinction coefficient values at 280 nm indicate the presence of high concentration of tyrosine residues in both IFITM1 and IFITM3 proteins. The high Grand Average hydropathy (GRAVY) indices of both proteins indicate they cannot interact well with water. This enables us to conclude that both the proteins are hydrophobic in nature. Detailed amino acid composition of IFITM1 and IFITM3 proteins are shown in Table 3.

ISSN: 0975-8585

Name of a.a	IFITM1		IFITM3	
	Number of a.a	Percentage of a.a	Number of a.a	Percentage of a.a
Ala (A)	9	7.2%	11	8.3%
Arg I	4	3.2%	3	2.3%
Asn (N)	4	3.2%	7	5.3%
Asp (D)	3	2.4%	3	2.3%
Cys I	3	2.4%	3	2.3%
Gln (Q)	3	2.4%	4	3.0%
Glu (E)	5	4.0%	5	3.8%
Gly (G)	8	6.4%	7	5.3%
His (H)	5	4.0%	5	3.8%
Ile (I)	12	9.6%	10	7.5%
Leu (L)	14	11.2%	12	9.0%
Lys (K)	5	4.0%	4	3.0%
Met (M)	4	3.2%	5	3.8%
Phe (F)	6	4.8%	7	5.3%
Pro (P)	4	3.2%	10	7.5%
Ser (S)	9	7.2%	9	6.8%
Thr (T)	8	6.4%	9	6.8%
Trp (W)	3	2.4%	2	1.5%
Tyr (Y)	4	3.2%	4	3.0%
Val (V)	12	9.6%	13	9.8%
Pyl (O)	0	0.0%	0	0.0%
Sec (U)	0	0.0%	0	0.0%

Table 3: Amino acid composition of IFITM1 and IFITM3 protein

Cysteine Recognition online tool showed the presence of 3 cysteine residues each in IFITM1 and IFITM3 proteins respectively with complete absence of disulphide bonds The probable patterns of the pairs are mentioned in Table 4.

Query Protein	No of Cys residues	Position	Scores
IFITM1	3	50 51 84	-22.9 -26.9 -34.9
IFITM3	3	71 72 105	-18.3 -23.9 -21.6

Table 4: Results of Cysteine recognition online tool for IFITM1 and IFITM3 protein

In the present study secondary structures of the query proteins were predicted using SOPMA analysis. IFITM1 protein exhibited high propensity for alpha helix and random coil conformation (44%, 30.40% respectively) (Table 6). However, the amino acids like tryptophan and methionine which have propensity for a helix were present in low quantities (2.4%, and 3.2% respectively). IFITM3 protein also showed high propensity for alpha helix and random coil conformation (42.11%, 38.35% respectively).

ISSN: 0975-8585

Characteristics	IFITM1		IFITM3	
	Number of a.a	% distribution	Number of a.a	% distribution
Alpha helix (Hh)	55	44	56	42.11
3_{10} helix (Gg)	0	0.00	0	0.00
Pi helix (Ii)	0	0.00	0	0.00
Beta bridge (Bb)	0	0.00	0	0.00
Extended strand (Ee)	26	20.80	23	17.29
Beta turn (Tt)	6	4.80	3	2.26
Bend region (Ss)	0	0.00	0	0.00
Random coil (Cc)	38	30.40	51	38.35
Ambigous states (?)	0	0.00	0	0.00
Other states	0	0.00	0	0.00

Table 5: Secondary structure analysis of IFITM1 protein and IFITM3 protein

Transmembrane Hidden Markov Model (TMHMM) tool is used to analyze the number of transmembrane domain in a given protein. Both IFITM1 & IFITM3 proteins show 2 transmembrane domains. The Exp number which stands for the value of the first 60 amino acids, if greater than 10 indicates the protein to be a signal peptide. The Exp number for IFITM1 is 22.49583 indicating that it is a signal peptide protein whereas IFITM3 with Exp number value 2.29236 is not a signal peptide.

Table 6: Transmembrane helix analysis of IFITM1 protein and IFITM3 protein

Parameters	IFITM1	IFITM3
Length	125	133
Number of predicted TMHs	2	2
Exp number of AAs in TMHs	46.86112	45.39718
Exp number, first 60 AAs	22.49583	2.29236
Total probability of N terminal trans membrane protein	0.12274	0.05272

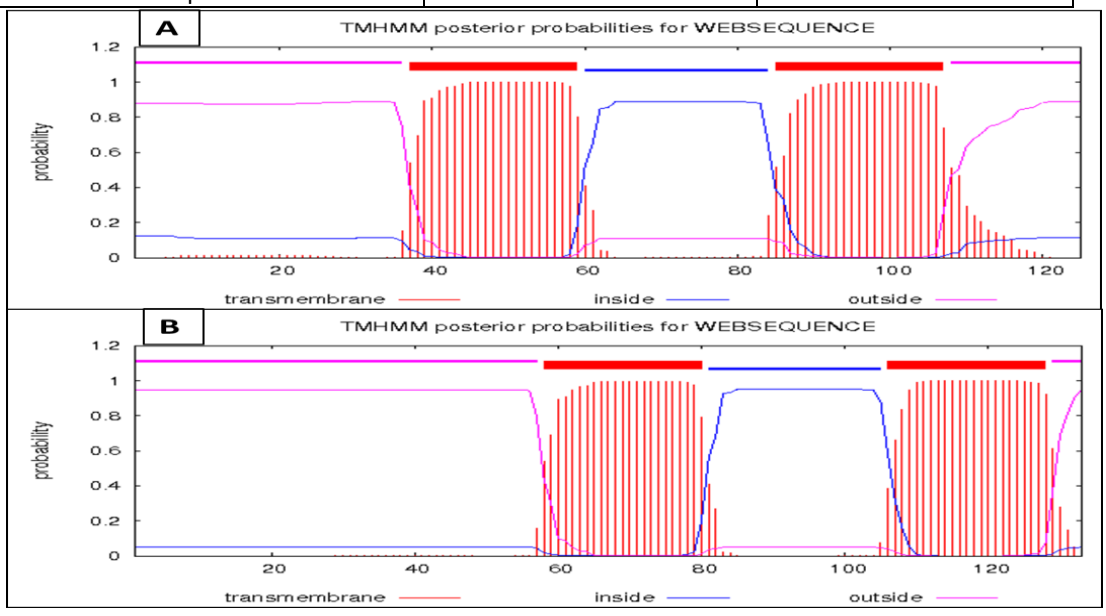


Figure 1: Screenshot of Transmembrane helix analysis of (A) IFITM1 protein sequence; (B) IFITM3 protein sequence

ISSN: 0975-8585

Conserved Domain Database analysis indicates that both the proteins belong to the human leukocyte antigen CD225 superfamily, which is an interferon inducible transmembrane protein family and is associated with interferon cell growth suppression. Possible conserved domains for the proteins were also observed by CDD analysis.

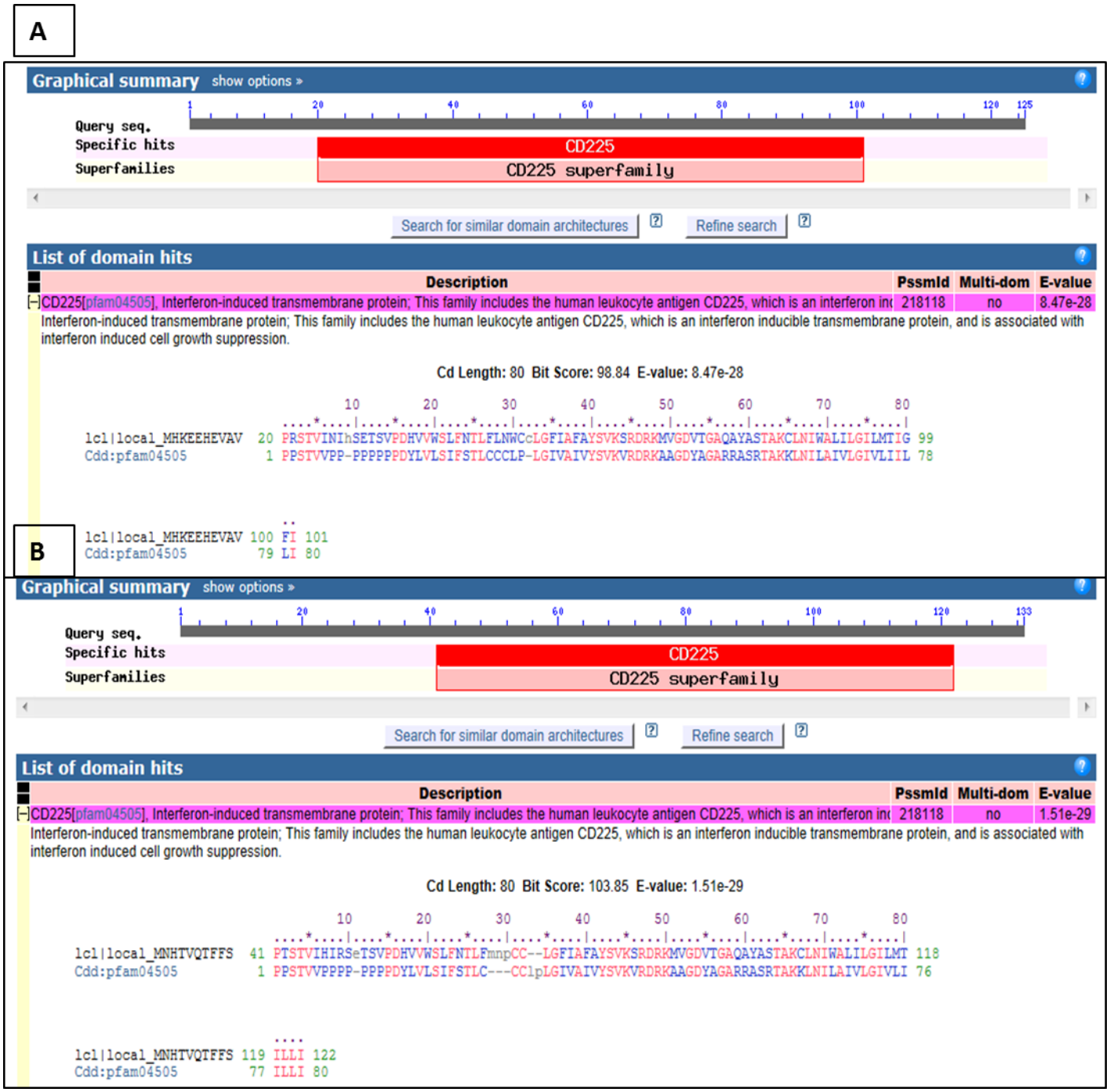


Figure 2: Screenshot of conserved domain database analysis of (A)IFITM1 protein sequence; (B)IFITM3 protein sequence

ISSN: 0975-8585

Index	Query	Status	Score	Sequences	Taxonomy	Taxid	Accessions
1	MHKEEHEVAVLGAPP	done	1	1738	cellular organisms	131567	pfam04505
1	MHKEEHEVAVLGAPP	done	1	42	Bacteria	2	pfam14237~pfam04505
1	MHKEEHEVAVLGAPP	done	1	20	Bacteria	2	cl11198~pfam04505
1	MHKEEHEVAVLGAPP	done	1	4	Oscillatoriophycideae	1301283	cl02429~pfam04505
1	MHKEEHEVAVLGAPP	done	1	3	Frankia	1854	cl02542~pfam04505
1	MHKEEHEVAVLGAPP	done	1	2	pseudoobscura subgroup	32358	pfam04505~pfam00379
1	MHKEEHEVAVLGAPP	done	1	2	Oryziaslatipes	8090	pfam04505~cl02808
1	MHKEEHEVAVLGAPP	done	1	2	Capitellateleta	283909	cl00117~pfam04505
1	MHKEEHEVAVLGAPP	done	1	2	Euarchontoglires	314146	pfam07966~cl11403~pfam04505

Table 7 A: Possible conserved domain matching the IFITM1.

Index	Query	Status	Score	Sequences	Taxonomy	Taxid	Accessions
1	MNHTVQTFFSPVNSG	done	1	1738	cellular organisms	131567	pfam04505
1	MNHTVQTFFSPVNSG	done	1	42	Bacteria	2	pfam14237~pfam04505
1	MNHTVQTFFSPVNSG	done	1	20	Bacteria	2	cl11198~pfam04505
1	MNHTVQTFFSPVNSG	done	1	4	Oscillatoriophycideae	1301283	cl02429~pfam04505
1	MNHTVQTFFSPVNSG	done	1	3	Frankia	1854	cl02542~pfam04505
1	MNHTVQTFFSPVNSG	done	1	2	pseudoobscura subgroup	32358	pfam04505~pfam00379
1	MNHTVQTFFSPVNSG	done	1	2	Oryziaslatipes	8090	pfam04505~cl02808
1	MNHTVQTFFSPVNSG	done	1	2	Capitellateleta	283909	cl00117~pfam04505
1	MNHTVQTFFSPVNSG	done	1	2	Euarchontoglires	314146	pfam07966~cl11403~pfam04505

Table 7 B: Possible conserved domain matching the IFITM3.

The query protein sequences were then subjected to Motif analysis using PRINTS42 and Matrix Blos62. The results for IFITM1 protein showed one fingerprint, of which the NRPEPTIDEWR domain fingerprint had two motif regions-motif 4(51-61) and

ISSN: 0975-8585

motif 6(81-94) drawn from conserved regions spanning the central portion of the alignment while that for IFITM3 protein also showed two fingerprints, of which the SUPERTUBBY domain fingerprint had two motif regions– motif 5(8-25) and motif 7(70-93).

A FOR IFITM1

Ten top scoring fingerprints for your query. Detailed by motif

FingerPrint Name	Motif Number	IdScore	PfScore	Pval	Sequence	Length	low	Pos	high
NRPEPTIDEWR	4 of 8	40.00	127	7.80e-02	CLGFIAFAYS	10	0	51	0
	6 of 8	26.92	177	6.60e-02	TAKCLNIWALILG	13	0	81	0

B FOR IFITM3

Ten top scoring fingerprints for your query. Detailed by motif

FingerPrint Name	Motif Number	IdScore	PfScore	Pval	Sequence	Length	low	Pos	high
SUPERTUBBY	5 of 7	17.65	151	5.25e-02	FFSPVNSGQPPNYEMLK	17	0	8	0
	7 of 7	26.88	143	9.11e-02	PCCLGFIAFAYSVKSRDRKMVGD	23	0	70	0
NRPEPTIDEWR	4 of 8	40.00	127	8.11e-02	CLGFIAFAYS	10	0	72	0
	6 of 8	26.92	177	6.86e-02	TAKCLNIWALILG	13	0	102	0

Figure 3: Screenshot of fingerprint analysis of query sequences (A) IFITM1 protein sequence; (B) IFITM3 protein sequence

The query proteins were subjected to BLASTP analysis & the alignments of similar sequences were computed on the basis of expectation (E) values. IFITM1 and IFITM3 proteins showed good identity percentage with its aligned sequences. Both proteins showed around 100% identity and significant e-value with interferon inducing proteins of Homo sapiens. The query sequences also showed 99% identity with some primate sequences. Therefore, for both IFITM1 and IFITM3 proteins the closest match was found with the interferon inducing protein groups.

Comparative modeling

Homology protein modeling uses experimentally determined protein structures (templates) to predict the 3-D structure of another protein that has a similar amino acid sequence (the target). This approach to modeling is possible since a small change in the protein sequence usually results in a small change in its corresponding 3-D structure.

Modeling of IFITM1 protein and model validation

The structure was modeled using I-TASSER by ab initio method. The templates for modeling were selected by LOMETS, a locally installed meta-threading approach. The modeled IFITM1 and the superimposed templates on the modeled structure are shown in the Fig 4.

ISSN: 0975-8585

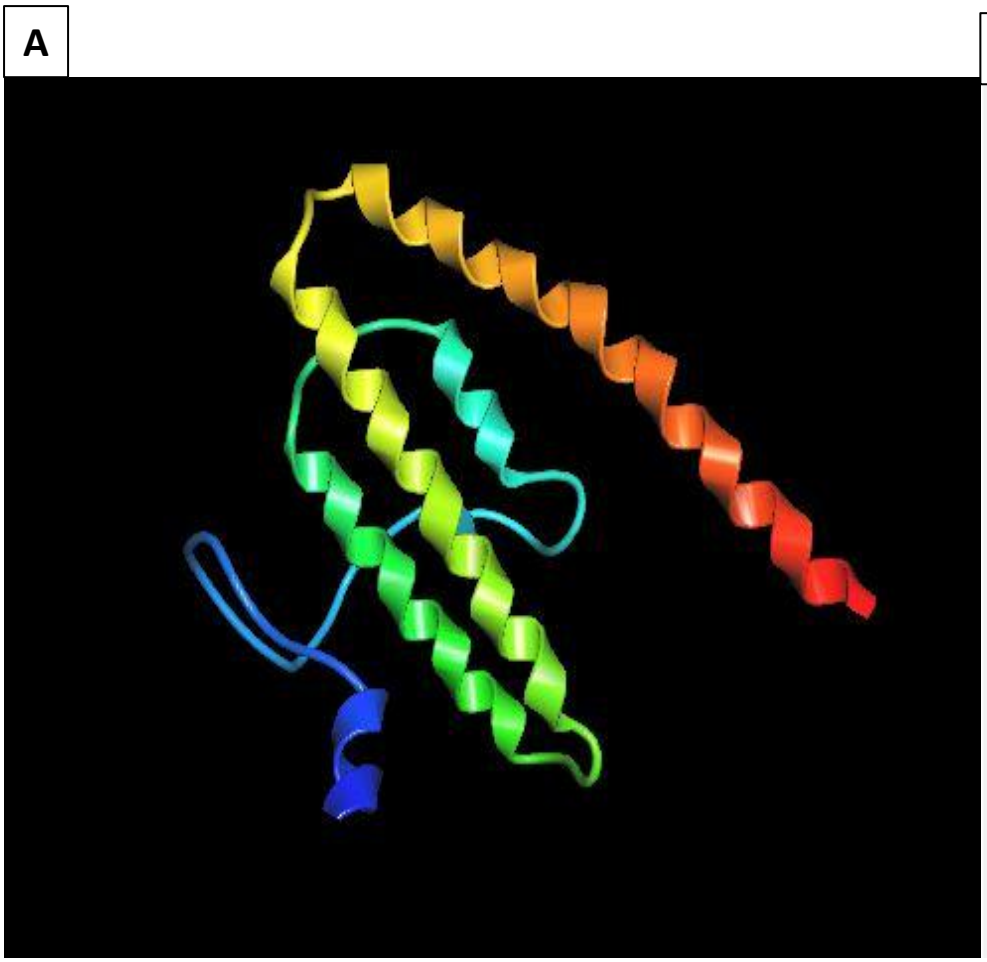


Figure 4: (A)Screenshot of I-TASSER 3-D Modeled structure of IFITM1 protein showing UCSF Chimera showing helix and strands and Coils; (B)Screenshot of templates used for modeling

Name	C-score	Exp.TM-Score	Exp.RMSD	No.of decoys	Cluster density
Model 1	**-3.89**	**0.30+-0.09**	**13.6+-4.0**	**1138**	**0.0263**
Model 2	-4.09			963	0.0215
Model 3	-4.61			643	0.0129
Model 4	-4.98			437	0.0089
Model 5	-4.96			404	0.0090

Table 8 A: I-TASSER modeling score

The first model was selected depending on the confidence score for estimating the quality of predicted models by I-TASSER. It was calculated based on the significance of threading template alignments and the convergence parameters of the structure assembly simulations. C-score is typically in the range of [-5 to 2], where a C-score of higher value signifies a model with a high confidence and vice-versa. Based on the number of decoys and clustering density the first model was selected (Table 7). Here we only report the quality prediction (TM-score and RMSD) for the first model, because we found that the correlation between C-score and TM-score is weak for lower rank models.

ISSN: 0975-8585

Modeling of IFITM3 protein and model validation

The structure was modeled using I-TASSER by ab initio method. The templates for modeling were selected by LOMETS, a locally installed meta-threading approach. The modeled IFITM3 and the superimposed templates on the modeled structure are shown in the Fig 5.

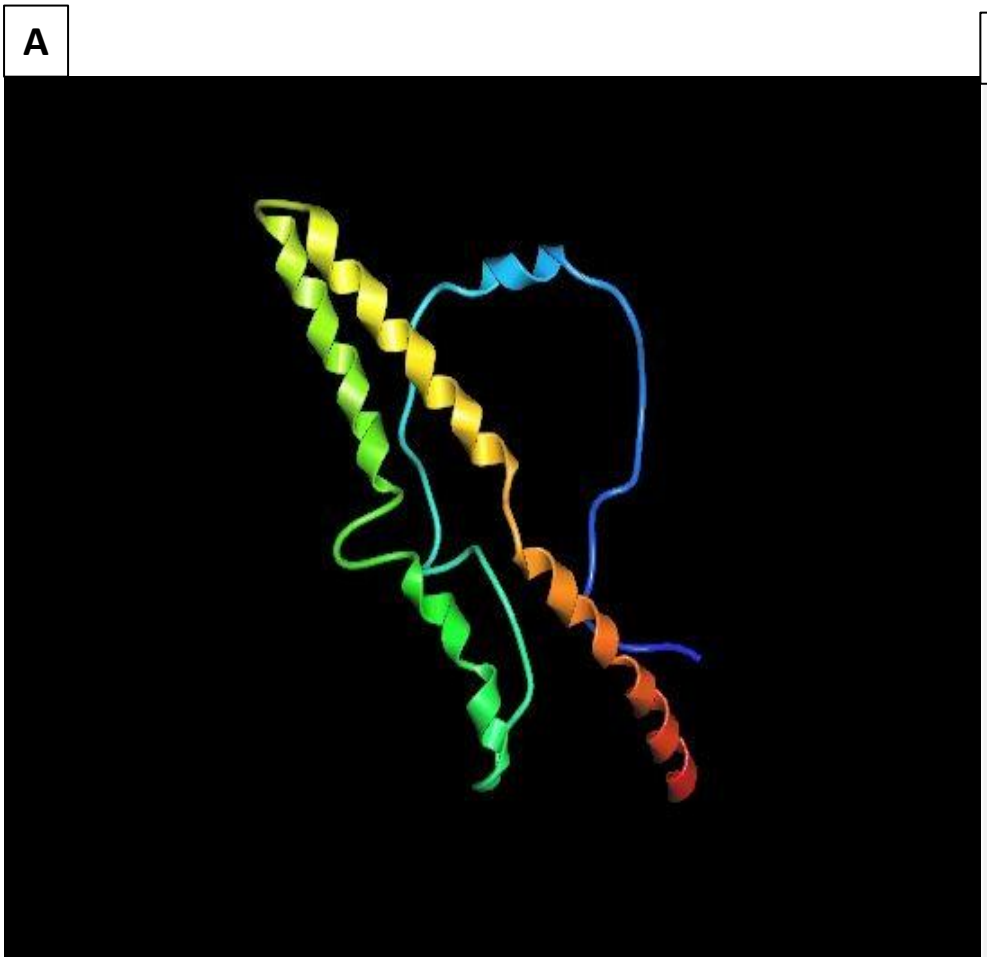


Figure 5: (A)Screenshot of I-TASSER 3-D Modeled structure of IFITM1 protein showing UCSF Chimera showing helix and strands and Coils; (B)Screenshot of templates used for modeling

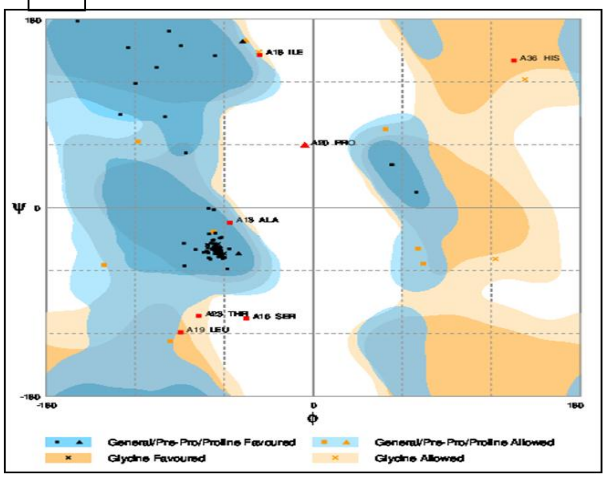
Name	C-score	Exp.TM-Score	Exp.RMSD	No.of decoys	Cluster density
Model 1	**-4.06**	**0.28+-0.09**	**13.9+-3.9**	**1271**	**0.0237**
Model 2	-4.37			988	0.0173
Model 3	-4.44			823	0.0162
Model 4	-4.71			628	0.0123
Model 5	-5.00			450	0.0089

Table 8 B: I-TASSER modeling score

The modeled structures of IFITM1 & IFITM3 proteins were validated by RAMPAGE. This tool provides detailed information about the target protein and the model building process, functional annotation, a detailed template selection log & the target-template alignment. Energy Minimization of the structure was done using ArgusLab [23]. The summary of the model building and model quality assessment are as shown in Fig 6.

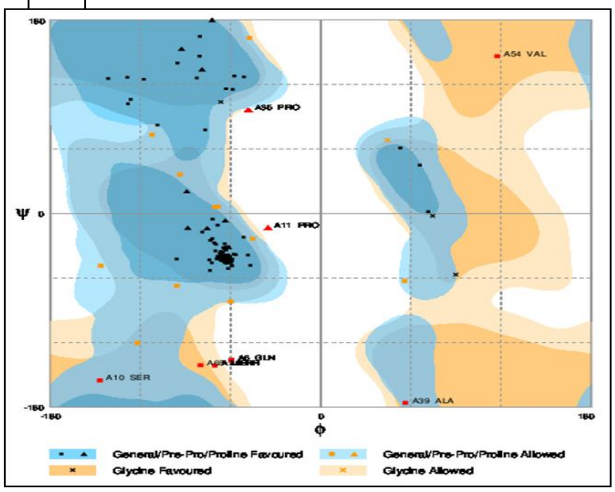
ISSN: 0975-8585

A



Plot Analysis	
Number of residues in favoured region	105 (85.4%)
Number of residues in allowed region	11 (8.9%)
Number of residues in outlier region	7 (5.7%)

B



Plot Analysis	
Number of residues in favoured region	111 (84.7%)
Number of residues in allowed region	12 (9.2%)
Number of residues in outlier region	8 (6.1%)

Figure 6: Ramchandran plot and plot statistics of the modeled protein; (A) IFITM1 (B) IFITM3

The models were also tested for ϕ and ψ torsion angles using the Ramchandran plot. Ramachandran plot analysis for IFITM1 showed that main-chain conformations for 85.4 % of amino acid residues are within the most favored or allowed region, 8.9 % in the allowed and 5.7% in the generously allowed region. A similar analysis for IFITM3 showed that 84.7 % of amino acid residues are within the most favored or allowed region, 9.2 % in the allowed and 6.1% in the generously allowed region. In general, a score close to 100% implies good stereochemical quality of both the models.

CONCLUSION

IFITM1 and IFITM3 proteins play a major role in host antiviral immune responses. This has been the idea behind the work to model these proteins using computataional methods.The proteins IFITM1 and IFITM3 were modeled using in silico approach and characterization was carried out for both which could be further used to carry out in vitro and in vivo study of these proteins.

ISSN: 0975-8585

REFERENCES

[1] Friedman RL, Manly SP, McMahon M, Kerr IM, Stark GR. Cell. 1984;38(3):745-55.
[2] Martensen PM, Justesen J. J Inter Cytokine Res 2004;24(1):1-19.
[3] Hickford D, Frankenberg S, Shaw G, Renfree MB. BMC Genomics 2012;13:155. Epub 2012/04/28.
[4] Brass AL, Huang I, Benita Y, John SP, Krishnan MN, Feeley EM, et al. Cell 2009;139(7):1243-54.
[5] Huang IC, Bailey CC, Weyer JL, Radoshitzky SR, Becker MM, Chiang JJ, et al. PLoS pathogens. 2011;7(1):e1001258.
[6] Eric M Feeley ET AL. PLoS Pathog 2011;7(10):17.
[7] Diamond MS, Farzan M. Nature reviews Immunology 2013;13(1):46-57.
[8] Fredy Siegrist ME, and Ulrich Certa. J Inter Cytokine Res 2011;31(1):16.
[9] http://linux1.softberry.com/berry.phtml?topic=cys_rec&group=programs&subgroup=propt.
[10] Marchler-Bauer A, Anderson JB, Chitsaz F, Derbyshire MK, DeWeese-Scott C, Fong JH, et al. Nucleic Acids Res 2009;37(suppl 1):D205-D10.
[11] Attwood TK, Beck ME, Flower DR, Scordis P, Selley J. Nucleic Acids Res 1998;26(1):304-8.
[12] Geourjon C, Deleage G. CABIOS 1995;11(6):681-4.
[13] Combet C, Blanchet C, Geourjon C, Deleage G. Trends Biochem Sci 2000;25(3):147-50.
[14] Remmert M, Biegert A, Hauser A, Soding J. Nature Methods 2012;9(2):173-5.
[15] Soding J. Bioinformatics 2005;21(7):951-60.
[16] Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. J Mol Biol 1990;215(3):403-10.
[17] Roy A, Kucukural A, Zhang Y. Nature Protocols 2010;5(4):725-38.
[18] Roy A, Yang J, Zhang Y. Nucleic Acids Res. 2012;40(Web Server issue):W471-7. E
[19] Zhang Y. BMC Bioinformatics. 2008;9(1):40.
[20] Laskowski RA, Rullmannn JA, MacArthur MW, Kaptein R, Thornton JM. J Biomol NMR. 1996;8(4):477-86. Epub 1996/12/01.
[21] Lovell SC, Davis IW, Arendall WB, 3rd, de Bakker PI, Word JM, Prisant MG, et al. Proteins 2003;50(3):437-50.
[22] Morris AL, MacArthur MW, Hutchinson EG, Thornton JM. Proteins 1992;12(4):345-64.
[23] http://www.arguslab.com/arguslab.com/ArgusLab.html.