

PROTEIN MODELLING AND IDENTIFICATION OF BINDING INTERACTIONS BETWEEN FAT10 AND MAD2

Anupama Pandrangi*

Osmania University, Hyderabad, INDIA

ABSTRACT

Chromosome instability involving gains and loss of chromosomes has been found to occur in most malignancies. Genes responsible for chromosomal instability in human cancers include mitotic "check point" genes that monitor the proper progression through the cell cycle. MAD2 is a key mitotic spindle checkpoint protein whose primary role is to ensure that all of the chromosomes are properly attached to the mitotic spindle before the onset of anaphase. FAT10, a member of the ubiquitin-like modifier family of proteins, may modulate tumorigenesis by noncovalent interaction to MAD2. The inhibition of MAD2 function has been associated with chromosomal instability, a characteristic of many cancers. This paper deals with the protein-protein interactions between FAT10 and MAD2 using protein-protein docking studies. The overall study identifies the amino acid residues which are important for binding in FAT10. Further studies are aimed at designing new chemical entities which may be used as potent therapeutic leads.

Received on: 26th-Apr-2014

Revised on: 09th-June-2014

Accepted on: 09th- July-2014

Published on: 24th-Aug-2014

KEY WORDS

Chromosome instability; check point; FAT10; MAD2; docking studies; Protein-protein docking.

*Corresponding author: Email: p_anupama2002@yahoo.com; Tel: +91-9966193890

[1] INTRODUCTION

Dysregulated cell-cycle control or apoptosis often leads to tumorigenesis. An instrumental role is played by ubiquitin and ubiquitin-related families of proteins in various cellular processes through modification of target proteins [1, 2].

Ubiquitin is a conserved 76 amino-acid polypeptide that plays role in conjugation of target proteins, thereby tagging "these proteins for degradation via the 26s proteasome pathway [3, 4]. This process of ubiquitylation, is an ATP dependent process that involves the sequential action of at least three different classes of enzymes. An E1 or ubiquitin activating enzyme, an E2 or ubiquitin conjugating enzyme, an E3 or ubiquitin protein ligase and in some cases an E4 chain elongation factor [5].

1.1. Ubiquitin-like proteins

A new group of proteins have been discovered recently that are related to ubiquitin or function similarly. These ubiquitin-like proteins can be divided into two groups, namely the ubiquitin - like modifiers (UBLs) and Ubiquitin-domain proteins (UDPs) [1].

Ubiquitin- like modifiers (UBL) modify other proteins by covalent formation of an isopeptide-bond with their target. These include Ubiquitin, SUMO-1/SMT3, SUMO-2, SUMO-3, NEDD8, ISG-15, FAT 10, HUB, Fau, An1a, An1b, APG12, URM1. Ubiquitin domain proteins (UDP) on the other hand are only structurally linked with ubiquitin and cannot become

covalently conjugated to target proteins [1]. These include RAD23/HHR23A, B; DSK2, PLIC-1, PLIC-2/Chap1, XDRP1, BAG-1, BAT3/Chap2, Scythe, Parkin, UIP28/RBCK1, UBP6, Elongin B, Gdx.

FAT10 or diubiquitin belongs to UBL class of ubiquitin - like proteins. The FAT10 gene is identified as one of the genes in the MHC class I HLA-F locus of chromosome 6 [6]. It encodes an 18kDa protein containing 165 amino acid residues that has two ubiquitin- like domains, which are separated by a linker of 5 amino acid. The N-terminal domain is with 29% identity and C-terminal domain is with 36% identity with ubiquitin [Figure-1].

Of the seven lysine-residues in ubiquitin four are conserved in both ubiquitin- like domains of FAT10. These lysine residues correspond with the Lys-27, Lys-33, Lys-48 and Lys-63 of ubiquitin. Among these lysine residues Lys-48 and Lys-63 of ubiquitin are important for polyubiquitin-chain formation. Also the diglycine motif of the very C-terminus of the C-terminal domain is conserved. Atypical of ubiquitin is the appearance of four cysteine residues in FAT10 [8].

Over-expression of the FAT10 gene was observed in the tumors of several cancers including gastrointestinal and gynaecological cancers [9]. Increased mitotic non-dysjunction and chromosome instability was observed in cells having high level expression of FAT10 protein [10]. Fat10 has been shown to bind

noncovalently to the human spindle assembly checkpoint protein, MAD2 [11], a protein responsible for maintaining spindle integrity during mitosis [12]. The inhibition of MAD2 function has been associated with chromosomal instability, a characteristic of many cancers [13,14].

Protein-protein interactions are integral to many cellular control mechanisms, because of the formation of protein-protein complexes in many biological events. Disruption of protein-

protein interactions can cause biochemical diseases [15, 16]. In order to identify the way how proteins interact with each other, the structural details of these complexes at atomic level and their docking studies will be particularly important for elucidating functional mechanisms and designing inhibitors, and thus is a major goal of structural biology [17-19]. In the present study, we have carried out protein-protein docking studies between the two proteins i.e., FAT10 and MAD2.

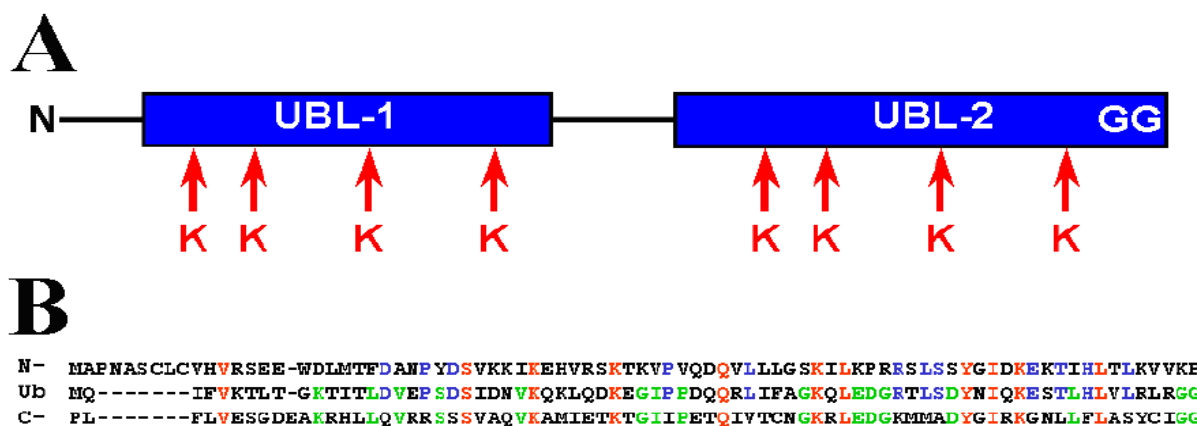


Fig: 1. Primary Structure of FAT10 and its sequence similarity to ubiquitin. (from Mark Steffen Hipp, 2005) [7]. (A) schematic diagram of FAT10 (B) sequence comparison of the N and C-terminal halves of FAT10 (N- / C-) with ubiquitin (Ub)

[II] MATERIALS AND METHODS

2.1. Sequence retrieval and template selection

The amino acid sequence of FAT10 is retrieved in FASTA format from SWISS PROT database followed by BLAST against PDB for template selection. The BLAST is used to find the similarity of the sequence to closest homologous proteins with known structures available in the PDB and identifies the structure with high identity and similarity to be employed as template for homology modeling.

2.2. Sequence alignment and model building

By extracting the sequence of template and target the alignment process is carried out by using ClustalW. Using "MODELLER 9.12" the 3D structure of FAT10 is generated. Energy minimization of the modeled structure is carried out by applying CHARMM force fields and steepest descent algorithm followed by conjugate gradient algorithm in DS until the convergence gradient is satisfied.

2.3. Model validation

2.3.1. Procheck

Procheck is used in validation of protein structure and models by verifying the parameters like Ramachandran plot quality, peptide bond planarity, bad non-bonded interactions, main chain hydrogen bond energy, C-alpha chirality and over-all G factor and the side chain parameters like standard deviations of chi1 gauche minus, trans and plus, pooled standard deviations of chi1 with respect to refined structures [20].

2.3.2. Prosa

This program compares Z scores between target and template structure which are a measure of compatibility between its sequence and structure. The Z score of the model should be comparable to the Z scores obtained from the template [21, 22].

2.3.3. Docking studies

Protein-protein docking studies of diubiquitin (FAT10) with its receptor, mitotic spindle check point protein (MAD2), were performed in GRAMM-X Protein-Protein Docking Web Server v.1.2.0 using the GRAMM global search algorithm [23]. The crystal structure of MAD2 with PDB entry 2V64 containing 205 amino acid residues at a resolution of 2.90 was taken from the PDB database.

[III] RESULTS

3.1. Sequence retrieval and template selection

The FASTA sequence of the FAT10 protein (1-165) taken from SWISSPROT database accession number: **O15205**, entry name: UBD_HUMAN, and protein name: **Ubiquitin D** was submitted to BLAST against PDB database. The BLAST results yield X-ray structure of 2ZVN from human having a sequence identity of 31% with a resolution of 3.00 Å. All the further procedures are carried out using MODELLER 9.12.

3.2. Sequence alignment and model building

An essential input to a homology modeling program is the sequence alignment between the template-target pair. Alignment of FAT10 protein with the extracted sequence of 2ZVN was carried out using Clustal W. On the basis of this alignment as input, model of the FAT10 is built using "MODELLER 9.12." Ten molecular models of FAT10 are generated. The refinement process is carried out using DS by applying CHARMM force field and steepest descent method is applied with 0.001 minimizing RMS gradient and 2000 minimizing steps followed by conjugant gradient method till it reaches the satisfactory results for minimization. The energy refinement method gives the best conformation to the model [Figure- 2].

3.3. Model validation

The final refined modeled structure of FAT10 protein is analyzed by the Procheck and Prosa.

3.4. Procheck

Procheck is used to evaluate the overall stereochemical quality of the modeled protein by analyzing the overall residue by residue geometry and psi and phi torsion angles of Ramachandran plot showing residues with most favorable region (88.4%), generously allowed region (0.0%), additionally allowed region (10.2%), and disallowed region (1.4%) in comparison to 2ZVN template (80.2%, 0.8%, 19.2%, and 0.0%), respectively. The plot statistics are given in [Table-1, Figure-3].



Fig: 2. Modeled FAT10 protein

Table: 1. The % of residues in the core region of the Ramachandran plot for the built FAT10 model and the template

Structure	Core	Allowed	Generous	Disallowed
2ZVN	80.2	19.2	0.8	1.4
FAT10	88.4	10.2	0.0	0.0

3.5. Prosa

Prosa was used for quality assessment of the model which revealed that the FAT10 model matched NMR region of the plot

[IV] DISCUSSION

4.1. Docking studies

Protein-protein docking has been performed using GRAMM-X docking server which employs smoothed potentials, refinement stage, and knowledge based scoring [23]. It uses smoothed Lennard-Jones potential on a fine grid during the global search FFT stage, followed by the refinement optimization in continuous coordinates and rescoring with several knowledge-based potential terms. The extracted PDB structures 2V64 of MAD2 and FAT10 are taken for docking studies. The top 4000 grid-based predictions are subjected to a conjugate gradient minimization in continuous 6D rigid body space with the soft potential. The minimization accumulates many points, initially located on the grid, in a fewer local minima. One representative prediction for each minimum is stored and the number of initial predictions falling into this minimum is marked as the volume

with Z score (-6.48) which is reliable to the Z score of the template 2ZVN (-7.37) . It signifies the quality of our model.

of the minimum. The average radius of such minima is 5 s. The local minimization of a smoothed landscape can be viewed as clustering on the original rugged Lennard-Jones landscape, and helps locate the protein binding funnel.

For each minimized prediction soft Lennard-Jones potential, evolutionary conservation of predicted interface, statistical residue-residue preference, volume of the minimum, empirical binding free energy and atomic contact energy are calculated. To eliminate predictions that are likely to be located far from the correct binding site, Support Vector Machine filter trained on a subset of the benchmark set is applied using the above mentioned set of potential terms. The remaining predictions are re-scored by a weighted sum of the potential terms. Figure- 4 illustrates the top ranking pose of predicted MAD2-FAT10 complex formation that reveals the correct binding mode of the protein.

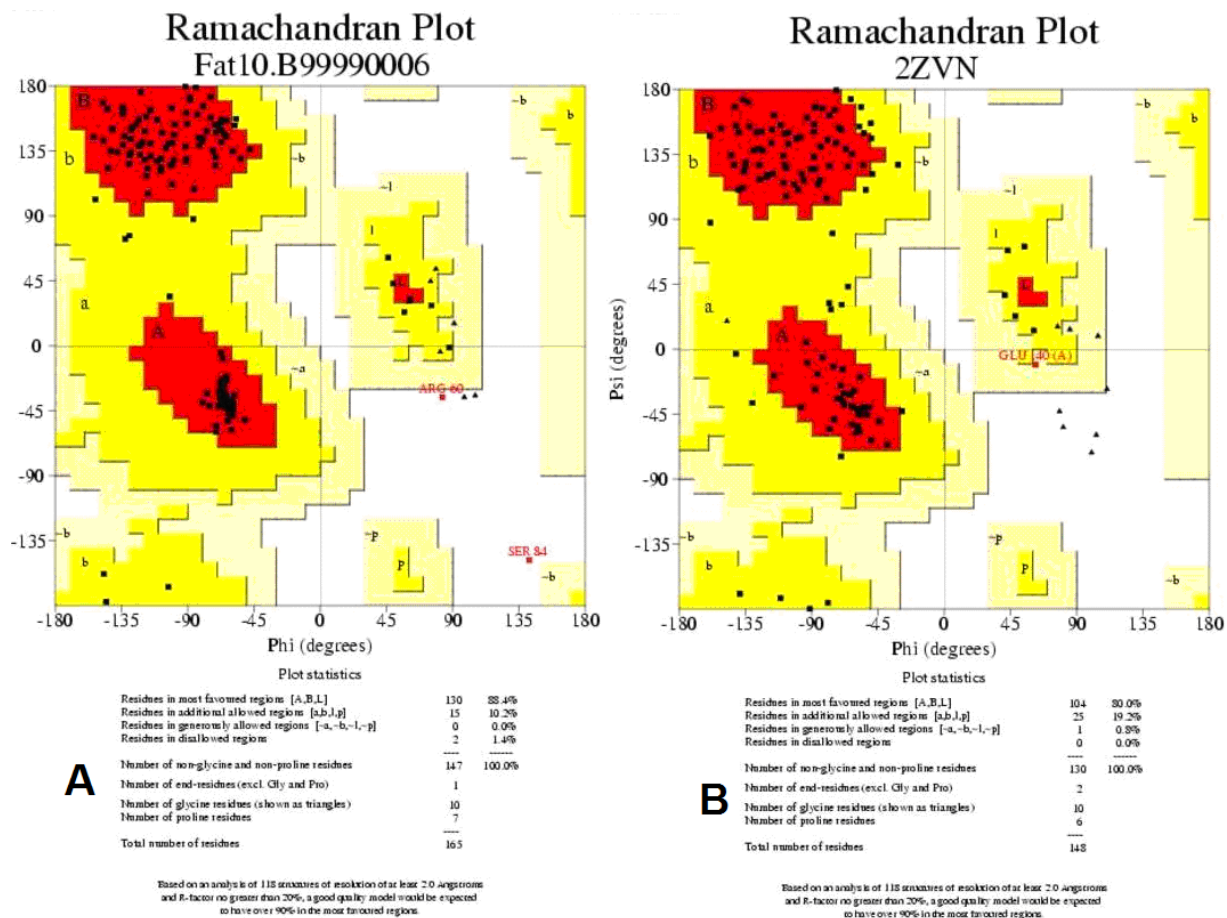


Figure: 3. A) Ramachandran's Map of FAT10 protein. B) Ramachandran's Map of 2ZVN protein

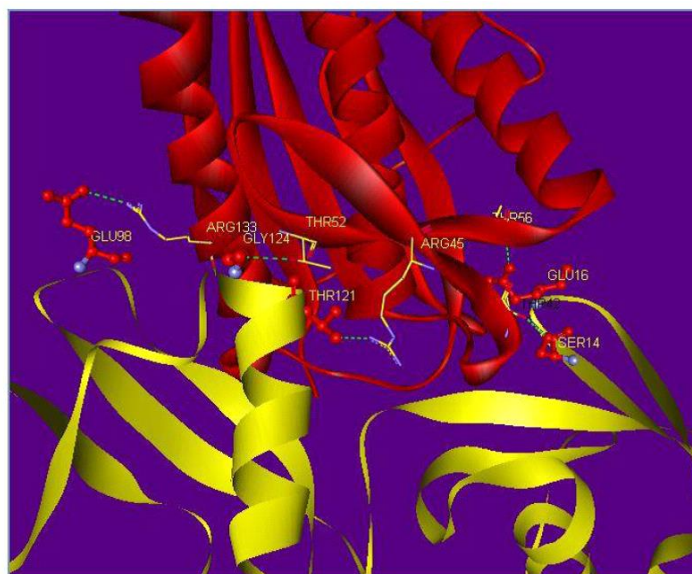


Fig: 4. The top ranking pose of predicted MAD2 (red colored flat ribbon shaped)-FAT10 (yellow colored flat ribbon shaped) complex formation illustrating intermolecular hydrogen bonding between the active site regions of the MAD2 and FAT10 complex. The amino acids on MAD2 involved in hydrogen bonding are represented in yellow colored stick shaped while the amino acids on FAT10 involved in hydrogen bonding are represented in red colored ball and stick shaped.

Several intermolecular hydrogen bonds consistently form between the active site regions of the MAD2 and FAT10 complexes which are listed in [Table-2]. All these binding interactions lead to stability of complex thus preventing the function of mitotic check point function of MAD2. MAD2 (mitotic arrest-deficient 2) is a key mitotic spindle checkpoint protein which ensures that all of the chromosomes are properly attached to the mitotic spindle before the onset of anaphase [24]. It is activated by associating with unattached kinetochores. Activated MAD2 binds to Cdc20 and prevents the anaphase-promoting complex from ubiquitylating securin. As a result, anaphase is delayed until all of the kinetochores are attached by microtubules and the chromosomes are properly aligned along the metaphase plate [12, 25, 26]

Table: 2. Various binding interactions between the MAD2 and FAT10 complex

Sl. No	Intermolecular Hydrogen bonds	
	MAD2	FAT10
1.	THR 42	SER 14
2.	ARG 45	THR 121
3.	THR 52	GLY 124
4.	THR 56	GLU 16
5.	ARG 133	GLU 98
6.	SER 14	THR 42

[IV] CONCLUSION

The present study demonstrated the binding interactions of the MAD2-FAT10 complex through docking analysis. Docking studies reveals the correct binding mode and interacting amino acids. This protein – protein docking studies of MAD2 and FAT10 provides a great assistance in understanding structural details and may lead to the establishments of therapeutic approaches in designing of effective drugs for gastrointestinal and gynaecological cancers.

ACKNOWLEDGEMENT

The author acknowledge Professor V. Uma, Department of Chemistry, Osmania University, for providing technical assistance in carrying out this work

CONFLICT OF INTEREST

Author declares no conflict of interest.

FINANCIAL DISCLOSURE

The work is not supported by any grant.

REFERENCES

- [1] Jentsch S, Pyrowolakis G. [2000] Ubiquitin and its kin: how close are the family ties? *Trends Cell Biol* 10:335–342.
- [2] Jesenberger V, Jentsch S. [2002] Deadly encounter: ubiquitin meets apoptosis. *Nat Rev MolCell Biol* 3:112–121.
- [3] Hochstrasser M. [1996] Protein Degradation or Regulation: Ub the Judge. *Cell* 84:813–815.
- [4] Hochstrasser M. [1996] Ubiquitin-dependent protein degradation. *Annu. Rev. Genet* 30:405–439.
- [5] Pickart CM. [2001] Mechanisms underlying ubiquitination. *Annu Rev Biochem* 70:503–533.
- [6] Fan W, Cai W, Parimoo S, Lennon GG, Weissman SM. [1996] Identification of seven new human MHC class I region genes around the HLA-F locus. *Immunogenetics* 44:97–103.
- [7] Mark Steffen Hipp. [2005] NUB1L and FAT10, two ubiquitin-like proteins involved in protein degradation. Dissertation.
- [8] Bates EFM, Ravel O, Dieu MC, et al. [1997] Identification and analysis of a novel member of the ubiquitin family expressed in dendritic cells and mature B cells *Eur J Immuno* 127:2471–2477.
- [9] Lee CG, Ren J, Cheong IS, et al. [2003] Expression of the FAT10 gene is highly upregulated in hepatocellular carcinoma and other gastrointestinal and gynecological cancers. *Oncogene* 22:2592–2603.
- [10] Ren J, A Kan A, SH Leong SH, et al. [2006] Fat10 plays a role in the regulation of chromosomal stability. *J Biol Chem* 281:11413 –11421.
- [11] LiuY, Pan J, Zhang C, et al. [1999] A MHC-encoded ubiquitin-like protein (FAT10) binds noncovalently to the spindle assembly checkpoint protein MAD2. *Proc Natl Acad Sci USA* 96:4313–4318.
- [12] Shah JV, Cleveland DW. [2000] Waiting for Anaphase: Mad2 and the Spindle Assembly Checkpoint. *Cell* 103:997–1000.
- [13] Wang X, Jin DY, Wong YC, et al. [2000] Correlation of defective mitotic checkpoint with aberrantly reduced expression of MAD2 protein in nasopharyngeal carcinoma cells. *Carcinogenesis* 21:2293–2297.
- [14] Gemma A, Hosoya Y, Seike M, Uematsu K, Kurimoto F. [2001] Genomic structure of the human MAD2 gene and mutation analysis in human lung and breast cancers. *Lung Cancer* 32:289–295.
- [15] Berman HM, Westbrook J, Feng Z, et al. [2000] The Protein Data Bank. *Nucleic Acids Research* 28:235–242.
- [16] Valdar WS, Thornton JM. [2001] Protein–protein interfaces: analysis of amino acid conservation in homodimers. *Proteins* 42:108–124.
- [17] Camacho CJ, Vajda S. [2002] Protein–protein association kinetics and protein docking. *Curr. Opin. Struct. Biol* 12:36–40.
- [18] Halperin I, Ma B, Wolfson H, Nussinov R. [2002] Principles of docking: An overview of search algorithms and a guide to scoring functions. *Proteins* 47:409–443.
- [19] Smith GR, Sternberg MJ. [2002] Prediction of protein–protein interactions by docking methods. *Curr Opin Struct Biol* 12:28–35.
- [20] Laskowski RA, MacArthur MW, Moss DS, Thornton JM. [1993] PROCHECK - a program to check the stereochemical quality of protein structures. *J App Cryst* 26: 283–291.

- [21] Wiederstein and Sippl. [2007] ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins. *Nucleic Acids Res* 35:W407–W410.
- [22] Sippl MJ. [1993] Recognition of Errors in Three-Dimensional Structures of Proteins. *Proteins* 17:355–362.
- [23] Tovchigrechko A , Vakser IA. [2006] GRAMM-X public web server for protein- protein docking. *Nucleic Acids Res* 34:W310–314.
- [24] Yong Li, Robert Benezra. [1996] Identification of a Human Mitotic Checkpoint Gene: hsMAD2. *Science* 274:246–248.
- [25] Yu H. [2002] Regulation of APC–Cdc20 by the spindle checkpoint. *Curr. Opin. Cell Biol* 14:706–714.
- [26] Musacchio A, and K.G. Hardwick KG. [2002] The spindle checkpoint: structural insights into dynamic signaling. *Nat Rev Mol Cell Biol* 3:731–741.

ABOUT AUTHOR

Dr. Anupama Pandrangi holds PhD from University of Hyderabad, Hyderabad, India. Her major research interest includes Drug designing using computational techniques.