International Journal of Computational Bioinformatics and In Silico Modeling
Vol. 3, No. 5 (2014): 466-472
Research Article
Open Access

ISSN: 2320-0634

Structural Characterization of Arg322Gln Mutation in Fat Mass and Obesity Associated Protein

Mazhar Hussain*, Ramla Ashfaq and Syeda Anum Zehra

National Centre of Excellence in Molecular Biology, University of the Punjab, Lahore.

*Corresponding author: Mazhar Hussain; e-mail: mazhar.hussain@cemb.edu.pk

Received: 06 August 2014 Accepted: 25 August 2014 Online: 01 September 2014

ABSTRACT

Fat mass and obesity associated gene has shown a strong association with adiposity in children and adults during several Genome Wide Association studies. It encodes an enzyme which catalyzes demethylation of single stranded DNA by using 2-oxoglutarate (2-OG) as a co-substrate. A mutation in the N - terminal of the enzyme encodes a glutamine in place of the arginine residue at position 322, which results in a complete disruption of protein function. Sequence analysis Arg322 is a highly conserved amino acid in the Fat mass obesity associated protein in animals. Homology modeling analysis reveals that 2-OG binding site of the enzyme which comprises various amino acids including Arg322. Molecular docking studies showed that Gln322 was unable to develop the hydrogen bonding and electrostatic interaction with the 2-OG. These interactions are important for extending the 2-OG to make its bonds more accessible to other amino acids. This study indicates that Arg322 is one of the most conserved amino acid in FTO protein which is directly involved in the catalysis of the 2-OG to succinate and CO_2.

Keywords: Fat mass obesity associated protein, 2-oxoglutarate, homology modeling, molecular docking, loss of function mutation

INTRODUCTION

Fat mass and obesity associated (*Fto*) gene encodes Fe(II) and 2-OG dependent dioxygenase, also known as fat mass and obesity associated (FTO) protein [1-3]. Genome wide association studies (GWAS) have revealed a strong association of single nucleotide polymorphisms (SNPs) in the first intron with adiposity in both children and adults [2]. About 16% of the study population (obese) were homozygous for the obesity associated risk alleles [4]. These associations are consistent among different ethnicities including Asians, Africans, Europeans and Hispanics. Hence, FTO protein can be a potential target for anti-obesity medicines [5].

FTO localizes to the nucleus of transfected cells [6]. Studies reveal that FTO mRNA is most abundant in the brain cells, particularly in hypothalamic nuclei [7]. Recent emerging evidence points to a role of FTO in sensing of nutrients, regulation of translation and growth. A deficiency in FTO leads to postnatal growth retardation and also pointing to some fundamental developmental role [8].

FTO catalyzes the Fe(II) and 2-OG dependent demethylation of 3-methylthymine, 3-methyluracil, and 6-methyladenine in single stranded DNA [2, 9]. The enzymatic activity stimulated by ascorbate concomitantly produces formaldehyde, succinate and CO_2 [10]. Differential scanning fluorometry and liquid chromatography based assay showed that both cyclic and acyclic analogues of 2-OG are potential inhibitors of FTO activity [9]. It shows that 2-OG is an essential co-substrate for the FTO enzymatic activity.

A non-synonymous mutation in an obese and a lean individual has discovered [3]. It substitutes an arginine (Arg) at position 322 to a glutamine (Gln) which results in complete inactivity to catalyze conversion of 2-OG to succinate. In this study, we are reporting structural characterization of this loss of function mutation and its effects on the function of FTO protein.

MATERIALS AND METHODS

Template selection and multiple sequence analysis

The FTO protein sequence was retrieved from UniProt with Entry ID: Q9C0B1. This sequence was used to query RCSB Protein Databank using E-value cutoff 1E-05 and BLOSUM62 scoring matrix by using SwissModel server [11]. Based on the query coverage, E-value and identity; crystal structure with RCSB PDB ID:4IE6 was selected. From UniProt, protein sequences of *Mus musculus*, *Rattus norvegicus*, *Xanopus laevis*, *Pongo abelii* and *Gallus gallus* were retrieved by Entry ID: Q8BGW1, Q2A121, Q68F54, Q5R7X0 and D7REI7, respectively. These sequences were subjected to multiple sequence alignment with human FTO using ClustalW [12] program to sort out the identical domains and insertions/deletions. Secondary structures of the protein were predicted by HHPred [13].

Construction of the homology models and energy minimization

The homology model of FTO was predicted using crystallographic structure 4IE6 as template using Modeller v9.10 program [14] which implements comparative protein structure modeling by satisfying the spatial constraints. A structure with R322Q substitution was generated by using mutation function in Swiss-pdb-viewer [15]. In this way, two structures were generated: (i) wild type was referred as FTO_w and (ii) R322Q mutant was referred as FTO_{R322Q}. These structures were energy minimized using Gromacs96 force field [16] by 1000 steps of Steepest Descent and 20000 steps of Conjugate Gradient until final structures reached an energy derivative of 0.001kcal/mol. Structural poses were visualized and rendered using Swiss-PDB-Viewer and VMD [17].

Assessment of the Homology Models

Predicted structures were evaluated based on their stereochemcial properties, folding energy and visual analysis. The stereochemical quality of the protein was checked by Ramachandran plot statistics implemented in PROCHECK which is based on the distribution of phi and psi torsion angles of backbone protein conformation [18]. The visual comparison of the predicted structures and template was performed using Swiss-pdb-viewer and Visual Molecular Dynamics (VMD) program. Root mean squared deviation (r.m.s) between template and query was measured using VMD MultiSeq tool. The probable protein folding energy of the theoretical model was studied by ProSa program [19] which compares energy criteria with the potential mean force derived from a large set of experimental structures.

Molecular Docking of Models with 2-OG

2-OG was extracted from the crystallographic structure of aspartate aminotransferase (PDB ID: 3ZZJ) using Swiss-Pdb-Viewer and saved as a separate file. For the prediction of ligand-binding sites, the surface analysis of protein structures was performed using PocketFinder [20], which maps the potential pockets and sort them out on the basis of pocket volume. AutoDock [21, 22] docking with a grid-point method was used, which calculates the interaction energy between probe atoms and receptor-binding sites before docking. The grid was centered on Fe(II) and was encompassed the residues in the potential pocket of the enzyme. The AutoGrid calculated gridpoint energy of every atom type in the 2-OG. During the docking procedure, 2-OG was configured as flexible while receptor was configured as fixed. The Lamarckian genetic algorithm was used to optimize 10 docking runs with initial population size of 150. Maximum number of generations and energy evaluations were set to 27000 and 2,500,000, respectively. Once docking was completed suitable docking results were selected based on energy and conformation profiles. The docked structures were analyzed using Molegro Viewer [23] and VMD.

RESULTS AND DISCUSSION

Sequence analysis of FTO protein

Human FTO, is a 505 amino acid enzyme, which consists of 39% helices, 27.5% sheets and 33.5% loops (Fig. 1). FTO enzyme is highly conserved in all higher eukaryotes and specifically mammals. It ranges in size from 501 amino acid residues in *Xenopus laevis* to 507 amino acid residues in *Gallus gallus*. FTO from *Pongo abelii* showed 98.61% homology while *Mus musculus* and *Rattus norvegicus* were 88% homologous to human FTO. Similarly, *Gallus* and *Xenopus* shared 58.22% and 53.49% homology with human FTO. Alignment study also revealed that Arg322 was conserved in all FTO protein sequences.

Homology Modeling and Structural analysis FTO protein

Crystal structure was selected on the basis of highest homology (92%) and lowest E-value. The superimposition of predicted FTO structure on the crystallographic structure before energy minimization and mutation, showed the r.m.s deviation of 0.60A and a Qh value of 0.8447 (Fig. 2). Ramachandran plot showed that in all cases more than 95% of residues were in allowed regions and 0.1% to 0.9% residues were in disallowed regions of the plot (Table 1). Analysis showed that all main chain and side chain were in “better” range. Moreover, the average G-factor, the measure of normality degree of protein properties was in permitted range for homology models. Bond lengths and angles were also in allowed range. The overall and local protein folding energies for FTO_w and FTO_{R322Q} were also within the permitted range with Z-scores of -10.68 and -11.11 (Fig. 3).

FTO protein has an N-terminal (1-326) and a C-terminal (327-505) domains (Fig. 4). N-terminal is catalytically active. Its active site residues make the largest pocket P1 on the surface of protein with an overall volume of 473A which is only 1% of the total protein size (Table 2). P1 contains Arg322 which dictates its importance in FTO- 2-OG interactions. Residues of loop L10 (H231 and D233) and sheet S13 (H307) carry an Fe(II) in a triangular symmetry (Fig.

5). There was no significant difference detected between FTO_W and FTO_{R322Q} structures, in terms of secondary structures and pocket sizes.

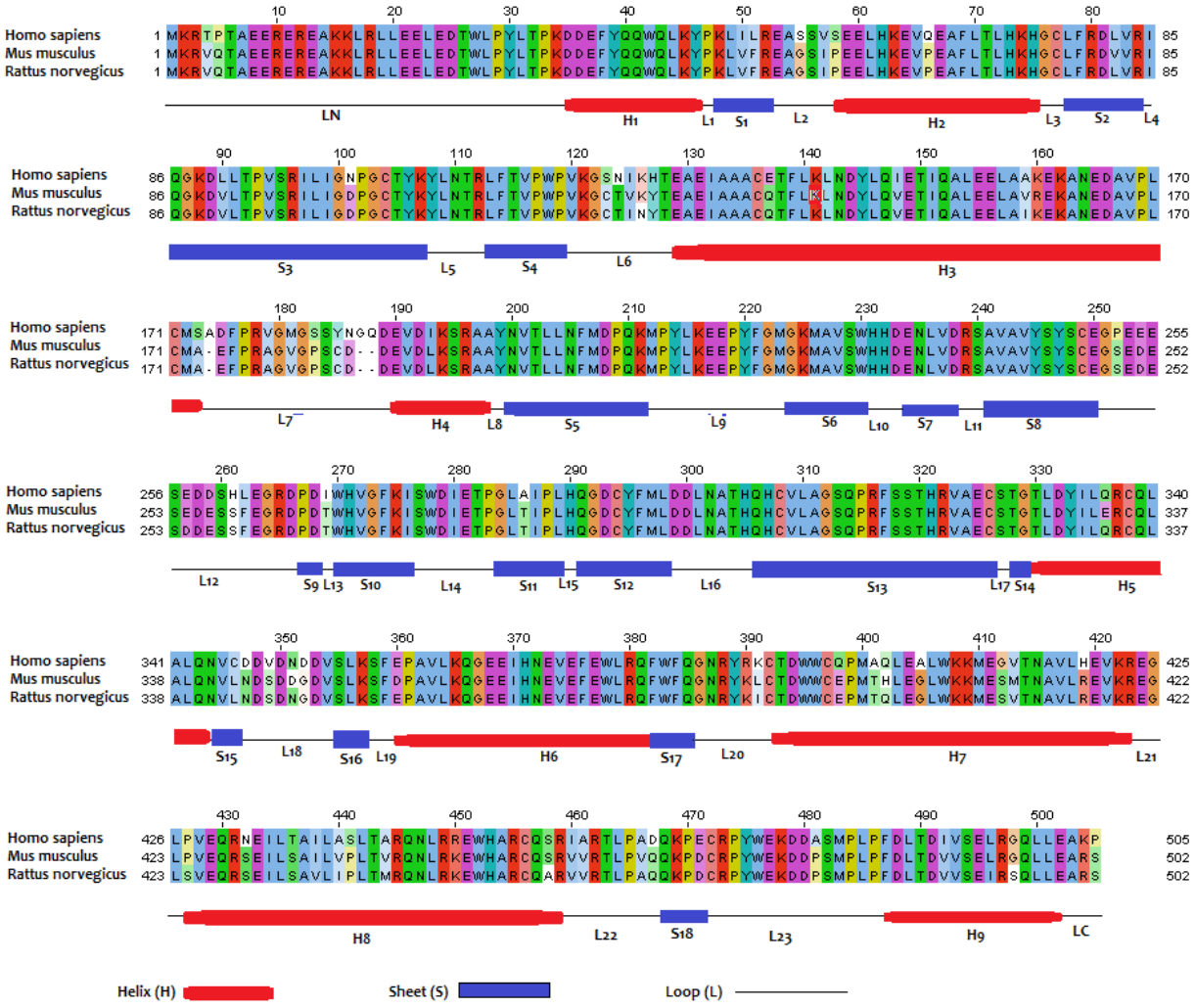

Figure 1. Multiple sequence alignment of Human FTO with *Mus musculus* and *Rattus norvegicus.* Alignment shows homology among 3 species and secondary structures are shown under the alignment.

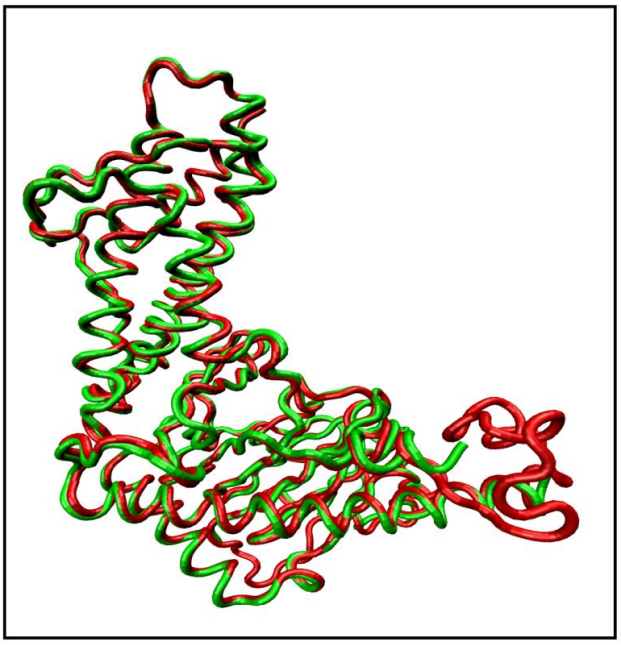

Figure 2. Tube representation of superimposition of the predicted FTO (Red) on the template (Green).

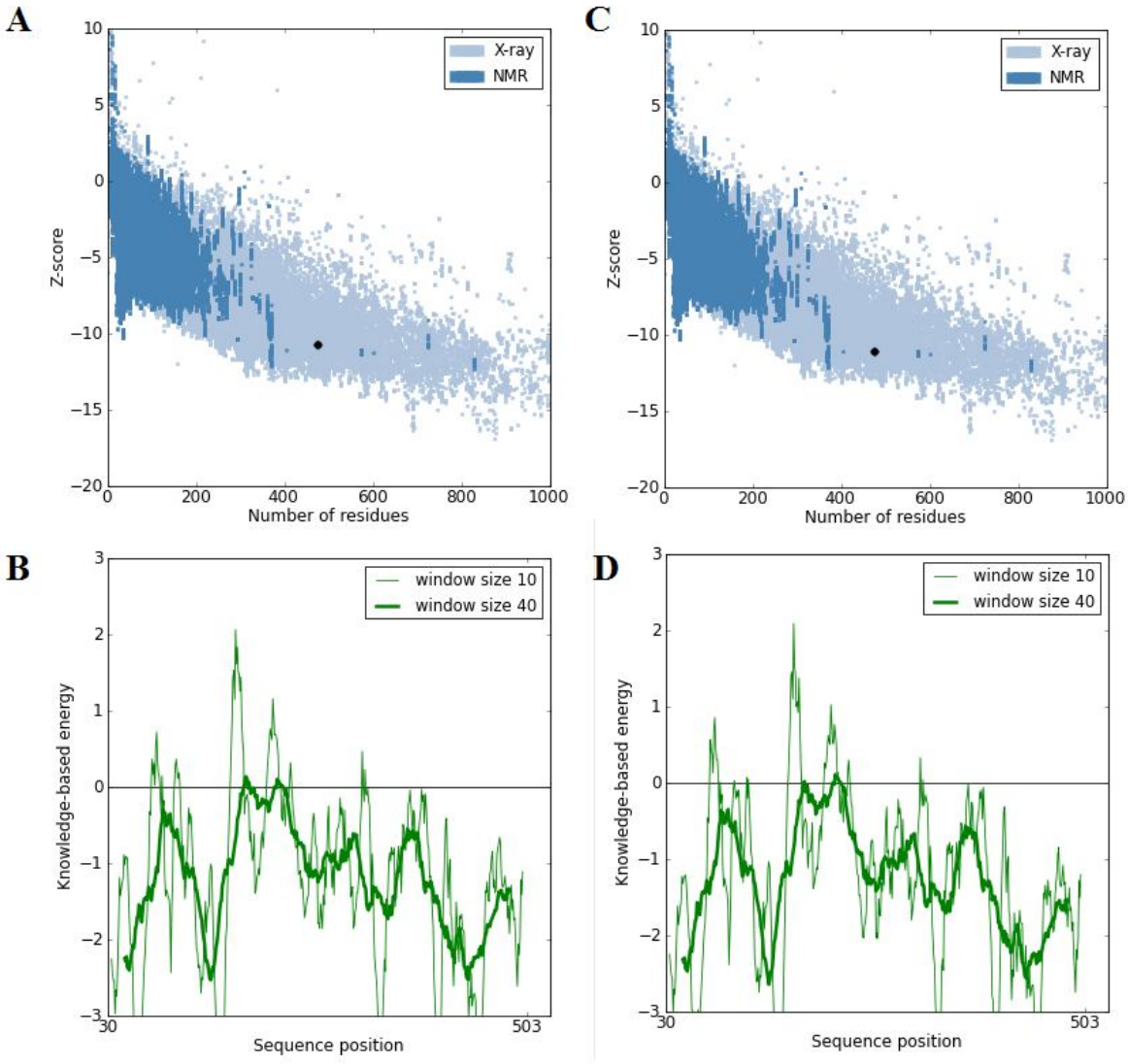

Figure 3. ProSA protein folding energy profiles of FTO_W (A and B) and FTO_{R322Q} (C and D). The Z-scores of the predicted models (A and C) were in the permitted range (Black dot), when compared with the high resolution X-ray and NMR models. Similarly, local energy profiles of both models (B and D) were also in the permitted range.

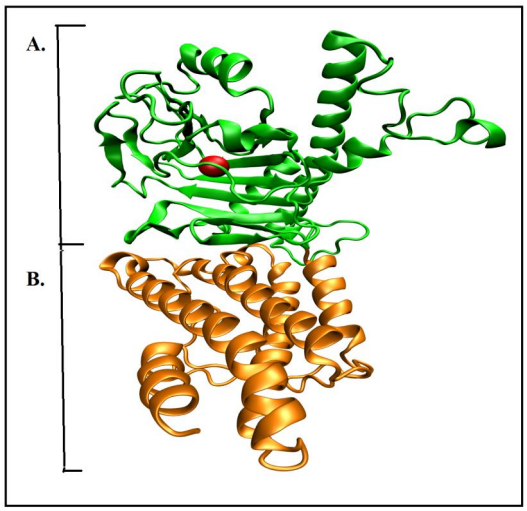

Figure 4. Ribbon representation of N-terminal and C-terminal of FTO enzyme. N-terminal (A) catalyzes 2-OG into Succinate and CO_2. Catalysis of 2-OG is coupled with the demethylase activity which resides in the C - terminal (B) of the FTO. Fe(II) is present in the N-terminal and represented as a red ball.

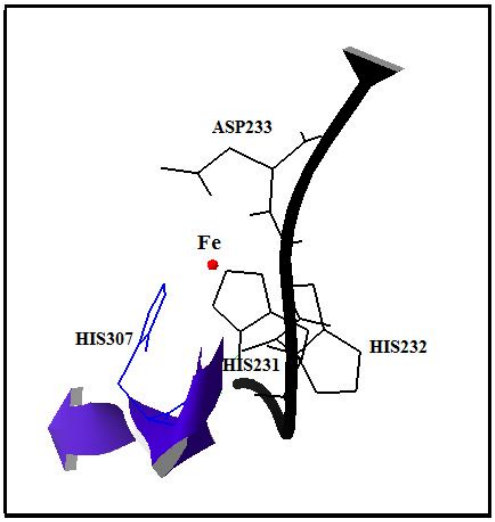

Figure 5. Fe(II) carrying motifs are shown in ribbon representation. Loop L10 is represented in black and Sheet S13 is represented in Blue with the side chains of residues HIS231, ASP233 and HIS307 interacting with Fe(II) . Together these residues act as a scaffold to capture metal ion in a triangular geometry.

Table 1. Ramachandran plot, bond length, bond angle and degree of normality analysis

Structure	Ramachandran plot distribution [Allowed (A), Generously allowed (GA), Disallowed (D)]	% Bond lengths (L) and % angles (A) within limits	Overall G-factor
4IE6 (Template)	A=99.5%, GA=0.4%, D=0.1%	L= 100% A= 99.4%	0.10 (allowed)
FTO_w	A=97.7%, GA=1.4%, D=0.9	L=100% A=96.5%	-0.11 (allowed)
FTO_{R322Q}	A=97.4%, GA=1.6%, D=0.9%D	L=100% A=96.4%	-0.09 (allowed)

Table 2. Surface analysis of FTO protein using PocketFinder

Pocket name	Volume angstrom (percent volume relative to protein)	Residues
P1	473 (1.02)	VAL83, ILE85, THR92, PRO93, VAL94, ARG96, TYR106, TYR108, LEU109, LEU113, LEU203, ASN205, MET207, ALA227, VAL228, SER229, HIS231, ASP233, GLU234, VAL244, TRP270, TYR295, MET297, HIS307, VAL309, ARG316, SER318, THR320, ARG322
P2	229 (0.49)	LYS48, LEU154, LEU157, ALA158, GLU161, LEU170, CYS171, ASP175, ASP192, ILE193, ARG196, ARG239, ASP299
P3	96 (0.21)	LYS162, CYS171, MET172, ALA174, ASP175, PHE176, GLN188, GLU190
P4	83 (0.18)	PHE79, ASP81, ARG96, ILE97, LEU98, LEU113, PHE114, VAL116, ARG388, TYR389, CYS392
P5	81 (0.17)	ARG84, LYS107, ASN110, GLN381, LEU448, GLU451, TRP452, ARG455

Molecular Docking analysis of FTO_w and FTO_{R322Q}

Molecular docking studies revealed interesting structural interaction differences between 2-OG and FTO_w and FTO_{R322Q}. In FTO_w side chains of Thr92, Arg96, Tyr106, Tyr108, Leu203, Asn205, Val228, His231, Asp233, Thr320 and Arg322 interact with 2-OG (Fig. 6A). Specially, O atoms of carboxylate groups make hydrogen bonds with Arg96 and Arg322. Residues Arg96 and Arg322 are also involved in developing a positively charged environment in the ligand-binding site which also assists in producing stable interactions with double negatively charged 2-OG (Fig. 7A). Analysis of docking results from FTO_{R322Q} mutant showed that replacement of Arg at position 322 with Gln, a polar, hydrophilic amino acid with shorter side chain, results lacking in hydrogen bonds and electrostatic interactions with 2-OG (Fig. 6B). In addition to this, Gln322 substitution also resulted in disruption of the positively charged polar environment created by Arg96 and Arg322 combined (Fig. 7B). Besides roughly double stability of binding energy and ligand efficiency, electrostatic interaction of FTO_w with 2-OG was 3 times more stable as compared to FTO_{R322Q} (Table 3).

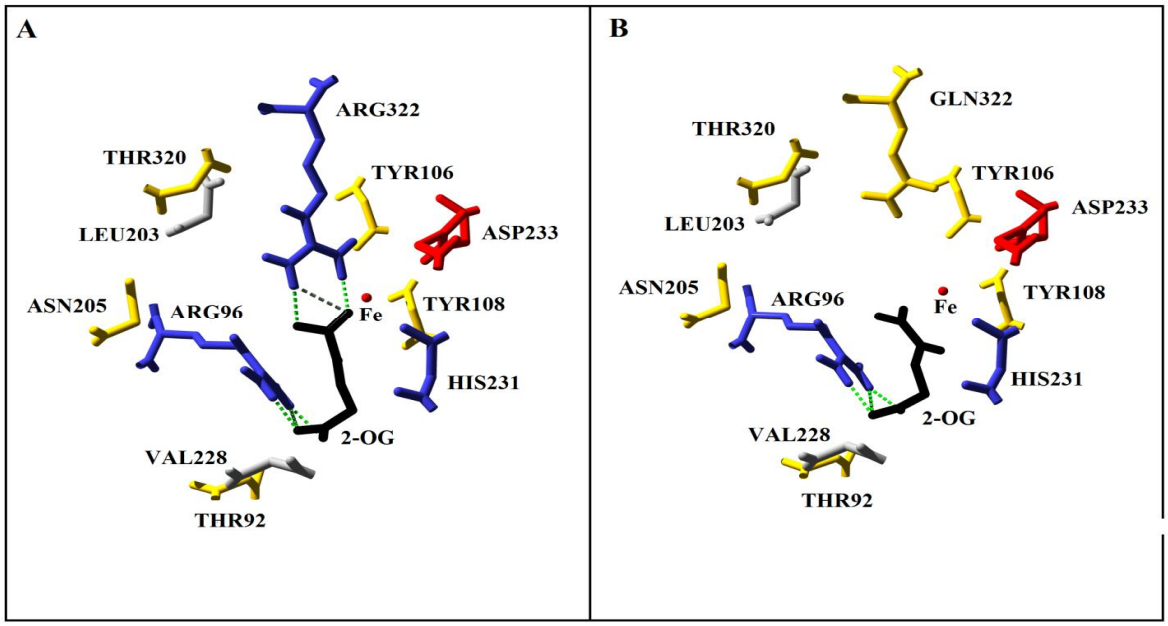

Figure 6. Interactional differences of 2-OG with FTO_w and FTO_{R322Q}. (A) FTO_w successfully develops hydrogen bonding and hydrostatic interactions with 2-OG. Residues ARG96 and ARG322 are shown with side chains making hydrogen bonds and hydrostatic interactions with 2-OG while the backbones of remaining amino acid residues are shown. (B) FTO_{R322Q} could make similar molecular-molecular interactions in terms of Arg96, but the short side chain of Gln322 is unable to develop contacts with the 2-OG. Hence 2-OG is in a relatively extended configuration in FTO_w which makes its bonds to be more accessible to the side chains of other amino acid residues.

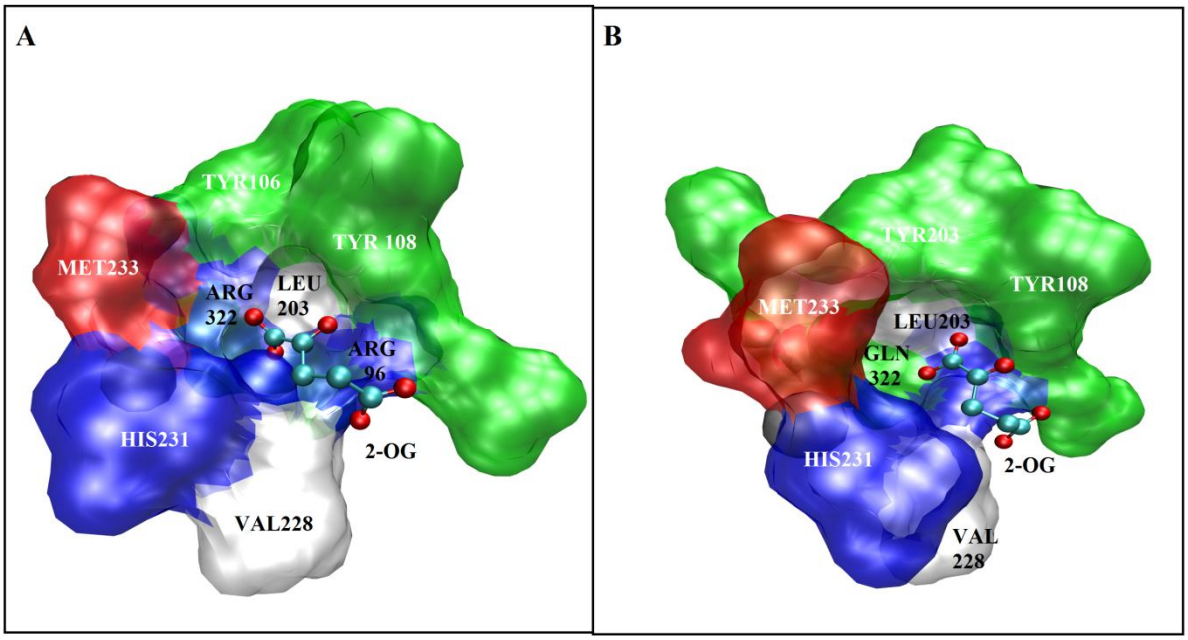

Figure 7. van der Waals surface representation of (A) FTO_w and (B) FTO_{R322Q}. In this figure, hydrophilic positively charged amino acids are shown blue; non-polar and hydrophobic amino acids are represented white; polar hydrophilic amino acids are shown in green; and non-polar hydrophobic amino acids are shown red. Positively charged, hydrophilic patch generated by Arg96 and Arg322 (A) makes stable hydrostatic interactions with 2-OG which stretches the backbone of substrate.

Table 3. Comparison of Molecular dockings of FTO_w and FTO_{R322Q} with 2-OG

	Binding Energy	Ligand Efficiency	Inhibition Constant (mM)	Intermolecular Energy	van der Waals, Hydrogen bonds energy	Electrostatic Energy
FTO_w	-3.72	-0.37	1.89	-4.98	-3.66	-1.33
FTO_{R322Q}	-1.89	-0.19	41.19	-3.19	-2.75	-0.45

CONCLUSION

Arg322 is a highly conserved residue which resides in the catalytic site of the FTO. Arg322 and Arg96 develop a positively charged surface with which 2-OG interacts by making hydrogen bonds and electrostatic interactions. On the contrary, substitution of Gln322 disrupted the positively charged environment and Gln short and neutral side chain failed to produce stable interactions with the 2-OG.

REFERENCES

1. Sanchez-Pulido L, Andrade-Navarro MA (1970). The FTO (fat mass and obesity associated) gene codes for a novel member of the non-heme dioxygenase superfamily. BMC Biochem. 8(1):23.
2. Gerken T, Girard CA, Tung YCL et al (2007) The obesity-associated FTO gene encodes a 2-oxoglutarate-dependent nucleic acid demethylase. Science. 318:1469-1472.
3. Meyre D, Proulx K, Kawagoe-Takaki H et al (2010). Prevalence of loss-of-function FTO mutations in lean and obese individuals. Diabetes. 59(1): 311-318.
4. Scuteri A, Sanna S, Chen WM et al (2007). Genome-wide association scan shows genetic variants in the FTO gene are associated with obesity-related traits, PLoS Genet. 3(7):115-121.
5. McMurray F, Church CD, Larder R et al (2013). Adult onset global loss of the Fto gene alters body composition and metabolism in the mouse. PLoS Genet., 9(1):e1003166.
6. Loenarz C., Schofield CJ (2008). Expanding chemical biology of 2-oxoglutarate oxygenases. Nat. Chem. Biol. 4(3):152-156.
7. Fredriksson R, Hägglund M, Olszewski PK et al (2008). The obesity gene, FTO, is of ancient origin, up-regulated during food deprivation and expressed in neurons of feeding-related nuclei of the brain. Endocrinol., 149(5) :2062-2071.
8. Gulati P, Yeo GS (2013). The biology of FTO: from nucleic acid demethylase to amino acid sensor. Diabetologia. 56(10):1-9.
9. Aik W, Demetriades M, Hamdan MK et al (2013). Structural Basis for Inhibition of the Fat Mass and Obesity Associated Protein (FTO). J. Med. Chem. 56(9):3680-3688.
10. Hess ME, Hess S, Meyer KD et al(2013). The fat mass and obesity associated gene (Fto) regulates activity of the dopaminergic midbrain circuitry. Nature Neurosci. 16(8):1042-1048.
11. Schwede T, Kopp J, Guex N et al (2003). SWISS-MODEL: an automated protein homology-modeling server. Nucl. Acid. Resear., 31(13):3381-3385.
12. Chenna R., Sugawara H., Koike T. et al (2003). Multiple sequence alignment with the Clustal series of programs, Nucl. Acid. Resear. 31:3497-3500.
13. Söding J, Biegert A, Lupas AN et al.(2005). The HHpred interactive server for protein homology detection and structure prediction, Nucl. Acid. Resear. 33(2):244-248.
14. Marti-Renom MA, Stuart A, Fiser A et al (2000). Comparative protein structure modeling of genes and genomes, Ann. Rev. Biophy. Biomol. Struc. 29:291-325.
15. Guex N., Peitsch MC (1997). SWISS - MODEL and the Swiss - Pdb Viewer: an environment for comparative protein modeling. Electrophoresis. 18(15):2714-2723.
16. Schuler LD, Daura X, Van Gunsteren WF (2001). An improved GROMOS96 force field for aliphatic hydrocarbons in the condensed phase. J. Comput. Chem. 22(11):1205-1218.
17. Humphrey W, Dalke A, Schulten K (1996). VMD: visual molecular dynamics, J. Mol. Graph. 14(1):33-38.
18. Laskowski RA, MacArthur MW, Moss DS et al (1993). PROCHECK: a program to check the stereochemical quality of protein structures, J. Appl. Crystallograph.. 26(2):283-291.
19. Güntert P, Dötsch V, Wider G et al (1992.) Processing of multi-dimensional NMR data with the new software PROSA. J. Biomol. 2(6):619-629.
20. Hendlich M, Rippmann F, Barnickel G. (1997) LIGSITE: automatic and efficient detection of potential small molecule-

binding sites in proteins. J. Molecul. Graph. Model. 15(6):359-363.
21. Goodsell DS, Morris GM, Olson AJ (1996). Automated docking of flexible ligands: applications of AutoDock. J. Molecul. Recog. 9(1):1-5.
22. Morris GM, Goodsell DS, Huey R et al (1996) Distributed automated docking of flexible ligands to proteins: parallel applications of AutoDock 2.4. J. Comp.-Aid. Molecul. Design. 10(4):293-304.
23. Thomsen R, Christensen MH (2006) MolDock: a new technique for high-accuracy molecular docking. J Med. Chem. 49(11):3315-3321.

*****