International Journal of Science and Research (IJSR)
ISSN (Online): 2319-7064
Index Copernicus Value (2013): 6.14 | Impact Factor (2013): 4.438

In Silico Sequence Analysis, Structure Prediction and Evaluation of Fatty Acid Synthase Protein (FAS) of *Gallus gallus* (Chicken) by Homology Modelling

Aman Singh Bhalla[1], Roma Lal[2]

[1]Research Assistant, Allele Life Sciences Pvt. Ltd, Noida, U.P., India
[2]Research Associate, Allele Life Sciences Pvt. Ltd, Noida, U.P., India

Abstract: ***FAS (Fatty acid synthase) an enzyme catalysing synthesis of fatty acid. The protein contains two subunits,namely A & B. Fatty acid synthase protein plays a key role in fat deposition in chicken as it was found that the protein is directly correlated with fat content of chicken. Administration of glucocorticoids changes the expression of FAS protein.3D structure of FAS protein was not available on RCSB Protein Databank. FAS protein of Human was available in the database which was selected as template for homology modelling.The present study deals with sequence analysis, secondary structure prediction and homology modelling of Fatty acid synthase of chicken. The secondary structure of protein shows that the protein is chiefly of α-helical structure followed by coils and extended strand. 3D structure of protein was predicted using SwissModel. The predicted structure was evaluated by Procheck, Errat, Ramachandran Plot, QMean, Z-Score using SAVES, an online structure evaluation server (http://services.mbi.ucla.edu/SAVES/).***

Keyword: Fatty acid synthase, secondary structure, 3D structure, Homology modelling

1. Introduction

Fatty acid synthase (FAS EC 2.3.1.85) is a multifunctional enzyme. It plays a key role in fatty acid chain synthesis. The de-novo synthesis of fatty acid occurs mainly in hepatic cell in birds. FAS protein consists of two identical subunits,namely A & B subunit. Each subunit consist of seven functional domains: Dehydratase (DH), Enoyl reductase (ER), β-Keto acyl reductase(KR), acyl carrier protein (ACP) &Thioesters(TE) [1]. Among all thioesters, TE domain plays an important role in regulating the final chain length of the product [2].

It was reported that fasting in animals or birds reduced the amount of FAS and then refeeding restored the amount of FAS activity [3,4]. Fasting in broiler chicken significantly decreases mRNA level of FAS gene. It was also discovered that administration of glucocorticoids affected the FASN enzyme by regulating its gene expression, though change in expression was of temporary duration and did not alter the basic activity of the enzyme. FAS plays an important role in regulating chicken body fat contentas it was discovered that fat percentage increased from youth to middle age[5].

3D structure information of FAS protein will help us to know the role of protein in fat content of chicken and interaction of their domains with their ligands. 3D structure prediction of protein require X-ray crystallography and NMR spectroscopy which is very time consuming, tedious method and generate a large amount of data creating a gap between available sequences and solved structure. In silico method of predicting 3D structure reduces this gap.

2. Material and Methods

Protein sequence for FASN gene (Accession No.: NP_990486.2) was downloaded from NCBI database.

Primary Structure analysis
Physicochemical properties like Molecular weight, Theoretical pI, % composition of amino acids, Extinction coefficient[6], estimated half life, Instability Index[7], Aliphatic Index[8], Grand Average of Hydropathicity(GRAVY)[9] of linear sequence of FASN gene was calculated using ExpasyProtPARAM server (http://expasy.org/cgi-bin/protpraram).

Secondary Structure prediction
Secondary structure of protein sequence was predicted using GORIV and SOPMA. GOR IV method is based on information theory and was developed by J.Garnier, D.Osguthorpe and B.Robson (J.Mol.Biol.120, 97, 1978).SOPMA (Self-Optimized Prediction Method with Alignment) is an improvement of SOPM method. These methods are based on the homologue method of Levin et al.

Protein motif identification
Motif of target protein was identified by FingerPRINTScan, an online tool (http://www.ebi.ac.uk/Tools/pfa/fingerprintscan/) that identifies the closest matching PRINTS sequence and motif fingerprints in a protein sequence.

Homology Modelling
Computational prediction of 3D structure of protein is based on Ab-Initio, Threading or Homology modelling. If sequence similarity search of target sequence is more than 60%, then homology modelling can be useful for 3D structure prediction of target protein. Homology modelingwas done using a template sequence whose structure has been solved by either X-ray diffraction or NMR technology. SwissModel is an online tool for 3D structure prediction based on Homology modeling.
Steps followed by Swiss model for homology modeling:

Template was searched against Swiss-model template library (SMTL) using Blast and HHBlits algorithms. A total of 266 templates were found. The templates with the highest quality have then been selected for model building. Models are built based on the target-template alignment. Model quality was estimated by assessing the QMEAN score and Z-Score.

Structure visualization

Modelled 3D structure of FAS protein was visualized by Pymol software. Pymol is open source software that can produce high quality 3D images of biomolecules.

3. Result

NCBI's BLAST tool for protein sequence was used to search similar sequences. The result predicted 100% similarity to Fatty Acid Synthase [*Homo sapiens*]. Some of the similar sequences with E value 0.00 are in table 1.

Table 1: BLAST result of FAS protein

Accession No.	Description	Total Score	E value
2VZ8 A	Chain A, Crystal structure of Mammalian Fatty Acid Synthase	3433	0.00
3HHDA	Chain A, Structure Of The Human Fatty Acid Synthase Ks-Mat Didomain As A Framework For Inhibitor Design.	1476	0.00
2JFK A	Chain A, Structure Of The Mat Domain Of Human Fas With Malonyl-Coa	606	0.00
2JFD A	Chain A, Structure Of The Mat DomainOf Human Fas	597	0.00

Primary Structure prediction

Primary structure of FAS protein was predicted and its physicochemical properties were analyzed by using Expasy'sProtParam server (http://expasy.org/cgi-bin/protparam). Result showed that FAS protein has 2512 amino acid residues and the estimated molecular weight is 274774.4. The calculated pI is 5.92 which shows that the protein is acidic in nature. The maximum and minimum number of amino acid present in sequence is Leucine (11.8%) and Trp (1.2%) respectively. The total number of positively charged residues (Arg + Lys) is 230 and total number of negatively charged residues (Asp + Glu) is 273. The estimated instability index of protein is 41.89 whichpredictsthat protein stability is low in test tube and thus the FAS protein is unstable. High aliphatic index (94.90) of protein predicts its stability under wide range of temperature. The negative value of Grand Average of Hydropathicity (GRAVY) (-0.093) indicates that protein is non-polar, hydrophilic (GRAVY typical value for hydrophilic protein is <-1) and better interaction of the protein with water.

Secondary Structure prediction

The secondary structure of FAS protein was predicted using GOR IV and SOPMA and the result showed that the sequence is mainly composed of Alpha helix and Beta sheets. The comparative analysis by both GOR IV and SOPMA is given in Table 2. From the result,it is predicted that the protein is chiefly composed of alpha helix (40.25%) followed by random coil (30.85%) and extended strand (18.51%). The result is graphically represented in Figure 1.

Table 2: Secondary structure prediction of FAS protein by GOR IV and SOPMA

Secondary Structure	GOR IV	SOPMA
Alpha helix (Hh)	35.47%	40.25%
3_{10} helix (Gg)	0.00%	0.00%
Pi helix (Ii)	0.00%	0.00%
Beta bridge	0.00%	0.00%
Extended strand (Ee)	18.99%	18.51%
Beta turn (Tt)	0.00%	0.00%
Bend region (Ss)	0.00%	0.00%
Random coil (Cc)	45.54%	30.85%
Ambiguous states (?)	0.00%	0.00%
Other states	0.00%	0.00%
Sequence length	2512bp	2512bp

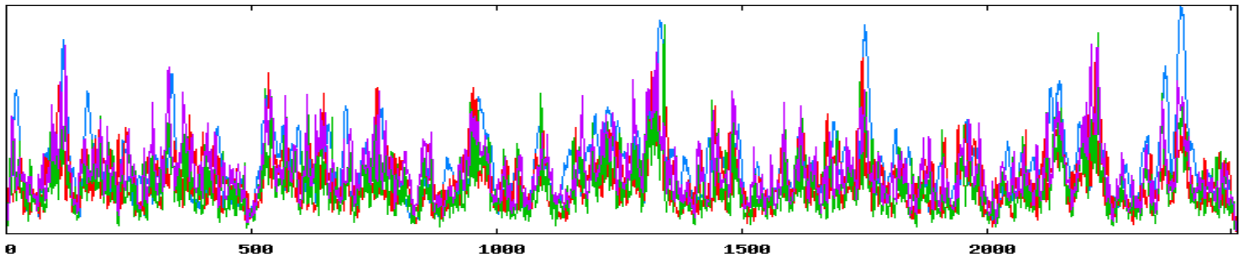


Figure 1: Graphical representation of secondary structure of FAS protein by SOPMA
helix sheet coil

Protein motif identification

Motif of FAS protein was identified by FingerPRINTScan tool. The result predicted 5 FingerPrints with number of

motifs as 2 for each. The result of FingerPRINTScan is tabulated in table 3.

Table 3: FingerPRINTScan result of FAS protein

FingerPRINT	No. Of Motifs
TROPOMYOSIN	2
AFETOPROTEIN	2
FADPNR	2
MELATONIN1XR	2
ZETATUBULIN	2

Tertiary Structure prediction

Tertiary structure or 3D structure prediction was done using Swiss model Automated mode for homology modeling. Swiss model server searched for the solved templates with similar sequences, the result of top five best templates are given in Table 4. Best templates were aligned with target amino acids sequence. Templates with best E-value, percentage similarities and maximum number of query sequence covered were selectedfor homology modeling.

Table 4: Template search result for FAS protein by SwissModel

10	Seq Identity	Oligo-state	Found by	Method	Resolution	Seq Similarity	Coverage	Description
2vz9.1.A	66	homo-dimer	BLAST	X-ray	3.30Å	0.5	1	FATTY ACID SYNTHASE
2vz9.1.B	66	homo-dimer	BLAST	X-ray	3.30Å	0.5	1	FATTY ACID SYNTHASE
2vz9.1.A	65.76	homo-dimer	HHblits	X-ray	3.30Å	0.49	1	FATTY ACID SYNTHASE
2vz9.1.B	65.76	homo-dimer	HHblits	X-ray	3.30Å	0.49	1	FATTY ACID SYNTHASE
3hhd.1.A	71.86	homo-dimer	BLAST	X-ray	2.15Å	0.52	0.38	Fatty acid synthase

Template 2vz9.1.B was selected for homology modelling. Generated model was subjected to evaluation programs Procheck, Errata and verify3D. The procheck result showsthat 76 residues out of 2116 amino acids were in disallowed region.Ramachandran plot of the predicted model showed 89.3% residues in most favoured region, 9.2% residues in additional allowed region, 0.9% residues in generously allowed region, 11 residues or 0.6% residues in disallowed region Figure 2.Errat result shows the overall quality factor of 69.282.The obtained 3D structure of FAS protein was visualized by Pymol (Refer Figure 4).

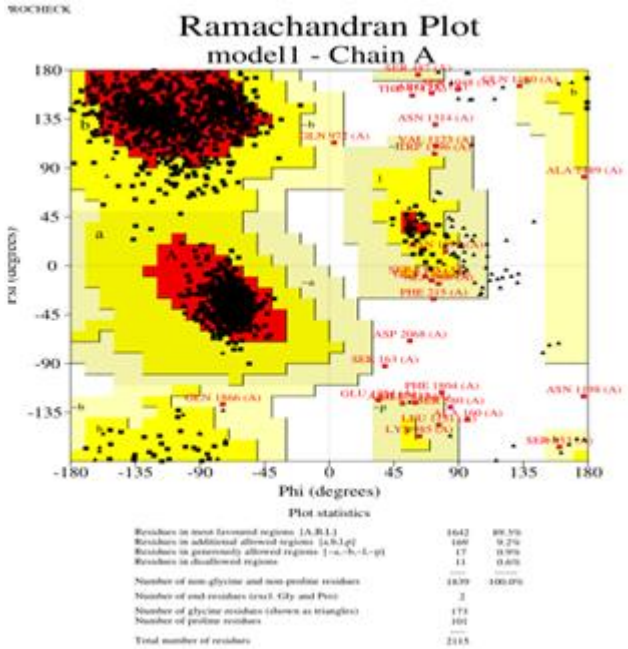
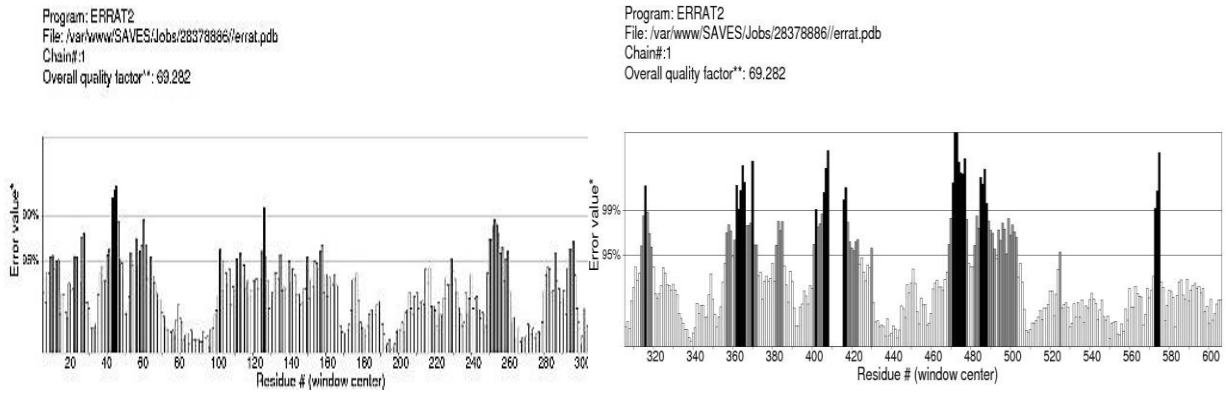


Figure 2: Graphical representation of Ramachandran Plot by SAVES



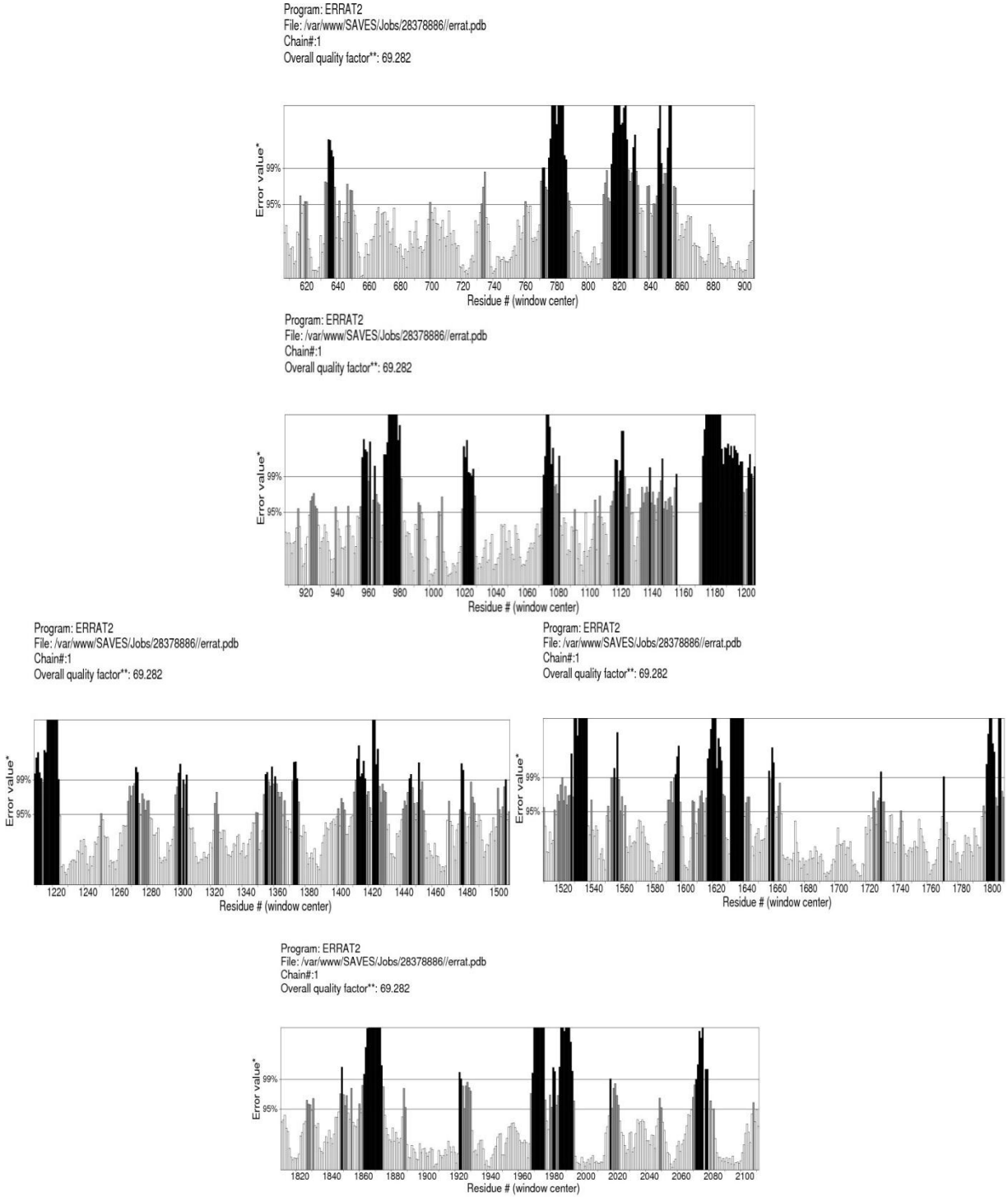


Figure 3: ERRAT result for Predicted model

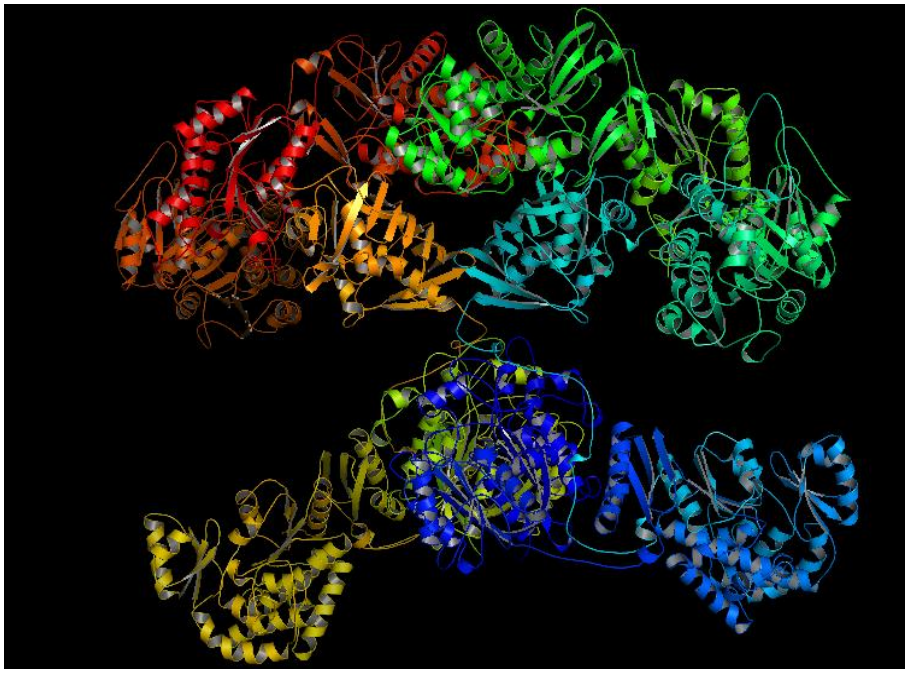


Figure 4: Three Dimensional structure of FAS protein by Pymol

4. Conclusion

Present study deals with the *insilico* sequence and structure analysis of FAS protein of *Gallus gallus* by various tools and softwares. Based on the finding, it could be concluded that the protein is non polar, hydrophilic in nature. Protein has 5 FingerPRINTS with 2 motifs. The predicted structure can be used to know more about interaction of FAS protein domains with glucocorticoids or ligands and their role in fat content of chicken.

References

[1] Smith, S.,Witkowski, A., Joshi, A.K., *Prog Lipid Res*(**2003**) 42, 289-317
[2] Zhang, S., Knight, T.J., Reecy, J.M., Beitz DC., *Anim Genet*(**2008**) 39, 62-70
[3] Bloch, K., Vance, D., *Ann. Rev. Biochem.,* (**1977**) 46, 263-298
[4] Vdpe, J.J., Vagelos, E.R., *Physiol. Rev.,* (**1976**) 56, 339-417
[5] Zhigang Song, Lei LiuYunshuang Yue, Hongchao Jiao, Hai Lin, ArdashirSheikhahmadi, Nadia Everaert, Eddy Decuypere, Johan Buyse ,J. Kyte, RF. Doolottle, *J. Mo Biol* (**1982**) 157, 105- 132.
[6] SC. Gill, PH.Von Hippel, *Anal. Biochem*, (**1989**) 189, 319.
[7] K.Guruprasad, BVP. Reddy, MW. Pandit, *Prot. Eng.* (**1990**) 4, 155 – 161.
[8] AJ. Ikai, *J. Biochem*, (**1980**) 88, 1895-1898.
[9] J. Kyte, RF. Doolottle, *J. Mo Biol.*(**1982**) 157, 105- 132.