ISSN: 0976-3104
REGULAR ISSUE
Kesharwani et al.

THE IIOAB JOURNAL

RESEARCH ARTICLE OPEN ACCESS

DESIGNING OF A MODEL FOR HUMAN AMELOGENIN FOR PREDICTING ITS ROLE IN MINERALIZATION DURING SYNTHESIS OF TOOTH ENAMEL, AN IN SILICO APPROACH

Rajesh Kumar Kesharwani, Vandhana Mahajan, and Krishna Misra*

Division of Applied Science and Indo-Russian Center for Biotechnology, Indian Institute of Information Technology, Deoghat Jhalwa campus, Allahabad-211012, INDIA

ABSTRACT

The Amelogenin protein found in developing tooth enamel is believed to organize enamel crystals during tooth development by acting as a crystallization centre for mineralization. The motive of the current study was to predict the 3-D structure of human amelogenin X and mechanism of its self-aggregation using computational methods. Homology modeling followed by threading and ab-initio methods were used for structure prediction to obtain a protein model suitable to study the self-aggregation of amelogenins in-silico. The model obtained from HHPred was selected, refined, and optimized using different bioinformatics servers and softwares. On analysis, the model gave acceptable Procheck, Verify-3D, Errat and ProQ results. The predicted model could be validated by studying its intermolecular interaction to form nanospheres which is in agreement with literature reports. This predicted protein model can be used further to study the protein-mineral-protein interactions taking place in the process of amelogenesis for hypothesizing on new restorative materials.

Received on: 3rd-March-2011
Revised on: 1st -April-2011
Accepted on: 18th -May-2011
Published on: 1st -Sept-2011

KEY WORDS

Tooth enamel; In silico; Homology Modeling; ramachandran plot; amelogenin

Corresponding author: Email: krishnamisra@hotmail.com; Tel: +91-9415247579, 261486; Fax: +91-532-2461376/2608469

[I] INTRODUCTION

Biological materials like enamel exhibit intricate architecture and outstanding physical properties, unobtainable by traditional methods of materials synthesis. As a result, despite enormous effort no ideally biocompatible artificial tooth enamel material has so far been developed [1]. In case the mechanism of the natural process of enamel formation is understood, better materials based on the same principle can be designed which would have properties closer to the natural enamel. Enamel is composed of proteins amelogenin, ameloblastin, enamelin and tuftelin. Of these more than 90% is constituted by amelogenin. Thus, it can be considered as major protein responsible for enamel formation and has been studied under the present work.

The protein has been found to exist in two major isoforms - AMELX (Amelogenin X isoform) and AMELY (Amelogenin Y isoform), encoded by AMGX and AMGY genes present on the short arms of the human X and Y chromosomes [2]. Amelogenin X was chosen for this study because it is reported that mutations in AMELX may cause Amelogenesis imperfecta, a disorder of tooth enamel development in which teeth may become usually small, discolored, pitted or grooved, and prone to rapid wear breakage [3]. Also, AMELX is the structural constituent of tooth enamel and is involved in hydroxyapatite binding and Amelogenin-Amelogenin protein binding [2].

The matrix-mediated enamel biomineralization involves secretion of the enamel specific amelogenin proteins that through self-assembly into nanosphere structures provide the framework within which the initial enamel crystallites are formed [4]. Amelogenin protein is found in the developing tooth enamel and belongs to the family of extracellular matrix (ECM) proteins. Its function is believed to be in organizing enamel crystals during tooth development. It has been found that amelogenin is primarily hydrophobic, rich in proline (25%), glutamine (14%), leucine (9%), and histidine (7%), which altogether account for more than 50% of the amino acids. The amino acid sequence can be divided into 3 domains, based on differences in composition of amino acids [1]:

a). The N-terminal domain with 45-amino-acids is rich in Tyr (TRAP),

THE IIOAB JOURNAL
ISSN: 0976-3104

b). The hydrophobic central region, primarily composed of Xxx-Yyy-Pro- repeat motifs (where Xxx and Yyy are primarily Gln) and
c). The 11-Amino-acid-long C-terminal domain, being charged and hydrophilic, contains a number of potential electrostatic binding sites—acidic aspartic acid and glutamic acid and basic lysine and arginine.

The entire sequence shows high levels of homology in short conserved regions of protein (motifs) as determined by the MEME motif discovery method. The list includes the following motifs: DKTKREEVD, SYGYEPMGGW, GYINFS/LYE, LKWYQSMIR, MGTWILFACLL LLGAAF, DLPLEAW, MMPVPGQ/HHSMTPTQHHQPN, LHHQIIPVL/VSQ, S/AHA/ TLQPHHHI/ LPV/MVPAQQPV, QQPFQPQ [5].

A phylogenetic tree was constructed using the amelogenin sequences of human, pig, horse, goat, bovine, guinea pig, hamster, mouse, rat, crocodile, xenopus, snake, dog, platypus, porcupine etc., which showed closest homology between pig and human amelogenin sequences [5]. Due to this fact most of the study related to this protein has been done using pig amelogenin, which is readily available from slaughterhouses.

Amelogenin is a unique biomineralization protein because it self-assembles to form supramolecular structures called "Nanospheres"- spherical aggregates of monomers that are 20−60 nm in diameter, consisting of spherical substructures that are 4-8 nm in diameter [1, 6]. Amelogenin nanosphere assembly proceeds through intermediate structures (perhaps represented in vivo by "stippled material") of some 4-5 nm hydrodynamic radius, and is computed to comprise 4-6 amelogenin monomers [7]. Nanospheres are best described by a dense, predominantly hydrophobic, protein core, surrounded by a loose shell comprised of the negatively charged hydrophilic C-terminus that is exposed to the aqueous environment. It has been suggested that this negatively charged surface prevents the aggregation of nanospheres at room temperature and below. Conservation of above mentioned functional sites may indicate a role for these residues in nanosphere formation [1].

Biomineralization is regulated by an interplay between hydrophobic and hydrophilic molecules, whereby the hydrophobic molecules provide a skeletal or space-filling structure and the hydrophilic (acidic) molecules are involved in the regulation of crystal nucleation and growth. Thus, amelogenin possesses both properties required for biomineralization of enamel [4].

Two key proteinases have been identified within the enamel matrix - MMP20 (Enamelysin) and Kallikrein (KLK4), along with a serine proteinase ESMP1. MMP-20 is expressed during the secretory stage, when the crystallites are growing predominantly in length. As MMP-20 slowly degrades these proteins, the crystallites grow in width and thickness so that there is a net replacement of protein by minerals [8].

KLK4 is a protease with broad target sequence specificity that can degrade enamel proteins. This protein is secreted during the transition and maturation stages of amelogenesis, immediately preceding the point where the quantity of enamel proteins in the matrix drops precipitously. KLK4 functions during the later stages of dental enamel formation to degrade the accumulated enamel matrix, allowing the crystals to grow in width and thickness until adjacent crystals contact. This activity is essential for hardening of the enamel [9,10].

Thus if this natural process of enamel formation can be simulated and additionally synthetic material based on parallel process can be generated, it may have desirable properties of a biocompatible restorative material. For working on this idea, our first objective was to predict the 3-D structure of the protein amelogenin X and study its role in enamel formation.
.

[II] MATERIALS AND METHODS

2.1. NCBI sequences

The protein sequence of amelogenin X (Genbank id: AAA51717.1) zz was downloaded from the NCBI database, 191 amino acids are found in total size of the sequence and the source of this protein is Homo sapiens in origin [11].

2.2. NCBI-Blast

The sequence obtained from the NCBI Reference Sequences database was submitted to NCBI-Blast and searched against the Protein Data Bank (PDB) database to extract suitable structural templates [12].

Since, on doing BLAST of query sequence with PDB database, none of the matches obtained had even up to 30% similarity score, it could be concluded that homology modeling cannot be performed as no appropriate template could be found. Therefore, secondary structure predictions (2.3), along with threading and ab-initio predictions (2.4) were used to obtain the suitable 3-D amelogenin X protein model.

2.3. Prediction of the 2D structure of Amelogenin X using the following list of the bioinformatics tools and servers

1. Genamics Expression (http://genamics.com/expression/strucpred.htm) - It is a windows application program for DNA and protein sequence analysis. It offers an interface to a large array of secondary structure prediction algorithms including DPM, DSC, GOR 4, HNN, MLRC, PHD, PREDATOR, SIMPA96, and SOPMA [13].

2. PSIPRED (http://bioinf.cs.ucl.ac.uk/psipred/psiform.html) - It is a protein structure prediction server that allows users to input a protein sequence, do a prediction and get the output of the prediction textually via e-mail as well as graphically via the web [14].

3. JPRED (http://www.compbio.dundee.ac.uk/www-jpred/) - It is not an individual routine but assembles various scores for secondary structure. The consensus of structure prediction algorithms is obtained as result [15].

4. GOR4 (http://npsa-pbil.ibcp.fr/cgi-in/npsa_automat.pl? page=npsa_gor4.html) - It is the fourth version of GOR. It is based on information theory and does not have a defined decision constant, i.e., the standard against which the values of pair frequencies have to be compared. GOR

THE IIOAB JOURNAL
ISSN: 0976-3104

IV employs all likely pair frequencies inside the window of 17 amino acid residues. A mean accuracy of 64.4% has been obtained for a three state prediction (Q3) after carrying out cross validation for a database having 267 proteins [16].

5. HNN (http://npsa-pbil.ibcp.fr/cgi-bin/npsa_automat.pl? page=npsa_nn.html) - Hierarchical Neural Network prediction method is constituted by two networks. One of these is a sequence-to-structure network and the other one is a structure-to-structure network. This makes only local information as the basis of prediction [17].

6. PROF pred (http://www.aber.ac.uk/~phiwww/prof/) - It is a service provided by Predict Protein serving sequence analysis, structure prediction and function prediction [18].

7. 3D-Jigsaw (http://bmm.cancerresearchuk.org/~3djigsaw/) – It is an automated, comparative modelling server which may be used for predicting the structure and/or function of a protein sequence [19].

8. PORTER (http://distill.ucd.ie/porter/) – It is a server used for protein secondary structure prediction. A collection of 45 bidirectional recurrent neural networks (BRNNs) forms the basis of its algorithm. It has been tested by a rigidly accurate 5-fold cross validation method and achieves 79% correct classification on the "hard" CASP 3-class assignment [20].

9. PORTER+ (http://distill.ucd.ie/porter+/) – It is a server for the prediction of local structural motifs. The motifs are constructed by using multidimensional scaling (MDS) and clustering for pair-wise angular distances of multiple Φ and Ψ dihedral angle values compiled from high-resolution protein structures. This method allowed visualization of the protein backbone fragments in a scaled down 3D conformational space from the earlier multiple dimensions and led to the recognition of a few conformational clusters which are populated by real or validated backbones. In Porter+ these clusters are mapped into a conformational alphabet of 14 letters which represent structural motifs for tetra-peptides. Porter+'s architecture is similar to Porter's. It classifies nearly 60% of residues as the right structural motif and around 30% to be above a base-line statistical predictor [20].

2.4. Prediction of 3D structure of amelogenin X using following bioinformatics servers and softwares

1. I-Tasser (http://zhang.bioinformatics.ku.edu/I-TASSER/) – It is an online service for predicting the protein structure and function. Based on multiple-threading alignments, models are constructed by LOMETS and I-TASSER simulations [21].

2. LOMETS (http://zhang.bioinformatics.ku.edu/LOMETS/) - Local Meta-Threading-Server is an internet service for protein structure prediction. It returns 3D models by assembling consensus target-to-template alignments from nine locally-installed threading programs which include FUGUE, PROSPECT2, PAINT, SPARKS, PPA-I, SP3, PPA-II, HHsearch, SAM-T02 [22].

3. Genesilico (http://genesilico.pl/) - It is a metaserver which comprises of many constituents including PSIPRED, PROF, HMMPFAM, TMHMM, local PDB-BLAST, several 3-D structure prediction methods (RAPTOR, 3DPSSM, FUGUE, mGENTHREADER, FFAS, SAM-T02 and BIOINBGU). Based on the target-template alignments obtained as FR results and backbone of the template, primary 3D models of query structure are constructed using SCWRL. All FR alignments received from various servers undergo united appraisal by energetic standards implemented in VERIFY3D alongwith the grading criterion proposed by the PCONS server [23].

4. Bioinfobank (http://meta.bioinfo.pl) – The structure prediction Meta Server provides access to various fold recognition, function prediction and local structure prediction method. It takes the amino acid sequence of the query protein. All results of fold recognition servers are translated into uniform formats. The information extracted from the raw output of the servers includes the PDB codes of the hits, the alignments and the similarity (reliability) scores specific for every server. Mapping of the hits to the SCOP and FSSP classifications are made either using known PDB representatives or alignment of the template sequence with the databases of proteins in both classifications. The secondary structure assignments for all hits are taken from the mapped FSSP [24].

5. Loop refiner (http://genesilico.pl/) [23].

6. HHPred (http://toolkit.tuebingen.mpg.de/hhpred) – It is a sensitive protein homology analysis and structure prediction utilizing HMM-HMM-comparison. HHpred constructs a profile HMM using a query sequence and equates it to a database of HMMs which is having footnoted protein families or domains having known structure (PDB, SCOP). Output obtained is a list of nearest homologs and their alignments [25].

7. MODELLER: Selected templates were used in modeling the protein to construct the theoretical three dimensional protein structure using the Modeller software version 9v7 [26].

8. SwissPdbViewer: The theoretical model was subjected to SwissPdbViewer for energy minimization and for correcting the stereochemistry of the model [27].

2.5. Evaluation of protein model

A) UCLA server (http://www.doe-mbi.ucla.edu/) - It consists of Procheck [28], Errat [29], Verify-3D [30]. Procheck assures the stereochemical character of a protein structure by examining its residue-by-residue as well as overall structure geometry, Errat examines the statistics of non-bonded interactions within different atom types. It also maps the location of a 9-residue sliding window value versus error function. The location of the 9-residue sliding window is computed by a search for similarity and differences with statistics of highly refined structures, Verify-3D assigns a structural class to the atomic model (3D) based on its environment and location (polar, nonpolar, loop, alpha, beta etc.) and compares the results to good structures. In this way it decides the harmony of the atomic model to its own amino acid sequence (1D). The results of different models obtained were compared using this server.

B) PROQ: it is a neural network-based method to predict the quality of a protein model that extracts structural features, such as frequency of atom–atom contacts, and predicts the quality of a model [31].

2.6. Docking of Amelogenin monomers

ClusPro docking server (http://nrc.bu.edu/cluster/) – It presents as the first fully automated, web-based program for computational docking of protein structures. Coordinate files of the proteins are uploaded. It assesses many acknowledged complexes, holding back a predetermined number of complexes with favorable surface complementarities. A filter is then employed for this set of structures. Structures having correct electrostatic and desolvation free energies are selected for further clustering. The output of the program is a short list of acknowledged complexes, i.e., complexes which have been produced by server after clustering and checking for feasible values of electrostatic and desolvation free energies, which are ranked in accordance with their clustering properties [32].

[III] RESULTS

3.1. Blast output results

The PDB Blast results indicate poor homology of known PDB structures to human Amelogenin X sequence [Figure-1]. The templates were 2CSD/A (Topoisomerase V), 2CSB/A (Topoisomerase V from Methanopyrus Kandleri), 2BLE/A

THE IIOAB JOURNAL
ISSN: 0976-3104

(Human GMP Reductase in complex with GMP), 2ESH/A (Consereved potential transcription factor) having identity of 17%, 17%, 29%, 26%, respectively and similarity of 39%, 36%, 43%, 37%, respectively. From these results it is clear that homology modeling will not yield good results. So, alternative methods were undertaken.

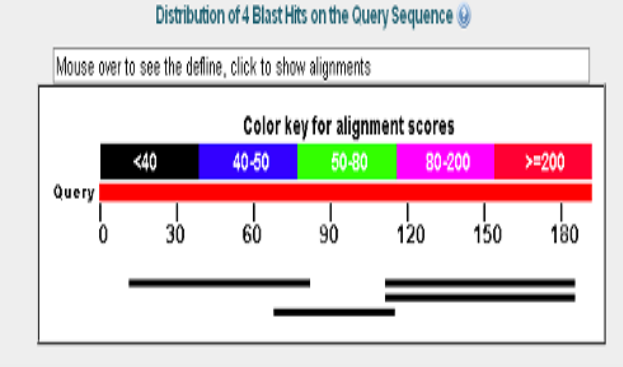

Fig: 1. Blast output result of Amelogenin X showing poor homology between the templates and Amelogenin amino acid sequence.

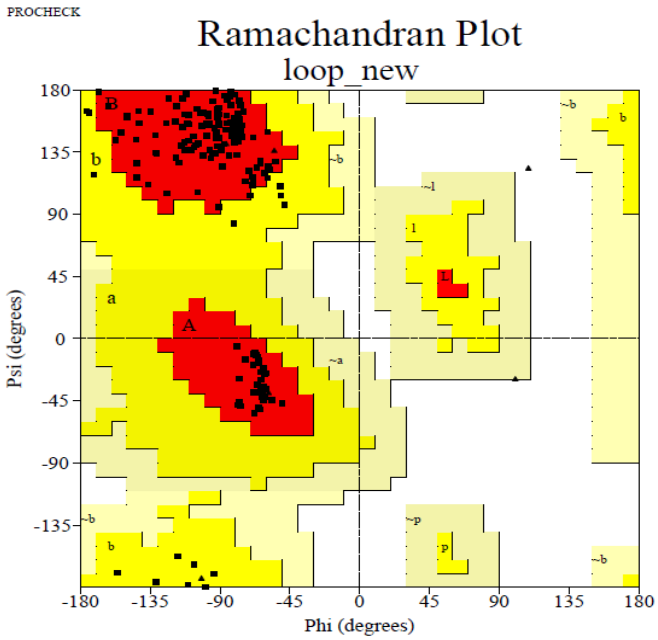

Fig: 2. Ramachandran plot for final model of amelogenin X showing most of the residues falling in the core and allowed regions of the plot.

3.2. Secondary structure prediction of Amelogenin X

Using the above mentioned tools (2.3) for secondary structure prediction, the predicted results were evaluated for the helix, strand and coil content of the protein. These results showed that most of the protein structure consists of coil and from the consensus it is evident that a very small part is constituted by helical elements. A still smaller or negligible part is constituted by strand **[Table-1]**.

3.3. *3D structure prediction of amelogenin X*

Using the various threading and ab-initio modeling servers (2.4) a number of models were obtained and analysed using various in-silico parameters like the Ramachandran scores, Verify 3D, Errat, ProQ scores and the energy values. The I-TASSER and Genesilico prediction servers gave low Ramachandran scores and high energy values. LOMETS though gave very good Ramachandran scores, the Verify-3D and Pro-Q scores were not in an acceptable range. The model predicted by Bioinfobank also gave unacceptable Verify-3D and LG scores. From this analysis it is concluded that HHPred gives the only reliable 3-D protein model on the basis of the Ramachandran scores (Core – 83.6%, Allowed – 12.5%, Generously allowed – 2.3% and Disallowed – 1.6%), energy (-1167.12) and other parameters in an acceptable range. This HHPred model was taken for further loop refinement, energy minimization and analysis **[Table-2]**.

THE IIOAB JOURNAL
ISSN: 0976-3104

3.4. Final model analysis

The model predicted by HHPred server was subjected to further refinement and energy minimization. It shows an improvement in Ramachandran scores and decrease in energy at each step, values of other scores also being acceptable, resulting in the 5th model as best approximation **[Table-3]**. Aiming for further refinement rendered the parameters become unfavorable, i.e., the Ramachandran score decreased, energy values increased, verify-3D, Errat and ProQ scores also entered unfavorable ranges. From this the 5th model was taken for further study as the values of all parameters analyzed were acceptable and the Ramachandran plot showed the phi-psi torsion angles for all residues in the structure (except those at the chain termini) in an acceptable area **[Figure-2]**. The secondary and tertiary structure of the predicted model obtained from HHPred **[Figure-3(a)]**, and the final predicted tertiary structure of Amelogenin X after refinement of model obtained from HHPred are shown in **[Figure-3(b)]**.

3.5. Self-aggregation of amelogenins

In order to further check the validity of the model predicted above, ClusPro was used. Its output gave ten configurations of the aggregation of amelogenin X to form clusters, and based on the energy values, the best configuration was selected **[Figure 4]**. Being evident from the results, the configurations match the literature findings that the hydrophobic central region of Amelogenin is forming the core and N- and C-termini are facing the outside in the formation of nanospheres upon interaction of the six monomers [7].

Table: 1. Secondary structure prediction of human Amelogenin X sequence using different servers and analyzing the helix, strand and coil content (in %) as predicted by these tools for reaching a consensus

Sl. NO.	Programs/Tools	Helix	Strand	Coil
1.	PSIPRED	3.60%	-	96.40%
2.	JPRED	10.99%	-	89%
3.	GOR4	-	17.80%	82.20%
4	HNN	6.28%	-	93.72%
5	SOPMA	10.47%	10.99%	71.73%
6	PROFPRED	5.76%	3.14%	91.10%
7	PORTER	6.80%	-	93.20%
8	PORTER+	9.95%	6.80%	83.25%
9	LOOPP	4.71%	1.05%	94.24%
10	I-TASSER	4.19%	-	95.81%
11	PHYRE	4.19%	-	95.81%
12	3D-PSSM	8.90%	-	91.10%
13	3D-Jigsaw	6.80%	-	93.20%
14	Prof prediction	8.38%	6.28%	85.34%
	Genamics Expression			
15	DPM	-	-	100%
16	DSC	8.90%	3.14%	87.96%
17	GOR4	-	17.80%	82.20%
18	HNN	5.76%	8.90%	85.34%
19	MLRC	14.66%	4.71%	80.63%
20	PHD	6.28%	9.95%	83.77%
21	Predator	6.28%	3.66%	90.06%
22	SIPMA96	15.70%	2.60%	81.70%
23	SOPMA	13.09%	10.99%	75.92%

[IV] DISCUSSION

The process of biomineralization of enamel is of great biological interest due to its uniqueness. This process is different from e.g. bone mineralization of bone, as in the latter process the protein collagen is an integral part of the mineralized tissue, whereas in mineralization of enamel the protein amelogenin is itself degraded in the process, ultimately leaving only the mineral content, thus making enamel the hardest tissue in humans. Therefore, for simulating this natural process of amelogenesis, the structure of amelogenin X protein was predicted, and by using a protein-protein docking approach we identified self aggregated amelogenins in form of cluster **[Figure-4]**. Next, binding interaction of amelogenin X to MMP20 and to the mineral Calcium hydroxyapatite needs to be done. These interactions provide an idea about the natural process of enamel formation, which can then be used to propose a material by substituting some components of this system by other components which may lead to a more biocompatible material that can be used for restoration.

THE IIOAB JOURNAL
ISSN: 0976-3104

Table: 2. Comparative result of Ramachandran Score (in %), Verify 3-D, Errat, ProQ and Energy of different models predicted by I-Tasser, LOMETS, Genesilico Model server, Hhpred and Bioinfobank server.

MODEL No.	Ramachandran Score (in %)				Verify 3-D	Errat	ProQ		Energy
	Core	Allowed	Gener	Disall		Overall quality factor	LG Score	SF Score	
I-Tasser Prediction									
1	51.50	30.30	14.40	3.80	10.42	0	51.50	30.30	14.40
2	77.30	11.40	9.10	2.30	0.52	0	77.30	11.40	9.10
3	75.80	12.90	0.80	4.50	0.00	12.766	75.80	12.90	0.80
4	44.70	28.00	18.20	9.10	48.96	0	44.70	28.00	18.20
5	54.50	28.30	15.20	1.50	24.48	0.637	54.50	28.30	15.20
LOMETS Prediction									
1	85.60	10.60	2.30	1.50	4.17	9.259	1.191	0.121	-2079.12
2	89.40	7.60	0.80	2.30	11.46	11.765	1.176	0.107	-1740.55
3	92.90	7.10	0.00	0.00	25.52	0	0.603	0.095	41.95
4	77.30	18.90	2.30	1.50	36.46	34.973	1.855	0.18	-1158.61
5	81.80	12.10	0.03	3.00	11.98	8.523	1.047	0.074	-534.403
6	92.90	7.10	0.00	0.00	29.69	0	0.286	0.048	192.27
7	100.00	0.00	0.00	0.00	0.00	0	0.91	0.081	-1745.75
8	85.00	15.00	0.00	0.00	0.00	0	1.073	0.094	-1531.43
9	75.00	20.50	3.00	1.50	18.23	12.429	0.937	0.089	-479.24
10	83.30	14.40	2.30	0.00	0.00	5.369	1.354	0.134	-1490.82
Gensilico Prediction									
Best model	59.80	28.80	6.80	4.50	33.33	2.21	1.714	0.105	16336.2
HHPred prediction									
	83.60%	12.50%	2.30%	1.60%	3.13%	13.376	0.479	0.057	-1167.12
Bioinfobank prediction									
Best model	73.90%	21.70%	0.00%	4.30%	29.17%	1.109	2.16	0.186	-2110.81

Table: 3. Showing comparative results of Ramachandran Score (in %), Verify 3-D, Errat, ProQ and Energy of different models obtained after loop refinement and energy minimization of the optimal protein model at various stages during refinement and energy minimization

Model No.	Ramachandran Score (in %)				Verify	Errat	ProQ		Energy
	Core	Allowed	Gener	Disall		Overall quality factor	LG Score	SF Score	
1	84.40%	12.50%	2.30	0.80%	14.06%	17.033	0.81	0.098	-1651.03
2	85.20	13.30	1.60	0.00	8.85	27.011	1.50	0.17	-3125.13
3	88.30	10.90	0.80	0.00	23.44	39.412	2.44	0.307	-3910.89
4	88.30	11.70%	0.00	0.00	15.63%	49.123	2.62	0.341	-3832.98
5	**88.30**	**11.70**	**0.00**	**0.00%**	**12.50**	**49.123**	**2.42**	**0.336**	**-3929.03**

[IV] CONCLUSION

The From the results obtained for secondary and tertiary structure prediction of human amelogenin X, a model could be constructed. Since this model is giving reliable results on validation and also the docking results obtained to check the proteins self-aggregation show similar results as given in literature [7], this model can be considered as the desired protein model for predicting the exact mechanism of enamel matrix formation, leading to more cost effective and biocompatible synthetic materials.

ACKNOWLEDGEMENT

We are thankful to the Director of Indian Institute of Information Technology, Allahabad for providing a supportive environment to carry out research activities.

THE IIOAB JOURNAL
ISSN: 0976-3104

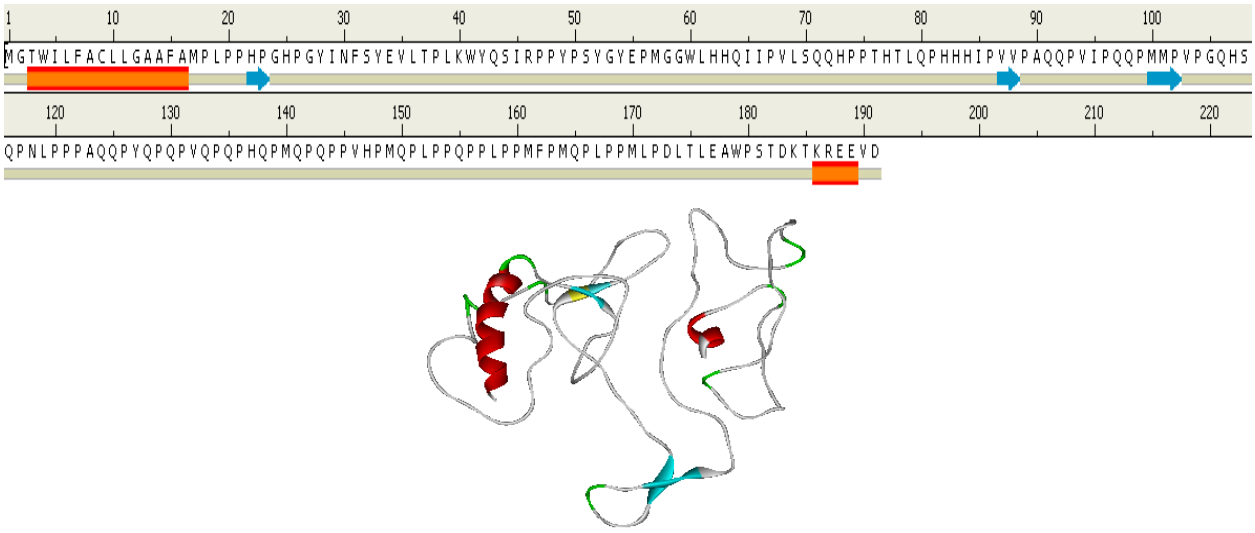

Fig: 3(a) Secondary and tertiary structure of the predicted model obtained from HHPred

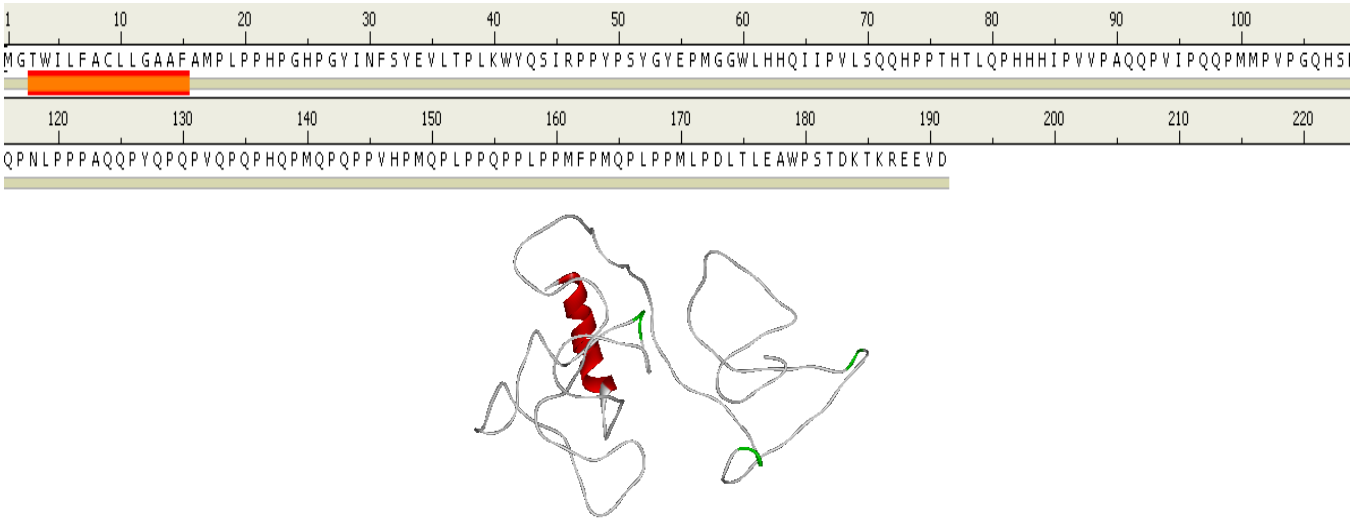

Fig: 3(b) Predicted final secondary and tertiary structure of amelogenin X after model refinement obtained from HHPred.

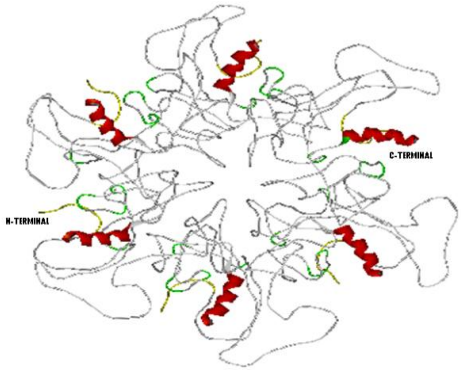

Fig: 4. Docking result of amelogenin with amelogenin to form aggregates; six amelogenin monomers interact with each other with the hydrophobic central region of Amelogenin forming the core and N- and C-termini facing outside**.**

THE IIOAB JOURNAL
ISSN: 0976-3104

REFERENCES

[1] Margolis HC, Beniash E, Fowle CE. [2006] Role of Macromolecular Assembly of Enamel Matrix Proteins in Enamel Formation. *Journal of Dental Research* 85(9):775–793.

[2] Salido EC, Yen PH, Koprivnikar K, Yu LC, Shapiro LJ. [1992] The human enamel protein gene amelogenin is expressed from both the X and the Y chromosomes. *American Journal of Human Genetic* 50:303–316.

[3] Crawford Peter JM, Aldred M, Agnes Bloch-Zupan A. [2007] Amelogenesis imperfecta. *Orphanet Journal of Rare Diseases* 17:62–65.

[4] J Moradian-Oldak J, Jimenez I, Maltby D, Fincham AG. [2001] Controlled Proteolysis of Amelogenins Reveals Exposure of Both Carboxy- and Amino-Terminal Regions. *Biopolymers* 58(7):606–616.

[5] Krishnaraju RK, Hart TC, Schleyer TK. [2003] Comparative Genomics and Structure Prediction of Dental Matrix Proteins. *Advances in Dental Research* 17:100–103.

[6] Tarasevich BJ, Lea S, Bernt W, Engelhard M, Shaw WJ. [2009] Adsorption of Amelogenin onto Self-Assembled and Fluoroapatite Surfaces. *Journal of Physical Chemistry B* 113(7):1833–1842.

[7] Fincham AG, Leung W, Tan J, Moradian-Oldak J. [1998] Does amelogenin nanosphere assembly proceed through intermediary-sized structures? *Connective Tissue Research* 38(1-4):237–240.

[8] Mahajan V, Kesharwani RK, Misra K. [2010] A structural study of matrix metalloproteinase 20 (MMP20): An in silico approach. *Advances in Applied Science Research* 1 (1): 158–165.

[9] Sun Z, Fan D, Fan Y, Du C, Moradian-Oldak J. [2008] Enamel Proteases Reduce Amelogenin-Apatite Binding. *Journal of Dental Research* 87(12):1133–1137.

[10] Nagano T, Kakegawa A, Yamakoshi Y, Tsuchiya S, Hu JC, Gomi K, et al. [2009] Mmp-20 and Klk4 Cleavage Site Preferences for Amelogenin Sequences. *Journal of Dental Research* 88(9):0823–0828.

[11] Pruitt KD, Tatusova T, Maglott DR. [2007] NCBI Reference Sequence (RefSeq): a curated non-redundant sequence database of genomes, transcripts and proteins. *Nucleic Acids Research* 35:(Database issue) D 61–65.

[12] Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, et al. [1997] Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Research* 25(17):3389–3402.

[13] King RD, Sternberg MJ. [1996] Identification and application of the concepts important for accurate and reliable protein secondary structure prediction. *Protein Science* 5(11): 2298–2310.

[14] McGuffin LJ, Bryson K, Jones DT. [2000] The PSIPRED protein structure prediction server. *Bioinformatics* 16(4):404–405.

[15] Cole C, Barber JD, Barton GJ. [2008] The Jpred 3 secondary structure prediction server. *Nucleic Acids Research* 36:W197–W201.

[16] Garnier J, Gibrat JF, Robson B. [1996] GOR secondary structure prediction method version IV. *Methods in Enzymology* 266:540–553.

[17] Asai K, Hayamizu S, Handa K. [1993] Prediction of protein secondary structure by the hidden Markov model. *Bioinformatics* 9:(2) pp.141–146.

[18] Quali M, King RD. [2000] Cascaded multiple classifiers for secondary structure prediction. *Protein Science* 9(6):1162–1176.

[19] Contreras-Moreira B, Bates PA. [2002] Domain Fishing: a first step in protein comparative modeling. *Bioinformatics* 18(8):1141–1142.

[20] Pollastri G, McLysaght A. [2008] Porter: a new, accurate server for protein secondary structure prediction. *Bioinformatics* 21(8):1719–1720.

[21] Zhang Y. [2008] I-TASSER server for protein 3D structure prediction. *BioMed Central Bioinformatics* 9:40.

[22] Wu S, Zhang Y. [2007] LOMETS: A local meta-threading-server for protein structure prediction. *Nucleic Acids Research* 35(10):3375–3382.

[23] Kurowski MA, Bujnicki JM. [2003] GeneSilico protein structure prediction meta-server. *Nucleic Acids Research* 31(13):3305–3307.

[24] Kaján L, Rychlewski L. [2007] Evaluation of 3D-Jury on CASP7 models. *BioMed Central Bioinformatics* 8:304.

[25] Söding J, Biegert A, Lupas AN. [2005] The HHpred interactive server for protein homology detection and structure prediction. Nucleic *Acids Research* 33:W244–W248.

[26] Martí-Renom MA, Stuart AC, Fiser A, Sánchez R, Melo F, Sali A. [2009] Comparative protein structure modeling of genes and genomes. *Annual Review of Biophysics and Biomolecular Structure* 29:291–325.

[27] Guex, N, Peitsch, MC. [1996] Swiss-PdbViewer: A Fast and Easy-to-use PDB Viewer for Macintosh and PC. Protein Data Bank

[28] Laskowski RA, MacArthur MW, Moss DS, Thornton JM. [1993] PROCHECK: a program to check the stereochemical quality of protein. *Journal of Applied Crystallography* 26(2): 283–291.

[29] Colovos C, Yeates TO. [1993] Verification of protein structures: patterns of nonbonded atomic interactions. *Protein Science* 2(9):1511–1509.

[30] Lüthy R, Bowie JU, Eisenberg D. [1992] Assessment of protein models with three-dimensional profiles. *Nature* 356:83–85.

[31] Wallner B, Elofsson A. [2003] Can correct protein models be identified? *Protein Science* 12(5):1073–86.

[32] Comeau SR, Gatchell DW, Vajda S, Camacho CJ. [2004] ClusPro: an automated docking and discrimination method for the prediction of protein complexes. *Bioinformatics* 20(1):45–50.

THE IIOAB JOURNAL
ISSN: 0976-3104

ABOUT AUTHORS

Mr. Rajesh Kumar Kesharwani *received his M.Tech Degree from Indian Institute of Information Technology, Allahabad, India. He is currently persuing Ph.D. in Bioinformatics from Indian Institute of Information Technology, Allahabad. He has qualified GATE Examination and got MHRD fellowship during M.Tech and also got research fellowship from government of India for persuing Ph.D. His research fields of interest are Medical Informatics, protein structure and function prediction, Computer aided drug designing and structural biology.*

Ms. Vandhana Mahajan *did her Bachelor's in Dental Surgery from Baba Farid University of Health Sciences, Punjab, India, affiliated by the Dental Council of India after which she did M.Tech in Bioinformatics from Indian Institute of Information Technology, Allahabad, India, due to bent towards research. Major fields of interest are Computer aided drug design, Chemoinformatics, Proteomics and Genomics, Dentistry, Data mining.*

Professor Krishna Misra*, Ph.D., FNASc, currently working as Honorary Professor, Bioinformatics division, Coordinator, Indo-Russian Centre for Biotechnology, Indian Institute of Information Technology, Allahabad-211012 and also Honorary professor of Center of Biomedical Magnetic Resonance Lucknow, India. Her Major fields of research are Chemistry of Natural products, Nucleic Acids, Chemoinformatics and drug designing. She has so far supervised more than 40 Ph.D. students and Published 170 papers in refereed journals, written about one dozen reviews and three books and five book chapters. She is life member of a number of societies.*