Indian Journal of Pharmaceutical Education and Research | Association of Pharmaceutical Teachers of India

In Silico Antigenicity Screening of Moringa Coagulant

Namdeo Jadhav*[1], Badrinath Bhawale[1], Santosh Bhende[1], Sushilkumar Patil[2] and Kailas Sonawane[3]

[1]Department of Pharmaceutics, Bharati Vidyapeeth College of Pharmacy, Kolhapur-416013, Maharashtra, India.

[2]Pharmacy Group, Birla Institute of Technology and Science, Pilani, Rajasthan.

[3]Department of Bioinformatics and Microbiology, Shivaji University, Kolhapur- 416013, Maharashtra, India.

ABSTRACT

Submitted: 03/05/2012 Revised: 05/09/2012 Accepted: 29/11/2012

Moringa coagulant (MC) obtained from the seeds of plant moringa oleifera (family- Moringaceae) has been substantially investigated for water purification, but, scarcely pharmaceutical applications. Recent investigations state its utility in dissolution enhancement of poorly soluble drugs, amorphous state stabilization and extended drug release. To broaden the applications of MC as a pharmaceutical excipient, need has been felt to test its antigenicity owing to its proteinic nature. In silico antigenicity screening has been carried out to enable its use in formulation of both parenteral and oral dosage form. Virtual screening carried out using online servers (ProtScale Analysis, BcePred, ABCPred, HLA-DR4Pred) predicted the probable antigenic epitopes of MC using hydrophobicity, hydrophilicity, exposed surface, secondary structure, antigenic propensity scale etc. Using these servers, the most probable antigenic regions in MC were observed to be 2-10, 29-34 and 47-53. The studies demonstrate that, MC may not pose difficulty in formulation of nonparenteral dosage forms, but, if intended for parenteral delivery, its antigenic domain needs to be exhaustively investigated.

Keywords: moringa coagulant, antigenicity and pharmaceutical excipient

INTRODUCTION

Now a days, use of protein/s as a polymer in design of drug delivery systems is widely attempted due to their biocompatibility and biodegradability in body. Mainly gelatin, arginine, glycine, lysine, human serum albumin, gliadin and legumin have been used.[1-2] Recently, attempt has been made to separate protein from seeds of plant Moringa oleifera, belonging to family Moringaceae, which has been used as a moringa coagulant for water purification.[3-5] It has been investigated that moringa coagulant improves dissolution rate of poorly soluble drugs like ibuprofen, diltiazem and hydrochlorthiazide [6] and also it has been investigated that moringa coagulant stabilizes poorly soluble drugs in their amorphous state.[7] Furthermore it has been observed that, in tabletted form, it extends the release of diclofenac, ibuprofen and diltiazem for about 18 hours.[8] But, literature has revealed its antigenicity potential.[9] After knowing the rich availability and safety of this protein (LD_{50} 550mg/kg/body weight), it has been thought to explore its applications in design of parenteral and nonparenteral dosage forms.[10] Especially, its use in the targetted drug delivery, development of synthetic vaccines, immunodiagnostic tests and antibody production can also be enabled. Hence, to test and predict antigenicity of moringa protein, it has become essential to virtually screen protein and carry out in vivo studies.

***Address for Correspondence:**

Namdeo Jadhav, Department of Pharmaceutics, Bharati Vidyapeeth College of Pharmacy, Kolhapur-416013, Maharashtra state, India

E-mail: nrjadhav18@rediffmail.com

Earlier, precise delineation of antigenic determinants i. e. epitopes was carried out by preparing large number of derived proteins and well characterized peptidic fragments, from the original protein antigen and then these derivatives were tested for probable immunological activity.[11] But, such studies are extremely laborious, and hence, researchers prefer to develop more rapid prediction methods, based only on the amino acids sequence of the protein in question.

An antigenic site has specific characteristics, like hydrophilicity, typical secondary structure characteristics (alpha helix and beta sheet). It is a site in protein structure comprising a fixed number and sequence of amino acids.[12] There are several methods that predict the epitopes position in whole amino acid sequence, based on some physicochemical properties of proteins such as hydrophilicity, mobility / flexibility, surface accessibility, helicity, structure etc.[13-15] Alternatively, a homologous series of proteins may be used to assess the influence of particular amino acid substitutions, thereby implicating certain regions as antigenic determinants.[16-17] But, many surface oriented regions are non-antigenic. Hopp and Woods adopted an empirical approach to analyze certain amino acids for hydrophilicity in order to find a particular kind of sequence that is favored for antibody binding. This report describes a system that uses a simplified method to successfully predict antigenic determinants, if amino acid sequence of a protein is given.[18]

The hydrophobicity is another main force contributing for antigen – antibody interactions. The hydrophobic interactions generally increase the enthalpy, leading to a poor stability of the formed complexes. Thus, the recognition sites i.e. epitopic

site of proteins and ligands are often more hydrophobic than the rest of protein molecule.[19]

The secondary structure of proteins contains two types of arrangement, α-*helix* and β-*sheet*. In addition, there are other intermediate structures derived from these two, such as β-*turn* or *coil*. Tertiary structure of protein is derived from folding of polypeptidic chains, when a hydrophobic area comes on inner side, to form core. The β-sheet and β-turn structure plays an important role for the formation of tertiary structure.[20-21] Generally the β-type structures appear on the surface molecule and so involved in the intermolecular interactions. This increases the probability of antigen – antibody interaction with greater probability at β-turn level than in the compact, rigid structure (α-helix).

In the present work, attempt has been made to test in vivo antigenicity of moringa protein in mice and virtually it has been screen for antigenicity using Protoscale, BcePred, ABCPred and HLA-DR4 antigenicity prediction online servers.

MATERIALS AND METHODS

Moringa seed powder was supplied by Veg India Export (Tamilnadu, India). All other chemicals were purchased from Sigma Aldrich and they were of analytical grade.

Virtual Screening of Moringa Protein for Antigenicity Prediction

1. Homology Modeling of Moringa Protein Sequence

The target sequence of moringa protein was extracted from NCBI's protein sequence database **(gi|998697|gb|AAB34890.1)**

[1]QGPGRQPDFQRCGQQLRNISPPQRCPSLRQAVQLTH QQQGQVGPQQVRQMYRVASNIPST[60]

Whereas template structure (2DS2_B.pdb) was taken from protein data bank. A pair-wise sequence alignment between target and template sequence was made and three-dimensional model of moringa protein was developed by using molecular operating environment (MOE) builder software 2009.10. [22-23]

2. In Silico Studies Using Online Servers

2.1. ProtScale Analysis

a) Hydropathic scale: Hydropathic scale combines the hydrophilic and hydrophobic properties of the 20 essential amino acids. In this method, for each amino acid, hydrophilic or hydrophobic index has been assigned. The hydrophobic index is a measure of the free energy of an amino acid, consequently deciding its transition from aqueous to organic solvents. Amongst numerous methods adopted for hydropathic scale preparation, Kyte & Doolittle is the one, widely used due to its highest accuracy, and hence used for our study. Table 1, showed the basic hydrophilic and hydropathic values for each amino acid .[24]

The hydropathic profile of a moringa protein was graphically expressed in a coordinate system in which averaged hydrophobic indices of amino acids are plotted versus each amino acid position in the protein sequence. The entire sequence was scanned and the average indices for each encountered amino acid were computed.

b) Hydrophilic scale: Hydrophilicity is another parameter, used to study antigenicity prediction. Table 1 has listed the hydrophilicity values assigned to the 20 amino acids commonly found in proteins. Presently, Hopp-Woods scale was used for predicting potential antigenic sites of moringa proteins, likely to be rich in charged and polar residues. Each amino acid in the sequence of the protein, with assigned hydrophilicity values were repetitively averaged down the length of the polypeptide chain, to generate a series of local hydrophilicity values. The number of hydrophilicity values that are averaged at each repetition is arbitrary, and we chosen groups of six, because, that is the approximate size of an antigenic determinant. Once the complete set of averaged value was obtained, the list was then scanned to locate the highest value. At each position, the mean hydrophobic index of the amino acids within the window was calculated and that value was plotted as the midpoint of the window. According to present scale studies, high point will invariably lie within or be immediately adjacent to one of that protein's antigenic determinants.[25]

c) Protein secondary structure: The probable profile of alpha-helix and beta-sheet frequencies was achieved using the Chou-Fasman algorithm by means of ProtScale software.[27] It takes into account the probability that each of the 20 amino acids, with basic individual values given in table 1, may be involved into an alpha or beta-type of structure, and computes the likelihood of amino acids to appear in a particular structure (propensity for secondary structure).

2.2. BcePred Prediction Server

BcePred evaluates the performance of existing linear B-cell epitope prediction methods based on physico-chemical properties on a non-redundant dataset. Bcipep database consists of 1029 B-cell epitope and equal number of non-epitopes obtained randomly from Swiss-Prot database. Using this server we were able to predict B-cell epitopes of moringa protein sequence using various physico-chemical properties like hydrophilicity, flexibility, accessibility, exposed surface and turns. The peak of the average amino acid residue segment having value higher than the threshold value i. e. more than default value of 2.38 is considered as probable B-cell epitope .[26]

Table 1 List of amino acids with their basic physicochemical values

Amino acids	Hydrophobicity (Kyte-Doolittle)	Hydrophilicity value	Alpha helix	Beta sheet	Beta turn
Alanine	1.8	-0.5	1.42	0.83	0.66
Arginine	-4.5	3.0	0.98	0.93	0.95
Asparagine	-3.5	0.2	0.67	0.89	1.56
Aspartic acid	-3.5	3.0	1.01	0.54	1.46
Cysteine	2.5	-1.0	0.70	1.19	1.19
Glutamine	-3.2	0.2	1.11	1.10	0.98
Glutamic acid	-3.2	3.0	1.51	0.37	0.74
Glycine	-0.4	0.0	0.57	0.75	1.56
Histidine	-3.2	-0.5	1.00	0.87	0.95
Isoleucine	4.5	-1.8	1.08	1.60	0.47
Leucine	3.8	-1.8	1.21	1.30	0.59
Lysine	-3.9	3.0	1.16	0.74	1.01
Methionine	1.9	-1.3	1.45	1.05	0.60
Phenylalanine	2.8	-2.5	1.13	1.38	0.60
Proline	-1.6	0.0	0.57	0.55	1.52
Serine	-0.8	0.3	0.77	0.75	1.43
Threonine	-0.7	-0.4	0.83	1.19	0.96
Tryptophan	-0.9	-3.4	1.08	1.37	0.96
Tyrosine	-1.3	-2.3	0.69	1.47	1.14
Valine	4.2	-1.5	1.06	1.70	0.50

General Basis of Predictions

In this server, hydrophilicity, flexibility, accessibility, turns, exposed surface, polarity and antigenic propensity of polypeptides chains have been correlated with the location of continuous epitopes in a few well-characterized proteins.[27-28] The calculations of probable antigenic sites are based on the different scales obtained by various methods. Each scale consists of 20 values assigned to each of the amino acid residues on the basis of their relative propensity to possess the property described by the scales obtained by various methods like, Parker Method, Karplus Method, Emini Method, Pellequer Method, Kolaskar Method, Exposed surface scale, and Polarity scale.[13-15,29]

2.3. ABCPred server

The ABCpred server predicts B cell epitope(s) in an antigen sequence, using artificial neural network. B-cell epitopes were obtained from B cell epitope database which contains 2479 continuous epitopes, including 654 immunodominant, 1617 immunogenic epitopes. The dataset prepared from swiss-prot consists of 700 B-cell epitopes and equal numbers of non-epitopes. The server is thus able to predict the epitopes of given sequence using various physicochemical properties like *Flexibility, Hydrophilicity, Polarity,* and *Surface* properties combined at a threshold of 2.38.[30-31]

2.4. HLA-DR4Pred

The HLA-DR4Pred is an SVM and ANN based HLA-DRB1*0401(MHC class II alleles) binding peptides prediction method. The MHCBN database was used as the source of data. The binding core of 9 amino acids was obtained from the binders without considering MHC binding motifs by using Matrix Optimization Techniques (MOT) package. All the duplicate entries were cropped from the dataset. The final dataset is consisted of 567 unique MHC binders of varying binding affinity.[32]

RESULTS AND DISCUSSION

Virtual Screening of Moringa Protein for Antigenicity Prediction

1. Homology Modeling of Moringa Protein Sequence

The 3-D structure of the moringa protein (ribbon model) has been given in Fig. 1, which shows folding of 60 amino acids with three alpha helices in its secondary structure.

2. In Silico Studies Using Online Servers

2.1. Protscale Analysis

a) Hydropathic scale

The output of ExPASy server as seen in Fig. 2 shows the two antigenic region in the moringa protein, from **29-34** and second from **51-56**. For searching them, hydropathic scale analysis has used Kyte and Doolitle scale as a base.

b) Hydrophilicity Scale

Hopp-Woods scale has been designed for predicting potential antigenic regions of polypeptides. Values greater than zero are considered as hydrophilic in nature and thus, likely to be

exposed on the surface of folded protein. From present scale analysis amino acids from **5-10** and **22-29** will be called as antigenic in nature (Fig. 3).

C) Protein Secondary Structure

i) Alpha helix and beta sheet

An alpha helix and beta sheet structures are characteristic for antigenic sites as per the criterion set by Chau and fasman. Each alpha-helix ascendant area corresponds to a descendent area of beta-sheet. On the basis of Chau and Fasman, and ExPASy server output, as seen in the form of Fig. 4 and Fig. 5, following locations on moringa protein sequence have been identified as antigenic determinants.

For alpha helix: **29-37** and **45-53** (Fig. 4),

For beta sheet: **12-18, 29-37** and **45-53** (2 peaks), (Fig. 5).

ii) Beta – turn

The beta-turn structures had a maximum frequency of 1.314 and the following antigenic sites have been identified, as shown in Fig. 6; 1**-6, 19-27** and **38-45**.

2.2 BcePred Prediction server

This server predicts B-cell epitopes using any of the physico-chemical properties (hydrophilicity, flexibility/mobility, accessibility, polarity, exposed surface and turns) or combination of properties. If the values of amino acid sequence are greater than the respective threshold, its considerd as an antigenic in nature. The details of probable antigenic regions with various physico-chemical properties have been given in Table 2, along with respective threshold.

The predicted B-cell epitopes are shown in blue color and underlined.

2.3 ABCPred server

The present server, predicts and ranks B cell epitopes according to their score obtained by trained recurrent neural network with respective threshold of 0.51. Higher score of the peptide means the higher probability of being an antigenic epitope. The peptides showing threshold value above chosen

Fig. 1: Ribbon model showing the residues of moringa protein

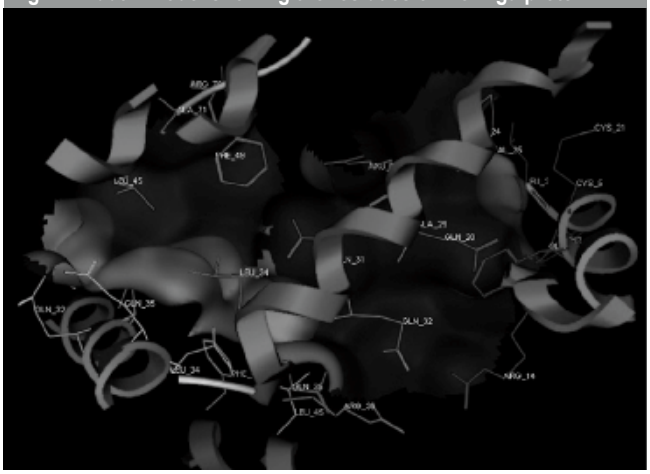

Fig. 2: Spectral presentation of the Kyte and Doolitle scale for moringa protein sequence

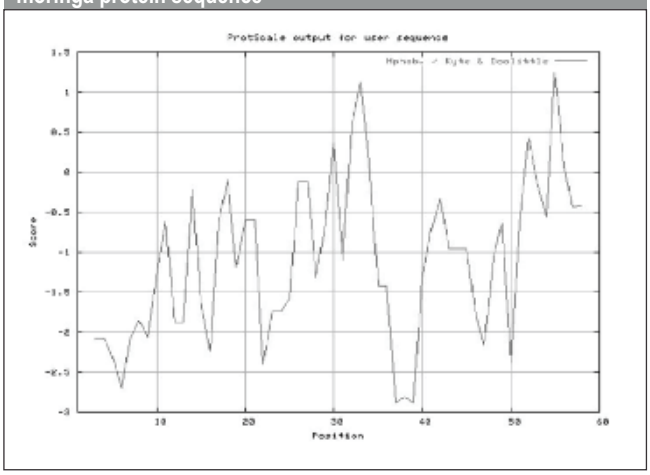

Fig. 3: Spectral presentation of the Hopp and woods scale for moringa protein sequence

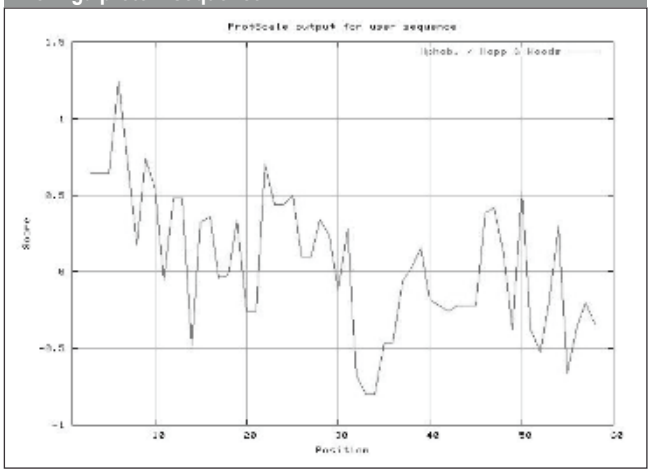

Fig. 4: Spectral presentation of the alpha helix scale for moringa protein sequence

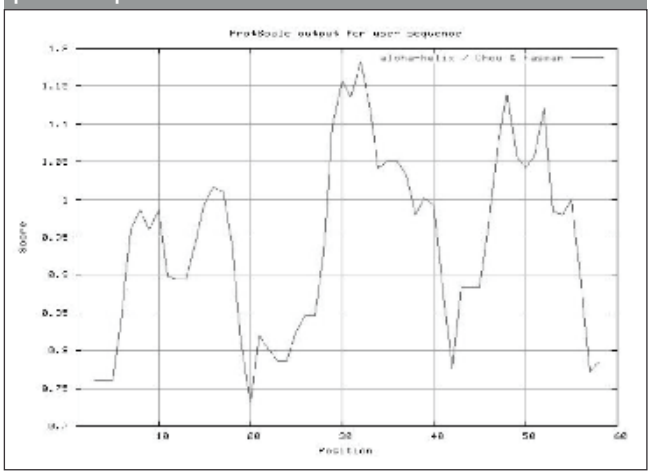

Table 2: Predicted antigenic sequence shown in blue color for moringa protein sequence

Sequence	Respective Threshold[1]	QGPGRQPDFQRCGQQLRNISPPQRCPSLRQAVQLTHQQQGQVGPQQVRQMY RVASNIPST[60]
Hydrophilicity	**1.9**[1]	QGPGRQPDFQRCGQQLRNISPPQRCPSLRQAVQL[34]THQQQGQVGPQQVRQMY RVASNIPST[60]
Exposed Surface	**2.3**[1]	QGPG[5]RQPDFQRCGQQLRNISPPQRCPSLRQAVQLTHQQQGQVGPQQVRQMY RVASNIPST[60]
Antigenic Propensity	**1.9**[1]	QGPGRQPDFQRCGQQLRNISPPQRCPSLRQAVQLTHQQQGQVGPQQVRQMY RVASNIPST[60]
Polarity	**1.8**[1]	QGPG[5]RQPDFQRCGQQLRNISPPQRCPSLRQAVQLTHQQQGQVGPQQVRQMY RVASNIPST[60]

Fig. 6: Spectral presentation of the beta turn scale for moringa protein sequence

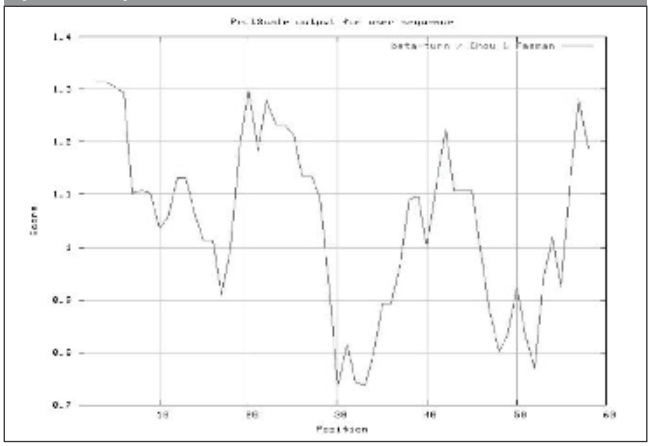

Fig. 5: Spectral presentation of the beta sheet scale for moringa protein sequence

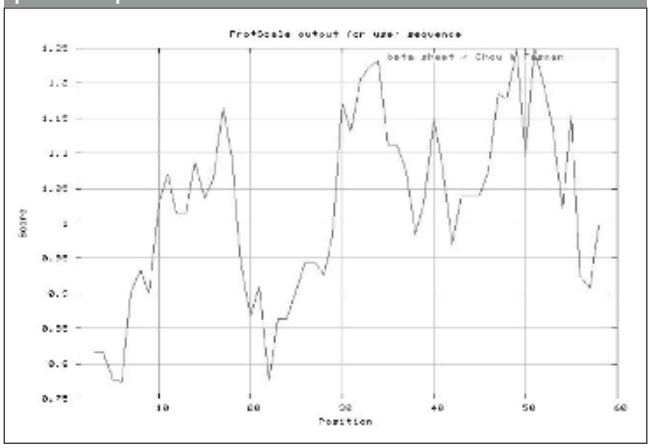

Table 3: Scoring for various amino acids in moringa protein sequence by ABCPred server

Rank	Start position	Score
1	1	0.83
2	20	0.81
3	45	0.78
4	7	0.70
5	26	0.57

Table 4: Predicted binding potential for amino acids in HLA-DR4P red server for moringa protein sequence

Peptide Rank	Start Position	Sequence	Score	Prediction
1	27	LRQAVQLTH	1.088	Binder
2	50	RVASNIPST	0.766	Binder
3	9	QRCGQQLRN	0.665	Binder
4	43	GPQQVRQMY	0.297	Non-binder
5	33	QLTHQQQGQ	0.061	Non-binder

(0.51) are considered as probable antigenic regions. Means, in moringa protein, region starting with amino acid sequence 1 having highest score (0.83) can be deemed as most probable antigenic epitope (Table 3).

2.4 HLA-DR4Pred

This server can be used to find binding and nonbinding potential of amino acid sequence in protein, depending on the score above or below threshold score (0.5). If the score exceeds threshold score, means amino acid region is probably antigenic and if the score is below 0.5, its non binding sequence and hence nonantigenic. The n=below mentioned moringa protein sequence with blue color antigenic regions is found to be binding. Details of amino acid sequence with their score have been shown in Table 4.**QGPGRQPDFQRCGQQLRNISPPQRCPSLRQAVQ LTHQQQGQVGPQQVRQMYRVASNIPST**

CONCLUSION

The antigenic nature of moringa protein was demonstrated from in silico studies. Studies have shown that, group of 6-12 amino acid decides the antigenic region, which depends on the average local value of physicochemical properties like hydrophilicity, hydrophobicity, secondary structure etc of individual amino acids. Using various online servers, the most probable antigenic region in moringa protein sequence was predicted as 2-10, 29-34 and 47-53. Since, moringa seeds are used as a food material, use of protein may not pose difficulty in formulation of nonparenteral dosage forms, but, if intended for parenteral delivery, and further exhaustive studies are warranted.

REFERENCES

1. Dong XL, Yu-Kyuong O, Soo-Jeong L, Jong OK, Ho JY, Jung HS, Chul SY, HanGon C. Novel gelatin microcapsule with bioavailability enhancement of ibuprofen using spray-drying technique. Int. J. Pharm. 2008; 355: 277-84.
2. Paola M, Francesca M. and Marzia Cirri. Ternary systems of naproxen with hydroxyl propyl-β-cyclodextrin and amino acids. Int. J. Pharm. 2003; 260, 293-302.
3. Ghebremichael KA. Moringa seed & Pumice as alternative natural materials for drinking water treatement [PhD thesis,] Uppasala Univesity stockhom. 2004.
4. Guo JH, Skinner, JW, Harcum WW. Pharmaceutical applications of naturally occurring water-soluble polymers. Pharm. Sci. & Tech. Today. 1998; 6, 254-61.
5. Ndabigengesere, A., Narasiah, K.S., Talbot, B.G., 1995. Active agents & mechanism of coagulation of turbid waters using Moringa oleifera. Wat. Res. 29, 703-10.
6. Jadhav NR, Bhawale BK, Paradkar AR. and Kadam SS. More HN. Novel pharmaceutical composition for solubility enhancement. Indian patent. 2010; 1883.
7. Jadhav NR, Bhende SA. and More HN. Novel stabilized composition of amorphous state solids. Indian patent. 2011; 2046.
8. More, H.N., Jadhav, N.R., Patil, SY. and Kadam, H.M., Extended release application of moringa coagulant. Indian patent. 2010; 1172.
9. Patil SY. Studies on extended release applications of polymer extracted from Moringa oleifera, [Thesis] Shivaji University, India, 2009.
10. Grabow WOK JL, Slabbert Morgan, WSG. Toxicity and mutagenicity evaluation of water coagulated with *Moringa oleifera* seed preparations using fish, protozoan, bacterial, coliphage, enzyme and Ames *Salmonella* assays; Water S.A.A. 1985; 11(1): 9.
11. Thomas P, Hopp Keneth R, Woods. Prediction of protein antigenic determinants from amino acid sequences. Immunology. 1981. 78, 3824-8.
12. Pellequer JL, Westhof E, Van Regenmortel MH. Correlation between the location of antigenic sites and the prediction of turns in proteins. Immunol. Lett. 1993; 36, 83-99.
13. Greenbaum JA, Andersen PH, Blythe M, Bui HH, Cachau RE, Crowe J, Davies M, Kolaskar AS, Lund O, Morrison S. Towards a consensus on datasets and evaluation metrics for developing B-cell epitope prediction tools. J. Mol. Recognit. 2007; 20: 75-82.
14. Karplus PA, Schulz GE. Prediction of chain flexibility in proteins a tool for the selection of peptide antigens. Naturwissenschaften. 1985; 72, 212-3.
15. Kolaskar AS, Tongaonkar PC. A semi-empirical method for prediction of antigenic determinants on protein antigens. FEBS. Lett. 1990; 276, 172-4.
16. Shide Liang. Prediction of antigenic epitopes on protein surfaces by consensus scoring. BMC Bioinformatics. 2009; 10, 302. doi:10.1186/1471-2105-10-302.
17. Liang S, Zhang C, Liu S, Zhou Y. Protein binding site prediction using an empirical scoring function. Nucleic Acids Res. 2006; 34, 3698-707.
18. Hopp TP, Woods KR. Prediction of protein antigenic determinants from amino acid sequences. Proc. Natl. Acad. Sci. USA. 1981; 78, 3824-8.
19. Peters B, Sidney J, Bourne P, Bui HH, Buus S, Doh G, Fleri W, Kronenberg M, Kubo R, Lund O. The immune epitope database and analysis resource: from vision to blueprint. PLOS Biol. 2005; 3, 379-81.
20. Allcorn LC, Martin AC. SACS--self-maintaining database of antibody crystal structure information. Bioinformatics. 2002; 18: 175-81.
21. Rapberger R, Lukas A, Mayer B. Identification of discontinuous antigenic determinants on proteins based on shape complementarities. J. Mol. Recognit., 2007; 20, 113-21.
22. Hideaki Y, Tatsuo A, Tao Y, Yumi K, Katsuyoshi K, Toshiro N, Hironori U. and Kazuo R. Homology modeling and structural analysis of 11b-hydroxysteroid dehydrogenase type 2. Eur. J. Med. Chem. 2011; 46, 1325.
23. Toshio I, Kazunori S, Khaled A, Daisuke K, Atsushi K, Kensaku S, Toshifumi Y. and Akira W. J. Inhibition of RuBisCO Cloned from Thermosynechococcus vulcanus and Expressed in Escherichia coli with Compounds Predicted by Molecular Operation Environment (MOE). Biosci Bioengi. 2008; 105(1): 26.
24. Welling GW, Weijer WJ, Van der zee R. and Welling-Wester S. Prediction of sequential antigenic regions in proteins. FEBS. Lett. 1985; 188, 215.
25. Chou PY, Fasman GD. Empirical predictions of protein conformation. Annu. Rev. Biochem. 1978; 47: 251-76.
26. Sweredoski MJ, Baldi P. PEPITO: improved discontinuous B-cell epitope prediction using multiple distance thresholds and half sphere exposure. Bioinformatics. 2008; 24, 1459-60.
27. Chen J, Liu H, Yang J, Chou KC. Prediction of linear B-cell epitopes using amino acid pair antigenicity scale. Amino Acids. 2007; 33: 423-8.
28. Saha S, Raghava GP. Prediction of continuous B-cell epitopes in an antigen using recurrent neural network. Proteins. 2006; 65, 40-8.
29. Kulkarni-Kale U, Bhosale S, Kolaskar AS. CEP: a conformational epitope prediction server. Nucleic Acids Res. 2005; 33, 168-71.
30. El-Manzalawy Y, Dobbs D, Honavar V. Predicting linear B-cell epitopes using string kernels. J. Mol. Recognit. 2008; 21: 243-55.
31. Rubinstein ND, Mayrose I, Pupko T. A machine-learning approach for predicting B-cell epitopes. Mol. Immunol. 2009; 46, 840-7.
32. Schonbach C, Koh JL, Sheng X, Wong L, Brusic V. FIMM, a database of functional molecular immunology. Nucleic Acids Res. 2000; 28, 222-4.

****