DOI: 10.24287/1726-1708-2020-19-1-35-46

In silico assessment of plant L-asparaginase and estimating its allergenicity in comparison to bacteria asparaginase

Milad Yazdi[1], Maryam Kolahi[2], Ali Mohamad Foroghmand[1], Mohammad Reza Tabandeh[3]

[1] *Department of Genetic, Faculty of Science, Shahid Chamran University of Ahvaz, Ahvaz, Iran.*
[2] *Department of Biology, Faculty of Science, Shahid Chamran University of Ahvaz, Ahvaz, Iran.*
[3] *Department of Biochemistry and Molecular Biology, Faculty of Veterinary Medicine, Shahid Chamran University of Ahvaz, Ahvaz, Iran.*

Received 30.10.2019
Accepted 23.12.2019

Correspondence:
Maryam Kolahi, Department of Biology, Faculty of Science, Shahid Chamran University of Ahvaz, Ahvaz, Iran.
E-mail: m.kolahi@scu.ac.ir

L-asparaginase is widely distributed among microorganisms, animals and plants. L-asparaginase has been utilized as a drug in the treatment of lymphoid malignancies and plays a crucial role in asparagine metabolism in plant stress response mechanisms. Multiple sequence alignment of Neighbor–Joining phylogenetic tree was executed utilizing Mega 4.0. Two plants asparaginase were identified whose three dimensional structures compared well with two bacterial samples of L-asparaginase used in humans as a therapeutic drug. Prediction of antigen cites, B-cell epitope identification and prediction of epitopes by use of Cytotoxic T-lymphocyte was performed using various in silico server resources. The survey showed that between the 40 plants, 2 identified items of human, 12 bacteria and 6 algae of asparaginase genes, generally two main branches created that samples of green algae is in the neighborhood of to the bacterial samples. Interestingly the data showed that the two bacterial samples of L-asparaginase used in medicine, when compared to plant asparaginase genes, have less similarity to asparaginase genes of human, while the two human asparaginase genes are located perfectly between the plant groups with their sequence revealing high similarity with plant species. Although there was some allergen epitope found in plant asparaginase, these are different from the allergen epitopes of microbial asparaginase that are used as a drug in humans with no common sequence being found between them. This manuscript provides evidence suggesting the potential utilization of *Phaseolus vulgaris* asparaginase, which has less epitopes, better predicting tool scores and high similarity, in drug design as an enzymetherapy in leukemia and other cancers.

Key words: *Allergenicity, L-asparaginase, Lupinus luteus, Phaseolus vulgaris*

Yazdi Milad, et al. Pediatric Hematology/Oncology and Immunopathology, 2020; 19 (1): 35–46.
DOI: 10.24287/1726-1708-2020-19-1-35-46

L-asparaginase (EC 3.5.1.1) catalyzes the hydrolysis of L-asparagine to yield L-aspartic acid and ammonia [1]. This enzyme is widely distributed among microorganisms, animal, plant, eukaryotic microorganisms such as filamentous fungi and yeasts and humans [1, 2]. L-asparaginase plays an essential role in the metabolism of asparagine in response to plant stress [3, 4]. Additionally, L-asparaginase can facilitate the translocation of nitrogen within sink organs of the plant [5]. L-asparaginase plays a vital role during development of plant nodule, seed development and germination due to its involvement in amide and amino acid metabolism [5–7]. Two asparaginase isoforms existing as K^+-dependent and K^+-independent plays physiological roles in many plants [8]. Recent structural and enzyme functionality data have allowed for the use of enzyme therapy as a tool in the treatment of many diseases such as cardiovascular, cancer, infection, autoimmune disorder [9]. L-asparaginase is currently used in enzyme therapy, with expanded attention in recent years due to its chief applications, which includes its use in the drug industry as a substitute for the treatment of different types of cancers such as critical lymphoblastic leukemia, malignant diseases of the lymphoid system and Hodgkin's lymphomas [1].

The use of asparaginases can result in a number of side effects. Hypersensitivity due to anti-asparaginase antibody assembly, have been detected in more than 60% of patients that use *Escherichia coli* asparaginase therapy [10]. The progress of these antibodies appears to be more commonly observed in patient that received native *Escherichia coli* asparaginase [11–13]. This enzyme is also utilized in the food industry to inhibit the formation of acrylamide when foods are processed at high temperatures. Due to acrylamide being a neurotoxin and carcinogenic, the utilization of this enzyme is imperative [14, 15].

Information about the structure and function of L-asparaginase from *Escherichia coli* and *Erwinia sp*. has led to the use of special resources for the preparation of the enzyme as a drug [14]. The use of *E. coli* L-asparaginase may be restricted due to serious adverse effects of which allergic reaction is the most common [16]. Nevertheless, the use of Erwinia asparaginase was extended by the Food and Drug Administration (FDA) agreement in 2011, with the aim of finding an alternative for patients with allergic reactions to *Escherichia coli*-derived asparaginase [17]. Some allergic responses involve complex interactions between the protein and the immune system, and are therefore especially difficult to predict. However, data has shown that various proteins are fundamentally more allergenic than others. Toxicologist scientists now have the challenge to overcome the issue of detecting those characteristics of proteins with the potential of encouraging allergic sensitization and allergic disease [18, 19]. In an attempt to determine the allergenic potential of peptides a comparison was made with those of known allergens and sources considered not to cause allergy in humans. Regions of homology with known protein allergens spanning eight or more consecutive amino acid residues indicates the possible existence of common epitopes, from a known allergenic source. If no noteworthy homology with known allergens is detected, this suggests that the protein probably won't be allergenic.

The asparaginase enzyme is extensively present in plants with superior properties to bacteria. The study was aimed at evaluating asparaginase from plant sources and comparing the homology of their sequence with human asparaginases and some bacteria asparaginases that are commonly used as drugs. Two known plant species were compared to two bacteria species that are used as drugs with allergenic side effects, with regards to the allergen properties of L-asparaginase peptide sequences utilizing various bioinformatics techniques.

MATERIAL AND METHODS

Phylogenetic analysis of the Asp genes and Protein Sequence Collection. For immunoinformatic and phylogenetics analysis asparaginase enzymes sequences were taken from the NCBI protein database using advanced search for the keywords, “L-asparaginase” “Asparaginase” “green plant” “Algae” and “Bacteria”. Multiple sequence alignment were made using muscle algorithm of Mega 7 software and the alignment results viewed using Jalview to detect conserved residues [20]. The phylogenetic tree was constructed using the Maximum likelihood method by mega7 with default setting [21].

Antigen prediction. Two methods were used for allergen prediction; the first searched for sequence similarity with alignment like SDAP (Structural Database of Allergenic Proteins); the second approach was based on detection of motifs that proved to be allergen like VaxiJen. Both of them depended on the sequence but the second was alignment-free. The first server that was used was Vaxijen. Vaxijen uses three databases for prediction: bacteria, virus and tumor. Antigen prediction for the tumor database, the used threshold was as default of database [22]. The AlgPred server was used for prediction of allergens based on IgE epitopes [23]. The SDAP database was another database used for prediction based on alignment of AllergenFP v1.0 [24, 25]. The amino acids in the protein sequence in data sets were described by five E-descriptors and the strings were transformed into uniform vectors by auto-cross covariance (ACC) transformation. The first principal component (E1) reflects the hydrophobicity of amino acids; the second (E2) – their size; the third (E3) – their helix-forming propensity; the fourth (E4) correlates with the relative abundance of amino acids; and the fifth (E5) is dominated by the β-strand forming propensity. The last server that was utilized was AllerTOP v.2 that is another free alignment tool [26].

Prediction of Immunogenic Epitopes

B-CELL epitope(s). To predict B-cell epitope(s) three linear software were utilized.

ABCpred. ABCpred server was the first one that was able to predict epitopes with 65.93% accuracy using recurrent neural network [27]. 16 was selected as the window length for prediction and 0.51 as the threshold.

BcePred Prediction Server. BcePred was the second tool used for predicting linear B-cell epitopes in a protein sequence [23]. This server allows users to predict B-cell epitopes using any of the physico-chemical properties (hydrophilicity, flexibility/mobility, accessibility, polarity, exposed surface and turns) or a combination of properties. The server was able to predict epitopes with 58.7% accuracy.

ElliPro Predicted Server. ElliPro predicts linear and discontinuous antibody epitopes based on a protein antigen's 3D structure. ElliPro accepts as an input, protein structure in PDB format. ElliPro associates each predicted epitope with a score, defined as a PI (Protrusion Index) value averaged over epitope residues. In this method, the protein's 3D shape is approximated by a number of ellipsoids, so that the ellipsoid with PI = 0.9 would include within 90% of the protein residues with 10% of the protein residues being outside of the ellipsoid. The default minimum score is 0.5 but for stronger

result 0.7 was chosen and a maximum distance of 6 (angstrom) [28].

T CELL epitope(s). NetCTL was used to predict T cell epitopes based of protein sequence. A threshold of 0.75 was chosen that is equal to 0.8 sensitivity and 0.970 specificity. The weight on C terminal cleavage and the weight on TAP transport efficiency were 0.15 and 0.05, respectively [28].

RESULTS

Protein Sequence Collection, Phylogenetic analysis of the Asp genes, Classification for L-asparaginase Genes. Initially 60 L-asparaginase sequences from plants (40 sequences), algae (6 sequences), bacteria (12 sequences) species, and two human L-asparaginase genome were sequenced. The outcome of the extensive database search is briefed in *Table 1*. The study aimed to compare 60 L-asparaginase genes of human, bacterial, algae and plants in an attempt to identify more appropriate sources of medicinal asparaginase based on the extent of the similarities and differences of the sequences. Phylogenetic analyses permitted the recognition of evolutionary conservation and gene divergence. A maximum likelihood tree was produced from amino acid sequences of the deduced full-length peptides with the best-fit evolutionary Mega 7 (*Figure 1*). The survey showed that between the 40 plants, 2 identified items of human, 12 bacteria and 6 algae of asparaginase genes, generally two main branches created that samples of green algae is in the neighborhood of to the bacterial samples.

These homologs and a few evolved in plant and human asparaginase genes indicated a common eukaryotic ancestor. Points to note in relation to the clade of plant asparaginase genes is that in some cases the species belonging to monocotyledons and dicotyledons would have been well separated, even with those species belonging to a family of plants put together, that indicates conserved regions between these parts of the gene. For example, species belongs to the family Fabaceae (marked with the blue square), Gramineae (marked with the brown diamond) and Ephorbiaceae (marked with the green triangle) would have been isolated. On the other hand some species belonging to one plant family are quite far from this tree (marked with the yellow triangle).

Allergenic Site Prediction. In order to study and predict the sites of the antigens in the asparaginase plant, 60 samples were studied using Phylogenetic analysis; two samples of *Lupinus luteus* and *Phaseolus vulgaris* which are known 3D plant asparaginase were compared in terms of the antigen peptides with two samples of bacteria asparaginase (*Erwinia chrysanthemi* and *Escherichia coli*) which are used as drugs were selected.

Vaxijen server. Vaxijen server was the main server utilized for alignment-independent prediction of protective antigens. Outcome displays that both bacterial samples are antigen and both plants are non-antigen but are close to the threshold (0.5) (*Table 2*).

AlgPred server. The AlgPred server predicts allergens based on several algorithms that include, Mapping of IgE epitopes and PID, MEME/MAST motif, SVM module based on amino acid composition, SVM module based on dipeptide composition, Blast search on allergen representative peptides (ARPs). Only *Phaseolus vulgaris* showed a non-allergen base all algorithms. Also both plant asparaginases were found as non-allergen SVM module based on dipeptide composition with appropriate and acceptable scores while both bacterial asparaginases were seen as allergens in this algorithm (*Table 3*).

Search of samples performed in the SDAP allergens database. Results show that the FASTA alignments between the plant and bacteria asparaginases and all SDAP allergens have an E score higher than 0.010000 that is, none of them are allergens. The outcome of AllergenFP v1.0 was that with the exception of *Lupinus luteus* other asparaginases are probable allergens. The results from AllerTOP v. 2.0 are that bacterial samples are probable allergens and plants are non-allergen. These two last servers supported the results of AlgPred and VaxiJen servers about being none allergen plant asparaginases.

Prediction of Immunogenic Epitopes

Linar-B-cell Epitopes. ABCpred: In this array the predicted B cell epitopes are ranked according to their score obtained by trained recurrent neural network. Higher scores of the peptide mean higher probability of being an epitope. All the peptides shown here are above the threshold value of 0.8 while the default of ABCpred is 0.5. In plant asparaginases, sixteen peptide sequence were selected for any plant with a score between 0.8–0.96, as the most probable sequence that six of them was almost same. The common sequences in both plants to arrange the score amount included: "EASIMDGNTMKCGAVS", "SGRIGDSPLIGAGTYA", "MGGWAIAVHGGAGVDP", "TPLIGAGTYANELCAV", "VRELETDPLFNSGRGS" and "TGGLMNKMSGRIGDSP" (*supplemental Table 1*) highlighted with the same color. The start position of these common sequences in two plant species is close to each other. Needless to say, these two species in terms of phylogenic and aliment results for asparaginases do not have high affinity together (*Figure 1*).

The servers showed fifteen peptide sequences for *Erwinia chrysanthemi* and fourteen peptide sequences

Table 1
Genomic characteristic, number of L-asparaginase sequences and levels of alternative transcripts in plants, human, algae and bacteria were studied

Organism	GenBank ID	Type	Length (as)
Homo sapiens	AAM28434.1	asparaginase like protein	308
Homo sapiens	AAA35903.1	glycosylasparaginase	346
Shewanella baltica	KZK65943.1	L-asparaginase 1	337
Dickeya dadantii	OOC15512.1	L-asparaginase	348
Oribacterium	SEA64443.1	L-asparaginase	339
Enterococcus faecalis	OYN35025.1	L-asparaginase	321
Burkholderia	AMM17015.1	L-asparaginase	369
Escherichia coli	AAA23445.1	L-asparaginase II (ansB)	348
Erwinia teleogrylli	WP_058909779.1	L-asparaginase 1	338
Bacillus sonorensis	EME73692.1	L-asparaginase	329
Algoriphagus machipongonensis	EAZ80348.1	L-asparaginase	355
Bacillus thuringiensis	YP_035712.1	L-asparaginase	324
Alcaligenes faecalis	OSZ31330.1	L-asparaginase	329
Zostera marina	KMZ61400.1	Isoaspartyl peptidase/L-asparaginase 1	309
Carica papaya	XP_021909378.1	Isoaspartyl peptidase/L-asparaginase 1	305
Prunus persica	XP_007209366.1	Isoaspartyl peptidase/L-asparaginase 1	320
Phalaenopsis equestris	XP_020599431.1	Isoaspartyl peptidase/L-asparaginase 1	320
Phaseolus vulgaris	ABC01060.1	L-asparaginase 2	326
Cajanus cajan	KYP47966.1	Isoaspartyl peptidase/L-asparaginase	321
Ananas comosus	OAY72256.1	Isoaspartyl peptidase/L-asparaginase	322
Populus tomentosa	APR64035.1	L-asparaginase family protein	328
Brassica napus	XP_013715619.1	Isoaspartyl peptidase/L-asparaginase 1	315
Hevea brasiliensis	XP_021674871.1	Isoaspartyl peptidase/L-asparaginase 1	322
Pinus sylvestris	CAK22360.1	L-asparaginase	375
Sesamum indicum	XP_011086915.1	Isoaspartyl peptidase/L-asparaginase 1	321
Sorghum bicolor	XP_002464265.2	Isoaspartyl peptidase/L-asparaginase 1	423
Glycine max	AAM23265.1	L-asparaginase	326
Glycine max	NP_001236606.2	L-asparaginase	326
Glycine max	AFA35112.1	asparaginase 2	327
Medicago truncatula	AES59377.1	Isoaspartyl peptidase/L-asparaginase	325
Arabidopsis lyrata	EFH49598.1	L-asparaginase	315
Arachis duranensis	XP_015943800.1	Isoaspartyl peptidase/L-asparaginase	323
Auxenochlorella protothecoides	KFM25994.1	L-asparaginase 1	338
Dendrobium catenatum	XP_020704297.1	isoaspartyl peptidase/L-asparaginase	326
Hordeum vulgare	AAG28786.1	asparaginase	333
Dorcoceras hygrometricum	KZV21706.1	L-asparaginase	321
Asparagus officinalis	XP_020270247.1	Isoaspartyl peptidase/L-asparaginase 1	326
Glycine soja	KHN15197.1	Isoaspartyl peptidase/L-asparaginase	322
Helianthus annuus	XP_022039003.1	Isoaspartyl peptidase/L-asparaginase 1	313
Cucumis sativus	KGN55235.1	L-asparaginase	294
Momordica charantia	XP_022156717.1	Isoaspartyl peptidase/L-asparaginase 1	319
Lupinus angustifolius	P30364.1	Isoaspartyl peptidase/L-asparaginase 1	325
Ostreococcus tauri	OUS46332.1	L-asparaginase 1	404
Lupinus angustifolius developing seed	CAA43099.1	developing seed L-asparaginase	325
Helicosporidium	KDD76357.1	Isoaspartyl peptidase/L-asparaginase 1	401
Herrania umbratica	XP_021286498.1	Isoaspartyl	320
Populus trichocarpa	EEE81702.1	L-asparaginase family protein	328
Jatropha curcas	XP_012080428.1	Isoaspartyl peptidase/L-asparaginase 1	321
Klebsormidium nitens	GAQ80197.1	asparaginase	451
Micromonas commoda	ACO68792.1	asparaginase	479
Micromonas pusilla	XP_003055370.1	asparaginase	431
Morus notabilis	EXB89117.1	Isoaspartyl peptidase/L-asparaginase 1	316
Spinacia oleracea	XP_021855270.1	Isoaspartyl peptidase/L-asparaginase 1	318
Prunus avium	XP_021809982.1	Isoaspartyl peptidase/L-asparaginase 1	320
Arabidopsis thaliana	BAB02681.1	L-asparaginase (L-asparaginea midohydrolase)	325
Manihot esculenta	XP_021614687.1	Isoaspartyl peptidase/L-asparaginase 1	323
Triticum urartu	EMS61236.1	Isoaspartyl peptidase/L-asparaginase	343
Lupinus albus	AAA33409.1	L-asparaginase	325
Lupinus luteus	AAD03742.1	L-asparaginase	325

Figure 1
Phylogenetic tree of the Asp ortholog genes in human, bacteria, algae and plants. The tree was generated using the neighbor joining method and Mega 7 software based on homology between amino acid sequences of the Asp plants (40 sequences), algae (6 sequences) and bacteria (12 sequences) species. The bootstrap values are at the nodes

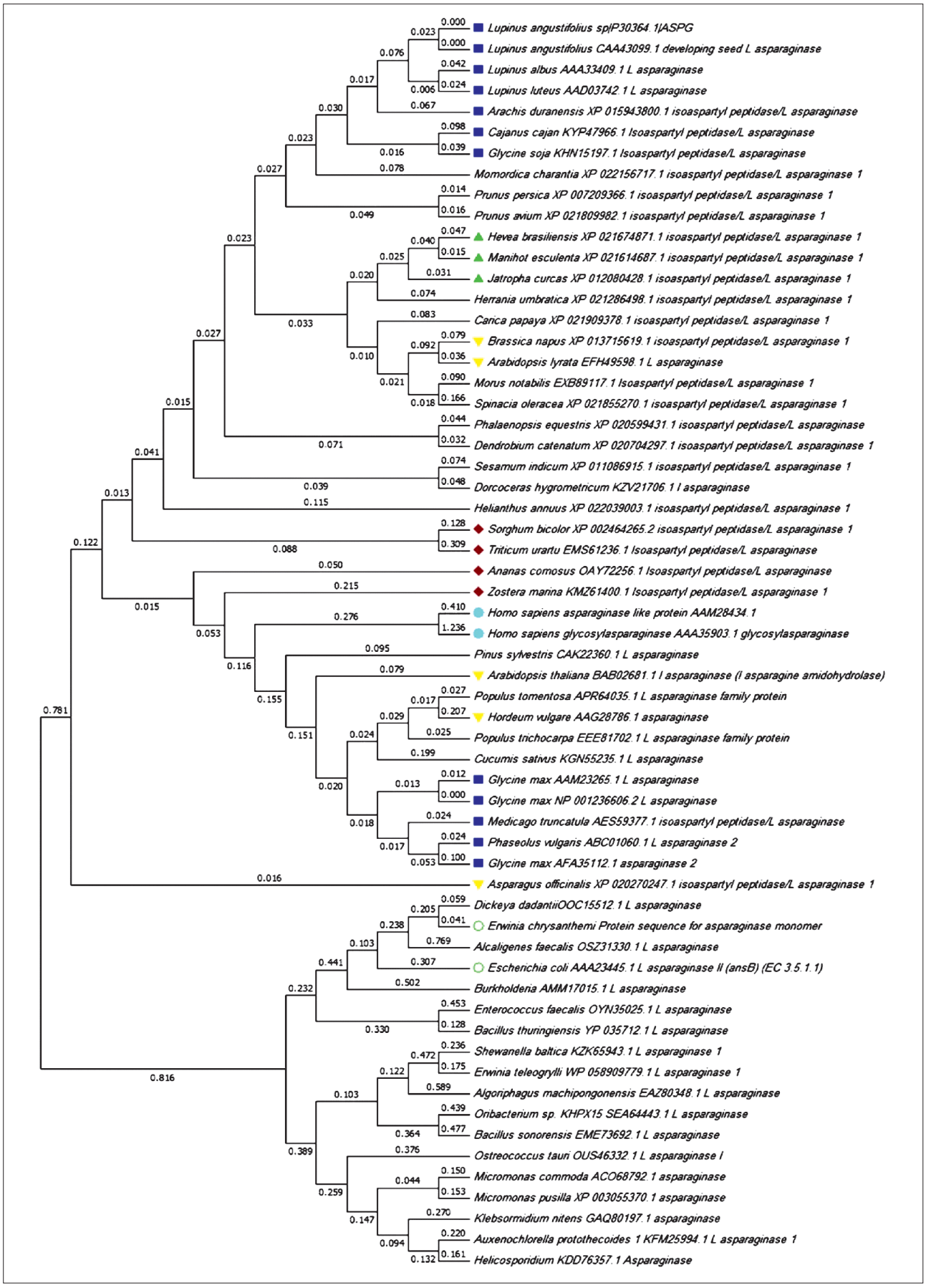


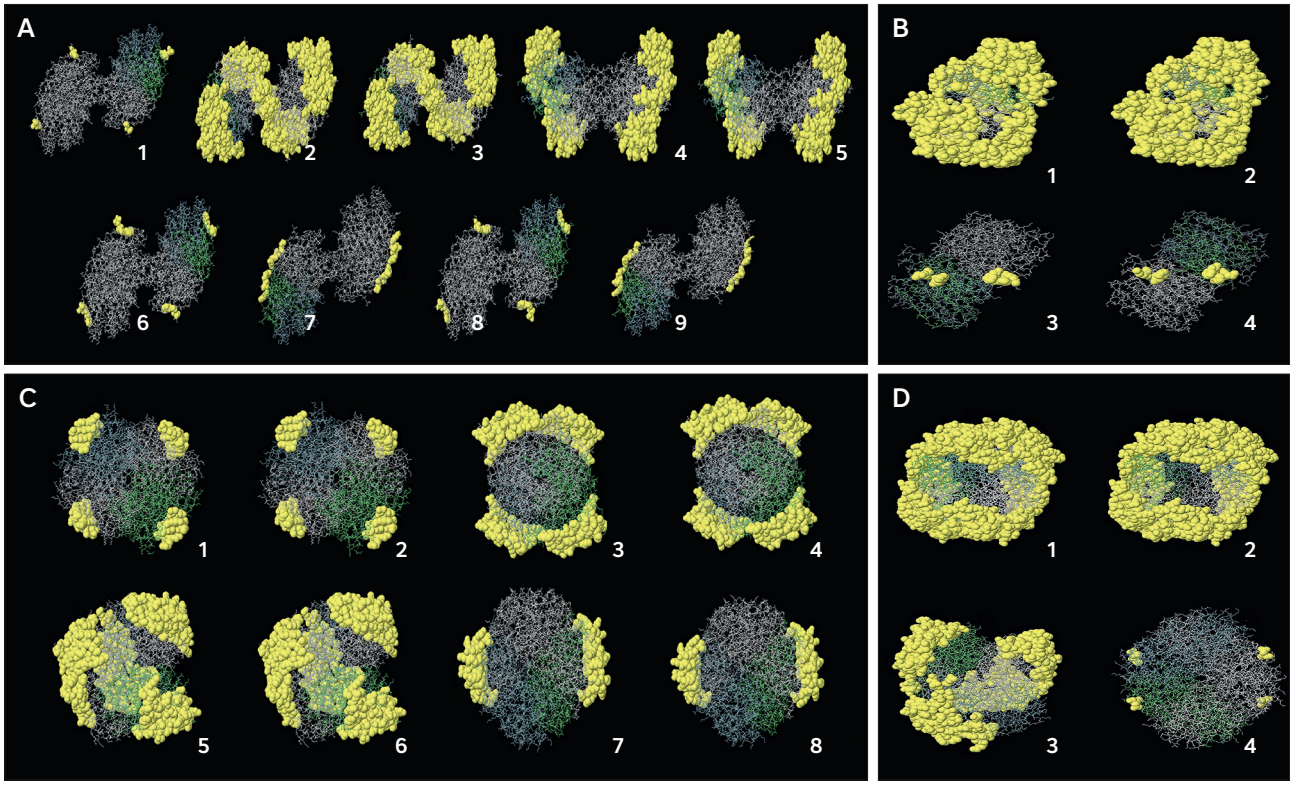
Table 2
VaxiJen models validation in two plants and two bacteria species

Organism	Overall Antigen Prediction	Type
Escherichia coli	0.5629	Antigen
Erwinia chrysanthemi	0.5574	Antigen
Phaseolus vulgaris	0.4961	Non-antigen
Lupinus luteus	0.4847	Non-antigen

Table 3
The possibility of allergens or non-allergens asparaginases in two plants and two bacteria species shown in terms of Mapping of IgE epitopes, PID, MEME/MAST motif and SVM module using AlgPred server

Organism	Mapping of IgE epitopes and PID	MEME/MAST motif	SVM module based on amino acid composition	Score Threshold= -0.4	SVM module based on dipeptide composition	Score Threshold= -0.2	Blast search on allergen representative peptides (ARPs)
Escherichia coli	Non-Allergen	Non- Allergen	Allergen	1.5802423	Allergen	1.3716288	Non-Allergen
Erwinia chrysanthemi	Non-Allergen	Non- Allergen	Allergen	-0.31327275	Allergen	0.13209808	Non-Allergen
Phaseolus vulgaris	Non-Allergen	Non- Allergen	Non- Allergen	-0.53965517	Non- Allergen	-0.49685789	Non-Allergen
Lupinus luteus	Non-Allergen	Non- Allergen	Allergen	-0.1842798	Non- Allergen	-0.37416826	Non-Allergen

Figure 2
3D Representation of the predicted discontinuous epitopes asparaginases protein of *Lupinus luteus* (**A**), *Phaseolus vulgaris* (**B**), *Erwinia chrysanthemi* (**C**) and *Escherichia coli* (**D**) as predicted by ElliPro



for *Escherichia coli* selected with a score of between 0.8–0.94 as the most probable sequences that four of them were almost the same but with different scores in each species such as: "TVKSDKPVVFVAAMRP", "PKVGIVYNYANASDLP", "PSTSMSADGPFNLYNA", "DGVVITHGTDTVEESA" and "TSLPKVDILYGYQDDP" (*supplemental Table 1*). These sequences had no similarity with the plant asparaginases. Based on the asparaginase bacteria being allergenic, the result of this can be compared, and we can postulate that the epitopes of bacterial allergens cannot be found in asparaginases plants whereas linar-B-cell epitopes, can be found in nearly every organism investigated, for the other enzyme, including animals such as mammals and insects as well as in fungi, plants, and bacteria.

BcePred Prediction Server. In this array, asparaginases peptide sequences are evaluated for the prediction of continuous B-Cell epitopes in antigenic sequences utilizing seven physico-chemical properties such as hydrophilicity, flexibility, accessibility, turns, exposed surface, polarity and antigenic propensity with 58.7% accuracy. Epitopes identified for every seven physico-chemical property for plant and bacteria asparaginases showed that these were comparable in terms of number, length and amino acid sequence.

Predicted Linear Epitope(s) with Ellipro. In the study of asparaginases with ElliPro epitope predictor program, 9 peptides with scores of 0.0744–0.788 and 6 peptides with scores of 0.0706–0.878 were detected in *Lupinus luteus* and *Phaseolus vulgaris* respectively whereas in *Escherichia coli* and *Erwinia chrysanthemi*, 7 and 5 peptides were distinguished with scores of 0.0702–0.868 and 0.08–0.912 respectively. Association of the obtained score for animal and plant species with regards to the 0.7 threshold of the server shows that the plant species in comparison to bacterial samples have less allergenicity, which is possibly due to the Ellipro server with *Phaseolus vulgaris* being better than any of the other plant species. Details shown in *Table 3 supplemental*, indicate that the starting point of the detected epitope sequence in plants and bacterial asparaginases are the same or are placed close to each other. The scores that are defined by the server for plant asparaginases shows that the rate of their allergies are close to the server threshold or slightly higher than it and that this rate in comparison to the score for bacterial epitopes did not show specified differences even though the distinguished epitopes is different between bacterial and plant species. *Figure 2* shows the results of the predicted discontinuous epitope(s) Ellipro, amino acid residues, the number of residues, sequence location as well as their scores which are tabulated in *Table 3 supplemental*. The graphical representation of the discontinuous epitopes is displayed in *Figure 2.* 3D Representation of the predicted discontinuous epitopes asparaginases protein predicted by ElliPro are presented based on the scores which are listed in *Table 3.* As seen in *Figure 2,* there are nine 3D structures for *Lupinus luteus*, four 3D structures for *Phaseolus vulgaris,* eight 3D structures for *Erwinia chrysanthemi* and four 3D structures for *Escherichia coli.* Their characteristics are presented in *Table 4 supplemental*.

Predicted with NetCTL-1.2. In the array of asparaginases with the NetCTL-1.2 program, 8 peptides with scores of 1.0281-3.1539 and 7 peptides with a score 0.7762-0.8558 were identified in *Lupinus luteus* and *Phaseolus vulgaris* respectively, whereas in *E. coli* and Erwinia 9 peptides were distinguished with scores of 1.1059-1.3235 and 0.9369-2.5751 respectively. The common sequences in both plants resulting in these scores includes: "VMDKSPHSY, EANTILFDY and ATEDGFMVV" as revealed in the *Table 4 supplemental*, highlighted with the same color. The start position of these common sequences in the two plant species is close to each other. The common sequence in the studied bacteria was presented with different scores in each species followed: "SADGPFNLY" and "ELPKVGIVY" (*supplemental Table 4)*. These sequences did not show any similarity between asparaginases of plants and bacteria. Based on the obtained results and scores for the detected peptides and the threshold number of 0.75 in NetCTL-1.2 programs it appears that *Phaseolus vulgaris* was the lowest-rated allergenic compared to bacterial and any other plant species.

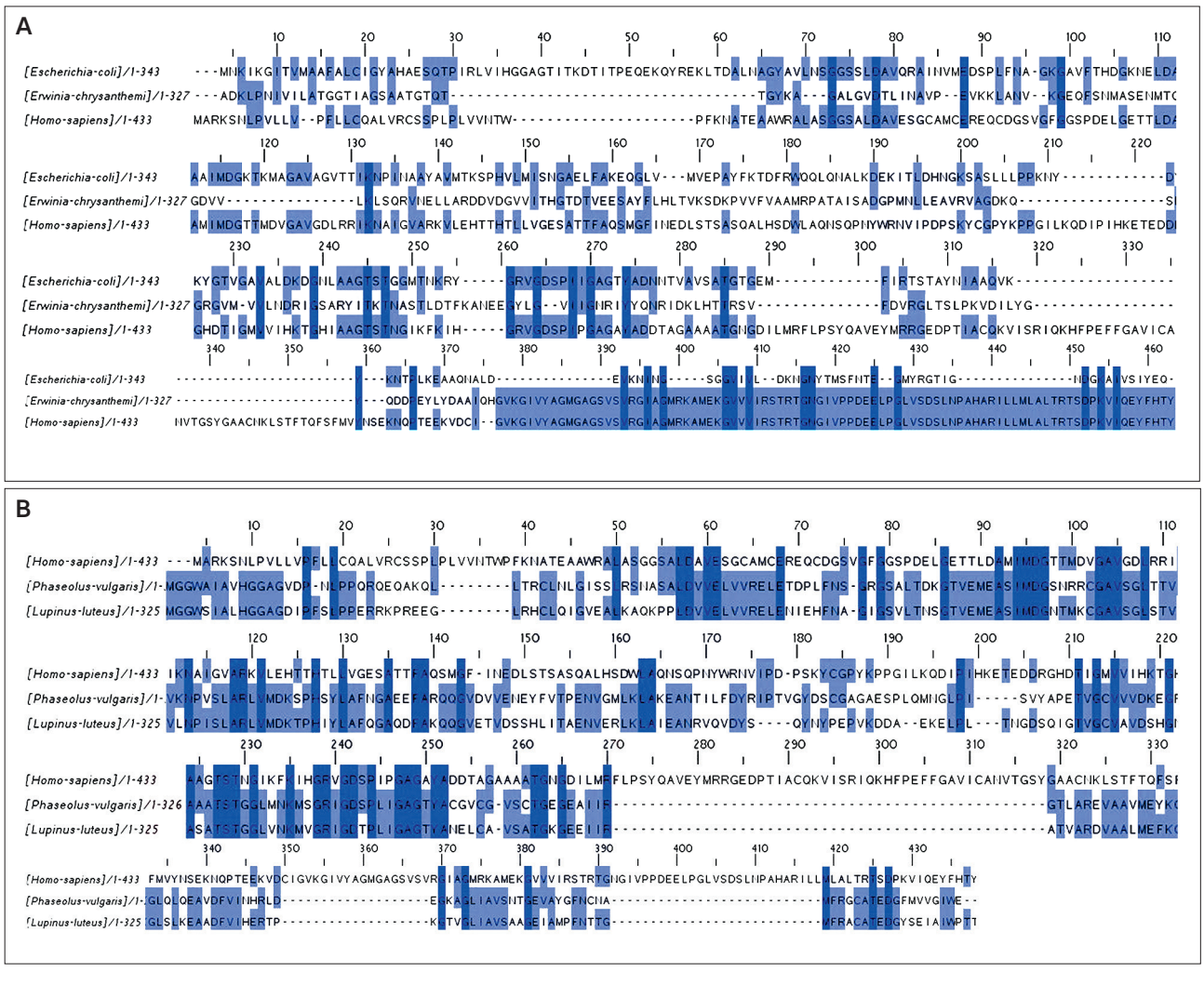
DISCUSSION

The survey showed that between the 40 plants, 2 identified items of human, 12 bacteria and 6 algae of asparaginase genes, generally two main branches created that samples of green algae is in the neighborhood of to the bacterial samples. Interestingly the data showed that the two bacterial samples of L-asparaginase (*Erwinia chrysanthemi* and *Escherichia coli*) (marked with the green hollow circle) used in humans as a therapeutic drug often resulting in allergies [29], in comparison to plant asparaginase genes, have less similarity to human asparaginase genes, while the two human asparaginase genes (marked with the green bold circle) are located perfectly between the plants group and their sequence reveals significant similarity with plant species (*Figure 1*). This leads to the point that various isoforms of asparaginase genes are present in plants. Database research confirmed the existence of isoforms of asparaginase genes in plants. Crystallography of this enzyme also supports two

types of isoforms of L-asparaginase enzymes which included potassium-independent plant asparaginase (in *Lupinus luteus*) and potassium-dependent plant-type L-asparaginase (in *Phaseolus vulgaris*) [30, 31]. A comparison of the three PvAspG1 structures setting the catalytic switch to ON (when K^+ is coordinated) or OFF (when Na^+ is coordinated) respectively allows or prevents anchoring of the reaction substrate/product in the active site [31]. Based on biochemical and molecular data, L-asparaginase has a variety of functions in plants, including catalytic activity, asparaginase activity, hydrolase activity (acting on carbon-nitrogen but not peptide bonds), in linear, amides, cellular component, intracellular, cell and cytoplasm [32]. Plant type L-asparaginases differ structurally and have a different evolutionary origin than bacterial L-asparaginases [30]. Alignment was performed between two plant L-asparaginases (*Lupinus luteus* and *Phaseolus vulgaris*), human L-asparaginase and L-asparaginases (*Erwinia chrysanthemi* and *Escherichia coli*) (*Figure 3*). Sequence alignment Protein BLAST results on NCBI between human asparaginase and two plant L-asparaginases (*Lupinus luteus* and *Phaseolus vulgaris*) and bacteria asparaginase (*Erwinia chrysanthemi* and *Escherichia coli*) showed *Phaseolus vulgaris* with identity 30%, Query cover 82% and E value 2e-25 having the highest similarity, *Lupinus luteus* with 30%, Query cover 76% and E value 4e-24 having the second similarity and *Erwinia chrysanthemi* with 33%, Query cover 6% and E value 4.3 having less similarity while *Escherichia coli* due to lack of similarity was removed from this alignment description.

Based on the x-ray structure, a three-dimensional model of the spatial conformation and catalytic site of the enzyme, Sanguinarine reductase in plant, has high structural similarity with enzymes of human and bacteria with similar functions as the plant homologs but bear little amino acid sequence similarity, in fact they have similar enzyme activities, i.e. mostly NAD-dependent sugar phosphate epimerases/dehy-

Figure 3
Multiple alignments of the amino acid sequences of Asps orthologous from (**A**) human and plants, (**B**) human and bacteria, using muscle alignment tool, as seen with Jalview tool. Residues shaded in blue are identical amino acids with other Asps protein sequences, focusing on the conserve motif regions *(*Accession numbers for the sequences shown are in *Table 1)*



dratases or hydroxysteroid dehydrogenases/isomerases [33].

Rosati et al. focused on the enzyme 5αR, pointing out new similarities between human and plant steroid metabolism. They showed the ability of the human isozymes of 5αR,to reduce campestenone, the natural substrate of the known plant enzyme DET2, the first direct evidence for human 5αR. Results of this research revealed that type 5αR human isozyme with higher affinity reduced DET2 as the plant substrate [34].

Although there was some allergen epitope found in plant asparaginase, these are different from the allergen epitopes of microbial asparaginase that are used as a drug in humans with no common sequence being found between them. VaxiJen server is a consistent and regular tool used for the forecasting of protective antigens. It can be used separately or in combination with other bioinformatics tools used for antigen site [22].

Study on plant allergens showed that dissimilar products can include similar allergenic proteins. Some allergen proteins that bound to IgE in allergic patients were identified including Jun a 3 (pollen of mountain cedar), Cap a 1 (bell pepper), Pru av 2 (cherry), Act c 2 (kiwi), Lyc e NP24 (tomato), and Mal d 2 (apple) that belonged to pathogenesis related proteins of group 5 (PR5) [35, 36]. SDAP covers data for above 800 allergens and wide bibliographic references in a relational database with contacts to other publicly available databases. SDAP is available without restrictions on the Web, allowing researchers to easily find structural and functional relations among known allergens and to recognize theoretically cross-reacting antigens [36]. Verma et al. processed the allergenicity prediction of transgenic proteins expressed in genetically modified crops using AlgPred server in order to tract the probability, reducing incorrect predictions to a great extent (74–78%) [37]. AlgPred server has been used for the calculation of new allergens in genetically modified (GM) foods and biopharmaceuticals [38].

Additionally, some enzymes have been identified as an allergen in fish, midges, crustaceans, and various plants based on *in silico* Prediction of T and B Cell Epitopes [39–42]. Research has been performed on predicting protective continuous B-cell epitopes on plant pathogen proteins using the ABCpred server [43].

The results (*supplemental Tables 2)* showed that some peptide sequences were classified on the basis of their physicochemical characteristics are the same in the two plant species and these sequences are different from the common sequences in bacteria asparaginases (*supplemental Tables 2)*. For example “TNSGTVE”, “ATSTGG” are common sequences in plant asparaginases and “RACATEDG” instead of “HGTDT” and “TKTN” are common sequences in bacteria asparaginases based on hydrophilicity features whereas “RELE”, “DKTPHIY” and “AQDFAKQQGVE” are common sequences in plants asparaginases instead of “ANVKGEQ”, “TVKSDKPVV” and “TKTN” which are common sequences in bacteria asparaginases based on accessibility features. The arrangement of antigenic elements in the epitopes and nonepitopes disregarding the antigen reconfiguration in Ag-Ab complex may not precisely reflect biological actuality [44]. The accurate identification of B-cell epitopes establishes a source for improvement of antibody therapeutics [45], and immunodiagnostic tools [46]. Scientists believe B-cell epitopes are arranged based on the three-dimensional structure as continuous (linear or sequential) and discontinuous (nonlinear or conformational) epitopes; in the latter case amino acid residues are in close contact due to the three-dimensional conformation. Amino acid sequence ranging from at least 20 to 400 amino acids is necessary for appropriate folding of the discontinuous epitope in native proteins. Many scientists assume that most of the identified linear antigenic determinants are a part of the conformational B-cell epitopes [47–49]. Descriptions are less stringent for stability, and confirmed that the majority of discontinuous epitopes (over 70%) have 1–5 linear segments with 1–6 amino acids [50]. Presently, epitopes play important roles in the course of cell development such as the human epidermal growth factor receptor 2 (HER2) and the extracellular domain of insulin-like growth factor-1 receptor which have been predicted by BcePred Prediction Server *in silico* methods [51, 52].

ElliPro together with other algorithms and web programs are able to predict the antibody epitopes in a protein sequence or structure. ElliPro presents the best epitope prediction in comparison to common methods based on structure. Another advantage of Ellipro is its ability to predict based on the geometrical properties of protein structure. Without involved preparation one can predict different kinds of protein-protein interactions [28]. Dall'Antonia et al., identified IgE-binding epitopes of two allergens by utilizing several servers, which expressed the highest specificity found by ElliPro for both studied allergens in comparison to other tools [53]. Jimenez-Lopez et al., were able to study Epitopes of β-conglutin proteins in *Lupinus angustifolius* with emphasis on cross-allergenic reactivity in comparison to some legumes [54]. Kundu et al. determined that two of these 10 fragments, exist in major immunogenic epitopes which are restricted on the outer surface of the protein MTC28 molecule of *Mycobacterium tuberculosis* which have key roles in the mycobacterial infection pathway. They were also able to predict the continuous epitope *in silico* with Ellipro [55].

The NetCTL 1.2 server can therefore be used for the prediction of epitopes by using cytotoxic T lymphocyte (CTL) that induced immunological tolerance in autoimmunity and allograft transplantation [56]. Larsen et al. (2007) showed that NetCTL-1.2 has an advanced predictive performance than other servers such as EpiJen, MAPPP, MHC-pathway, and WAPP [57]. CTL epitope prediction was an essential *in silico* tool in the allergenic research as it reduces the requirement for *in vitro* tests [56]. NetCTL-1.2 has been used as an integrative method for prediction of 9-mer CTL epitopes [58].

CONCLUSION

The data interestingly showed that the two bacterial samples of L-asparaginase used in humans as therapeutic drugs that often cause allergies compares well with the plant asparaginase genes and have less similarity to the human asparaginase. The two human asparaginase genes are located perfectly between the plant groups in the phylogenic tree. Sequence alignment results from the Protein BLAST on NCBI between human, two plants and two bacteria asparaginase showed that *Phaseolus vulgaris* has the highest similarity with human, while *Escherichia coli* due to lack similarity was removed from the alignment description. Results of Allergenic Sites Prediction showed that both asparaginase bacterial samples are antigen and both asparaginase plants are non-antigen in Vaxijen and AlgPred server, in the majority of tests. Some common sequences were detected as epitope by all of the three servers *ABCpred,* BcePred and *ElliPro* linear that was based on B-cells epitopes designed in asparaginases which is as follows: “IPFSLPPERRKPREEGLRHCL”, “ALKAQKPPLDV”, “NIEHFN”, “SHLITAENVERLKLAI” and “RACATE” epitopes in Lupinus luteus; “VDPTLPLERQEE” epitope in *Phaseolus vulgaris*; “VKCDKP” and “TSDTPFDVSKLNELPK” “VFDVRGLTSLPK” epitope in *Escherichia coli* and at least “LGVDTLINAVPEVKKLANV”, “DAAIQH”, “GVK” epitopes in *Erwinia chrysanthemi*. Predictions based on T-cells in NetCTL-1.2 server also showed that the *Phaseolus vulgar* has the lowest-rated allergies compared to bacterial and the other plant species. The data shows consistent scoring of various structures indicating a better presentation than any single term. On a whole it appears that the difference in allergenic scores between plants and microbial asparaginases rating based on the used software is remarkable. Although some allergen epitope are found in plant asparaginase, these are absolutely different from the allergen epitopes in microbial asparaginase that is used as a drug and results in allergies in humans. No common sequence could be found between them. The study of plant asparaginases indicates that the type of allergenic is completely different in plant asparaginase, it can be used as drug independent or use after emergence allergy of Erwinia asparaginase in patient. This manuscript affords an outline of available evidence proposing that *Phaseolus vulgaris* asparaginase has less epitopes and better scores in predicting tools and high similarity with the potential to be considered for drug design as enzymthrophy in leukemia and other cancers.

ACKNOWLEDGEMENTS
The authors gratefully acknowledge the support of Mohadeseh Dadmehr for making the graphic abstract.

CONFLICT OF INTEREST
The authors declare that there is no conflict of interest.

ORCID
Yazdi M. ORCID: http://orcid.org/0000-0003-3374-667X

FUNDING
This work was funded by a grant from the Shahid Chamran University of Ahvaz Research Council (Grant NO: 94/3/02/31580).

AUTHORSHIP
Kolahi carried genetic feature analysis and drafted the manuscript. Yazdi studied bioinformatics aspects. Foroghmand performed some analysis and visualization of results. Tabandeh provided suggestions for editing and revising the manuscript. All authors read and approved the final manuscript.

The online version of this article contains a date supplement.

References

1. Cachumba J.J.M., Antunes F.A.F., Peres G.F.D., Brumano L.P., Santos J.C.D., Da Silva S.S. Current applications and different approaches for microbial l-asparaginase production. Braz J Microbio 2016; 47: 77–85.
2. Ruma K., George T.K., Aswani P., Jisha M.S. Production and Optimization of Extra Cellular L-asparaginase by Fusarium solani Isolated from Withania sominifera. J Bio Active Pro Nat 2017; 7: 81–8.
3. Cho C., Lee H., Chung E., Kim K.M., Heo J.E., Kim J., et al. Molecular characterization of the soybean L-asparaginase gene induced by low temperature stress. Mol Cells 2007; 23: 280–6.
4. Lea P.J., Sodek L., Parry M.A.J., Shewry P.R., Halford N.G. Asparagine in plants. Ann Appl Biol 2007; 150: 1–26.
5. Canas R.A., Quillere I., Lea P.J., Hirel B. Analysis of amino acid metabolism in the ear of maize mutants deficient in two cytosolic glutamine synthetase isoenzymes highlights the importance of asparagine for nitrogen translocation within sink organs. Plant Biotechnol J 2010; 8: 966–78.

6. Vincze E., Reeves J.M., Lamping E., Farnden K.J., Reynolds P.H. Repression of the L-asparaginase gene during nodule development in Lupinus angustifolius. Plant Mol Biol 1994; 26: 303–11.
7. Fait A., Angelovici R., Less H., Ohad I., Urbanczyk-Wochniak E., Fernie A.R., Galili G. Arabidopsis seed development and germination is associated with temporally distinct metabolic switches. Plant physiology. J Plant Physiol 2006; 142: 839–54.
8. Credali A., García-Calderón M., Dam S., Perry J., Díaz-Quintana A., Parniske M., et al. The K+-dependent asparaginase, NSE1, is crucial for plant growth and seed production in Lotus japonicus. Plant Cell Physiol 2013; 54: 107–18.
9. UmaMaheswari T., Hemalatha T., Sankaranarayanan P., Puvanakrishnan R. Enzyme Therapy: Current Perspectives. Indian J Exp Biol 2016; 54: 7–16.
10. Panosyan E.H., Seibel N.L., Martin-Aragon S., Gaynon P.S., Avramis I.A., Sather H., et al. Asparaginase antibody and asparaginase activity in children with higher-risk acute lymphoblastic leukemia: Children's Cancer Group Study CCG-1961. J Pediatr Hematol Oncol 2004; 26: 217–26.
11. Larson R.A., Fretzin M.H., Dodge R.K., Schiffer C.A. Hypersensitivity reactions to L-asparaginase do not impact on the remission duration of adults with acute lymphoblastic leukemia. Leukemia 1998; 12: 660–5.
12. Woo M.H., Hak L.J., Storm M.C., Sandlund J.T., Ribeiro R.C., Rivera G.K., et al. Hypersensitivity or development of antibodies to asparaginase does not impact treatment outcome of childhood acute lymphoblastic leukemia. J Clin Oncol 2000;18: 1525–32.
13. Avramis V.I., Sencer S., Periclou A.P., Sather H., Bostrom B.C., Cohen L.J., et al. A randomized comparison of native Escherichia coli asparaginase and polyethylene glycol conjugated asparaginase for treatment of children with newly diagnosed standard-risk acute lymphoblastic leukemia: a Children's Cancer Group study. Blood 2002; 99: 1986–94.
14. Batool T., Makky E.A., Jalal M., Yusoff M.M. A Comprehensive Review on L-Asparaginase and Its Applications. Appl Biochem Biotechnol 2016; 178: 900–23.
15. Hung Wei-Lun, Joon Hyuk Suh, Yu Wang. Chemistry and health effects of furanocoumarins in grapefruit. J Food Drug Anal 2016; 25 (1): 71–83.
16. Henriksen L.T., Harila-Saari A., Ruud E., Abrahamsson J., Pruunsild K., Vaitkeviciene G., et al. PEG-asparaginase allergy in children with acute lymphoblastic leukemia in the NOPHO ALL2008 protocol. Pediatr Blood Cancer 2015; 62: 427–33.
17. Salzer W., Seibel N., Smith M. Erwinia asparaginase in pediatric acute lymphoblastic leukemia. Expert Opin Biol Ther 2012; 12: 1407–14.
18. Huby R.D.J., Dearman R.J., Kimber I. Why Are Some Proteins Allergens? Toxicol Sci 2000; 55: 235–46.
19. Lin C.H., Chen C.H., Lin Z.C., Fang J.Y. Recent advances in oral delivery of drugs and bioactive natural products using solid lipid nanoparticles as the carriers. J Food Drug Anal 2017; 25 (2): 219–34.
20. Kumar S., Stecher G., Tamura K. MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. Mol Biol Evol 2016; 33: 1870–4.
21. Waterhouse A.M., Procter J.B., Martin D.M., Clamp M., Barton G.J. Jalview Version 2-a multiple sequence alignment editor and analysis workbench. Bioinformatics 2009; 25: 1189–91.
22. Doytchinova I.A., Flower D.R. VaxiJen: a server for prediction of protective antigens, tumour antigens and subunit vaccines. BMC Bioinformatics 2007; 8: 4.
23. Saha S., Raghava G.P.S. AlgPred: prediction of allergenic proteins and mapping of IgE epitopes. Nucleic Acids Res 2006; 34: W202–W209.
24. Ivanciuc O., Mathura V., Midoro-Horiuti T., Braun W., Goldblum R.M., Schein C.H. Detecting potential IgE-reactive sites on food proteins using a sequence and structure database, SDAP-food. Journal of agricultural and food chemistry 2003; 51 (16): 4830–7.
25. Dimitrov I., Naneva L., Doytchinova I., Bangov I. AllergenFP: allergenicity prediction by descriptor fingerprints. Bioinformatics 2014; 30: 846–51.
26. Dimitrov I., Flower D.R., Doytchinova I. AllerTOP – a server for in silico prediction of allergens. BMC Bioinformatics 2013; 14: S4.
27. Saha S., Raghava G.P. Prediction of continuous B-cell epitopes in an antigen using recurrent neural network. Proteins 2006; 65: 40–8.
28. Ponomarenko J., Bui H.H., Li W., Fusseder N., Bourne P.E., Sette A., Peters B. ElliPro: a new structure-based tool for the prediction of antibody epitopes. BMC Bioinformatics 2008; 9: 514.
29. Aghaiypour K., Bahreini E., Jafari S. Bioinformatic Analysis of L-Asparaginase II from Citrobacter Freundii 1101, Erwinia Chrysanthemi DSM 4610, E. coli BL21 and Klebsiella Pneumoniae ATCC 10031. Int J Med Lab 2017; 4: 123–34.
30. Michalska K., Bujacz G., Jaskolski M. Crystal structure of plant asparaginase. J MOL BIOL 2006; 360: 105–16.
31. Bejger M., Imiolczyk B., Clavel D., Gilski M., Pajak A., Marsolais F., Jaskolski M. Na(+)/K(+) exchange switches the catalytic apparatus of potassium-dependent plant L-asparaginase. Acta Crystallogr D Biol Crystallogr 2014; 70: 1854–72.
32. Kersey P.J., Allen J.E., Allot A., Barba M., Boddu S., Bolt B.J., et al. Ensembl Genomes 2018: an integrated omics infrastructure for non-vertebrate species. Nucleic Acids Res 2017: gkx1011–gkx1011.
33. Vogel M., Lawson M., Sippl W., Conrad U., Roos W. Structure and Mechanism of Sanguinarine Reductase, an Enzyme of Alkaloid Detoxification. J Biol Chem 2010; 285: 18397–406.
34. Rosati F., Danza G., Guarna A., Cini N., Racchi M.L., Serio M. New Evidence of Similarity between Human and Plant Steroid Metabolism: 5α-Reductase Activity in Solanum malacoxylon. Endocrinology 2003; 144: 220–9.
35. Mari A. Importance of databases in experimental and clinical allergology. Int Arch Allergy Immunol 2005; 138: 88–96.
36. Schein C.H., Ivanciuc O., Braun W. Bioinformatics Approaches to Classifying Allergens and Predicting Cross-Reactivity. Immunol Allergy Clin North Am 2007; 27.
37. Verma A.K., Misra A., Subash S., Das M., Dwivedi P.D. Computational allergenicity prediction of transgenic proteins expressed in genetically modified crops. Immunopharmacol Immunotoxicol 2011; 33: 410–22.

38. Sircar G., Saha B., Bhattacharya S.G., Saha S. In silico prediction of allergenic proteins. Methods Mol Biol 2014; 1184: 375–88.
39. Sander I., Flagge A., Merget R., Halder T.M., Meyer H.E., Baur X. Identification of wheat flour allergens by means of 2-dimensional immunoblotting. J Allergy Clin Immunol Pract 2001; 107: 907–13.
40. Sudha V.T., Arora N., Gaur S.N., Pasha S., Singh B.P. Identification of a serine protease as a major allergen (Per a 10) of Periplaneta americana. Allergy 2008; 63: 768–76.
41. Li X., Yang H.W., Chen H., Wu J., Liu Y., Wei J.F. In silico prediction of T and B cell epitopes of Der f 25 in Dermatophagoides farinae. Int J Genomics 2014.
42. Mehla K., Ramana J. Surface proteome mining for identification of potential vaccine candidates against Campylobacter jejuni: an in silico approach. Funct Integr Genomics 2017; 17: 27–37.
43. Sollner J., Grohmann R., Rapberger R., Perco P., Lukas A., Mayer B. Analysis and prediction of protective continuous B-cell epitopes on pathogen proteins. Immunome Res 2008; 4: 1–17.
44. Zhao L., Wong L., Li J. Antibody-Specified B-Cell Epitope Prediction in Line with the Principle of Context-Awareness. IEEE/ACM Trans Comput Biol Bioinform 2011; 8: 1483–94.
45. Van Regenmortel MHV. Immunoinformatics may lead to a reappraisal of the nature of B cell epitopes and of the feasibility of synthetic peptide vaccines. J Mol Recognit 2006; 19: 183–7.
46. Leinikki P., Lehtinen M., Hyöty H., Parkkonen P., Kantanen M.L., Hakulinen J. Synthetic peptides as diagnostic tools in virology. Adv Virus Res 42 1993; 42: 149–86.
47. Laver W.G., Air G.M., Webster R.G., Smith-Gill S.J. Epitopes on protein antigens: Misconceptions and realities. Cell 1990; 61: 553–6.
48. Van Regenmortel M.H.V. Mapping Epitope Structure and Activity: From One-Dimensional Prediction to Four-Dimensional Description of Antigenic Specificity. Methods 1996; 9: 465–72.
49. Rubinstein N.D., Mayrose I., Halperin D., Yekutieli D., Gershoni J.M., Pupko T. Computational characterization of B-cell epitopes. Mol Immunol 2008; 45: 3477–89.
50. Potocnakova L., Bhide M., Pulzova L.B. An Introduction to B-Cell Epitope Mapping and In Silico Epitope Prediction. J Immunol Res 2016; 26 (4): 479–90.
51. Mahdavi M., Mohabatkar H., Keyhanfar M., Dehkordi A.J., Rabbani M. Linear and conformational B cell epitope prediction of the HER 2 ECD-subdomain III by in silico methods. Asian Pac J Cancer Prev 2012; 13 (7) : 3053–9.
52. Bayrami V., Keyhanfar M., Mohabatkar H., Mahdavi M., Moreau V. In silico prediction of B cell epitopes of the extracellular domain of insulin-like growth factor-1 receptor. Mol Biol Res Commun 2016: 5 (4): 201–14.
53. Dall'Antonia F., Gieras A., Devanaboyina S.C., Valenta R., Keller W. Prediction of IgE-binding epitopes by means of allergen surface comparison and correlation to cross-reactivity. J Allergy Clin Immunol Pract 2011; 128: 872–9 e878.
54. Jimenez-Lopez J.C., Lima-Cabello E., Melser S., Foley R.C., Singh K.B., Juan D. A. Lupin Allergy: Uncovering Structural Features and Epitopes of β-conglutin Proteins in Lupinus Angustifolius L. with a Focus on Cross-allergenic Reactivity to Peanut and Other Legumes. In: Ortuño F, Rojas I, eds. Bioinformatics and Biomedical Engineering: Third International Conference, IWBBIO 2015, Granada, Spain, April 15–17, 2015, Proceedings, Part I. Cham: Springer International Publishing, 2015: 96–107.
55. Kundu P., Biswas R., Mukherjee S., Reinhard L., Dutta A., Mueller-Dieckmann J., et al. Structure-based Epitope Mapping of Mycobacterium tuberculosis Secretary Antigen MTC28. J Biol Chem 2016; 291: 13943–54.
56. Mirza M.U., Rafique S., Ali A., Munir M., Ikram N., Manan A., et al. Towards peptide vaccines against Zika virus: Immunoinformatics combined with molecular dynamics simulations to predict antigenic epitopes of Zika viral proteins. Sci Rep 6 (2016): 37313.
57. Larsen M.V., Lundegaard C., Lamberth K., Buus S., Lund O., Nielsen M. Large-scale validation of methods for cytotoxic T-lymphocyte epitope prediction. BMC Bioinformatics 2007; 8: 424.
58. Larsen M.V., Lundegaard C., Lamberth K., Buus S., Brunak S., Lund O., Nielsen M. An integrative approach to CTL epitope prediction: a combined algorithm integrating MHC class I binding, TAP transport efficiency, and proteasomal cleavage predictions. Eur J Immunol 2005; 35: 2295–303.