International Journal of Medical Laboratory 2017;4(2): 123-134.

International Journal of Medical Laboratory
Shahid Sadoughi University of Medical Sciences

Original Article

Bioinformatic Analysis of L-Asparaginase II from *Citrobacter Freundii* 1101, *Erwinia Chrysanthemi* DSM 4610, *E.coli* BL21 and *Klebsiella Pneumoniae* ATCC 10031

Khosrow Aghaiypour[1]Ph.D., Elham Bahreini[2*]Ph.D., Shiva Jafari[1]M.Sc.

[1]Gene-Bank Department, Razi Vaccine and Serum Research Institute (RVSRI), Karaj, Iran.
[2]Department of Biochemistry, Faculty of Medicine, Iran University of Medical Sciences, Tehran, Iran.

A B S T R A C T

Article history
Received 2 Apr 2017
Accepted 14 May 2017
Available online 25 Jun 2017

Key words
Asparaginase
Bioinformatic
Citrobacter freundii
E.coli
Erwinia chrysanthemi
Klebsiella pneumoniae

Background and Aims: L-Asparaginase II is a cornerstone of treatment protocols for acute lymphoblastic leukemia. Only asparaginase II obtained from *E.coli* K12 and *Erwinia chrysanthemi* have been used in human as therapeutic drug. The therapeutic effects of asparaginase II from *E.coli* K12 and *Erwinia chrysanthemi* are accompanied by side effects. It is desirable to search for other asparaginase II sources with novel properties that could be therapeutic and produce an enzyme with less adverse effects.

Materials and Methods: Previously, we performed the *in vitro* studies, including cloning, sequencing and expression of L-asparaginase II genes (*ansB*) from *Citrobacter freundii* 1101, *Erwinia chrysanthemi* DSM 4610, *E.coli* BL21 and *Klebsiella pneumoniae* ATCC 10031. In this article, the obtained results were compared bioinformatically. The nucleotide and amino acid sequence alignments were carried out by ClustalW2. Protein localization and signal peptides were predicted by PSORT and SIG-Pred softwares, respectively. Percentages of hydrophobic and hydrophilic residues were calculated by Genscript software. The physicochemical parameters were computed using Expasy's ProtParam prediction server. The secondary and 3D structures were predicted by SOPMA and the online server Phyre2, respectively. The antigenicity of the asparaginase IIs was predicted using Semi-empirical method.

Results: *E.coli* BL21 and *Citrobacter freundii* 1101 had the most similarity in physicochemical parameters and antigenicity with *E.coli* K12. Also, *Erwinia chrysanthemi* DSM 4610 had the most similarity in physicochemical parameters and antigenicity with *Erwinia chrysanthemi*.

Conclusions: In spite of these similarities with drug types, the potentiality of other low-similar asparaginase IIs should also be determined and compared with drug types.

*C**orresponding Author**: Department of Biochemistry, Faculty of Medicine, Iran University of Medical Sciences, Tehran, Iran. **Tel:** +989352461622, **Fax:** +982188622742, **Email:** Elhambahreini@yahoo.com

Introduction

Significant increase in the amount of data available on the internet and in public databases combined with the increasing processing speed of workstations has created new opportunities for researchers to make scientific discoveries. The vast of biological data that has thus become accessible to the biological research community has considerably changed biologists' way of doing science. Recently, "In silico" methods and their utility are widely practical in protein and genome sequence analysis [1]. Computational software provides researchers to understand physicochemical and structural properties of protein [2]. A large number of online tools and servers are available from different sources for making prediction regarding the identification and structure of proteins. The various parameters like sequence length, number of amino acids and the physicochemical properties of a protein such as molecular weight (MW), atomic composition, extinction coefficient (EC), isoelectric point (PI), grand average of hydropathicity (GRAVY), aliphatic index (AI), instability index, etc. could be computed by various computational tools for the prediction and characterization of protein structure [3, 4]. The amino acid sequence provides most of the information required for determining and characterizing the molecule's function, physical and chemical properties.

L-asparaginase/L-glutaminase is a generic denomination for enzymes that catalyze the transformation of L-asparagine or L-glutamine into their respective acids and ammonia [5, 6]. These enzymes can be specific for asparagine, with negligible activity against glutamine, and thus termed asparaginases (EC 3.5.1.1), or can catalyze both asparagine and glutamine conversion, in which case they receive the denomination of glutaminase-asparaginases (EC 3.5.1.38) [7]. Based on the sequence homology analysis, as well as on biochemical and crystallographic data, available asparaginase sequences can be divided into three families. The first family corresponds to the bacterial-type asparaginases, the second to plant-type asparaginases and the third to enzymes similar to *Rhizobium etli* asparaginase [8].

Bacterial-type L-asparaginases can be further classified into two subtypes: type I and type II. Type I L-asparaginase was found to be expressed constitutively, whereas type II is induced by anaerobiosis. Only the type II Lasparaginases presents tumor inhibitory activity and, for this reason, have been extensively studied [7].

Bacterial L-asparaginases II have been used as therapeutic agents in treatment of acute childhood lymphoblastic leukemia. Its antileukemic effect is believed to be resulted from the depletion of circulating asparagine, which is not essential for normal cells, but essential for most malignant lymphoblastic cells [7, 9, 10]. Tumor-inhibitory asparaginases have also been isolated from a number of bacterial sources (such as *Proteus vulgaris, Corynebacterium glutamicum, Pseudomonas putida*, *Serratia marcescens* and others), but only the enzymes from *E.coli* strain K12 and *Erwinia chrysanthemi* strain (previously known as *Erwinia carotovora*) have been and are being frequently used in cancer therapy; because serious side effects like neurotoxicity, hepatitis and other dysfunctions due to intrinsic glutaminase activity restrict their clinical applications except for *E.coli* strain K12 and *Erwinia chrysanthemi* as they possess strong preference to asparagine over glutamine and show less severe immune related side effects [11-15]. Using a strategy based on the polymerase chain reaction, *ansB* gene of *Citrobacter freundii* strain 1101 [16], *Erwinia chrysanthemi* strain DSM 4610, *E.coli* strain BL21 and *Klebsiella pneumoniae* strain ATCC 10031 were cloned, sequenced and recorded in GenBank by Aghaiypour et al. in our previous study. Most of the available researches on asparaginase II are focused on identification, purification and application of the protein. In this article, it was used some bioinformatics software and servers to characterize asparaginase II from these bacteria species and compared with two therapeutic asparaginase IIs from *E.coli* strain K12 and *Erwinia chrysanthemi* [17-19]. This bioinformatics and in silico study could be a rapid method to analyse, compare, predict and estimate before starting *in vivo* study and drug design.

Materilas and Methods

Sequences *ansB* genes and protein construction

The complete nucleotide and related amino acid sequences of *ansB genes* from *Citrobacter freundii* 1101, *Erwinia chrysanthemi* DSM 4610, *E.coli* BL21, *Klebsiella pneumoniae* ATCC 10031, E.coli K12 and *Erwinia chrysanthemi* were retrieved from GeneBank using their accession numbers of EU624347, JF972567, FJ643626, FJ189504, P00805 and P06608, respectively. These protein sequences were retrieved in FASTA format and used for further analysis. The sequence alignment for the mentioned genes was performed by ClustalW2, a multiple sequence alignment program. The alignment results were compared with *E.coli* K12 and *Erwinia chrysanthemi*.

Computational tools and servers

The amino acid composition of the obtained sequences were analyzed using the bioinformatic tools. Protein localization was predicted by PSORT software. SIG-Pred online software used for predicting signal peptides and possible cleavage positions. Percentages of hydrophobic and hydrophilic residues were calculated from the primary structure analysis by Genscript software. The physicochemical parameters, theoretical PI, MW, total number of positive (+R) and negative (-R) residues, EC, half-life, instability index and GRAVY were computed using Expasy's ProtParam prediction server. GRAVY value for a peptide or protein is calculated as the sum of hydropathy values of all the amino acids, divided by the number of residues in the sequence. The self optimized prediction method with alignment (SOPMA) method was used for the secondary structure prediction. The modelled 3D structure was generated using the online server Phyre2.

Results

Sequence alignments of *ansB*-nucleotide and asparaginase II-amino acid are given in table 1 and table 2, respectively. The scores describe the percent of similarity in sequences. Table 3 shows the amino acid composition (in%) and table 4 determines the percent of the hydrophilic and hydrophobic residue contents in the asparaginase II proteins using ProtParam and Genscript softwares, respectively. Parameters computed using Expasy's ProtParam tool were represented in table 5. ProtParam tool computes different physicochemical parameters depending on the queries submitted to the databases. The formula (n.atoms), MW, absorption 0.1% (or 1g/l) 280 nm, theoretical PI, number of +R (positive residues: Arg+Lys) and -R (negative residues: Asp+Glu), EC, half-life, instability index, GRAVY and AI were depicted in this table. Table 6 shows signal peptide sequences predicted by SIG-Pred software. The results of secondary structure of asparaginase II proteins predicted by SOPMA are represented in table 7. The 3D structure of asparaginase IIs′ monomer was predicted by the online server Phyre2 was illustrated in figure 1. The antigenicity of the asparaginase IIs from was predicted using semi-empirical method. The obtained antigenicity scores were 1.0327 for *Erwinia chrysanthemi*, 1.0270 for *E.coli* K12, 1.0279 for *Citrobacter freundii* 1101, 1.0380 for *Klebsiella pneumoniae* ATCC 10031, 1.0314 for *Erwinia chrysanthemi* DSM 4610 and 1.0253 for *E.coli* BL21, respectively. The possible localizations of all asparaginase IIs are given in table 8. As predicted by PSORT server the studied asparaginase IIs are periplasmic protein.

Table 1. Score table of multiple nucleotide sequence alignment.

Name	Length	Name	Length	Score
***Cit.f.* 1101**	1047	***Er.c. DSM***	1039	53.0
***Cit.f.* 1101**	1047	***E.coli* BL21**	1125	56.0
***Cit.f.* 1101**	1047	***Kle.pne***	1005	61.0
***Cit.f.* 1101**	1047	***E.coli.* K12**	1045	82.0
***Cit.f.* 1101**	1047	***Erw.C***	1044	61.0
***Er.c.* DSM**	1039	***E.coli* BL21**	1125	60.0
***Er.c.* DSM**	1039	***Kle.pne***	1005	49.0
***Er.c.* DSM**	1039	***E.coli* K12**	1045	51.0
***Er.c.* DSM**	1039	***Erw.c***	1044	51.0
***E.coli* BL21**	1125	***Kle.pne***	1005	53.0
***E.coli* BL21**	1125	***E.coli* K12**	1045	53.0
***E.coli* BL21**	1125	***Erw.c***	1044	53.0
Kle.pne	1005	***E.coli* K12**	1045	62.0
Kle.pne	1005	***Erw.c***	1044	70.0
***E.coli* K12**	1045	***Erw.c***	1044	61.0

Cit.f. 1101= *Citrobacter freundii* 1101; *Er.c.* DSM= *Erwinia chrysanthemi* DSM 4610; *Kle.pne*= *Klebsiella pneumoniae* ATCC 10031. Score describes the percent of similarity in nucleotide sequence

Table 2. Score table of multiple protein sequence alignment

Name	Length	Name	Length	Score
***Cit.f.* 1101**	348	***Erw.c.DSM***	348	43.0
***Cit.f.* 1101**	348	***E.coli* BL21**	348	93.0
***Cit.f.* 1101**	348	***Kle. pne***	334	45.0
***Cit.f.* 1101**	348	***E.coli* K12**	348	92.0
***Cit.f.* 1101**	348	***Erw.c***	348	46.0
***Erw.c.*DSM**	348	***E.coli* BL21**	348	43.0
***Erw.c.*DSM**	348	***Kle. pne***	334	69.0
***Erw.c.*DSM**	348	***E.coli* K12**	348	43.0
***Erw.c.*DSM**	348	***Erw.c***	348	92.0
***E.coli* BL21**	348	***Kle. pne***	334	46.0
***E.coli* BL21**	348	***E.coli* K12**	348	98.0
***E.coli* BL21**	348	***Erw.c***	348	45.0
Kle. pne	334	***E.coli* K12**	348	45.0
Kle. pne	334	***Erw.c***	348	70.0
***E.coli* K12**	348	***Erw.c***	348	45.0

*Cit.f.*1101= *Citrobacter freundii* 1101; *Erw.c.* DSM= *Erwinia chrysanthemi* DSM 4610; *Kle.pne*= *Klebsiella pneumoniae* ATCC 10031. Score describes the percent of similarity in amino acid sequence.

Table 3. Amino acid composition (in %) of the asparaginase IIs using ProtParam software

Amino acid composition	*E.coli* K12	*Erw.c.*	*Cit.f.* 1101	*Erw.c.* DSM 4610	*E.coli* BL21	*Kle. pne.* ATCC 10031
Ala (A)	38(10.9%)	33(9.5%)	39(11.2%)	34(9.8%)	39(11.2%)	40 (12.0%)
Arg (R)	8(2.3%)	19(5.5%)	9(2.6%)	20(5.7%)	8(2.3%)	16(4.8%)
Asn (N)	23(6.6%)	14(4.0%)	20(5.7%)	16(4.6%)	23(6.6%)	11(3.3%)
Asp (D)	27(7.8%)	21(6.0%)	24(6.9%)	21(6.0%)	28(8.0%)	20(6.0%)
Cys (C)	2(0.6%)	0(0.0%)	2(0.6%)	0(0.0%)	2(0.6%)	0(0.0%)
Gln (Q)	13(3.7%)	8(2.3%)	10(2.9%)	8(2.3%)	13(3.7%)	14(4.2%)
Glu (E)	7(2.0%)	15(4.3%)	11(3.2%)	13(3.7%)	7(2.0%)	12(3.6%)
Gly (G)	30(8.6%)	32(9.2%)	31(8.9%)	33(9.5%)	30(8.6%)	28(8.4%)
His (H)	3(0.9%)	6(1.7%)	4(1.1%)	7(2.0%)	3(0.9%)	7(2.1%)
Ile (I)	13(3.7%)	20(5.7%)	12(3.4%)	19(5.5%)	13(3.7%)	22(6.6%)
Leu (L)	26(7.5%)	31(8.9%)	26(7.5%)	33(9.5%)	26(7.5%)	25(7.5%)
Lys (K)	24(6.9%)	18(5.2%)	24(6.9%)	16(4.6%)	24(6.9%)	14(4.2%)
Met (M)	8(2.3%)	10(2.9%)	9(2.6%)	11(3.2%)	8(2.3%)	8(2.4%)
Phe (F)	11(3.2%)	11(3.2%)	11(3.2%)	11(3.2%)	11(3.2%)	8(2.4%)
Pro (P)	13(3.7%)	12(3.4%)	13(3.7%)	13(3.7%)	13(3.7%)	16(4.8%)
Ser (S)	17(4.9%)	20(5.7%)	16(4.6%)	19(5.5%)	16(4.6%)	20(6.0%)
Thr (T)	35(10.1%)	26(7.5%)	36(10.3%)	25(7.2%)	35(10.1%)	28(8.4%)
Trp (W)	1(0.3%)	1(0.3%)	1(0.3%)	1(0.3%)	1(0.3%)	0(0.0%)
Tyr (Y)	12(3.4%)	13(3.7%)	12(3.4%)	14(4.0%)	12(3.4%)	10(3.0%)
Val (V)	37(10.6%)	38(10.9%)	38(10.9%)	34(9.8%)	36(10.3%)	35(10.5%)

*Cit.f.*1101= *Citrobacter freundii* 1101; *Erw.c.*= *Erwinia chrysanthemi*; *Kle.pne*= *Klebsiella pneumoniae* ATCC 10031.

Table 4. Hydrophilic and hydrophobic residue contents in asparaginase IIs using Genscript software

Bacteria	Percentage of hydrophilic residue	Percentage hydrophobic of residue	Other	Net hydrophobic content
***Cit.f.*1101**	21%	43%	36%	High
***Erw.c.* DSM**	22%	45%	33%	High
***E.coli* BL21**	20%	42%	38%	High
Kle. pne	21%	46%	33%	High
***E.coli* K12**	20%	42%	38%	High
Erw.c	23%	45%	32%	High

*Cit.f.*1101= *Citrobacter freundii* 1101; *Erw.c.* DSM= *Erwinia chrysanthemi* DSM 4610; *Kle.pne*= *Klebsiella pneumoniae* ATCC 10031.

Table 5. The computed parameters using Expasy's ProtParam tool

Bacteria	Formula(n.atoms)	MW	Abs. 0.1% (280 nm) (1cm)	Theo. pI	-R	+R	EC $M^{-1}cm^{-1}$	II	GRAVY	AI
***Cit.f.*1101**	$C_{1626}H_{2602}N_{438}O_{513}S_{11}$ (5190)	147.39	0.635	6.12	35	33	23380	23.16	-0.076	85.46
***Erw.c.* DSM**	$C_{1674}H_{2681}N_{463}O_{499}S_{11}$ (5328)	150.52	0.701	8.56	34	36	26360	17.58	0.009	96.38
***E.coli* BL21**	$C_{1622}H_{2591}N_{439}O_{518}S_{10}$ (5180)	147.40	0.634	5.68	35	32	23380	18.33	-0.128	84.91
Kle. pne	$C_{1563}H_{2531}N_{435}O_{482}S_{8}$ (5019)	141.54	0.421	6.39	32	30	14900	29.12	0.050	97.25
***E.coli* K12**	$C_{1623}H_{2595}N_{439}O_{517}S_{10}$ (5184)	147.40	0.634	5.96	34	32	23380	18.27	-0.128	85.46
Erw.c	$C_{1673}H_{2692}N_{458}O_{502}S_{10}$ (5335)	150.3	0.662	7.84	36	37	24870	17.20	0.042	98.30

*Cit.f.*1101= *Citrobacter freundii* 1101; *Erw.c.* DSM= *Erwinia chrysanthemi* DSM 4610; *Kle.pne*= *Klebsiella pneumoniae* ATCC 10031; MW= Molecular weight; Abs= Absorption; pI= Theoretical isoelctric point; +R= positive residues: Arg+Lys; -R= negative residues: Asp+Glu; EC= Extinction coefficient; II= Instability index; GRAVY= Grand average of hydropathicity; AI= Aliphatic index

Table 6. Signal peptide sequence predicted by SIG-Pred software

Bactria	Number of aa	Signal peptide sequence	Active mature enzyme (one sub U)
***Cit.f.*1101**	19 or 22	MEFFKRTALAALVMGFSGA|ALA	329 or 326 aa
***Erw.c.* DSM**	21	MERWFKSLFVMVLFFVFTANA	327 aa
***E.coli* BL21**	19 or 22	MEFFKKTALAALVMGFSGA|ALA	329 or 326 aa
Kle. pne	29	MSSLAFSETRLPHIVILATGGTIAGSAAS	305 aa
***E.coli* K12**	19 and 22	MEFFKKTALAALVMGFSGA|ALA	329 or 326 aa
Erw.c	21 and 22	MERWFKSLFVLVLFFVFTA|SAA	327 or 328 aa

*Cit.f.*1101= *Citrobacter freundii* 1101; *Erw.c.* DSM= *Erwinia chrysanthemi* DSM 4610; *Kle.pne*= *Klebsiella pneumoniae* ATCC 10031; aa= Amino acid

Table 7. The percentage of the secondary structures in asparaginase IIs

Bactria	Alpha helix	Beta sheet	Beta turn	Random coil
***Cit.f.*1101**	39.66%	18.97%	5.46%	35.92%
***Erw.c.*DSM**	38.22%	17.53%	4.60%	39.66%
***E.coli* BL21**	39.08%	19.25%	5.46%	36.21%
Kle. pne	35.33%	18.56%	5.39%	40.72%
***E.coli* K12**	37.93%	19.54%	4.31%	38.22%
Erw.c	33.33%	20.40%	6.32%	39.94%

*Cit.f.*1101= *Citrobacter freundii* 1101; *Erw.c.* DSM= *Erwinia chrysanthemi* DSM 4610; *Kle.pne*= *Klebsiella pneumoniae* ATCC 10031

Table 8. Prediction of asparaginase IIs′ localization by PSORT

Bactria	Periplasmic space, Certainty	Outer membrane, Certainty	Inner membrane, Certainty	Cytoplasm, Certainty
***Cit.f.*1101**	0.568	0.371	0.000	0.000
***Erw.c.* DSM**	0.649	0.324	0.000	0.000
***E.coli* BL21**	0.383	0.361	0.000	0.000
Kle. pne	0.100	0.000	0.000	0.000
***E.coli* K12**	0.396	0.371	0.000	0.000
Erw.c	0.637	0.390	0.000	0.000

*Cit.f.*1101= *Citrobacter freundii* 1101; *Erw.c.* DSM= *Erwinia chrysanthemi* DSM 4610; *Kle.pne*= *Klebsiella pneumoniae* ATCC 10031

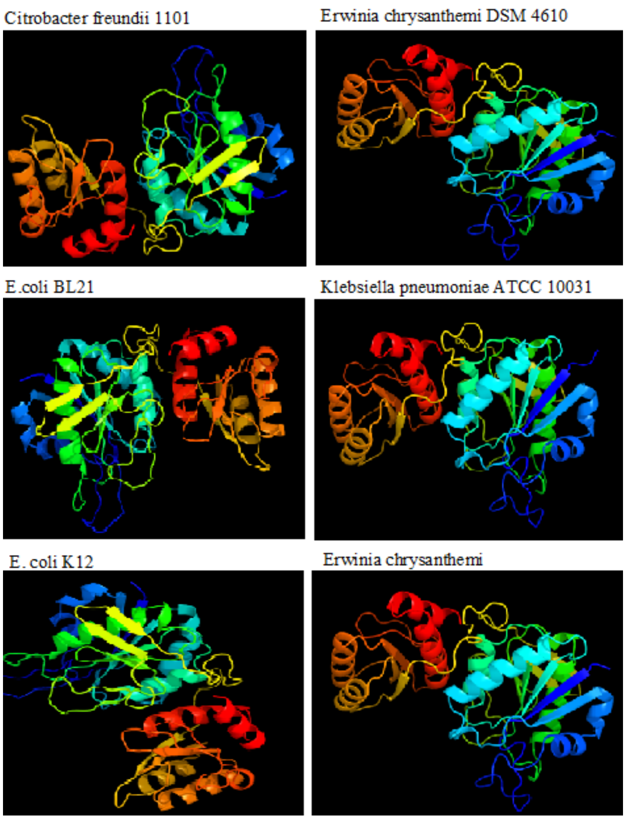

Fig.1. 3D structures of asparaginase IIs, predicted by online server Phyre2

Discussion

Asparaginase II from various sources has been studying for a long time in different fields such as structure and its potency as an antitumor agent. For example, Swain et al. designed the crystal structure of L-asparginase [20]; Guo et al. compared the antitumor activity and the effect of recombinant enzyme both *in vitro* and *vivo* [21]. Mohamed et al. isolated *E.coli* MG27 from the River Nile, amplified and cloned *asnB* and then characterized it by DNA sequencing and bioinformatics software [22]. In addition to such experimental studies that are high cost, time-consuming and may be impossible to apply for various organisms at the same time, bioinformatics provide an inexpensive and rapid approach to study biomolecules from various organisms without any vital changes in protein structure and activity. In this in silico study, we analyzed and compared asparaginase II from four bacteria that were cloned, sequenced and expressed in our lab and recorded in GenBank, previously.

Despite the high differences in the alignment scores of nucleotide sequences, the alignment scores of amino acid sequences were not very different among the mentioned bacteria (Tables 1 and 2). Because of variation in amino acid codons, assessment of amino acid sequence could be more reliable than nucleotide sequences. *Citrobacter freundii* 1101 and *E.coli* BL21 has the highest score alignment with *E.coli* K12 in amino acid sequences. *Erwinia chrysanthemi* DSM 4610 hass the highest score alignment with *Erwinia chrysanthemi* in amino acid sequences. The amino acid sequence provides most of the information required for determining and characterizing the molecule's function, physical and chemical properties.

It may be seemed from the results of the primary analysis (Tables 3 and 4) that all asparaginase II proteins are hydrophobic in primary nature due to the presence of high non-polar residues, but below GRAVY analysis show other results. The low presence of cysteine residues or its absence in asparaginase II (<0.6%) indicates the lack of disulphide bridges ("SS" bonds).

The parameters in table 5 were computed using Expasy's ProtParam tool. The estimated molecular weights are in the range of 141.54- 150.52 KD with the lowest MW for *Klebsiella pneumoniae* ATCC 10031 and the highest MW for *Erwinia chrysanthemi* DSM 4610, respectively. The isoelctric point is the pH at which the surface protein is covered with charge, but the net charge of protein is zero. The pI calculation is based on the peptide sequence alone without considering the effect of modifications. It is calculated using pKa values of amino acids,s which depend on their side chain [23]. The pI of a protein greater than seven (pI>7) indicates that they are basic and lower than seven (pI<7) indicates that they are acidic. Asparaginase II from *E.coli* BL21 with the lowest computed pI=5.68 is acidic and asparaginase II from *Erwinia chrysanthemi* DSM 4610 with the highest computed pI=8.56 is basic. The computed pI has useful application such as developing buffer system for separation by isoelectric focusing method and purification preparation by chromatography technique [24]. The EC illustrates how much light a protein absorbs relative to their composition at a certain wavelength. ProtParam computes EC for a range of (276 nm, 278 nm, 219 nm, 280 nm and 282 nm) wave length, 280 nm is favorable because proteins absorb strongly there with respect to the concentration of Cys, Trp and Tyr, while other substances commonly in protein solution do not [25]. EC of asparaginase IIs at 280 nm were ranging from 14900 $M^{-1}cm^{-1}$ to 26360$M^{-1}cm^{-1}$ with the lowest range for *Klebsiella pneumoniae* ATCC 10031 and the highest range for *Erwinia chrysanthemi* DSM 4610. The computed protein EC helps in the quantitative study of protein-protein and protein-ligand interactions in solution. The instability index provides an estimate of the stability of a protein in a test tube. This method assigns a weight value of instability. A protein whose instability index is smaller than 40 is predicted as stable, a value above 40 predicts that the protein may be unstable [26]. The instability index for all mentioned asparaginase IIs is smaller than 40.

GRAVY that is defined at one specific position in a sequence is the mean value of the hydrophobicity of the amino acids within a window, usually 19 residues long, around each position [27]. Hydrophobic-hydrophilic interactions have a strong impact on the three dimensional structure a protein will adopt. Because structure, not amino acid sequence order, carries out certain functions it is important to understand how these forces affect the protein folding process. GRAVY is the average value of the hydropathy index at each position. The hydropathy values range from -2 to +2 for most proteins, with the positively rated proteins being more hydrophobic [28]. GRAVY scores of some asparaginase IIs are negative and others approximately around the zero in the range of -0.128 to 0.050. Then, among the study asparaginase IIs, ones from *E.coli* strains of K12, BL21 and *Citrobacter freundii* 1101 with the negative scores are more hydrophilic and asparaginase IIs from *Erwinia chrysanthemi*, *Erwinia chrysanthemi* DSM 4610 and *Klebsiella pneumoniae* ATCC 10031 with the positive score are more hydrophobic molecules. AI described as the relative volume occupied by aliphatic side chains of the amino acids such as alanine, valine, isoleucine and leucine, is regarded as a positive factor for the increase of thermal stability of globular proteins. The AI of proteins from thermophilic bacteria was found to be significantly higher than that of ordinary proteins and hence, it can serve as a measure of thermostability of proteins [29, 30]. High aIvalues of the analyzed asparaginase IIs showed an increase in the thermostability of the proteins. The half-life is a prediction of the time it takes for half of the amount of protein in a cell to disappear after its synthesis in the cell. The ProtParam biocomputed half-life of all six asparaginase II is 30 hours *in vitro* and more than 10 hours *in vivo*.

Signal peptides occur in bacterial periplasmic and secretory proteins. The primary analysis by SIG-Pred software suggested a signal peptide at the N terminal and the subsequent fragment as active unit in the mature enzyme (Table 6). Secondary structure features as predicted using SOPMA indicated whether a given amino acid lied in a helix, strand or coil and also the percentage of amino acids located in these structures. The results revealed that alpha helix and random coils were dominated among secondary structure elements followed by beta sheet and beta turns for all sequences. All known types of asparaginase II are active as homotetramers [31, 32]. The 3D structure of asparaginase IIs′ monomer was predicted by the online server Phyre2 (Fig. 1). In the case of 3D structure, hydrophilic domains tend to be on exterior surface, while hydrophobic domains avoids external environment and forms internal core of the protein. The closest interactions between two pairs of subunits lead to the formation of two intimate dimmers within which the four non-allosteric catalytic centers are created. Such formation of tetramers, for reasons that are not completely clear, appear to be essential for the catalytic ability of asparaginase II [33, 34]. PSORT server is the most precise bacterial localization prediction tool available [35]. The localization of all asparaginase IIs were predicted by PSORT server as periplasmic protein (Table 8).

Antigenicity is a local property of the protein sequences and that protein sequence properties of composition, secondary structure, solvent accessibility and evolutionary conservation are the determinants of antigenicity and specificity in immune response [30]. The predicted antigenicity scores of all asparaginase IIs were comparable with each other and with therapeutic asparaginase IIs from *Erwinia chrysanthemi* and *E.coli* K12. Although these similarities suggest that asparaginase IIs from *E.coli* BL21, *Cit.f.*1101 and *Erwinia chrysanthemi* DSM 4610 would be the choice for *in vivo* study, the potentiality of other non-similar asparaginase IIs with drug types should also be determined and compared; maybe others being more potential than drug types.

Conclusions

We compared bioinformatics data from six asparaginase IIs by online predictor softwares. The aim of the present study was to get how much non-therapeutic asparaginase IIs that we had studied, sequenced and cloned in our previous researches would be similar to the therapeutic asparaginase IIs from *E.coli* K12 and *Erwinia chrysanthemi*. *E.coli* BL21 and *Citrobacter freundii* 1101 had the most similarity in amino acid sequence (98% and 92%, respectively), GRAVY, MW, AI and antigenicity with *E.coli* K12. Also, *Erwinia chrysanthemi* DSM 4610 had the most similarity in amino acid sequence (92%), GRAVY, MW, AI and antigenicity with *Erwinia chrysanthemi*. Although such in silico analysis provides valuable information, more study is needed to reduce the drug's toxicity and improve its potency without any vital changes in protein structure and activity by means of bioinformatics because it is possible to reduce the costs of study and produce new variants of drug.

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgments

The basic of our study was supported by Razi Vaccine and Serum Research Institute (Iran). Thanks to the members of the biotechnology department of Razi Vaccine and Serum Research Institute for their favors in improving this study.

References

[1]. Ekins S, Mestres J, Testa B. In silico pharmacology for drug discovery: methods for virtual ligand screening and profiling. Br J Pharmacol. 2007; 152(1): 9-20.

[2]. Gentleman RC, Carey VJ, Bates DM, Bolstad B, Dettling M, Dudoit S, et al. Bioconductor: open software development for computational biology and bioinformatics. Genome Biol. 2004; 5(10): R80.

[3]. Katoh K, Misawa K, Kuma KI, Miyata T. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. Nucleic Acids Res. 2002; 30(14): 3059-3066.

[4]. Gasteiger E, Hoogland C, Gattiker A, Duvaud SE, Wilkins MR, Appel RD, et al. Protein identification and analysis tools on the ExPASy server. Humana Press; 2005.

[5]. Pasut G, Sergi M, Veronese FM. Anti-cancer PEG-enzymes: 30 years old, but still a current approach. Adv Drug Delivery Rev. 2008; 60(1): 69-78.

[6]. Narta UK, Kanwar SS, Azmi W. Pharmacological and clinical evaluation of L-asparaginase in the treatment of leukemia. Critic Rev Oncol/Hematol. 2007; 61(3): 208-21.

[7]. Sanches M, Krauchenco S, Polikarpov I. Structure, substrate complexation and reaction mechanism of bacterial asparaginases. Curr Chem Biol. 2007; 1(1): 75-86.

[8]. Borek10 D, Jaskólski M. Sequence analysis of enzymes with asparaginase activity. Acta Biochim Pol. 2001; 48: 893-902.

[9]. Cedar H, Schwartz JH. Localization of the two L-asparaginases in anaerobically grown Escherichia coli. J Biologic Chem. 1967; 242(16): 3753-755.

[10]. Graham ML. Pegaspargase: a review of clinical studies. Adv Drug Delivery Rev. 2003; 55(10): 1293-302.

[11]. Ramya LN, Doble M, Rekha VP, Pulicherla KK. L-Asparaginase as potent anti-leukemic agent and its significance of having reduced glutaminase side activity for better treatment of acute lymphoblastic leukaemia. Appl Biochem Biotechnol. 2012; 167(8): 2144-159.

[12]. Kumar DS, Sobha K. L-Asparaginase from microbes: a comprehensive review. Adv Biores. 2012; 3(12): 137-57.

[13]. El-Bessoumy AA, Sarhan M, Mansour J. Production, isolation, and purification of L-asparaginase from Pseudomonas aeruginosa 50071 using solid-state fermentation. BMB Reports. 2004; 37(4): 387-93.

[14]. Hüser A, Klöppner U, Röhm KH. Cloning, sequence analysis, and expression of *ansB* from Pseudomonas fluorescens, encoding periplasmic glutaminase/asparaginase. FEMS Microbiol Lett. 1999; 178(2): 327-35.

[15]. Heinemann B, Howard AJ. Production of tumor-inhibitory L-asparaginase by submerged growth of Serratia marcescens. Appl Microbiol. 1969; 18(4): 550-54.

[16]. Bahreini E, sabaghi A, Aghaeipour Kh. Cloning, sequencing and expression of L-Asparaginase II gene from *Citrobacter freundii* 1101. International Journal of Biology Research 2017; 2(2): 7-12

[17]. Choi JH, Lee SY. Secretory and extracellular production of recombinant proteins using Escherichia coli. Appl Microbiol Biotech. 2004; 64(5): 625-35.

[18]. Kotzia GA, Labrou NE. L-Asparaginase from *Erwinia chrysanthemi* 3937: cloning, expression and characterization. Journal of biotechnology 2007; 127(4): 657-69.

[19]. Harms E, Wehner A, Jennings MP, Pugh KJ, Beacham IR, Rohm KH. Construction of expression systems for Escherichia coli asparaginase II and two-step purification of the recombinant enzyme from periplasmic extracts. Protein Expr Purifi. 1991; 2(2-3): 144-50.

[20]. Swain AL, Jaskolski M, Housset D, Rao JK, Wlodawer A. Crystal structure of Escherichia coli L-asparaginase, an enzyme used in cancer therapy. Proc Natl Acad Sci USA 1993; 90(4): 1474-478.

[21]. Guo QL,Wu MS, Chen Z. Comparison of antitumor effect of recombinant L-asparaginase with wild type one *in vitro* and *in vivo*. Acta Pharmacol Sin. 2002; 23(10): 946-51.

[22]. Mohamed ZK, Elnagdy SM, Seufi AE, Gamal M. Cloning and molecular analysis of Lasparaginase II gene (*ansB*). J BioSci Biotechnol. 2015; 4(3): 291-302.

[23]. Pace CN, Grimsley GR, Scholtz JM. Protein ionizable groups: pK values and their contribution to protein stability and solubility. J Biologic Chem. 2009; 284(20): 13285-3289.

[24]. Gasteiger E, Gattiker A, Hoogland C, Ivanyi I, Appel RD, Bairoch A. ExPASy: the proteomics server for in-depth protein knowledge and analysis. Nucleic Acids Res. 2003; 31(13): 3784-788.

[25]. Schmid FX. Biological Macromolecules: UV-visible Spectrophotometry. Encyc Life Sci (eLS) 2001; 3: 240-43.

[26]. Guruprasad K, Reddy BB, Pandit MW. Correlation between stability of a protein and its dipeptide composition: a novel approach for predicting *in vivo* stability of a protein from its primary sequence. Protein Eng. 1990; 4(2): 155-61.

[27]. Kyte J, Doolittle RF. A simple method for displaying the hydropathic character of a protein. J Mol Biol. 1982; 157(1): 105-32.
[28]. Dehouck Y, Gilis D, Rooman M. Database-derived potentials dependent on protein size for In silico folding and design. Biophys J. 2004; 87(1): 171-81.
[29]. Atsushi IK. Thermostability and aliphatic index of globular proteins. J Biochem. 1980; 88(6): 1895-898.
[30]. Ferron F, Longhi S, Canard B, Karlin D. A practical overview of protein disorder prediction methods. Proteins 2006; 65(1): 1-4.
[31]. Greenquist AC, Wriston JC. Chemical evidence for identical subunits in L-asparaginase from Escherichia coli B. Arch Biochem Biophys. 1972; 152(1): 280-86.
[32]. Swain AL, Jaskólski M, Housset D, Rao JK, Wlodawer A. Crystal structure of Escherichia coli L-asparaginase, an enzyme used in cancer therapy. Proceed Nati Acad Sci. 1993; 90(4): 1474-478.
[33]. Mezentsev YV, Molnar AA, Sokolov NN, Lisitsina VB, Shatskaya MA, Ivanov AS, et al. Specificity of molecular recognition in oligomerization of bacterial L-asparaginases. Biochemistry (Moscow) Supplement Series B: Biomed Chem. 2011; 5(2): 124.
[34]. Michalska K, Jaskolski M. Structural aspects of L-asparaginases, their friends and relations. ACTA Biochemical Polonica-English Edition 2006; 53(4): 627.
[35]. Gardy JL, Laird MR, Chen F, Rey S, Walsh CJ, Ester M, et al. PSORTb v. 2.0: expanded prediction of bacterial protein subcellular localization and insights gained from comparative proteome analysis. Bioinformatics 2005; 21(5): 617-23.