

RESEARCH COMMUNICATION

In silico identification and characterization of antineoplastic asparaginase enzyme from endophytic bacteria

Elaheh Zadeh Hosseingholi¹ | Neda Neghabi¹ | Ghader Molavi^{2,3} |
Seyed Mohammad Gheibi Hayat⁴  | Hamed Shahriarpour¹

¹Department of Biology, Faculty of Basic Sciences, Azarbaijan Shahid Madani University, Tabriz, Iran

²Drug Applied Research Center, Tabriz University of Medical Sciences, Tabriz, Iran

³Department of Molecular Medicine, Faculty of Advanced Medical Sciences, Tabriz University of Medical Sciences, Tabriz, Iran

⁴Department of Medical Genetics, School of Medicine, Shahid Sadoughi University of Medical Science, Yazd, Iran

Correspondence

Elaheh Zadeh Hosseingholi, Department of Biology, Faculty of Basic Sciences, Azarbaijan Shahid Madani University, P. O. Box: 5375171379, 35 km Road Tabriz-Maragheh, Tabriz, East Azarbaijan, Iran.
Email: zhosseingholi@gmail.com

Abstract

It is generally accepted that L-asparagine is an important amino acid required for the fast growth of cells. Cancerous cells receive this amino acid from extracellular sources. The depletion of L-asparagine from its surrounding environments by asparaginase enzyme can be used as a therapeutic strategy in cancer patients. This therapeutic enzyme is produced commercially mainly from bacteria such as *Escherichia coli* and *Erwinia chrysanthemi*. The side effects of such drugs have persuaded scientists to find new enzyme sources. In this study, in silico approach was applied to investigate L-asparaginase producing endophytic bacteria that produce more compatible enzymes within the body. Protein-protein basic local alignment search tool with *E. coli* and *E. chrysanthemi* asparaginase enzyme sequences against 262 endophytic bacteria were performed. The results with identity more than 35%, coverage more than 80%, and *E*-value less than 10^{-4} were selected. Then, some of bioinformatics tools were used to characterize them. A total of nine sequences consisting of seven known and two hypothetical proteins were identified in six bacterial species. The results showed that some of the asparaginase enzymes produced by endophytic bacteria possess more suitable immunological indices compared with asparaginase enzymes of *E. coli* and *E. chrysanthemi*. *Herbaspirillum rubrisubalbicans* was predicted to produce a nonallergen and nonantigen asparaginase enzyme. The number of antigenic determinants was predicted to be lower in asparaginase enzymes produced by *Bacillus amyloliquefaciens*, *H. rubrisubalbicans*, and *H. seropedicae*. Moreover, the number of high-scored B-cell epitopes was lower in enzyme sequences related to the mentioned bacteria and *Paenibacillus polymyxa*. The number of discontinuous epitopes and the number of

Abbreviations: 3D, three-dimensional; BLASTp, protein-protein basic local alignment search tool; HP, hypothetical protein; IC50, 50% inhibitory concentration; MHCI, major histocompatibility complex class I; MHCII, major histocompatibility complex class II; nr, nonredundant; PGPR, plant growth-promoting rhizobacteria; ProSA, protein structure analysis; TMHMM, transmembrane helices hidden Markov model.

T-cell epitopes were lower in *B. amyloliquefaciens* produced enzymes. Therefore, the therapeutic use of these enzymes is possible.

KEYWORDS

bacterial endophytes, bioinformatic software, cancer, immunoinformatics, in silico, L-asparaginase

1 | INTRODUCTION

Cancer is a serious disease and one of the most important causes of death in the world.¹ One of the problems in the treatment of this disease is the unwanted effects of drugs on the normal cells of the body.² Cancer cells differ from the normal ones in some metabolic pathways. The inability to synthesize certain amino acids and the need to obtain them from the surrounding environment is one of these differences.³ Therefore, depleting the essential amino acids from the environment surrounding the cancerous cells can serve as a simple strategy to destroy these cells through delayed protein, DNA and RNA synthesis, and apoptosis of the cells, without harming normal cells.⁴ Diets alone do not fulfill this goal. Therefore, the injection of amino acid degrading enzymes can be helpful. One example is the application of L-asparaginase (EC 3.5.1.1) as a chemotherapeutic agent in the treatment of human cancers, especially acute lymphoblastic leukemia⁵ and non-Hodgkins lymphoma.⁶ L-asparaginase can affect the survival of abnormal cells by converting asparagines to L-aspartic acid and ammonium. Asparaginase is a member of the amidase or amidohydrolase (EC 3.5.1.4) enzymes.⁷ Many researchers have reported the wide distribution of L-asparaginase in plants, animals, rodents, and microorganisms, including bacteria, fungi, and actinomycetes.⁸ Currently, most of the patients with lymphomas receive the first phase of the commercial L-asparaginase enzyme produced by *Escherichia coli* (EcAII), approved by the food and drug administration, as a medicine. In half of the patients, severe immune responses to this enzyme or even types of conjugated with polyethylene glycol have been seen.⁹ As a solution, these patients have been provided with the L-asparaginase enzyme extracted from *Erwinia chrysanthemi* which has no cross-reactivity with antibodies produced against EcAII. However, it is immunogenic and also less effective than the previous enzyme.¹⁰ So far, almost all enzyme engineering experiments have focused on these two enzymes.¹¹ The microbial world is a rich source of enzymes with different structural and functional properties, and it is possible to isolate L-asparaginase with better therapeutic properties and provide ideas for better enzyme engineering.¹² Endophytes are a diverse group of microorganisms located in the internal tissues of leaves, roots, or stems in healthy plants. More than 200 bacterial genera have been identified to be endophytes. Endophytes have attracted scientists'

attention in the biotechnological sector recently.¹³ It is reported that the endogenous natural products produced by endophytes have unique structures and high biological activity, and these bacteria are a huge source of useful materials in medicine, agriculture, and the industry.^{14,15} Since the discovery of taxol production by an endophytic fungus, *Taxomyces andrenae*,¹⁶ more studies are conducted on the use of endogenous-derived bioactive substances in medicine.¹⁷ To the best of our knowledge, there is no report about in silico investigation of finding more human body compatible asparaginase enzymes from endophytic bacteria source. Therefore, in this study immunological properties of enzymes from this origin were explored through the application of some immunoinformatics tools. The obtained results were compared with ones related to two commercially available therapeutic enzymes. The main goal of this study was to provide a prospective in probable production of the new recombinant asparaginase as a drug instead of ones produced by *E. coli* and *E. chrysanthemi*.

2 | METHODS

2.1 | Sequence retrieval

The amino acid sequences of *E. coli* asparaginase (UniProt: P00805) and *E. chrysanthemi* (UniProt: P06608) was used as a query in protein-protein basic local alignment search tool (BLASTp) (<http://blast.ncbi.nlm.nih.gov/Blast.cgi>)¹⁸ against nonredundant (nr) protein sequence database with endophytic bacteria as a selected organism. The list of 262 selected endophytic bacteria used in this study was obtained through a literature review (S1).^{19,20} The results with identity more than 35%, coverage more than 80%, and E-value less than 10^{-4} were selected. Since the identification of some endophytic bacteria had been reported in the genus level in previous studies, the isolation sources of bacteria which met defined criteria were checked using BioSample database (<https://www.ncbi.nlm.nih.gov/biosample/>),²¹ and the sequences of endophytic species were considered as a positive result. The analyses described hereunder were performed for selected sequences and *E. coli* and *E. chrysanthemi* asparaginase sequences simultaneously, to introduce the best candidate by comparing the results.

2.2 | Physicochemical characterization

Basic physicochemical properties such as molecular weight, extinction coefficient, isoelectric point, aliphatic index, grand average of hydropathicity, and instability index were estimated by ProtParam tool (<http://web.expasy.org/protparam/>).²²

2.3 | Complementary sequence analysis for hypothetical proteins

If a hypothetical protein (HP) protein exists among chosen sequences, further analysis will be required to ensure its potential to have asparaginase activity. Therefore, for HPs, the following analyses were performed. The prediction of expression probability of HP sequences via codon adaptation index was revealed using CAIcal tool.²³ Subcellular localization of proteins was predicted by CELLO (<http://cello.life.nctu.edu.tw/>)²⁴ and PSLpred (<http://www.imtech.res.in/raghava/>).²⁵ Involvement of HPs in nonclassical protein secretion was predicted using SecretomeP 2.0 (<http://www.cbs.dtu.dk/services/SecretomeP/>).²⁶ To predict probable transmembrane helices in HPs, transmembrane helices hidden Markov model (TMHMM) (<http://www.cbs.dtu.dk/services/TMHMM/>)²⁷ server was used. To protein domains annotation and domain architectures analysis, simple modular architecture research tool (<http://smart.embl-heidelberg.de/>)²⁸ was used.

2.4 | Tertiary structure prediction, structure validation, and quality prediction

The three-dimensional (3D) structure of a protein affects its shape, size, and function. For Three-dimensional protein structure predictions, (PS)2-v3: template-based protein structure prediction server (<http://ps2v3.life.nctu.edu.tw/>)²⁹ was used. Ramachandran plot (<http://mordred.bioc.cam.ac.uk/~rapper/rampage.php>)³⁰ and the Z-score calculated by protein structure analysis (ProSA) web (<https://prosa.services.came.sbg.ac.at/prosa.php>)³¹ were checked to measure the quality of the generated models.

2.5 | Immunological properties

Allergenicity of selected sequences was predicted by Allpred sever (<http://www.imtech.res.in/raghava/allpred/submission.html>) prediction.³²

Antigenic regions of sequences were analyzed using Bioinformatic Antigenic Peptide Prediction (BPAP) tool (<http://imed.med.ucm.es/Tools/antigenic.html>).³³ The

VaxiJen v2.0 server (<http://www.ddg-pharmfac.net/vaxijen/VaxiJen/VaxiJen.html>) was used to predict the productivity of proteins. Sequences were subjected to predict linear B-cell epitopes using ABCreds (<http://crdd.osdd.net/raghava/abcpred/>).³⁴ Conformational discontinuous B-cell epitopes were predicted by ElliPro tool (<http://tools.iedb.org/ellipro/>)³⁵ using three-dimensional protein structures. MHC class I and class II binding regions within sequences were revealed by ProPredI (<http://crdd.osdd.net/raghava/propred1/>)³⁶ and ProPred (<http://crdd.osdd.net/raghava/propred/>)³⁷ servers predictions, respectively. The 50% inhibitory concentration (IC₅₀) of epitopic regions within the protein sequences which bind to highest number of major histocompatibility complex class I (MHCI) and major histocompatibility complex class II (MHCII) alleles was calculated using quantitative prediction of binding affinity of peptide-MHC binding site, MHCpred (<http://www.ddg-pharmfac.net/mhcpred/MHCpred/>).³⁸ IC₅₀ values less than 100 nM for allele HLA-A* 0201 in MHCI binding epitopes³⁹ and allele DRB1*0101 in MHCII binding epitopes⁴⁰ were selected as a strong binder epitope.

3 | RESULTS AND DISCUSSION

3.1 | Sequence retrieval

The results of BLASTp of the amino acid sequence of *E. coli* and *E. chrysanthemi* asparaginase sequences against nonredundant protein sequences of endophytic bacteria showed that nine protein sequences from six endophytic bacteria five different genera met determined criteria. The accession number of selected proteins and the name of the bacteria they belong to are given in Table 1. As can be seen from the table, two gram-positive

TABLE 1 List of bacteria and their related protein candidates which met determined criteria in BLASTp

Bacteria	Protein accession number
<i>Azorhizobium caulinodans</i>	BAF90379.1
<i>Bacillus amyloliquefaciens</i>	ERK81845.1 (hypothetical), KNX35781.1
<i>Bradyrhizobium japonicum</i>	BAL10002.1 (hypothetical), AHY50093.1
<i>Herbaspirillum rubrisubalbicans</i>	ALU91046.1, ALU91730.1
<i>Herbaspirillum seropedicae</i>	ADJ65404.1
<i>Paenibacillus polymyxa</i>	KOS01895.1

Abbreviation: BLASTp, protein-protein basic local alignment search tool.

TABLE 2 List of predicted physicochemical parameters by ExPASy's ProtParam tool for asparaginases of *E. coli* and *E. chrysanthemi* and candidate proteins

Protein name	Length (amino acid)	Grand average of hydropathicity	Extinction coefficient	Aliphatic index	Instability index	Half-life	Theoretical isoelectric point	Total number of positively charged residues (Arg + Lys)	Total number of negatively charged residues (asp + Glu)	Molecular mass
<i>E. coli</i> asparaginase	348	-0.114	23,505 23,380	85.46	18.27	30 hr (mammalian reticulocytes, in vitro), >20 hr (yeast, in vivo), >10 hr (<i>E. coli</i> , in vivo).	5.96	32	34	36,850.69
<i>E. chrysanthemi</i> asparaginase	327	-0.042	19,370	97.77	16.37	4.4 hr (mammalian reticulocytes, in vitro), >20 hr (yeast, in vivo), >10 hr (<i>E. coli</i> , in vivo).	7.23	35	35	35,054.09
BAF90379.1	359	0.238	22,585 22,460	97.74	34.20	30 hr (mammalian reticulocytes, in vitro), >20 hr (yeast, in vivo), >10 hr (<i>E. coli</i> , in vivo).	5.88	27	31	36,699.97
ERK81845.1	265	-0.362	16,390	83.96	8.77	30 hr (mammalian reticulocytes, in vitro), >20 hr (yeast, in vivo), >10 hr (<i>E. coli</i> , in vivo).	9.11	34	29	28,692.52
KNX35781.1	371	-0.225	17,880 17,880	83.72	19.76	30 hr (mammalian reticulocytes, in vitro), >20 hr (yeast, in vivo), >10 hr (<i>E. coli</i> , in vivo).	9.15	43	36	39,509.05
AHY50093.1	361	-0.050	29,450	91.97	27.72	30 hr (mammalian reticulocytes, in vitro), >20 hr (yeast, in vivo), >10 hr (<i>E. coli</i> , in vivo).	5.39	35	40	37,909.94
BAL10002.1	361	-0.056	29,450	91.16	27.88	30 hr (mammalian reticulocytes, in vitro), >20 hr (yeast, in vivo), >10 hr (<i>E. coli</i> , in vivo).	5.39	35	40	37,942.00
ALU91046.1	275	-0.072	22,920	97.49	17.20	30 hr (mammalian reticulocytes, in vitro), >20 hr (yeast, in vivo), >10 hr (<i>E. coli</i> , in vivo).	6.16	28	29	29,837.17
ALU91730.1	349	0.192	17,210 16,960	101.55	39.81	30 hr (mammalian reticulocytes, in vitro), >20 hr (yeast, in vivo), >10 hr (<i>E. coli</i> , in vivo).	5.63	31	35	37,219.91
ADJ65404.1	275	-0.069	22,920	98.55	15.13	30 hr (mammalian reticulocytes, in vitro), >20 hr (yeast, in vivo), >10 hr (<i>E. coli</i> , in vivo).	6.75	29	29	29,812.15
KOS01895.1	388	-0.031	24,870	90.03	20.72	30 hr (mammalian reticulocytes, in vitro), >20 hr (yeast, in vivo), >10 hr (<i>E. coli</i> , in vivo).	6.98	31	31	40,499.62

selected bacteria were from *Bacillus* and *Paenibacillus* genera.

Endophytic *Bacillus* species have been shown to benefit plants through a variety of mechanisms, including promoting plant growth, protecting the plant from phytopathogenic bacteria, fungi, insects, and nematodes, and inducing systematic resistance. *B. amyloliquefaciens* is one of the most studied endophytic *Bacillus* species. Various bioactive compounds and novel molecules are predicted to be produced by *Bacillus* spp. strains.⁴¹ The production of L-asparaginase by *B. amyloliquefaciens* has been mentioned previously.⁴² *Ophiopogon japonicas* is a medicinal plant common in South-East Asia with anticancer activity⁴³ and isolation of *B. amyloliquefaciens* from this herb has been reported.⁴⁴ It is assumed that some of the anticancerous attributes of medicinal plants are derived from endosymbiont bacteria in the plant.⁴⁵ *Paenibacillus* spp. are predominant endophytic bacteria isolated from woody plants.⁴⁶ *Paenibacillus polymyxa* (previously known as *Bacillus polymyxa*) is a plant growth-promoting rhizobacteria (PGPR) with wide ranges of plant hosts.⁴⁷ Screening, production, and optimization of L-asparaginase from this bacterium have been studied.⁴⁸ Endophytic *P. polymyxa* has been isolated from medicinal plant *Lonicera japonica*.⁴⁹ This plant has been used in herbal medicine due to its anticancer effects on the respiratory system.⁵⁰ Moreover, isolation of this bacterium from Ginseng (*Panax ginseng* CA Meyer), a well-known anticancer medicinal plant, has been reported.⁵¹ Three-gram negative selected bacteria were from *Azorhizobium*, *Bradyrhizobium*, and *Herbaspirillum* genera (Table 1). *Azorhizobium caulinodans* is a diazotrophic endophyte usually associated with cereal crops.⁵² Also, isolation of this bacterium from *Allium ursinum* plant has been reported.⁵³ *Allium* herbs have beneficial antimicrobial and antitumor effects.⁵⁴ *Bradyrhizobium japonicum* is an endophytic bacteria that usually has been isolated from soybean (*Glycine ma* L. Merr.) varieties.⁵⁵ Strong anticancer properties have been discovered in soybeans.⁵⁶ L-asparaginase production by *Bradyrhizobium* sp. (*Lupinus*) has been shown,⁵⁷ but there is no report showing that *B. japonicum*. *Herbaspirillum rubrisubalbicans* is an endophytic diazotroph of sugarcane (*Saccharum* spp.) and *Herbaspirillum seropedicae* colonizes in different species of the *Gramineae*

family.⁵⁸ *Herbaspirillum* spp. has been isolated from tea (*Camellia sinensis*) plant.⁵⁹ The tea plant is considered one of the medicinal plants with anticarcinogenic properties.⁶⁰ There is no report about L-asparaginase production by these species.

3.2 | Physicochemical analyses

ProtParam parameters shown in Table 2 reveal the physicochemical properties of the selected protein. The values of aliphatic index are high among studied proteins and they range from 83.72 to 101.55. All studied proteins are predicted to have high extinction coefficient (16,390 to 29,450). Interestingly, according to prediction results, *E. chrysanthemi* asparaginase possesses the shortest half-life in mammalian reticulocytes and in vitro among studied proteins. ProtParam classifies all studied proteins as stable (Instability index < 40).

3.3 | Complementary analyses for hypothetical proteins

As can be seen from Table 1, among nine candidate proteins, two proteins, that is, ERK81845.1 and BAL10002.1, hereafter called HP1 and HP2, were hypothetical. To ensure their functional annotation and their ability to be asparaginase, the following analyses were performed. There is heterogeneity in the codon adaptation among different genes of various species. The rate of this index varies from 0 to 1 and is related to the amount of gene expression high probability of the gene expression is indicated by an index greater than 0.5. The index values were 0.768 and 0.902 for HP1 and HP2, respectively. Therefore, both of them had a high probability of expression. Both subcellular localization prediction sites viz., CELLO and PSLpred, identified HP1 as an extracellular and HP2 as a periplasmic protein.

Confidently predicted domains, repeats, motifs, and features of HPs are given in Table 3. Asparaginase domain is predicted to present in both of them with acceptable *E*-value. Moreover, HP2 is predicted to

TABLE 3 The results of identification and annotation of HPs domains by simple modular architecture research tool

Protein name	Identifiers	Confidently predicted domains, repeats, motifs, and features	<i>E</i> -value	Start	End
HP1		Asparaginase	1.21e-86	2	259
HP2	G7D7T8_BRAJP, G7D7T8	Signal peptide Asparaginase	N/A 9.07e-123	1 35	21 353

Abbreviation: HP, hypothetical protein.

TABLE 4 The summarized results of used softwares for prediction of immunological properties

Prediction softwares and parameters proteins	Vaxijen	Algpred	Antigenic peptide prediction (number of antigenic determinants)	ABCpred (number of epitopes with score of >0.8)	Ellipro (number of discontinuous epitopes)	ProPredI (number of epitopes which bind to highest number of alleles; highest number of alleles which a epitope can bind)	ProPred (number of epitopes) number of epitopes which bind to highest number of alleles; highest number of alleles which a epitope can bind)
<i>E. coli</i> Asparaginase	Antigen	Potential allergen	17	13	5	(85) 4;6	(63) 3;51
<i>E. chrysanthemi</i> asparaginase	Antigen	Allergen	15	15	9	(89) 2;6	(88) 2;51
BAF90379.1	Antigen	Nonallergen	17	16	6	(90) 1;7	(87) 5;51
ERK81845.1 hypothetical	Antigen	Potential allergen	8	9	6	(84) 1;7	(62) 1;51
KNX35781.1	Antigen	Potential allergen	11	15	4	(95) 3;5	(88) 4;51
AHY50093.1	Antigen	Allergen	17	15	6	(98) 1;9	(80) 3;51
BAL10002.1 hypothetical	Antigen	Allergen	18	15	6	(97) 1;9	(80) 4;51
ALU91046.1	Nonantigen	Potential allergen	11	9	8	(87) 1;7	(70) 1;51
ALU91730.1	Nonantigen	Nonallergen	21	11	7	(101) 1;6	(88) 1;51
ADJ65404.1	Antigen	Potential allergen	11	9	9	(88) 1;7	(64) 2;51
KOS01895.1	Antigen	Potential allergen	17	11	7	(88) 1;7	(98) 4;51

possess a signal peptide sequence. This enhances the possibility of extracellular prediction of the HP by subcellular localization prediction sites. When a protein is not predicted to contain a signal peptide, it may secrete nonclassically if it obtains a SecP score exceeding the threshold of 0.5. Predicted SecP score of HP1 was 0.844078 which is in line with the previous result of periplasmic localization of this protein. There is a pool of enzymes in the periplasm compartment that carry on the verity of functions such as protein folding,

degrading and oxidation. These residing enzymes are exported across the inner membrane mainly by classical secretory pathways of Sec and twin-arginine translocation pathways. However, some proteins lack signal peptides and are translocated across the inner membrane by the nonclassical secretion pathway.⁶¹ According to the TMHMM prediction, both HPs did not have transmembrane helices, which confirms the prediction results of subcellular localization and SecretomeP 2.0.

TABLE 5 IC₅₀ values of epitopic regions within the protein sequences which bind to highest number of MHCI and MHCII alleles

Protein	Epitope sequence which bind to highest number of MHCI alleles	IC ₅₀ value	Epitope sequence which bind to highest number of MHCII alleles	IC ₅₀ value
<i>E. coli</i> Asparaginase	FDVSKLNEL	724.44	VVNIGSQDM	1,606.94
	YNYANASDL	638.26	VVMVGAMRP	57.28
	ARVLLQLAL	1,233.1	MRPSTMSA	26.36
	DPQQIQQIF	Nonbinder		
<i>E. chrysanthemi</i> asparaginase	KLSQRVNEL	4,355.12	IVILATGGT	195.88
	LPGLVSDSL	Nonbinder	MRPATAISA	4.95
BAF90379.1	RPPEGRHL	Nonbinder	VARITTLQA	95.94
			VVMVGAMRP	57.28
			VVLNDTIHG	1,563.15
			FRSLNTGPA	9.38
			LQLLLANGH	6.89
ERK81845.1 hypothetical	NPQKARILL	Nonbinder	VVVGSMRP	158.12
KNX35781.1	KQFMPFITF	3,326.6	VVNIGSQNM	1,023.29
	KLAKRVNKL	2,710.19	VVVGSMRP	158.12
	ARILLMLAL	425.60	VVLNDRIAS	138.04
			ILLMLALTK	7.00
AHY50093.1	YPLPAGEQL	182.81	VVIVGSMRP	340.41
			MRPATAVSA	3.59
			IVLAGVGDG	229.61
BAL10002.1 hypothetical	YPLPAGEQL	182.81	VVIVGSMRP	340.41
			MRPATAVSA	3.59
			IVLAGVGDG	229.61
			LLIAGGMTA	594.29
ALU91046.1	KLAKRVNTL	1,674.94	VVIVGAMRP	78.89
ALU91730.1	LPRVEIIPM	4,655.86	IRKAPVPAL	16.63
ADJ65404.1	KLGKRVNTL	1,725.84	VVIVGAMRP	78.89
			FVVSDDLNA	1.12
KOS01895.1	IPLLASFIL	Nonbinder	MTMIPLLAS	15.00
			FILGSTLWS	1.99
			YFLNLVIKS	1.93
			MRPATAISA	4.95

Abbreviations: MHCI, major histocompatibility complex class I; MHCII, major histocompatibility complex class II.

3.4 | Tertiary structure prediction and validation

Three-dimensional model structures of protein sequences were generated by (PS)2-v3 and structures quality was validated by Ramachandran plot and ProSA (Table S2). ProSA-web Z-score plot can be used to check whether the Z-score of the input structure is within the range of the source typically found for the native proteins of a similar size. The Z-score values of all protein structures checked in this study were highlighted as a black dot which indicates being in the range of native conformations.³¹ The Ramachandran plot is a plot of the torsional angles of amino acids contained in a peptide. The Ramachandran plot analysis of the generated models revealed that 90% of all residues proteins were located in the favored and allowed regions, indicating that the predicted models had good qualities.⁶² Generated 3D structures were used to predict conformational discontinuous B-cell epitopes.

3.5 | Immunological properties prediction

The results of the software used for the prediction of immunological properties are summarized in Table 4. As can be seen, upon Algpred prediction, two of the investigated endophytic asparaginase sequences (BAF90379.1, ALU91730.1) were nonallergen, whereas *E. coli* and *E. chrysanthemi* asparaginases, which are used as a drug, were predicted to be potential allergen and allergen, respectively. A spectrum of immunologically mediated hypersensitivity reactions with varying mechanisms and clinical presentations is called allergy. Drug allergy is an unpredictable reaction that leads to a decline in the quality of patient life, probably delayed treatment, and even death. Given the myriad of symptoms associated with the condition, diagnosis is often challenging.⁶³ Therefore, allergenicity prediction of drug candidates is an important subject.⁶⁴ The administration of therapeutic protein products to patients may be associated with eliciting anti-drug antibodies.⁶⁵ There are many experimental and protein engineering methods for assessment of the immunological properties of biological drugs and reducing the immunogenicity of biotherapeutics.^{66,67} However, there are some in silico methods which play significant roles in predicting unwanted immune responses.⁶⁸ According to vaxijen server prediction (Table 4), all of the studied proteins predicted to be protective antigen except for endophytic sequences of ALU91046.1, and ALU91730.1. Referring to predictive results of antigenic regions (Table S3), the number and sequence of B-cell (Table S4), T-cell epitopes (Tables S5 and S6), IC50 values of epitopic regions within the protein sequences which bind to highest number of MHCI and MHCII alleles (Table 5), and

summarized results of used software for prediction of immunological properties (Table 4), the following information was obtained. Four proteins with the accession numbers of ERK81845.1, KNX35781.1, ALU91046.1, and ADJ65404.1 possessed less antigenic regions than *E. coli* and *E. chrysanthemi* asparaginases. B-cell epitope prediction results were indicative of four proteins with the accession numbers of ERK81845.1, ALU91046.1, ALU91730.1, and ADJ65404.1 that had less high-scored epitopic region than sequences currently used as a medication. All studied asparaginase sequences from endophytic bacteria are predicted to own conformational B-cell epitopes equal or less than *E. chrysanthemi* asparaginase and sequences of KNX35781.1 showed to have the least conformational epitope number of 4. ERK81845.1 predicted to have the least number of MHCI (84) and MHCII (62) binding epitopes among all studied sequences. Moreover, predicted IC50 values (Table 5) showed a weak binding affinity for two epitopic regions within this sequence which bound to the highest number of MHCI and MHCII alleles. The ability of epitopes to bind to multiple alleles is considered an advantage in the development of an epitope-based vaccines⁶⁹ whereas it is a negative factor in drug studies. Table 5 also indicates that just two ERK81845.1 and KOS01895.1 proteins have binding epitopes stronger than *E. coli* and *E. chrysanthemi* asparaginases among epitopes which bind to the highest number of MHCII alleles.

4 | CONCLUSION

L-asparaginase is an enzyme catalyzing the conversion of L-asparagine amino acid to L-aspartic acid and ammonia. This amino acid is necessary for the function of most neoplastic cells, and it is introduced as a chemotherapy drug. Stimulation of immune response of currently used asparaginases in some patients has motivated scientists to find new sequences with asparaginase activity but fewer side effects. Given that endophytes have been proven to synthesize pharmaceuticals useful to health, this study identified and immunologically characterized 9 endophyte originated asparaginase sequences selected from 262 endophytic bacteria by in silico approaches. The obtained immunological analyses results with prediction software indicated that some endophytic bacteria such as *B. amyloliquefaciens* and *H. rubrisubalbicans* could produce enzymes with suitable immunological properties and therapeutic use of these enzymes is possible. However, advanced researches using specialized databases, more innovative software and experimental investigations are required to characterize the predicted sequences more precisely in the future.

CONFLICT OF INTEREST

The authors declared no conflict of interests.

ORCID

Sayed Mohammad Gheibi Hayat  <https://orcid.org/0000-0002-1378-118X>

REFERENCES

- Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin*. 2018;68:394–424.
- Nurgali K, Jagoe RT, Abalo R. Adverse effects of cancer chemotherapy: Anything new to improve tolerance and reduce sequelae? *Front Pharmacol*. 2018;9:245.
- Fu Y-M, Meadows GG. Specific amino acid dependency regulates the cellular behavior of melanoma. *J Nutr*. 2007;137:1591S–1596S.
- Ueno T, Ohtawa K, Mitsui K, et al. Cell cycle arrest and apoptosis of leukemia cells induced by L-asparaginase. *Leukemia*. 1997;11:1858–1861.
- Dinndorf PA, Gootenberg J, Cohen MH, Keegan P, Pazdur R. FDA drug approval summary: Pegaspargase (Oncaspar®) for the first-line treatment of children with acute lymphoblastic leukemia (ALL). *Oncologist*. 2007;12:991–998.
- Muss HB, Spell N, Scudieri D, et al. A phase II trial of PEG-L-asparaginase in the treatment of non-Hodgkins lymphoma. *Invest New Drugs*. 1990;8:125–130.
- Sharma M, Sharma NN, Bhalla TC. Amidases: Versatile enzymes in nature. *Rev Environ Sci Biotechnol*. 2009;8:343–366.
- Batool T, Makky EA, Jalal M, Yusoff MM. A comprehensive review on L-asparaginase and its applications. *Appl Biochem Biotechnol*. 2016;178:900–923.
- Zeidan A, Wang ES, Wetzler M. Pegasparginase: Where do we stand? *Expert Opin Biol Ther*. 2009;9:111–119.
- Duval M, Suci S, Ferster A, et al. Comparison of *Escherichia coli*—Asparaginase with *Erwinia*-asparaginase in the treatment of childhood lymphoid malignancies: Results of a randomized European Organisation for Research and Treatment of Cancer—Children's leukemia group phase 3 trial. *Blood*. 2002;99:2734–2739.
- Shrivastava A, Khan AA, Khurshid M, Kalam MA, Jain SK, Singhal PK. Recent developments in l-asparaginase discovery and its potential as anticancer agent. *Crit Rev Oncol Hematol*. 2016;100:1–10.
- Krishnapura PR, Belur PD, Subramanya S. A critical review on properties and applications of microbial l-asparaginases. *Crit Rev Microbiol*. 2016;42:720–737.
- Gouda S, Das G, Sen SK, Shin H-S, Patra JK. Endophytes: A treasure house of bioactive compounds of medicinal importance. *Front Microbiol*. 2016;7:1538.
- Zhang HW, Song YC, Tan RX. Biology and chemistry of endophytes. *Nat Prod Rep*. 2006;23:753–771.
- Huang W, Cai Y, Hyde K, Corke H, Sun M. Biodiversity of endophytic fungi associated with 29 traditional Chinese medicinal plants. *Fungal Diversity*. 2008;33:61–75.
- Stierle A, Strobel G, Stierle D. Taxol and taxane production by *Taxomyces andreanae*, an endophytic fungus of Pacific yew. *Science*. 1993;260:214–216.
- Singh M, Kumar A, Singh R, Pandey KD. Endophytic bacteria: A new source of bioactive compounds. *3 Biotech*. 2017;7:315.
- Madden T. The NCBI handbook [Internet]. 2nd ed. USA: National Center for Biotechnology Information, 2013.
- Rosenblueth M, Martínez-Romero E. Bacterial endophytes and their interactions with hosts. *Mol Plant Microbe Interact*. 2006;19:827–837.
- Boyle CJC, Schulz BJE, Sieber TN. Microbial root endophytes. Heidelberg, Germany: Springer, 2006.
- Barrett T, Clark K, Gevorgyan R, et al. BioProject and BioSample databases at NCBI: Facilitating capture and organization of metadata. *Nucleic Acids Res*. 2011;40:D57–D63.
- Gasteiger E, Hoogland C, Gattiker A, Wilkins MR, Appel RD, Bairoch A. The proteomics protocols handbook. Totowa, New Jersey: Humana Press, 2005;p. 571–607.
- Puigbò P, Bravo IG, Garcia-Valle S. CAIcal: A combined set of tools to assess codon usage adaptation. *Biol Direct*. 2008;3:38.
- Yu CS, Lin CJ, Hwang JK. Predicting subcellular localization of proteins for gram-negative bacteria by support vector machines based on n-peptide compositions. *Protein Sci*. 2004;13:1402–1406.
- Bhasin M, Garg A, Raghava G. PSLpred: Prediction of subcellular localization of bacterial proteins. *Bioinformatics*. 2005;21:2522–2524.
- Bendtsen JD, Kierner L, Fausbøll A, Brunak S. Non-classical protein secretion in bacteria. *BMC Microbiol*. 2005;5:58.
- Chen Y, Yu P, Luo J, Jiang Y. Secreted protein prediction system combining CJ-SPHMM, TMHMM, and PSORT. *Mamm Genome*. 2003;14:859–865.
- Letunic I, Bork P. 20 years of the SMART protein domain annotation resource. *Nucleic Acids Res*. 2017;46:D493–D496.
- Huang T-T, Hwang J-K, Chen C-H, Chu C-S, Lee C-W, Chen C-C. 2: Protein structure prediction server version 3.0. *Nucleic acids research*. 2015;43:W338–W342.
- Lovell SC, Davis IW, Arendall WB III, et al. Structure validation by C α geometry: ϕ , ψ and C β deviation. *Proteins Struct Funct Bioinf*. 2003;50:437–450.
- Wiederstein M, Sippl MJ. ProSA-web: Interactive web service for the recognition of errors in three-dimensional structures of proteins. *Nucleic Acids Res*. 2007;35:W407–W410.
- Saha S, Raghava G. AlgPred: Prediction of allergenic proteins and mapping of IgE epitopes. *Nucleic Acids Res*. 2006;34:W202–W209.
- Kolaskar A, Tongaonkar PC. A semi-empirical method for prediction of antigenic determinants on protein antigens. *FEBS Lett*. 1990;276:172–174.
- Saha S, Raghava G. Prediction of continuous B-cell epitopes in an antigen using recurrent neural network. *Proteins Struct Funct Bioinf*. 2006;65:40–48.
- Ponomarenko J, Bui H-H, Li W, et al. ElliPro: A new structure-based tool for the prediction of antibody epitopes. *BMC Bioinf*. 2008;9:514.
- Singh H, Raghava G. ProPred1: Prediction of promiscuous MHC class-I binding sites. *Bioinformatics*. 2003;19:1009–1014.
- Singh H, Raghava G. ProPred: Prediction of HLA-DR binding sites. *Bioinformatics*. 2001;17:1236–1237.
- Guan P, Doytchinova IA, Zygouri C, Flower DR. MHCpred: A server for quantitative prediction of peptide-MHC binding. *Nucleic Acids Res*. 2003;31:3621–3624.
- Du Q-S, Wei Y-T, Pang Z-W, Chou K-C, Huang R-B. Predicting the affinity of epitope-peptides with class I MHC molecule

- HLA-A* 0201: An application of amino acid-based peptide prediction. *Protein Eng Des Sel*. 2007;20:417–423.
40. Wang P, Sidney J, Dow C, Mothe B, Sette A, Peters B. A systematic assessment of MHC class II peptide binding predictions and evaluation of a consensus approach. *PLoS Comput Biol*. 2008;4:e1000048.
 41. Lopes R, Tsui S, Gonçalves PJ, de Queiroz MV. A look into a multifunctional toolbox: Endophytic bacillus species provide broad and underexploited benefits for plants. *World J Microbiol Biotechnol*. 2018;34:94.
 42. Abdelrazek NA, Elkhatab WF, Raafat MM, Aboulwafa MM. Experimental and bioinformatics study for production of L-asparaginase from bacillus licheniformis: A promising enzyme for medical application. *AMB Express*. 2019;9:39.
 43. Chen J, Yuan J, Zhou L, et al. Regulation of different components from *Ophiopogon japonicus* on autophagy in human lung adenocarcinoma A549Cells through PI3K/Akt/mTOR signaling pathway. *Biomed Pharmacother*. 2017;87:118–126.
 44. Chen Y-T, Yuan Q, Shan L-T, Lin M-A, Cheng D-Q, Li C-Y. Antitumor activity of bacterial exopolysaccharides from the endophyte *Bacillus amyloliquefaciens* sp. isolated from *Ophiopogon japonicus*. *Oncol Lett*. 2013;5:1787–1792.
 45. Chow Y, Ting AS. Endophytic L-asparaginase-producing fungi from plants associated with anticancer properties. *J Adv Res*. 2015;6:869–876.
 46. Ulrich K, Stauber T, Ewald D. Paenibacillus—A predominant endophytic bacterium colonising tissue cultures of woody plants. *Plant Cell Tiss Org Cult*. 2008;93:347–351.
 47. Timmusk S, Grantcharova N, Wagner EGH. *Paenibacillus polymyxa* invades plant roots and forms biofilms. *Appl Environ Microbiol*. 2005;71:7292–7300.
 48. Wakil SS, Adelegan AA. Screening, production and optimization of L-asparaginase from soil bacteria isolated in Ibadan, South-Western Nigeria. *J Basic Appl Sci*. 2015;11:39–51.
 49. Zhao L, Xu Y, Lai X-H, Shan C, Deng Z, Ji Y. Screening and characterization of endophytic bacillus and Paenibacillus strains from medicinal plant *Lonicera japonica* for use as potential plant growth promoters. *Braz J Microbiol*. 2015;46:977–989.
 50. Park KI, Park H, Nagappan A, et al. Polyphenolic compounds from Korean *Lonicera japonica* Thunb. Induces apoptosis via AKT and caspase cascade activation in A549 cells. *Oncol Lett*. 2017;13:2521–2530.
 51. Gao Y, Liu Q, Zang P, et al. An endophytic bacterium isolated from *Panax ginseng* CA Meyer enhances growth, reduces morbidity, and stimulates ginsenoside biosynthesis. *Phytochem Lett*. 2015;11:132–138.
 52. Liu H, Wang X, Qi H, et al. The infection and impact of *Azorhizobium caulinodans* ORS571 on wheat (*Triticum aestivum* L.). *PLoS One*. 2017;12:e0187947.
 53. Beiranvand M, Amin M, Hashemi-Shahraki A, Romani B, Yaghoubi S, Sadeghi P. Antimicrobial activity of endophytic bacterial populations isolated from medical plants of Iran. *Iran J Microbiol*. 2017;9:11.
 54. Ivanova A, Mikhova B, Najdenski H, Tsvetkova I, Kostova I. Chemical composition and antimicrobial activity of wild garlic *Allium ursinum* of Bulgarian origin. *Nat Prod Commun*. 2009;4:1059–1062.
 55. Egamberdieva D, Jabborova D, Wirth SJ, Alam P, Alyemeni MN, Ahmad P. Interactive effects of nutrients and Bradyrhizobium japonicum on the growth and root architecture of soybean (*Glycine max* L.). *Front Microbiol*. 2018;9:1000.
 56. Jayachandran M, Xu B. An insight into the health benefits of fermented soy products. *Food Chem*. 2019;271:362–371.
 57. Kretovich V, Sidel'nikova L, Ivanushkin A, Karyakina T. Localization of aspartase, asparaginase, and glutaminase in intact bacteroids of Rhizobium lupini. *Applied Biochem Microbiol*. 1985;20(4):362–364.
 58. Baldani J, Pot B, Kirchhof G, et al. Emended description of *Herbaspirillum*; incursion of [*pseudomonas*] *rubrisubalbicans*, a mild plant pathogen, as *Herbaspirillum rubrisubalbicans* comb. Nov.; and classification of a Group of Clinical Isolates (EF Group 1) as *Herbaspirillum* species 3. *Int J Syst Evol Microbiol*. 1996;46:802–810.
 59. Yan X, Wang Z, Wang L, et al. Isolation, diversity, and growth-promoting activities of endophytic bacteria from tea cultivars of Zijuan and Yunkang-10. *Front Microbiol*. 2018;9:1848.
 60. Wheeler DS, Wheeler WJ. The medicinal chemistry of tea. *Drug Dev Res*. 2004;61:45–65.
 61. Wang G, Chen H, Xia Y, et al. How are the non-classically secreted bacterial proteins released into the extracellular milieu? *Curr Microbiol*. 2013;67:688–695.
 62. Laskowski R, de Beer T. (2004). Ramachandran Plot. *Dictionary of Bioinformatics and Computational Biology*.
 63. Warrington R, Silviu-Dan F, Wong T. Drug allergy. *Allergy Asthma Clin Immunol*. 2018;14:1–11.
 64. Blumenthal KG. The role of the clinical history in drug allergy prediction. *J Allergy Clin Immunol Pract*. 2018;6:149–150.
 65. Pratt K. Anti-drug antibodies: Emerging approaches to predict, reduce or reverse biotherapeutic immunogenicity. *Antibodies*. 2018;7:19.
 66. Onda M. Reducing the immunogenicity of protein therapeutics. *Curr Drug Targets*. 2009;10:131–139.
 67. Rosenberg AS, Sauna ZE. Immunogenicity assessment during the development of protein therapeutics. *J Pharm Pharmacol*. 2018;70:584–594.
 68. Bryson CJ, Jones TD, Baker MP. Prediction of immunogenicity of therapeutic proteins. *BioDrugs*. 2010;24:1–8.
 69. Bui H-H, Sidney J, Dinh K, Southwood S, Newman MJ, Sette A. Predicting population coverage of T-cell epitope-based diagnostics and vaccines. *BMC Bioinf*. 2006;7:153.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of this article.

How to cite this article: Zadeh Hosseingholi E, Neghabi N, Molavi G, Gheibi Hayat SM, Shahriarpour H. In silico identification and characterization of antineoplastic asparaginase enzyme from endophytic bacteria. *IUBMB Life*. 2020;72:991–1000. <https://doi.org/10.1002/iub.2237>

SUPPORTING INFORMATION TABLE S1. Criteria for risk group classification in protocol ALL-SEHOP-PETHEMA 2013

STANDARD RISK: The patient must meet each and every one of the following criteria:

- Age ≥ 1 and < 10 years.
- Leukocytes $< 20 \times 10^9/l$ at diagnosis.
- Non-T immunophenotype.
- Absence of CNS or testicular infiltration
- Good risk cytogenetics
 - o High hyperdiploidy (51-67 chromosomes), DNA index 1.10-1.44 (always confirmed by other cytogenetic techniques). or
 - o t(12;21) / ETV6-RUNX1 rearrangement
- Absence of t(1;19) / TCF3/PBX1 rearrangement.
- No MLL rearrangement.
- Good prednisone response ($< 1,000$ blasts/mm³ in peripheral blood on day +8 of Induction IA)
- Good early response ($< 5\%$ de blasts and $< 0.1\%$ MRD in bone marrow on day +15 of Induction and at the end of induction IA)

HIGH RISK: The existence of any of the following criteria determines the inclusion of the patient in this High-risk group:

- t(4;11) (MLL/AF4).
- Hypodiploidy < 44 chromosomes or DNA index < 0.81 (confirmation by other techniques is required).
- Poor prednisone response ($> 1,000$ blasts in peripheral blood on day +8 of Induction,)
- M2 ($> 25\%$ blasts) and $> 10\%$ MRD on day +15 of Induction, in bone marrow.
- MRD $> 1\%$ on day +33 of Induction, in bone marrow.
- MRD $> 0.1\%$ before consolidation, in bone marrow.
- Patients with BCR-ABL1 rearrangement (Ph+ ALL).

INTERMEDIATE RISK: Those patients who do not meet the criteria for Standard Risk or High Risk.

SUPPORTING INFORMATION TABLE S2. Treatment protocol for standard-risk patients in ALL-SEHOP-PETHEMA 2013

Treatment phase/drug	Single or daily dose	Days of application per phase ^a
<u>Induction</u>		
<i>IA</i>		
Prednisone/Prednisolone (PO/IV)	60 mg/m ² /d	1-28 ^c
Vincristine (IV)	1.5 mg/m ² /dose (max. 2 mg)	8, 15, 22, 29
Daunorubicin (PI over 1 h)	30 mg/m ² /dose	8, 15
L-Asparaginase (IM)	10 000 IU/m ² /dose	12, 15, 18, 21, 24, 27, 30, 33
Methotrexate/ARA-C/Prednisone (IT)	12/30/20 mg/dose ^b	1,12, 19, 26, 33 ^{d,e}
<i>IB</i>		
Cyclophosphamide (PI over 1 h)	1000 mg/m ² /dose	36, 64
Cytarabine (PI over 1 h)	75 mg/m ² /dose	38-41, 45-48, 52-55, 59-62
6-Mercaptopurine (PO)	60 mg/m ² /d	36-63
Methotrexate/ARA-C/Prednisone (IT)	12/30/20 mg/dose ^b	45, 59
<u>Consolidation</u>		
Methotrexate (PI over 24 h) ^f	5000 mg/m ² /dose	8, 22, 36, 50
6-Mercaptopurine (PO)	25 mg/m ² /d	1-56
Methotrexate/ARA-C/Prednisone (IT)	12/30/20 mg/dose ^b	8, 22, 36, 50
<u>Reinduction</u>		
Dexamethasone (PO/IV)	10 mg/m ² /d	1-21 ^c
Vincristine (IV)	1.5 mg/m ² /dose (max. 2 mg)	8, 15, 22, 29
Doxorubicin (PI over 1 h)	30 mg/m ² /dose	8, 15, 22, 29
L-Asparaginase (IM)	10 000 IU/m ² /dose	8, 11, 15, 18
Cyclophosphamide (PI over 1 h)	1000 mg/m ² /dose	36
Cytarabine (IV)	75 mg/m ² /dose	38-41, 45-48
6-Thioguanine (PO)	60 mg/m ² /d	36-49
Methotrexate/ARA-C/Prednisone (IT) ^g	12/30/20 mg/dose ^b	1, 17, 38, 45
<u>Maintenance</u>^h		
Methotrexate (PO)	20 mg/m ² /dose ⁱ	once a week
6-Mercaptopurine (PO)	50 mg/m ² /d ⁱ	daily
Methotrexate/ARA-C/Prednisone (IT)	4x 12/30/20 mg/dose ^b	monthly

^aAdjustments of schedule were allowed if clinical condition and bone marrow recovery were inadequate.

^bDoses of IT drugs were adjusted for children <3 years of age.

^cSteroids were tapered over 9 additional days.

^{***}Differences of induction treatment given after amendment on November 2016:

- Asparaginase given in induction is PEG-asparaginase IM at dose 1000 IU/m² on days 12 and 26.

^dIT therapy on days 19 and 26 was only administered to patients with CNS-2 at diagnosis or T-ALL and > 100 000 leukocytes/mm³ at diagnosis.

^ePatients with CNS status CNS 3 received 5 weekly doses of IT therapy on days 1, 8, 15, 22, and 29 in Induction Protocol A.

^fA loading dose of 10% was infused over 30 min, the remaining 90% over 23.5 h. Leucovorin rescue was given at hours 42, 48, and 54 (each 15 mg/m²). Doses of leucovorin rescue were adjusted, if MTX levels were > 1.0 µmol/l at hour 42 or later. If the MTX level at hour 54 was > 0.25 µmol/l, the rescue was continued at six-hour intervals until MTX levels were ≤ 0.25 µmol/l.

^{***}Differences of reinduction treatment given after amendment on November 2016:

- Asparaginase given in reinduction is PEG-asparaginase IM at dose 1000 IU/m² on day 1.
- Dexamethasone given from day 1 to 14.

^gIT therapy on days 1 and 17 in reinduction was only administered to patients with CNS-2 or CNS-3 at diagnosis or T-ALL with > 100 000 leukocytes/mm³ at diagnosis.

^hDoses were adjusted to WBC (target range 2.0-3.0 ×10⁹/L).

ⁱMaintenance was given from the end of intensive chemotherapy until 104 weeks after diagnosis.

PO indicates orally; IV, intravenous push; PI, intravenous infusion; IT, intrathecal; IM intramuscular

SUPPORTING INFORMATION TABLE S3. Treatment protocol for intermediate-risk patients in ALL-SEHOP-PETHEMA 2013.

Treatment phase/drug	Single or daily dose	Days of application per phase ^a
<u>Induction</u>		
<i>IA</i>		
Prednisone/Prednisolone (PO/IV)	60 mg/m ² /d	1-28 ^c
Vincristine (IV)	1.5 mg/m ² /dose (max. 2 mg)	8, 15, 22, 29
Daunorubicin (PI over 1 h)	30 mg/m ² /dose	8, 15, 22, 29
L-Asparaginase (IM)	10 000 IU/m ² /dose	12, 15, 18, 21, 24, 27, 30, 33
Methotrexate/ARA-C/Prednisone (IT)	12/30/20 mg/dose ^b	1,12, 19, 26, 33 ^{d,e}
<i>IB</i>		
Cyclophosphamide (PI over 1 h)	1000 mg/m ² /dose	36, 64
Cytarabine (PI over 1 h)	75 mg/m ² /dose	38-41, 45-48, 52-55, 59-62
6-Mercaptopurine (PO)	60 mg/m ² /d	36-63
Methotrexate/ARA-C/Prednisone (IT)	12/30/20 mg/dose ^b	45, 59
<u>Consolidation</u>		
Methotrexate (PI over 24 h) ^f	5000 mg/m ² /dose	8, 22, 36, 50
6-Mercaptopurine (PO)	25 mg/m ² /d	1-56
Methotrexate/ARA-C/Prednisone (IT)	12/30/20 mg/dose ^b	8, 22, 36, 50
<u>Reinduction</u>		
Dexamethasone (PO/IV)	10 mg/m ² /d	1-21 ^c
Vincristine (IV)	1.5 mg/m ² /dose (max. 2 mg)	8, 15, 22, 29
Doxorubicin (PI over 1 h)	30 mg/m ² /dose	8, 15, 22, 29
L-Asparaginase (IM)	10 000 IU/m ² /dose	8, 11, 15, 18
Cyclophosphamide (PI over 1 h)	1000 mg/m ² /dose	36
Cytarabine (IV)	75 mg/m ² /dose	38-41, 45-48
6-Thioguanine (PO)	60 mg/m ² /d	36-49
Methotrexate/ARA-C/Prednisone (IT) ^g	12/30/20 mg/dose ^b	1, 17, 38, 45
<u>Maintenance</u>^h		
Methotrexate (PO)	20 mg/m ² /dose ⁱ	once a week
6-Mercaptopurine (PO)	50 mg/m ² /d ⁱ	daily
PEG-asparaginase (IM)	1000 IU/m ²	biweekly
Methotrexate/ARA-C/Prednisone (IT)	6x 12/30/20 mg/dose ^b	monthly

^aAdjustments of schedule were allowed if clinical condition and bone marrow recovery were inadequate.

^bDoses of IT drugs were adjusted for children <3 years of age.

^cSteroids were tapered over 9 additional days.

^{***}Differences of induction treatment given after amendment on November 2016:

- Asparaginase given in induction is PEG-asparaginase IM at dose 1000 IU/m² on days 12 and 26.

^dIT therapy on days 19 and 26 was only administered to patients with CNS-2 at diagnosis or T-ALL and > 100 000 leukocytes/mm³ at diagnosis.

^ePatients with CNS status CNS 3 received 5 weekly doses of IT therapy on days 1, 8, 15, 22, and 29 in Induction Protocol A.

^fA loading dose of 10% was infused over 30 min, the remaining 90% over 23.5 h. Leucovorin rescue was given at hours 42, 48, and 54 (each 15 mg/m²). Doses of leucovorin rescue were adjusted, if MTX levels were > 1.0 µmol/l at hour 42 or later. If the MTX level at hour 54 was > 0.25 µmol/l, the rescue was continued at six-hour intervals until MTX levels were ≤ 0.25 µmol/l.

^{***}Differences of reinduction treatment given after amendment on November 2016:

- Asparaginase given in reinduction is PEG-asparaginase IM at dose 1000 IU/m² on day 1.
- Dexamethasone given from day 1 to 14.

^gIT therapy on days 1 and 17 in reinduction was only administered to patients with CNS-2 or CNS-3 at diagnosis or T-ALL with > 100 000 leukocytes/mm³ at diagnosis.

^hDoses were adjusted to WBC (target range 2.0-3.0 x10⁹/L).

ⁱMaintenance was given from the end of intensive chemotherapy until 104 weeks after diagnosis.

PO indicates orally; IV, intravenous push; PI, intravenous infusion; IT, intrathecal; IM intramuscular.

SUPPORTING INFORMATION TABLE S4. Treatment protocol for high-risk patients in ALL-SEHOP-PETHEMA 2013.

Treatment phase/drug	Single or daily dose	Days of application per phase ^a
<u>Induction</u>		
<i>IA</i>		
Prednisone/Prednisolone (PO/IV)	60 mg/m ² /d	1-28 ^c
Vincristine (IV)	1.5 mg/m ² /dose (max. 2 mg)	8, 15, 22, 29
Daunorubicin (PI over 1 h)	30 mg/m ² /dose	8, 15, 22, 29
L-Asparaginase (IM)	10 000 IU/m ² /dose	12, 15, 18, 21, 24, 27, 30, 33
Methotrexate/ARA-C/Prednisone (IT)	12/30/20 mg/dose ^b	1,12, 19, 26, 33 ^{d,e}
 <i>IB</i>		
Cyclophosphamide (PI over 1 h) ^f	1000 mg/m ² /dose	36, 64
Cytarabine (PI over 1 h)	75 mg/m ² /dose	38-41, 45-48, 52-55, 59-62
6-Mercaptopurine (PO)	60 mg/m ² /d	36-63
Methotrexate/ARA-C/Prednisone (IT)	12/30/20 mg/dose ^b	45, 59
 <u>Consolidation</u>		
<i>HR1'</i>		
Dexamethasone (PO/IV)	20 mg/m ² /d	1-5
Vincristine (IV)	1.5 mg/m ² /dose (max. 2 mg)	1, 6
Cytarabine (IV)	2x 2000 mg/m ² /dose	5
Methotrexate (PI over 24 h) ^g	5000 mg/m ² /dose	1
Cyclophosphamide (PI over 1 h)	5x 200 mg/m ² /dose	2-4
PEG-Asparaginase (IM)	1000 IU/m ² /dose	7
Methotrexate/ARA-C/Prednisone (IT)	12/30/20 mg/dose ^b	1
 <i>HR2'</i>		
Dexamethasone (PO/IV)	20 mg/m ² /d	1-5
Vindesine (IV)	3.0 mg/m ² /dose (max. 5 mg)	1, 6
Daunorubicin (IV)	30 mg/m ² /dose	5
Methotrexate (PI over 24 h) ^g	5000 mg/m ² /dose	1
Ifosfamide (PI over 1 h)	5x 800 mg/m ² /dose	2-4
PEG-Asparaginase (IM)	1000 IU/m ² /dose	7
Methotrexate/ARA-C/Prednisone (IT)	12/30/20 mg/dose ^b	1
 <i>HR3'</i>		

Treatment phase/drug	Single or daily dose	Days of application per phase ^a
Dexamethasone (PO/IV)	20 mg/m ² /d	1-5
Cytarabine (IV)	4x 2000 mg/m ² /dose	1-3
Etoposide (IV)	5x 100 mg/m ² /dose	3-5
PEG-Asparaginase (IM)	1000 IU/m ² /dose	6
Methotrexate/ARA-C/Prednisone (IT)	12/30/20 mg/dose ^b	5

Delayed Intensification

Reinduction 1^h

Dexamethasone (PO/IV)	10 mg/m ² /d	1-15 ^c
Vincristine (IV)	1.5 mg/m ² /dose (max. 2 mg)	1, 8
Doxorubicin (PI over 1 h)	30 mg/m ² /dose	1, 8
PEG-Asparaginase (IM)	1000 IU/m ² /dose	1,15
Cyclophosphamide (PI over 1 h)	500 mg/m ² /dose	15
Cytarabine (IV)	75 mg/m ² /dose	17-20, 24-37
6-Thioguanine (PO)	60 mg/m ² /d	15-28
Methotrexate/ARA-C/Prednisone (IT)	12/30/20 mg/dose ^b	1 ⁱ , 17, 24

Reinduction 2

Dexamethasone (PO)	10 mg/m ² /d	1-15 ^c
Vincristine (IV)	1.5 mg/m ² /dose (max. 2 mg)	1, 8
Doxorubicin (PI over 1 h)	30 mg/m ² /dose	1, 8
PEG-Asparaginase (IM)	1000 IU/m ² /dose	1,15
Cyclophosphamide (PI over 1 h)	500 mg/m ² /dose	15
Cytarabine (IV)	75 mg/m ² /dose	17-20, 24-27
6-Thioguanine (PO)	60 mg/m ² /d	15-28
Methotrexate/ARA-C/Prednisone (IT)	12/30/20 mg/dose ^b	1 ⁱ , 17, 24

Reinduction 3

Dexamethasone (PO)	10 mg/m ² /d	1-15 ^c
Vincristine (IV)	1.5 mg/m ² /dose (max. 2 mg)	1, 8
Doxorubicin (PI over 1 h)	30 mg/m ² /dose	1, 8
PEG-Asparaginase (IM)	1000 IU/m ² /dose	1,15
Cyclophosphamide (PI over 1 h)	500 mg/m ² /dose	15
Cytarabine (IV)	75 mg/m ² /dose	17-20, 24-27
6-Thioguanine (PO)	60 mg/m ² /d	15-28
Methotrexate/ARA-C/Prednisone (IT)	12/30/20 mg/dose ^b	1 ⁱ , 17, 24

Interim maintenance

Methotrexate (PO)	20 mg/m ² /dose ^j	once a week
6-Mercaptopurine (PO)	50 mg/m ² /dose ^j	daily

Maintenance^k

Treatment phase/drug	Single or daily dose	Days of application per phase ^a
Methotrexate (PO)	20 mg/m ² /dose ⁱ	once a week
6-Mercaptopurine (PO)	50 mg/m ² /d ⁱ	daily
Methotrexate/ARA-C/Prednisone (IT)	6x 12/30/20 mg/dose ^b	monthly

^aAdjustments of schedule were allowed if clinical condition and bone marrow recovery were inadequate.

^bDoses of IT drugs were adjusted for children <3 years of age.

^cSteroids were tapered over 9 additional days.

^{***}Differences of treatment given after amendment on November 2016:

- Asparaginase given in induction is PEG-asparaginase IM at dose 1000 IU/m² on days 12 and 26,
- In induction, patients Ph+ only receive a single dose of 1000 IU/m² of PEG-asparaginase on day 12.

^dIT therapy on days 19 and 26 was not administered to patients with BCP ALL who were included in the HR group only because of a poor prednisone response at day 8.

^ePatients with CNS status CNS 3 received 5 weekly doses of IT therapy on days 1, 8, 15, 22, and 29 in Induction Protocol A.

^fT-ALL patients with poor prednisone response at day 8 received one cyclophosphamide at dose 1000 mg/m² on day 10 in induction protocol A.

^gA loading dose of 10% was infused over 30 min, the remaining 90% over 23.5 h. Leucovorin rescue was given at hours 42, 48, and 54 (each 15 mg/m²). Doses of leucovorin rescue were adjusted, if MTX levels were > 1.0 µmol/l at hour 42 or later. If the MTX level at hour 54 was > 0.25 µmol/l, the rescue was continued at six-hour intervals until MTX levels were ≤ 0.25 µmol/l.

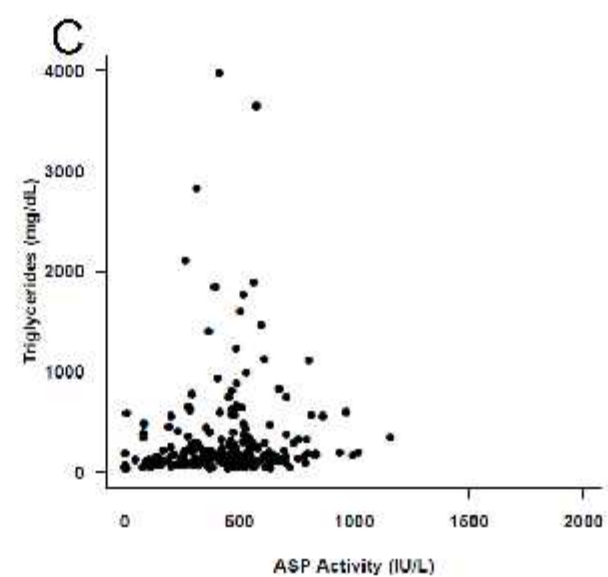
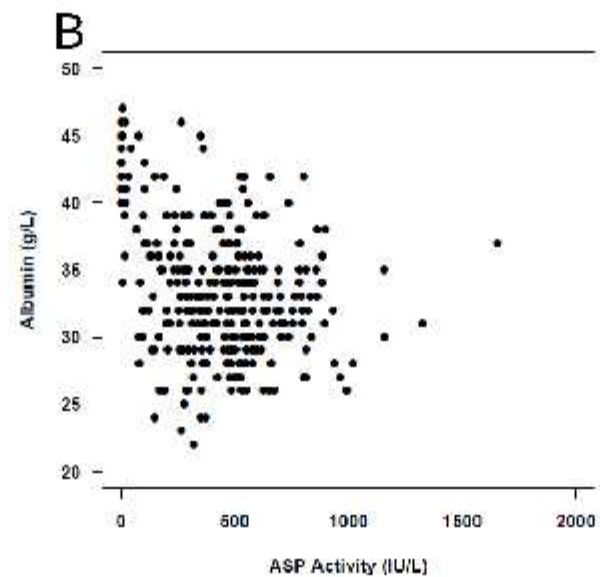
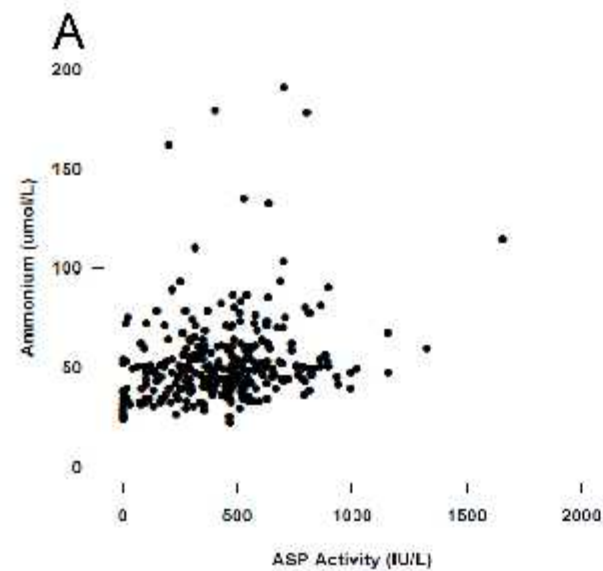
^hIn T-ALL patients with CNS status CNS 3 and without HSCT indication, delayed intensification reinduction 2 was substituted by an additional HR1' and HR2'.

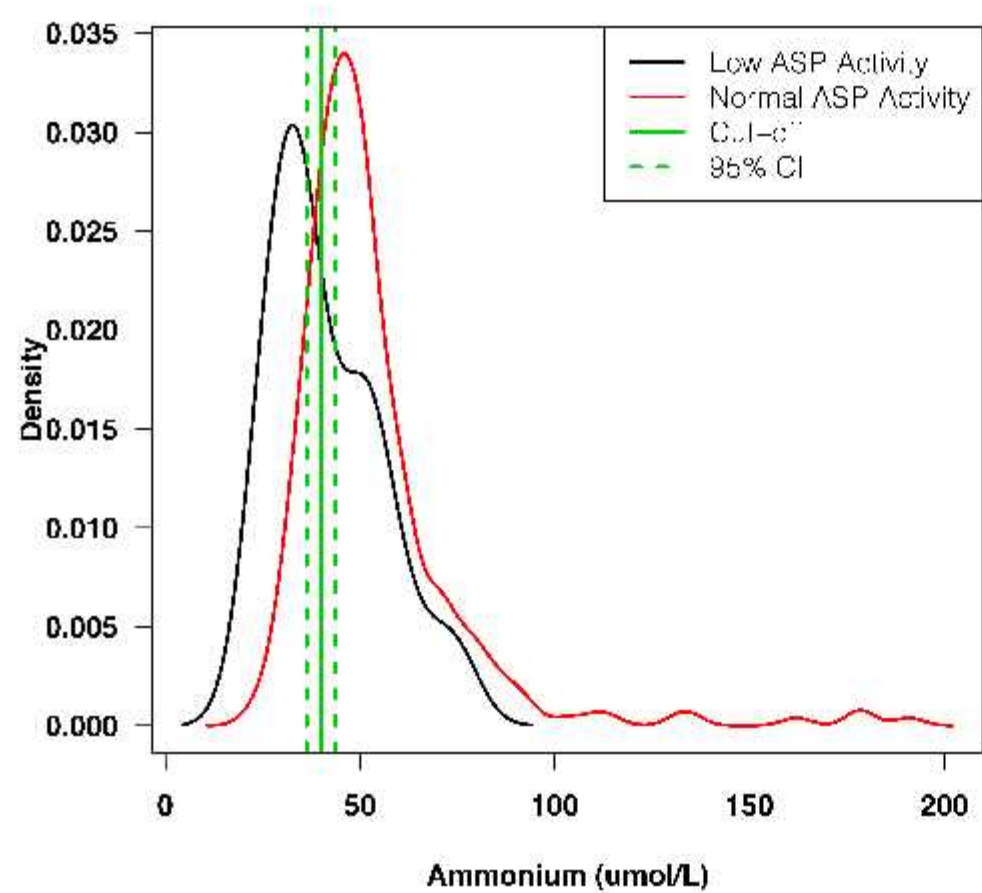
ⁱIT therapy on day 1 in delayed intensification reinduction 1, 2, and 3 was not administered to patients with BCP ALL who were included in the HR group only because of a poor prednisone response at day 8 of induction.

^jDoses were adjusted to WBC (target range 2.0-3.0 x10⁹/L).

^kMaintenance was given from the end of intensive chemotherapy until 104 weeks after diagnosis.

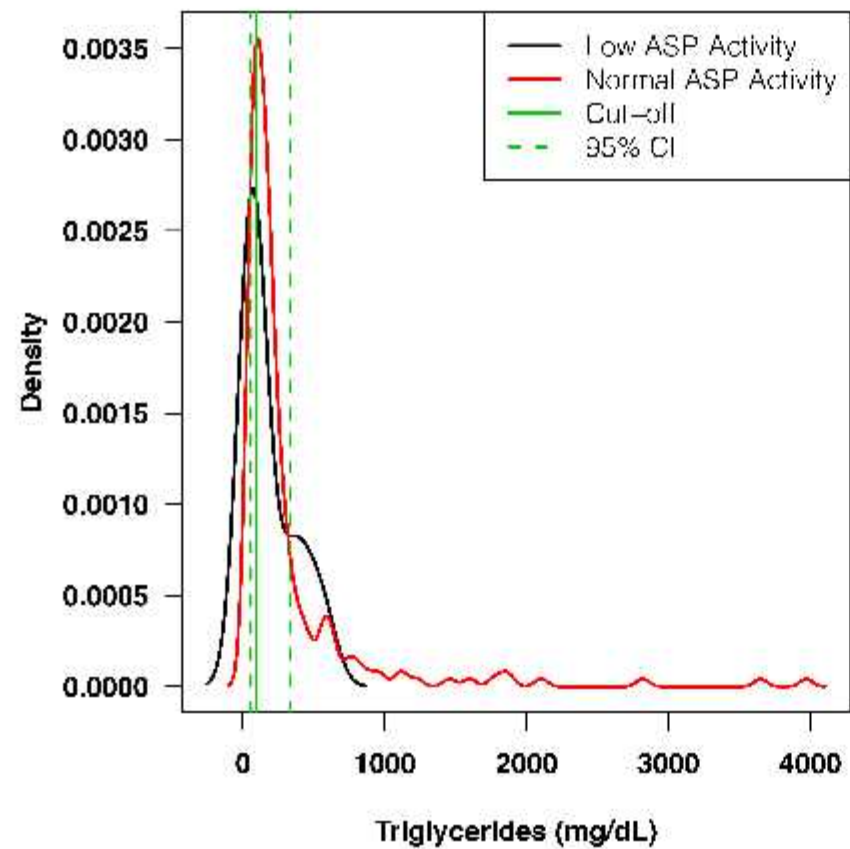
PO indicates orally; IV, intravenous push; PI, intravenous infusion; IT, intrathecal; IM intramuscular.





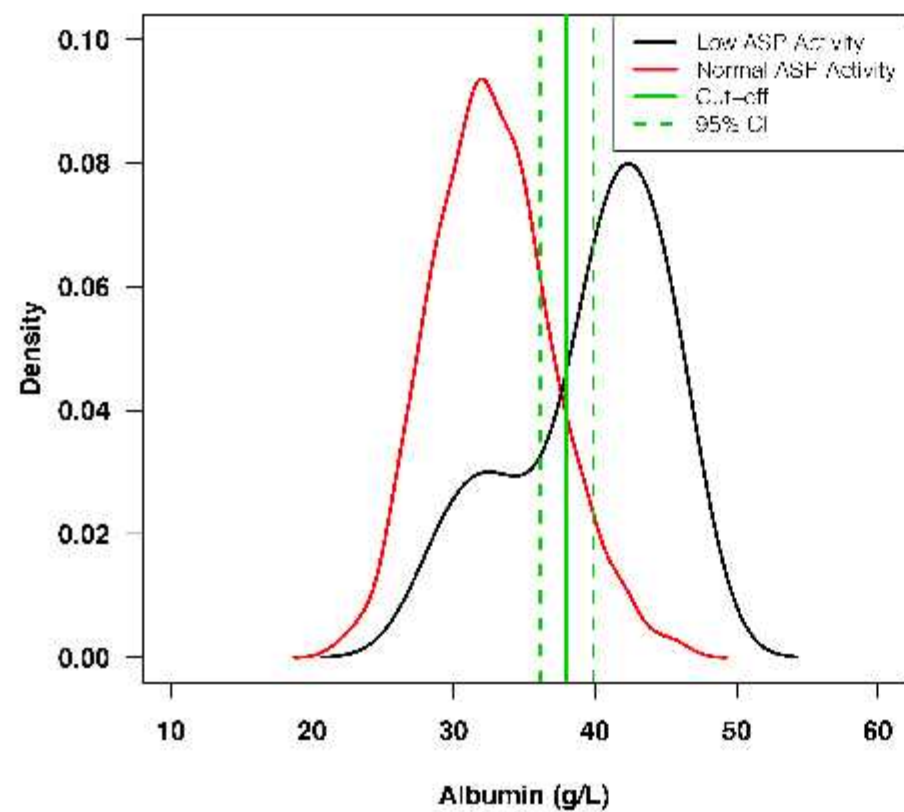
		ASP activity		
		Low	Normal	Total
Ammonium categorized	< 40	18	55	73
	≥ 40	11	220	231
Total		29	275	304

	Estimate	95% CI	
Sensitivity	0.621	0.424	0.787
Specificity	0.800	0.747	0.845
Pos. Pred. Val.	0.247	0.156	0.364
Neg. Pred. Val.	0.952	0.914	0.975



		ASP activity		
		Low	Normal	Total
Triglycerides categorized	< 95	8	53	61
	≥ 95	6	160	166
Total		14	213	227

	Estimate	95% CI	
Sensitivity	0.571	0.296	0.812
Specificity	0.751	0.687	0.807
Pos. Pred. Val.	0.131	0.062	0.248
Neg. Pred. Val.	0.964	0.919	0.985



		ASP activity		
		Low	Normal	Total
Albumin categorized	≥ 38	22	35	57
	< 38	8	242	250
Total		30	277	307

	Estimate	95% CI	
Sensitivity	0.733	0.538	0.870
Specificity	0.874	0.827	0.909
Pos. Pred. Val.	0.386	0.263	0.524
Neg. Pred. Val.	0.968	0.936	0.985