RESEARCH ARTICLE

Cloning and molecular analysis of L-asparaginase II gene (*ansB*)

Zeinat K. Mohamed [1]
Sherif M. Elnagdy [1]
AlaaEddeen Seufi [2]
Mohamed Gamal [1]

Authors' addresses:
[1] Botany and Microbiology Department, Faculty of Science, Cairo University, Egypt.
[2] Entomology Department, Faculty of Science, Cairo University, Egypt.

Correspondence:
Sherif M. Elnagdy
Botany and Microbiology Department, Faculty of Science, Cairo University, Egypt.
Tel.: +20-1008843357
e-mail: sh.elnagdy@gmail.com

Article info:
Received: 19 April 2015
Accepted: 14 August 2015

ABSTRACT

The deamination of L-asparagine to L-aspartic acid and ammonia is catalyzed by L-asparaginases (L-asparagine amino hydrolase). The enzyme L-asparaginase is widely distributed in nature from different living organisms, starting from bacteria till mammals and plants. It has been recently thought to be a therapeutic agent in treatment of various lymphoblastic leukemia diseases. There have been many attempts to isolate microorganisms that produce L-asparaginase. L-ASNase producing bacteria, *Escherichia coli* MG27, was previously isolated from the River Nile and identified. In this study, *ansB* gene, encoding L-ASNase II from *E. coli* MG27, was amplified by PCR, cloned and characterized by DNA sequencing. The DNA sequence was then analyzed using bioinformatics analysis and translated into amino acid sequence. Identification of highly conserved amino acid sequence motifs was conducted by comparison against the InterPro database. Analysis revealed that the protein sequence had a catalytic domain of L-asparaginase type II (IPR004550) that belong to asparaginase/glutaminase family (IPR006034) and has asparaginase/glutaminase conserved site (IPR020827). According to results predicted using PSIpred tool, *ansB* consists of eight α-helices and 13 β-strands.

Key words: L-asparaginase, cloning, *Escherichia coli*, leukemia, sequencing

Introduction

The enzyme L-asparaginase is widely distributed in nature from bacteria, yeast, filamentous fungi, mammals and plants. The amino acid sequence of L-asparaginase II was determined by protein sequencing in the seventies and the nucleotides sequence of the *ansB* gene was reported in 1990. The L-asparaginase II precursor has a 22-residue N-terminal secretory signal peptide which is cleaved between alanine and leucine residues (amino acids 22 and 23, numbering from the initiating methionine residue) to yield mature protein with N-terminal leucine residues. The 22-residue amino acid secretory signal peptide directs the translocation of the protein through the membrane and has a charged N-terminal region, amino acids from 1 to 6, followed by hydrophobic core, amino acids from 7 to 16, and terminated with more polar region, amino acids from 17 to 22. The mature enzyme consists of 326 amino acids located in periplasm as a homo tetramer and has molecular weight of 141 kDa and its synthesis is 100 to 1000 fold induced in anaerobic cultures. Each one of the four active sites is located between the N and C-terminal domains of two adjacent monomers. Thus, the L-asparaginase II tetramer can be treated as a dimer of dimers. Despite this fact, the active enzyme is always a tetramer (Aung et al., 2000; Kozak et al., 2002; Khushoo et al., 2004).

The production of *E. coli* L-asparaginase II is regulated by two pleitropic regulatory proteins, the oxygen-sensitive FNR protein, which activates a number of genes during anaerobiosis (Partridge et al., 2008; Tolla & Savageau, 2011; Shan et al., 2012) and the cyclic AMP receptor protein CRP, which controls the initiation of transcription of genes in various catabolic pathways (Beatty et al., 2003; Uppal et al., 2011; Chen et al., 2012; Kraxenberger et al., 2012).

The information from X-ray crystallography and extensive site-directed mutagenesis studies on *E. coli* asparaginase II, revealed a number of amino acid residues

DOI: 10.69085/jbb20153291

RESEARCH ARTICLE

that are important for catalysis and substrate binding. Among them Thr 12, Tyr 25, Thr 89 and Lys 162 which play a catalytic role, while the positions of Ser 58, Asp 90, Asn 248 and Glu 283 assist substrate binding (Wehner et al., 1994; Palm et al., 1996; Aung et al., 2000).

Many attempts have been made to clone L-asparaginase (*ansB*) gene from different bacterial and fungal species. Hüser et al. (1999) cloned a class II glutaminase/asparaginase coding gene (*ansB*) from *Pseudomonas fluorescens* into pASK-C expression vector that was transformed into *E. coli* CU1783. The expressed fusion protein with a C-terminal His_8-tag was purified by affinity chromatography on Talon-Spin columns.

DNA fragment coding for L-asparaginase from *E. coli* AS1.357 was generated using polymerase chain reaction (PCR) technology. It was cloned into the expression vector pBV220 and transformed into *E. coli* strains JM105, JM109, TG1, DH5α and AS1.357 (Wang et al., 2001). Khushoo et al. (2004) fused the gene coding L-asparaginase II of *E. coli* to an efficient *pelB* leader sequence and an N-terminal 6x histidine tag. The fused gene has been cloned downstream of the T7*lac* promoter in pET22b expression vector.

Saccharomyces cerevisiae gene coding for the periplasmic L-asparaginase II was cloned into pPIC9 expression vector in-frame with the *S. cerevisiae* α-factor secretion signal under the control of the *AOX1* gene promoter and expressed in the methylotrophic yeast *Pichia pastoris* (Ferrara et al., 2006). Furthermore, Kotzia & Labrou (2007) reported the cloning and expression of L-asparaginase from *Erwinia chrysanthemi* 3937 in *E. coli* BL21 (DE3)pLysS, using pCR®T7/CT-TOPO® expression vector.

Youssef & Al-Omair (2008) reported the cloning of L-asparaginase II gene from *E. coli* W3110 into pGEX-2T expression vector in-frame with the glutatione S-transferase fusion protein (GST) in *E. coli* BL21 (DE3) cells. Also, Cappelletti et al. (2008) cloned *ansB* gene from the pathogenic strain *Helicobacter pylori* CCUG 17874. The gene was isolated by PCR using specific primers and the PCR product was cloned into pCR2.1-TOPO cloning vector and sequenced.

Vidya et al. (2011) isolated L-asparaginase II-encoding gene (*ansB*) with excluding the native signal sequence from *E. coli* MTCC 739 by PCR technique. The 981-bp amplicon was cloned into pET20b expression vector and expressed in *E. coli* DE3 cells. Vidya & Pandey (2012) isolated L-asparaginase II gene from a moderately thermotolerant bacterium belonging to Enterobacteriaceae by PCR. They used specific primers that were designed in such a way that the native signal sequence was excluded and the mature gene sequence was cloned into pET20b expression vector with a six histidine sequences at the C-terminal end transformed to competent BL21 DE3 cells. Also, Pokrovskaya et al. (2012) cloned *ansB* gene from *Yersinia pseudotuberculosis* and constructed a stable inducible expression system that overproduce L-asparaginase in *E. coli* BL21 (DE3) cells.

This study targeted *ansB* gene, encoding L-ASNase II from *E. coli* MG27 amplification by PCR, cloning and characterization by DNA sequencing. The DNA sequence was then analyzed using bioinformatics analysis and translated into amino acid sequence.

Materials and Methods

Amplifying of ansB gene by PCR

Genomic DNA preparation

Genomic DNA was prepared from *E. coil* MG27 cells (previously isolated from the River Nile and identified) using GeneJET™ Genomic DNA Purification Kit (Thermo Scientific, USA) and according to its protocol. The purified genomic DNA was store at -20°C.

Primers design for amplification of ansB gene

Primers were designed based on the sequence of *E. coli* K-12 from GenBank accession number M34277 that yielded a single 1047-bp ORF supposed to encode for L-asparaginase II. Primers were designed for amplification of *ansB* gene by excluding the native signal sequence.

A DNA fragment coding for the predicted mature part of L-asparaginase II (amino acid residues 23-346) was amplified using primers *ansB*-F (forward, 5'-GGTGGATCC TTACCCAATATCACCATTTTAG-3') and *ansB*-R (reverse, 5'-GGGAAGCTTTTAGTACTGATTGAAGATCTG-3').

*Bam*HI and *Hin*dIII restriction sites (underlined) were incorporated into the primers at their 5' ends, respectively, to facilitate the directional cloning of the structural asparaginase II gene.

Polymerase chain reaction (PCR)

The genomic DNA isolated from *E. coli* strain MG27 was used as template for the amplification of *ansB* gene. It was amplified by polymerase chain reaction without its native signal sequence using *ansB*-F and *ansB*-R primers.

The PCR reaction was carried out in a total volume of 25µl. The PCR conditions were as follows: initial

DOI: 10.69085/jbb20153291

RESEARCH ARTICLE

denaturation at 94°C for 4 min, denaturation at 95°C for 40 s, annealing at 52°C for 40 s, extension at 72°C for 60 s for 35 cycles, and a final extension at 72°C for 10 min. The PCR product was then analyzed on 1% agarose gel electrophoresis.

Agarose gel electrophoresis

For analyzing DNA samples, agarose gel electrophoresis was used. The DNA bands in the gel were visualized using short wave ultraviolet light provided by a transilluminator and photographed.

Elution and purification of PCR fragments

To insure high purity of PCR fragments, the amplified DNA bands were eluted and purified form agarose gels using QIAquick Gel Extraction Kit according to manufacturer's protocol.

TA-cloning of ansB gene

Ligation

The gel purified PCR product was ligated into the pGEM-T Easy vector (Promega crop. Madison, Wi, USA) using T4 DNA ligase provided in the kit according to the manufacturer's instructions. The concentration of the gel purified PCR product was measured prior to ligation. One µl of the resultant recombinant construct was used to transform *E. coli* JM109 competent cells

Preparation and Transformation of $CaCl_2$ competent cells

For preparation and transformation of competent cells, $CaCl_2$-treatment was performed (Sambrook & Russel, 2001).

Screening of positive colonies

E. coli JM109 colonies harboring the recombinant plasmid were screened for the presence of insert in pGEM-T easy vector by blue/white color selection on plate containing ampicillin, 5-bromo-4-chloro-3-indolyl-β-D-galactosidase (X-Gal) and isopropyl-β-D-thiogalactopyranoside (IPTG). White colonies were randomly selected. The presence of insert was verified by PCR and restriction endonuclease digestion of plasmids isolated from these white colonies.

Plasmid DNA preparation

Plasmid DNA was prepared from white colonies of transformed *E. coli* JM109 cells using QIAprep Spin Miniprep Kit. The purified plasmid DNA was stored at -20°C until used for confirmation of cloning by PCR amplification and restriction digestion.

Verification tests for the recombinant ansB clones

Polymerase chain reaction (PCR) screening

Plasmid DNA prepared from white clones was used as template for confirmation of transformation using PCR. The presence of *ansB* gene was checked by PCR using the insert specific primers *ansB*-F and *ansB*-R.

The PCR reaction was carried as previously described. Plasmids from blue clones served as negative control. The PCR products were analyzed on 1% agarose gel electrophoresis for selection of right clones.

Restriction digestion

Cloning of the *ansB* gene into pGEM-T Easy vector was confirmed by single and double restriction enzyme digestion of the recombinant plasmid with specific restriction enzymes. Single restriction of the recombinant plasmid was done with *Bam*HI while double restriction was carried out with *Bam*HI and *Hin*dIII.

For linearization of plasmids, one microgram of plasmids was digested with FastDigest *Bam*HI restriction enzyme in 20 µl volume with FastDigest buffer. Digestion reactions were incubated at 37°C for 5 minutes. For release the insert, double digestion was carried out using FastDigest *Bam*HI and FastDigest *Hin*dIII restriction enzymes with FastDigest buffer. Reactions were incubated at 37°C in heat blocks for 5 minutes, and then electrophoresed into 1% agarose gel. Agarose gels were visualized by ultraviolet transilluminator. The released inserts were eluted and purified form agarose gels using QIAquick Gel Extraction Kit as described before. Gel purified inserts were used for subcloning into pQE-30 expression vector.

Nucleotide sequencing

The nucleotide sequence of the insert was determined according to Sanger dideoxy chain-termination method at GATC Biotech (Konstanz, Germany), using M13 forward and reverse primers. The DNA sequence was determined by automated DNA sequencing method using ABI 3730xI sequence analyzer (Applied Biosystems, Foster City, CA, USA). The forward and reverse DNA sequence reads were assembled to obtain the consensus sequence by using DNA Baser Sequence Assembler software v.3.5.3.

Bioinformatics analysis

The DNA sequence was analyzed and translated into amino acid sequence using the BioEdit program (Hall, 1999). Restriction analysis was done using CodonCode Aligner (version 3.5.6). ProtParam tool (Gasteiger et al., 2005) was used for computing physicochemical properties that can be deduced from a protein sequence query, such as molecular weight, theoretical pI, amino acid composition, instability index, aliphatic index and grand average of hydropathicity.

DOI: 10.69085/jbb20153291

RESEARCH ARTICLE

Nucleotide and amino acid sequence data were analyzed against all sequences in the GenBank using the basic local alignment search tool (BLAST) (Altschul et al., 1997). Sequence alignments were performed using CLC Main Workbench (version 6.5).

The Conserved Domain Database tool (CDD) (Marchler-Bauerwas et al., 2009) was used to search for proteins with similar domain architecture and superfamilies that show specific sequence matches complementary to the amino acid query sequence. PROSITE tool (De Castro et al., 2006) was used for the detection of known motifs for the classification of a protein into families sharing functional attributes derived from a common ancestor. PROSITE is a database of protein families, domains and associated patterns as well as of signatures and functional sites. INTERPRO tool (Quevillon et al., 2005) was used for classification and characterization of the protein sequence based on consensus sequences between known families, domains and models and the amino acid query sequence. Protein secondary structure of the putative L-asparaginase II was predicted using PSIpred tool (Jones, 1999).

Results

Amplifying of ansB gene by PCR

Amplification of 981-bp fragment was performed using pair of specific primers *ansB*-F and *ansB*-R incorporating the sequence for the restriction endonucleases *Bam*HI and *Hin*dIII, respectively (Figure 1).

TA-Cloning of ansB into pGEM-T Easy vector

Three white colonies were picked and subjected to confirmation procedures to detect the recombinant clones harboring the putative gene encoding L-ASNase II from local isolate *E. coli* strain MG27(Figure 2).

Verification tests for the recombinant ansB clones

Three white colonies designated W1, W2 and W3 were picked from LB/ampicillin/IPTG/X-gal plate. These clones were plated on LB-ampicillin plates, as well as being cultured in LB-ampicillin broth for Plasmid DNA minipreps. The presence of insert was confirmed by PCR and restriction digestion.

Polymerase chain reaction (PCR) screening

Figure 3 showed the amplified products of clones W1, W2 and W3 which had the same expected size (981 bp) for PCR product of *ansB* gene as in positive control (lane 5). Positive control was conducted by using genomic DNA of *E. coli* MG27 as DNA template. Negative control was conducted by using plasmid from a blue colony as DNA template and showed no band indicating no recombination (lane 4).

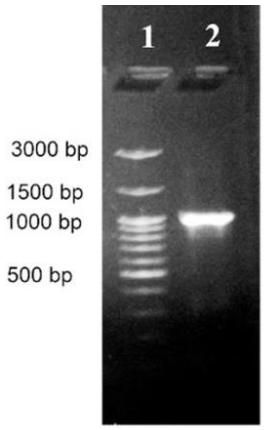

Figure 1. *Agarose gel electrophoresis of PCR amplification of putative ansB gene amplified from genomic DNA of E. coli MG27. Lane 1, DNA marker; lane 2, PCR amplicon of putative ansB gene. Genomic DNA isolated from E. coli strain MG27 was used as a template for PCR amplification of ansB gene without its native signal sequence using ansB-F and ansB-R specific primers. The PCR product analyzed using 1% agarose gel electrophoresis showed a fragment of the expected size (981 bp) of ansB gene.*

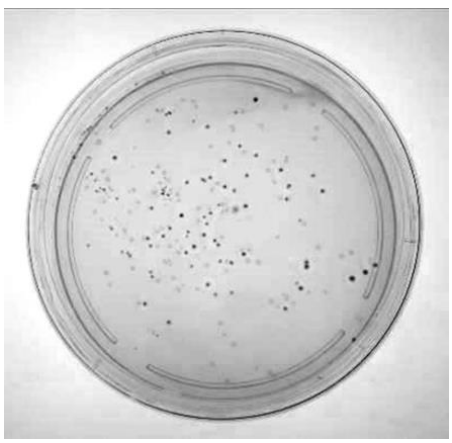

Figure 2. *Blue/White screening of transformants. LB/ampicillin/IPTG/X-gal plate was plated with 50 µl of transformation mixture and incubated at 37°C for 18-24 h. White colonies represent recombinant clones while blue colonies represent empty clones.*

DOI: 10.69085/jbb20153291

RESEARCH ARTICLE

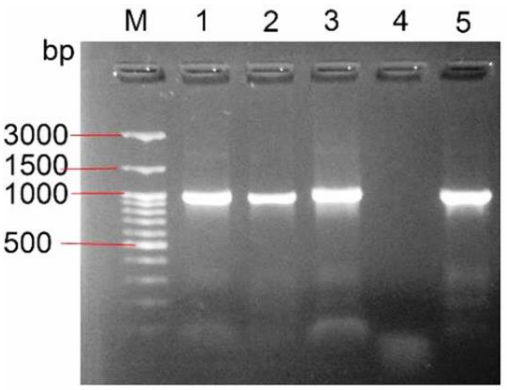

Figure 3. *Screening of white clones for recombinant pGEM-ansB plasmid by PCR. Lane M, DNA marker; lanes 1-3, PCR products from white clones W1, W2 and W3, respectively. Lane 4, negative control; lane 5, positive control. Plasmid DNA was prepared from white clones W1, W2 and W3 and used as DNA template for PCR reactions using insert-specific primers ansB-F and ansB-R. Negative control was conducted by using plasmid from a blue colony as DNA template. Positive control was conducted by using genomic DNA of E. coli MG27 as DNA template. PCR amplicons were analyzed on 1% agarose gel electrophoresis. White clones W1, W2 and W3 generated fragments of expected size 981 bp confirming the presence of putative ansB gene.*

Confirmation of transformation by restriction digestion

Linearized vector in lane 3 showed one band at the expected size (~4 Kb) while two bands were observed in case of double digested vector (lane 4). A DNA fragment about 981 bp was released from the vector backbone (~3.0 kb). Plasmid extracted from the clone W3 was designated pGEM-*ansB* and subjected to the nucleotide sequence analysis (Figure 4).

Nucleotide sequencing

Nucleotide sequence of the putative *ansB* gene of *E. coli* MG27 was submitted to the NCBI database and an accession number KC416966 was assigned while the deduced amino acid sequence was submitted under accession number AGE81914.

Bioinformatics analysis

Results revealed that, sequence of putative *ansB* consists of 981 bp codes for 326 amino acids (Figure 5). Restriction map of putative *ansB* gene generated using CodonCode aligner program was represented in Figure 6.

Using ProtParam tool of ExPASy, The results showed that the protein contains 20 amino acids with valine in the highest percentage (10.7%) while tryptophan was the lowest (0.3%) as detailed in Table 1. Other physiochemical properties obtained from ProtParam analysis has been tabulated (Table 2). The instability index of the protein calculated using ProtParam tool was 19.61. The predicted theoretical isoelectric point (pI) value was 5.66 and its molecular weight was estimated to be 34.5788 kDa. The number of negatively charged residues (Asp + Glu) was greater than the number of positively charged residues (Arg + Lys). ProtParam results showed that 33 residues are negatively charged and 30 residues are positively charged. Additionally the grand average of hydropathicity (GRAVY) and aliphatic index were computed to be -0.214 and 84.33, respectively.

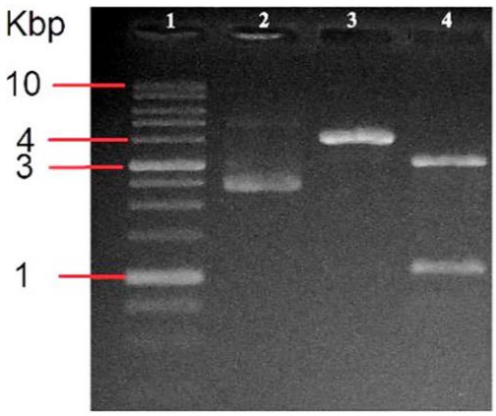

Figure 4. *Restriction analysis of constructed pGEM-ansB vector. Lane 1, 1 kb DNA marker; lane 2, undigested vector; lane 3, Linearized vector digested with BamHI restriction enzyme; lane 4, vector digested with BamHI and HindIII restriction enzymes. Plasmid DNA prepared from clone W3 was single and double digested and analyzed on 1% agarose gel electrophoresis. Single digestion with BamHI resulted in a linearized vector at the expected size (~4 kb). Two bands were observed when vector was double digested with BamHI and HindIII. Bands in lane 4 represent the expected vector backbone (~3.0 kb) and the released insert (~981 bp).*

The deduced amino acid sequence was utilized for similarity search through BLAST at NCBI selecting non-redundant database. BLAST analysis on the deduced amino acid sequence of putative *ansB* gene from *E. coli* MG27 showed 100% identity with L-asparaginases II from *Shigella sonnei* Ss046 (accession number YP_312053.1), *Escherichia coli* OK1357 (accession number WP_001345951.1) and *E. coli* MS 79-10 (accession number WP_001012363.1). In

DOI: 10.69085/jbb20153291

RESEARCH ARTICLE

addition, significant homology with 99-87% similarity was found with several bacterial L-asparaginases including L-asparaginase II from *Escherichia coli* KTE66 (accession number WP_001559780.1), *Shigella boydii* 5216-82 (accession number WP_000394146.1), *Citrobacter freundii* (accession number ACC85692.1), *C. youngae* ATCC 29220 (accession number WP_006686837.1), *Salmonella enterica* subsp. *enterica* (accession number WP_000394193.1), *Enterobacter cloacae* SCF1 (accession number YP_003940358.1) and *Serratia marcescens* VGH107 (accession number WP_004928279.1). Multiple sequence alignment was conducted on these sequences using CLC program. Alignment results revealed that several highly conserved domains were extended along *ansB* sequences.

1	**TTA CCC AAT ATC ACC ATT TTA GCA ACC GGC GGG ACC ATT GCC GGT**	45
1	Leu Pro Asn Ile Thr Ile Leu Ala Thr Gly Gly Thr Ile Ala Gly	15
46	**GGT GGT GAC TCC GCA ACC AAA TCT AAC TAC ACA GCG GGT AAA GTT**	90
16	Gly Gly Asp Ser Ala Thr Lys Ser Asn Tyr Thr Ala Gly Lys Val	30
91	**GGC GTA GAA AAT CTG GTT AAT GCG GTG CCG CAA CTA AAA GAC ATT**	135
31	Gly Val Glu Asn Leu Val Asn Ala Val Pro Gln Leu Lys Asp Ile	45
136	**GCG AAC GTT AAA GGC GAG CAG GTA GTG AAT ATC GGC TCC CAG**	180
46	**GAC**	60
	Ala Asn Val Lys Gly Glu Gln Val Val Asn Ile Gly Ser Gln Asp	
181	**ATG AAC GAT AAT GTC TGG CTG ACA CTG GCG AAA AAA ATT AAC ACC**	225
61	Met Asn Asp Asn Val Trp Leu Thr Leu Ala Lys Lys Ile Asn Thr	75
226	**GAC TGC GAT AAA ACC GAC GGC TTC GTC ATT ACC CAC GGT ACC GAC**	270
76	Asp Cys Asp Lys Thr Asp Gly Phe Val Ile Thr His Gly Thr Asp	90
271	**ACG ATG GAA GAA ACC GCT TAC TTC CTC GAC CTG ACG GTG AAA TGC**	315
91	Thr Met Glu Glu Thr Ala Tyr Phe Leu Asp Leu Thr Val Lys Cys	105
316	**GAC AAA CCG GTG GTG ATG GTC GGC GCA ATG CGC CCG TCC ACG**	360
106	**TCC**	120
	Asp Lys Pro Val Val Met Val Gly Ala Met Arg Pro Ser Thr Ser	
361	**ATG AGC GCA GAC GGT CCA TTC AAC CTG TAT AAC GCG GTA GTG ACC**	405
121	Met Ser Ala Asp Gly Pro Phe Asn Leu Tyr Asn Ala Val Val Thr	135
406	**GCA GCT GAT AAA GCC TCC GCT AAT CGT GGC GTG CTG GTG GTG**	450
136	**ATG**	150
	Ala Ala Asp Lys Ala Ser Ala Asn Arg Gly Val Leu Val Val Met	
451	**AAC GAC ACC GTA CTG GAC GGT CGC GAT GTC ACC AAA ACC AAC ACC**	495
151	Asn Asp Thr Val Leu Asp Gly Arg Asp Val Thr Lys Thr Asn Thr	165
496	**ACC GAC GTA GCG ACC TTC AAG TCT GTT AAC TAC GGT CCT CTG GGA**	540
166	Thr Asp Val Ala Thr Phe Lys Ser Val Asn Tyr Gly Pro Leu Gly	18
541	**TAC ATT CAC AAC GGT AAG ATT GAC TAC CAA CGT ACC CCG GCA CGT**	585
181	Tyr Ile His Asn Gly Lys Ile Asp Tyr Gln Arg Thr Pro Ala Arg	195
586	**AAG CAC ACC AGC GAT ACG CCA TTC GAT GTC TCT AAG CTG AAT GAG**	630
196	Lys His Thr Ser Asp Thr Pro Phe Asp Val Ser Lys Leu Asn Glu	210
631	**CTG CCG AAA GTC GGC ATC GTT TAT AAC TAC GCT AAC GCA TCC GAT**	675
211	Leu Pro Lys Val Gly Ile Val Tyr Asn Tyr Ala Asn Ala Ser Asp	225
676	**CTT CCG GCT AAA GCA CTG GTA GAT GCG GGC TAT GAT GGC ATC GTT**	720
226	Leu Pro Ala Lys Ala Leu Val Asp Ala Gly Tyr Asp Gly Ile Val	240
721	**AGC GCT GGT GTG GGT AAT GGT AAC CTG TAT AAA TCC GTG TTC GAC**	765
241	Ser Ala Gly Val Gly Asn Gly Asn Leu Tyr Lys Ser Val Phe Asp	255
766	**ACC CTG GCA ACC GCC GCG AAA AAC GGC ACT GCA GTA GTG CGT**	810
256	**TCT**	270
	Thr Leu Ala Thr Ala Ala Lys Asn Gly Thr Ala Val Val Arg Ser	
811	**TCC CGC GTA CCG ACG GGT GCT ACC ACT CAG GAT GCT GAA GTG**	855
271	**GAT**	285
	Ser Arg Val Pro Thr Gly Ala Thr Thr Gln Asp Ala Glu Val Asp	
856	**GAT GCG AAA TAC GGC TTC GTC GCC TCT GGC ACG CTG AAC CCG CAA**	900
286	Asp Ala Lys Tyr Gly Phe Val Ala Ser Gly Thr Leu Asn Pro Gln	300
901	**AAA GCG CGC GTC CTG CTG CAG CTG GCT CTG ACG CAA ACC AAA**	945
301	**GAT**	315
	Lys Ala Arg Val Leu Leu Gln Leu Ala Leu Thr Gln Thr Lys Asp	
946	**CCG CAG CAG ATC CAG CAG ATC TTC AAT CAG TAC TAA**	981
316	Pro Gln Gln Ile Gln Gln Ile Phe Asn Gln Tyr End	326

Figure 5. *Nucleotide sequence of putative L-asparaginase II gene (ansB) and its deduced amino acid sequence. The sequence extends, 981 nucleotid length and the translation product of the ansB gene was shown below the nucleotide sequence.*

DOI: 10.69085/jbb20153291

RESEARCH ARTICLE

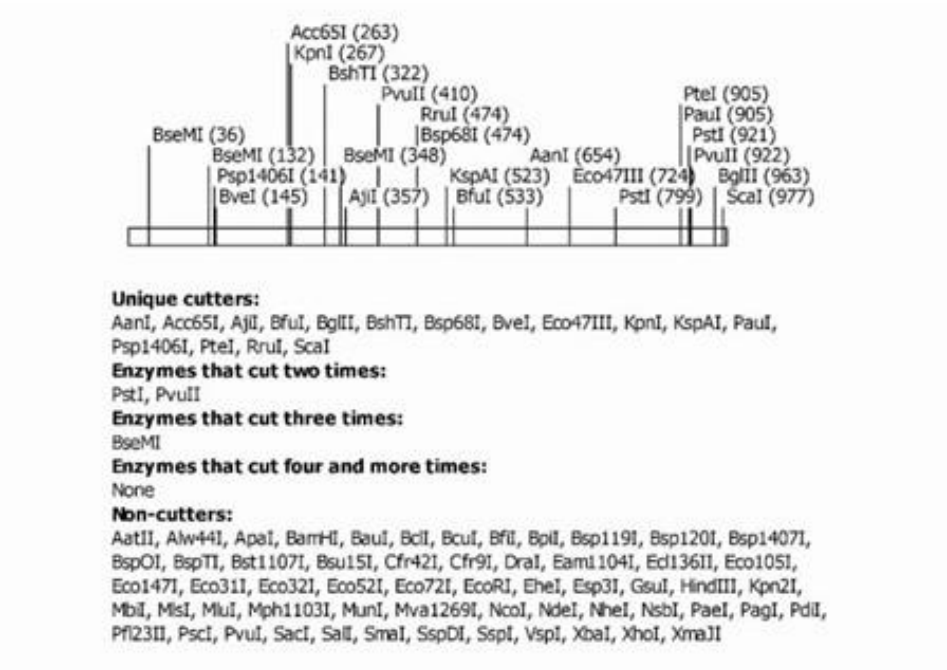

Figure 6. *Restriction map of the putative ansB gene.*

Table 1. *Amino acid composition of E. coli MG27 putative ansB calculated using the ProParam tool of ExPASy.*

Amino acid	Number of residues	Percentage
Alanine	33	10.1
Arginine	8	2.5
Asparagine	24	7.4
Aspartic acid	27	8.3
Cysteine	2	0.6
Glutamine	13	4.0
Glutamic acid	6	1.8
Glycine	28	8.6
Histidine	3	0.9
Isoleucine	13	4.0
Leucine	23	7.1
Lysine	22	6.7
Methionine	6	1.8
Phenylalanine	8	2.5
Proline	13	4.0
Serine	16	4.9
Threonine	33	10.1
Tryptophan	1	0.3
Tyrosine	12	3.7
Valine	35	10.7
pyrrolysine	0	0.0
Selenocysteine	0	0.0

Table 2. *Physicochemical parameters of E. coli MG27 putative ansB computed using ExPASy's ProtParam tool.*

Parameter	Value
Number of amino acids	326
The instability index	19.61
Theoretical isoelectric point (pI)	5.66
Molecular weight	34.5788 kDa
Total number of negatively charged residues (Asp + Glu)	33
Total number of positively charged residues (Arg + Lys).	30
Grand average of hydropathicity (GRAVY)	-0.214
Aliphatic index	84.33

The deduced amino acid sequence of *E. coli* MG27 *ansB* gene was scanned for conserved residues by the Conserved Domain Database (CDD). The CDD results revealed that the mature protein contains a conserved domain of L-asparaginase-like superfamily (CDD accession: cl00216) and L-asparaginase-like domain (CDD accession: cd00411) as visualized in Figure 7.

DOI: 10.69085/jbb20153291

RESEARCH ARTICLE

Figure 7. *Prosite analysis of the deduced amino acid sequence of putative ansB showing two active domains. Active site signatures are ASN_GLN_ASE_1 (PS00144) and ASN_GLN_ASE_2 (PS00917) at amino acids 6-14 and 82-92, respectively.*

Analysis with PROSITE program revealed that L-asparaginase II contains two recognizable structural domains and their locations are marked below (shadowed). The first structural domain (residues 6-14) is located near the N-terminal section while the second structural domain located within residues 82 to 92.

LPNITILATGGTIAGGGDSATKSNYTAGKVGVENLVNAVPQ
LKDIANVKGEQVVNIGSQDMNDNVWLTLAKKINTDCDKTDG
FVITHGTDTMEETAYFLDLTVKCDKPVVMVGAMRPSTSMSA
DGPFNLYNAVVTAADKASANRGVLVVMNDTVLDGRDVTKTN
TTDVATFKSVNYGPLGYIHNGKIDYQRTPARKHTSDTPFDV
SKLNELPKVGIVYNYANASDLPAKALVDAGYDGIVSAGVGN
GNLYKSVFDTLATAAKNGTAVVRSSRVPTGATTQDAEVDDA
KYGFVASGTLNPQKARVLLQLALTQTKDPQQIQQIFNQY

The two asparaginase/glutaminase active site signatures are ASN_GLN_ASE_1 (PS00144) and ASN_GLN_ASE_2 (PS00917) having active site consensus pattern [LIVM]-x-{L}-T-G (2)-T-[IV]-[AGS] and [GA]-x-[LIVM]-x (2)-H-G-T-D-T-[LIVM]. The amino acids 12 and 89 are the active site residues, respectively. The results of the InterPro database are summarized below (Table 3; Figure 8). InterProScan analysis revealed that the protein sequence had a catalytic domain of L-asparaginase type II (IPR004550) that belong to asparaginase/glutaminase family (IPR006034) and has Asparaginase/glutaminase conserved site (IPR020827).

Table 3. *Protein signatures and functional domains of ansB protein identified using InterProScan.*

Source database accession	InterPro accession	Signature ID	Amino acids	E-value
TGRFAMs / TIGR00520	IPR004550	asnASE_II	1-326	3.7e-166
PIRSF / PIRSF001220	IPR006034	L-ASNase_gatD	1-326	1.8e-94
SUPERFAMILY / SSF53774		Asp/Glutamnse	1-326	4.0e-111
PROSITE patterns / PS00144	IPR020827	ASN_GLN_ASE_1	6-14	1.0
PROSITE patterns / PS00917		ASN_GLN_ASE_2	82-92	1.0

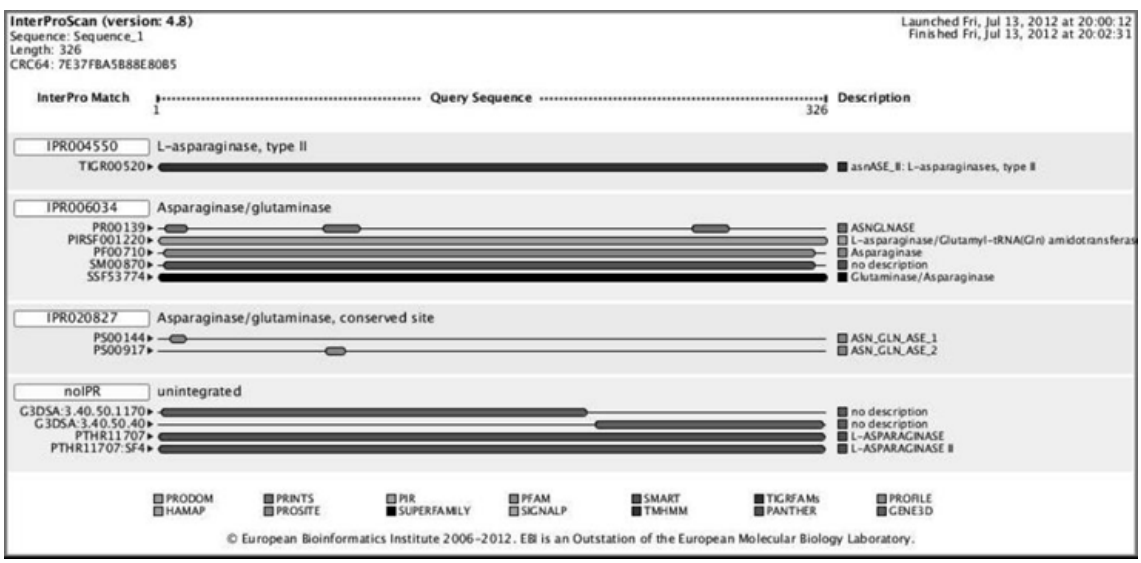

Figure 8. *Graphical view of InterProScan showing functional domains from the deduced amino acid sequence of L-asparaginase II. Characteristic signatures of L-ASNase II which were found after analysis of the 326 amino acid sequence in the integrative protein signature database InterPro.*

DOI: 10.69085/jbb20153291

RESEARCH ARTICLE

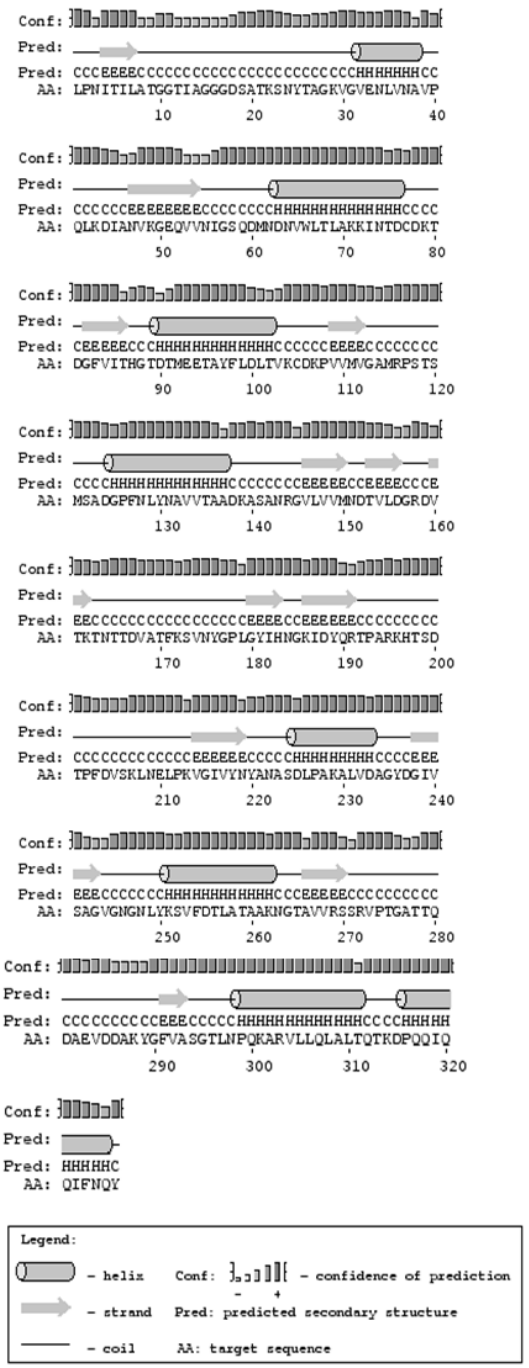

Figure 9. *Predicted secondary structure of L-asparaginase II. Secondary structure prediction was performed based on position-specific scoring matrices using the PSIpred method. The sequences marked as 'H', 'E' and 'C' correspond to helix, strand and coil, respectively. According to the PSIpred prediction, the protein has 8 α-helices and 13 β-strands, as shown.*

The PSIpred database predicted the following secondary structures of the L-asparaginase II from the amino acid sequence as depicted in Figure 9. PSIpred program revealed that *ansB* consists of eight α-helices and 13 β-strands. *ansB* was predicted to be formed of approximately 27% of α-helices (88 residues) and 20% of β-strands (63 residues) while random coils represents 53% (175 residues).

Discussion

Studying L-Asparaginase (L-ASNase) has recently gained much attention for its anti-carcinogenic potential. Several authors documented the use of L-ASNase in cancer therapy (Avramis & Panosyan 2005; Narta et al., 2007; Pieters et al., 2011; Tong et al., 2013). Although L-ASNases are present in many plants, mammalian and bacterial species, only the enzymes from *Escherichia coli* and *Erwinia chrysanthemii* have been produced on industrial scale as chemotherapeutics in acute lymphoblastic leukemia. This is due to their high catalytic activity and specificity towards L-asparagine (Müller & Boos 1998; Aghaiypour et al., 2001; Duval et al., 2002). Apart from the therapeutic use, L-ASNase has a potent application in food industry to reduce acrylamide formation in heat-processed products (Friedman & Levin, 2008; Pedreschi et al., 2008; Kukurová et al., 2013).

The mature sequence of *ansB* gene was amplified from the genomic DNA of a moderately thermotolerant bacterium belonging to Enterobacteriaceae by PCR and the amplicon of ~980 bp was cloned into pET20b vector (Vidya & Pandey, 2012). Furthermore, a putative L-asparaginase gene consisting of 981 bp was amplified by PCR from the *Pyrococcus furiosus* genomic DNA and the PCR product was cloned into a pET14b vector (Bansal et al., 2010). Many investigators cloned L-asparaginase coding genes from various bacteria such as *Pseudomonas fluorescens* (Hüser et al., 1999), *E. coli* (Wang et al., 2001), *Erwinia carotovora* (Kotzia & Labrou, 2005), *E. crysanthemi* (Kotzia & Labrou, 2007), *Yersinia pseudotuberculosis* (Pokrovskaya et al., 2012) and *Bacillus subtilis* (Jia et al., 2013).

In the present study, the clone harboring pGEM-*ansB* construct was selected for nucleotide sequencing using M13 forward and reverse primers. Nucleotide sequence of the putative *ansB* gene of *E. coli* MG27 was submitted to the NCBI database and an accession number KC416966 was assigned while the predicted protein sequence was submitted under accession number AGE81914. Based on the instability index of the deduced amino acids predicted by ProtParam

DOI: 10.69085/jbb20153291

RESEARCH ARTICLE

tool, the protein is stable with a value of 19.61.

The grand average of hydropathicity (GRAVY) of the deduced amino acids was computed to be -0.214. This negative value of GRAVY suggests the hydrophilicity of the protein. The aliphatic index of the deduced amino acids predicted using ProtParam tool was 84.33. This high aliphatic index indicates that the protein can be stable within a wide range of temperature.

Based on BLAST analysis, the deduced amino acid sequence of mature L-ASNase II showed 100% identity with L-ASNase II from *Escherichia coli* OK1357 (WP_001345951.1), *E. coli* MS 79-10 (WP_001012363.1). BLAST analysis of the deduced amino acid sequence revealed significant similarity (99- 87%) with L-ASNase II of *Shigella boydii* 5216-82 (WP_000394146.1), *Citrobacter freundii* (ACC85692.1), *C. youngae* ATCC 29220 (WP_006686837.1), *Salmonella enterica* subsp. *enterica* (WP_000394193.1), *Enterobacter cloacae* SCF1 (YP_003940358.1) and *Serratia marcescens* VGH107 (WP_004928279.1).

The amino acid sequence alignment of putative L-ASNase II from *E. coli* MG27 with L-ASNase II sequences from other 10 strains of bacteria revealed that the sequence of this enzyme is highly conserved especially with Thr-12, Tyr-25, Thr-89, Asp-90, and Lys-162. It was suggested that these residues are essential for reaction the enzymatic activity (Wehner et al., 1994). Thr-12 and Thr-89 are able to act as primary nucleophiles (Harms et al., 1991; Palm et al., 1996). Thr-12 and the adjacent Tyr-25 are components of a mobile loop that closes over the active site during catalysis while Thr-89, Asp-90 and Lys-162 are all located in a rigid part of the structure (Aung et al., 2000; Derst et al., 2000). In this study, the amino acid sequence differs in two positions from the sequence published by Jennings & Beacham (1990). These changes are of substitution type, where alanine is present instead of valine and asparagine is present instead of threonine at residues 27 and 263, respectively. These differences may result in enzymes with different activities between different strains.

NCBI conserved-domain search of the deduced protein revealed the presence of a conserved L-asparaginase-like superfamily domain (CDD accession: cl00216) and L-asparaginase-like domain (CDD accession: cd00411). As predicted by the Prosite program, the deduced amino acid sequence contains two recognizable structural domains. These asparaginase/glutaminase active site signatures are ASN_GLN_ASE_1 (PS00144) and ASN_GLN_ASE_2 (PS00917) located within residues 6 to14 and 82 to 92, respectively. Prosite analysis revealed that two threonine residues at 12 and 89 are the active site residues. Harms et al. (1991) provided an evidence for the importance of threonine-12 for catalytic activity of L-ASNase II that lost its activity when Thr-12 is mutated to alanine. In addition, Thr-89 was postulated to play a catalytic role in ASNase II activity (Swain et al., 1993; Palm et al., 1996).

In the present study, highly conserved amino acid sequence motifs were identified by comparison against the InterPro database. InterProScan analysis revealed that the protein sequence had a catalytic domain of L-asparaginase type II (IPR004550) that belong to asparaginase/glutaminase family (IPR006034) and has Asparaginase/glutaminase conserved site (IPR020827). According to results predicted using PSIpred tool, *ansB* consists of eight α-helices and 13 β-strands. C-terminal domain (residues 213-326) was predicted to be consisted of four β-strands and four α-helices. These results agree with Swain et al. (1993).

References

Aghaiypour K, Wlodower A, Lubkowski J. 2001. Structural basis for the activity and substrate specificity of *Erwinia chrysanthemi* L-asparaginase. Biochemistry, 40: 5655-5664.

Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res., 25: 3389-3402.

Aung HP, Bocola M, Schleper S, Röhm KH. 2000. Dynamics of a mobile loop at the active site of *Escherichia coli* asparaginase. Biochim. Biophys. Acta, 1481: 349-359.

Avramis VI, Panosyan EH. 2005. Pharmacokinetic /pharmacodynamic relationships of asparaginase formulations: the past, the present and recommendations for the future. Clin. Pharmacokinet., 44: 367-393.

Bansal S, Gnaneswari D, Mishra P, Kundu B. 2010. Structural stability and functional analysis of L-asparaginase from *Pyrococcus furiosus*. Biochemistry (Moscow), 75: 375-381.

Beatty CM, Browning DF, Busby SJ, Wolfe AJ. 2003. Cyclic AMP receptor protein-dependent activation of the *Escherichia coli* acsP2 promoter by a synergistic class III mechanism. J. Bacteriol., 185: 5148-5157.

Cappelletti D, Chiarelli LR, Pasquetto MV, Stivala S, Valentini G, Scotti C. 2008. *Helicobacter pylori* L-asparaginase: A promising chemotherapeutic agent. Biochemical and Biophysical Research Communications, 377: 1222-1226.

Chen YP, Lin HH, Yang CD, Huang SH, Tseng CP. 2012. Regulatory role of cAMP receptor protein over *Escherichia coli* fumarase genes. J. Microbiol., 50: 426-433.

DOI: 10.69085/jbb20153291

RESEARCH ARTICLE

De Castro E, Sigrist CJA, Gattiker A, Bulliard V, Langendijk-Genevaux PS, Gasteiger E, Bairoch A, Hulo N. 2006. ScanProsite: detection of PROSITE signature matches and ProRule-associated functional and structural residues in proteins. Nucleic Acids Res., 34 (Web Server issue): W362-365.

Derst C, Henseling J, Röhm KH. 2000. Engineering the substrate specificity of *Escherichia coli* asparaginase II. Selective reduction of glutaminase activity by amino acid replacements at position 248. Protein Science, 9: 2009-2017.

Duval M, Suciu S, Ferster A, Rialland X, Nelken B, Lutz P, Benoit Y, Robert A, Manel AM, Vilmer E, Otten J, Philippe N. 2002. Comparison of *Escherichia coli*-asparaginase with *Erwinia*-asparaginase in the treatment of childhood lymphoid malignancies: results of a randomized European Organization for Research and Treatment of Cancer-Children's Leukemia Group phase 3 trial. Blood, 99: 2734-2739.

Ferrara MA, Severino NMB, Mansure JJ, Martins AS, Oliveira EMM, Siani AC, Pereira N, Torres FAG, Bon EBS. 2006. Asparaginase production by a recombinant *Pichia pastoris* strain harbouring *Saccharomyces cerevisiae* ASP3 gene. Enzyme and Microbial Technology, 39: 1457-1463.

Friedman M, Levin E. 2008. Review of methods for the reduction of dietary content and toxicity of acrylamide. Journal of Agricultural and Food Chemistry, 56: 6113-6140.

Gasteiger E, Hoogland C, Gattiker A, Duvaud S, Wilkins MR, Appel RD, Bairoch A. 2005. In: John M. Walker (ed): Protein Identification and Analysis Tools on the ExPASy Server. The Proteomics Protocols Handbook, Humana Press. pp. 571-607.

Hall TA. 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucl. Acids. Symp., Ser. 41: 95-98.

Harms E, Wehner, A, Aung HP, Röhm KH. 1991. A catalytic role for threonine-12 of *E. coli* asparaginase II as established by site-directed mutagenesis. FEBS Lett., 285: 55-58.

Hüser A, Klöppner U, Röhm KH. 1999. Cloning, sequence analysis, and expression of ansB from *Pseudomonas fluorescens*, encoding periplasmic glutaminase/asparaginase. FEMS Microbiol. Lett., 178: 327-335.

Jennings MP, Beacham IR. 1990. Analysis of the *E. coli* gene encoding L-asparaginase II, ansB, and its regulation by cyclic AMP receptor and FNR proteins. J. Bacteriol., 172: 1491-1498.

Jia M, Xu M, He B, Rao Z. 2013. Cloning, expression, and characterization of L-asparaginase from a newly isolated *Bacillus subtilis* B11-06. J. Agric. Food Chem., 61: 9428-9434.

Jones DT. 1999. Protein secondary structure prediction based on position-specific scoring matrices. J. Mol. Biol., 292: 195-202.

Khushoo A, Pal Y, Singh BN, Mukherjee KJ. 2004. Extracellular expression and single step purification of recombinant *Escherichia coli* L-asparaginase II. Protein Expression and Purification, 38: 29-36.

Kotzia GA, Labrou NE. 2005. Cloning, expression and characterization of *Erwinia carotovora* L-asparaginase. J. Biotechnol., 119: 309-323.

Kotzia GA, Labrou NE. 2007. l-Asparaginase from *Erwinia chrysanthemi* 3937: Cloning, expression and characterization. J. Biotechnol., 127: 657-669.

Kozak M, Borek D, Janowski R, Jaskólski M. 2002. Crystallization and preliminary crystallographic studies of five crystal forms of *Escherichia coli* L-asparaginase II (Asp90Glu mutant). Acta Crystallogr. D. Biol. Crystallogr., 58: 130-132.

Kraxenberger T, Fried L, Behr S, Jung K. 2012. First insights into the unexplored two-component system YehU/YehT in *Escherichia coli*. J. Bacteriol., 194: 4272-4284.

Kukurová K, Ciesarová Z, Mogol BA, Açar ÖÇ, Gökmen V. 2013. Raising agents strongly influence acrylamide and HMF formation in cookies and conditions for asparaginase activity in dough. Eur. Food Res. Technol., 237: 1-8.

Marchler-Bauer A, Anderson JB, Chitsaz F, Derbyshire MK, DeWeese-Scott C, Fong JH, Geer LY, Geer RC, Gonzales NR, Gwadz M, He S, Hurwitz DI, Jackson JD, Ke Z, Lanczycki CJ, Liebert CA, Liu C, Lu F, Lu S, Marchler GH, Mullokandov M, Song JS, Tasneem A, Thanki N, Yamashita R A, Zhang D, Zhang N, Bryant SH. 2009. CDD: specific functional annotation with the conserved domain database. Nucleic Acids Res., 37: (D) 205-210.

Müller HJ, Boos J. 1998. Use of L-asparaginase in childhood ALL. Crit. Rev. Oncol. Hemat., 28: 97-113.

Narta UK, Kanwar SS, Azmi W. 2007. Pharmacological and clinical evaluation of L-asparaginase in the treatment of leukemia. Crit. Rev. Oncol/Hematol., 61: 208-221.

Palm GJ, Lubkowski J, Derst C, Schleper S, Röhm KH, Wlodawer A. 1996. A covalently bound catalytic intermediate in *Escherichia coli* asparaginase: crystal structure of a Thr-89-Val mutant. FEBS Lett., 390: 211-216.

Partridge JD, Browning DF, Xu M, Newnham LJ, Scott C, Roberts RE, Poole RK, Green J. 2008. Characterization of the *Escherichia coli* K-12 ydhYVWXUT operon: regulation by FNR, NarL and NarP. Microbiology, 154: 608-618.

Pedreschi F; Kaack K, Granby K. 2008. The effect of asparaginase on acrylamide formation in French fries. Food Chemistry, 109: 386-392.

Pieters R, Hunger SP, Boos J, Rizzari C, Silverman L, Baruchel A, Goekbuget N, Schrappe M, Pui CH. 2011. L-asparaginase treatment in acute lymphoblastic leukemia: a focus on *Erwinia asparaginase*. Cancer, 117: 238-249.

Pokrovskaya MV, Aleksandrova SS, Pokrovsky VS, Omeljanjuk NM, Borisova AA, Anisimova NY, Sokolov NN. 2012. Cloning, expression and characterization of the recombinant *Yersinia pseudotuberculosis* L-asparaginase. Protein Expression and Purification, 82: 150-154.

Quevillon E, Silventoinen V, Pillai S, Harte N, Mulder N, Apweiler R, Lopez R. 2005. InterProScan: protein domains identifier. Nucleic Acids Res., 33 (Web Server issue): W116-120.

Sambrook J, Russell DW. 2001. Molecular cloning: A laboratory manual. 3rd ed. Cold Spring Harbor (NY): Cold Spring Harbor Laboratory Press.

Shan Y, Pan Q, Liu J, Huang F, Sun H, Nishino K, Yan A. 2012. Covalently linking the *Escherichia coli* global anaerobic regulator FNR in tandem allows it to function as an oxygen stable dimer. Biochem. Biophys. Res. Commun., 419: 43-48.

Swain AL, Jaskólski M, Housset D, Rao JK, Wlodawer A. 1993. Crystal structure of *Escherichia coli* L-asparaginase, an enzyme used in cancer therapy. Proc. Natl. Acad. Sci. U S A, 90: 1474-1478.

Tolla DA, Savageau MA. 2011. Phenotypic repertoire of the FNR regulatory network in *Escherichia coli*. Mol. Microbiol., 79: 149-165.

Tong WH, van der Sluis IM, Alleman C, van Litsenburg RR, Kaspers GJ, Pieters R, Uyl-de Groot CA. 2013. Cost-analysis of treatment of childhood acute lymphoblastic leukemia with

DOI: 10.69085/jbb20153291

RESEARCH ARTICLE

asparaginase preparations: the impact of expensive chemotherapy. Haematologica, 98: 753-759.

Uppal S, Maurya SR, Hire RS, Jawali N. 2011. Cyclic AMP receptor protein regulates cspE, an early cold-inducible gene, in *Escherichia coli*. J. Bacteriol., 193: 6142-6151.

Vidya J; Vasudevan UM; Soccol CR, Pandey A. 2011. Cloning, functional expression and characterization of L-asparaginase II from *E. coli* MTCC 739. Food Technol. Biotechnol., 49: 286-290.

Vidya J, Pandey A. 2012. Recombinant expression and characterization of L-asparaginase II from a moderately thermotolerant bacterial isolate. Appl. Biochem. Biotechnol., 167: 973-980.

Wang Y ,Qian S, Meng G, Zhang S. 2001. Cloning and expression of L-asparaginase gene in *Escherichia coli*. Appl. Biochem. Biotechnol., 95: 93-101.

Wehner A, Derst C, Specht V, Aung HP, Röhm KH. 1994. The catalytic mechanism of *Escherichia coli* asparaginase II. Hoppe Seylers Z. Physiol. Chem., 375: 108.

Youssef MM, Al-Omair MA. 2008. Cloning, purification, characterization and immobilization of L-asparaginase II from *E. coli* W3110. Asian Journal of Biochemistry, 3: 337-350.

DOI: 10.69085/jbb20153291