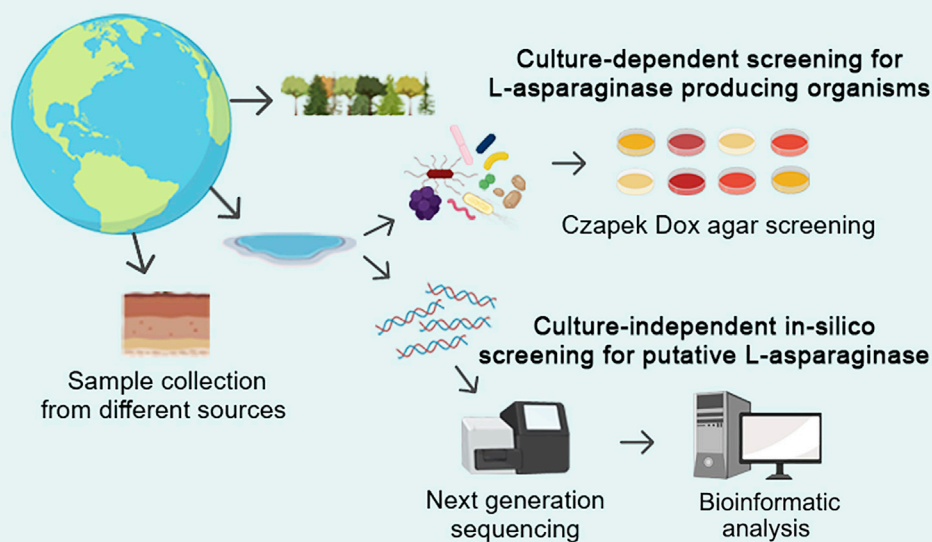## Article

Metagenomic discovery and functional validation of L-asparaginases with anti-leukemic effect from the Caspian Sea

Motahareh Sobat, Sedigheh Asad, Mahboubeh Kabiri, Maliheh Mehrshad

asad@ut.ac.ir

Highlights

In silico screening finds 87 putative L-asparaginases from brackish Caspian Sea

Three candidate genes, cloned and expressed in *E. coli*, showed the desired activity

In silico screening coupled with functional validation bypasses cultivation bottleneck

Article

Metagenomic discovery and functional validation of L-asparaginases with anti-leukemic effect from the Caspian Sea

Motahareh Sobat,[1] Sedigheh Asad,[1,3,*] Mahboubeh Kabiri,[1] and Maliheh Mehrshad[2]

Summary

By screening 27,000 publicly available prokaryotic genomes, we recovered ca. 6300 type I and ca. 5200 type II putative L-asparaginase highlighting the vast potential of prokaryotes. Caspian water with similar salt composition to the human serum was targeted for *in silico* L-asparaginase screening. We screened ca. three million predicted genes of its assembled metagenomes that resulted in annotation of 87 putative L-asparaginase genes. The L-asparagine hydrolysis was experimentally confirmed by synthesizing and cloning three selected genes in *E. coli*. Catalytic parameters of the purified enzymes were determined to be among the most desirable reported values. Two recombinant enzymes represented remarkable anti-proliferative activity (IC50 <1IU/ml) against leukemia cell line Jurkat while no cytotoxic effect on human erythrocytes or human umbilical vein endothelial cells was detected. Similar salinity and ionic concentration of the Caspian water to the human serum highlights the potential of secretory L-asparaginases recovered from these metagenomes as potential treatment agents.

Introduction

L-asparaginase (EC 3.5.1.1) is an amidohydrolase catalyzing the breakdown of L-asparagine into aspartic acid and ammonia. For over 30 years, L-asparaginase has been widely used as an antineoplastic agent for treatment of acute lymphoblastic leukemia (ALL) and non-Hodgkin lymphoma (Hill et al., 1967; Oettegen, 1967; Covini et al., 2012). Its mode of function is that it selectively kills leukemic cells by depleting patient's serum of L-asparagine, an amino acid that tumor cells essentially depend on for the malignant growth (Covini et al., 2012). The L-asparaginases from *Escherichia coli* and *Erwinia chrysanthemi* have been extensively used for medical purposes; however, arising complications such as hypersensitivity, antigenicity, short half-life, temporary blood clearance, and non-desirable L-glutaminase-dependent neurotoxicity necessitate continuous explorations seeking more suitable substitutes (Narta et al., 2007; Duval et al., 2002; El-Naggar et al., 2014).

Various strategies have been adopted with the common aim of achieving L-asparaginases of optimal characteristics for therapeutic purposes (Figure 1). While some researchers focus on sequence engineering to enhance enzyme's affinity or extinguish immunogenic epitopes (Derst et al., 2000; Cantor et al., 2012), others aim to protect the enzyme from the immune system by conjugating it to a variety of organic or inorganic compounds or using carriers for drug delivery (Asselin et al., 1993; Willer et al., 2011; Agrawal et al., 2013). Since L-asparaginase gene is distributed in all domains of life, various sources including bacteria, fungi, plants, and animals have been screened in hope of finding enzymes with superior features (Benchamin et al., 2019; Beulah and Hemalatha, 2019; Kumar et al., 2019). Prokaryotic L-asparaginase, specifically the secretory type, is involved in recycling the organically bound nitrogen through the ammonification process, thus playing a notable role in nitrogen biogeochemical cycling (Hill et al., 1967). The high metabolic diversity of prokaryotes renders them a promising resource for L-asparaginase screening campaigns. However, most screening efforts so far have only targeted cultivated prokaryotes (Ghasemi et al., 2017; Shirazian et al., 2016). One major limitation of this approach is due to cultivation bottleneck; the majority of prokaryotes (up to 99%) are still evading the bound of culture (Steen et al., 2019), consequently; their novel genetic contents remain inaccessible. Additionally, prokaryotes represent selective expression for different types of L-asparaginases under certain environmental condition (Herrero et al., 2019; Silpa et al., 2017), thus further limiting the culture/expression-dependent screening approaches. For example,

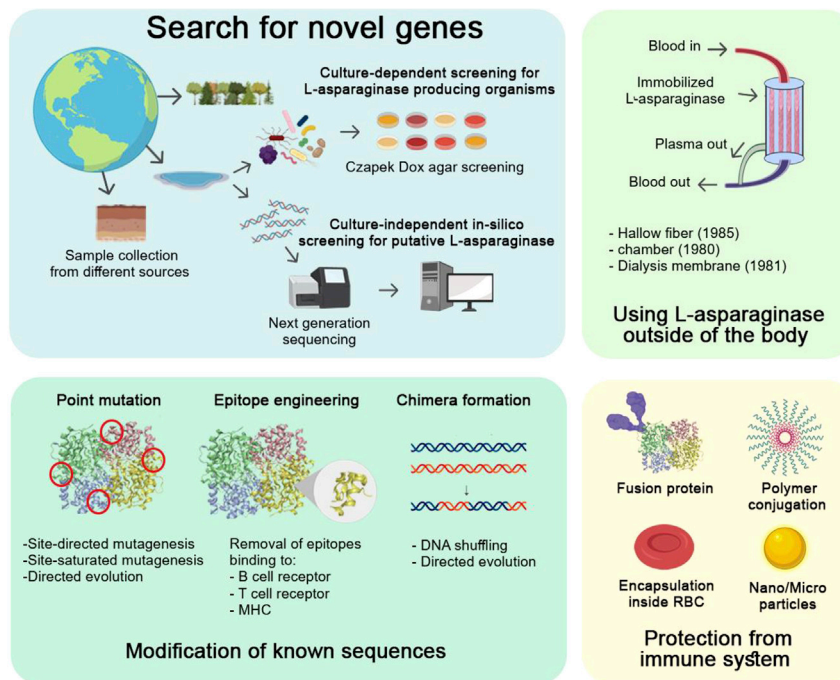[1]Department of Biotechnology, College of Science, University of Tehran, Tehran, Iran

[2]Department of Ecology and Genetics, Limnology and Science for Life Laboratory, Uppsala University, Uppsala, Sweden

[3]Lead contact

*Correspondence: asad@ut.ac.ir

Figure 1. Different approaches for achieving desirable L-asparaginases
Schematic overview of approaches used to tackle the immunogenicity and low activity of the L-asparaginase enzyme in respect to its use for acute lymphoblastic leukemia treatment.

model organism *E. coli*'s type I L-asparaginases is a cytoplasmic low-affinity protein continuously expressed and required for growth, whereas its type II enzyme is periplasmic with high affinity toward L-asparagine, expressed under anaerobic and starvation conditions acting as a scavenger to absorb nitrogen from the environment (Herrero et al., 2019).

One way to bypass the limitations caused by the culture/expression-dependent screening approaches is to explore prokaryotes' genomic potential through direct screening of sequenced environmental DNA, i.e., metagenomics (Madhavan et al., 2017; Wong et al., 2019; Arjun et al., 2018). In this study, we leverage the availability of metagenomic data sets derived from the brackish samples of the Caspian Sea for in silico screening of putative L-asparaginase genes and further verify their activity via cloning and expression analysis. As the salinity and osmolality of brackish Caspian water are akin to blood serum (130–145 and 132.44 mM sodium in the human serum and the Caspian Sea water, respectively, and 3.5–5.3 and 3.04 mM potassium in the human serum and the Caspian Sea water, respectively, [Worth, 1985; Mehrshad et al., 2016; Ghasemi et al., 2017]), it is expected that secretory proteins screened from this ecosystem would represent optimal activity in physiologic condition. Additionally, culture-independent in silico screening of metagenomes provides access to an untapped reservoir of unexplored genes already available in nature (Hess et al., 2011).

We screen the assembled metagenomes of the Caspian Sea depth profile for putative L-asparaginase genes. The enzymatic activity of three selected genes recovered through this screening was verified by cloning and expression of the synthetized gene in the *E. coli* host. The enzymes were purified to homogeneity. Kinetic and stability parameters and their anticancer activity were then studied.

Results and discussion

L-asparaginase is widely distributed in prokaryotes

Prokaryotic L-asparaginase is ecologically involved in transformation of organic nitrogen to release ammonia via the ammonification process (Herrero et al., 2019). Nitrogen rich compounds, peptides, and amino acids are among available nitrogen sources in aquatic habitats where type II secretory L-asparaginase can be utilized for scavenging nitrogen by a range of prokaryotes (Herrero et al., 2019; Silpa et al.,
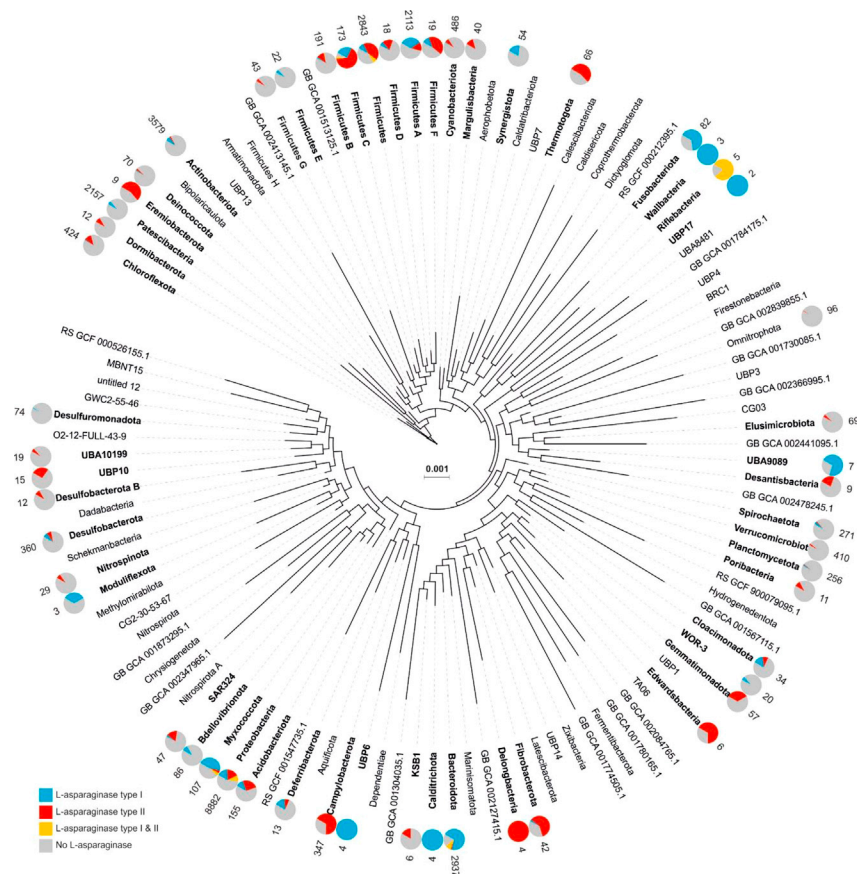
Figure 2. Analysis of bacterial L-asparaginases distribution
Phylogenetic distribution of L-asparaginase containing representatives across bacterial phyla. Phyla with representatives containing L-asparaginase genes are shown in bold. The pie charts show the distribution of genomes containing genes annotated as secretory or periplasmic L-asparaginase or those containing both L-asparaginase types simultaneously. The number next to the pie chart represents the number of genomes present in the phyla. See also Figure S1 for distribution of type II (TIGR00520) and type I (TIGR00519) L-asparaginase genes in different bacterial phyla. The list of genomes/MAGs screened for L-asparaginase is mentioned in Table S1.

2017). The secretory L-asparaginase acts as a public good, providing surrounding cells with nitrogen source (Smith and Schuster, 2019). To explore the phylogenomic distribution of L-asparaginase containing bacteria, we perform an extensive screening of 27,000 publicly available bacterial genomes/metagenome-assembled genomes (MAGs) via AnnoTree (Mendler et al., 2019) (list of genomes/MAG identifiers in the Table S1). Our results show a vast distribution of L-asparaginase genes in bacteria where 54 out of 112 bacterial phyla contain the genetic potential for L-asparaginase activity. Many of these bacterial phyla have no representative in culture and consequently remain inaccessible to culture/expression-dependent screening approaches (Figure 2). This reiterates the importance of cultivation-independent screening for finding new potential L-asparaginases as chemotherapy agents.

A total of 6329 bacterial genomes containing genes annotated as TIGR00519 (L-asparaginase, type I) were recovered. These genomes are affiliated to 36 different bacterial phyla. Five thousand two hundred ninety-four bacterial genomes affiliated to 42 different phyla contain genes annotated as TIGR00520 (L-asparaginase, type II) (Figure 2). Representatives of eight bacterial phyla contain both type I and type II L-asparaginase genes in their genomes (*Proteobacteria* 817, *Bacteroidota* 249, *Firmicutes* 175, *Firmicutes_A* 8, *Myxococcota* 5, *Firmicutes_C* 5, *Riflebacteria* 4, and *Firmicutes_B* 2). Phyla *Proteobacteria* and *Firmicutes* contain more representatives harboring the secretory L-asparaginase compared to the cytoplasmic one, whereas phylum *Bacteroidota* represents the opposite with higher frequency of representatives containing cytoplasmic L-asparaginase (Figure S1). These screened genomes/MAGs originate from different habitats

ranging from aquatic, soil, to host-associated microbes showing the vast distribution of this enzyme among bacterial representatives.

In silico screening of the Caspian Sea metagenomes results in 87 putative L-asparaginase coding genes

Screening novel environments by tapping into the vast metabolic reservoir of yet to be cultured prokaryotic majority can serve as a highly promising resource (Figures 1 and 2). Current therapeutic L-asparaginase enzymes originating from *E. coli* and *Erwinia chrysanthemi*, respectively, lose 40 and 80% of their activity in blood salinity (0.9%) (Werber et al., 1997; Ghasemi et al., 2017). Here, we target the brackish microbiome of the Caspian Sea for in silico metagenomic exploration of putative L-asparaginase. The brackish salinity of the Caspian Sea represents a highly similar salinity and main ionic concentration to the human serum (130–145 and 132.44 mM sodium in the human serum and the Caspian Sea water, respectively, and 3.5– 5.3 and 3.04 mM potassium in the human serum and the Caspian Sea water, respectively) (Worth, 1985; Mehrshad et al., 2016) (Table S2). Therefore, we hypothesize that secretory enzymes in the brackish microbiome of the Caspian See could potentially represent higher stability in physiologic conditions of the human serum, thus promising desirable therapeutic applications.

A total of 703171, 1169997, and 1214607 predicted open reading frames were screened for L-asparaginase activity, respectively, from 15, 40, and 150 m metagenomic data sets of the Caspian Sea. A total of 175, 296, and 284 putative L-asparaginase genes were annotated as TIGR00519 or TIGR00520 Hidden Markov Model (HMMs) from 15, 40, and 150m depth metagenomes, respectively. The putative annotation of these genes was further evaluated by inspecting the conserved regions and protein secondary structure (as explained in the Transparent Methods section). Amongst them, 87 genes were verified for L-asparaginase activity after *in silico* evaluations (18, 30, and 39, respectively, from 15, 40, and 150 m metagenomes). Recovered L-asparaginase genes had the length in the range of 245–489 amino acids (median 336 amino acids) with the highest sequence identity to representatives of bacterial taxa *Verrucomicrobiota*, *Gemmatimonadota*, *Alphaproteobacteria*, *Acidobacteriota*, *Chloroflexota*, *Bacteroidota*, *Gammaproteobacteria*, *Patescibacteria*, *Actinobacteriota*, *Cyanobacteria*, *Firmicutes_C*, and "*Candidatus* Rokubacteria" (n = 84), a single gene showing 75.7% identity to an archaeal L-asparaginase and two genes with 86.2 and 87% identity to single-cell eukaryotes affiliated to *Bathycoccus* and *Micromonas*, respectively (Table S3).

To assess the activity and other characteristics of recovered enzymes, we have selected three candidates to synthesis the enzyme coding sequence and experimentally verify their activity, Km, survival percentage at human serum salinity, and cellular analysis. Sequences of secretory enzymes (Table S4) related to fastidious bacterial taxa were preferably selected for cloning and enzyme activity verification. Additionally, bacteria spotted in all three depths of 15, 40, and 150 meter were assumed to be more desirable as are expected to be adapted to sustain wide range of environmental conditions. Three genes were selected: CAspI from 40m metagenomics data set and affiliated to phylum *Acidobacteriota* (79.9% sequence identity), CAspII also from 40m metagenomic data set and affiliated to phylum *Gemmatimonadota* (51.8% sequence identity), and CAspIII from 15m metagenomics data set affiliated to phylum "*Candidatus* Rokubacteria" (50.7% sequence identity) based on their best Basic Local Alignment Search Tool (BLAST) hit against National Center for Biotechnology Information- Non Redundant protein sequences (NCBI NR) database (Table S3). Phylogenetic reconstruction of L-asparaginase genes annotated from the Caspian Sea metagenomes together with the secretory L-asparaginase genes recovered from screening bacterial genomes through AnnoTree is shown in Figure 3. Highlighting the enzyme phylogeny based on genome taxonomy brings forward the possibility of several horizontal transfer events for the type II L-asparaginase gene between different taxa.

Selected L-asparaginase genes have low BLAST similarities with the most used therapeutic enzymes

Pairwise nucleotide sequence alignment was performed between *E. coli* and *Erwinia chrystanthemi* L-asparaginases and selected genes from the Caspian Sea metagenomes. The maximum similarity was between CAspI and *Erwinia* (49.2% similarity, 34.3% identity) which is lower than the similarity between *E. coli* type II and *Erwinia* L-asparaginase (62.8% similarity, 42.3% identity). *Erwinia* L-asparaginase is used as the second line of chemotherapy in cases that show intense immune reaction against *E. coli* L-asparaginase. As antibodies against *E. coli* L-asparaginase do not cross-react with *Erwinia* considering their
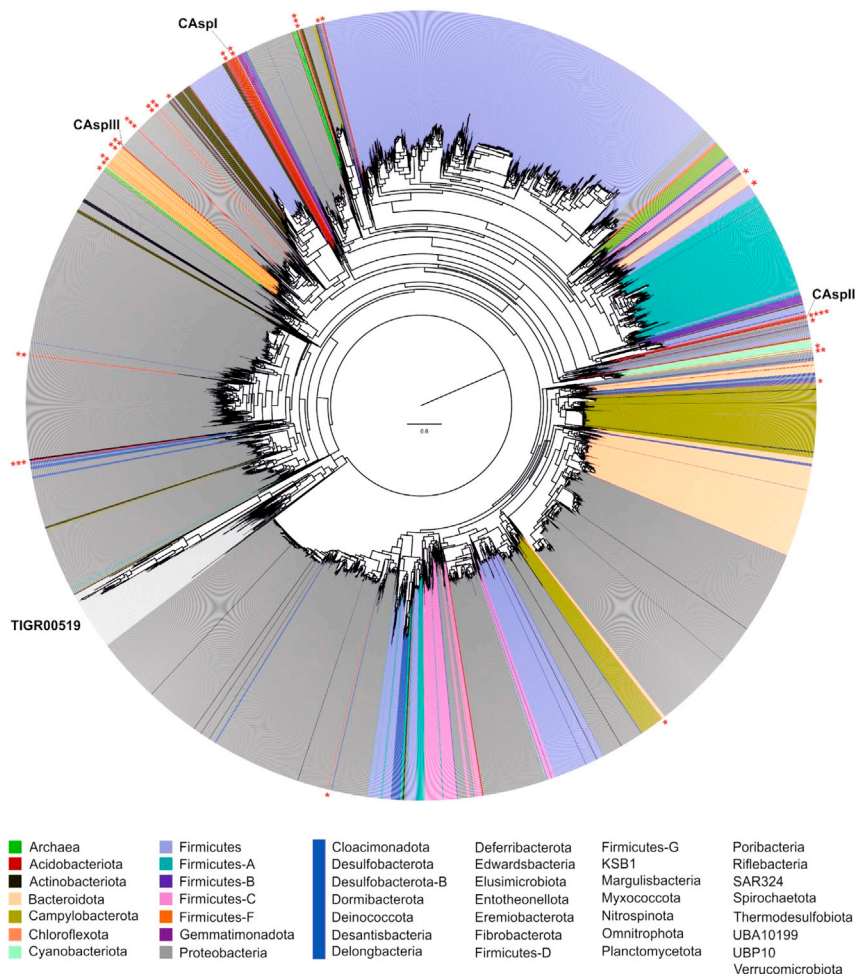
Figure 3. Phylogenetic analysis of L-asparaginases predicted in Caspian Sea metagenomes

Phylogeny of the secretory (TIGR00520) L-asparaginase protein sequences recovered from Caspian Sea metagenomes together with L-asparaginase genes recovered from bacterial genomes annotated in the AnnoTree database together with the reference L-asparaginase protein sequences recovered from UniProt. The phylogenetic position of the Caspian Sea recovered L-asparaginases is shown with asterisk. The periplasmic L-asparaginase (TIGR00519) sequences were used to root the tree. See also Figure S4 for workflow of bioinformatic analysis pipeline for screening.

sequence similarity of 62.8%, we estimated that our genes might be safe against *E. coli* and *Erwinia* antibodies based on their lower sequence similarity (Table 1).

Three L-asparaginases screened from the Caspian Sea metagenomes named CAspI, CAspII, and CAspIII (1092, 1218, and 1011 bp, respectively) were codon optimized for expression in *E. coli*, synthesized and cloned in pET21a (+) vector in fusion with hexahistidine tag at the C-terminus (Table S5).

Recombinant L-asparaginases were purified to homogeneity after expression optimization

Subsequent to plasmid transformation and confirmation by re-sequencing, enzyme expression and activity were examined. Successfully all three recombinant enzymes showed L-asparaginase activity while the specific activity of the crude enzyme from untransformed *E. coli* was negligible, indicating that our computational metagenome-wide screening method is a promising approach for gene screening. After culture condition optimization, maximum specific activity of crude extract was achieved by induction of mid-log phase transformed cells for three hours at 30°C using Luria-Bertani (LB) as culture medium. Optimum Isopropyl β-d-1-thiogalactopyranoside (IPTG) concentration was found 0.3 mM for CAspI and 0.2 mM for CAspII and CAspIII (Figure S2).

Table 1. Pairwise nucleotide alignment results of the selected L-asparaginases from the enzymes from *E. coli* and *E. chrysanthemi*

| Pairwise sequence alignment (%)	*E. coli* type I	*E. coli* type II	*E. chrysanthemi*
CAspI	S: 36.5 I: 21.9	S: 47.0 I: 31.7	S: 49.2 I: 34.3
CAspII	S: 32.5 I: 19.7	S: 46.5 I: 31.1	S: 47.4 I: 34.0
CAspIII	S: 35.6 I: 22.5	S: 48.3 I: 32.6	S: 49.9 I: 32.3
E. chrysanthemi	S: 35.6 I: 19.0	S: 62.8 I: 47.3	
E. coli type II	S: 36.6 I: 23.6		

See Tables S3 and S4 for more detailed BLAST analyses of all 87 candidate genes. Also, see Table S5 for statistics of selected L-asparaginase coding genes.

Recombinant CAspI, CAspII, and CAspIII were purified to homogeneity by nickel-agarose affinity chromatography. The protein preparations were enriched and not completely pure before making the measurements (Figure S3). Molecular mass of the monomeric enzymes using their amino acid sequences was predicted to be 38.2, 43.3, and 35.6 kDa, respectively, which was consistent with protein migration on Sodium Dodecyl Sulfate PolyAcrylamide Gel Electrophoresis (SDS-PAGE). The molecular weight of L-asparaginase protein varies according to the enzyme source, i.e., for bacterial L-asparaginases, the molecular weight according to SDS-PAGE analysis is usually in the range of 35–40 kDa (Aghaiypour et al., 2001; Kozak and Jurga, 2002; Prakasham et al., 2010). However, there are reported cases of very low molecular weight of 11.2 kDa for *Streptobacillus* sp. KK2S4 (Makky et al., 2014) and the high molecular weight (near 97.4 kDa) of purified L-asparaginase from *Streptomyces tendae* (Kavitha and Vijayalakshmi, 2010). The specific activities of the purified enzymes were 700, 240, and 100 U/mg, respectively, for CAspI, CAspII, and CAspIII while the specific activity of the commercial L-asparaginases from *E. coli* and *E. chrysanthemi* is between 280 and 400 U/mg and 650–700 U/mg, respectively (Narta et al., 2007). Phetsri et al. reported maximum specific activity of 113 U/mg for *Streptococcus thermophiles* among four species of lactic acid bacteria tested (Phetsri et al., 2019). Specific activities of 833 and 155 U/mg are also stated for the L-asparaginases purified from *Thermococcus kodakarensis* and *Acinetobacter soli*, respectively (Jiao et al., 2020; Chohan et al., 2020).

Kinetic parameters of recombinant L-asparaginases were among the most desirable reported values of microbial L-asparaginases

Kinetic parameters of the enzymes were calculated according to the classical Michaelis-Menten equation using L-asparagine as a substrate (Figure 4 and Table 2). Km values of 10, 0.35, and 0.15 mM were achieved for CAspI, CAspII, and CAspIII, respectively. Enzyme affinity toward its substrate is reflected by the value of the Km. The lower the Km value, the better the binding ability of the enzyme. While the Km value of all three enzymes is higher than that of *E. coli* type II and *E. chrysanthemi* L-asparaginases (Narta et al., 2007), our data suggest that CAspIII Km is among the lowest reported L-asparaginase Km values to date.

The Km value of CAspIII is lower than that of the L-asparaginases produced by *Streptomyces fradiae* NEAE-82, *Halomonas elongata*, and *Enterobacter cloacae* (El-Naggar et al., 2016; Ghasemi et al., 2017; Husain et al., 2016). This gene was recovered from a 1161 bp long contig assembled from the 15m depth metagenomes of the Caspian Sea showing the highest protein sequence identity (50.7%) to the reconstructed MAG ''*Candidatus* Rokubacteria'' bacterium AR30 (The BioSample accession number SAMN08911936). This MAG was reconstructed from meadow soil samples at 30-40cm depth, Angelo Coast Range Reserve, CA, USA (The BioSample accession number SAMN08902845). Representatives of this candidate phylum so far have evaded the bound of culture and remained inaccessible to the culture/expression-screening campaigns (Hug et al., 2016; Diamond et al., 2019; Becraft et al., 2017). Representatives of this *Candidatus* phylum are shown to be involved in biogeochemical cycling of elements in the soil (Hug et al., 2016; Becraft et al., 2017; Diamond et al., 2019) and harbor a vast potential for secondary metabolites biosynthesis (Crits-Christoph et al., 2018). While the CAspIII is only distantly related to the L-asparaginase gene of the ''*Candidatus* Rokubacteria'' bacterium AR30, it most probably belongs to the
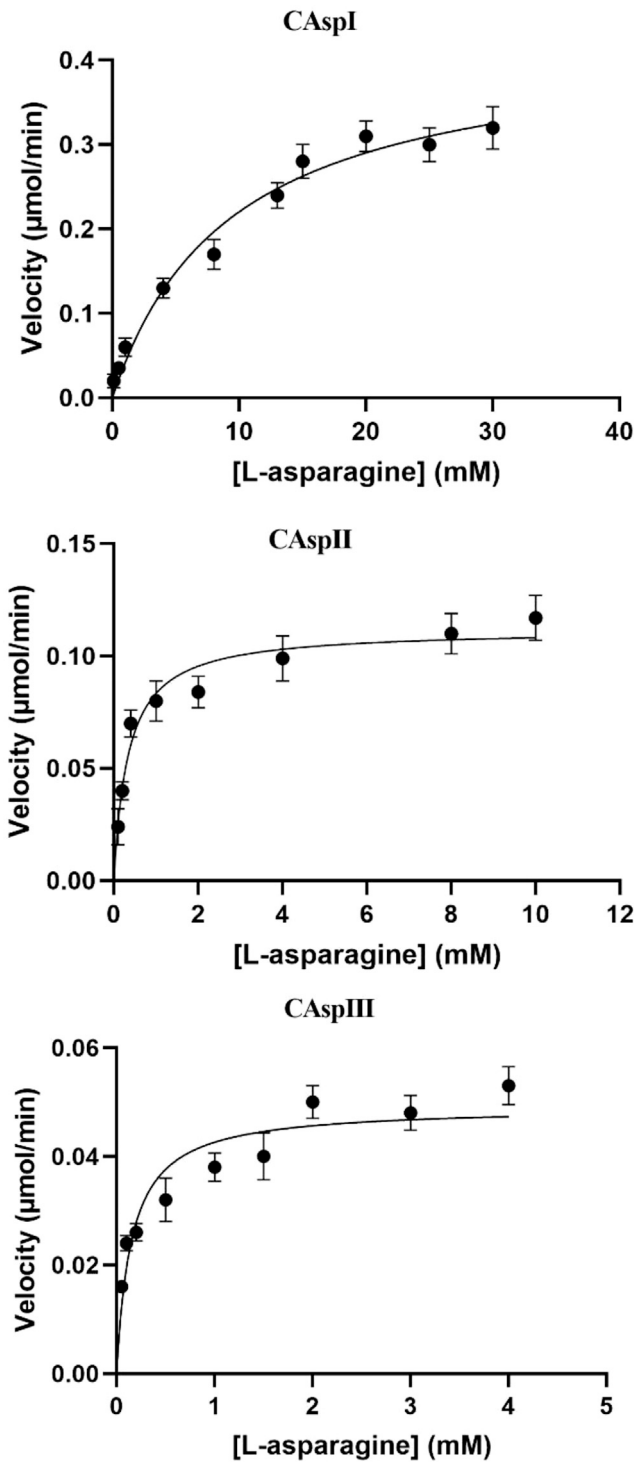
Figure 4. Enzyme kinetic characterization

Michaelis-Menten kinetics of recombinant L-asparaginases toward L-asparagine as a substrate. See Figure S2 to find the results of the culture condition optimization for each enzyme and also Figure S3 for SDS-PAGE analysis of expression and purification steps. Data are represented as mean $\pm$ Standard Error of the Mean (SEM).

Table 2. Biochemical characteristics of the recombinant L-asparaginases

L-asparaginase	$kcat$ (s^{-1})	Km (mM)	$kcat/Km$ (mM^{-1}s^{-1})	$Vmax$ (μmol min^{-1})
CAspI	446 ± 20	10 ± 1	44.6	0.35 ± 0.02
CAspII	174 ± 8	0.35 ± 0.02	497.14	0.12 ± 0.01
CAspIII	59.3 ± 4	0.15 ± 0.02	395.4	0.05 ± 0.004

Data are represented as mean ± SEM.

rare and fastidious microbiome of the Caspian Sea; however, the taxonomic affiliation cannot be assured only based on the BLAST identity. Additionally, the representatives of the "*Candidatus* Rokubacteria" are ubiquitous in a diverse range of terrestrial ecosystems and subsurface habitats with no reported marine representatives (Becraft et al., 2017). CAspII and CAspIII also revealed catalytic efficiencies of approximately 10-fold higher than CAspI suggesting that they could metabolize asparagine more efficiently (Beckett and Gervais, 2019).

Enzymes showed optimal activity at physiologic pH and salinity

The activity of the purified enzymes was studied in the pH range of 5–8 (Figure 5A). All recombinant enzymes displayed maximal L-asparaginase activity at pH 7.5 that is favorable feature of these enzymes where maintaining optimal enzymatic activity at the physiological pH is one of the perquisites for antitumor activity (Siddalingeshwara and Lingappa, 2011). A sharp decrease in activity was observed at more acidic or basic pH in the case of CAspII. CAspI activity was increased gradually up to pH 7 and retained its maximum activity up to pH 8. CAspIII showed maximum activity at pH 6.5–7.5. The amidases enzymes such as L-asparaginases are mostly active and stable at neutral and alkaline pH ranges of 5–9 (Ohshima et al., 1976). L-asparaginase, purified from alkaliphilic *Streptomyces fradiae* NEAE-82, exhibited maximum activity at pH 8.5 (El-Naggar et al., 2016). The optimal L-asparaginase activity from *Halomonas elongata* was reported to be at pH 6–9 (Ghasemi et al., 2017). Maximum activity at pH 8 was obtained for purified L-asparaginase from *Pyrococcus furiosus* (Saeed et al., 2020).

Effect of NaCl concentration on the enzyme activity is shown in Figure 5B. As expected, it can be observed that 140 mM NaCl (equal to physiologic 0.9% saline solution) had no adverse effect on enzymatic activity of the recombinant L-asparaginases; furthermore, it is consistent with blood Na$^+$ concentration that ranges between 130 and 145 mM (Worth, 1985). It should be noted that both commercial L-asparaginases and most reported enzymes show decreased activity when subjected to physiological salinity (Werber et al., 1997). This reiterates the adaptive advantage of secretory enzymes recovered from the brackish waters of the Caspian Sea to retain their activity in the salinity of human serum due to similar ionic concentration. L-asparaginase isolated from the halophilic *H. elongata* also retained its maximum activity at physiologic salinity (Ghasemi et al., 2017).

Anti-leukemic assessment on Jurkat cell line demonstrated CAspII and CAspIII as potential chemotherapeutic agents

L-asparaginases have been isolated from various sources but all do not have cytotoxic effects on cancerous cells. The cytotoxicity of recombinant L-asparaginases was examined on human lymphoblastic leukemia
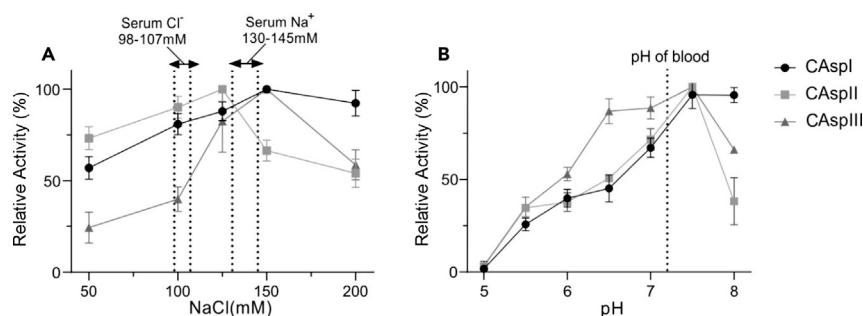
Figure 5. Optimum condition for enzyme activity

Effect of pH (A) and NaCl (B) concentration on enzyme activity. Dashed lines show the blood serum conditions. Data are represented as mean ± Standard Error of the Mean (SEM).
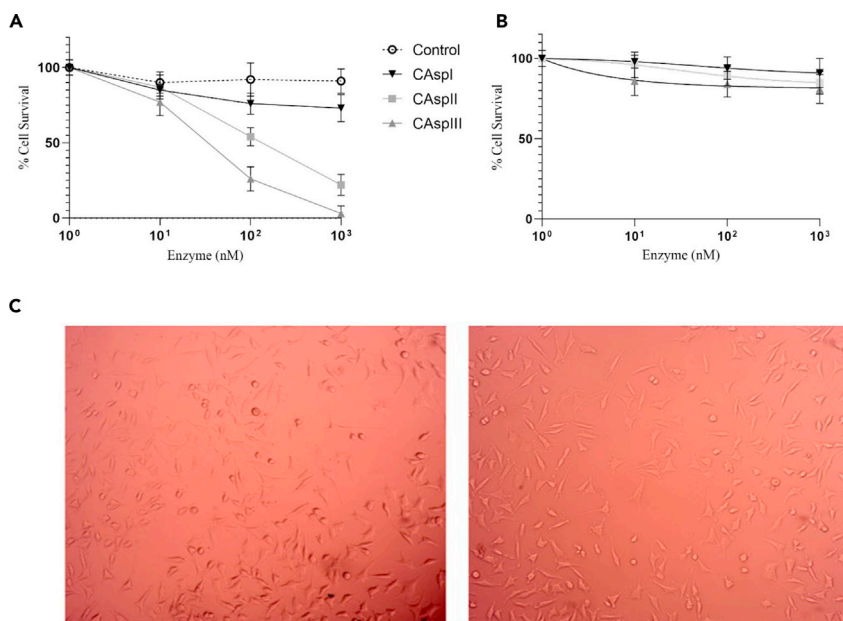
Figure 6. Anti-proliferative and cytotoxic activity assessment

Effect of purified L-asparaginases on Jurkat (A) and HUVEC (B) cell lines (data are represented as mean ± Standard Error of the Mean),

(C) Enzymes effect on HUVEC cell line morphology. The cells were incubated for 24 hr with purified enzyme (right), and cell morphology was compared with untreated cells (left).

cell line, Jurkat (widely used cell lines in L-asparaginase cytotoxicity studies), by Methyl-Thiazolyl-Tetrazolium (MTT) assays (Figure 6A). After 24 hr of incubation, CAspIII and II proved to be highly effective against the leukemic cell line with IC50 of 120 and 33 nM which are equal to 0.6 and 0.06 IU/mL of the enzymes, respectively. However, the commercial L-asparaginase from *E. coli* has IC50 of 1.0 IU/mL and that of *Erwinia* has been reported to have IC50 of 7.5–10.0 IU/mL (Abakumova et al., 2012). These results indicate that the purified recombinant CAspII and CAspIII could be considered as effective chemotherapeutic agents in killing human leukemic cell line "Jurkat" primarily due to depletion of the asparagine pool. Although asparagine is a nonessential amino acid, it is vital for some leukemia and cancer cells for two reasons. Firstly, asparagine is required for the synthesis of glycoproteins and other cellular proteins, and secondly, these cells have low expression levels of L-asparagine synthetase for *de novo* synthesis of asparagine (Saeed et al., 2020). However, further studies, especially *in vivo* tests, are needed to justify this claim. The IC50 value for CAspI was higher than 1000 nM (13 IU/mL), the highest concentration used in the dose-response curve. As seen in Figure 6A, no tested concentration of CAspI could reduce cell viability to less than 60%; therefore, the IC50 value is approximated (calculated based on the non-linear regression model).

Enzymes revealed no cytotoxicity on human erythrocytes and Human Umbilical Vein Endothelial Cells (HUVEC)

As a chemotherapeutic agent, asparaginase is routinely administered intravenously and thus it would come into contact with both leukemic and non-cancerous or non-blood cell types. As such, any probable anti-proliferative or cytotoxic effect of enzyme on both target and other cell types should be carefully assessed. Thus, we inspected the cytotoxicity of the purified enzymes with non-leukemic, non-myeloid cell line (HUVEC), along with the dominant cells of the blood, erythrocytes, to monitor any possible side effects of the enzymes on other cell types. No detectable adverse effect was observed for HUVEC cells (Figures 6B and 6C). Additionally, no sign of erythrocyte hydrolysis was observed for any of the *in vitro* hemolysis test experiments conducted with the CAspI, II, and III enzymes (data not shown).

L-asparaginase has been widely used as antineoplastic agent to deplete patient's serum from L-asparagin, essentially needed by auxotrophic cancer cells. Its therapeutic potential for ALL and non-Hodgkin lymphoma has fueled the search for novel L-asparaginases to combat the complications associated with the existing

therapeutic L-asparaginases. Harnessing the vast potential of prokaryotes via culture-independent methods is a promising approach for finding novel enzymes. Prokaryotes express L-asparaginase as a common good to scavenge for nitrogen sources. Consequently, screening the oligotrophic waters of the Caspian Sea, mimicking the salinity and ionic concentration of blood serum can potentially lead to finding enzymes that can attenuate the complications with hypersensitivity, antigenicity, short half-life, and temporary blood clearance. Recognizing that cancer has the properties of ecological systems have been suggested to be the way forward in devising sustainable treatments (Pienta et al., 2008). Leveraging the knowledge of ecology and tapping into the genetic potential of ecosystems similar to cancer ecosystem can facilitate the biodiscovery of treatment agents.

Limitations of the study

To select putative L-asparaginase genes for synthesis, all the screened genes were thoroughly analyzed regarding their active sites and secondary structure to make sure of their function, and we were left with 87 candidates. For these candidates, we had proof of L-asparaginase function based on all in silico methods tested. Owing to funding constraints, only 3 candidate genes were synthesized and evaluated in this study. The other designed genes will be studied in later projects. Although the cytotoxicity of recombinant L-asparaginases was examined on human lymphoblastic leukemia cell line, Jurkat (the main cell lines for L-asparaginase cytotoxicity assessment), the enzymes efficacy on other leukemia cell lines and also *in vivo* test is necessary to assure their anti-leukemic effect. These are important questions that need to be further explored in future studies.

Resource availability

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Sedigheh Asad (asad@ut.ac.ir).

Material availability

All reagents generated in this study are available from the Lead Contact with a completed Materials Transfer Agreement.

Data and code availability

The published article includes all data sets generated during this study. The putative L-asparaginase gene sequences together with the *in vitro* examined codon-optimized genes are provided in the Tables S3 and S5.

Methods

All methods can be found in the accompanying Transparent Methods supplemental file.

Supplemental Information

Supplemental Information can be found online at https://doi.org/10.1016/j.isci.2020.101973.

Author contributions

S.A. and M.M. devised the research. M.S. and M.M. performed the bioinformatics analysis. M.S., M.K., and S.A. carried out the experiments. M.S., M. M., and S. A. analyzed the results and drafted the manuscript. All authors read and approved the manuscript.

Declaration of interests

The authors declare no competing interests.

References

Abakumova, O.Y., Podobed, O., Karalkin, P., Kondakova, L., and Sokolov, N. (2012). Antitumor activity of L-asparaginase from *Erwinia carotovora* against different human and animal leukemic and solid tumor cell lines. Biochem. (Mosc.), Suppl. Ser. B Biomed. Chem. *6*, 307–316.

Aghaiypour, K., Wlodawer, A., and Lubkowski, J. (2001). Do bacterial l-asparaginases utilize a catalytic triad Thr-Tyr-Glu? *BBA-Protein*. Struct. M. *1550*, 117–128.

Agrawal, V., Hee Woo, J., Borthakur, G., Kantarjian, H., and E Frankel, A. (2013). Red blood cell-encapsulated L-asparaginase: potential therapy of patients with asparagine synthetase deficient acute myeloid leukemia. Protein Pept. Lett. *20*, 392–402.

Arjun, J.K., Aneesh, B.P., Kavitha, T., and Harikrishnan, K. (2018). Characterization of a novel asparaginase from soil metagenomic libraries generated from forest soil. Biotechnol. Lett. *40*, 343–348.

Asselin, B.L., Whitin, J.C., Coppola, D.J., Rupp, I.P., Sallan, S.E., and Cohen, H.J. (1993). Comparative pharmacokinetic studies of three asparaginase preparations. J. Clin. Oncol. *11*, 1780–1786.

Beckett, A., and Gervais, D. (2019). What makes a good new therapeutic L-asparaginase? World. J. Microb. Biot. *35*, 152.

Becraft, E.D., Woyke, T., Jarett, J., Ivanova, N., Godoy-Vitorino, F., Poulton, N., Brown, J.M., Brown, J., Lau, M.C.Y., Onstott, T., et al. (2017). Rokubacteria: genomic Giants among the uncultured bacterial phyla. Front. Microbiol. *8*, 2264.

Benchamin, D., Sreejai, R., Jensy Roshan, F., Sujitha, S., Kurup, B.S., and Albert, C. (2019). Morphological and Molecular Identification of L-Asparaginase Producing Fungus from Annona Muricata. Pharma. Innovation. *8*, 18–20.

Beulah, K., and Hemalatha, K. (2019). Screening of medicinal plants for potential source of l-asparaginase and optimization of conditions for maximum extraction and assay of l-asparaginase from *Asparagus racemosus*. Iran. J. Sci. Technol. A *43*, 1–6.

Cantor, J.R., Panayiotou, V., Agnello, G., Georgiou, G., and Stone, E.M. (2012). Engineering reduced-immunogenicity enzymes for amino acid depletion therapy in cancer. Methods Enzymol. *502*, 291–319.

Chohan, S.M., Sajed, M., Un Naeem, S., and Rashid, N. (2020). Heterologous gene expression and characterization of TK2246, a highly active and thermostable plant type l-asparaginase from *Thermococcus kodakarensis*. Int. J. Biol. Macromol. *147*, 131–137.

Covini, D., Tardito, S., Bussolati, O., R Chiarelli, L., V Pasquetto, M., Digilio, R., Valentini, G., and Scotti, C. (2012). Expanding targets for a metabolic therapy of cancer: L-asparaginase. Recent Pat. Anti Cancer Drug Discov. *7*, 4–13.

Crits-Christoph, A., Diamond, S., Butterfield, C.N., Thomas, B.C., and Banfield, J.F. (2018). Novel soil bacteria possess diverse genes for secondary metabolite biosynthesis. Nature *558*, 440–444.

Derst, C., Henseling, J., and Röhm, K.-H. (2000). Engineering the substrate specificity of *Escherichia coli* asparaginase II. Selective reduction of glutaminase activity by amino acid replacements at position 248. Protein Sci. *9*, 2009–2017.

Diamond, S., Andeer, P.F., Li, Z., Crits-Christoph, A., Burstein, D., Anantharaman, K., Lane, K.R., Thomas, B.C., Pan, C., and Northen, T.R. (2019). Mediterranean grassland soil C–N compound turnover is dependent on rainfall and depth, and is mediated by genomically divergent microorganisms. Nat. Microbiol. *4*, 1356–1367.

Duval, M., Suciu, S., Ferster, A., Rialland, X., Nelken, B., Lutz, P., Benoit, Y., Robert, A., Manel, A.-M., and Vilmer, E. (2002). 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 J. Hematol. *99*, 2734–2739.

El-Naggar, N.E.-A., Deraz, S.F., Soliman, H.M., El-Deeb, N.M., and El-Ewasy, S.M. (2016). Purification, characterization, cytotoxicity and anticancer activities of L-asparaginase, anti-colon cancer protein, from the newly isolated alkaliphilic *Streptomyces fradiae* NEAE-82. Sci. Rep. *6*, 32926.

El-Naggar, N.E.-A., El-Ewasy, S.M., and El-Shweihy, N.M. (2014). Of acute lymphoblastic leukemia: the pros and cons. Int. J. Pharmacol. *10*, 182–199.

Ghasemi, A., Asad, S., Kabiri, M., and Dabirmanesh, B. (2017). Cloning and characterization of *Halomonas elongata* L-asparaginase, a promising chemotherapeutic agent. Appl. Microbiol. Biot. *101*, 7227–7238.

Herrero, A., Flores, E., and Imperial Ródenas, J. (2019). Nitrogen Assimilation in Bacteria. In Encyclopedia of Microbiology, 4th ed. V3 (Academic Press).

Hess, M., Sczyrba, A., Egan, R., Kim, T.-W., Chokhawala, H., Schroth, G., Luo, S., Clark, D.S., Chen, F., and Zhang, T. (2011). Metagenomic discovery of biomass-degrading genes and genomes from cow rumen. Science *331*, 463–467.

Hill, J.M., Roberts, J., Loeb, E., Khan, A., Maclellan, A., and Hill, R.W. (1967). L-asparaginase therapy for leukemia and other malignant neoplasms: remission in human leukemia. JAMA *202*, 882–888.

Hug, L.A., Thomas, B.C., Sharon, I., Brown, C.T., Sharma, R., Hettich, R.L., Wilkins, M.J., Williams, K.H., Singh, A., and Banfield, J.F. (2016). Critical biogeochemical functions in the subsurface are associated with bacteria from new phyla and little studied lineages. Environ. Microbiol. *18*, 159–173.

Husain, I., Sharma, A., Kumar, S., and Malik, F. (2016). Purification and characterization of glutaminase free asparaginase from *Enterobacter cloacae*: in-vitro evaluation of cytotoxic potential against human myeloid leukemia HL-60 cells. PLoS. One *11*, e0148877.

Jiao, L., Chi, H., Lu, Z., Zhang, C., Chia, S.R., Show, P.L., Tao, Y., and Lu, F. (2020). Characterization of a novel type I l-asparaginase from *Acinetobacter soli* and its ability to inhibit acrylamide formation in potato chips. J. Biosci. Bioeng. *129*, 672–678.

Kavitha, A., and Vijayalakshmi, M. (2010). Optimization and purifi cation of L-asparaginase produced by *Streptomyces tendae* TK-VL_333. Z. Naturforsch. C. *65*, 528–531.

Kozak, M., and Jurga, S. (2002). A comparison between the crystal and solution structures of *Escherichia coli* asparaginase II. Acta Biochim. Pol. *49*, 509–513.

Kumar, M., Muralidharan, A., Pai, A., Subrahmanyam, V., and Kamath, V. (2019). L-asparaginase from soil actinomycetes: optimization of process variables using plackett burman method. RJPT *12*, 3487–3490.

Madhavan, A., Sindhu, R., Parameswaran, B., Sukumaran, R.K., and Pandey, A. (2017). Metagenome analysis: a powerful tool for enzyme bioprospecting. Appl. Biochem. Biotech. *183*, 636–651.

Makky, E.A., Loh, Y.C., and Karim, M.R. (2014). Purification and partial characterization of a low molecular weight L-asparaginase produced from corn cob waste. Biocatal. Agric. Biotechnol. *3*, 265–270.

Mehrshad, M., Amoozegar, M.A., Ghai, R., Shahzadeh Fazeli, S.A., and Rodriguez-Valera, F. (2016). Genome reconstruction from metagenomic data sets reveals novel microbes in the brackish waters of the Caspian Sea. Appl. Environ. Microb. *82*, 1599–1612.

Mendler, K., Chen, H., Parks, D.H., Lobb, B., Hug, L.A., and Doxey, A.C. (2019). AnnoTree: visualization and exploration of a functionally annotated microbial tree of life. Nucleic Acids Res. *47*, 4442–4448.

Narta, U.K., Kanwar, S.S., and Azmi, W. (2007). Pharmacological and clinical evaluation of L-

asparaginase in the treatment of leukemia. Crit. Rev. Oncol. Hemat. *61*, 208–221.

Oettegen, H. (1967). Inhibition of leukemia in man by L-asparaginase. Cancer Res. *27*, 2619–2626.

Ohshima, M., Yamamoto, T., and Soda, K. (1976). Further characterization of glutaminase isozymes from *Pseudomonas aeruginosa.* Agric. Biol. Chem. *40*, 2251–2256.

Phetsri, K., Furukawa, M., Yamashiro, R., Kawamura, Y., Hayashi, J., Tobe, R., Toyotake, Y., and Wakayama, M. (2019). Comparative biochemical characterization of L-asparaginases from four species of lactic acid bacteria. J. Biomed. Biotechnol. *2*, 112–124.

Pienta, K.J., Mcgregor, N., Axelrod, R., and Axelrod, D.E. (2008). Ecological therapy for cancer: defining tumors using an ecosystem paradigm suggests new opportunities for novel cancer treatments. Transl. Oncol. *1*, 158.

Prakasham, R., Hymavathi, M., Rao, C.S., Arepalli, S., Rao, J.V., Kennady, P.K., Nasaruddin, K., Vijayakumar, J., and Sarma, P. (2010). Evaluation of antineoplastic activity of extracellular asparaginase produced by isolated *Bacillus circulans.* Appl. Biochem. Biotech. *160*, 72.

Saeed, H., Hemida, A., El-Nikhely, N., Abdel-Fattah, M., Shalaby, M., Hussein, A., Eldoksh, A., Ataya, F., Aly, N., and Labrou, N. (2020). Highly efficient *Pyrococcus furiosus* recombinant L-asparaginase with no glutaminase activity: expression, purification, functional characterization, and cytotoxicity on THP-1, A549 and Caco-2 cell lines. Int. J. Biol. Macromol. *156*, 812–828.

Shirazian, P., Asad, S., and Amoozegar, M.A. (2016). The potential of halophilic and halotolerant bacteria for the production of antineoplastic enzymes: L-asparaginase and L-glutaminase. EXCLI J. *15*, 268.

Siddalingeshwara, K., and Lingappa, K. (2011). Production and characterization of L-asparaginase-a tumour inhibitor. Int. J. Pharm. Tech. Res. *3*, 314–319.

Silpa, S., Bhattacharya, S., and Venkatanagaraju, E. (2017). Overview on L-asparaginase. World J. Pharm. Pharm. Sci. *6*, 561–601.

Smith, P., and Schuster, M. (2019). Public goods and cheating in microbes. Curr. Biol. *29*, R442–R447.

Steen, A.D., Crits-Christoph, A., Carini, P., Deangelis, K.M., Fierer, N., Lloyd, K.G., and Thrash, J.C. (2019). High proportions of bacteria and archaea across most biomes remain uncultured. ISME J. *13*, 3126–3130.

Werber, G., Ahlke, E., Nowak-Göttl, U., Jürgens, H., Verspohl, E., and Boos, J. (1997). Asparaginase activities in vitro are highly sensitive to different buffer conditions. In Acute Leukemias VI, *38*Acute Leukemias VI (Springer), pp. 512–516.

Willer, A., Gerß, J., König, T., Franke, D., Kühnel, H.-J., Henze, G., Von Stackelberg, A., Möricke, A., Schrappe, M., and Boos, J. (2011). Anti–*Escherichia coli* asparaginase antibody levels determine the activity of second-line treatment with pegylated *E coli* asparaginase: a retrospective analysis within the ALL-BFM trials. Blood J. Hematol. *118*, 5774–5782.

Wong, D.W., Chan, V.J., and Liao, H. (2019). Metagenomic discovery of feruloyl esterases from rumen microflora. Appl. Microbiol. Biot. *103*, 8449–8457.

Worth, H.G.J. (1985). A comparison of the measurement of Sodium and potassium by flame photometry and ion-selective electrode. Ann. Clin. Biochem. *22*, 343–350.

Supplemental Information

Metagenomic discovery and functional

validation of L-asparaginases

with anti-leukemic effect from the Caspian Sea

Motahareh Sobat, Sedigheh Asad, Mahboubeh Kabiri, and Maliheh Mehrshad
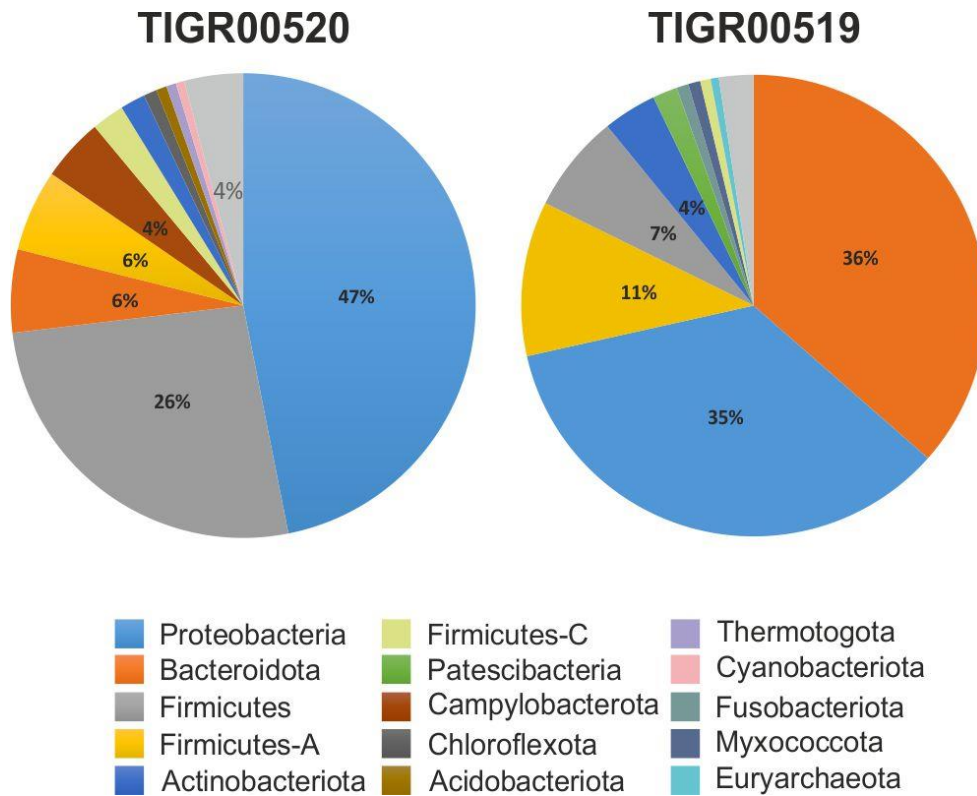
Figure S1, Related to Fig. 2- Distribution of Type II (TIGR00520) and Type I (TIGR00519) L-asparaginase genes in different bacterial phyla according to the results recovered from the Annotree database.
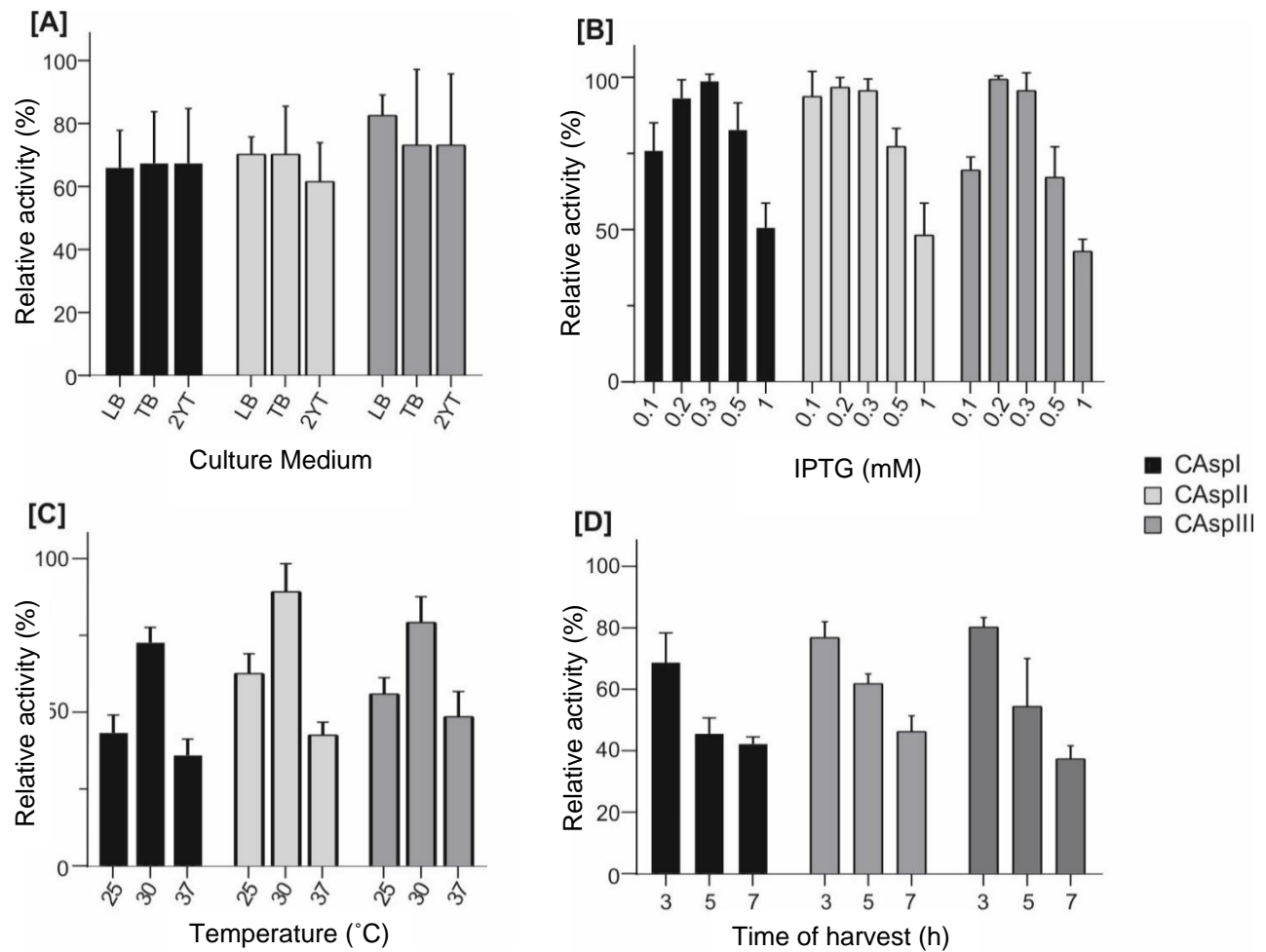
Figure S2, Related to Fig.4- Culture condition Optimization using one-factor-at-a time screening: A) Medium effect: Relative activity was not significantly different among three different culture media LB, TB, 2YT. LB was selected for the rest of the study. B) IPTG concentration effect: Optimal IPTG concentration for all three recombinant genes was determined to be between 0.2-0.3 mM. C) Temperature effect D) Harvest time duration effect. Data are represented as mean +/- SEM.

[A]

[B]

[C]
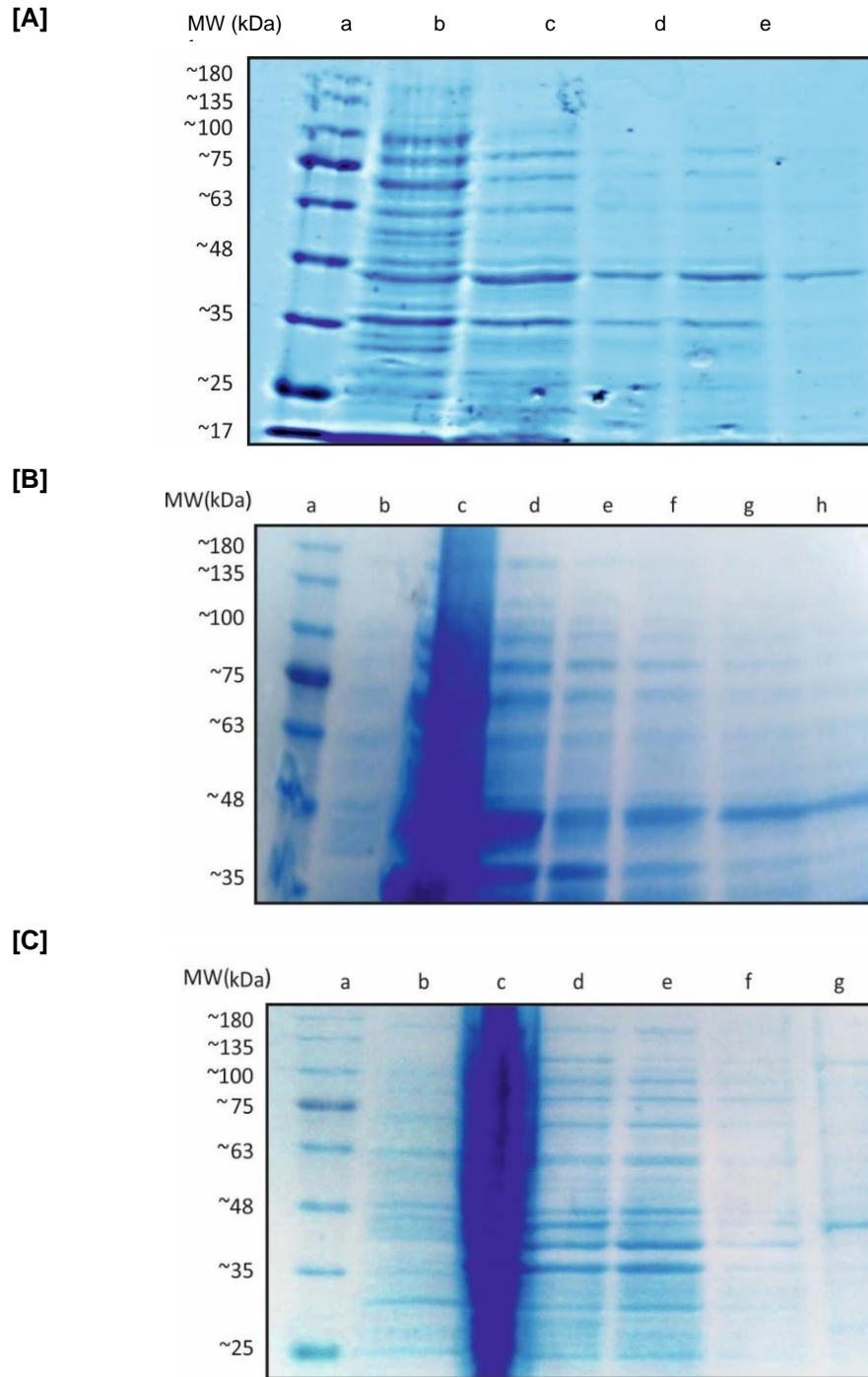
Figure S3, Related to Fig.4- Expression of recombinant L-asparaginases in *E. coli* BL21 (DE3). Expression of CAspI [A]: protein ladder (Lane a), pellet fraction of total proteins after induction (lane b), total soluble proteins of recombinant strain after induction (lane c), lane d to f: different fractions of column elution. Expression of CAspII [B] and CAspIII [C]: protein ladder (Lane a), total proteins before induction (lane b), total proteins after induction (lane c), total soluble proteins after induction (lane d), lane e,f and g: different fractions of affinity column after elution with imidazole.
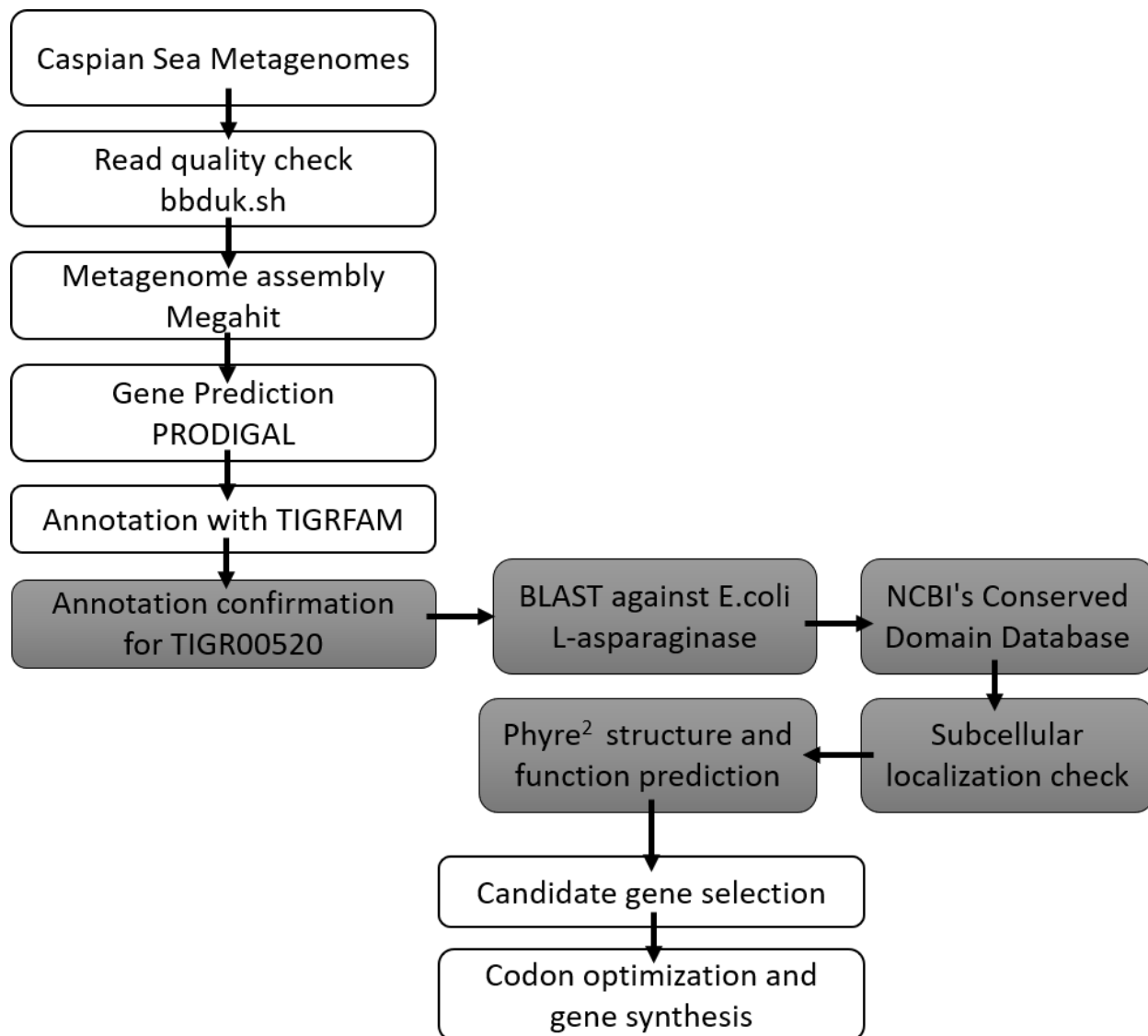
Figure S4, Related to Fig. 3- the flowchart of bioinformatic analysis pipeline for screening, in-silico function validation and synthesis of the candidate L-asparaginase genes from the Caspian Sea metagenomes.

Table S2, Related to Fig. 3- physicochemical features of human serum in comparison with the Caspian Sea water.

Sample	Na$^+$ (mM)	K$^+$ (mM)	pH	Salinity
Human serum	130 – 145	3.5 – 5.3	7.4	0.9
Caspian Sea	132.44	3.04	8	1.1

Transparent Methods
Chemicals and cell lines
All chemicals were of analytical grade purity. L-asparagine, L-glutamine, Dithiothreitol (DTT) and isopropyl-β-D-thiogalactopyranoside (IPTG) were purchased from Sigma-Aldrich (USA). All other chemicals were obtained from Merck (Germany) unless stated differently. The Jurkat and human umbilical vein endothelial cells (HUVEC) were procured from the Iranian Biological Resource Center (IBRC C10155) and Pasteur Institute of Iran (CRL-1730), respectively.

In-silico **L-aspariginase screening**
Metagenomics datasets derived from three different depths of the Caspian Sea (15, 40 and 150 meter) (Mehrshad et al., 2016) were used for screening putative L-asparaginase enzymes. The sequenced reads of these datasets were quality checked using bbduk.sh (https://sourceforge.net/projects/bbmap/) and assembled using Megahit (Li et al., 2015). Open reading frames (ORFs) of the assembled contigs ≥1kb were predicted using PRODIGAL (Hyatt et al., 2010). Identified ORFs were annotated against TIGRFAM (Haft et al., 2003)comprising curated multiple sequence alignments, Hidden Markov Models (HMMs) for 4488 protein families using hmmscan (Haft et al., 2003). ORFs annotated as putative L-asparaginase were affiliated to either of TIGR00519 or TIGR00520 HMMs. Putative L-asparaginase ORFs were further evaluated using NCBI's Conserved Domain Database (CDD) (Marchler-Bauer et al., 2016) and cross checked with both hallmark L-asparaginase genes of *E.coli* using BLAST. Sequences affiliated to the TIGR00520 were selected for further analysis for the purpose of cloning and expression. Subcellular localization of sequences was predicted using 5 different online servers: Phobius, mPLoc (Shen and Chou, 2009, Chou and Shen, 2008, Shen and Chou, 2007), SecretomeP 2.0 Server (Bendtsen et al., 2005a), SignalP 4.1 Server (Petersen et al., 2011), TatP 1.0 Server (Twin-arginine signal peptide) (Bendtsen et al., 2005b), if these servers enquired cell wall type of the organism containing these genes (Gram staining positive or negative) as an input for predictions, the sequence was examined twice with both possible assumptions. Structure and function of selected sequences were predicted using Phyre^2 web-based service (Kelley et al., 2015). (**Fig. S4**) Predicted secretory proteins were marked to be more desirable for further analysis. Sequences affiliated to fastidious bacteria were preferably selected for cloning. Additionally, genes with best BLAST hit to species spotted in all three depths of 15, 40 and 150 m were assumed to be more desirable as are adapted to sustain wide range of environmental conditions (**Table S3** shows the best BLAST hits for the recovered genes).

Phylogenetic analysis
Reviewed L-asparaginase sequences were collected from the UniProt (n=138). All open reading frames annotated as putative L-asparaginase (TIGR00519 and TIGR00520) in the Annotree database were also collected (n=12553) together with the 87 verified L-asparaginase genes identified from the Caspian Sea metagenomes. The retrieved sequences were aligned using Kalign (Lassmann and Sonnhammer, 2005) and the maximum likelihood phylogeny was reconstructed using FastTree2 (WAG+CAT model, gamma approximation, 100 bootstrap replicates) (Price et al., 2010).

Gene cloning
Three of the candidate genes annotated as TIGR00520 were codon optimized for expression in *E. coli*. These genes were synthetized and cloned into the pET21a (+) vector between *Nde*I and *Xho*I restriction sites in fusion with a histidine tag at C terminus (General Biosystems, USA). Following transformation into competent *E. coli* BL21 (DE3) cells, bacterial pools were plated on Luria-Bertani (LB) medium supplemented with ampicillin (100 µg/ml). Cloning accuracy was further confirmed by plasmid extraction (Yekta Tajhiz Azma Kit, Iran) and sequencing (Bioneer, Korea).

Protein expression and purification
To select the optimal expression condition, different concentrations of IPTG (0.1, 0.2, 0.3, 0.5, 1 mM), induction duration (3, 5 and 7 h), culture media (Luria Bertani (LB), Terrific Broth (TB) and Yeast Extract Tryptone (2xYT)) as well as the induction temperature (20, 30 and 37 °C) were tested. At last, pursuant to overnight pre-culture of bacteria in 10 ml LB medium (100 µg/ml ampicillin, 37 °C, 150 rpm), pre-culture was inoculated (1 % v/v) into 250 ml of fresh LB medium and left to grow at 37 °C, 150 rpm. Reaching the optical density of 0.5-0.6 at 600 nm, protein expression was induced by addition of 0.2 mM IPTG for CAspI and 0.3 mM for CAspII and III besides temperature decrement to 30 °C. Over a period of 3 h, cells were spun down (9000 g, 10 min), concentrated in 5 ml of ice-cold lysis buffer (50 mM potassium phosphate buffer, pH 7.8, NaCl 150 mM) and lysed by sonication (recurrent periods of 10s pulsing and 5s resting for 10 minutes, SYCLON Ultra Sonic Cell Sonicator SKL950-IIDN). Lysate was centrifuged (9000g, 30 min, 4°C) and supernatant was loaded on Ni-NTA agarose column pre-equilibrated with lysis

buffer. Elution of bound proteins was accomplished by the buffer containing 100mM of imidazole (pH 6.8). Purified protein solution was dialyzed against phosphate buffer (pH 7.2, 4°C) overnight to remove imidazole. All protein purification steps were carried out at 4°C.

SDS-PAGE (12% w/v acrylamide) was used to confirm the purity of the enzyme. Gels were stained by Coomassie Brilliant Blue R-250 (Bio-Rad, USA). Protein concentration was estimated by BCA method using bovine serum albumin as standard (Walker, 2009).

Enzyme activity assessment and biochemical characterization

1 µg/ml of purified enzymes was incubated with 4mM of L-asparagine for 2 min in 0.5 ml of 50 mM phosphate buffer (pH 7.2) and activity was determined by real time measurement of ammonia production via colorimetric assay (Nesslerization) (Kanazawa and Kiyota, 1995). Absorbance of Nessler reagent complexed with ammonia was determined by Perkin Elmer lambda 25 UV/VIS spectrophotometer at 410nm. Reaction rate was examined at different substrate concentrations (1.25 to 30 mM L-asparagine) using a standard curve of known concentrations of ammonium chloride. The data were fitted to generate a Michaelis–Menten equation, and apparent kinetic values were calculated. One unit (U) of enzyme's catalytic activity was defined as the amount of enzyme which converts one micromole of L-asparagine to ammonia per min at the condition of assay. The specific activity was expressed as U/mg protein. All assays were performed in triplicate, the means and standard deviations were presented.

Effect of pH and NaCl concentration on enzyme activity

With the intention of finding optimal activity conditions, enzyme activities were studied in various NaCl concentrations (50, 100, 150, 200 mM in 50mM potassium phosphate buffer, pH 7.2). Optimal pH for enzyme activity was studied by measuring activity at the pH range of 5-8 in a mixed buffer of 20 mM Tris-HCl (pKa 8.06), 20 mM Bis-Tris (pKa 6.46), and 20 mM sodium acetate (pKa 4.76) possessing buffering range of 3.5 – 9.2 (Brooke et al., 2015).

Anti- proliferation studies

Jurkat cells, clone E6-1, acute T cell leukemia cell line, were used for examining the anti-leukemic effect of our purified enzymes. The cells were cultured in RPMI-1640 medium (Gibco, USA) supplemented with 10% heat-inactivated fetal bovine serum (Gibco, USA) and 1% Pen/Strep (100 U/ml penicillin and 100 µg/ml streptomycin) and incubated under humidified 5% CO_2 atmosphere at 37°C. 100 µl of culture medium containing 5×10^5 cells/ml was added to 96 well plates. Purified recombinant enzymes were concentrated to various concentrations using Amicon Ultra-15 centrifugal filter. The concentrated solutions were sterilized using hydrophilic 0.22 µm Millipore filters. 100 µl of sterile enzyme solutions (10, 100 and 1000 nM) was added to each well and incubated for 24h at 37°C. All three concentrations of each of CAspI, CAspII and CAspIII enzymes, were evaluated in quadruplicate. 100 µl enzyme-free PBS solution was used to serve as negative control. As described previously (Kabiri et al., 2015), MTT (3-(4,5-dimethyl-2-thiazolyl)-2,5-diphenyl-2-tetrazolium bromide) assay was performed after 24h in order to quantify the amount of cell death in enzyme-treated groups in comparison with the untreated control group.

Enzyme Cytotoxicity assessment on non-leukemic cells

In order to assess any possible cytotoxic effect of our purified L-asparaginases on non-leukemic cells, HUVEC (human umbilical vein endothelial cells) were exposed to purified L-asparaginases (10, 100 and 1000 nM). MTT assay was performed after 24h of enzyme addition, and the results were compared against negative controls, as described above.

In vitro hemolysis test was also performed according to available protocols with minor modifications (Evans et al., 2013, Aslani et al., 2019). The test was done based on standard phlebotomy procedures. Briefly, 10 ml of human blood was collected in K2-EDTA-coated Vacutainer tubes and centrifuged at 500 g for 5 min. Hematocrit was washed twice with 150 mM NaCl solution and resuspended in PBS, pH 7.2. 100µl of sterile enzyme solution with different concentrations (10, 100 and 1000 nM) were mixed with 100 µl of blood in V-bottom 96-well plates and incubated at 37 °C for one hour in orbital shaker. Serial dilutions of Triton X-100 (0 to 20%) in PBS were used as controls. Plates were centrifuged at 500 g for 5 min to pellet intact erythrocytes. Then supernatants were transferred to new 96 well plate and absorbance was read at 541 nm. Hemolysis percentage was normalized against controls.

Supplemental References

ASLANI, S., KABIRI, M., KEHTARI, M. & HANAEE-AHVAZ, H. 2019. Vascular tissue engineering: Fabrication and characterization of acetylsalicylic acid-loaded electrospun scaffolds coated with amniotic membrane lysate. *Journal of cellular physiology,* 234**,** 16080-16096.

BENDTSEN, J. D., KIEMER, L., FAUSBØLL, A. & BRUNAK, S. 2005a. Non-classical protein secretion in bacteria. *BMC microbiology,* 5**,** 58.

BENDTSEN, J. D., NIELSEN, H., WIDDICK, D., PALMER, T. & BRUNAK, S. 2005b. Prediction of twin-arginine signal peptides. *BMC bioinformatics,* 6**,** 167.

BROOKE, D., MOVAHED, N. & BOTHNER, B. 2015. Universal buffers for use in biochemistry and biophysical experiments. *AIMS Biophysics,* 2**,** 336-342.

CHOU, K.-C. & SHEN, H.-B. 2008. Cell-PLoc: a package of Web servers for predicting subcellular localization of proteins in various organisms. *Nature protocols,* 3**,** 153.

EVANS, B. C., NELSON, C. E., SHANN, S. Y., BEAVERS, K. R., KIM, A. J., LI, H., NELSON, H. M., GIORGIO, T. D. & DUVALL, C. L. 2013. Ex vivo red blood cell hemolysis assay for the evaluation of pH-responsive endosomolytic agents for cytosolic delivery of biomacromolecular drugs. *JoVE (Journal of Visualized Experiments),* e50166.

HAFT, D. H., SELENGUT, J. D. & WHITE, O. 2003. The TIGRFAMs database of protein families. *Nucleic Acids Research,* 31**,** 371-373.

HYATT, D., CHEN, G.-L., LOCASCIO, P. F., LAND, M. L., LARIMER, F. W. & HAUSER, L. J. 2010. Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC bioinformatics,* 11**,** 119.

KABIRI, M., ORAEE-YAZDANI, S., DODEL, M., HANAEE-AHVAZ, H., SOUDI, S., SEYEDJAFARI, E., SALEHI, M. & SOLEIMANI, M. 2015. Cytocompatibility of a conductive nanofibrous carbon nanotube/poly (L-Lactic acid) composite scaffold intended for nerve tissue engineering. *EXCLI journal,* 14**,** 851.

KANAZAWA, S. & KIYOTA, H. 1995. Estimation of L-glutaminase and L-asparaginase activities in soils by the indophenol method. *Soil Science and Plant Nutrition,* 41**,** 305-311.

KELLEY, L. A., MEZULIS, S., YATES, C. M., WASS, M. N. & STERNBERG, M. J. 2015. The Phyre2 web portal for protein modeling, prediction and analysis. *Nature protocols,* 10**,** 845.

LASSMANN, T. & SONNHAMMER, E. L. 2005. Kalign–an accurate and fast multiple sequence alignment algorithm. *BMC bioinformatics,* 6**,** 298.

LI, D., LIU, C.-M., LUO, R., SADAKANE, K. & LAM, T.-W. 2015. MEGAHIT: an ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph. *Bioinformatics,* 31**,** 1674-1676.

MARCHLER-BAUER, A., BO, Y., HAN, L., HE, J., LANCZYCKI, C. J., LU, S., CHITSAZ, F., DERBYSHIRE, M. K., GEER, R. C., GONZALES, N. R., GWADZ, M., HURWITZ, D. I., LU, F., MARCHLER, G. H., SONG, J. S., THANKI, N., WANG, Z., YAMASHITA, R. A., ZHANG, D., ZHENG, C., GEER, L. Y. & BRYANT, S. H. 2016. CDD/SPARCLE: functional classification of proteins via subfamily domain architectures. *Nucleic Acids Research,* 45**,** D200-D203.

MEHRSHAD, M., AMOOZEGAR, M. A., GHAI, R., SHAHZADEH FAZELI, S. A. & RODRIGUEZ-VALERA, F. 2016. Genome Reconstruction from Metagenomic Data Sets Reveals Novel Microbes in the Brackish Waters of the Caspian Sea. *Applied and Environmental Microbiology,* 82**,** 1599-1612.

PETERSEN, T. N., BRUNAK, S., VON HEIJNE, G. & NIELSEN, H. 2011. SignalP 4.0: discriminating signal peptides from transmembrane regions. *Nature methods,* 8**,** 785.

PRICE, M. N., DEHAL, P. S. & ARKIN, A. P. 2010. FastTree 2–approximately maximum-likelihood trees for large alignments. *PloS one,* 5.

SHEN, H.-B. & CHOU, K.-C. 2007. Gpos-PLoc: an ensemble classifier for predicting subcellular localization of Gram-positive bacterial proteins. *Protein Engineering, Design & Selection,* 20**,** 39-46.

SHEN, H.-B. & CHOU, K.-C. 2009. Gpos-mPLoc: a top-down approach to improve the quality of predicting subcellular localization of Gram-positive bacterial proteins. *Protein and peptide letters,* 16**,** 1478-1484.

WALKER, J. M. 2009. The bicinchoninic acid (BCA) assay for protein quantitation. *The protein protocols handbook.* Springer.