Send Orders for Reprints to reprints@benthamscience.ae

Current Proteomics, 2017, *14*, 000-000

RESEARCH ARTICLE

Evaluation of Different Signal Peptides for Secretory Production of Recombinant Bovine Pancreatic Ribonuclease A in Gram Negative Bacterial System: an *In silico* Study

Ali Forouharmehr[1], Mohammadreza Nassiri[1,*], Shahrokh Ghovvati[2] and Ali Javadmanesh[1]

[1]*Department of Animal Science, Faculty of Agriculture, Ferdowsi University of Mashhad, Mashhad, Iran;* [2]*Department of Biotechnology, Faculty of Agriculture, University of Guilan, Rasht, Guilan, P.O. Box: 41635-1314, Iran*

Abstract: ***Background***: Prokaryotic systems such as *E. coli* are among the most affordable and simplest hosts which are being employed to express recombinant proteins, nevertheless without appropriate signal peptide these systems cannot be used for secretory proteins. Bovine pancreatic ribonuclease A is a protein with four disulfide bonds which might be used as an immunoenzyme for immunotherapy. Consequently, the production of this recombinant protein, using prokaryotic system, requires a suitable signal peptide to protect disulfide bonds and to prevent misfolding.

Objective: This study was designed to predict the best signal peptides to express bovine pancreatic ribonuclease A protein in *E. coli*.

Method: In this study, 42 signal sequences were selected from data bases and the most important features of them were evaluated. First, n, h and c regions of signal peptides and their probability were investigated by signalP software's version 3 and 4.1. Then, physico–chemical features of them were evaluated by Portparam and SOLpro. Also, secretion sorting and sub-cellular localization sites were evaluated by PRED-TAT and ProtcompB software programs.

Results: The results showed that among all studied signal peptides only 28 out of 41 remained signal peptides could be considered as appropriate secretory signal peptides.

Conclusion: Finally, Phage shock protein E, ranked as the best signal peptide and after that, pectate lyase B, F41 fimbrial protein and Lipopolysaccharide export system protein lptA considered as the next best signal peptides which were approved by *in silico* tools as the most appropriate secretory signal peptides in *E. coli*. However, further experiments are required to validate these *in silico* results.

ARTICLE HISTORY

Received: February 14, 2017
Revised: May 29, 2017
Accepted: June 02, 2017

DOI:
10.2174/1570164614666170725144424

Keywords: Bioinformatics, Bovine Pancreatic Ribonuclease A, *E. coli*, *In silico*, Signal peptide.

1. INTRODUCTION

Bovine pancreatic ribonucleic A (RNase A) (UniProtKB - P61823) which is secreted by exocrine cells of bovine pancreas, has 124 amino acids with a molecular mass of 13.7kD [1, 2]. Nowadays, most studies on RNase A have focused on its possible application as an immuneRNase, because it can degrade the transcriptome of the target cells, which could consequently kill them [3]. Therefore, the expression of a biologically active RNaseA as a recombinant protein is extremely crucial in cancer treatment [4]. Disulfide bonds in secretory proteins are essential for the proper folding and function [5]. Cytoplasm of eukaryote's cells has reductive environment which leads to the reduction and later destruction of the disulfide bonds, on the other hand, their rough endoplasmic reticulum (RER) environment provides an oxidative condition to maintain the disulfide bonds. Despite the fact that prokaryotes have reductive cytoplasm, they do not have RER, in fact periplasmic compartment of prokaryote does resemble RER, since it also has oxidative environment with the ability to provide safe condition to protect disulfide bonds [6-8], consequently in eukaryotes, a nascent secretory protein is directed to RER and in prokaryotes, it is directed to periplasmic compartment. Prokaryotic system in particular, *Escherichia coli (E. coli)* is being used for production of recombinant protein, in fact *E. coli* is the best host for the expression of recombinant proteins because it not only is less expensive than eukaryotic system like mammalian cells but

*Address correspondence to this author at the Department of Animal science, Faculty of Agriculture, Ferdowsi university of Mashhad, Mashhad, Iran; Tel: +98 511 8795616-18; Fax: +98 511 8787430;
E-mail: nassiryr@um.ac.ir

1570-1646/17 $58.00+.00

also is very simple to apply [9, 10]. Although RNase A as a recombinant protein can be expressed in E.coli, an important issue is to be considered here; there are four beta sheets and three alpha helixes in RNase A that are joined together by four disulfide bonds [11]. These disulfide bonds can be reduced by reductive environment of *E. coli* cytoplasm, therefore to keep up protein folding; we need a tool to direct RNase A to periplasmic compartment of *E. coli* [1, 5, 7]. The best approach to transfer RNase A to periplasmic compartment is to use an appropriate signal peptide (SP) [12]. As a matter of fact, in bacteria these tools can naturally translocate proteins to periplasmic compartment through different pathways. In general, in bacteria there are three important pathways for translocation of a secretory protein to periplasmic compartments that have been classified to the general secretion pathway(Sec-pathway);the twin arginine translocation (TAT-pathway) and the signal recognition particle pathway (SPR pathway). Furthermore, among these pathways TAT pathway can transfer folded proteins to periplasmic compartment, whereas Sec and SPR pathway transfer unfolded proteins to preplasmic compartment and rescue them [7, 13], therefore the researchers are widely using these tools to express secretory protein in which the identification of suitable SP for each protein seems very indispensable to express [7, 14, 15]. There have been some differences especially in the composition and length of SPs, but in general all of them are n-terminal peptide with three important regions; N-terminal region (n-region) with positive charge, a hydrophobic region (h-region) and a cleavable region (c-region). The h-region usually has 7-15 residues whereas n and c regions have fewer (3-5 residues in length). N and h- regions have vital role in transferring proteins to periplasmic space [16-18], whereas c-region plays a key role as a cleavable site which can be recognized by signal peptidase. Despite SPs important role in the production of recombinant protein, there has been no universal approach to detect them [16, 19]. In recent decades with the increase in biological data, biologists are mostly applying intelligence method such as machine learning to analyze the data [20], as in today, *in silico* analysis and bioinformatics tools have attracted special attention in biology, because they not only decrease the high cost of experiments but also provide reliable results [17]. In this study in order to identify an appropriate SP for secretory expression of RNase A protein in gram negative bacteria, most important features of 42 SPs sequences from gram negative bacteria were evaluated and compared using *in silico* methods and the best of which are introduced for experimental applications.

2. MATERIALS AND METHODS

2.1. Signal Sequence Collection and Study Design

In this study, amino acid sequences of 42 SPs were taken from national center of biotechnology information (NCBI) as shown in Table **1**. *In Silico* methods such as machine learning techniques were employed to analyze and characterize the collected signal sequences. Eventually, after trimming and prediction of subcellular localization site and also after excluding inappropriate signal peptides, the selected signal peptides were then evaluated to observe whether they have gained high level of secretory expression of RNase A protein in gram negative bacteria (*E. coli*).

2.2. *In Silico* Prediction of n, h and c Regions and Signal Peptide Probability

In order to predict n, h and c regions and signal peptide probability SignalP server version 4.1 (http://www.cbs.dtu.dk/services/SignalP/) was used [21, 22]. In order to use the server, each SP was connected to N-terminal of RNase A amino acid sequence and methionine residues were inserted between each SP and RNase A amino acid sequence. These are based on a combination of several artificial neural networks and hidden Markov models [21, 22].

2.3. Analysis of Physico – Chemical, Solubility, Secretion Sorting and Identification of Sub-cellular Localization Site of Signal Peptides

In silico analysis of physico-chemical features of signal peptides such as amino acid composition, molecular weight, theoretical PI, solubility index, aliphatic index, positively and negatively charged residues and grand average of hydropathicity (GRAVY) were all evaluated by protParam server [23] (http://web.expasy.org/cgi-bin/protparam/protparam). Prediction of signal peptide solubility was investigated by SOLpro server; in fact this server predicts the propensity of a protein to be soluble upon overexpression in *E. coli* using a two-stage SVM architecture based on multiple representations of the primary sequence [10] (http://scratch.proteomics.ics.uci.edu/). In order to sort SPs based on the secretion properties, PRED-TAT server [24] was applied (http://www.compgen.org/tools/PRED-TAT/submit). PRED-TAT operates based on hidden Markov models [24]. For analysis of signal peptides sub-cellular location, ProtCompB server was used, ProtComp combines several methods of protein localization prediction - neural networks-based prediction; direct comparison with updated base of homologous proteins of known localization; comparisons of pentamer distributions calculated for query and DB sequences (http://www.softberry.com). Average accuracy of ProtCompB is 86-100 % which depends on compartment of subcellular location, for instance this accuracy in plasma membrane is 100% but in extracellular is 86%. In order to apply SOLpro, PRED-TAT and ProtCompB, each SP was jointed to N – terminal of RNase A amino acid sequence and methionine residues were also inserted between SPs and RNase A amino acid sequence [10, 17].

3. RESULTS

3.1. *In Silico* Prediction of n, h and c-regions and Signal Peptide Probability

The results showed that SPs' D-scores were between 0.627 (phoA) and 0.894 (pspE) (Table **2**). The most important parameter for the diagnosis of a SP is the discriminating score (D-score) which is usually described with a cut-off value of 0.5. In fact only when an SP sequence has a D-score above 0.50, it is considered. The in silico analysis results of SignalP server has also indicated that the highest D-score belonged to pspE, PelB, FimF41a and lptA, respectively. As it was mentioned before that n and h regions are important in

Table 1. List of signal peptides present in various proteins of *Escherichia coli.*

No.	Full Name	Signal Peptide	Accession Number	Amino Acid Sequence
1	Periplasmic appA protein	appA	EHN88412	**MKAI**LIPFLSLLIPLT**PQSAFA**
2	Cytochrome c-type biogenesis protein	ccmH	AEJ57359	**MRF**LLGVLMLMI**SGSALA**
3	Protein cexE	cexE	WP_001687026	**MKKY**ILGVILAMG**SLSAIA**
4	Thiosulfate-binding protein	cysP	WP_033801079	**MAVNLLKKNS**LALVASLLL**AGHVQA**
5	Dr hemagglutinin structural subunit	draA	P24093	**MKKL**AIMAAASMVFA**VSSAHA**
6	Thiol:disulfide interchange protein dsbD	dsbD	WP_058033897	**MAQR**IFTLILLLC**STSVFA**
7	Thiol:disulfide interchange protein dsbG	dsbG	ETJ26382	**MLKK**ILLLALLPAI**AFA**
8	K88 fimbrial protein AD	faeG	WP_001380745	**MKKT**LIALAIAASAA**SGMAHA**
9	Iron(III) dicitrate-binding periplasmic protein	fecB	KDW96130	**MLAFIR**FLFAGLLLV**ISHAFA**
10	F107 fimbrial protein	fedA	ACY05963	**MKRL**VFISFVALSM**TAGSAMA**
11	F41 fimbrial protein	FimF41a	AAA23421	**MKKT**LIALAVAASAAV**SGSVMA**
12	Flagellar P-ring protein	flgI	EFJ97486	**MVIK**FLSALILLLV**TTAAQA**
13	Protein transport protein hofQ	hofQ	EDV85112	**MKQW**IAALLLML**IPGVQA**
14	Outer-membrane lipoprotein carrier protein	lolA	WP_016247003	**MKKI**AITCALLSSLV**ASSVWA**
15	Lipopolysaccharide export system protein lptA	lptA	EHV68281	**MKFKTNKLSLN**LVLASSLLAA**SIPAFA**
16	Maltose-binding periplasmic protein	malE	ELC32968	**MKIKTGAR**ILALSALTTM**MFSASALA**
17	Penicillin-insensitive murein endopeptidase	mepA	WP_001043836	**MNKT**AIALLALLA**SSVSLA**
17	Nickel-binding periplasmic protein	nikA	WP_021568845	**MLSTLRR**TLFALLACA**SFIVHA**
19	Cytochrome c-552	nrfA	CTU12334	**MTRIKINARR**IFSLLIPFF**FFTSVHA**
20	Outer membrane protein A	ompA	WP_021544610	**MKKT**AIAIAVALAG**FATVAQA**
21	Outer membrane protease ompP	ompP	WP_041124237	**MQTK**LLAIMLAAPVV**FSSQEASA**
22	Outer membrane protein W	ompW	EKW81199	**MKKLT**VAALAVTTLL**SGSAFA**
23	Fimbrial adapter papK	papK	WP_020239066	**MIKST**GALLLFAAL**SAGQAIA**
24	D-alanyl-D-alanine endopeptidase	pbpG	WP_032295491	**MPKFRV**SLFSLALMLAVP**FAPQAVA**
25	pectate lyase B	PelB	AAA24848	**MKYLL**PTAAAGLLLLA**AQPAMA**
26	Alkaline phosphatase	phoA	AAA24362	**MKQST**IALALLPLL**FTPVTKA**
27	Outer membrane pore protein E	phoE	EIO69468	**MKKST**LALVVMGIVA**SASVQA**
28	Protein prsK	prsK	EQN57820	**MIKST**GALLLFAAL**SAGQAMA**
29	Phage shock protein E	pspE	KDU08780	**MFKK**GLLALALVF**SLPVFA**
30	Protease 3	ptrA	EIL66839	**MPRSTWFK**ALLLLVALW**APLSQA**
31	S-fimbrial adhesin protein sfaS	sfaS	WP_021524832	**MKLK**AIILATGLINCI**AFSAQA**
32	Taurine-binding periplasmic protein	tauA	WP_032218149	**MAISSRN**TLLAALAFI**AFQAQA**
33	Thiamine-binding periplasmic protein	thiB	WP_032307836	**MLKK**CLPLLLLC**TAPVFA**
34	Periplasmic protein torT	torT	WP_029487908	**MRVL**LFLLLSLFM**LPAFS**
35	sn-glycerol-3-phosphate-binding periplasmic protein ugpB	ugpB	ELJ77555	**MKPLHYT**ASALALGLAL**MGNAQA**
36	D-xylose-binding periplasmic protein	xylF	EOV74805	**MKIKNI**LLTLCTSLLL**TNVAAHA**
38	Uncharacterized protein yfeK	yfeK	WP_053887217	**MKKI**ICLVITLLM**TLPVYA**
39	UPF0379 protein yhcN	yhcN	WP_058905387	**MKIKTT**VAALSVLSVL**SFGAFA**
40	Uncharacterized protein yncJ	yncJ	EYB53638	**MFTK**ALSVVLLTCAL**FSGQLMA**
41	UPF0482 protein ynfB	ynfB	WP_000705210	**MKITLSKRIG**LLAILLPCALAL**STTVHA**
42	Zinc resistance-associated protein	zraP	WP_042082503	**MKRNTKI**ALVMMALSAMA**MGSTSAFA**

In amino acid sequence of SPs, contrary to h region, n and c regions have been shown in bold.

Table 2. ***In silico*** **analysis of the signal peptide sequences by SignalP version 4.1.**

No.	Signal Peptides	n-Region	h-Region	c-Region	Cleavage Site	Cleavage Site Probability %	C-Score	Y-Score	S-Score	S-Mean	D-Score
1	appA	1-4(4)	5-16(12)	17-22(6)	AFA	98	0.609	0.696	0.927	0.832	0.76
2	ccmH	1-3(3)	4-12(9)	13-18(6)	ALA	90	0.708	0.808	0.943	0.906	0.854
3	cexE	1-4(4)	5-13(9)	14-19(6)	AIA	58	0.478	0.66	0.951	0.886	0.766
4	cysP	1-10(10)	11-19(9)	20-25(6)	VQA	92	0.517	0.657	0.904	0.857	0.751
5	draA	1-4(4)	5-15(11)	16-21(6)	AHA	93	0.448	0.642	0.965	0.929	0.777
6	dsbD	1-4(4)	5-13(9)	14-19(6)	VFA	80	0.771	0.843	0.953	0.906	0.873
7	dsbG	1-4(4)	5-14(10)	15-17(3)	AFA	43	0.284	0.518	0.959	0.936	0.714
8	faeG	1-4(4)	5-15(11)	16-21(6)	AMA	97	0.561	0.705	0.963	0.901	0.797
9	fecB	1-6(6)	7-15(9)	16-21(6)	AFA	79	0.555	0.721	0.958	0.928	0.818
10	fedA	1-4(4)	5-14(10)	15-21(7)	AMA	92	0.504	0.679	0.966	0.921	0.793
11	FimF41a	1-4(4)	5-16(12)	17-22(6)	VMA	96	0.793	0.843	0.972	0.914	0.876
12	flgI	1-4(4)	5-14(10)	15-20(6)	AQA	83	0.647	0.782	0.967	0.938	0.856
13	HLA-A	1-8(8)	9-19(11)	20-24(5)	TWA	96	0.69	0.777	0.964	0.875	0.83
14	hofQ	1-4(4)	5-13(9)	14-18(5)	VQA	59	0.534	0.68	0.94	0.839	0.755
15	lolA	1-4(4)	5-15(11)	16-21(6)	VWA	99	0.575	0.73	0.966	0.926	0.822
16	lptA	1-11(11)	12-21(10)	22-27(6)	AFA	90	0.726	0.816	0.983	0.938	0.874
17	malE	1-8(8)	9-18(10)	19-26(8)	ALA	93	0.461	0.653	0.983	0.937	0.787
18	mepA	1-4(4)	5-13(9)	14-19(6)	SLA	85	0.628	0.77	0.962	0.934	0.847
19	nikA	1-7(7)	8-16(9)	17-22(6)	VHA	94	0.67	0.787	0.941	0.918	0.849
20	nrfA	1-10(10)	11-19(9)	20-26(7)	VHA	88	0.421	0.596	0.934	0.87	0.725
21	ompA	1-4(4)	5-14(10)	15-21(7)	AQA	93	0.6	0.736	0.945	.0904	0.815
22	ompP	1-4(4)	5-15(11)	16-23(8)	ASA	55	0.363	0.505	0.9	0.766	0.628
23	ompW	1-5(5)	6-15(10)	16-21(6)	AFA	96	0.619	0.759	0.95	0.928	0.838
24	papK	1-5(5)	6-14(9)	15-21(7)	AIA	97	0.481	0.641	0.923	0.869	0.748
25	pbpG	1-6(6)	7-18(120	19-25(7)	AVA	84	0.424	0.601	0.981	0.904	0.744
26	PelB	1-6(6)	7-16(10)	17-22(6)	AMA	99	0.779	0.848	0.96	0.924	0.884
27	phoA	1-5(5)	6-14(9)	15-21(7)	TKA	51	0.285	0.476	0.888	0.797	0.627
28	phoE	1-5(5)	6-15(10)	16-21(6)	VQA	80	0.536	0.685	0.937	0.863	0.769
29	prsK	1-5(5)	6-14(9)	15-21(7)	AMA	99	0.672	0.771	0.939	0.89	0.827
30	pspE	1-4(4)	5-13(9)	14-19(6)	VFA	94	0.787	0.862	0.965	0.931	0.894
31	ptrA	1-8(8)	9-17(9)	18-23(6)	SQA	82	0.615	0.765	0.971	0.944	0.849
32	sfaS	1-4(4)	5-16(12)	17-22(6)	AQA	82	0.466	0.63	0.935	0.857	0.737
33	tauA	1-7(7)	8-16(9)	17-22(6)	AQA	93	0.659	0.737	0.917	0.848	0.789
34	thiB	1-4(4)	5-12(8)	13-18(6)	VFA	52	0.341	0.563	0.959	0.919	0.73
35	torT	1-4(4)	5-13(9)	14-18(5)	AFS	42	0.367	0.601	0.937	0.951	0.765
36	ugpB	1-7(7)	8-17(10)	18-23(6)	AQA	95	0.653	0.741	0.913	0.85	0.792
37	xylF	1-6(6)	7-16(10)	17-23(7)	AHA	94	0.456	0.649	0.959	0.918	0.775
38	yfeK	1-4(4)	5-13(9)	14-19(6)	VYA	77	0.605	0.754	0.937	0.908	0.826
39	yhcN	1-6(6)	7-16(10)	17-22(6)	AFA	94	0.523	0.7	0.961	0.937	0.811
40	yncJ	1-4(4)	5-15(11)	16-22(7)	LMA	99	0.582	0.732	0.945	0.914	0.818
41	ynfB	1-10(10)	11-22(12)	23-28(6)	VHA	99	0.779	0.828	0.983	0.922	0.872
42	zraP	1-7(7)	8-18(11)	19-26(8)	AFA	96	0.551	0.708	0.991	0.94	0.817

translocating a protein, while C region has an important role in cleaving SPs from protein, therefore a reliable SP sequence should have clear n, h and c regions. In the current study the collected SPs' n-region length was between 4 and 11, h region length was between 9 and 12, and c region length was between 3 and 5 amino acids. It seemed all SP sequences in our study not only had D-score above 0.50, but also contained obvious n, h and c regions.

3.2. Analysis of Physico – Chemical Properties and Solubility of Signal Peptides

The *in silico* results demonstrated that the studied SPs length variation was between 17 (dsbG) and 28 (ynfB) amino acid, the lowest and the highest Mw belonged to dsbG (3167.8) and ynfB (2948.7), respectively (Table **3**). The results also showed that the range of Net positive charge was between 0 and 4, whereas the range of PI was between 5.75 (ompP) and 12.3 (nrfA). The grand average of hydropathy score (GRAVY) is used to compare SPs overall hydropathy, in fact this parameter is defined as the sum of hydropathy of amino acids [17]. As it is observed the lowest GRAVY belonged to zraP (0.746) and the highest GRAVY belonged to fecB (2.076). Another factor used to show hydrophobicity is aliphatic index, this parameter is defined as the relative volume occupied by aliphatic side chain in an amino acid sequence. According to *in silico* results, the variation in range of aliphatic index was between 79.23 (zraP) and 207.06 (dsbG). Instability index evaluated as another parameter too, in general when instability is more than 40, possible proteins is considered unstable, whereas when instability is less than 40, it indicates the stability of the protein [17]. The instability of signal peptides alone and also in connection with RNase A was evaluated by instability index. The *in silico* analysis results showed that the variation in range of Instability index was between -2.6 (papK) and 65.64 (thiB). As a matter of fact the analysis results demonstrated that papK and yhcN(2.03) were the most stable signal peptides among the 42 studied signal peptides, respectively (-The most unstable signal peptides in connection with RNase A were thiB (65.6), nikA (60.45) and pbpG (57.99), respectively). The SOlpro server was applied for characterization of RNase A solubility in connection with the 42 studied signal peptides. Consequently, the results have indicated that all SPs connected to RNase A protein could make an insoluble protein, theoretically.

3.3. Secretion Sorting and Sub-cellular Localization

In this study, Sec, SRP and TAT pathways were evaluated by PRED-TAT software and the results demonstrated that all 42 studied SPs belonged to Sec – pathway. It seems all the SPs used in this study could transfer the expressed RNase A recombinant protein to the periplasmic compartment. Sub-cellular localization analysis showed (by ProtCompB server) that among 42 SPs, 29 SPs can localize RNase A in periplasmic space, 11 SPs can transfer this heterologous protein into extracellular space, one SP can localize this heterologous protein into cytoplasm (ompA) and finally, only one of the 42 studied SPs can transfer the expressed RNase A protein to plasma membrane (tauA). As it is observed the results of sub-cellular location (Table **4**) are largely confirming the results of pervious analysis (type of secretion pathways).

4. DISCUSSION

To keep appropriate folding and adequate accumulation, secretory proteins should be evaded from reductive environment of cytoplasm [5, 25, 26]. Eukaryote cells which use co-translation-translocation process evade nascent protein from cytoplasm and would direct it to RER, in fact this process depends on signal recognition particle complex (SRP) which can distinguish SP from nascent protein and translocate that to RER through transposon [27]. Prokaryote cells do not have RER but they have periplasmic compartment which provides oxidative environment to fold nascent protein. Sec, TAT and SPR are the main pathways in prokaryote cells directing nascent protein to periplasmic compartment. Furthermore these pathways operate based on signal peptide recognition, hence it is easily inferred that signal peptides play an important role in folding secretory protein in both eukaryote and prokaryote cells [28, 29]. As mentioned earlier, *E. coli* is the cheapest and simplest host to express recombinant proteins however the success in using it completely depends on employing the suitable SPs [30]. Consequently, the identification of suitable SPs is one of the most crucial steps to produce secretory proteins as a recombinant protein in *E. coli*. Today bioinformatics tools are widely being used in various parts of biological studies mainly because they reduce the cost of experiments and they also provide more accurate results [17, 31-32]. As it is observed in this study, it was attempted to employ the most accurate and recent version of bioinformatics tools to predict the variety of SP features. Among various features of SP, Net positive charge, aliphatic index, GRAVY, D-score, h-region length, cleavable site and Sub-cellular location are more important (Table **5**). Accordingly, these features were expected to make the final decision of selecting the best possible SPs. D-score is the first parameter in diagnosing an SP, therefore, SPs have all been sorted on the basis of D -score. In fact when D score is above 0.50, a signal sequence can be considered SP [17]. Since all SPs' D-score in this study is above 0.50, thereby all of them could be SP but for optimum screening, other features of selection should be considered. N – region is a crucial area in an SP which interferes translocation of a secretory protein, in fact for maintaining its function, n- region requires a positive charge and this charge is directly linked to the presence of one or more basic residues such as lysine at the beginning of an SP [17]. It is believed that switching the basic residues with neutral or acidic residues have an impact on translocation of nascent protein because of the significant role of this positive charge in interacting between SP of nascent protein and membrane phospholipid of RER [33]. As the results show, the range of net positive charge is calculated between 1 and 4, thereby it seems in this stage we do not have enough justification to decide whether to select any SP since all the selected SPs have appropriate net positive charge. Another important region which plays vital role in translocation is h-region, in fact the most important factor enabling h- region, is hydrophobicity. It has been reported this factor extremely relies on

Table 3. Physico-chemical properties of the signal peptides determined by ProtParam and SOLpro.

No.	Signal Peptides	Amino Acid Length	MW (Da)	PI	Net Positive Charge	GRAVY	Aliphatic Index	Instability	Solubility
1	appA	22	2384.9	8.5	1	1.405	155.45	53.16	INSOLUBLE
4	ccmH	18	1923.4	9.5	1	1.828	157.22	5.26	INSOLUBLE
5	cexE	19	1979.5	9.7	2	1.411	154.21	29.75	INSOLUBLE
6	cysP	25	2575.1	10	2	1.064	164	11.14	INSOLUBLE
7	draA	21	2135.6	10	2	1.162	98.1	16.49	INSOLUBLE
8	dsbD	19	2127.6	8	1	1.632	148.95	26.11	INSOLUBLE
9	dsbG	17	1839.4	10	2	2.018	207.06	33.41	INSOLUBLE
10	faeG	21	2027.4	10	2	1.005	112.38	11.36	INSOLUBLE
11	fecB	21	2350.9	9.52	1	2.076	162.86	9.52	INSOLUBLE
12	fedA	21	2231.7	11	2	1.29	102.38	29.55	INSOLUBLE
13	FimF41a	22	2090.5	10	2	1.355	124.55	15.15	INSOLUBLE
14	flgI	20	2116.6	8.5	1	1.935	185.5	10.64	INSOLUBLE
16	HLA-A	24	2528.1	9.5	1	1.258	142.5	5.55	INSOLUBLE
17	hofQ	18	1996.5	8.5	1	1.322	162.78	21	INSOLUBLE
18	lolA	21	2192.7	9.31	2	1.324	139.52	16.67	INSOLUBLE
19	lptA	27	2849.4	10.3	3	0.881	130.37	17.32	INSOLUBLE
20	malE	26	2698.3	11.17	3	1.012	113.08	2.85	INSOLUBLE
21	mepA	19	1887.3	8.5	1	1.479	164.74	32.07	INSOLUBLE
22	nikA	22	2434.9	10.35	2	1.35	137.37	60.45	INSOLUBLE
23	nrfA	26	3126.8	12.3	4	0.792	108.85	30.31	INSOLUBLE
24	ompA	21	2046.5	10	2	1.295	121.43	9.52	INSOLUBLE
25	ompP	23	2406.8	5.75	0	0.904	114.78	44.47	INSOLUBLE
26	ompW	21	2093.5	10	2	1.21	125.71	1.44	INSOLUBLE
27	papK	21	2047.4	8.5	1	1.39	140	-2.6	INSOLUBLE
28	pbpG	25	2705.3	11	2	1.228	117.2	57.99	INSOLUBLE
29	PelB	22	2228.7	8.34	1	1.191	138.18	41.42	INSOLUBLE
30	phoA	21	2256.8	10	2	0.971	139.52	56.02	INSOLUBLE
31	phoE	21	2104.5	10	2	1.195	130	1.44	INSOLUBLE
32	prsK	21	2065.5	8.5	1	1.267	121.43	3.27	INSOLUBLE
33	pspE	19	2065.6	10	2	1.711	148.95	17.37	INSOLUBLE
34	ptrA	23	2613.2	11	2	0.857	131.74	51.93	INSOLUBLE
35	sfaS	22	2290.8	9.31	2	1.314	146.82	5.41	INSOLUBLE
36	tauA	22	2308.7	9.5	1	1.055	120.45	34.41	INSOLUBLE
37	thiB	18	1974.6	8.89	2	1.589	157.22	65.64	INSOLUBLE
38	torT	18	2111.7	9.5	1	2.061	173.33	26.66	INSOLUBLE
39	ugpB	23	2342.8	8.37	1	0.622	110.87	18.01	INSOLUBLE
40	xylF	23	2482	9.31	2	1.083	161.3	33.61	INSOLUBLE
41	yfeK	19	2163.8	9.19	2	1.742	179.47	42.39	INSOLUBLE
42	yhcN	22	2254.7	10	2	1.418	128.64	-2.03	INSOLUBLE
43	yncJ	22	2344.9	7.98	1	1.541	128.64	15.15	INSOLUBLE
44	ynfB	28	2948.7	10.06	3	1.239	163.93	29.32	INSOLUBLE
45	zraP	26	2733.3	11.17	3	0.746	79.23	28.75	INSOLUBLE

Table 4. Secretion sorting and sub-cellular location of SPs.

				Sub-cellular Location Score				
No.	**Signal peptides**	**Type of SP**	**Reliability Score %**	**Cytoplasmic**	**Periplasmic**	**Secreted (Extracellular)**	**Membrane**	**Final Prediction Site**
1	appA	Sec	100	1.73	8.27	0.00	0.00	Periplasmic
2	ccmH	Sec	97	3.68	6.32	0.00	0.00	Periplasmic
3	cexE	Sec	97	2.59	6.54	0.87	0.00	Periplasmic
4	cysP	Sec	99	2.43	0.00	7.42	0.14	Extracellular
5	draA	Sec	100	1.45	5.44	3.11	0.00	Periplasmic
6	dsbD	Sec	97	3.16	6.13	0.70	0.00	Periplasmic
7	dsbG	Sec	99	2.80	7.20	0.00	0.00	Periplasmic
8	faeG	Sec	99	3.41	6.59	0.00	0.00	Periplasmic
9	fecB	Sec	99	0.31	0.00	9.69	0.00	Extracellular
10	fedA	Sec	99	1.79	3.05	5.17	0.00	Extracellular
11	FimF41a	Sec	99	3.95	6.05	0.00	0.00	Periplasmic
12	flgI	Sec	99	0.14	0.00	9.07	0.78	Extracellular
13	HLA-A	Sec	99	3.17	5.65	1.18	0.00	Periplasmic
14	hofQ	Sec	99	4.27	5.73	0.00	0.00	Periplasmic
15	lolA	Sec	99	0.69	9.05	0.00	0.26	Periplasmic
16	lptA	Sec	99	1.26	7.71	0.00	1.02	Periplasmic
17	malE	Sec	99	1.48	8.52	0.00	0.00	Periplasmic
18	mepA	Sec	97	0.64	0.00	9.13	0.22	Extracellular
19	nikA	Sec	98	1.32	7.26	1.41	0.00	Periplasmic
20	nrfA	Sec	99	0.29	9.32	0.39	0.00	Periplasmic
21	ompA	Sec	99	4.71	4.45	0.84	0.00	Cytoplasmic
22	ompP	Sec	99	4.39	5.28	0.00	0.33	Periplasmic
23	ompW	Sec	99	2.04	3.34	4.62	0.00	Extracellular
24	papK	Sec	99	3.44	6.56	0.00	0.00	Periplasmic
25	pbpG	Sec	99	1.24	8.76	0.00	0.00	Periplasmic
26	PelB	Sec	100	4.02	5.98	0.00	0.00	Periplasmic
27	phoA	Sec	99	1.50	8.50	0.00	0.00	Periplasmic
27	phoE	Sec	99	3.72	6.28	0.00	0.00	Periplasmic
29	prsK	Sec	99	4.32	5.68	0.00	0.00	Periplasmic
30	pspE	Sec	98	2.51	7.49	0.00	0.00	Periplasmic
31	ptrA	Sec	99	1.77	8.23	0.00	0.00	Periplasmic
32	sfaS	Sec	99	0.00	6.87	3.13	0.00	Periplasmic
33	tauA	Sec	99	3.65	0.00	1.02	5.33	Membrane
34	thiB	Sec	98	2.17	7.65	0.00	0.18	Periplasmic
35	torT	Sec	98	0.87	7.97	0.00	1.16	Periplasmic
36	ugpB	Sec	99	0.64	9.36	0.00	0.00	Periplasmic
37	xylF	Sec	100	0.08	4.75	5.17	0.00	Extracellular
38	yfeK	Sec	99	0.50	1.54	7.97	0.00	Extracellular
39	yhcN	Sec	99	0.23	0.00	9.77	0.00	Extracellular
40	yncJ	Sec	99	4.42	0.00	5.34	0.24	Extracellular
41	ynfB	Sec	99	2.03	0.00	7.55	0.41	Extracellular
42	zraP	Sec	100	0.33	9.67	0.00	0.00	Periplasmic

Reliability score is defined as a parameter that shows amount of our trust on accuracy of data.

Table 5. Sorting the signal peptides according to aliphatic index, GRAVY, h-region length and D-score respectively.

No.	Signal Peptides	Net Positive Charge	Aliphatic Index	Gravy	D-score	h-Region Length	Final Prediction Site	Signal Peptide Statues
1	pspE	2	148.95	1.711	0.894	5-13(9)	Periplasmic	Confirmed
2	PelB	1	138.18	1.191	0.884	7-16(10)	Periplasmic	Potentially
3	FimF41a	2	124.55	1.355	0.876	5-16(12)	Periplasmic	Confirmed
4	lptA	3	130.37	0.881	0.874	12-21(10)	Periplasmic	Confirmed
5	dsbD	1	148.95	1.632	0.873	5-13(9)	Periplasmic	Potentially
6	ynfB	3	163.93	1.239	0.872	11-22(12)	Extracellular	Potentially
7	flgI	1	185.5	1.935	0.856	5-14(10)	Extracellular	Potentially
8	ccmH	1	157.22	1.828	0.854	4-12(9)	Periplasmic	Potentially
9	nikA	2	137.37	1.35	0.849	8-16(9)	Periplasmic	Confirmed
10	ptrA	2	131.74	0.857	0.849	9-17(9)	Periplasmic	Confirmed
11	mepA	1	164.74	1.479	0.847	5-13(9)	Extracellular	Potentially
12	ompW	2	125.71	1.21	0.838	6-15(10)	Extracellular	Confirmed
13	prsK	1	121.43	1.267	0.827	6-14(9)	Periplasmic	Potentially
14	yfeK	2	179.47	1.742	0.826	5-13(9)	Extracellular	Potentially
15	lolA	2	139.52	1.324	0.822	5-15(11)	Periplasmic	Potentially
16	yncJ	1	128.64	1.541	0.818	5-15(11)	Extracellular	Potentially
17	fecB	1	162.86	2.076	.0818	7-15(9)	Extracellular	Potentially
18	zraP	3	79.23	0.746	0.817	8-18(11)	Periplasmic	Confirmed
19	ompA	2	121.43	1.295	0.815	5-14(10)	Cytoplasmic	Confirmed
20	yhcN	2	128.64	1.418	0.811	7-16(10)	Extracellular	Potentially
21	faeG	2	112.38	1.005	0.797	5-15(11)	Periplasmic	Confirmed
22	fedA	2	102.38	1.29	0.793	5-14(10)	Extracellular	Confirmed
23	ugpB	1	110.87	0.622	0.792	8-17(10)	Periplasmic	Confirmed
24	tauA	1	120.45	1.055	0.789	8-16(9)	Plasma Membrane	Confirmed
25	malE	3	113.08	1.012	0.787	9-18(10)	Periplasmic	Confirmed
26	Astn1	0	121.43	1.948	0.782	4-13(10)	Periplasmic	Potentially
27	draA	2	98.1	1.162	0.777	5-15(11)	Periplasmic	Confirmed
28	xylF	2	161.3	1.083	0.775	7-16(10)	Extracellular	Confirmed
29	phoE	2	130	1.195	0.769	6-15(10)	Periplasmic	Confirmed
30	cexE	2	154.21	1.411	0.766	5-13(9)	Periplasmic	Confirmed
31	torT	1	173.33	2.061	0.765	5-13(9)	Periplasmic	Confirmed
32	appA	1	155.45	1.405	0.76	5-16(12)	Periplasmic	Confirmed
33	hofQ	1	162.78	1.322	0.755	5-13(9)	Periplasmic	Potentially
34	cysP	2	164	1.064	0.751	11-19(9)	Extracellular	Confirmed
35	papK	1	140	1.39	0.748	6-14(9)	Periplasmic	Potentially
36	pbpG	2	117.2	1.228	0.744	7-18(12)	Periplasmic	Confirmed
37	sfaS	2	146.82	1.314	0.737	5-16(12)	Periplasmic	Confirmed
38	thiB	2	157.22	1.589	0.73	5-12(8)	Periplasmic	Confirmed
39	nrfA	4	108.85	0.792	0.725	11-19(9)	Periplasmic	Confirmed
40	dsbG	2	207.06	2.018	0.714	5-14(10)	Periplasmic	Confirmed
41	ompP	0	114.78	0.904	0.628	5-15(11)	Periplasmic	Confirmed
42	phoA	2	139.52	0.971	0.627	6-14(9)	Periplasmic	Confirmed

Sorting of this table was done according to D-score (from highest to lowest score).

the length of h-region. In the other word, the increase in the length of h-region would improve the level of hydrophobicity [34]. Accordingly, there has not been a significant diversity in the length of SPs h-region (9 to 12) thereby other important parameters were used such as aliphatic index and GRAVY in recognition of hydrophobicity. Aliphatic index and GRAVY are the two parameters with direct relationship with hydrophobicity, in the other words the increase in these parameter, leads to the increase of hydrophobicity [17, 33]. As it has been reported in Table **5**, among 42 SPs only zraP has low aliphatic index (79.23) and GRAVY (0.746) while in the case of other SPs, no significant difference was observed; therefore, it seems zraP is not a suitable SP to express RNase A protein. C-region, particularly the three terminal residues that are also named -3, -2, -1 box, are extremely significant in detaching SP and the secretory protein after translocation, in fact -3, -2, -1 boxes are recognized and cleaved by the signal peptidase. Previous studies have indicated that there are usually Small or neutral residues such as alanine in -1 and -3 positions, whereas there are often big residues in -2 position which is different with the residues in -1 and -3 positions, this residue is illustrated with X [12, 17, 35]. As shown in Table **3** all SPs are following this rule and are almost similar to AXA box, therefore we have avoided mentioning this parameter in Table **5**. In general the bacteria which uses Sec and SPR pathways translocate unfolded proteins to periplasmic compartment where folding and accumulation are both occurring, on the contrary by the use of TAT pathway they tend to fold secretory proteins in cytoplasm compartment and then translocate the folded proteins to periplasmic compartment for accumulation [7], it seems Sec and SPR pathways are more Vital than TAT Pathway because folding and purification of secretary proteins in periplasmic are easier than in cytoplasm. Since degradation of secretary proteins is less than cytoplasm, it can be concluded that the SPs using these pathways can be more appropriate than SPs which use TAT pathways [36, 37]. As it is shown in Table **4**, all SPs in this study belonged to Sec pathway and none could be deleted using this analysis, subsequently other analysis was performed here (it has been reported in previous sections). Finally, it was clarified that among 41 SPs (without zarP), 28 of them can translocate RNase A protein to periplasmic compartment which could confirm the previous analysis (sec pathway), however 13 of them translocate it to other compartments (Extracellular, cytoplasmic and membrane compartment) which could not confirm the previous result; therefore, it seems only these 28 signal sequences can be introduced as reliable SP. There has been no big difference among 28 SP; therefore, according to D-score (the most important feature), Phage shock protein E (pspE), pectate lyase B (PelB), F41 fimbrial protein (FimF41a), and Lipopolysaccharide export system protein lptA (lptA) were introduced (respectively) as the best signal peptides to express RNase A protein in *E. coli*. PleB which is the most famous signal peptide and is considered the second best SP in this analysis, is a well-founded proof in approving our data since it is commercially being used to express secretory protein in *E. coli* in which the results obtained by this SP were satisfactory [4, 38].

CONCLUSION

Recently, the emergence of *in silico* approach such as bioinformatics and data analysis in theoretical biology accelerated the process of analyzing SPs for the production of recombinant proteins. Moreover, it has reduced the costs of the expression and purification of recombinant proteins as well the time required for the process. So, predicting the best SPs by *in silico* approach would help biologist and protein engineers to accelerate and facilitate the vital projects. The aim of this study was to Predict the candidate SPs to express RNaseA, and it was attempted to employ the most accurate softwares to evaluate the most important features of SPs. Eventually, Phage shock protein E (pspE), pectate lyase B (PelB), F41 fimbrial protein (FimF41a), Lipopolysaccharide export system protein lptA(lptA)and Thiol:disulfide interchange protein dsbD(dsbD) were introduced (respectively) as the best signal peptides to express RNase A protein in *E.coli*. Although the confirmation of these results requires experimental evaluation.

LIST OF ABBREVIATIONS

c-region	=	Cleavable region
D-score	=	Discriminating score
GRAVY	=	Grand average of hydropathicity
h-region	=	Hydrophobic region
NCBI	=	National Center of Biotechnology Information
n-region	=	N-terminal region
RER	=	Rough endoplasmic reticulum
RNase A	=	Bovine pancreatic ribonuclease A
Sec-pathway	=	General secretion pathway
SP	=	Signal peptide
SPR pathway	=	Signal recognition particle pathway
SRP	=	Signal recognition particle complex
TAT-pathway	=	Twin arginine translocation pathway

ETHICS APPROVAL AND CONSENT TO PARTICIPATE

Not applicable.

HUMAN AND ANIMAL RIGHTS

No Animals/Humans were used for studies that are base of this research.

CONSENT FOR PUBLICATION

Not applicable.

CONFLICT OF INTEREST

The authors declare no conflict of interest, financial or otherwise.

ACKNOWLEDGEMENTS

This study was financially supported by the Ferdowsi University of Mashhad (grant number: 3/37489).

REFERENCES

[1] Breukelman, H. J.; van der Munnik, N.; Kleineidam, R. G.; Furia, A.; Beintema, J. J. Secretory ribonuclease genes and pseudogenes in true ruminants., *Gene* **1998**, *212*, 259-268.

[2] Vasantha, N.; Filpula, D. Expression of bovine pancreatic ribonuclease A coded by a synthetic gene in Bacillus subtilis., *Gene* **1989**, *76*, 53-60.

[3] Leich, F.; Stöhr, N.; Rietz, A.; Ulbrich-Hofmann, R.; Arnold, U. Endocytotic internalization as a crucial factor for the cytotoxicity of ribonucleases. *J. Biol. Chem.,* **2007**, *282*, 27640-27646.

[4] Riccio, G.; D'Avino, C.; Raines, R. T.; De Lorenzo, C. A novel fully human antitumor ImmunoRNase resistant to the RNase inhibitor. *Protein Eng Des Sel.*, **2012**, *26*, 243-248.

[5] Kadokura, H.; Katzen, F.; Beckwith, J. Protein disulfide bond formation in prokaryotes. *Annu rev biochem.,* **2003**, *72*, 111-135.

[6] Berndt, C.; Lillig, C. H.; Holmgren, A. Thioredoxins and glutaredoxins as facilitators of protein folding. *BBA-Mol Cell Res.,* **2008**, *1783*, 641-650.

[7] De Marco, A. Strategies for successful recombinant expression of disulfide bond-dependent proteins in Escherichia coli. *Microbcell fac.,* **2009**, *8*, 1.

[8] Messens, J.; Collet, J. F. Pathways of disulfide bond formation in Escherichia coli. *The Int J Biochem Cell B.,* **2006**, *38*, 1050-1062.

[9] Idicula-Thomas, S.; Balaji, P.V. Understanding the relationship between the primary structure of proteins and its propensity to be soluble on overexpression in *Escherichia coli*. *Protein Science : Protein Sci,* **2005**, *14*, 582-592.

[10] Magnan, C. N.; Randall, A.; Baldi, P. SOLpro: accurate sequence-based prediction of protein solubility. *Bioinformatics.,* **2009**, *25*, 2200-2207.

[11] Shin, H. C.; Scheraga, H. A. Catalysis of the oxidative folding of bovine pancreatic ribonuclease A by protein disulfide isomerase. *J. Mol. Biol.,* **2000**, *300*, 995-1003.

[12] Choi, J.; Lee, S. Secretory and extracellular production of recombinant proteins using Escherichia coli. *Appl Microbiol Biot .,* **2004**, *64*, 625-635.

[13] Natale, P.; Brüser, T.; Driessen, A. J. Sec-and Tat-mediated protein secretion across the bacterial cytoplasmic membrane—distinct translocases and mechanisms. *BBA-Biomembranes.,* **2008**, *1778*, 1735-1756.

[14] Gardy, J. L.; Brinkman, F. S. Methods for predicting bacterial protein subcellular localization. *Nature Rev. Microbiol.,* **2006**, *4*, 741-751.

[15] Müller, M.; Bernd Klösgen, R. The Tat pathway in bacteria and chloroplasts (review). *Mol. Membr. Biol.,* **2005**, *22*, 113-121.

[16] Emanuelsson, O.; Brunak, S.; Von Heijne, G.; Nielsen, H. Locating proteins in the cell using TargetP, SignalP and related tools. *Nat. protocols.,* **2007**, *2*, 953-971.

[17] Zamani, M.; Nezafat, N.; Negahdaripour, M.; Dabbagh, F.; Ghasemi, Y. In silico evaluation of different signal peptides for the secretory production of human growth hormone in E. coli. *Int J Pept Res Thers.,* **2015**, *21*, 261-268.

[18] Zimmermann, R.; Eyrisch, S.; Ahmad, M.; Helms, V. Protein translocation across the ER membrane. *(BBA)-Biomembrane.,* **2011**, *1808*, 912-924.

[19] Zhang, C.; Marcia, M.; Langer, J. D.; Peng, G.; Michel, H. Role of the N-terminal signal peptide in the membrane insertion of Aquifex aeolicus F1F0ATP synthase c-subunit. *FEBS Journal .,***2013**, *280*, 3425-3435.

[20] Ezziane, Z. Applications of artificial intelligence in bioinformatics: A review. *Expert Systems with Applications.,* **2006**, *30*, 2-10.

[21] Bendtsen, J. D.; Nielsen, H.; von Heijne, G.; Brunak, S. Improved prediction of signal peptides: SignalP 3.0. *J. mol. biol .,* **2004**, *340*, 783-795.

[22] Petersen, T. N.; Brunak, S.; von Heijne, G.; Nielsen, H. SignalP 4.0: discriminating signal peptides from transmembrane regions. *Nat. methods.,***2011**, *8*, 785-786.

[23] Walker, J.M., *The Proteomics Protocols Handbook*. Humana Press: 2005.

[24] Bagos, P. G.; Nikolaou, E. P.; Liakopoulos, T. D.; Tsirigos, K. D. Combined prediction of Tat and Sec signal peptides with hidden Markov models. *Bioinformatics.,* **2010**, *26*, 2811-2817.

[25] Roszczenko, P.; Radomska, K.A.; Wywial, E.; Collet, J.-F.; Jagusztyn-Krynicka, E.K. A Novel Insight into the Oxidoreductase Activity of Helicobacter pylori HP0231 Protein. *PLoS ONE.,* **2012**, *7*, e46563.

[26] Seras-Franzoso, J.; Affentranger, R.; Ferrer-Navarro, M.; Daura, X.; Villaverde, A.; García-Fruitós, E. Disulfide bond formation and activation of Escherichia coli β-galactosidase under oxidizing conditions. *Appl. Environ. Microbiol.,* **2012**, *78*, 2376-2385.

[27] Luirink, J.; Sinning, I. SRP-mediated protein targeting: structure and function revisited. *BBA-Mol Cell Res.,* **2004**, *1694*, 17-35.

[28] Baneyx, F.; Mujacic, M. Recombinant protein folding and misfolding in Escherichia coli. *Nat. Biotechnol.,* **2004**, *22*, 1399-1408.

[29] Keller, R.; de Keyzer, J.; Driessen, A. J.; Palmer, T. Co-operation between different targeting pathways during integration of a membrane protein. *J. Cell Biol.,* **2012**, *199*, 303-315.

[30] Rosano, G. L.; Ceccarelli, E. A. Recombinant protein expression in Escherichia coli: advances and challenges. *Front Microbiol.,* **2014**, 14,130-137.

[31] Ghasemi, Y.; Dabbagh, F.; Rasoul-Amini, S.; Haghighi, A. B.; Morowvat, M. H. The possible role of HSPs on Behçet's disease: a bioinformatic approach. *Comput. Biol. Med.,* **2012**, *42*, 1079-1085.

[32] Nezafat, N.; Ghasemi, Y.; Javadi, G.; Khoshnoud, M. J.; Omidinia, E. A novel multi-epitope peptide vaccine against cancer: an in silico approach. *J. Theor. Biol.,* **2014**, *349*, 121-134.

[33] Low, K. O.; Mahadi, N. M.; Illias, R. M. Optimisation of signal peptide for recombinant protein secretion in bacterial hosts. *Appl. Microbiol. Biotechnol.,* **2013**, *97*, 3811-3826.

[34] Chen, H.; Kim, J.; Kendall, D. A. Competition between functional signal peptides demonstrates variation in affinity for the secretion pathway. *J Bacteriol.,***1996**, *178*, 6658-6664.

[35] Payne, S. H.; Bonissone, S.; Wu, S.; Brown, R. N.; Ivankov, D. N.; Frishman, D.; Paša-Tolić, L.; Smith, R. D.; Pevzner, P. A. Unexpected diversity of signal peptides in prokaryotes. *MBio.,* **2012**, *3*, e00339-12.

[36] Pugsley, A. P.; Schwartz, M. Export and secretion of proteins by bacteria. *FEMS Microbiol. Lett.,* **1985**, *1* (1), 3-38.

[37] Talmadge, K.; Gilbert, W. Cellular location affects protein stability in Escherichia coli. *Proc. Natl. Acad. Sci. U.S.A.,* **1982**, *79*, 1830-1833.

[38] Sockolosky, J. T.; Szoka, F. C. Periplasmic production via the pET expression system of soluble, bioactive human growth hormone. *Protein Expr. Purif.,* **2013**, *87*, 129-135.