university of groningen

University Medical Center Groningen

University of Groningen

Analysis of different signal peptides for the secretory production of Ama r 2 in gram-positive systems (Lactococcus lactis)

Vasiee, Alireza; Norouzi, Neda; Yazdi, Farideh Tabatabaee; Mortazavi, Seyed Ali; Sankian, Mojtaba; Mahmoudi, Mahmoud; Shahidi, Fakhri

Published in:
Microbial Pathogenesis

DOI:
10.1016/j.micpath.2019.103819

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2020

Link to publication in University of Groningen/UMCG research database

Citation for published version (APA):
Vasiee, A., Norouzi, N., Yazdi, F. T., Mortazavi, S. A., Sankian, M., Mahmoudi, M., & Shahidi, F. (2020). Analysis of different signal peptides for the secretory production of Ama r 2 in gram-positive systems (Lactococcus lactis). *Microbial Pathogenesis*, *138*, [103819]. https://doi.org/10.1016/j.micpath.2019.103819

Copyright
Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the “Taverne” license. More information can be found on the University of Groningen website: https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment.

Take-down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): http://www.rug.nl/research/portal. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

Microbial Pathogenesis 138 (2020) 103819

Contents lists available at ScienceDirect

Microbial Pathogenesis

journal homepage: www.elsevier.com/locate/micpath

Analysis of different signal peptides for the secretory production of Ama r 2 in gram-positive systems (*Lactococcus lactis*)

Alireza Vasiee[a], Neda Norouzi[b], Farideh Tabatabaee Yazdi[a], Seyed Ali Mortazavi[a,*], Mojtaba Sankian[c], Mahmoud Mahmoudi[c], Fakhri Shahidi[a]

[a] Department of Food Science and Technology, Faculty of Agriculture, Ferdowsi University of Mashhad, Mashhad, Iran
[b] University Medical Center Groningen, Antonius Deusinglaan 1, 9713, AV, the Netherlands
[c] Immunology Research Center, Bu-Ali Research Institute, School of Medicine, University of Medical Sciences, Mashhad, Iran

ARTICLE INFO

Keywords:
Gram positive system
Bioinformatics
Ama r 2 pollen
Signal peptide

ABSTRACT

Prokaryotic systems have been considered the most affordable and simplest hosts which are being employed to express recombinant proteins such as allergens; nevertheless, without appropriate signal peptide (SP), these systems cannot be used for secretory proteins. Recently, a lot of effort has been put into assessing the potential of gram-positive strains such as lactic acid bacteria for new applications in the production of heterologous proteins. Ama r 2 is a respiratory allergen from *Amaranthus retroflexus,* whose recombinant production in the probiotic host could be introduced as a specific and effective way to rapid diagnosis and immunotherapy of this allergy. Consequently, the production of this recombinant protein using the prokaryotic system, requires a suitable SP to protect disulfide bonds and to prevent misfolding. This study was designed to predict the best SPs for the expression of Ama r 2 protein in *Lactococcus lactis* as the host. In this study, 42 signal sequences were selected from SP databases and the most important features of them were evaluated. First, n, h and c regions of the SPs and their probabilities were investigated by signalP software version 4.1. Then, their physicochemical properties were evaluated by Portparam and SOLpro. Moreover, the secretion sorting and sub-cellular localization sites were evaluated by PRED-TAT and ProtcompB software programs. The results revealed that yjgB, entC2 (Entrotoxine type C-2), ent B (Entrotoxine type), blaZ (Beta lactamase), dex (number 21), blm (Beta lactamase 2), dex (Dextranase; number 20) and number 26 were introduced theatrically as the best SPs to express Ama r 2 in *Lactococcus lactis*.

1. Introduction

Amaranthaceae family, short-lived plants, grow in different countries with dry climate such as Iran, Kuwait, India, the Mediterranean area and Saudi Arabia [1–4]. *Amaranthus retroflexus* is a native and well-known species in this family whose pollen can play a major role in seasonal respiratory allergy among the people who live in the regions with such climatic conditions [4]. *A. retroflexus* is a causative agent of rhinitis allergic in Iran and up to 69% of patients are sensitive to its pollens [5,6]. Protein profile characterization of crude *A. retroflexus* extract revealed different components ranging from 10 to 85 kDa with main IgE-reactive bands at 10, 15, 18, 25, 39, 45, 50, 66, and 85 kDa [1,3,4]. According to the published literature, Ama r 2 is one of the two identified allergens of *A. retroflexus* pollen, including Ama r 1 [2] and Ama r 3 [1]. [6] designated and reported Ama r 2 with a molecular weight of14.2 kDa as the first allergen from *A. retroflexus* pollen [3].

Improvement in diagnostic and immunotherapy strategies can be simplified through our knowledge of the allergenic profile characterization from a biological source and how to gain the recombinant allergens. Although huge efforts have been made on natural allergens preparation, there have been several disadvantages, including the production of heterogeneous products (non-allergic proteins and other macromolecules), risk of being contaminated with allergens from other sources or proteolytic enzymes, and laboriousness to standardize a mixture of different proteins even using advanced techniques [7]. Over the past 10 years, the production of recombinant allergens as well as introducing a specific and effective way to the rapid diagnosis and immunotherapy of allergy, has been able to dissolve these drawbacks. There have been remarkable progresses in the molecular biology of allergens as several publications have cloned and expressed numerous allergens in different prokaryote or eukaryote cells [8–10].

Prokaryote cells have been considered better hosts for the

* Corresponding author.
E-mail address: Morteza@um.ac.ir (S.A. Mortazavi).

https://doi.org/10.1016/j.micpath.2019.103819
Received 4 September 2019; Received in revised form 22 October 2019; Accepted 23 October 2019
Available online 25 October 2019
0882-4010/ © 2019 Published by Elsevier Ltd.

Table 1
Allergens from the *Amaranthus retroflexu* that have been identified, characterized, and expressed as recombinant proteins.

Organism	Allergen	Protein familly	MW (kDa)	Accession number	Source of recombinant expression	Reference
Amaranthus retroflexus	Ama r 1	Ole e1-like	18	A0A0K1SC10	*Escherichia coli*	[2]
	Ama r 2	Profilin	14	C3W2Q7	*Escherichia coli*	[3]
	Ama r 3	Polcalcin	12	–	*Escherichia coli*	[1]

biotechnological production of heterologous proteins and clinical delivery of therapeutic molecules since they are not only more economic, but also entirely manageable to be applied. The use of prokaryote systems, in particular gram-negative strains (e.g. *E. coli*), as a tool has been investigated in the production of several recombinant allergens (Table 1) [1–3]. Recently, several efforts have been made to assess the potential of gram-positive strains such as lactic acid bacteria (LAB) for new applications in the production of heterologous proteins [11–14]. This research area has drawn a lot of attention because of the easy protein secretion into the medium by most LAB, which streamlines the often expensive and inefficient downstream purification process, besides other advantages like simple and well-known metabolism and the availability of various genetic tools [13]. Furthermore, LAB, in particular *Lactococcus lactis,* have been generally regarded as safe (GRAS) and of major interest as potential food-grade cell factories, whereas do not produce any endotoxin or lipopolysaccharide compounds [12,15]. Interestingly, no experimental evidence of using gram-positive bacteria for cloning and expression of one of *A. retroflexus* allergens as a recombinant protein has been obtained up to now, which may be crucial in immunotherapy of allergy. Therefore, *Lactococcus lactis* can be represented as the best-characterized species of the Gram-positive system, and figures as the model organism for recombinant allergens production, because of its easy manipulation, sequenced genome and the development of several genetic tools.

Proteins are synthesized as precursors containing the mature fraction of the protein with an amino-terminal extension sequence called signal peptide (SP) in both prokaryote and eukaryote systems [13]. Generally, SPs (with 15–30 amino acids) are located in the n-terminal of nascent protein. they contain three structures, including a positively-charged N terminus (n-region, 1–5 residues), a hydrophobic core (h-region, 7–15 residues), and a neutral and polar C terminus (c-region) containing the SP cleavage site (3–7 residues) which can be recognized by the SP [16–19]. Following the cleavage of the SP by signal peptidases, the mature proteins translocate extracellularly. Therefore, c-region has the main role as a cleavable site, whereas N and h-regions play key roles in transferring the proteins to beyond the cell membrane [17]. While these structural features are conserved between bacterial SPs, their sequences and lengths may vary [20]. The primary step in order to have a suite translocation of a recombinant protein, increase the protein yield and get the protein with appropriate solubility and folding, is to use an appropriate SP. Since SP plays an important role in targeting the protein entry into the secretory pathway which identification and selection of a proper SP seems requisite to express each selective protein [19].

There are three pathways for the translocation of secretory proteins to periplasmic space or medium in prokaryote systems, named Sec (Secretion), TAT (Twin Arginine Translocation) and SPR (Signal Recognition Particle) pathways [17]. In recent decades, *in vitro* and *in vivo* efforts have been made to evaluate numerous SPs for several secretory proteins in the periplasmic space of prokaryote cells; however, these are very time-consuming, laborious and expensive approaches. Consequently, these disadvantages made biologists apply Intelligence methods such as *in silico* analysis and bioinformatics tools to analyze the data, as in today, have attracted special attention in biology, because they not only decrease the high cost of experiments, but also provide reliable outcomes and moreover, enable scientists to have a clearer view of the results previous to the real experimental study [16,17].

To date, heterologous proteins such as bovine plasmin [21], bovine beta-lactoglobulin (BLG [22] and bovine rotavirus nonstructural protein 4 (NSP4 [23]) have been fused into *lactococcal* SPs to direct their secretion in the medium. On the other hand, Table 1 provides a short overview of the studies on the expression of *A. retroflexus* allergens in Gram-negative sources. Interestingly, to the best of our knowledge, there is experimental evidence for neither the secretory production nor the theoretical prediction of suitable SPs for Ama r 2 allergen in *L. lactis*. In this study, we used an *in silico* approach to evaluate and predict the most appropriate SP candidates among 200 SPs for the secretory production of Ama r 2 in *Lactococcus lactis* NZ1330.

2. Materials and methods

2.1. Signal peptide sequence collection

Firstly, the amino acid sequence of Ama r 2 was retrieved from https://www.uniprot.org with the accession number of C3W2Q7. The amino acid sequences of the SPs were also provided from SP database server (http://www.signalpeptide.de/). In the second step, the important features of the SP sequences were analyzed using *in silico* methods. The amino acid sequence of each SP was fused into Ama r 2 sequence. Finally, after trimming and predicting the subcellular localization site, as well as excluding inappropriate SPs, the selected SPs were compared and evaluated to achieve a high level of Ama r 2 secretion in the gram-positive hosts (Table 2).

2.2. Prediction of signal peptide regions and probability

In this study Signal P server version 4.1 (http://www.cbs.dtu.dk/services/SignalP-4.1/) was employed as a computational tool for the prediction of n, h and c regions, in addition tosignal peptide probability. Among all available prediction servers, Signal P server has been known as the best server in distinguishing between signal and non-signal peptides as well as identifying signal peptides cleavage sites in different organisms, with an overall accuracy of 87% [24]. In order to perform this server, each signal sequence and the N-terminal of Ama r 2 amino acid sequences were connected with methionine residues. The probability of each signal peptide sequence was determined based on the discrimination score called "D-score" briefly.

2.3. In silico evaluation of physicochemical parameters

To calculate and *in silico* estimate the different physical and chemical properties of the signal sequences (SSs), including molecular weight, amino acid composition, PI, aliphatic index, solubility index, positively- and negatively-charged residues and grand average of hydropathicity (GRAVY), ProtParam online server available at http://web.expasy.org/protparam/was applied.

2.4. Prediction of solubility, secretion sorting and sub-cellular localization site of SPs

The analysis of protein solubility upon expression was performed via PROSO II server available at http://mips.helmholtz-muenchen.de/ and SOLpro server at http://scratch.proteomics.ics.uci.edu/applying two-stage Support Vector Machine (SVM) architecture based on

Table 2
Signal sequences (SS) and their source list retrieved from http://www.signalpeptide.de/.

No.	Accession Number	Full Name	Signal peptide	Source	Amino acid sequence
1	P16271	PI-type proteinase	prtP	*Lactococcus lactis* subsp. *cremoris*	MQRKKKGLSILLAGTVALGALAVLPVGEIQAKA
2	Q9AIQ2	n/a	prtP	*Lactococcus lactis* subsp. *cremoris*	MQRKKKGLSFLLAGTVALGALAVLPVGEIQAKA
3	O66086	n/a	–	*Lactococcus lactis*	MKKINLALLTLATLMGVSSTAVVFA
4	Q8KKF8	n/a	yjgB	*Lactococcus lactis*	MLKKIIISAALMASLSAAMIANPAKA
5	Q9RLV2	n/a	310	*Lactococcus lactis*	MKFNKKRVAIATFIALIFVSFFTISSIQDNQTNA
6	Q9AIQ2	n/a	prtP	*Lactococcus lactis* subsp. *cremoris*	MQRKKKGLSFLLAGTVALGALAVLPVGEIQAKA
7	Q05044	S-layer protein	–	*Lactobacillus brevis*	MQSSLKKSLYLGLAALSFAGVAAVSTTASA
8	P38059	S-layer protein	slpH	*Lactobacillus helveticus*	MKKNLRIVSAAAAALLAVAPIAATAMPVNA
9	Q9S398	n/a	S-layer	*Lactobacillus helveticus*	MKKNLRIVSVAAAALLAVAPVAATAMPVNA
10	Q9XB19	n/a	S-layer	*Lactobacillus helveticus*	MKKNLRIVSAAAAALLAVAPVAATAMPVNA
11	A1IHD1	n/a	glcNAcase	*Lactobacillus casei*	MKKHGFRFWLLAIVGLLGMVLSLSQPPKQVAA
12	Q02470	PII-type proteinase	prtP	*Lactobacillus paracasei*	MQRKKKGLSILLAGTVALGALAVLPVGEIQAKA
13	Q5QSY9	n/a	cpf	*Lactobacillus coryniformis subsp. coryniformis*	MKKLVRNSMLVGAAVLGIGSLGTVAQA
14	Q9FCU5	n/a	mapA	*Lactobacillus reuteri*	MKFWKKALLTIVALTVGTPAGITSVSA
15	B1NRV2	n/a	bacA	*Enterococcus faecalis*	MKKKLVKGLVICGMIGIGFTALGTNAEA
16	Q0Z8B6	Bacteriocin hiracin-JM79	hirJM79	*Enterococcus hirae*	MKKKVLKHCVILGILGTCLAGIGTGIKVDA
17	P34071	Enterotoxin type C-2	entC2	*Staphylococcus aureus*	MNKSRFISCVILIFALILVLFTPNVLA
18	P00807	Beta-lactamase	blaZ	*Staphylococcus aureus*	MKKLIFLIVIALVLSACNSNSSHA
19	Q2FHS7	FPRL1 inhibitory protein	flr	*Staphylococcus aureus* (strain USA300)	MKKNITKTIIASTVIAAGLLTQTNDAKA
20	P39653	Dextranase	dex	*Streptococcus downei*	MNNRMLSFPSMLFLLAFGIVLSVSAGTTHA
21	Q8VLP4	n/a	dex	*Streptococcus downei*	MLSLPSMLCLLAFGMVFSISAKPAHA
22	P00779	Streptokinase C	skc	*Streptococcus equisimilis*	MKNYLSFGMFALLFALTFGTVNSVQA
23	Q54099	n/a	fnb	*Streptococcus equisimilis*	MKNKVLKTLVLFLFAVLGIFAMQTVEA
24	Q9LCB8	n/a	ily	*Streptococcus intermedius*	MKTKQNIARKLSRVVLLSTLVLSSAAPISAAFA
25	Q54727	Sialidase B	nanB	*Streptococcus pneumoniae*	MNKRGLYSKLGISVVGISLLMGVPTLIHA
26	Q54892	n/a	–	*Streptococcus pyogenes*	YSLRKLKTGTASVAVALTVLGAGFANQTEVKA
27	Q48R29	Streptopain	speB	*Streptococcus pyogenes* serotype M28	MNKKKLGIRLLSLLALGGFVLANPVFA
28	P04190	Beta-lactamase 2	blm	*Bacillus cereus*	MKKNTLLKVGLCVGLLGTIQFVSTISSVQA
29	Q9RED0	n/a	slpA	*Bacillus thuringiensis*	MAKTNSYKKVIAGTMTAAMVAGVVSPVAA
30	P09333	Outer cell wall protein	–	*Brevibacillus brevis*	MNKKVVLSVLSTTLVASVAASAFA
31	Q4JVC3	n/a	jk1070	*Corynebacterium jeikeium* (strain K411)	MPRRFSSFGRPLARFLAASGMVAAVSVGLGAPASA
32	Q9Z489	n/a	i-inlF	*Listeria ivanovii*	MRKNDWLKNVLITILVTVFVVCVNMSLETKAQA
33	Q01839	Protein p60	iap	*Listeria welshimeri*	MNMKKATIAATAGIAVTAFAAPTIASA
34	Q70YZ7	n/a	padA	*Streptococcus dysgalactiae*	MKKYVKILGLSSLAGLMLMASLVGNEASA
35	Q53974	n/a	mag	*Streptococcus dysgalactiae*	MEKEKKVKYFLRKSAFGLASVSAAFLVGTAVVNA
36	Q54099	n/a	fnb	*Streptococcus equisimilis*	MKNKVLKTLVLFLFAVLGIFAMQTVEA
37	Q9S399	n/a	S-layer	*Lactobacillus helveticus*	MKKNLRIVSVAAAALLAVAPVAATAMPVNA
38	P09332	Exfoliative toxin B	etb	*Staphylococcus aureus*	MDKNMFKKIILAASIFTISLPVIPFESTLQA
39	P01552	Enterotoxin type B	entB	*Staphylococcus aureus*	MYKRLFISHVILIFALILVISTPNVLA
40	Q2YU84	Uncharacterized leukocidin-like protein 1	SAB1875c	*Staphylococcus aureus* (strain bovine RF122/ET3-1)	MIKQLYKNITICSLTISTALTVFPATSYA
41	Q53691	n/a	hlgC-like ORF	*Staphylococcus aureus*	MLKNKILATTLSVSLLAPLANPLLENAKA
42	Q8NUI5	Lipase 1	lip1	*Staphylococcus aureus* (strain MW2)	MKSQNKYSIRKFSVGASSILIATLLFLSGGQAQA

multiple representations of the primary sequence. The solubility prediction was performed for the signal peptides linked to the N-terminal region of Ama r 2 sequence. The overall accuracy of SOLpro was estimated to be over 74% [25]. PRED-TAT server operating based on hidden Markov models [26], was applied for the secretion sorting of the signal peptides (http://www.compgen.org/tools/PRED-TAT/submit). In order to analyze the signal peptides sub cellular localization, we used ProtCompB server (http://www.softberry.com). The average accuracy of ProtCompB is 86–100% which depends on the compartment of the subcellular location. For instance, this accuracy is 100% in plasma membrane, but 86% in extracellular translocation. In order to perform SOLpro, PRED-TAT and ProtCompB, each SP was jointed to the N–terminal of Ama r 2 amino acid sequence, and methionine residues were also inserted between SPs and Ama r 2 amino acid sequence [17,25].

3. Results

3.1. Prediction of signal peptide regions and probability

We tried to use the prediction potency of bioinformatics tools to find the best signal peptide candidates for Ama r 2 in gram-positive hosts theoretically. Table 3 represents several SPs features, including different scores (C, Y, S, S-mean and D), cleavage site and common regions (n-, h- and c-regions) obtained via applying signal software version 4.1. In general, D-score is the most important factor in determining the potential signal peptide, which is described with a default cut-off value of 0.5; therefore, a signal sequence with a D-score above 0.5, can be considered to be a signal peptide. The results showed that the highest D-score belonged to yigB, dex, blaZ, iap and cpf respectively among the 42 collected SPs. It has been reported that regarding Gram-negative hosts, the length of h-region should be between 7 and 15 residues, whereas the length of c- and n-regions should be between 3 and 5 residues, otherwise, the SP cannot play its role well [27]. It should be noted that among 200 SPs, those having n-, h- and c-regions could not figure as a proper SP. It can be consisted by this claim that the SPs of gram-positive bacteria are longer than those of gram-negative ones [18]. In this investigation, the *in silico* analysis results of signalP server indicated that the SP's n-, h- and c-region lengths were between 4 and 14, 11 and 18, and 5 and 11 residues respectively.

3.2. In silico evaluation of physicochemical parameters and solubility of SPs

In this study, as mentioned earlier, Protparam server was applied to predict the physical and chemical properties of the signal peptides as shown in Table 4. The results demonstrated that the amino acid length of the SP ranged from 24 (blaZ and outer cell wall protein) to 35 (jk1070), while the lowest and highest molecular weight belonged to

Table 3
In silico analysis of the signal peptides sequences by SignalP 4.1.

No.	Signal peptide	n-Region	h-Region	c-Region	Cleavage Site	C-score	Y-score	S-score	S-mean	D-score
1	prtP	1, 6	7, 24	25, 33	AKA	0.427	0.47	0.84	0.681	0.552
2	prtP	1, 6	7, 24	25, 33	AKA	0.428	0.476	0.858	0.694	0.561
3	–	1, 5	6, 19	20, 25	VFA	0.277	0.457	0.932	0.813	0.596
4	yjgB	1, 5	6, 20	21, 26	AKA	0.833	0.811	0.949	0.856	0.829
5	310	1, 8	9, 24	25, 34	TNA	0.373	0.442	0.826	0.666	0.53
6	prtP	1, 6	7, 24	25, 33	AKA	0.428	0.476	0.858	0.694	0.561
7	–	1, 10	11, 24	25, 30	ASA	0.574	0.671	0.925	0.819	0.729
8	slpH	1, 6	7, 19	20, 30	VNA	0.459	0.58	0.959	0.844	0.683
9	S-layer	1, 6	7, 19	20, 30	VNA	0.392	0.546	0.971	0.864	0.67
10	S-layer	1, 6	7, 19	20, 30	VNA	0.409	0.555	0.967	0.859	0.673
11	glcNAcase	1, 8	9, 23	24, 32	VAA	0.6	0.63	0.843	0.615	0.625
12	prtP	1, 6	7, 24	25, 33	AKA	0.427	0.47	0.84	0.681	0.552
13	cpf	1, 7	8, 21	22, 27	AQA	0.653	0.682	0.872	0.754	0.71
14	mapA	1, 6	7, 20	21, 27	VSA	0.348	0.486	0.897	0.754	0.59
15	bacA	1, 8	9, 22	23, 28	AEA	0.358	0.463	0.835	0.678	0.547
16	hirJM79	1, 8	9, 23	24, 30	VDA	0.285	0.368	0.794	0.622	0.467
17	entC2	1, 6	7, 20	21, 27	VLA	0.702	0.621	0.778	0.632	0.626
18	blaZ	1, 4	5, 17	18, 24	SHA	0.522	0.668	0.92	0.867	0.745
19	flr	1, 7	8, 21	22, 28	AKA	0.749	0.659	0.802	0.647	0.655
20	dex	1, 7	8, 22	22, 30	THA	0.429	0.551	0.898	0.785	0.642
21	dex	1, 6	7, 19	20, 26	AHA	0.647	0.729	0.916	0.838	0.771
22	skc	1, 6	7, 20	21, 26	VQA	0.597	0.626	0.779	0.678	0.646
23	fnb	1, 7	8, 21	22, 27	VEA	0.404	0.512	0.866	0.73	0.597
24	ily	1, 13	14, 26	27, 33	AFA	0.404	0.572	0.95	0.86	0.684
25	nanB	1, 9	10, 23	24, 29	IHA	0.428	0.477	0.65	0.562	0.51
26	–	1, 8	9, 23	24, 32	VKA	0.6	0.575	0.864	0.681	0.616
27	speB	1, 9	10, 20	21, 27	VFA	0.613	0.605	0.842	0.705	0.644
28	blm	1, 8	9, 21	22, 30	VQA	0.571	0.606	0.747	0.672	0.632
29	slpA	1, 9	10, 24	25, 29	VAA	0.277	0.448	0.912	0.74	0.562
30	Outer cell wall protein	1, 4	5, 17	19, 24	AFA	0.353	0.55	0.963	0.896	0.685
31	jk1070	1, 14	15, 29	30, 35	ASA	0.269	0.428	0.929	0.793	0.57
32	i-inlF	1, 9	10, 25	26, 33	AQA	0.503	0.557	0.767	0.544	0.552
33	iap	1, 6	7, 20	21, 27	ASA	0.761	0.72	0.89	0.743	0.729
34	padA	1, 7	8, 22	23, 29	ASA	0.589	0.643	0.903	0.781	0.697
35	mag	1, 13	14, 27	28, 34	VNA	0.219	0.335	0.913	0.689	0.473
36	fnb	1, 7	8, 21	22, 27	VEA	0.404	0.512	0.866	0.73	0.597
37	S-layer	1, 6	7, 19	20, 30	VNA	0.392	0.546	0.971	0.864	0.67
38	etb	1, 8	9, 22	23, 31	LQA	0.26	0.369	0.768	0.604	0.461
39	entB	1, 9	10, 20	21, 27	VLA	0.523	0.508	0.718	0.547	0.523
40	SAB1875c	1, 8	9, 22	23, 29	SYA	0.337	0.434	0.845	0.651	0.519
41	hlgC-like ORF	1, 7	8, 20	21, 29	AKA	0.324	0.452	0.911	0.768	0.575
42	lip1	1, 12	13, 27	28, 34	AQA	0.611	0.631	0.764	0.66	0.642

blaZ (2573.15) and 310 (3865.55) respectively. PI was in the range of 6–12. Furthermore, the Net positive charge of all the selected SPs was between 1 and 4. Evaluation of the instability of the SPs alone was also presented by the instability index which indicates the stability of amino acid sequences when it is more than 40, SP or protein is known as unstable molecule, whereas a protein with an instability index of less than 40, is predicted to be stable [19]. The results showed that the instability index of all the 43 SPs was between −8.88 (bacA) and 53.64 (jk1070), while the most stable SPs were bacA, hirJM79 (−6.9) and blm (−5.92). In general, GRAVY is a measure of SP hydrophobicity defined as the sum of amino acids hydropathy [17]. A positive GRAVY score indicates hydrophobicity and a negative one is considered a hydrophilicity indicator. Moreover, aliphatic index plays an important role in the hydrophobicity of SPs [28]. Based on *in silico* observations, flr (0.229) and entB (1.663) had the lowest and highest GRAVY respectively, while the variation in the range of the aliphatic index was between 81.14 (jk1070) and 184.07 (entB). Since insoluble proteins tend to aggregate in inclusion bodies, protein solubility is one of the major properties of the passenger proteins to be translocated to the periplasmic space [28]. The solubility scores of the SP sequences (presented in Fig. 1) demonstrated that 12 SPs were soluble among all of the 43 SPs. Moreover, the maximum solubility scores belonged to glCNAcase, outer cell wall protein, blm and flr respectively.

3.3. *Secretion sorting and sub-cellular localization site of SPs*

Gram-positive bacteria possess only one lipid bilayer and a very thick cell wall, considerably thicker than that of Gram-negative ones. Because of these differences in the basic cell structure, it is not surprising that Gram-positive bacteria differ from Gram-negative organisms in their mechanisms of extracellular protein secretion. Like Gram-negative organisms, Gram-positive bacteria operate at both the Tat and Sec pathways to transport proteins across the cytoplasmic membrane [29]. Based on PRED-TAT results, all of the 42 studied SP belonged to the Sec pathway. Furthermore, sub-cellular localization analysis was done by ProtCompB server.

4. Discussion

As mentioned earlier, both Tat and Sec pathways contribute to protein transportation beyond the cell cytoplasm in Gram-positive bacteria, while the difference in the cell wall structure may cause differences in their mechanisms of extracellular protein secretion in these systems [29]. For example, the monolayer cell wall in LAB permits the direct secretion of the desired protein to the extracellular environment. However, in many cases, this transport is not sufficient to deliver proteins to their final objectives. Indeed, SP plays an important role in targeting the protein to the cytoplasmic membrane, where nascent protein is subsequently translocated by the Sec machinery. Following

Table 4
Physicochemical properties of the SP determined by ProtParam and SOLpro.

No.	Signal peptide	Amino acid length	MW (Da)	PI	Net Positive Charge	GRAVY	Aliphatic index	Instability	Solubility
1	prtP	33	3346.12	10.46	4	0.661	139.09	18.18	INSOLUBLE
2	prtP	33	3380.14	10.46	4	0.609	127.27	12.35	SOLUBLE
3	–	25	2593.22	10	2	1.38	144.4	13.91	INSOLUBLE
4	yjgB	26	2629.32	10.3	3	1.165	135.77	14.13	SOLUBLE
5	310	34	3865.55	10.29	3	0.474	97.65	34.58	INSOLUBLE
6	prtP	33	3380.14	10.46	4	0.609	127.27	12.35	SOLUBLE
7	–	30	2944.44	9.7	2	0.837	107.67	25.39	INSOLUBLE
8	slpH	30	2948.62	11.17	3	1.087	130.67	40.74	INSOLUBLE
9	S-layer	30	2962.65	11.17	3	1.157	133.67	47.16	INSOLUBLE
10	S-layer	30	2934.6	11.17	3	1.077	127.33	47.16	INSOLUBLE
11	glcNAcase	32	3537.37	11.26	4	0.669	121.88	28.1	SOLUBLE
12	prtP	33	3346.12	10.46	4	0.661	139.09	18.18	INSOLUBLE
13	cpf	27	2685.28	11.17	3	0.9	130	−1.52	SOLUBLE
14	mapA	27	2804.43	10.3	3	0.944	119.26	4.19	INSOLUBLE
15	bacA	28	2864.56	9.63	3	0.779	115	−8.88	SOLUBLE
16	hirJM79	30	3080.89	9.59	4	0.89	139.67	−6.9	INSOLUBLE
17	entC2	27	3036.81	9.5	2	1.73	169.63	49.08	INSOLUBLE
18	blaZ	24	2573.15	9.31	2	1.188	150.42	28.84	SOLUBLE
19	flr	28	2916.47	10	3	0.229	111.79	17.27	INSOLUBLE
20	dex	30	3226.86	9.52	1	0.99	107.31	14.02	SOLUBLE
21	dex	26	2736.4	8	1	1.338	116.54	33.08	INSOLUBLE
22	skc	26	2871.4	8.34	1	0.942	93.85	11.1	INSOLUBLE
23	fnb	27	3025.79	9.7	2	1.256	140.74	6.17	INSOLUBLE
24	ily	33	3486.22	12.02	5	0.648	127.27	45.12	SOLUBLE
25	nanB	29	3068.77	10.29	3	0.831	141.03	25.32	INSOLUBLE
26	–	32	3265.8	10	3	0.375	103.75	5.2	INSOLUBLE
27	speB	27	2884.61	11.26	4	0.948	148.15	22.75	SOLUBLE
28	blm	30	3149.84	9.79	3	0.863	133	−5.92	INSOLUBLE
29	slpA	29	2868.46	10	3	0.731	87.59	23.67	INSOLUBLE
30	Outer cell wall protein	24	2407.89	10	2	1.258	130	15.91	INSOLUBLE
31	jk1070	35	3507.14	12.48	4	0.634	81.14	53.64	INSOLUBLE
32	i-inlF	33	3778.59	9.19	2	0.639	129.7	22.87	INSOLUBLE
33	iap	27	2593.1	10	2	0.941	91.11	10.47	INSOLUBLE
34	padA	29	2996.68	9.52	2	0.834	127.93	17.42	SOLUBLE
35	mag	34	3661.36	10	4	0.406	94.71	7.03	INSOLUBLE
36	fnb	27	3025.79	9.7	2	1.256	140.74	6.17	INSOLUBLE
37	S-layer	30	2962.65	11.17	3	1.157	133.67	47.16	INSOLUBLE
38	etb	31	3468.21	8.25	1	0.742	119.68	40	SOLUBLE
39	entB	27	3085.87	9.99	2	1.663	184.07	44.74	INSOLUBLE
40	SAB1875c	29	3192.8	9.1	2	0.734	114.48	51.49	INSOLUBLE
41	hlgC-like ORF	29	3048.72	9.7	2	0.624	148.28	24.26	INSOLUBLE
42	lip1	34	3616.24	10.46	4	0.276	100.59	32.3	INSOLUBLE

the cleavage of the SP, the mature protein is released extracellularly [13]. Therefore, it is critical to predict precisely the physicochemical and structural characteristics of the SP, particularly the area of the SP cleavage sites when designing constructs for producing recombinant secreted proteins and the functionality of secretion. In order to identify the best SPs to express Ama r 2 in Gram-positive prokaryote systems, we used bioinformatics as an affordable, reliable and potent tool [15].

In this study, all the *Lactococcus* SPs which are available in the SP database, were screened. Prediction of the SP probability and cleavage site of all the 43 SPs was performed using SignalP 4.1, by calculating D-score; however, only a limited number of them had a rational D-score and proper cleavage site to express Ama r 2; therefore, the SPs from other Gram-positive organisms such as *Lactobacillus*, *Streptococcus*, etc. were also investigated. Among around 200 verified signal sequences, 42 ones were screened and selected for further investigation. As mentioned earlier, any signal sequence with a D-score of more than 0.5, is regarded as a SP, otherwise, it should be eliminated. According to our results, all the SP D-scores were above 0.5 (except etb and hirJM79); as a result, all

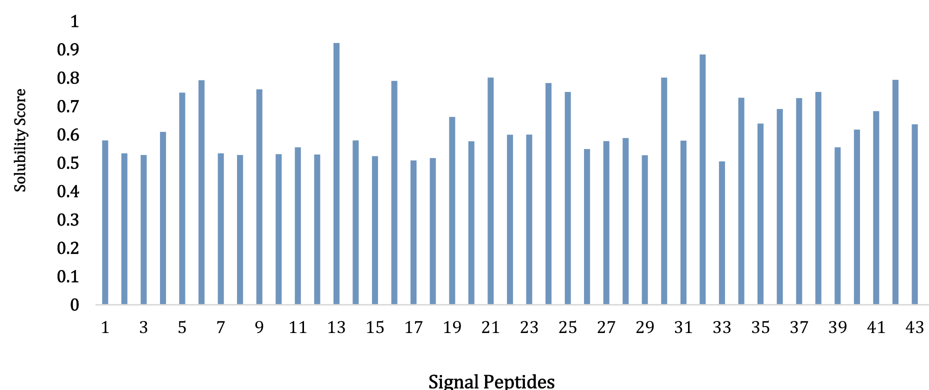

Fig. 1. The solubility scores of SP sequences.

Table 5
Secretion sorting and sub-cellular location of SPs.

No.	Signal peptide	Type of SP	Reliability score	Cytoplasmic	Membrane	Secreted (Extracellular)	Final Prediction Site
1	prtP	Sec	0.977	0.38	0.19	9.12	Extracellular
2	prtP	Sec	0.977	0.52	0.44	8.35	Extracellular
3	–	Sec	0.941	0.51	3.35	5.19	Extracellular
4	yjgB	Sec	0.98	0.77	0.37	5.68	Extracellular
5	310	Sec	0.977	1.81	4.38	3.53	Membrane
6	prtP	Sec	0.977	0.52	0.44	8.35	Extracellular
7	–	Sec	0.96	–	–	–	N.D.[a]
8	slpH	Sec	0.996	–	–	–	N.D.
9	S-layer	Sec	0.996	–	–	–	N.D.
10	S-layer	Sec	0.996	–	–	–	N.D.
11	glcNAcase	Sec	0.988	1.89	7.61	0	Membrane
12	prtP	Sec	0.977	0.38	0.19	9.12	Extracellular
13	cpf	Sec	0.988	–	–	–	N.D.
14	mapA	Sec	0.973	–	–	–	N.D.
15	bacA	Sec	0.97	–	–	–	N.D.
16	hirJM79	Sec	0.918	2.77	0.37	6.6	Extracellular
17	entC2	Sec	0.987	0	0.11	9.68	Extracellular
18	blaZ	Sec	0.982	2.49	1.76	5.25	Extracellular
19	flr	Sec	0.991	–	–	–	N.D.
20	dex	Sec	0.995	0	0.13	9.74	Extracellular
21	dex	Sec	0.967	0.98	3.99	3.99	Extracellular
22	skc	Sec	0.973	0.93	8.78	0	Membrane
23	fnb	Sec	0.96	1.22	8.41	0	Membrane
24	ily	Sec	0.979	–	–	–	N.D.
25	nanB	Sec	0.798	2.3	1.58	6.02	Extracellular
26	–	Sec	0.969	1.02	0	8.98	Extracellular
27	speB	Sec	0.982	–	–	–	N.D.
28	blm	Sec	0.915	2.95	0.61	4.62	Extracellular
29	slpA	Sec	0.984	–	–	–	N.D.
30	Outer cell wall protein	Sec	0.976	–	–	–	N.D.
31	jk1070	Sec	0.976	–	–	–	N.D.
32	i-inlF	Sec	0.809	2.72	5.27	1.59	Membrane
33	iap	Sec	0.995	2.59	0	0.93	Cytoplasmic
34	padA	Sec	0.979	1.44	1.57	0.698	Cytoplasmic
35	mag	sec	0.976	0.63	0.3	8.91	Extracellular
36	fnb	sec	0.96	1.22	8.41	0	Membrane
37	S-layer	sec	0.996	1.99	0	0.39	Cytoplasmic
38	etb	sec	0.849	5.38	4.58	0	Cytoplasmic
39	entB	sec	0.964	0	0.19	9.66	Extracellular
40	SAB1875c	sec	0.982	2.28	0	7.33	Extracellular
41	hlgC-like ORF	sec	0.954	1.76	0	8.24	Extracellular
42	lip1	sec	0.987	3.11	0	5.7	Extracellular

[a] N.D. = Not determined.

Table 6
Sorting the SPs according to amino acid length, net positive charge, GRAVY, aliphatic index, D-score, h-region length and SP final prediction site respectively.

No.	Signal peptide	AA length	Net Positive Charge	GRAVY	Aliphatic index	D-score	h-Region Lengh	Final Prediction Site
4	yjgB	26	3	1.165	135.77	0.829	6, 20 (15)	Extracellular (5.68)
7	–	30	2	0.837	107.67	0.729	11, 24 (14)	N.D.
8	slpH	30	3	1.087	130.67	0.683	7, 19 (13)	N.D.
9	S-layer	30	3	1.157	133.67	0.67	7, 19 (13)	N.D.
10	S-layer	30	3	1.077	127.33	0.673	7, 19 (13)	N.D.
11	glcNAcase	32	4	0.669	121.88	0.625	9, 23 (15)	Membrane
12	cpf	27	3	0.9	130	0.71	8, 21 (14)	N.D.
17	entC2	27	2	1.73	169.63	0.626	7, 20 (14)	Extracellular (9.68)
18	blaZ	24	2	1.188	150.42	0.745	5, 17 (13)	Extracellular (5.25)
20	dex	30	1	0.99	107.31	0.642	8, 22 (15)	Extracellular (9.74)
21	dex	26	1	1.338	116.54	0.771	7, 19 (13)	Extracellular (3.99)
22	skc	26	1	0.942	93.85	0.646	7, 20 (14)	Membrane
24	ily	33	5	0.648	127.27	0.684	14, 26 (13)	N.D.
26	–	32	3	0.375	103.75	0.616	9, 23 (15)	Extracellular (8.98)
27	speB	27	4	0.948	148.15	0.644	10, 20 (11)	N.D.
28	blm	30	3	0.863	133	0.632	9, 21 (13)	Extracellular (4.62)
30	Outer cell wall protein	24	2	1.258	130	0.685	5, 17 (13)	N.D.
33	iap	27	2	0.941	91.11	0.729	7, 20 (14)	Cytoplasmic
34	padA	29	2	0.834	127.93	0.697	8, 22 (15)	Cytoplasmic
37	S-layer	30	3	1.157	133.67	0.67	7, 19 (13)	Cytoplasmic
42	lip1	34	4	0.276	100.59	0.642	13, 27 (15)	Extracellular (5.7)

Table 7
Arrangement of SPs based on the most determinant features.

No.	Signal peptide	AA lengh	Net Positive Charge	GRAVY	Aliphatic index	D-score	h-Region Length	Final Prediction Site
1	yjgB	26	3	1.165	135.77	0.829	15	Extracellular
2	entC2	27	2	1.73	169.63	0.626	14	Extracellular
3	entB	27	2	1.663	184.07	0.523	11	Extracellular
4	blaZ	24	2	1.188	150.42	0.745	13	Extracellular
5	dex	26	1	1.338	116.54	0.771	13	Extracellular
6	blm	30	3	0.863	133	0.632	13	Extracellular
7	n/a (number 26)	32	3	0.375	103.75	0.616	15	Extracellular
8	dex	30	1	0.99	107.31	0.642	15	Extracellular

of them could be regarded as a SP, but for the optimum screening, other features, in particular the c-region or cleavage site of the selection should have been considered. Consequently, 21 SPs whose D-scores were above 0.6 were selected and compared based on the most important features, including Net positive charge, aliphatic index, GRAVY, D-score, h-region length, cleavable site and sub-cellular location (Table 6).

As observed in this study, the amino acid length of the SPs ranged between 24 and 35 residues, which is in agreement with previous studies aiming to analyze the variation in the length of the known SPs and revealed that the SP from gram-positive bacteria tended to be longer (mean length 29–31 residues) than the one from gram-negative bacteria (mean length 24 residues) [18]. Nevertheless, there has been no experimental evidence that shorter SPs are incapable of directing secretion in gram-positive organisms [20]. Furthermore, based on the literature, the median values of Net charge, aliphatic index, PI and GRAVY score for Gram-positive SPs are +3, 75–200, 10.3 and 93.5% respectively [30]. SP n-region interferes in the translocation of a secretory protein; therefore, for maintaining the SP function, n-region requires a positive charge [19]. As shown in Table 6, the Net positive charge of the SPs was calculated in the range of +1 to +4, and all the selected SPs had an appropriate net positive charge. Another important parameter to trim SPs is hydrophobicity which is generally introduced by three parameters, including GRAVY, aliphatic index and length of h-region [30]. In other words, the SPs with high scores for these parameters can be considered to be hydrophobic and consequently, an appropriate SP to be applied. To obtain an efficient SP processing and protein translocation, a sufficient hydrophobicity level of h-region is required. It has been also believed the rate of the protein secretion can be improved by enhancing the hydrophobicity levels and length of the h-region. According to Table 6, there has not been a significant diversity in the lengths of the SP h-regions (13–15 AA). Therefore, other important parameters were used such as aliphatic index and GRAVY in the diagnosis of hydrophobicity. Among the 21 SPs, iap (91.11) and skc (93.85) had the lowest aliphatic index. They also had unacceptable GRAVYs (0.941 and 0.942 respectively). Since the SPs with high scores for both parameters can be considered hydrophobic and consequently, appropriate to be applied, it seems these SPs are not suitable to express Ama r 2 protein. In the case of the other SPs, no significant difference in the aliphatic index was observed; hence, to trim data GRAVY was considered. According to the GRAVY results, the lowest GRAVY belonged to flr (0.229) and lip1 (0.276) respectively. They were then deleted; nevertheless, they had an aliphatic index in range. Among all the remaining SPs, yjgB, slpH, s-layer (numbers 9,10 and 37), entC2, blaZ, dex (number 21) and outer cell wall protein were acceptable in terms of this parameters, among which entC2 with the highest scores for both of the parameters (GRAVY and aliphatic index) could be considered the best SPs. Other important parameters such as the rational length of D-score, n- and h-regions in addition to net positive charge confirmed that this SP had a potential to be introduced as a proper SP to express Ama r 2 protein in prokaryote systems.

Regarding the literature, not all the secreted proteins in Gram-positive bacteria will remain embedded in the cell wall, but rather many proteins exported across the cytoplasmic membrane by the Sec or Tat pathways, will eventually be released into the extracellular environment, often by passive diffusion through the peptidoglycan layer [29]. In general, Gram-negative bacteria which use Sec and SPR pathways, translocate unfolded proteins to the periplasmic compartment where folding and accumulation are both occurring. Conversely, by the use of TAT pathway, they tend to fold secretory proteins in the cytoplasm compartment and then translocate the folded proteins to the periplasmic area for accumulation. It seems Sec and SPR pathways are more vital than TAT Pathway, because folding and purification of secretory proteins are easier in the periplasmic area than in the cytoplasm [17]. The Sec secretion pathway is one of the most conserved mechanisms of protein export and is found in all the classes of bacteria [29], and as shown in Table 5, all the 42 SPs in this study belonged to Sec pathway. S-layer (number 37), SP plus iap and padA were deleted, because they were not able to translocate the folded protein across the cell membrane. It was clarified that among the SPs which have determined final secretion site, yjgB, entC2 (Entrotoxine type C-2), blaZ (Beta lactamase), dex (Dextranase; number 20), blm (Beta lactamase 2), number 26 and dex (number 21) which were respectively ranked based on their final prediction site, could be introduced as the most appropriate SPs to express Ama r 2 in Gram-positive systems and translocate this protein to the extracellular environment. Eventually, by comparing the other important features (Table 6), it must be pointed out that entB SP out of all the 42 SPs, regardless of its D-score < 0.6, with high values of aliphatic index and GRAVY besides being able to secrete the recombinant protein into the extracellular compartment, could be considered the third best SP in this analysis. Table 7 comprises the arrangement of the given SPs based on the most determinant features. C-region or cleavage site was considered to be the final parameter to introduce the most appropriate SPs. In our study, C terminus in all the 8 SPs followed this rule and contained the correct SP cleavage site recognized by signal peptidase; therefore, we have avoided mentioning this parameter in Table 7.

5. Conclusion

Recently, *in silico* approaches such as bioinformatics have been emerged as a way to accelerate the process of analyzing SPs for the production of recombinant proteins. Moreover, it has concurrently improved the production yield, minimized the cost of the expression and purification of recombinant proteins as well as reducing the time required for the process. At the same time, it forecasts reliable outcomes before actual experimental work. Therefore, predicting the best SPs by *in silico* approaches would help biologists and protein engineers accelerate and facilitate the vital projects. The aim of this study was to predict the candidate SPs to express Ama r 2, and it was attempted to evaluate the most important features of the SPs based on the most accurate softwares. Eventually, yjgB, entC2 (Entrotoxine type C-2), ent B (Entrotoxine type), blaZ (Beta lactamase), dex (number 21), blm (Beta lactamase 2), dex (Dextranase; number 20) and number 26 were introduced as the best SPs to express Ama r 2 in *Lactococcus lactis*. Although it should be noted that experimental evaluation is required to

confirm these results.

Declaration of competing interest

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

Acknowledgments

The authors wish to express their profound gratitude sincerely to the Research Deputy of Ferdowsi University of Mashhad for funding this project with the code of 3/42029. In addition, a part of this study was financially supported by grant No: 951110 of the Biotechnology Development Council of the Islamic Republic of Iran.

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.

References

[1] M.V. Moghaddam, M. Fallahpour, M. Mohammadi, F.R. Varaee, K. Mokhtarian, M. Khoshmirsafa, R. Jafari, N. Shirzad, R. Falak, Identification of polcalcin as a novel allergen of Amaranthus retroflexus pollen, Allergol. Immunopathol. 47 (4) (2019) 357–364.
[2] P. Morakabati, M.-A. Assarehzadegan, G.R. Khosravi, B. Akbari, F. Dousti, Cloning and expression of Ama r 1, as a novel allergen of Amaranthus retroflexus pollen, J. Allergy (2016) 4092817.
[3] M. Tehrani, M. Sankian, M.A. Assarehzadegan, R. Falak, R. Noorbakhsh, M. Moghadam, F. Jabbari, A. Varasteh, Identification of a new allergen from Amaranthus retroflexus pollen, Ama r 2, Allergol. Int. 60 (2011) 309–316.
[4] M. Villalba, R. Barderas, S. Mas, C. Colás, E. Batanero, R. Rodriguez, Amaranthaceae pollens: review of an emerging allergy in the mediterranean area, J Investig. Allergol. Clin. Immunol. 24 (2014) 371–381.
[5] M. Fereidouni, R.F. Hossini, F.J. Azad, M.A. Assarezadegan, A. Varasteh, Skin prick test reactivity to common aeroallergens among allergic rhinitis patients in Iran, Allergol. Immunopathol. 37 (2009) 73–79.
[6] M. Tehrani, M. Sankian, M.A. Assarehzadegan, R. Falak, F. Jabbari, A. Varasteh, Immunochemical characterization of Amaranthus retroflexus pollen extract: extensive cross-reactive allergenic components among the four species of Amaranthaceae/Chenopodiaceae, Iran. J. Allergy, Asthma Immunol. (2010) 87–95.
[7] M.D. Chapman, A.M. Smith, L.D. Vailes, L.K. Arruda, V. Dhanaraj, A. Pomés, Recombinant allergens for diagnosis and therapy of allergic disease, J. Allergy Clin. Immunol. 106 (2000) 409–418.
[8] Z. Ghasemi, A.-R. Varasteh, M. Moghadam, F. Sedghi, F. Jabbariazad, M. Sankian, Production of recombinant protein of salsola kali (sal k1) pollen allergen in Lactococcus lactis, Iran. J. Allergy, Asthma Immunol. 17 (2018) 134–143.
[9] J. Glenting, L.K. Poulsen, K. Kato, S.M. Madsen, H. Frøkiær, C. Wendt, H.W. Sørensen, Production of recombinant peanut allergen Ara h 2 using Lactococcus lactis, Microb. Cell Factories 6 (2007) 28.
[10] L.R. Nasiraie, F. Tabatabaie, M. Sankian, F. Shahidi, A. Varasteh, Construction of a recombinant allergen-producing probiotic bacterial strain: introduction of a new line for a live oral vaccine against Chenopodium album pollen allergy, Reports.Biochem.Mol. Biol. 2 (2013) 16.
[11] G. Liu, H. Wang, M.W. Griffiths, P. Li, Heterologous extracellular production of enterocin P in Lactococcus lactis by a food-grade expression system, Eur. Food Res. Technol. 233 (2011) 123–129.
[12] T. Maischberger, I. Mierau, C.K. Peterbauer, J. Hugenholtz, D. Haltrich, High-level expression of Lactobacillus β-galactosidases in Lactococcus lactis using the food-grade, nisin-controlled expression system NICE, J. Agric. Food Chem. 58 (2010) 2279–2287.
[13] D.T. Ng, C.A. Sarkar, Engineering signal peptides for enhanced protein secretion from Lactococcus lactis, Appl. Environ. Microbiol. 79 (2013) 347–356.
[14] I. Quintana, M. Espariz, S.R. Villar, F.B. González, M.F. Pacini, G. Cabrera, I. Bontempi, E. Prochetto, J. Stülke, A.R. Perez, Genetic engineering of Lactococcus lactis co-producing antigen and the mucosal adjuvant 3′ 5′-cyclic di adenosine monophosphate (c-di-AMP) as a design strategy to develop a mucosal vaccine prototype, Front. Microbiol. 9 (2018).
[15] A. Baradaran, C.C. Sieo, H.L. Foo, R.M. Illias, K. Yusoff, R.A. Rahim, Cloning and in silico characterization of two signal peptides from Pediococcus pentosaceus and their function for the secretion of heterologous protein in Lactococcus lactis, Biotechnol. Lett. 35 (2013) 233–238.
[16] M.S. Dastjerdeh, M. Marashian, M.B. Boroujeni, H. Rahimi, In silico analysis of different signal peptides for the secretory production of recombinant human keratinocyte growth factor in Escherichia coli, Comput. Biol. Chem. 80 (2019) 225–233.
[17] A. Forouharmehr, M. Nassiri, S. Ghovvati, A. Javadmanesh, Evaluation of different signal peptides for secretory production of recombinant bovine pancreatic ribonuclease A in Gram negative bacterial system: an in silico study, Curr. Proteomics 15 (2018) 24–33.
[18] Y. Le Loir, S. Nouaille, J. Commissaire, L. Bretigny, A. Gruss, P. Langella, Signal peptide and propeptide optimization for heterologous protein secretion in Lactococcus lactis, Appl. Environ. Microbiol. 67 (2001) 4119–4127.
[19] M. Zamani, N. Nezafat, M. Negahdaripour, F. Dabbagh, Y. Ghasemi, In silico evaluation of different signal peptides for the secretory production of human growth hormone in E. coli, Int. J. Pept. Res. Ther. 21 (2015) 261–268.
[20] M. van Asseldonk, W.M. de Vos, G. Simons, Functional analysis of the Lactococcus lactis usp45 secretion signal in the secretion of a homologous proteinase and a heterologous α-amylase, Mol. Gen. Genet. MGG 240 (1993) 428–434.
[21] J. Arnau, E. Hjerl-Hansen, H. Israelsen, Heterologous gene expression of bovine plasmin in Lactococcus lactis, Appl. Microbiol. Biotechnol. 48 (1997) 331–338.
[22] J.-M. Chatel, P. Langella, K. Adel-Patient, J. Commissaire, J.-M. Wal, G. Corthier, Induction of mucosal immune response after intranasal or oral inoculation of mice with Lactococcus lactis producing bovine beta-lactoglobulin, Clin. Diagn. Lab. Immunol. 8 (2001) 545–551.
[23] S. Drouault, G. Corthier, S.D. Ehrlich, P. Renault, Expression of the Staphylococcus hyicus Lipase inLactococcus lactis, Appl. Environ. Microbiol. 66 (2000) 588–598.
[24] T.N. Petersen, S. Brunak, G. Von Heijne, H. Nielsen, SignalP 4.0: discriminating signal peptides from transmembrane regions, Nat. Methods 8 (2011) 785.
[25] C.N. Magnan, A. Randall, P. Baldi, SOLpro: accurate sequence-based prediction of protein solubility, Bioinformatics 25 (2009) 2200–2207.
[26] P.G. Bagos, E.P. Nikolaou, T.D. Liakopoulos, K.D. Tsirigos, Combined prediction of Tat and Sec signal peptides with hidden Markov models, Bioinformatics 26 (2010) 2811–2817.
[27] F. Vahedi, M. Nassiri, S. Ghovvati, A. Javadmanesh, Evaluation of different signal peptides using bioinformatics tools to express recombinant erythropoietin in mammalian cells, Int. J. Pept. Res. Ther. (2018) 1–7.
[28] M. Negahdaripour, N. Nezafat, N. Hajighahramani, S. Soheil Rahmatabadi, M. Hossein Morowvat, Y. Ghasemi, In silico study of different signal peptides for secretory production of interleukin-11 in Escherichia coli, Curr. Proteomics 14 (2017) 112–121.
[29] E.R. Green, J. Mecsas, Bacterial secretion systems–an overview, Microbiol. Spectr. 4 (2016).
[30] K.H. Choo, S. Ranganathan, Flanking signal and mature peptide residues influence signal peptide cleavage, BMC Bioinf. 9 (2008) S15.