In Silico Analysis of a Chimeric Protein as Alternative Antimicrobial Against Zoonotic Pathogenic Bacteria

Khadeeja S. Madhi, Mohammed H. Khudor*, Rasha M. Othman
College of Veterinary Medicine, University of Basrah, Basrah, Iraq

*Corresponding Author(s):
Name: Mohammed H. Khudor
Email Address: mohamed.khudor@uobasrah.edu.iq
ORCID ID: https://orcid.org/0000-0002-1456-3344

Received: Oct. 1, 2021; Accepted: Nov. 10, 2021

Abstract

The focus of the present study was to characterize chimeric synthetic plantaricin F which naturally produced by Lactobacillus plantarum against zoonotic pathogenic bacteria Staphylococcus aureus and Escherichia coli as antibacterial peptide. The synthetic bacteriocin by bioinformatics revealed higher stability under studied parameter, hence was taken up for further investigation. The amino acids of bacteriocin from L. plantarum were analyzed by SnapGene. Further, synthetic PLNF was characterized in silico. The translated partial amino acid sequence of the synthetic PLNF gene displayed 253 amino acids for whole and 148 without tag. The predicted properties of the peptide included theoretical isoelectric point (pI) and hydrophobicity was highly acidic. Molecular weight was 27.2KDa for whole protein and 15.8 KDa for without tag. Predication the molecular approach of using SnapGene software and the protein was having antingcity against bacteria and has B-cell epitope on the surface of protein. Prediction data base on characterization of bacteriocin is novel and predicts synthetic PLNF to be a peptide responsible for antimicrobial activity. The study provides information about a broad spectrum bacteriocin in native probiotic culture and paves way towards its application as alternative natural antimicrobial agent against zoonotic pathogenic bacteria. Finally, the 3D peptide structure analysis in present study showed that the predict structure of model and has more functional properties and probably the form most suitable for binding to bacterial cell walls.

Keywords: Lactobacillus plantarum, plantaricin, antimicrobial agents, bioinformatics
Introduction

The constant development of bacterial resistance to antibiotic is widely recognized as a potential threat to public health. The antimicrobial peptides are broad spectrum and small molecular weight compounds with antagonistic activity against viruses, bacteria, and fungi (1). Bacteriocins are antimicrobial peptides presently and beneath consideration equally actual alternatives or complements to community antibiotics (2). Bacteriocin of *Lactobacillus* bacteria are possessive antibacterial activity against species closely related to the producer strains or a wider range of microorganisms (3). The bacteriocin peptides that are produced by *Lactobacillus plantarum* is recognized as plantaricin and is commonly reported as a class II bacteriocin, a very wide-ranging class with a variety of bactericidal and/or bacteriostatic mechanisms (4). Plantaricins (A, E, F, J and K) produced by *L. plantarum* (5). Plantaricin F (pln F) is mostly used to prevent microbial contamination (6). Artificial construction gene is the process of synthesizing a gene *in vitro* without the requirement for initial template DNA samples, rapidly step with cost in effect alternative to cloning traditional technique and other molecular biology procedures (7). Moreover, the use of synthetic genes may constitute a successful approach for the heterologous production and functional expression of bacteriocins by recombinant technique (8). The detection of efficient, safe, and stable agents able to control pathogenic microorganisms is urgent (9). Bioinformatics is the use of computers in the field of biology to analyze amounts of data and generate new hypothesis. Protein sequence, structure, function and interactions can be studied with some current bioinformatics programs (10). The aim of the present study was bioinformatics analysis of chimeric recombinant plantaricin F that was designed and produced by amino acid optimization approach as antibacterial agent against zoonotic pathogenic bacteria.

Materials and Methods

In the present study, different bioinformatics tools and databases were used for molecular modeling of engineered chimeric protein plantaricin F. Homology modeling procedure was performed in four basic sequential steps: Template selection, target template alignment, model construction, and model assessment. ProSA tool was used to (11).

Design and Construction of Engineered Chimeric Protein

We ordered the nucleotide sequences of desired protein synthesized of plantaricin F (PLNF) gene according to the amino acids sequencing:

MGSSHHHHHHHHSSGLVPRGSHMASMTG
GQQMGRGSEFDKIIIILTDDSDFDTDV
LKADGAILVDFAEWCGPCKMIAPILD
GIRGIPTLLFLKNGEVAATKVGALSKG

Physical and Chemical Parameters

The physical and chemical properties of synthetic plantaricin F, including its theoretical isoelectric point (pI), estimated half-life, instability index, aliphatic index and lower grand average of hydropath city (GRAVY). They were predicted by using ExPASy ProtParam online software http://web.expasy.org/protparam/ (14).

B-Cell Epitope Predication

Predict linear B-cell epitopes using IEDB Analysis Resource at http://tools.iedb.org/bcell/. It is possible to predict linear B-cell epitopes when threshold at 0.5 as mention previously (15).

Predication Antigenicity

Predict protein antigenicity of protein was done by sing VaxiJen server http://www.ddgpharmfac.net/vaxijen/VaxiJen/VaxiJen.html (16).

Protein Solubility Predication

Predict protein solubility of synthetic plantaricin F was done by using http://service.tartaglialab.com/new_submission/ccsol (17).

Three Dimensional Strecture Analyses

The 3D of 253 different amino acids of synthetic PLNF model was done by homology modelling using chimera (1.15rc-software) (18). Inundation, the World wide Protein Data Bank (wwPDB) which have maintained the largest, public and free archive of macromolecular structural data (www.wwpdb.org) beside the RCSB Protein Data Bank: which is a powerful new tools for exploring 3D structures of biological macromolecules (19).

Results

Protein production was confirmed by using SnapGene for determine the amino acids sequencing and fusion tags (Figure 1).
Figure 1: Nucleotide and amino Acid sequence of fused Recombinant synthetic PLNF analysis by SnapGene, The Gray color is thioredoxin-(His)6tag fusion partner protein followed with S-Tag and enterokinase cleavage site (1-165 aa), the Red color is leader peptide of pln F (166-183 aa), the Brown color is mature peptide of pln F (184-217 aa), the asterix is stop codon.

Physical and Chemical Parameters

Synthetic plantaricin was predicted to be a stable, molecular 27.2 KDa according to SnapGene, chimeric protein with an estimated half-life of more than 10 h in E. coli. The aliphatic and grand averages of hydropathcity (GRAVY) indices were to be for whole bacteriocin - 0.364 and -0.228- respectively - The physicochemical properties of the protein and the results showed in table 1.

B-Cell Epitope Prediction for Synthetic Plantaricin F

The protein sequence was analyzed using Bepipred linear epitope prediction, surface accessibility and to assess the antigenicity respectively. The average for the whole protein was 0.520 and without tag was 0.489. All values equal or greater than the default threshold 0.5 were predicted to be a potential B cell binder (Figure 2 and 3). The
yellow region was a positive prediction of B cell epitope and the green region was negative.

Predict The Antigenicity

VaxiJen v2.0 was used as bioinformatics tool to analyze antigenicity index of synthetic PLNF protein. The threshold for these proteins was considered 0.4 and the synthetic was more than 0.4 (Table 2).

Table 1: The analysis result physicochemical properties of synthetic PLNF protein.

<table>
<thead>
<tr>
<th>Properties</th>
<th>Whole construct</th>
<th>Without tags</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of amino acids</td>
<td>253</td>
<td>148</td>
</tr>
<tr>
<td>Molecular weight</td>
<td>27285.80</td>
<td>15878.06</td>
</tr>
<tr>
<td>Theoretical pI</td>
<td>6.64</td>
<td>6.11</td>
</tr>
<tr>
<td>Formula</td>
<td>C_{1192}H_{1861}N_{351}O_{360}S_{1}</td>
<td>C_{699}H_{1096}N_{196}O_{21}</td>
</tr>
<tr>
<td>Total number of atoms</td>
<td>3777</td>
<td>2210</td>
</tr>
<tr>
<td>Yeast in vitro Estimated half-life</td>
<td>&gt;20 hours</td>
<td>&gt;20 hours</td>
</tr>
<tr>
<td>Escherichia coli Instability index (state)</td>
<td>&gt;10 hours</td>
<td>&gt;10 hours</td>
</tr>
<tr>
<td>Aliphatic index</td>
<td>73.72</td>
<td>81.15</td>
</tr>
<tr>
<td>Grand average of hydropathcity GRAVY</td>
<td>-0.364</td>
<td>-0.228</td>
</tr>
</tbody>
</table>

Table 2: Antigenicity index synthetic PLNF protein

<table>
<thead>
<tr>
<th>Properties</th>
<th>Whole construct</th>
<th>Without tags</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall Prediction for the protective antigen (antigen/non-antigen)</td>
<td>0.4867 (Probable ANTIGEN)</td>
<td>0.5303 (Probable ANTIGEN)</td>
</tr>
</tbody>
</table>
Predict The Solubility

Predict the solubility propensity of the purified the ccSol was applied. The possible expect results were demonstrated -solubility of synthetic PLNF was 84% with tag and 96% without tag (Figures 4and 5). Moreover, the ccsol omics bioinformatics website was gaven the probable results to predict the solubility of expression PLNF was 92% with tag and 98% without tag (Figures 6 and 7).

Three Dimensional Structures

The amino acid sequences of 253 chimeric synthetic plantaricin F was obtained using an in silico sequence conversion tool. The amino acid sequences were analyzed by SnapGene software. The 3D structures of 253 chimeric proteins were compared with the 3D structures using Chimera (Figure 8). The 3D qualified structure prediction of the chimeric protein was detected using the Robetta server tool [https://robetta.bakerlab.org/](https://robetta.bakerlab.org/) (Figure 9). Figure 10 and 11 were provided qualified 3D structures of the synthetic PLNF protein. The UCSC Chimera was used to measure the root-mean-square deviation of atomic positions (RMSD) between the whole protein and the recombinant synthetic PLNF protein without the mentioned tags (Figure 12).
Figure 2: The linear B-cell Epitopes results for the whole protein. Average: 0.520  Minimum: 0.253  Maximum: 0.713.

Figure 3: The linear B-cell epitopes results for protein without tags. Average: 0.489 , Minimum: 0.238  Maximum: 0.706

Figure 4: Solubility propensity for synthetic PLNF without the tags.

Figure 5 : Solubility propensity results for the whole Synthetic PLNF

Figure 6: The probable solubility for synthetic PLNF with Tags. Inset: Overall score distribution for soluble (red) and insoluble (blue) proteins

Figure 7: The probable solubility for synthetic PLNF without tags. Inset: Overall score distribution for soluble (red) and insoluble (blue) proteins

Figure 9: Protein 3D structure alignments generated by the UCSF Chimera package: The 3D structure in cartoon mode whole synthetic PLNF as modeled by Robetta server. The cartoon structure color was cyan.

Figure 8: The 3D structure of surface mode for whole synthetic PLNF as modeled by Robetta server. The structure was staining in cyan color.

Figure 10: The top structure in cartoon mode and the bottom structure is Ribbon mode explains the 3D structure of the synthetic PLNF as modeled by Robetta server. His-tag in red, thrombin site in smudge, T7 tag in deep purple, S-tag in white, enterokinase site in cyan, and the rest of the protein in green color.
Figure 11: The 3D structure of the synthetic PLNF without the tags as modeled by Robetta server. The (A) structure was in cartoon mode. However, the (B) structure represented the surface mode. The color is cyan.

Figure 12: The figures were the same but in different angles. The measured RMSD1 was 1.12 angstrom which was indicative of no substantial changes in the 3D Structure of the whole protein while harbouring the Tags. The Structures are in Ribbon mode, and the whole protein represented in white while the protein without the tags is in green. In bioinformatics, the root-mean-square deviation of atomic positions, or simply root-mean-square deviation (RMSD), is the measure of the average distance between the atoms (usually the backbone atoms) of superimposed proteins.
Discussion:

Recombinant DNA technical methods have permitted fusions of genes in a simple way (20). Designing recombinant protein via bioinformatics approach is comparatively stable, safe, inexpensive, specific and more effective. Thus, herein, we made a chimeric protein plantaricin F for the 1st time. The designed fusion synthetic plantaricin F gene was constructed and cloned into pUC57 vector and transformed into cloning host cell E. coli BL21(21).

The results of protein analysis of sequence of amino acid of synthetic PLNF were investigated by using the SnapGene software, SnapGene allows one to view the default sequence, commercially tag sequence classes, and change the view settings(Figure1). The allowance of such capabilities might be integrated with much larger data sets or genotypes and parallel processing may help offload latency times for creating output visualization (22). The mature sequence VFHAYSARGVRNNYKSAVGPADWIISAV RGFIHG of synthetic PLNF was similarity 100% to pervious study of (23). However, fusion tags include His tag, T7 tag, Trx tag and S tag used in E. coli to improve protein production yields, solubility and folding, and to facilitate protein purification. They can also confer specific properties for target proteins characterization. Fusion partners can also be used when producing toxic proteins for production of antimicrobial peptides by E. coli agree with Costa et al., (24). On the other hand, the ProtParam analysis showed that detected the synthetic PLNF sequence, in which the molecular weight was calculated by the addition of average isotopic masses of amino acids, the molecular weight was 27.2 KDa which provides information on the full-length of protein expression, expression of modified, and cleavage products(Table1). The instability index showed stable in nature considered to be stable (instability index <40). The theoretical isoelectric point (pI) was less than 7 (pI < 7) which indicates synthetic PLNF was considered highly acidic in nature. The computed isoelectric point will be useful for developing a buffer system for purification by isoelectric focusing method (25). The hydrophilic nature of the protein and the possibility of better interaction with water indicated by GRAVY indices value (-0.364, -0.288) was highly acidic (Table2) (26).

The results antigenicity revealed all values were more than the threshold 0.4 (Table3). This finding agreement with previously study (27). It has been found that synthetic PLNF have antigenicity against bacteria. While, the synthetic PLNF showed a high solubility(Figures 3 and 4) which indicate another positive feature of the synthetic PLNF that prevents the accumulation of protein in the inclusion and was used to confirm the full-
length of synthetic PLNF suitability for expression in *E. coli* and solubility propensity also indicated that a less accumulation of proteins in the inclusion bodies with the addition of restriction enzymes and a N-terminal his tag, as shown in figures 5 and 6, which also revealed the probable solubility was high heterologous expression for synthetic PLNF (red) soluble in compare with insoluble (blue) protein, and this finding agree with (20,28). On other hand, the graphic revealed all values equal to or greater than the threshold at 0.5. In which the Y-axis defines for each amino acid residual score, however, X is the position of the amino acid sequence residue(Figures 7 and 8) The results were showed specific areas in proteins that bind to the B cell receptor (29). Additionally, UCSF chimera was an extensible program for interactive visualization and analysis of molecular structures. The 3D peptide structure analysis in present study showed that the predicted structure of model has more functional properties and probably the form most suitable for binding to bacterial cell walls (30).

**Conclusion**

Bioinformatics approaches have ability to analyze bacteriocin avoiding the disadvantages of the conventional methods reduce time, cost required, enhance the safety and efficacy of the predicted epitopes. Several epitopes were predicted in the current study to study chimeric synthetic PLNF against zoonotic pathogenic bacteria using in silico prediction tools. Bioinformatics is recommended for further *in vitro* studies for confirmation and synthetic PLNF can be used as alternative antibacterial agents against pathogenic bacteria in medicine and veterinary medicine.

**Acknowledgements**

Authors thankful to the staff of Department of Medical Biotechnology, College of Medical Science, University of Tarbiat Modares, Tehran, Iran.

**Conflict of Interest**

The authors state that there is no conflict of interest.

**References**


13. Khalil, M. I. (2020, November). CRISPRCas9 design to knockout and


23. Widyasari, E. M., Mustopa, A. Z., Daruwati, I., Sriyani, M. E., Sugiharti, R. J., and


تحليل السيليكى لبروتين مهجن كمضاد بكتيري بديل ضد البكتريا المسببة للأمراض في الإنسان و الحيوان

خديجة سامي ماضي ، محمد حسن خضر ، رشا منذر عثمان
كلية الطب البيطري ، جامعة البصرة ، البصرة ، العراق

الخلاصة

تركزت الدراسة الحالية على توصيف بلانتيرسين F الكيميائي الاصطناعي الذي ينتج بشكل طبيعي بواسطة Staphylococcus aureus ضد البكتريا المسببة للأمراض في الإنسان والحيوان Lactobacillus plantarum كبيتيد مضاد. أظهرت النتائج من خلال دراسة المعلوماتية الحيوية توافقًا أعلى للبيتيد مع الاسم Escherichia coli المستخدم. تم تحليل الأحماض الأمينية للبكتريوسين من PLNF بواسطة L. plantarum بواسطة SnapGene. تم اجراء تسلسل للأحماض الأمينية الجزئية المترجمة لجين PLNF تميز الاصطناعي. فكانت 253 حامض اميني للبيتيد الكامل مع الاسم و 148 حامض اميني للبيتيد الكامل بدون اسم. أظهرت النتائج أنه بيتيد شديدة الحموضة. كان الوزن الجزيئي للبكتريوسين الكامل 27.2 كيلو دالتنو مع الاسم اما بدونه فكان 15.8 كيلو دالتنو. يتوقع أن يكون PLNF الأصطناعي عبارة عن بيتيد مسؤول عن نشاط مضاد واسع للميكروبات. توفر هذه الدراسة معلومات حول مجموعة واسعة من البكتريوسين التي لها فعالية حيوية ضد الميكروبات وتمهد الطريق نحو استخدام هذا البيتيد كعامل طبيعي بديل مضاد للميكروبات ضد البكتريا المسببة للأمراض في الإنسان والحيوان. أخيرًا، أظهرت التحليل ثلاثي الابعاد للبيتيد في دراستنا أن له خصائص وظيفية متفاوتة وربما يكون الشكل الأكثر ملاءمة للارتباط بجدران الخلايا البكتيرية ومن ثم دمثيرها.

الكلمات المفتاحية: بلانتيرسين، العوامل المضادة للميكروبات، المعلوماتية

Lactobacillus plantarum