

Research Article

Lactobacillus acidophilus: Boosting Immunity Through IL-6 Induction

Isaac Oluseun Adejumo¹

1. University of Ibadan, Nigeria

Probiotics, live microorganisms that promote health when consumed in adequate amount and ensure the balance of bacterial composition in the digestive system, suppress harmful pathogenic bacteria, with overall implications for animal and human health, welfare and performance. However, a lot remains unclear about their functional mechanisms. In this study, 14 uncharacterized proteins of *Lactobacillus acidophilus* were analyzed for subcellular localization, structural and safety profiling and interleukin-6- (IL-6)-inducing potential. Aliphatic index scores were generally high, ranging between 138.39 (LBA1705) and 78.39 (LBA1825). The instability index scores were less than 40 for all the query proteins except for LBA0995. All the proteins produce immunogenic IL-6-inducing peptides except for LBA0037, LBA1825 and LBA1788. The findings provide insight into understanding the functional mechanism of probiotic *Lactobacillus*, laying a strong foundation for more experimental studies.

Corresponding author: Isaac Oluseun Adejumo, smogisaac@gmail.com

Introduction

Animal and human nutrition is confronted with many challenges. The rising cost of feed ingredients is a major challenge confronting animal production [1], while diseases remain the main challenge for animal production and human existence. Over the years, several attempts have been made to overcome these challenges, including the use of antibiotics, although the use of antibiotics in feed is no longer widely accepted owing to the potential risk of antibiotic resistance as well as the concerns of food safety arising from contamination or residual effects of chemicals and drugs in animal foods [2][3]. This has necessitated the shift of attention to probiotics to enhance growth performance in livestock, especially poultry, as well as to eliminate the risk of health hazards of antimicrobial drugs and vaccines [4].

Probiotics are live microorganisms that, when administered in adequate amounts, confer a health benefit on the host [5][6]. They promote health and ensure the balance of bacterial composition in the digestive system, leading to the suppression of harmful pathogenic bacteria [7][8], with overall implications for animal health, welfare and performance. The documented benefits of probiotics include combating enteric pathogenic bacteria, improving growth performance, modulating the intestinal microflora, improving immune responses and improving feed conversion efficiency [9][10]. A high proportion of probiotic bacteria in the intestinal tract can form colonies that may suppress pathogenic bacteria, thereby protecting the intestinal tract from invasion by pathogenic bacteria [11][12]. The benefits probiotics offer have placed pressure on their demands for use in animal nutrition and as human supplements, for cosmetic and therapeutic purposes.

Among the lactic acid bacteria, *Lactobacillus* has been proposed as the main candidate strain for probiotic potential [13][14]. A recent study reported that bacterial isolates belonging to *Lactobacillus plantarum* and *Bacillus subtilis* were the most promising probiotic agents in chicken nutrition among others tested [15], although further studies to assess the safety and protective benefits of these probiotic candidates are recommended [15].

Improved nutrient absorption and digestibility, enterotoxin neutralization, improved immune responses, reduced gastrointestinal colonization by foodborne pathogenic organisms (*Campylobacter*, *Salmonella*, *Clostridium* and *Escherichia coli*) and improved growth performance are some of the benefits of *Lactobacillus* probiotic strains [16][17][18]. However, to ensure the safety and sustainability of probiotics in nutrition and therapeutic applications, it is necessary to unravel most of the doubts associated with the constituents of the microorganism. In addition, many proteins from the *Lactobacillus acidophilus* genome that may play important roles and assist in the biological understanding of the organism are uncharacterized. Uncharacterized proteins are proteins that are predicted to be expressed in an organism but to which no proper function has been assigned or known [19]. Uncharacterized proteins have not been assigned any functions experimentally because they have not been studied or their functions are not yet known.

A lot remains unclear about the characteristics, safety profile and functional mechanisms of probiotics, owing in parts to many of their proteins being uncharacterized or hypothetical. Understanding the proteins of *Lactobacillus acidophilus* genome will shed more light on its functional mechanisms, safety profile and sustainable applications in animal and human nutrition and therapeutic and cosmetic products. Therefore, this study was carried out to stimulate interests in and lay foundation for in vivo functional characterization of these proteins in order to better understand their functional mechanisms and thereby ensuring their

sustainable applications in animal and human nutrition, cosmetic and therapeutic products. Investigating the selected uncharacterized proteins of uncharacterized proteins of a reference genome of *Lactobacillus acidophilus* may provide insights that may find applications in the animal nutrition, health, food production and nutraceutical industries.

Materials and Methods

Sequence retrieval and analysis of protein families

The reference genome of *Lactobacillus acidophilus* (CP005926.2) was retrieved from NCBI and the protein families were determined in Pathosystems Resource Integration Center (PATRIC). The protein families were sorted into different classes and 14 uncharacterized proteins were randomly selected. The proteins are: uncharacterized membrane protein YfhO-like (LBA1510), uncharacterized metal-dependent hydrolase YcfH (LBA0214), uncharacterized MFS-type transporter (LBA1705), uncharacterized MFS-type transporter (LBA1446), uncharacterized DUF1113 membrane protein family (LBA0037), uncharacterized protein CAC3725 (LBA1825), uncharacterized GabP-family amino acid permease LBA0995 (LBA0995), uncharacterized transcriptional regulator YozG, Cro/CI family (LBA1788), uncharacterized transporter YbxG (yifk), uncharacterized transporter YbxG (LBA0899), uncharacterized UPF0750 membrane protein SPy2155 (LBA1209), uncharacterized UPF0750 membrane protein YpjC (LBA0972), uncharacterized substrate:H⁺ symporter, LctP family (LctP) and uncharacterized UPF0721 integral membrane protein (LBA0338). Their nucleotide and amino acid sequences were retrieved from PATRIC. The selected proteins are presented in Table 1.

Designation	Description	AA Mean	Std Dev
LBA1510	Uncharacterized membrane protein YfhO-like	859	0
LBA0214	Uncharacterized metal-dependent hydrolase YcfH	255	0
LBA1705	Uncharacterized MFS-type transporter	93	0
LBA1446	Uncharacterized MFS-type transporter	490	0
LBA0037	Uncharacterized DUF1113 membrane protein family	265	0
LBA1825	Uncharacterized protein CAC3725	87	0
LBA0995	Uncharacterized GabP-family amino acid permease LBA0995	490	0
LBA1788	Uncharacterized transcriptional regulator YozG, Cro/CI family	69	0
yifk	Uncharacterized transporter YbxG (peg.1837)	460	3
LBA0899	Uncharacterized transporter YbxG (peg.869)	463	0
LBA1209	Uncharacterized UPF0750 membrane protein SPy2155	292	0
LBA0972	Uncharacterized UPF0750 membrane protein YpjC	295	0
LctP	Uncharacterized substrate:H ⁺ symporter, LctP family	504	0
LBA0338	Uncharacterized UPF0721 integral membrane protein	252	0

Table 1. Protein family identities, description and statistics of query uncharacterized proteins

AA = amino acid; std = standard deviation

Blast and subcellular localization analysis

The retrieved sequences of the selected protein were evaluated for similarity check using NCBI BLAST (https://blast.ncbi.nlm.nih.gov/Blast.cgi#alnHdr_1854278943). The subcellular localization of the query proteins was analysed using PSORTb v3.0.2, <https://www.psorth.org/psorth/> [20].

Structural determination, physico-chemical properties and protein protein network

The secondary structures and three-dimensional tertiary structures of the query proteins were determined using the SOPMA server [21] and AlphaFold 2.0, <https://alphafold.ebi.ac.uk/> [22], respectively. SWISS-MODEL

Interactive Workspace (<https://swissmodel.expasy.org/assess>) [22] was used to evaluate the obtained three-dimensional tertiary structures. ExPASy ProParam, www.web.expasy.org/protparam [23] was used for the analysis of the physicochemical properties of the query proteins. Protein protein network interactions of the query proteins were analysed using STRING, a database of known and predicted protein interactions, drawn from genomic context, high-throughput experiments, conserved coexpression, and previous PubMed literature [24].

Allergenicity, immunogenicity, and toxicity analyses

Allergenicity evaluation of the query proteins used in this study was performed with AllerCatPro 2.0, allercatpro.bii.a-star.edu.sg [25], while immunogenicity and toxicity evaluations of the query proteins of the reference genome were performed using VaxiJen 3.0, <https://www.ddg-pharmfac.net/vaxijen3/> [26] and ToxinPred, (<https://webs.iitd.edu.in/raghava/toxinpred2/>) [27]. AllerCatPro 2.0 is a version of the previously established AllerCatPro method (AllerCatPro 1.7) [28]. The tool predicts protein allergenicity potential using the similarity of both amino acid sequences and 3D structures to the most comprehensive datasets (4979 protein allergens, 162 low allergenic proteins, and 165 autoimmune allergens) of reliable proteins associated with allergenicity from the WHO/IUIS, COMPARE, FARRP, UniProtKB and Allergome databases [25].

Analysis of active binding site of query proteins, gene expression and multiple sequence alignment

CASTp 3.0, <http://sts.bioe.uic.edu/castp/> [29] was used to determine the active binding sites of the query proteins used in the study. The CASTp server is a recognized bioinformatic tool used for identifying the identities of all interior cavities, surface pockets and cross channels in protein structures. It also provides a detailed delineation of all atoms participating in their formation [29]. IL-6-inducing potential of selected proteins was evaluated using il6pred (<https://webs.iitd.edu.in/raghava/il6pred/display.php?ran=36676>). The top 10 based on their prediction values were further evaluated for immunogenicity [26]. Gene expressions of IL-6 across various conditions in *Homo sapiens* (platform: HS_AFFY_U133PLUS_2-0) and *Mus musculus* (platform: MM_AFFY_430_2-1) was performed using GENEVESTIGATOR.

Results

The features, descriptions and designations of the query proteins of the reference genome of *Lactobacillus acidophilus* used in this study are presented in Table 1. The amino acid length ranged between 69 (LBA1788)

and 859 (LBA1510). The BLAST results revealed that the selected proteins are similar, showing almost 100% identity across different *Lactobacillus acidophilus* strains.

Physicochemical properties of query proteins

The atomic formulae of the selected proteins were $C_{4693}H_{7164}N_{1136}O_{1222}S_{17}$ (LBA1510), $C_{1326}H_{2055}N_{349}O_{391}S_7$ (LBA0214), $C_{507}H_{792}N_{110}O_{115}S_3$ (LBA1705), $C_{2448}H_{3927}N_{603}O_{652}S_{23}$ (LBA1446), $C_{1462}H_{2229}N_{363}O_{366}S_9$ (LBA0037), $C_{444}H_{734}N_{130}O_{120}S_9$ (LBA1825), $C_{2635}H_{3966}N_{622}O_{656}S_{31}$ (LBA0995), $C_{343}H_{574}N_{88}O_{103}S_5$ (LBA1788), $C_{2361}H_{3654}N_{554}O_{598}S_{16}$ (yifk), $C_{2397}H_{3700}N_{566}O_{607}S_{19}$ (LBA0899), $C_{1546}H_{2369}N_{385}O_{400}S_{14}$ (LBA1209), $C_{1524}H_{2429}N_{389}O_{399}S_{14}$ (LBA0972), $C_{2446}H_{3978}N_{586}O_{657}S_{26}$ (LCTP) and $C_{1313}H_{2034}N_{298}O_{320}S_{10}$ (LBA0338) (Table 2a). The total negatively charged residues ranged between 4 (LBA1705) and 56 (LBA1510). The total number of positively charged residues ranged between 7 (LBA1705) and 78 (LBA1510). LBA1510 (14232) has the largest total number of atoms while LBA1788 (1113) has the lowest total number of atoms (Table 2a).

Proteins	-R (ASP +Glu)	+R (Arg + Lys)	# atoms	Atomic Number
LBA1510	56	78	14232	C:4693, H:7164, N:1136, O:1222, S:17
LBA0214	41	35	4128	C:1326, H:2055, N:349, O:391, S:7
LBA1705	4	7	1527	C:507, H:792, N:110, O:115, S:3
LBA1446	25	36	7653	C:2448, H:3927, N:603, O:652, S:23
LBA0037	25	32	4429	C:1462, H:2229, N:363, O:366, S:9
LBA1825	9	18	1437	C:444, H:734, N:130, O:120, S:9
LBA0995	27	31	7910	C:2635, H:3966, N:622, O:656, S:31
LBA1788	11	11	1113	C:343, H:574, N:88, O:103, S:5
yifk	24	31	7183	C:2361, H:3654, N:554, O:598, S:16
LBA0899	22	32	7289	C:2397, H:3700, N:566, O:607, S:19
LBA1209	19	24	4714	C:1546, H:2369, N:385, O:400, S:14
LBA0972	20	31	4755	C:1524, H:2429, N:389, O:399, S:14
LctP	23	26	7693	C:2446, H:3978, N:586, O:657, S:26
LBA0338	6	15	3975	C:1313, H:2034, N:298, O:320, S:10

Table 2a. Physicochemical properties of query uncharacterized proteins

-R (ASP +Glu) = Total number of negatively charged residues (ASP +Glu); +R (Arg + Lys) = Total number of positively charged residues (Arg + Lys); # atoms = total number of atoms

The theoretical isoelectric points ranged between 5.61 (LBA0214) and 9.80 (LBA1825). LBA1510 has the highest amino acid length {m/859/} and molecular weight (99596.40) while LBA1788 has the lowest amino acid length {m/69/} and molecular weight (7739.16) (Table 2b).

Proteins	pI	#AA	MW	AI	GRAVY	II
LBA1510	9.43	859	99596.40	110.70	0.183	31.82
LBA0214	5.61	255	29366.43	87.88	-0.484	36.26
LBA1705	9.60	93	10364.71	138.39	0.978	36.99
LBA1446	9.66	490	52976.14	121.20	0.723	27.99
LBA0037	9.21	265	31035.53	104.38	0.172	34.85
LBA1825	9.80	87	10102.05	78.39	-0.546	31.49
LBA0995	8.64	490	55848.11	104.71	0.553	42.10
LBA1788	6.55	69	7739.16	104.49	-0.090	30.46
yifk	9.25	457	49881.30	122.04	0.816	21.95
LBA0899	9.49	463	50768.31	117.49	0.775	26.15
LBA1209	9.10	292	33198.00	105.21	0.406	29.50
LBA0972	9.43	295	33034.24	121.97	0.447	24.35
LctP	8.41	505	52941.58	135.28	1.036	28.50
LBA0338	9.61	252	27435.00	128.10	1.045	23.94

Table 2b. Physicochemical properties of query uncharacterized proteins (cont'd)

pI = theoretical isoelectric point, #AA = number of amino acids; MW = molecular weight; AI = aliphatic index; GRAVY = Grand average of hydropathicity; II = instability index

The aliphatic index scores were generally high, ranging between 138.39 (LBA1705) and 78.39 (LBA1825). LBA1510, LBA1705, LBA1446, LBA0037, LBA0995, yifk, LBA0899, LBA1209, LBA0972, lctP and LBA0338 had positive GRAVY values while negative values were obtained for the remaining proteins. The instability index scores were less than 40 for all the query proteins except for LBA0995, which had a value of 42.10 (Table 2b). Only LBA0995 was observed to be unstable in nature. The estimated half-life was the same for all the proteins, 30 hours for mammalian reticulocytes, *in vitro*, >20 hours for yeast *in vivo* and >10 hours for

Escherichia coli, *in vivo* (Table 2c). The extinction coefficient (EC) was highest for LBA1510 (150580), followed by LBA0995 (119080) while LBA1788 (1615) had the lowest value.

Amino acid compositions of query proteins

The amino acid compositions of the query proteins are shown in Figure 1. The prominent amino acids for LBA1510 are Leu, Ile and Phe while the least prominent ones included His and Trp (Figure 1a). The most prominent amino acids of LBA0214 include Val, Lys and Leu while the least prominent amino acids included Try and Cys (Figure 1b). Leu and Ile are the most prominent amino acids of LBA1705, while Trp and Gln are the least prominent amino acids (Figure 1c). Leu and Ile, Gly and Ala, and Cys and Trp are the most and least prominent amino acids of LBA1446 respectively (Figure 1d). The most and least prominent amino acids of LBA0037 are Ile and Leu and Cys and Gln respectively (Figure 1e). The most prominent amino acids of LBA1825 are Lys and Val while the least prominent are Ser and Trp (Figure 1f). Leu and Ile were the most prominent amino acids, while Cys and Arg are the least prominent amino acids in LBA0995 (Figure 1g).

Figure 1h shows that Leu and Lys are the most prominent amino acids of LBA1788, while Asn, Gln, His and Tyr are the least prominent. Ile and Leu were the most abundant in yifk (Figure 1i) and LBA0899 (Figure 1j), while Gln and Asp were the least abundant in yifk, and Gln and His (LBA0899) were the least abundant amino acids. The most prominent amino acids were Ile, Phe, Gly and Ala for LBA1209 (Figure 1k), Leu, Ile and Gly for LBA0972 (Figure 1l); Leu, Ile and Ala for LctP (Figure 1m) and Ile, Leu and Ala for LBA0338 (Figure 1n).

Sub-cellular localization and properties of the secondary structure of query proteins

The sub-cellular localization is presented in Table 3. With the exception of LBA1825 whose localization is not known and LBA1788 whose localization is cytoplasmic, the rest proteins are found in the cytoplasmic membrane.

Subcellular localization scores					
Protein	Cytoplasmic	Cytoplasmic membrane	Cell wall	Extracellular	Prediction
LBA1510	0.00	10.00	0.00	0.00	Cytoplasmic membrane
LBA0214	9.97	0.00	0.01	0.02	Cytoplasmic membrane
LBA1705	0.32	9.55	0.12	0.01	Cytoplasmic membrane
LBA1446	0.00	10.00	0.00	0.00	Cytoplasmic membrane
LBA0037	0.00	10.00	0.00	0.00	Cytoplasmic membrane
LBA1825	2.50	2.50	2.50	2.50	Unknown
LBA0995	0.00	10.00	0.00	0.00	Cytoplasmic membrane
LBA1788	7.50	1.15	0.62	0.73	Cytoplasmic
yifk	0.00	10.00	0.00	0.00	Cytoplasmic membrane
LBA0899	0.00	10.00	0.00	0.00	Cytoplasmic membrane
LBA1209	0.00	10.00	0.00	0.00	Cytoplasmic membrane
LBA0972	0.00	10.00	0.00	0.00	Cytoplasmic membrane
LctP	0.00	10.00	0.00	0.00	Cytoplasmic membrane
LBA0338	0.00	10.00	0.00	0.00	Cytoplasmic membrane

Table 3. Subcellular localization results of query uncharacterized proteins

The properties of the secondary structure of the query proteins of the reference genome of the probiotic *Lactobacillus acidophilus* are presented in Table 4. Figures 2 and 3 show the pictorial representation of secondary structure properties of query proteins and a chart showing the percentages of alpha helices, extended strands, beta turns and random coils of the proteins of the reference genome of *Lactobacillus acidophilus* respectively.

Secondary structure elements (%)										
Proteins	AH	3 ₁₀ helix	Pi hel	BB	ES	BT	BR	RC	AS	OS
LBA1510	41.56	0.00	0.00	0.00	19.56	4.77	0.00	34.11	0.00	0.00
LBA0214	45.88	0.00	0.00	0.00	18.43	4.71	0.00	30.98	0.00	0.00
LBA1705	39.78	0.00	0.00	0.00	23.66	5.38	0.00	31.18	0.00	0.00
LBA1446	51.84	0.00	0.00	0.00	18.78	5.31	0.00	24.08	0.00	0.00
LBA0037	64.15	0.00	0.00	0.00	13.58	1.51	0.00	20.75	0.00	0.00
LBA1825	28.74	0.00	0.00	0.00	26.44	4.60	0.00	40.23	0.00	0.00
LBA0995	52.45	0.00	0.00	0.00	14.69	5.10	0.00	27.76	0.00	0.00
LBA1788	53.62	0.00	0.00	0.00	14.49	11.59	0.00	20.29	0.00	0.00
yifk	50.55	0.00	0.00	0.00	17.94	6.35	0.00	25.16	0.00	0.00
LBA0899	48.60	0.00	0.00	0.00	17.06	7.13	0.00	27.21	0.00	0.00
LBA1209	46.92	0.00	0.00	0.00	23.97	7.53	0.00	21.58	0.00	0.00
LBA0972	42.03	0.00	0.00	0.00	23.39	8.47	0.00	26.10	0.00	0.00
LctP	56.15	0.00	0.00	0.00	15.08	5.75	0.00	23.02	0.00	0.00
LBA0338	58.33	0.00	0.00	0.00	17.86	2.78	0.00	21.03	0.00	0.00

Table 4. Properties of secondary structure of query uncharacterized proteins

AH = Alpha helix; BB = Beta bridge; ES = Extended strand; BT = Beta turn; BR = Bend region; RC = Random coil; AS = Ambiguous states; OS = Other states

The alpha helix scores (%) ranged between 28.74 (LBA1825) and 64.15 (LBA0037), the extended strand values ranged between 14.49% (LBA1788) and 26.44% (LBA1825), the beta turn (%) values ranged between 1.51 (LBA0037) and 11.59 (LBA1788), and random coil values (%) ranged between 20 (LBA1788) and 40 (LBA1825).

Three-dimensional structure of query proteins

The three-dimensional tertiary structures and Ramachandran plots as well as terms in the SWISS-MODEL Server for the query proteins of the reference genome of *Lactobacillus acidophilus* are presented in Figures, 4 and 5 and Table 5. No experimental structure was found for any of the proteins in question and the probable biological functions for the proteins were as follows: undefined catalytic activity (LBA1510, LBA0214, LBA0037, LBA1825 and LBA1788); no probable function was available for LBA1705, LBA1446, LBA0995, yifk, LBA0899, LBA1209, LBA0972 and LBA0338; for LctP, uptake of L-lactate across the membrane was predicted as was the possibility of transporting D-lactate and glycolate.

The Ramachandran plots validated the three-dimensional structures obtained (Figures 5). No transmembrane segment was found in LBA0214, LBA1825 or LBA1788. The favoured area of Ramachandran (%) ranged between 99.59 (LBA1446) and 92.31 (LBA1705) (Table 5). The Q Mean scores ranged between -3.39 (LBA1705) and 0.84 (LBA0214), the torsion angle energy values ranged between -3.23 (LBA1446) and 0.46 (LBA1788), the solvation energy values ranged between -0.42 (LBA1825) and 4.28 (LBA1446), and the C_b interaction energy values ranged between -2.61 (LBA1510) and 2.17 (LBA0214), as shown in Table 5.

Proteins	Transmembrane segment prediction	QMEAN score	C _b interaction energy	All atoms pairwise energy	Solvation energy	Torsion angle energy	Ramachandran favoured (%)	Ramachandran outliers (%)
LBA1510	P	-1.64	-2.61	0.13	1.01	-1.52	97.55	0.58, A560 PRO, A535 ASN, A43 VAL, A9 ASN, A8 HIS
LBA0214	NP	0.84	2.17	1.68	1.78	-0.25	99.21	0.00
LBA1705	P	-3.39	-1.85	-0.40	-2.16	-2.22	92.31	1.10 A2 LEU
LBA1446	P	-1.49	0.85	0.94	4.28	-3.23	99.59	0.00
LBA0037	P	-0.56	0.62	1.80	2.44	-1.70	98.48	0.00
LBA1825	NP	-0.92	-0.06	-0.81	-0.42	-0.73	92.94	0.00
LBA0995	P	-2.35	-2.25	0.59	2.16	-2.81	97.75	0.41, A6 GLU, A8 ILE
LBA1788	NP	0.75	-0.17	0.34	0.84	0.46	97.01	0.00
yifk	P	-1.69	-0.54	0.63	2.22	-2.41	96.48	0.00
LBA0899	P	-1.27	-0.94	0.45	2.53	-1.99	97.40	0.22, A3 LYS
LBA1209	P	-1.06	1.31	1.73	1.49	-1.99	96.90	1.03, A291 VAL, A142 GLY, A204 HIS
LBA0972	P	-1.79	-0.04	0.89	1.89	-2.61	96.25	0.68, A293 ILE, A285 ASN
LctP	P	-0.89	0.41	0.86	3.32	-2.16	98.61	0.00
LBA0338	P	-1.91	-0.23	0.46	2.75	-3.01	97.20	0.00

Table 5. MolProbity, QMEAN, Solvation energy and Torsion angle energy scores for query uncharacterized proteins

Protein protein interactions of query proteins

The predicted protein protein interaction network for the query proteins of the reference genome of the *Lactobacillus acidophilus* is presented in Figure 6. Protein scores less than 0.5 were not included in this analysis. The functional partners with their corresponding confidence scores for LBA1510 included LBA1283 (0.765), LBA1509 (0.738), LBA1927 (0.671), LBA0860 (0.618), LBA1511 (0.523) and LBA1194 {m/517/} (Figure 6a). For LBA0214, the functional partners and corresponding confidence scores were metG (0.980), rsmA (0.968), mmV (0.904), LBA0217 (0.783), tmk (0.722), LBA1271 (0.596), LBA0212 (0.562), polA (0.562), purR (0.543) and LBA0008 (0.521) as shown in Figure 6b.

The predicted functional partners of LBA1705 include LBA1471 (0.754), LBA0170 (0.570) and pepF (0.559) (Figure 6c). LBA1447 (0.920), LBA1679 (0.831), LBA1429 (0.830), LBA1444 (0.785), LBA0017 (0.534) and LBA0552 (0.524) are the functional partners of LBA1446 (Figure 6d). LBA0036 (0.548) and pbpX-2 (0.528) are the functional partners of LBA0037 as revealed in Figure 6e.

The functional partners of LBA1825 included gldB (0.913), ychF (0.909), parA (0.869), parB (0.863), parB-2 (0.863), LBA1823 (0.790), LBA0418 (0.689), LBA1258 (0.629) and LBA1830 (0.572) (Figure 6f). The functional partners of LBA0995 included LBA0096 (0.843), LBA0997 (0.711), fabG (0.565) and pepD (0.526) (Figure 6g). LBA1789 (0.800), LBA1790 (0.773) and LBA1787 (0.508) are the functional partners of LBA1788 (Figure 6h). Only aapA was the functional partner of yifk (Figure 6i) based on the specified condition, with a confidence score 0.710. Both thyA (0.536), and dfrA (0.507) are functional partners of LBA0899 (Figure 6j).

Based on the specified conditions, msrA was the functional partner of LBA1209 (Figure 6k) and its confidence score is 0.602. LBA0426 (0.883) and cca (0.611) are the functional partners of LBA0972 (Figure 6l). The functional partners of LctP include ldhD (0.807), LBA1598 (0.741), galM (0.702), fni (0.669), fruA (0.609), adhE (0.542), ldh1 (0.511) and citH (0.511) as shown in Figure 6m. For LBA0338, LBA0339 (0.688), LBA0340 (0.587), LBA0337 (0.574), pheS (0.540), LBA0336 (0.505) and htrA (0.502) were detected (Figure 6n).

Active site analysis of the query proteins

Images representing the results of the active site analysis of query proteins of the *Lactobacillus acidophilus* reference genome are presented in Figure 7. There are 98 amino acid residues in the active site (shaded) of LBA1510 (Figure 7a). Figure 7b shows the 26 amino acid residues (His, Asp, Gln, Phe, Cys, Pro, Asp, Glu, Asp, Trp, Asp, Glu, Gln, Ile, His, Ser, Arg, Phe, Leu, His, Asn, Phe, Asn, Asp, Try and Leu) in the active site (shaded)

of LBA0214. For LBA1705, there are 20 amino acid residues (Leu, Ile, Ile, Gln, Leu, Leu, Thr, Ile, Pro, Asn, Gly, Glu, Phe, Ile, Phe, Try, Met, Leu, Ala, and Leu) in its active site (shaded) as revealed in Figure 7c.

Figure 7d shows the 74 amino acid residues (Thr, Val, Gly, Thr, Leu, Ser, Thr, Leu, Pro, Met, Thr, ... Phe) in the active site (shaded) of LBA1446. Figure 7e reveals the 51 amino acid residues (Try, Ile, Gly, Trp, Glu, Try, Cys... Leu) in the active site (shaded) of LBA0037. The 11 amino acid residues (Leu, Leu, Arg, Met, Lys, Met, Ile, Val, Arg, Pro and Phe) in the active site of LBA1825 (shaded) are shown in Figure 7f.

There are 89 amino acid residues (Ile, Glu, Thr, Ile, Asn, Phe... Lys) in the active site (shaded) of LBA0995 (Figure 7g). There are seven amino acid residues (Ile, Asn, Leu, Gly, Val, Arg and Thr) in the active site (shaded) of LBA1788 (Figure 7h). The 35 amino acid residues (Ala, Lys, Lys, Ala, Pro, Glu...Pro) in the active site of yifk are shaded (Figure 7i). The 26 amino acid residues (Met, Gly, Lys, Lys, Asn, Val...Asn) in the active site (shaded) of LBA0899 are shown in Figure 7j.

There are 72 amino acid residues (Met, Leu, Asp, Asn, Arg, Try, Asn, Phe, Ser, Lys, Ser, Ser, Gly...Val) in the active site (shaded) of LBA1209 (Figure 7k). There are 60 amino acid residues (Asn, His, Thr, Ile, Glu, Arg, Try, Phe...Ile) in the active site (shaded) of LBA0972 (Figure 7l). Figure 7m shows the 53 amino acid residues (Glu, Met, Trp, Pro, Ile, Val, Leu, Leu, Thr, Gly, Glu, Val, Thr, Asn, Leu...Lys) in the active site (shaded) of LctP. Figure 7n indicates that there are eight amino acid residues (Ile, Ile, Gly, Lys, Lys, Arg, Ile and Ala) in the active site (shaded) of LBA0338.

Immunogenicity, allergenicity and toxicity evaluations of query proteins

Table 6 shows the results of the immunogenicity, allergenicity and toxicity analyses of the query proteins of the reference genome of *Lactobacillus acidophilus*. All the query proteins were non-antigenic, non-allergic and non-toxic.

Protein	Allergenicity	Antigenicity	Toxicity
LBA1510	Non-antigenic	Non-allergenic	Non-toxic
LBA0214	Non-antigenic	Non-allergenic	Non-toxic
LBA1705	Non-antigenic	Non-allergenic	Non-toxic
LBA1446	Non-antigenic	Non-allergenic	Non-toxic
LBA0037	Non-antigenic	Non-allergenic	Non-toxic
LBA1825	Non-antigenic	Non-allergenic	Non-toxic
LBA0995	Non-antigenic	Non-allergenic	Non-toxic
LBA1788	Non-antigenic	Non-allergenic	Non-toxic
yifk	Non-antigenic	Non-allergenic	Non-toxic
LBA0899	Non-antigenic	Non-allergenic	Non-toxic
LBA1209	Non-antigenic	Non-allergenic	Non-toxic
LBA0972	Non-antigenic	Non-allergenic	Non-toxic
LctP	Non-antigenic	Non-allergenic	Non-toxic
LBA0338	Non-antigenic	Non-allergenic	Non-toxic

Table 6. Immunogenicity, allergenicity and toxicity evaluation of query uncharacterized proteins

IL-6-inducing peptides and gene expression analyses

The proteins were scanned for regions producing IL-6-inducing peptides and the selected immunogenic IL-6-inducing peptides are presented in Table 7.

Source	Start	Sequence	IL-6 Score	Prediction	Immunogenicity score
LBA1510	175	LNGKKNHLILITFLL	0.54	IL-6 inducer	1.2859
LBA1510	85	YYLLSPFNLLLFVFP	0.43	IL-6 inducer	1.9628
LBA1510	321	LLLIFLIASLFWTPL	0.53	IL-6 inducer	1.7229
LBA1510	176	NGKKNHLILITFLLW	0.46	IL-6 inducer	1.6838
LBA0214	64	CPDIADYDQKAEDE	0.21	IL-6 inducer	1.3600
LBA0214	63	YCPDIADYDQKAED	0.14	IL-6 inducer	1.2623
LBA0214	68	AKDYDQKADELKQ	0.13	IL-6 inducer	1.2104
LBA0214	65	PDIADYDQKADEL	0.22	IL-6 inducer	1.1326
LBA1705	74	FFLIFIFYMALAYKL	0.31	IL-6 inducer	2.9050
LBA1705	73	NFFLIFIFYMALAYK	0.26	IL-6 inducer	2.4815
LBA1705	75	FLIFIFYMALAYKLP	0.25	IL-6 inducer	2.4695
LBA1705	72	NNFFLIFIFYMALAY	0.27	IL-6 inducer	2.1912
LBA1446	461	SIIFALIALVLSFFL	0.36	IL-6 inducer	1.5321
LBA1446	462	IIFALIALVLSFFLK	0.33	IL-6 inducer	1.4800
LBA1446	286	VGIEMVLPLYIQNLR	0.26	IL-6 inducer	1.4267
LBA1446	285	MVGIEMVLPLYIQNL	0.27	IL-6 inducer	1.2747
LBA0995	52	WIFLLFAYVIPYALM	0.29	IL-6 inducer	2.5476
LBA0995	53	IFLLFAYVIPYALMC	0.29	IL-6 inducer	2.0570
LBA0995	423	LIMGWWCLIFTFICA	0.29	IL-6 inducer	1.9308
LBA0995	422	ALIMGWWCLIFTFIC	0.28	IL-6 inducer	1.8432
ylfk	413	FALIMLVVIVIFMFI	0.24	IL-6 inducer	2.1249
ylfk	376	MVAWFVILLAE LRFR	0.29	IL-6 inducer	1.5889
ylfk	123	SDWIAGIIILFLLI	0.23	IL-6 inducer	1.4489
ylfk	377	VAWFVILLAE LRFR	0.25	IL-6 inducer	1.2896
LBA0899	415	FAFLMLLVIVIFMFI	0.33	IL-6 inducer	2.6087

Source	Start	Sequence	IL-6 Score	Prediction	Immunogenicity score
LBA0899	414	YFAFLMLLVIVIFMF	0.29	IL-6 inducer	2.4398
LBA0899	416	AFLMLLVIVIFMFIN	0.28	IL-6 inducer	1.9764
LBA0899	375	LPGMIPWVFVILLAEL	0.28	IL-6 inducer	1.4713
LBA1209	69	MYFILNFPLFILAWF	0.29	IL-6 inducer	3.0628
LBA1209	68	FMYFILNFPLFILAW	0.29	IL-6 inducer	3.0552
LBA1209	71	FILNFPLFILAWFKI	0.3	IL-6 inducer	2.7935
LBA1209	70	YFILNFPLFILAWFK	0.28	IL-6 inducer	2.5148
LBA0972	82	TIWGTLCLSFFLWFW	0.34	IL-6 inducer	2.3249
LBA0972	86	TLCLSFFLWFWRVSP	0.3	IL-6 inducer	2.1876
LBA0972	87	LCLSFFLWFWRVPI	0.35	IL-6 inducer	2.1674
LBA0972	83	IWGTLCLSFFLWFWR	0.41	IL-6 inducer	2.1239
LctP	3	IKFAMALIPHIWLII	0.2	IL-6 inducer	2.3257
LctP	5	FAMALIPHIWLIIISL	0.18	IL-6 inducer	1.9856
LctP	1	MWIKFAMALIPHIWL	0.18	IL-6 inducer	1.9148
LctP	4	KFAMALIPHIWLIIIS	0.17	IL-6 inducer	1.9062
LBA0338	130	PMWLKIVYLVALLIM	0.28	IL-6 inducer	1.2871
LBA0338	129	QPMWLKIVYLVALLI	0.29	IL-6 inducer	1.1688
LBA0338	66	ELKGHWKETWFYTIF	0.38	IL-6 inducer	1.0227
LBA0338	131	MWLKIVYLVALLIMG	0.29	IL-6 inducer	0.8849

Table 7. Selected immunogenic IL-6-inducing peptides from uncharacterized proteins from *L.acidophilus* reference genome

The importance of these regions is that they may be altered, amended or removed for vaccine development or for therapeutic purposes. In this study, all the proteins produce peptides which induce IL-6, suggesting their potential for effective therapeutic purposes, in addition to being effective feed additives and food

supplements. Immunogenic IL-6-inducing peptides from proteins of the *L. acidophilus* reference genome according to their IL-6-inducing prediction values and immunogenicity values when the nonimmunogenic ones (LBA0037, LBA1825) and non-IL-6-inducing peptides from LBA1788 have been removed are presented in Figure 8.

Gene expression levels of IL-6 across various conditions in *Homo sapiens* and *Mus musculus* are shown in Figure 9. IL-6 is upregulated in urinary bladder, anterior prostate, blood vessel, liver, lung, synovium and breast in *Homo sapiens*. In *Mus musculus*, IL-6 is regulated in extraocular muscles, gonadal adipose tissue, bone marrow, lung, skeletal muscle and aorta among others. According to cancer categories in *Homo sapiens*, IL-6 is upregulated by prostate, neoplasm, malignant stroma and malignant lymphoma NOS. on top of the list of the genes positively co-expressed with IL-6 according to cancer categories in *Mus musculus* are paired related homeobox 2 (Prrx2), twist basic helix-loop-helix transcription factor 1 (Twist1), EH-domain containing 2 (Ehd2), brain derived neurotrophic factor (Bdnf), 5-hydroxytryptamine (serotonin) receptor 1B (Htr1b) and microtubule associated serine/threonine kinase family (Mast4).

Discussion

Previous authors have documented the importance of exploring hypothetical proteins in different organisms of interest [30][31][32][33][34][35][36]. However, this is the first study in this area to focus on the structural, functional relationship, safety profile and IL-6-inducing capacity of uncharacterized proteins of probiotic *Lactobacillus acidophilus* for sustainable applications in animal and human nutrition as well as for therapeutic and cosmetics purposes.

The importance of proteins as the 'building blocks' for the body cannot be over-emphasized. Proteins play significant roles in the structural formation and functions of organisms, performing key roles in the cell, including protecting the body from pathogens and carrying out chemical reactions. However, proteins' discrete native structure influences its roles, exposes several channels, receptors and binding sites thereby controlling how they bind and interact with other molecules as well as how they form complexes used for regulatory and structural functions. A protein attains its functional native structure and reaches its final form (three-dimensional structure) by folding. Amino acid composition influences protein classification into different folding types, structural groups and functions. The physicochemical properties of a protein encoded in its amino acid determine its folding.

The physicochemical properties usually used in characterizing uncharacterized proteins include the molecular weight, theoretical isoelectric point (pI), amino acid composition, instability index, aliphatic

index and grand average of hydropathicity (GRAVY) [23]. In this study, the results of the physicochemical properties of the query proteins of the *Lactobacillus acidophilus* reference genome revealed that LBA1446 had the highest pI. LBA0214 and LBA1788 had values less than 7.00 while the remaining query proteins had values greater than 7.00, indicating whether they are basic or acidic.

A protein has no net charge at the isoelectric point [23]. At that point, the protein is not repelled or attracted by charged molecules. In a polar solvent, a charged molecule seems to be readily soluble. A protein with a pH above the pI is negatively charged and will not be attracted by a negatively charged surface, thereby decreasing its solubility. The pI plays an important role in protein purification, representing the pH during which solubility is minimal. The pI also talks about the basic or acidic nature of a protein. The pI of the amino acid an acidic protein will be lower and vice versa for a basic protein. Molecular weight is also useful in protein purification and separation.

The extinction coefficient has been defined as a measurement of how strongly a protein absorbs light at a given wavelength. Two values are usually produced by ProtParam for proteins measured in water at 280 nm. The first indicates that the computed value is based on the assumption that all cysteine residues appear as half cystines, that is, all pairs of Cys residues form cystines, while the second value is based on the assumption that no cysteine appears as half cystine, that is, the assumption that all Cys residues are reduced [23][34]. In this study, the extinction coefficient values are very high for LBA1510, LBA0214, LBA1446, LBA0037, LBA0995, yifk, LBA0899, LBA1209, LBA0972, LBA099513 and LBA0338 and high for LBA1705 and LBA1825, except for LBA1788. The high extinction coefficient may be an indication of the presence of a high concentration of Cys, Trp, and Tyr in the query proteins of *Lactobacillus acidophilus*, except for LBA1788, which does not contain Trp residues.

In this study, the half-life estimation was the same for all the query proteins. The estimated half-life is a prediction of the time it takes for half of the amount of protein in a cell to disappear after its synthesis in the cell, relying on the "N-end rule", which relates the half-life of a protein to the identity of its N-terminal residue, in this case, for humans (*in vitro*), yeast (*in vivo*) and *Escherichia coli* (*in vivo*) [23].

An instability index less than 40 indicates that the protein is stable while a value above 40 indicates that the protein may be unstable [23]. In the present study, all the query proteins had instability index values less than 40, which implies that they are stable, except for LBA0995 which may be regarded as unstable because it has an instability index greater than 40. The aliphatic index of a protein can be described as the relative volume occupied by aliphatic side chains, that is, alanine, isoleucine, leucine and valine), which may be regarded as a positive factor for the increase in thermostability of globular proteins [23].

The importance of a protein thermostability cannot be over-emphasized. A thermostable protein is preferred for industrial applications and food processing. The ability to withstand extreme industrial and preservation conditions in food processing and storage is beneficial for probiotic bacteria. Hence, the ability of probiotic bacteria to withstand harsh environmental conditions, having its proteins being thermostable is an advantage to food, cosmetic and therapeutic industries benefiting from the health-promoting probiotics in the era of climate change. Proteins with high aliphatic indices are thermostable under a wide temperature range. In this study, the aliphatic indices are generally high for all the query proteins, although the values for LBA1825 and LBA0214 are not as high as the values obtained for the other query proteins. The high aliphatic index values obtained by the query proteins of the *Lactobacillus acidophilus* reference genome used in this study indicate that the proteins are stable over a wide temperature range.

The importance of instability index is mainly in the storage of proteins in the correct solvent. Highly stable proteins are easy to store. For instance, insulin monomer with an instability index of 43 is considered unstable and may macroscopically aggregate in aqueous solution during storage, which may consequently result in decline in biological activities.

The calculation of the grand average of hydropathy value for a peptide or protein is based on the sum of hydropathy values of all the amino acids, divided by the number of residues in the sequence [23][37]. In the present study, only LBA0214, LBA1825 and LBA1788 obtained negative GRAVY values, which is an indication of being hydrophilic, while the other query proteins had positive GRAVY values, which implies that they are hydrophobic [38]. Another importance of GRAVY index is that it is used to determine whether a protein can be visualized on 2-D gel. Proteins with GRAVY score greater than 0.4 are usually difficult to view on 2-D gel because it lies outside the solubility range.

Proteins intended for use as food products or for use in consumer products are expected to be analysed for their allergenic reaction potential before they are introduced into the market to determine and mitigate the risk of inducing an immediate type I (IgE-mediated) allergic response. Allergenicity or allergenic potential in this study is defined as the potential of a protein to cause or elicit immediate type (IgE-mediated) allergic reactions in humans [39]. It is not certain whether probiotics fed to animals or offered for commercial sale are subjected to allergenic potential evaluations before they are introduced into the market. To ascertain the safety and sustainability of these probiotics, their safety profile including toxicity, allergenicity and antigenicity may be necessary, revealing the novelty of this study.

Bioinformatics tools such as AllerCatPro 2.0, VaxiJen ToxinPred [25][26][27], for quick and cost-saving approaches have been recommended. The findings from the present study revealed that none of the query

proteins of the *Lactobacillus acidophilus* reference genome studied were antigenic, allergenic or toxic for human or animal consumption.

The benefits of computational prediction of bacterial protein subcellular localization include the provision of rapid and cost-effective approaches for gaining insight into protein function, verifying experimental results, annotating newly sequenced bacterial genomes, detecting potential cell surface or secreted drug targets, and identifying microbial biomarkers. One of the most precise tools recommended for computational prediction of bacterial protein subcellular localization is PSORTb version 2.0 [40][41], which generates prediction results for five major locations for gram-negative bacteria (cytoplasmic, inner membrane, periplasmic, outer membrane and extracellular) and four locations for gram-positive bacteria (cytoplasmic, cytoplasmic membrane, cell wall and extracellular). However, the PSORTb version 3.0 used in this study is an updated version of 2.0 [20] with more features. In this study, LBA1788 is cytoplasmic, LBA1510, LBA0214, LBA1705, LBA1446, LBA0037, LBA0995 and yifk-LBA0338 are found in the cytoplasmic membrane while the subcellular localization of LBA1825 is not known.

The functional association of the query proteins of the *Lactobacillus acidophilus* reference genome are presented in the Results section (Figure 7). LBA1283 is a glycosyl transferase, LBA1509 is a penicillin-binding protein and LBA1927 and LBA0860 are hypothetical proteins. metG is a met-rRNA synthetase, that is not only required for the elongation of protein synthesis but also for the initiation of all mRNA translation. rsmA is a dimethyladenosine transferase and mmV is a putative promase like that is required for the correct processing of both the 5' and 3' ends of the 5S rRNA precursor. LBA0217 is a COG4466 uncharacterized protein conserved in bacteria while tmk is a thymidylate (dTTP) kinase; the phosphorylation of dTTP results in the formation of dTDP in both the de novo and salvage pathways of dTTP synthesis. LBA1471 is a putative multidrug efflux permease, LBA0170 is a putative 6-pyruvoyl-tetrahydropterin synthase and pepF is an oligopeptidase, endopeptidase F.

LBA1447 is a hypothetical protein, LBA1679 is an ABC transporter permease protein, LBA1429 is a putative transporter-membrane protein, LBA1444 is a transcriptional regulator family, MerR; LBA0017 is a putative general stress response and COG3237 is an uncharacterized protein conserved in bacteria belonging to the UPF0337 family. LBA0036 is a collagen-binding protein Cne precursor while pbpX-2 is a putative penicillin-binding protein. gldB is a glucose inhibited division protein B that specifically methylates the N7 position of a guanine in 16S rRNA, ychF is a GTP-binding protein and parA, parB and parB-2 are chromosome partitioning proteins. Both parB and parB-2 are predicted transcriptional regulators of COG1475 and belong to the Par family.

LBA0996 is a COG1982 arginine-lysine-ornithine decarboxylase, LBA0997 is an aluminum resistance protein, fabG is a 3-oxoacyl-(acyl-carrier protein) reductase and pepD is an aminoacyl-histidine dipeptidase, PepD, carnosinase. LBA1789 and LBA1787 are hypothetical proteins while LBA1790 is a vacuolar sorting receptor protein homologue of the PV72-cucur bit. The only functional associate of yifk, aapA, is an amino acid permease. The functional associate of LBA0899, thyA, is a thymidylate synthase, and dfrA is a dihydrofolate reductase, which is a key enzyme in folate metabolism.

LBA0426 is a hypothetical protein, while cca is a tRNA nucleotidyltransferase, polyA. LBA1598 is a glycolate oxidase, L-lactate dehydrogenase, cytochrome-type (lldD), ldhD is a D-lactate-dehydrogenase that belongs to the D-isomer specific 2-hydroxyacid dehydrogenase family, galM is a galactose-1-epimerase, mutarotase and fni is an isopentenyl diphosphate isomerase, involved in the biosynthesis of isoprenoids. LBA0339, the functional associate of LBA0338 is a phosphoglycerate mutase, COG0406 fructose-2,6-bisphosphatase, which belongs to the phosphoglycerate mutase family. LBA0340 is a cation efflux protein that is a cation diffusion facilitator (CDF) transporter (TC 2.A. 4) Family. LBA0337 is a hypothetical protein while pheS is a phenylalanyl-tRNA synthetase, beta subunit, SyfB and COG0073 EMAP domain.

IL-6 is mainly produced by T lymphocytes, B lymphocytes, dendritic cells microglia, fibroblasts, keratinocytes, mast cells mesangial cells and vascular endothelial cells, binding to either soluble IL-6 receptors (sIL-6R) or the membrane bound IL-6 receptors (mIL-6R) ^[42]. It is a pleiotropic cytokine which plays several key roles in the body such as influencing embryonic development, acute phase reactant pathways, B and T lymphocytes, synovial inflammation, blood brain barrier permeability and hematopoiesis (44, review). It is known to play key functions in the normal adaptive, innate and autoimmunity. It is primarily expressed when tumor necrosis factor- α (TNF α) and interleukin 1 β (IL-1 β) are activated. Prostaglandins, other cytokines, stress response, toll-like receptor activation (TLRs), and adipokines can also enhance its synthesis ^[43]. When dysregulated, IL-6 can contribute to inflammatory cascades, indicating pathophysiology of many autoimmune conditions.

At the initial phase of inflammation, after IL-6 is produced, it moves to the liver (via the bloodstream), which is followed by the quick induction of an extensive range of acute phase proteins including serum amyloid A (SAA), C-reactive protein (CRP), α 1-antichymotrypsin fibrinogen and haptoglobin ^[44]. IL-6 also reduces albumin, transferrin and fibronectin production. A persistent high-level concentration of SAA may result in several chronic inflammatory conditions with a serious complication through amyloid A amyloidosis generation ^[45], leading to deposition of amyloid fibril and progressive deterioration in organs ^[46].

In addition, IL-6 regulates serum zinc and iron levels by regulating their transporters. It initiates the production of hepcidin that hinders the iron transporter ferroportin 1 action on gut, thereby lowering the iron levels in the serum ^[47], indicating the importance of IL-6-hepcidin axis in regulating hypoferremia and anemia associated with chronic inflammation. In the bone marrow, IL-6 promotes the maturation of megakaryocyte, thereby releasing platelets ^[48]. Changes in acute phase protein levels and red blood cell and platelet counts can be used for determining the severity of inflammation in clinical laboratory examinations ^[46].

IL-6 axis has been successfully modulated (therapeutically) resulting in the approval of multiple therapeutic agents, currently being investigated (44, review). IL-6 has been suggested to have potential applications in neuro-inflammatory conditions (not yet FDA approved), based on its ability to form complex with other immunosuppressive medications (azathioprine and mycophenolate mofetil) (44, review).

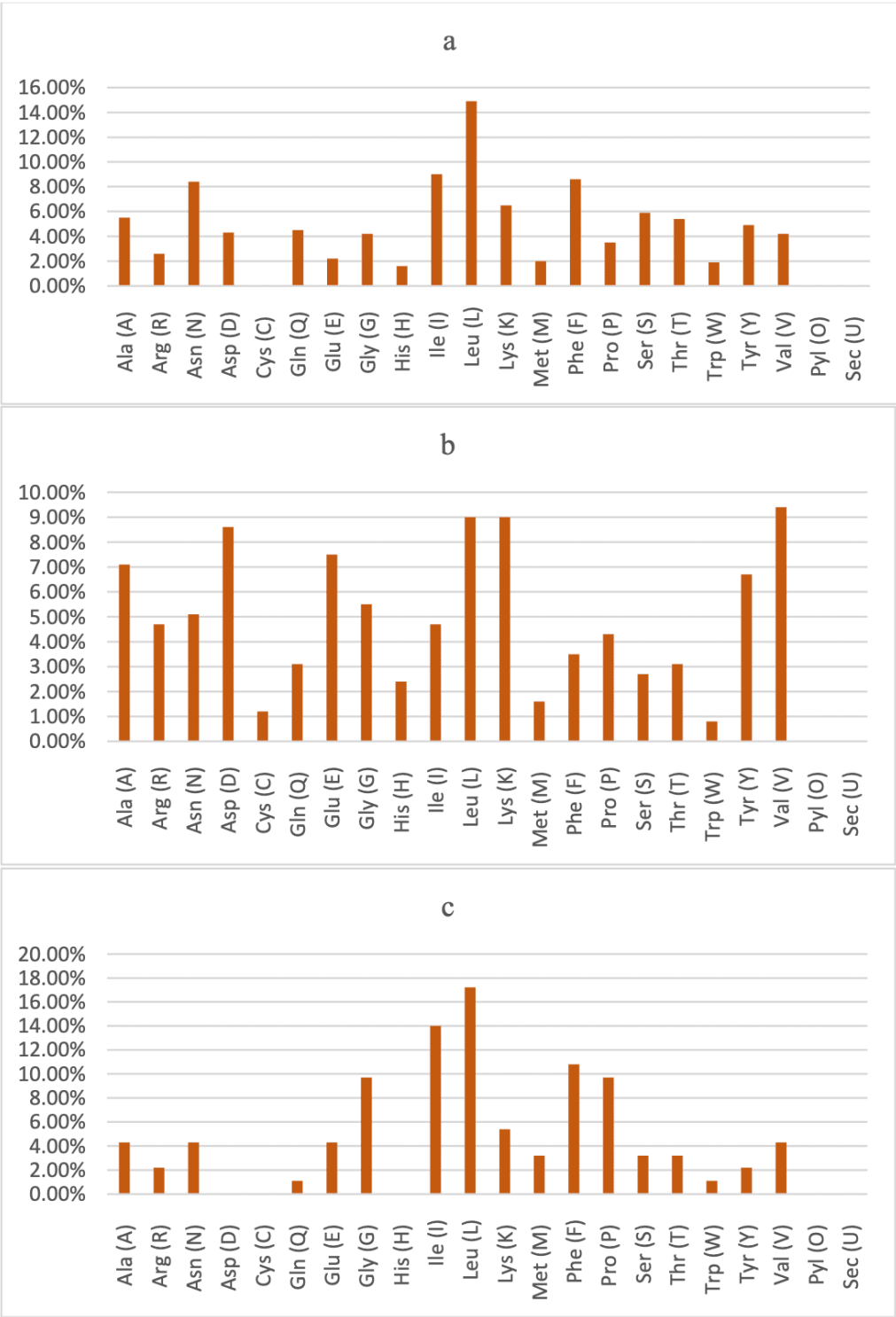
IL-6 has been reported to be produced by senescent cells and involved in senescence-induced inflammation and age-dependent pathologies and cancer ^[49]. It has been described as a crucial factor in inflammation, cancer and autoimmunity, performing its roles primarily via the IL-6–signal transducer and activator of transcription 3 (STAT3) pathway ^{[43][50][51]}. In addition, IL-6 amplifier (IL-6 Amp) is an amplification mechanism of IL-6, growth factors, chemokines and cytokines production via a synergic interaction between STAT3 and nuclear factor-kappa B (NF-κB), which also perform important functions in inflammatory conditions such as cancer, autoimmunity and cytokine storm syndromes ^{[52][53][54]}.

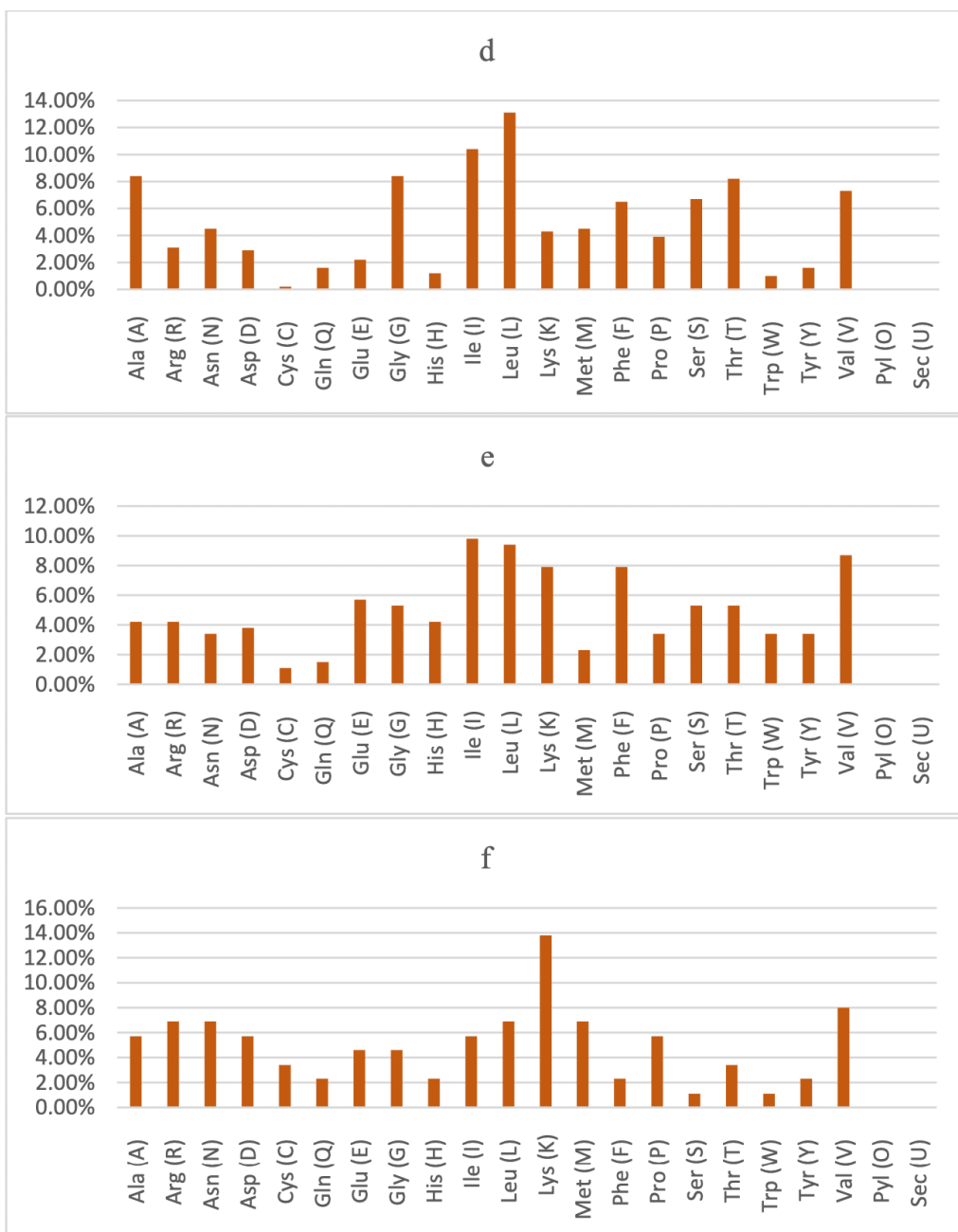
IL-6 inhibitors are now used clinically for conditions including rheumatoid arthritis (RA), Castleman's disease and considered for treating COVID-19 ^{[55][56]}, despite the involvement of many cytokines in inflammation-related diseases, based on the understanding that IL-6 is the main stimulator of STAT3 in inflammation. STAT3 performs crucial functions in inflammation and oncogenesis together with NF-κB, expressing IL-6 as a target, indicating STAT3 and NF-κB as the basis of the IL-6 Amp. Injury, infection, obesity, stressors, senescence, smoking, pre-neoplastic mutation and cell death are factors which may activate the IL-6 Amp ^[57]. The author further noted the importance of developing highly specific and effective small peptides against diseases with a requirement for high-resolution structural analysis of the ligand-receptor complex.

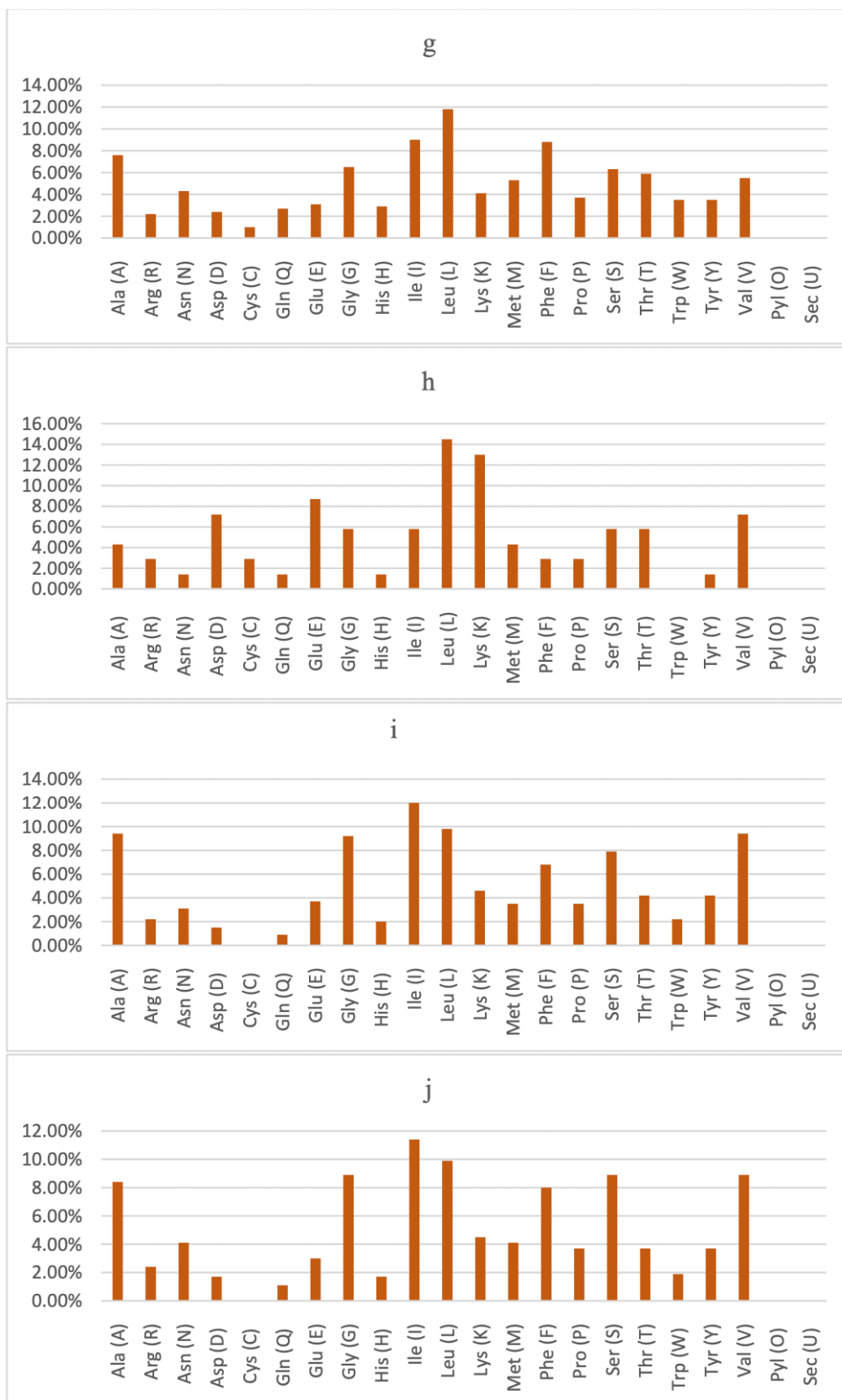
Studies are on-going to provide potent and health-promoting alternatives to synthetic antibiotics in animal and human nutrition without any adverse effect to animals' and human's well-being ^{[58][59][60][61]}, underscoring the importance of the present study, which attempted to unravel the functional mechanisms of one of such alternatives. Different analyses were carried out to characterize and investigate the safety

profiles of query proteins in the *L. acidophilus* reference genome. The findings of the study revealed that none of the query proteins were allergenic or toxic, attesting to safety profile of LAB probiotics. All the query proteins are stable under a wide range of temperatures, indicating they can withstand industrial processing and that climate change may not easily affect LAB probiotics. LBA0214, LBA1825 and LBA1788 are hydrophilic while the remaining ones are hydrophobic. Only LBA0995 has been determined to not be stable in nature, indicating that most of the proteins are stable, which is an advantage for their preservation and functionality. LBA1788 is cytoplasmic, while the rest are found in the cytoplasmic membrane. All the proteins produce peptides which induce IL-6 except LBA0037, LBA1825 and LBA1788 which are either non-immunogenic or non-IL-6-inducing. The immunogenic IL-6-inducing peptides are potential candidates for vaccine development and therapeutic purpose, further confirming the health-benefits of LAB probiotics in animal and human nutrition. The findings in this study are based on in silico analysis, hence, further studies adopting experimental approach (wet lab) are strongly recommended to validate the claims in this preliminary study.

Figures







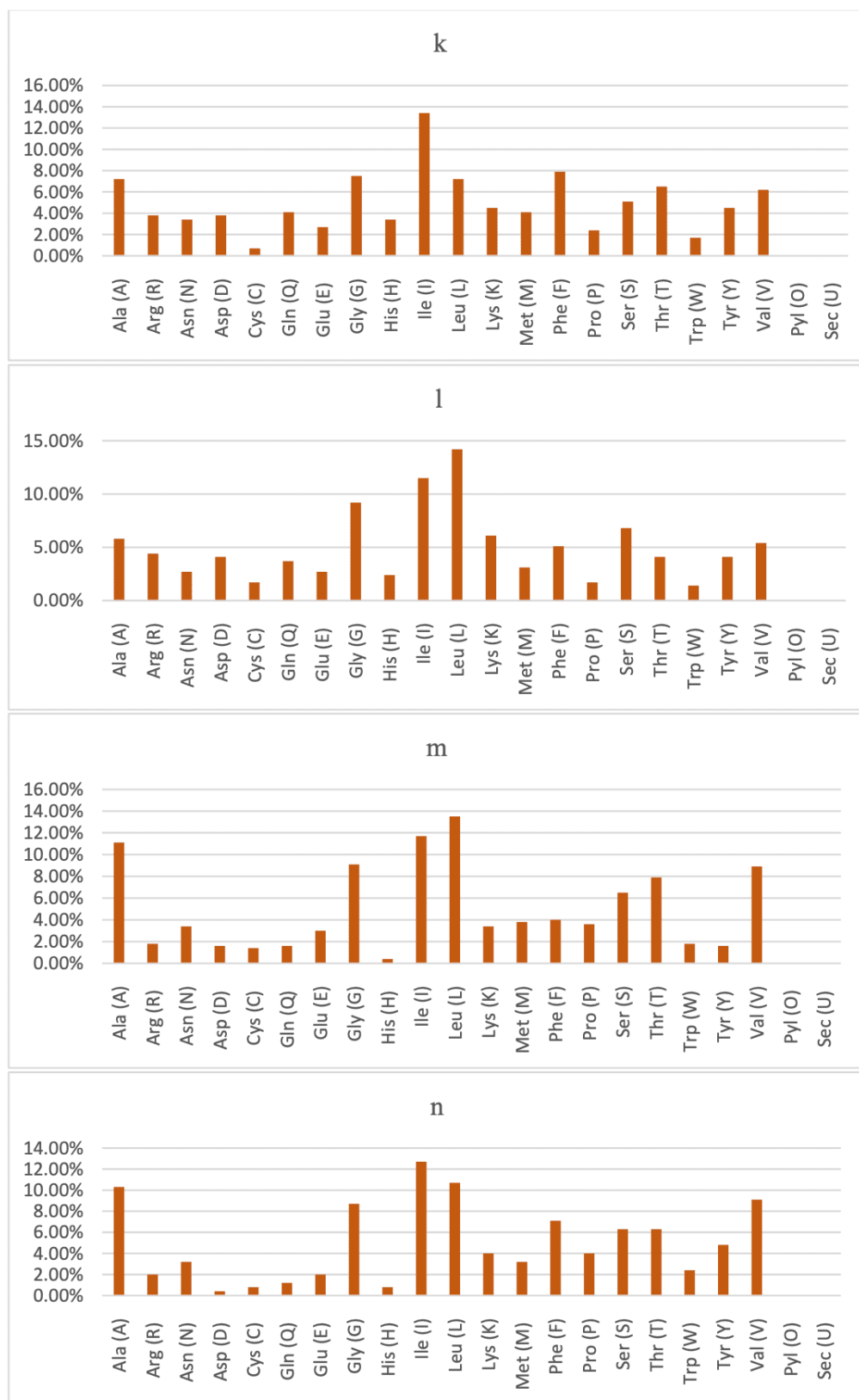


Figure 1. Charts showing amino acid compositions of query uncharacterized proteins as predicted for LBA1510 (a), LBA0214 (b), LBA1705 (c), LBA1446(d), LBA0037 (e), LBA1825 (f), LBA0995 (g),

LBA1788 (h), YIFK (i), LBA0899 (j), LBA1209 (k), LBA0972 (l), LctP (m) and LBA0338 (n)

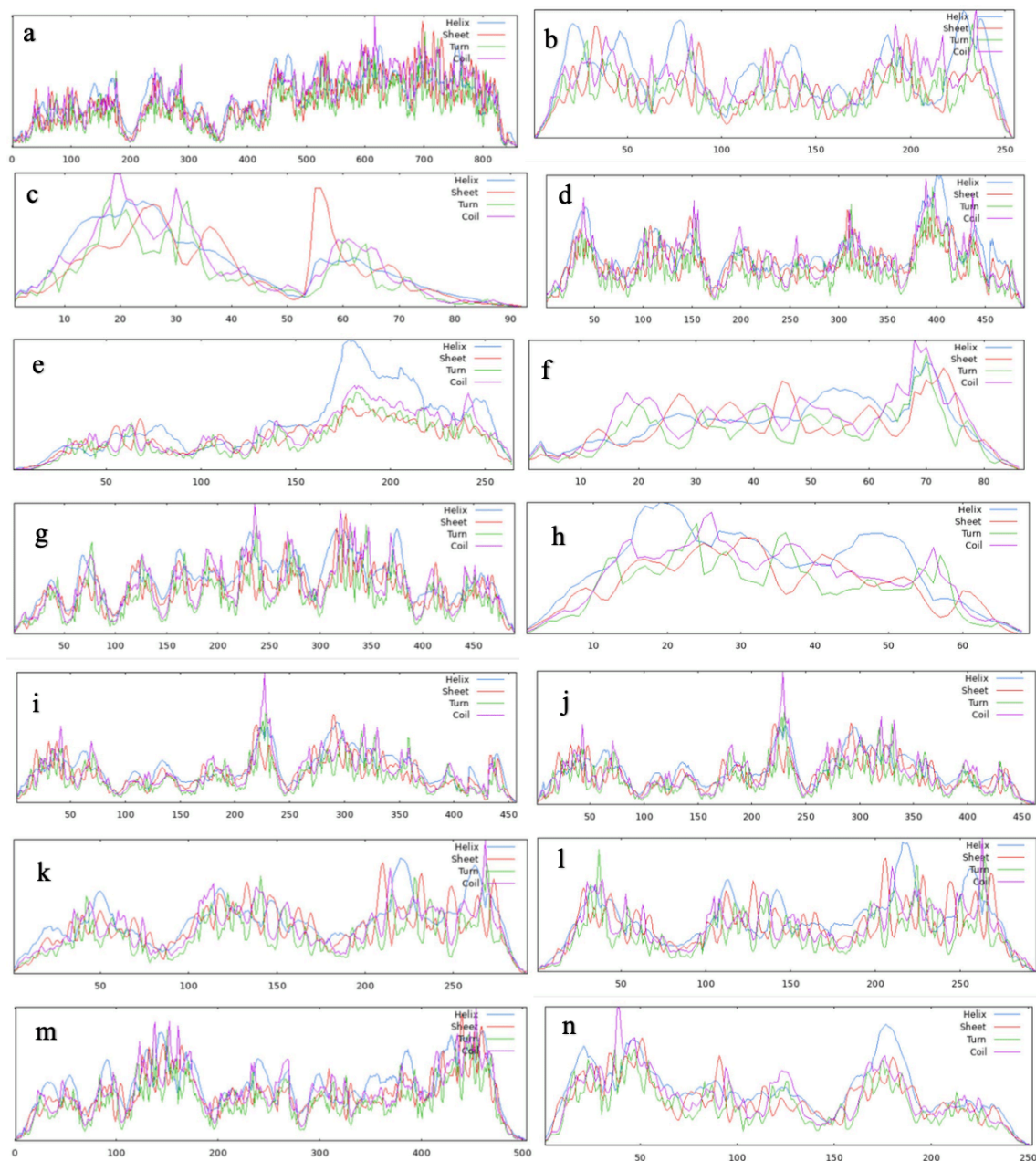


Figure 2. Pictorial representation of secondary structure properties of query uncharacterized proteins for LBA1510(a), LBA0214 (b), LBA1705 (c), LBA1446(d), LBA0037 (e), LBA1825 (f), LBA0995 (g), LBA1788 (h), YIFK (i), LBA0899 (j), LBA1209 (k), LBA0972 (l), LctP (m) and LBA0338 (n)

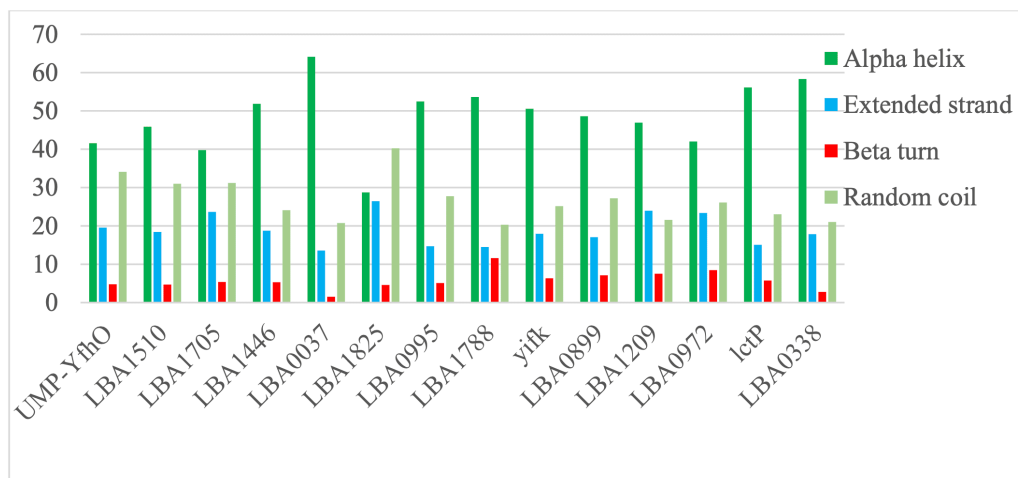


Figure 3. Chart showing the proportions in percentage of alpha helix, extended strand, beta turn and random coil of query uncharacterized proteins of reference genome of *Lactobacillus acidophilus*

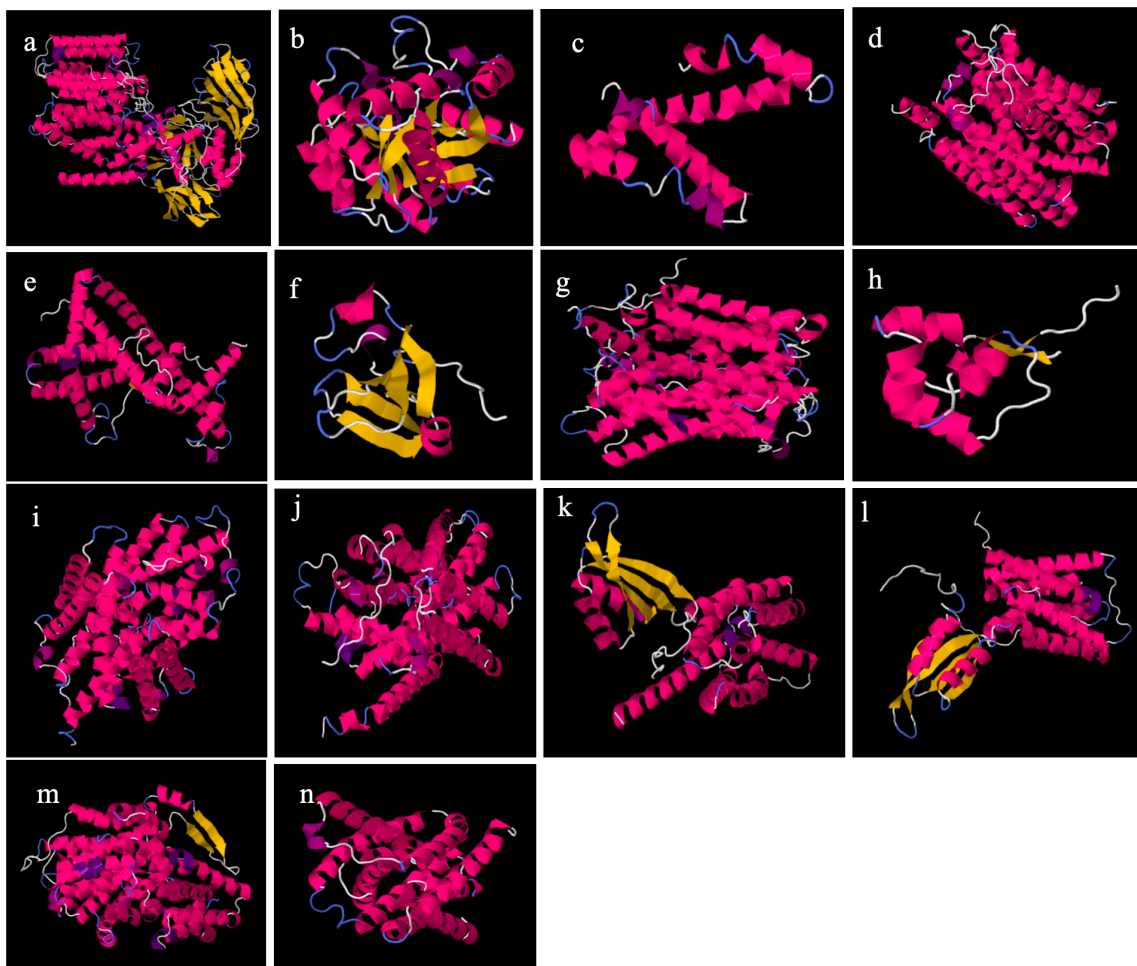
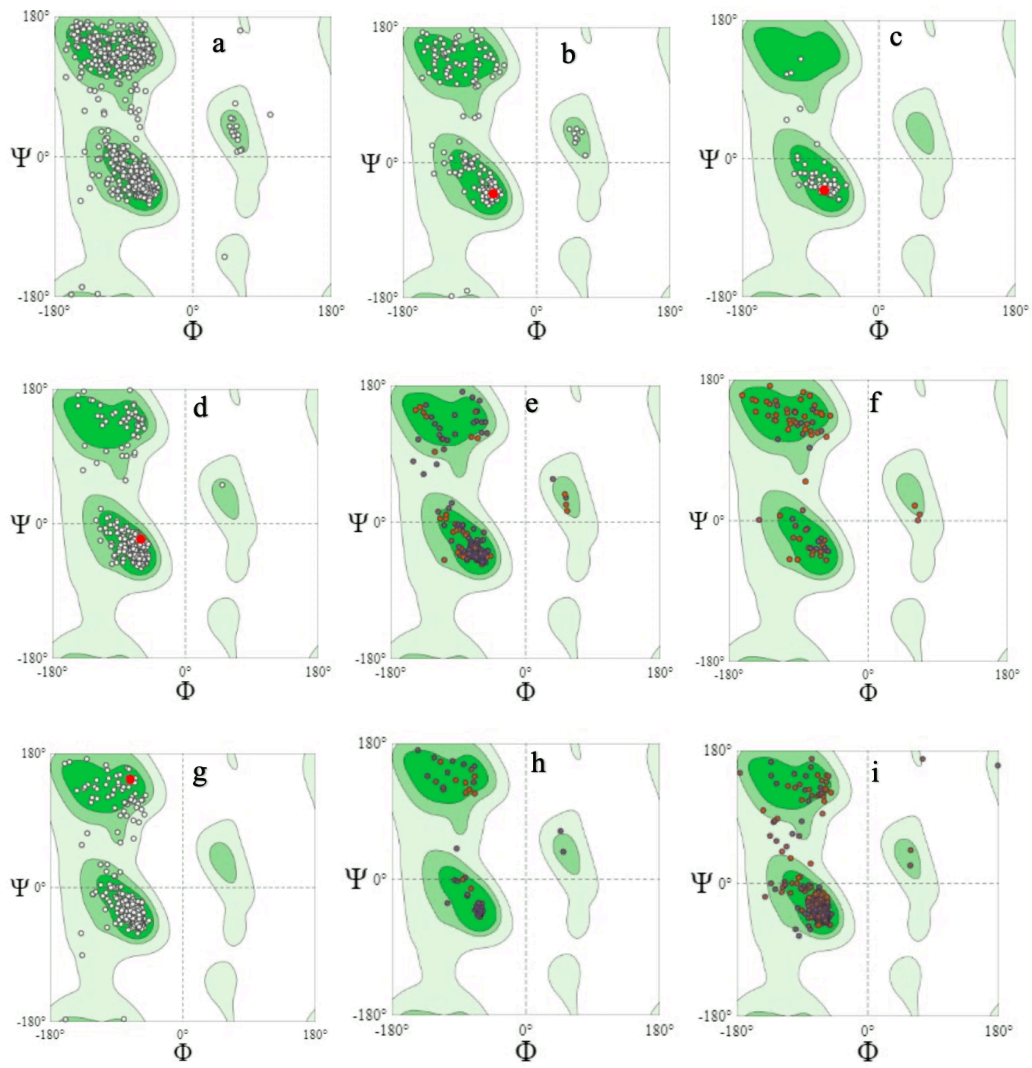


Figure 4. Three-dimensional tertiary structures of query uncharacterized proteins of reference genome of *Lactobacillus acidophilus* for LBA1510 (a), LBA0214 (b), LBA1705 (c), LBA1446 (d), LBA0037 (e), LBA1825 (f), LBA0995 (g), LBA1788 (h), yifk (i), LBA0899 (j), LBA1209 (k), LBA0972 (l), LctP (m) and LBA0338 (n)



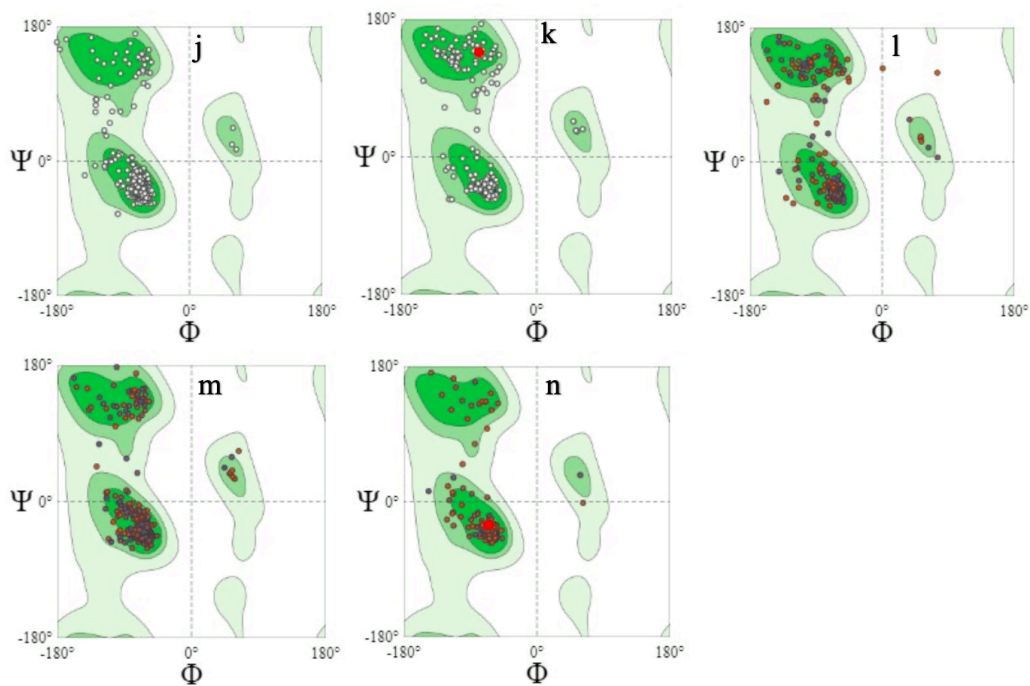
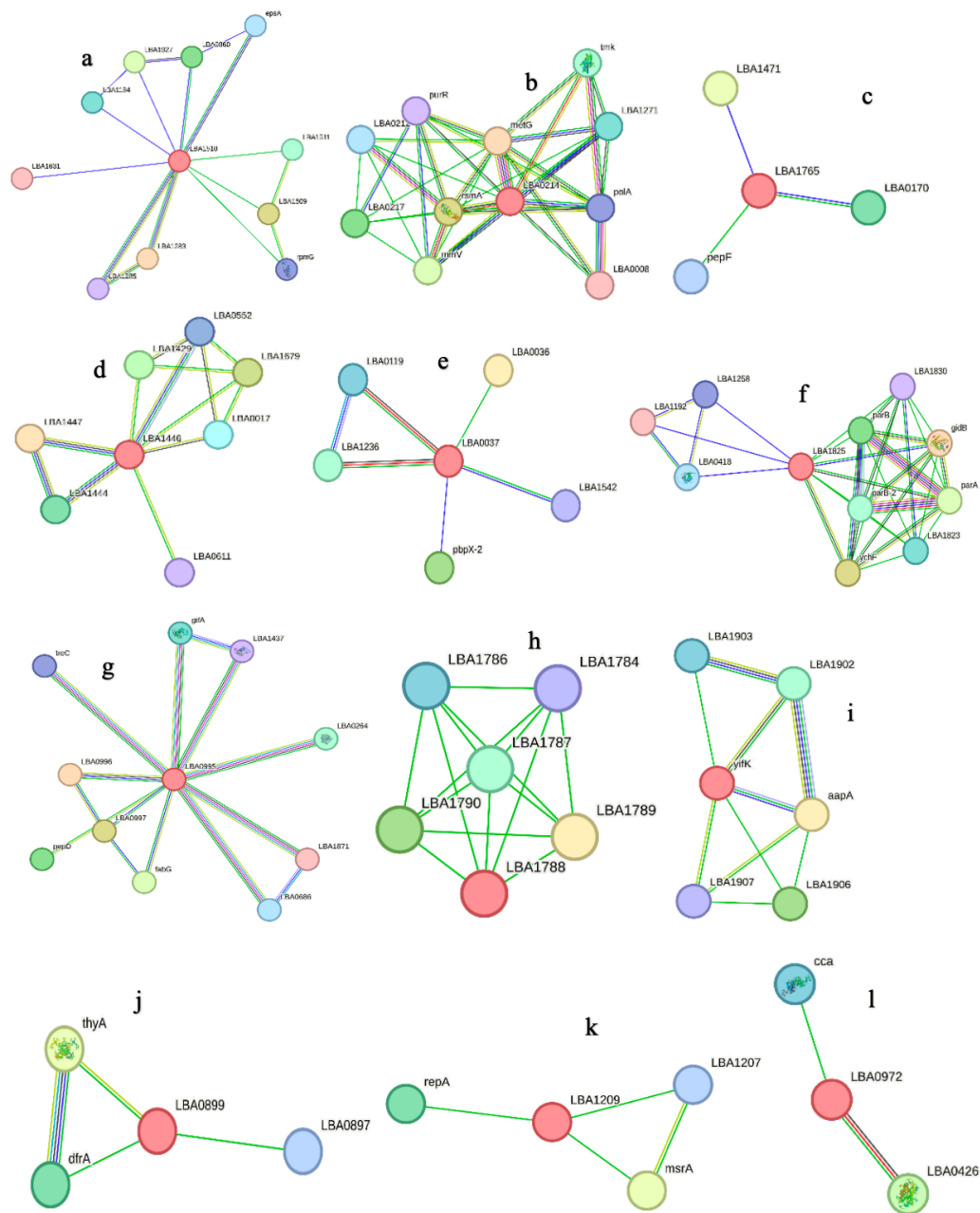
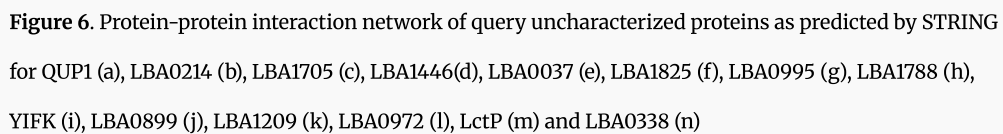
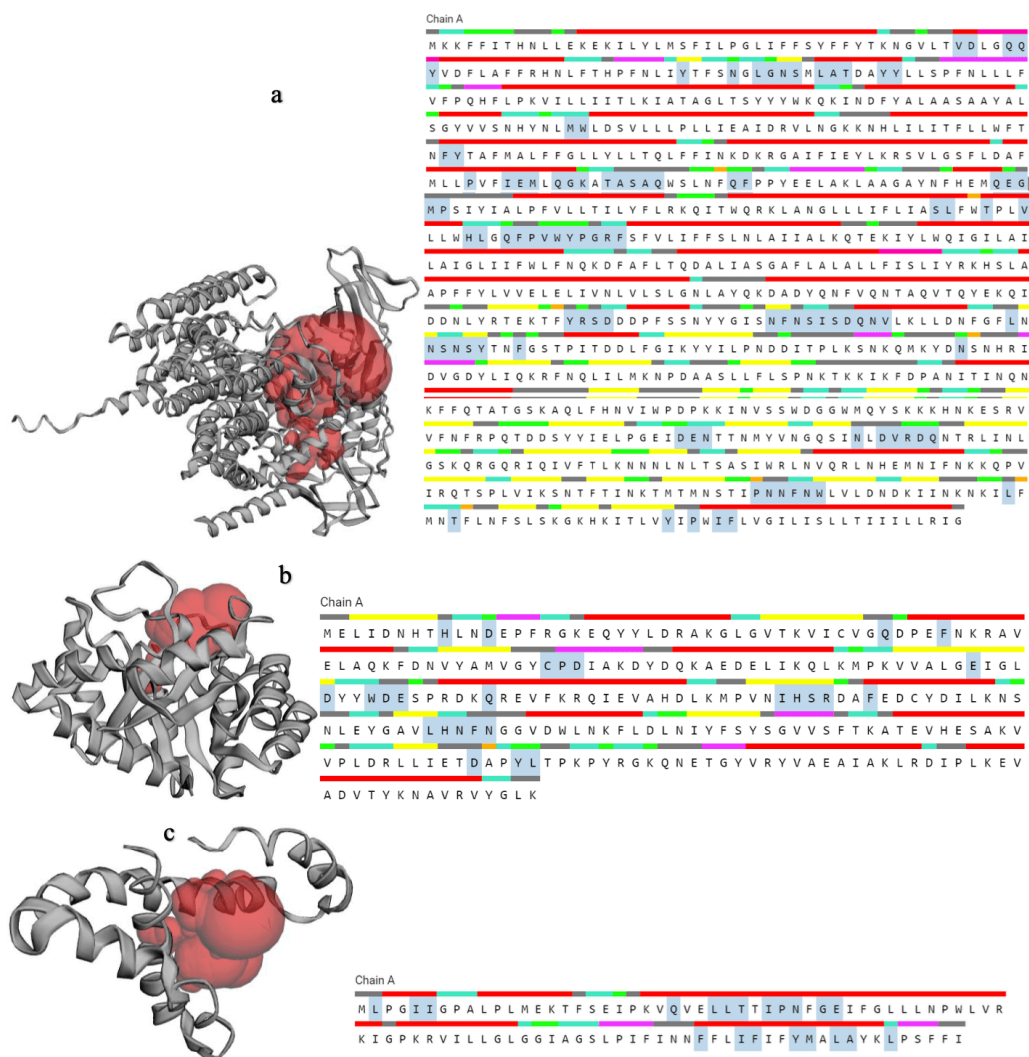
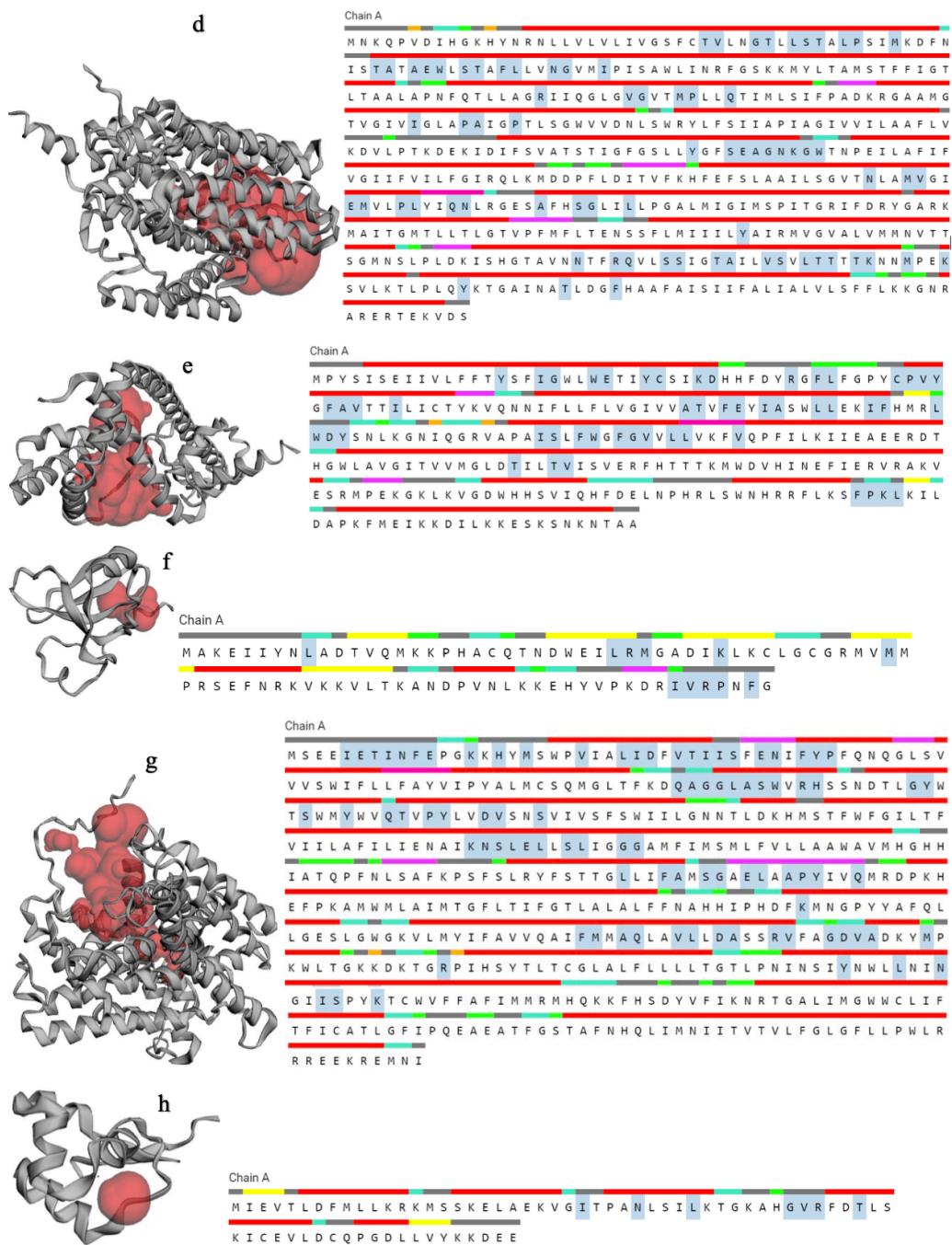


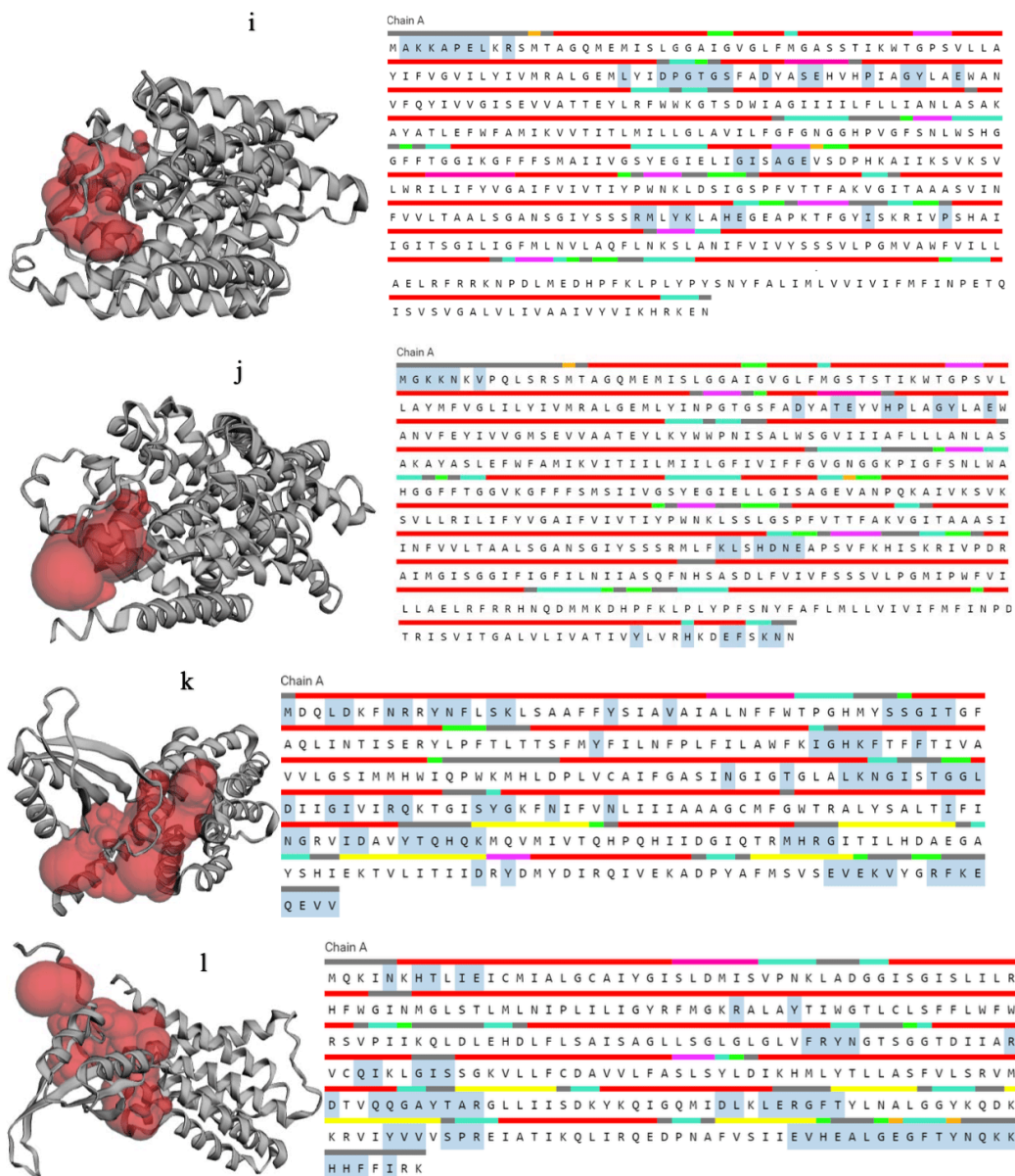
Figure 5. Ramachandran plots of query uncharacterized proteins as predicted for LBA1510 (a), LBA0214 (b), LBA1705 (c), LBA1446(d), LBA0037 (e), LBA1825 (f), LBA0995 (g), LBA1788 (h), YIFK (i), LBA0899 (j), LBA1209 (k), LBA0972 (l), LctP (m) and LBA0338 (n)











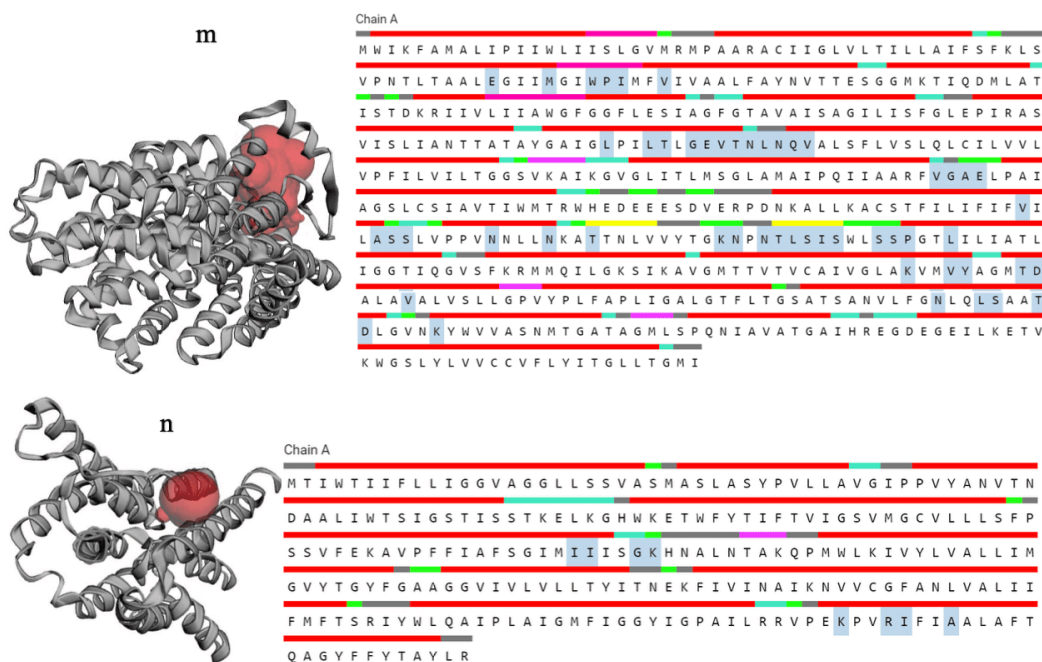


Figure 7. Active binding sites of query uncharacterized proteins for LBA1510(a), LBA0214 (b), LBA1705 (c), LBA1446(d), LBA0037 (e), LBA1825 (f), LBA0995 (g), LBA1788 (h), YIFK (i), LBA0899 (j), LBA1209 (k), LBA0972 (l), LctP (m) and LBA0338 (n). The red spots denote the active sites while chain A graphics show the amino residues

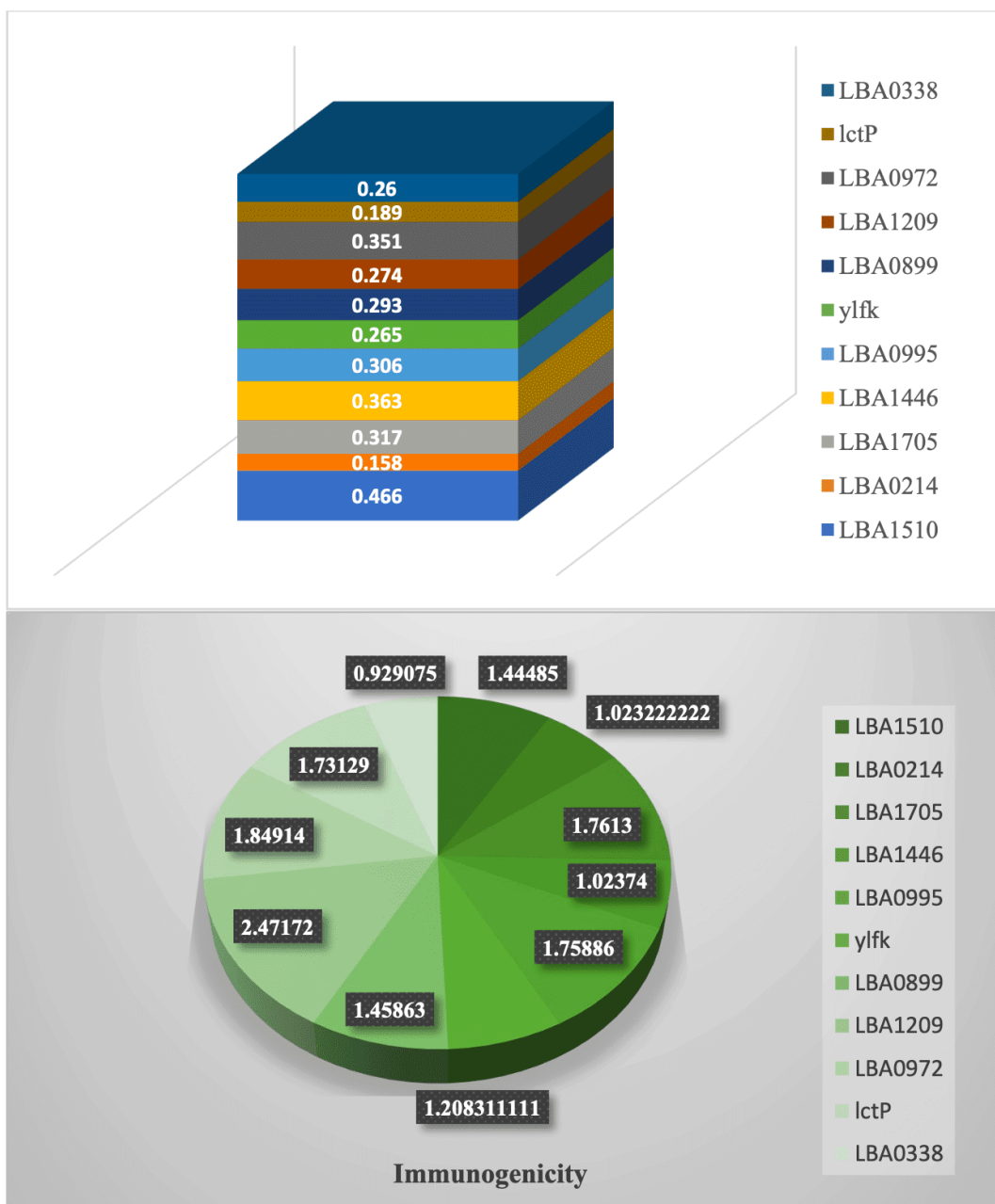


Figure 8. Immunogenic IL-6-inducing peptides from uncharacterized proteins of *L. acidophilus* reference genome: a. according to their IL-6-inducing prediction values; b. according to their immunogenicity values

This figure can be downloaded from the Supplementary Data section.

Figure 9. Gene expression levels of IL-6 and positively co-expressed genes with IL-6 in *Homo sapiens* and *Mus musculus* across various conditions: a. a scatter plot (Log2 scale) showing IL-6 expression across various anatomical parts and cell types in *Homo sapiens*; b. a scatter plot (Log2 scale) showing IL-6 expression across various anatomical parts and cell types in *Mus musculus*; c. a heatmap showing IL-6 expression (linear scale) in *Homo sapiens* according to various cancer categories; d. a circular view of positively co-expressed genes with IL-6 (top 50) according to cancer categories in *Mus musculus*.

Statements and Declarations

Ethics

Ethics approval was not applicable to this in silico study as no human or animal subjects were involved.

Conflicts of Interest

None.

References

1. [△]Zampiga MJ, Flees A, Meluzzi SD, Sirri F. Application of omics technologies for a deeper insight into qualitative production traits in broiler chickens: a review. *J. Anim. Sci. Biotechnol.* (2018) 61:1–18.
2. [△]Jadhav K, Sharma KS, Katoch S, Sharma VK, Mane BG. Probiotics in broiler poultry feeds: a review. *Int. J. Anim. Vet. Sci.* (2015) 2:4–16.
3. [△]Vieco-Saiz N et al. Benefits and inputs from LAB and their bacteriocins as alternatives to antibiotic growth promoters during food–animal production. *Front. Microbiol.* (2019) 10: 57–65.
4. [△]Mahfuz SU, Nahar MJ, Chen M, Zhang G, Zhongjun L, Hui S. Inclusion of probiotic on chicken performance and immunity: a review. *Int. J. Poult. Sci.* (2017) 16:328–335.
5. [△]FAO/WHO. Guidelines for the Evaluation of Probiotics in Food. Paris: FAO (2002) 1–11.
6. [△]Hill C, Guarner F, Reid G, Gibson GR, Merenstein DJ, Pot B, et al. Expert consensus document. The International Scientific Association for Probiotics and Prebiotics consensus statement on the scope and appropriate use of the term probiotic. *Nat. Rev. Gastroenterol. Hepatol.* (2014) 11:506–514. doi: 10.1038/nrgastro.2014.66

7. [△]Ashrafuzzaman M, Momtaz F, Foysal J, Ali Md H, Rahman MM. Optimization of probiotic activity of *Lactobacillus acidophilus* characterized from commercial yoghurt. *Int. J. Biosci.* (2015) 7:94–99.
8. [△]Anadon AM, Larrañaga RM, Ares I, Martínez M. Probiotics: safety and toxicity considerations. *Nutraceuticals* (2016) 65:1081–1105.
9. [△]Dhama K, Verma V, Sawant PM, Tiwari R, Vaid RK, Chauhan RS. Applications of probiotics in poultry: enhancing immunity and beneficial effects on production performances and health- a review. *J. Immunol. Immunopathol.* (2011) 13:1–19.
10. [△]Denli M, Demirel R. Replacement of antibiotics in poultry diets. *CAB Rev* 2018;13:35
11. [△]Shang YS, Kumar S, Oakley B, Woo KK. Chicken gut microbiota: importance and detection technology. *Front. Vet. Sci.* (2018) 5:254.
12. [△]Sugiharto YT, Isroli I, Widiastuti, Wahyuni HI. Hematological parameters and selected intestinal microbiota populations in the Indonesian indigenous crossbred chickens fed basal diet supplemented with multi strain probiotic preparation in combination with vitamins and minerals. *Vet. World* (2018) 11:874–882.
13. [△]Noohi N, Papizadeh M, Rohani M, Talebi M, Pourshafie MR. Screening for probiotic characters in lactobacilli isolated from chickens revealed the intra-species diversity of *Lactobacillus brevis*. *Anim. Nutr.* (2021) 7:119–26. <https://doi.org/10.1016/j.aninu.2020.07.005>.
14. [△]Sirisopapong M, Shimosato T, Okrathok S, Khempaka S. Assessment of lactic acid bacteria isolated from the chicken digestive tract for potential use as poultry probiotics. *Animal Bioscience* (2023) 3:1209–1220. <https://doi.org/10.5713/ab.22.0455>.
15. [△]^bKassa G, Alemayehu D, Andualemy B. Isolation identification and molecular characterization of probiotic bacteria from locally selected Ethiopian free range chickens gastrointestinal tract. *Poultry Science* (2024) 103:03311. <https://doi.org/10.1016/j.psj.2023.103311>
16. [△]Kizerwetter-Świda M, Binek M. Assessment of potentially probiotic properties of *Lactobacillus* strains isolated from chickens. *Pol. J. Vet. Sci.* (2016) 19:15–20. <https://doi.org/10.1515/pjvs-2016-0003>
17. [△]Al-Khalaija H, Al-Nasser A, Al-Surayee T et al. Effect of dietary probiotics and prebiotics on the performance of broiler chickens. *Poult. Sci.* (2019) 98:4465–79 <https://doi.org/10.10382/ps/pez282>
18. [△]Ahmad R. et al Influence of heat stress on poultry growth performance intestinal inflammation and immune function and potential mitigation by probiotics. *Animals* (2022) 12: 2297. <https://doi.org/10.3390/ani12172297>.
19. [△]Suravajhala P, Benso A, Valadi JK. Annotation and curation of uncharacterized proteins: systems biology approaches. *Front. Genet.* (2015) 6:224; doi:10.3389/fgene.2015.00224.

20. ^a ^bYu NY. et al. PSORTb 30: improved protein subcellular localization prediction with refined localization subcategories and predictive capabilities for all prokaryotes. *Bioinformatics* (2010) 26 (13):1608-1615.
21. ^ΔGeourjon C, Deleage G. SOPMA: significant improvements in protein secondary structure prediction by consensus prediction from multiple alignments. *Comput. Appl. Biosci.* (1995) 11:681-684.
22. ^a ^bWaterhouse A. et al. SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic Acids Res.* (2018) 46:W296-W303.
23. ^a ^b ^c ^d ^e ^f ^g ^hGasteiger E. In: *The Proteomics Protocols Handbook* (ed. Walker, J.M.) Humana Press (2005). p. 571–607
24. ^ΔSzklarczyk D. et al. STRING v10: protein-protein interaction networks, integrated over the tree of life. *Nucleic Acids Res.* (2015) 43:D447-52. Doi:101093/nar/gku1003.
25. ^a ^b ^cNguyen MN, Krutz NL, Limviphuvadh V, Lopata AL, Gerberick GF, Maurer-Stroh S. AllerCatPro 20: a web server for predicting protein allergenicity potential. *Nucleic Acids Res.* (2022) 50 Web Server issue <https://doi.org/101093/nar/gkac446>.
26. ^a ^b ^cDimitrov I, Zaharieva N, Doytchinova I. Bacterial immunogenicity prediction by machine learning methods. *Vaccines* (2020) 8:709. doi:103390/vaccines8040709.
27. ^a ^bSharma N, Naorem LD, Jain S, Raghava GPS. ToxinPred2: an improved method for predicting toxicity of proteins. *Briefings in Bioinformatics* (2022) 23(5):1–12; <https://doi.org/101093/bib/bbac174>.
28. ^ΔMaurer-Stroh S. et al. AllerCatPro–prediction of protein allergenicity potential from the protein sequence. *Bioinformatics* (2019) 35:3020–3027.
29. ^a ^bTian W, Chen C, Lei X, Zhao J, Liang J. CASTp 30: computed atlas of surface topography of proteins. *Nucleic Acids Res.* (2018) 46:W363–W367; <https://doi.org/101093/nar/gky473>.
30. ^ΔKolker E. et al. Identification and functional analysis of 'hypothetical' genes expressed in *Haemophilus influenza*. *Nucleic Acids Res.* (2004) 32(8):2353–2361; <https://doi.org/101093/nar/gkh555>.
31. ^ΔMazandu GK, Mulder NJ. Function prediction and analysis of *Mycobacterium tuberculosis* hypothetical proteins. *Int. J. Mol. Sci.* (2012) 13:7283–302; Doi: 103390/ijms13067283.
32. ^ΔMohan R, Venugopal S. Computational structure and functional analysis of hypothetical proteins of *Staphylococcus aureus*. *Bioinformation* (2012) 8(15):722–8; Doi:106026/97320630008722.
33. ^ΔPark SJ, Son WS, Lee BJ. Structural analysis of hypothetical proteins from *Helicobacter pylori*: an approach to estimate functions of unknown or hypothetical proteins. *Int. J. Mol. Sci.* (2012) 13(6):7109–7137; Doi:103390/ijms13067109.

34. ^{a, b}Enany S. Structural and functional analysis of hypothetical and conserved proteins of *Clostridium tetani*. *Journal of Infection and Public Health* (2014) 7:296–307.
35. ^ΔRahman MF. et al. A bioinformatics approach to characterize a hypothetical protein Q6S8D9 SARS of SARS-CoV-2. *Genomics Inform.* (2023) 21: (1):e3; <https://doi.org/10.5808/gi22021>
36. ^ΔAdejumo IO. Hypothetical proteins of chicken-isolated *Limosilactobacillus reuteri* subjected to in silico analysis induce IL-2 and IL-10. *Genes & Nutrition.* (2024) 19:21; <https://doi.org/10.1186/s12263-024-00755-4>
37. ^ΔKyte J, Doolittle RF. A simple method for displaying the hydropathic character of a protein. *J. Mol. Biol.* (1982) 157, 105–132.
38. ^ΔFernandez-Fernandez AD, Corpas FJ. In silico analysis of *Arabidopsis thaliana* peroxisomal 6-phosphogluconate dehydrogenase. *Scientifica (Cairo)* (2016) 3482760.
39. ^ΔBlackburn K, N'jai AU, Dearman RJ, Kimber I, Gerberick GF. Respiratory allergenic potential of plant-derived proteins: understanding the relationship between exposure and potency for risk assessments. *Crit. Rev. Toxicol.* (2015) 45(9):799–811.
40. ^ΔGardy JL. PSORTb 2.0: expanded prediction of bacterial protein subcellular localization and insights gained from comparative proteome analysis. *Bioinformatics* (2005) 21:617–623.
41. ^ΔGardy JL, Brinkman FSL. Methods for predicting bacterial protein subcellular localization. *Nat. Rev. Microbiol.* (2006) 4:741–751.
42. ^ΔKamimura D, Ishihara K, Hirano T. IL-6 signal transduction and its physiological roles: the signal orchestration model. *Rev. Physiol. Biochem. Pharmacol.* (2003) 149:1–38. doi: 10.1007/s10254-003-0012-2
43. ^{a, b}Hunter CA, Jones SA. IL-6 as a keystone cytokine in health and disease. *Nat. Immunol.* (2015) 16:448–57. doi: 10.1038/ni.3153
44. ^ΔHeinrich PC, Castell JV, Andus T. Interleukin-6 and the acute phase response. *Biochem. J.* (1990) 265:621–636.
45. ^ΔGillmore JD, Lovat LB, Persey MR, Pepys MB, Hawkins PN. Amyloid load and clinical outcome in AA amyloidosis in relation to circulating concentration of serum amyloid A protein. *Lancet* (2001) 358:24–29.
46. ^{a, b}Tanaka T, Narazaki M, Kishimoto T. IL-6 in Inflammation, Immunity, and disease. *Cold Spring Harb. Perspect. Biol.* (2014) 6:a016295
47. ^ΔNemeth E, Rivera S, Gabayan V, Keller C, Taudorf S, Pedersen BK, Ganz T. IL-6 mediates hypoferremia of inflammation by inducing the synthesis of the iron regulatory hormone hepcidin. *J. Clin. Invest.* (2004) 113:1271–1276.
48. ^ΔIshibashi T, Kimura H, Shikama Y, Uchida T, Kariyone S, Hirano T, Kishimoto T, Takatsuki F, Akiyama Y. Interleukin-6 is a potent thrombopoietic factor in vivo in mice. *Blood* (1989) 74:1241–1244.

49. ^ΔKuilman T, Michaloglou C, Vredevelde LC. et al. Oncogene-induced senescence relayed by an interleukin dependent inflammatory network. *Cell* (2008) 133:1019.
50. ^ΔGrivennikov S, Karin E, Terzic J. et al. IL-6 and Stat3 are required for survival of intestinal epithelial cells and development of colitis-associated cancer. *Cancer Cell* (2009) 15:103.
51. ^ΔYu H, Pardoll D, Jove R. STATs in cancer inflammation and immunity: a leading role for STAT3. *Nat. Rev. Cancer* (2009) 9:798.
52. ^ΔGrivennikov SI, Karin M. Dangerous liaisons: STAT3 and NF-kappaB collaboration and crosstalk in cancer. *Cytokine Growth Factor Rev.* (2010) 21:11.
53. ^ΔYoon S, Woo SU, Kang JH. et al. NF-κB and STAT3 cooperatively induce IL6 in starved cancer cells. *Oncogene* (2012) 31:3467.
54. ^ΔHirano T, Murakami M. COVID-19: a new virus, but a familiar receptor and cytokine release syndrome. *Immunity* (2020) 52:731.
55. ^ΔJones G, Sebba A, Gu J. et al. Comparison of tocilizumab monotherapy versus methotrexate monotherapy in patients with moderate to severe rheumatoid arthritis: the AMBITION study. *Ann. Rheum. Dis.* (2010) 69:88.
56. ^ΔXu X, Han M, Li T. et al. Effective treatment of severe COVID-19 patients with tocilizumab. *Proc. Natl Acad. Sci. USA* (2020) 117:10970.
57. ^ΔHirano T. IL-6 in inflammation, autoimmunity and cancer. *International Immunology* (2020) 33:127–148. doi:10.1093/intimm/dxaa078
58. ^ΔAdebisi FG, Ologhobo AD, Adejumo IO. Efficacy of *Allium sativum* as growth promoter, immune booster and cholesterol-lowering agent on broiler chickens. *Asian Journal of Animal Science* (2017) 11: 202-213
59. ^ΔOloghobo AD, Akangbe E, Adejumo IO, Ere R, Agboola B. Haematological and Histological Evaluation of African Mistletoe (*Viscum album*) Leaf Meal as Feed Additive for Broilers. *Annual Research & Review in Biology* (2017) 15(3): 1-7
60. ^ΔAdejumo IO, Adetunji CO, Olopade CO, George KO. Effect of *Moringa oleifera* leaf meal and differently processed seed meal as a replacement for synthetic antibiotics in broiler diet. *Journal of Animal Production Research* (2016) 28(2):200-206
61. ^ΔOloghobo AD, Adejumo IO. Haematological response and serum biochemical profile of broiler finishers fed with oxytetracycline and stonebreaker (*Phyllanthus amarus*) leaf meal. *Biotechnology Journal International* (2015) 7(1):51-56, DOI: 109734/BBJ/2015/10304

Supplementary data: available at <https://doi.org/10.32388/MR9OH9>

Declarations

Funding: No specific funding was received for this work.

Potential competing interests: No potential competing interests to declare.