Biotechnology in Animal Husbandry 33 (3), p 309-319 , 2017
Publisher: Institute for Animal Husbandry, Belgrade-Zemun
ISSN 1450-9156
UDC 614.91'639
https://doi.org/10.2298/BAH1703309Y

PROTEINS SEQUENCE ANALYSIS OF CONTAGIOUS CAPRINE PLEUROPNEUMONIA

Ayuba Dauda [1],Abdulmojeed Yakubu[2], Ihe Ndu Dim[1], Deeve Sebastian Gwaza [1]

[1]University of Agriculture Makurdi, P. M. B 2373 Makurdi, Nigeria
[2]State University Lafia Campus Nasarawa State, Nigeria
Corresponding author: ayubadauda87@gmail.com

Abstract. A total of twenty (20) contagious bovine pleuropneumonia (CCPP) proteins were retrieved from the GenBank (www.ncbi.nlm.nih.gov). The proteins sequences were used to investigate the molecular identity of various CCPP proteins. The physico-chemical properties of CCPP proteins were performed using protparam tool. Isoelectric point (pI), molecular weight (MW), extinction coefficient (EC); instability index (II), aliphatic index (AI) and grand average of hydropathicity (GRAVY) were computed. The study revealed that the pI of CCPP proteins were acidic and basic in nature. The EC and II of CCPP proteins indicate better stability which is an indication of resistant to mutation and thermally stable. The GRAVY of CCPP proteins revealed some are positive while some are negative. The positive value indicates solubility (hydrophilic) in water while negative is not soluble (hydrophobic) in water. The amino acid composition of CCPP proteins indicates that they are rich in isoleucine, leucine and lysine. The three dimensional structures (3D) of the CCPP proteins were determine using Phyre2 server. The amino acid sequences of CCPP proteins were subjected to secondary structure prediction using ExPASy's SOPMA tool. The proteins are more of alpha helix structure. The genetic information eminating from this study may bring insight into mutagenesis and pharmacogenetic.

Keywords: protein, caprine pleuropneumonia, Sequence

https://doi.org/10.2298/BAH1803369E

Introduction

Contagious Caprine Pleuropneumonia (CCPP) is a devastating disease of goats included in the list of notifiable diseases of the Organization for Animal Health (OIE).The first description of the disease dates back to 1873, in Algeria *(Thomas, 1873).*CCPP is a contagious disease of goats, which occurs in per acute,

acute or chronic forms and is characterized by fibrinous pneumonia, pleurisy and profuse pleural exudates *(Edelsten et al., 1990).* Mortality rates of 60–100% are common *(Edelsten et al., 1990).* The disease is reported to occur in many countries in West and Eastern Africa and in Pakistan and India *(OIE, 2001)*. The infectious agent *Mycoplasma capricoleum* subspecies *capripneumoniae*, formerly known as the F38-like group, is difficult to isolate and has only been identified in a few of the countries where the disease has been reported *(Bolske et al., 1995a).*

Materials and Methods

A total of twenty (20) CCPP proteins of goat were retrieved from the GenBank (www.ncbi.nlm.nih.gov). The Genbank accession numbers of the sequences and sequence variations are shown in Table 1. ProtParam Tool was used for the computation of various physical and chemical properties of the CCPP proteins using amino acid sequences. The computated parameters were molecular weight, theoretical pI (isoelectric point), amino acid composition, extinction coefficient, estimated half-life, instability index, aliphatic index and grand average of hydropathicity (GRAVY) *(Gasteiger, 2005).* The amino acid sequences of CCPP proteins were subjected to secondary structure prediction using ExPASy's SOPMA tool. It predicts 69.5% of amino acids for a 3 state description of the secondary structure (a helix, b sheets and coil). The Phyre2 server was used to predict the 3D structure of CCPP proteins. These servers predict the three-dimensional structure of a protein sequence using the principles and techniques of homology modeling *(Kelley and Sternberg, 2009).* Currently, the most powerful and accurate methods for detecting and aligning remotely related sequences rely on profiles or Hidden Markov Models (HMMs). 3DligandSite was used to predict the binding site of the 3D structure of the CCPP proteins. Phyre2 is coupled to the 3DligandSite server for protein binding site prediction *(Wass et al., 2010).*

Results

Physico-chemical characteristics of CCPP proteins predicted by protparam are shown in Table 2. The computed isoelectric point (pI) values of CCPP proteins in the study revealed Phosphoglycerate kinase, Glycyl-tRNA Synthetase, ATP-dependent protease La, GTP-Binding protein, tRNA Modification GTpass, Lysine-tRNA ligase and Chromosome segregation ATPase are acidic which have (pI<7) while the rest appeared to be basic in nature with (pI>7). The net charge of CCPP protein revealed only Phosphoglycerate kinase is neutral (no charge). Glycyl-tRNA Synthetase, ATP-dependent protease La, GTP-Binding protein, tRNA

Modification, Lysine-tRNA ligase and Chromosome segregation ATPase are negatively (-) charge while the rest of the protein are positively (+) charge. The extinction coefficient of a protein at 280 nm depends almost exclusively on the number of aromatic residues, particularly tryptophan *(Gill et al., 1989).* Extinction coefficient values for CCPP proteins at 280 nm ranged from 8940 to 143950 (Signal recognition particle protein is lowest and Prolipoprotein diacylglyceryl tranferase is highest) respectively.

Table 1: Protein name, accession, and number of goat CCPP

S/N	NAME OF PROTEIN	ACCESSION No	AMINO ACID No
1	Phosphoglycerate kinase	KEY8461	404
2	Chaperone protein Dnaj	KEY84219	372
3	Amino acid permease	KEY84758	515
4	Glycyl-tRNA Synthetase	KEY84567	456
5	GTP pyrophosphokinase	KEY84560	754
6	ATP-dependent protease La	KEY84722	779
7	DNA-primase	KEY84560	604
8	Histidy-tRNA Synthetase	KEY84179	414
9	GTP-Binding protein	KEY84763	364
10	Excinuclease ABC Subunit B	KEY84755	665
11	tRNA Modification GTpass	KEY84779	452
12	Tyrosine-TRNA ligase	KEY84654	414
13	PTS system-IIBC component	KEY84580	602
14	Hypothetical Protein Mccp_2340	KEY84577	820
15	Dihydro folate-foly poly glutamate synthase	KEY84753	369
16	Lysine-tRNA ligase	KEY84440	500
17	Prolipoprotein diacylglyceryl tranferase	KEY84751	526
18	Chromosome segregation ATPase	KEY84561	988
19	Cell division protein FtsY	KEY84539	424
20	Signal recognition particle protein	KEY84596	447

The half life of protein is the time it takes for half of the amount of protein in a cell to disappear after its synthesis in the cell of the proteins. In this study the half life of all the CCPP proteins is 30 hours.The instability index provides an estimate of the stability of the protein in a test tube. A protein whose instability index is smaller than 40 is predicted as stable, a value above 40 predicts that the protein will be unstable *(Guruprasad et al., 1990).* The result from this study shown that ATP-dependent protease La, Excinuclease ABC Subunit B, Prolipoprotein diacylglyceryl tranferase and Cell division protein FtsY protein have value >40 while the rest of protein are have<40.The aliphatic index of a protein is defined as the relative volume occupied by aliphatic side chains (alanine, valine, isoleucine, and leucine). The result revealed that Amino acid permease, DNA-primase, GTP-Binding protein, tRNA Modification GTpass, PTS system-

IIBC component, Hypothetical Protein Mccp 3340, Dihydro folate-foly poly glutamate synthase and Prolipoprotein diacylglyceryl tranferase proteins from this study have AI>100 while the rest of the CCPP protein have AI<100. The grand average hydropathicity **(**GRAVY) of the CCPP protein revealed that Amino acid permease, tRNA Modification GTpass and PTS system-IIBC component have positive while the rest of CCPP protein have negative value.

Table 2: Physico-chemical characteristic of proteins of CCPP predicted by protoparam

Protein	AA	Mol Wt	PI	Q	EC	Half Life	II	AI	GRAVY
Phosphoglycerate kinase	404	44636.4	6.94	Neu	38055	30hrs	25.67	97.97	-0.188
Chaperone protein Dnaj	370	42107.8	8.75	+	22850	30hrs	34.01	83.31	-0.545
Amino acid permease	515	57092.7	9.54	+	55725	30hrs	31.24	118.72	0.685
Glycyl-tRNA Synthetase	456	53446.9	6.58	-	76110	30hrs	35.91	84.65	-0.510
GTP pyrophosphokinase	754	87111.5	9.11	+	90900	30hrs	30.95	98.41	-0.409
ATP-dependent protease La	779	88439.6	5.62	-	65210	30hrs	47.76	98.70	-0.400
DNA-primase	604	69930.8	8.84	+	44155	30hrs	30.70	106.71	-0.404
Histidy-tRNA Synthetase	414	48625	8.61	+	41510	30hrs	38.56	93.94	-0.526
GTP-Binding protein	364	40671	5.23	-	27705	30hrs	28.80	105.27	-0.079
Excinuclease ABC Subunit B	665	76544.7	7.61	+	44030	30hrs	47.91	99.71	-0.395
tRNA Modification GTpass	452	50443.1	7.81	+	19035	30hrs	33.55	124.65	0.001
Tyrosine-TRNA ligase	414	47159.2	8.53	+	42650	30hrs	33.25	96.11	-0.252
PTS system-IIBC component	602	66660.3	9.43	+	47120	30hrs	22.68	134.07	0.715
Hypothetical Protein Mccp 3340	870	95388.3	8.95	+	105785	30hrs	31.37	100.11	-0.436
Dihydro folate-foly poly glutamate synthase	369	42100.3	8.78	+	45060	30hrs	18.78	112.74	-0.097
Lysine-tRNA ligase	500	58160.4	5.49	-	40925	30hrs	36.18	90.42	-0.465
Prolipoprotein diacylglyceryl tranferase	526	63020.6	9.43	+	143950	30hrs	40.46	104.89	-0.025
Chromosome segregation ATPase	988	112835.3	5.71	-	52845	30hrs	39.46	96.13	-0.514
Cell division protein FtsY	424	48485.7	8.80	+	29910	30hrs	87.87	85.54	-0.720
Signal recognition particle protein	447	50114.3	9.39	+	8940	30hrs	39.54	96.62	-0.364

AA=amino acid; pI=isoelectric point; Q=net charge; II=instability index; AI=alphatic index; GRAVY= grand average of hydropathicity ; EC= extinction coefficient; Mol wt=molecular weight, +=amino acid resides that positively charge, -= amino acid resides that negatively charge, Neut= amino acid resides that are neutral

The prediction of secondary structure of CCPP proteins is shown in Table 3. The result revealed that Signal recognition particle protein showed the highest alpha helix (53.91%) and the lowest is Chaperone protein Dnaj (26.88%). The extended strand prediction, Dihydro folate-foly poly glutamate synthase gives

highest value (28.18%) and the lowest is Signal recognition particle protein (13.87%). The beta turn prediction of secondary structure revealed that Chaperone protein Dnaj gives the highest value (14.25%) and Chromosome segregation ATPase is the lowest (5.67%). The random coil prediction of secondary structure revealed that Chaperone protein Dnaj gives the highest value (35.22%) and Chromosome segregation ATPase showed the lowest value (20.04%). All the CCPP proteins are having higher value in alpha helix structure. The amino acid composition percentage of CCPP protein is shown in Table 4. All the CCPP proteins used for this study have similar amino acid composition of all the CCPP protein with higher percentage in isoleucine, leucine and lysine. Isoleucine and leucine are aliphatic amino acid and lysine is polar amino amino acid. All the CCPP proteins have zero percentage composition of selenocystein and pyrrolysine amino acids.

Table 3: Prediction of secondary structures of CCPP protein

Protein	Alpha Helix (%)	Extended Strand (%)	Beta Turn (%)	Random Coil (%)
Phosphoglycerate kinase	42.57	21.78	12.62	23.02
Chaperone protein Dnaj	26.88	23.66	14.25	35.22
Amino acid permease	39.03	27.57	10.29	23.11
Glycyl-tRNA Synthetase	37.06	20.61	10.53	31.80
GTP pyrophosphokinase	50.13	20.56	7.43	21.88
ATP-dependent protease La	43.21	17.33	8.34	28.11
DNA-primase	47.02	21.36	7.78	23.84
Histidy-tRNA Synthetase	37.68	20.77	8.94	32.61
GTP-Binding protein	49.45	19.23	10.16	21.15
Excinuclease ABC Subunit B	31.13	17.14	10.23	21.50
tRNA Modification GTpass	45.80	22.35	7.74	24.12
Tyrosine-TRNA ligase	48.79	18.60	10.87	21.74
PTS system-IIBC component	40.37	26.08	10.63	22.92
Hypothetical Protein Mscp 3340	51.59	16.22	7.44	24.76
Dihydro folate-foly poly glutamate synthase	41.19	28.18	8.67	21.95
Lysine-tRNA ligase	40.60	22.40	7.60	29.40
Prolipoprotein diacylglyceryl tranferase	38.40	24.52	10.08	27.00
Chromosome segregation ATPase	58.20	16.09	5.67	20.04
Cell division protein FtsY	52.83	15.09	9.20	22.88
Signal recognition particle protein	53.91	13.87	8.41	24.16

Parameters:
Window width: 17
Similarity threshold: 8
Number of states: 4

Table 4: Amino acid composition (%) of CCPP proteins

Protein	A	R	N	D	C	Q	E	G	H	I	L	K	M	F	P	S	T	W	Y	V	O	U
Phosphoglycerate kinase	7.2	2.0	5.9	6.2	0.5	2.7	7.4	7.7	1.0	8.7	8.9	11.6	1.5	4.7	2.5	5.7	5.2	1.2	1.7	7.7	0.0	0.0
Chaperone protein Dnaj	3.0	2.4	9.9	6.5	2.2	4.6	5.6	6.7	1.3	10.2	7.0	11.6	1.3	5.1	1.9	9.1	3.0	0.0	4.0	4.6	0.0	0.0
Amino acid permease	7.6	2.7	4.7	3.3	1.2	3.3	4.7	7.6	1.6	11.7	11.7	5.6	2.9	9.1	1.9	8.5	3.9	1.2	2.9	7.0	0.0	0.0
Glycyl-tRNA Synthetase	3.3	4.2	8.1	6.1	1.1	5.2	7.2	5.0	1.1	8.3	9.4	9.0	1.8	7.0	2.9	6.4	3.7	2.2	3.1	4.2	0.0	0.0
GTP pyrophosphokinase	5.0	3.6	6.5	5.3	0.7	4.1	7.7	3.7	1.5	12.6	7.7	11.8	1.6	3.3	2.1	6.8	5.6	0.9	4.6	4.9	0.0	0.0
ATP-dependent protease La	4.7	3.2	4.5	5.8	0.0	3.0	10.5	5.1	1.8	9.2	9.9	10.8	1.9	2.8	4.0	6.8	5.3	0.5	3.7	6.7	0.0	0.0
DNA-primase	3.1	1.3	10.9	6.1	1.0	3.6	6.5	2.3	1.2	13.6	10.6	12.9	0.8	4.1	2.5	6.3	4.8	0.3	3.6	3.1	0.0	0.0
Histidy-tRNA Synthetase	2.4	4.3	9.7	6.0	1.2	4.6	7.2	4.1	0.5	10.4	10.4	10.1	1.7	4.8	3.4	5.3	5.1	0.2	5.8	3.6	0.0	0.0
GTP-Binding protein	6.9	2.5	5.5	6.9	1.6	4.1	7.1	6.3	0.9	8.8	11.5	9.6	1.1	4.9	2.5	5.2	4.7	0.5	3.0	6.6	0.0	0.0
Excinuclease ABC Subunit B	5.6	5.6	7.4	5.7	0.8	5.6	7.4	3.5	1.7	9.7	11.1	7.7	1.8	3.8	2.6	5.4	6.8	0.3	3.3	5.0	0.0	0.0
tRNA Modification GTpass	5.3	2.7	10.2	6.6	0.7	3.3	7.7	5.1	0.7	13.3	12.4	7.7	1.5	2.7	1.1	6.2	4.0	0.2	2.0	6.6	0.0	0.0
Tyrosine-TRNA ligase	5.6	1.9	7.0	6.5	1.0	6.3	4.8	5.1	1.7	8.2	11.6	10.4	1.7	6.3	1.4	6.3	6.0	1.2	2.4	4.6	0.0	0.0
PTS system-IIBC component	6.1	1.3	6.5	3.2	0.8	3.0	2.2	8.6	1.5	14.1	13.1	7.4	1.8	6.6	3.0	6.1	4.8	0.8	2.2	7.5	0.0	0.0
Hypothetical Protein Mccp 3340	2.7	0.2	12.6	5.9	0.2	5.7	5.1	1.7	0.2	9.4	17.4	12.3	0.6	6.0	1.3	8.4	5.9	1.2	4.1	2.9	0.0	0.0
Dihydro folate-foly poly glutamate synthase	1.6	1.4	7.9	6.0	1.1	3.5	5.4	4.1	2.2	13.3	11.4	11.7	1.1	5.1	2.4	6.0	4.1	0.8	5.1	5.1	0.0	0.0
Lysine-tRNA ligase	4.4	6.0	5.6	7.2	0.4	3.2	9.4	5.0	2.6	7.6	9.4	7.6	3.0	5.0	3.4	4.6	4.4	0.4	4.0	6.8	0.0	0.0
Prolipoprotein diacylglyceryl tranferase	2.5	2.7	9.7	2.1	1.0	4.6	4.8	3.8	2.3	12.9	10.6	8.0	1.5	8.6	3.8	5.1	2.9	3.4	5.7	3.6	0.0	0.0
Chromosome segregation ATPase	5.7	4.4	7.9	6.1	0.2	4.7	9.4	3.8	0.8	9.8	9.6	9.9	1.5	3.5	1.5	6.0	6.5	0.2	2.8	5.1	0.0	0.0
Cell division protein FtsY	6.4	1.9	6.4	7.1	0.2	5.9	9.7	4.2	0.2	6.1	9.2	16.3	2.8	3.1	0.9	4.0	6.1	0.7	2.1	6.4	0.0	0.0
Signal recognition particle protein	5.6	4.3	6.9	4.7	0.0	5.1	8.1	7.4	0.4	7.4	11.6	11.2	4.5	3.6	2.2	5.4	4.5	0.0	1.3	5.8	0.0	0.0

A=Alanine, Arginine=R, Asparagini=N, Aspartic acid=D, cysteine=C, Glutamic acid=E, Glutamin=Q, Glycine=G, Histidine=H, Isoleucine=I, Leucine=L, Lysine=K, Methionine=M, Phenylalanine=F, Proline=P, Serine=S, Theonine=T, Tryptophan=W, Tyrosi

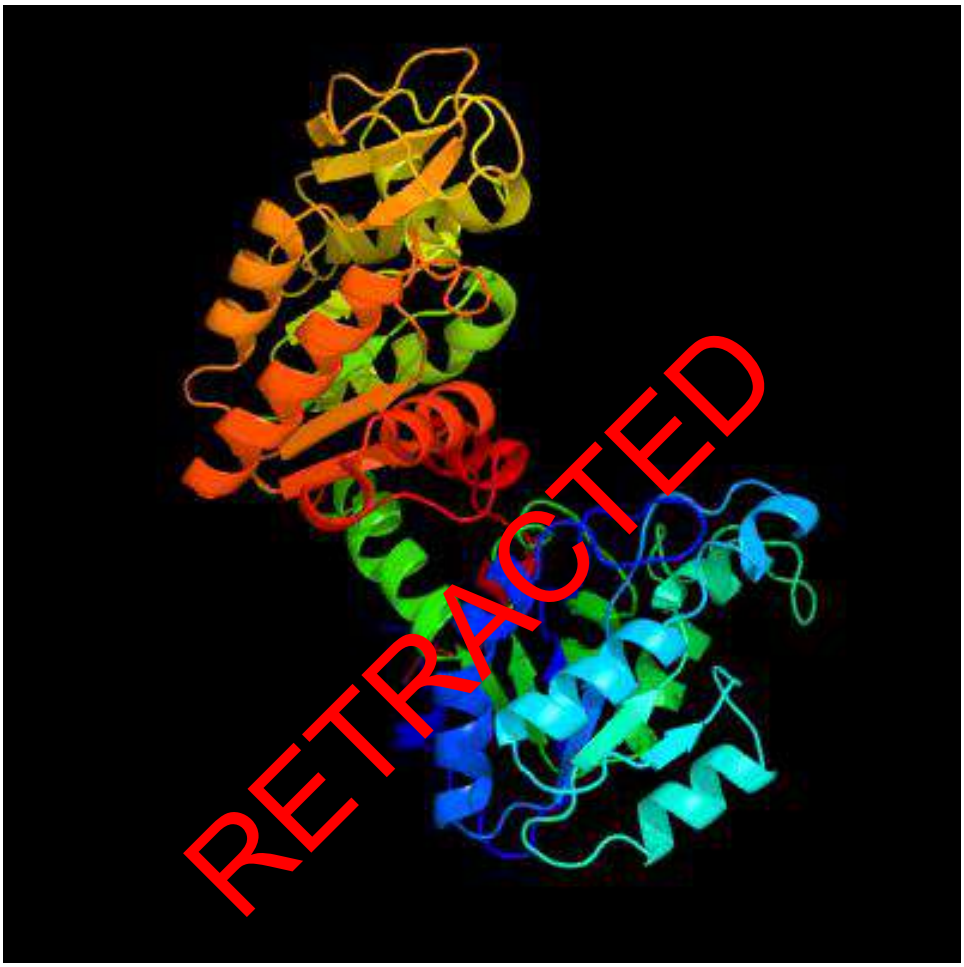

Image coloured by rainbow N → C terminus
Model dimensions (Å): X:69.827 Y:47.914 Z:70.606
Figure 1: Schematic 3D structure of goat Phosphoglycerate_kinase-caprine protein

Discussion

CCPP diseases disease notifiable to the World Organization for Animal Health (OIE) since it has a major impact on livestock production and a potential for rapid spread across national borders. As a result, CCPP-infected countries are excluded from international trade. At present, the disease causes vast problems in

Africa with severe socio-economical consequences. The computed isoelectric points (pI) for both CCPP will be useful for developing buffer system for purification by isoelectric focusing method. The isoelectric point is of significance in protein purification because it is the pH at which solubility is always minimal and at which mobility in an electro focusing system is zero and therefore the point at which the protein will accumulate *(Fennema, 2008).*

The extinction coefficient of a protein at 280 nm depends almost exclusively on the number of aromatic residues, particularly tryptophan *(Gill and Von-Hippel, 1989).* This indicates that the higher the EC value of the CCPP proteins, the higher the number of aromatic residues *(Gasteiger 2003; Munduganore et al., 2012)*. In particular, hydrophobic amino acids can be involved in binding/recognition of hydrophobic ligands such as lipids *(Betts et al., 2003*). All the CCPP proteins have zero selenocystein and pyrrolysine which is interpret as stop codons (protein cannot conclusively determine the identity of a residue) *(Sucharek et al., 2005).*

Many important biological processes such as cell signaling, transport of membrane-impermeable molecules, cell–cell communication, cell recognition and cell adhesion are mediated by membrane proteins *(Jones, 2007).* Although there has been some recent progress in predicting the full 3-D structure of transmembrane proteins (e.g. *Yarov-Yarovoy et al., 2006),* the most widely applied prediction technique for these proteins is to determine the transmembrane topology, i.e. the inside–outside location of the N and C termini relative to the cytoplasm, along with the number and sequence locations of the membrane spanning regions. This will facilitate the understanding of the structure and function of CCPP proteins.

Determining the structure and function of a novel protein is a cornerstone of many aspects of modern biology. The accuracy of protein structure prediction depends critically on sequence similarity between the query and template as observed in the present study. If a template is detected with >30% sequence identity to the query, then usually most or all of the alignment will be accurate and the resulting relative positions of structural elements in the model will be reliable *(Kelley et al., 2015).* The practical applications of CCPP protein structure prediction include guiding the development of functional hypotheses about hypothetical proteins, improving phasing signals in crystallography and selecting sites for mutagenesis *(Qian et al., 2007; Rava and Hussain, 2007).*

Conclusion

The physico-chemical properties, amino acid composition, and secondary structure of CCPP proteins indicated physical, chemical and thermal stability of the protein molecules. These indicated that the proteins are resistant to mutation and can withstand wide range of temperature. Genetic data revealed from this study

will bring new insights into epidemiological questions. Molecular typing has been instrumental in determining the population structure and evolution of pathogens. Since CCPP has both economical and nutritional consequences, efforts should be intensified towards finding sustainable genomic solutions to these deadly diseases which continue to ravage the livestock industry. New typing tool may help improve the surveillance and control of the disease, as well as to trace new epidemics.

Analiza sekvence proteina zarazne pleuropneumonije koza

Ayuba Dauda, Abdulmojeed Yakubu, Ihe Ndu Dim, Deeve Sebastian Gwaza

Rezime

Ukupno od dvadeset (20) proteina zarazne pleuropneumonije pluća goveda (contagious bovine pleuropneumonia - CCPP) je preuzeto iz GenBank-a (www.ncbi.nlm.nih.gov). Sekvence proteina korišćene su za ispitivanje molekularnog identiteta različitih CCPP proteina. Fizičko-hemijske osobine CCPP proteina su analizirane korišćenjem protparam alata. Isoelektrična tačka (pI), molekularna masa (MW), koeficijent ekstinkcije (EK); indeks nestabilnosti (II), alifatski indeks (AI) i veliki prosek hidropatičnosti (GRAVY). Studija je otkrila da su pI CCPP proteina bili kiseli i bazni po svojoj prirodi. EC i II proteina CCPP ukazuju na bolju stabilnost koja je indikacija otpornosti na mutaciju i toplotnu stabilnost. GRAVY CCPP proteina je otkrio da su neki pozitivni, dok su neki negativni. Pozitivna vrednost ukazuje na rastvorljivost (hidrofilni) u vodi, dok negativni nije rastvorljiv (hidrofobni) u vodi. Sastav amino kiselina proteina CCPP-a ukazuje na to da su bogati isoleucinom, leucinom i lizinom. Trodimenzionalne strukture (3D) proteina CCPP su određene pomoću Phyre2 servera. Aminokislinske sekvence CCPP proteina su podvrgnute predviđanju sekundarne strukture korišćenjem ExPASy's SOPMA alata. Proteini su više strukture alfa heliksa. Genetske informacije koje su rezultat ove studije mogu doneti uvid u mutagenezu i farmakogenetiku.

Ključne reči: protein, pleuropneumonija koza, sekvenca

References

BÖLSKE G., JOHANSSON K.E., HEINONEN R., PANVUGA P.A., TWINAMASIKO E. (1995): Contagious caprine pleuropneumonia in Uganda and isolation of *Mycoplasma capricolum* subspecies *capripneumoniae* from goats and sheep.Veterinary Record, 137: 594.

EDELSTEN R.M., GOURLAY R.N., LAWSON G.K.H., MORROW A.N., RAMACHANDRAN S. (1990): Diseases caused by bacteria. In: Sewell M.M.H. and Brockles by D.W. (eds), Handbook on animal diseases in thetropics. Baillière Tindall, London, UK.

GASTEIGER E. (2005): The Proteomics Protocols Handbook Humana Press, 571-607.

GASTEIGER E., GATTIKER A., HOOGLAND C., IVANYI I., APPEL R. D., BAIROCH A. (2003): ExPASy—the proteomics server for in-depth protein knowledge and analysis. Nucleic Acids Research, 31, 3784–3788.

GILL S. C., VON-HIPPEL P. H. (1989): Calculation of protein extinction coefficients from amino acid sequence data. Analytical Biochemistry, 182, 319–326.

GURUPRASAD K., REDDY B.V., PANDIT M.W. (1990): Correlation between stability of a protein and its dipeptide composition: a novel approach for predicting in vivo stability of a protein from its primary sequence. Protein Engeenering, 2, 155–161.

JONES D.T. (2007): Improving the accuracy of transmembrane protein topology prediction using evolutionary information. Bioinformatics, 3 (5), 538–544.

KELLEY L. A. (2015): The Phyre2 web portal for protein modeling, prediction and analysis. Nature Protocols, 10, 845-858.

KELLEY L.A., STERNBERG, M.J.E. (2009): Protein structure prediction on the Web: a case study using the Phyre server. Nature Protocols, 4 (3), 363–371

MANSO-SILVÁN L., DUPUY V., CHU Y., THIAUCOURT F. (2011): Multi-locus sequence analysis of Mycoplasmacapricolum subsp.capripneumoniae for the molecular epidemiology of contagious caprine pleuropneumonia. Veterinary Research, 42: 86. http://www.veterinaryresearch.org/content/42/1/86

OIE (Office International des Epizooties) (2001): Handistatus II. OIE, Paris, France.www.oie.int.

QIAN B. D. (2007): High-resolution structure prediction and the crystallographic phase problem. Nature, 450, 259–264.

RAVA P., HUSSAIN M.M. (2007): Acquisition of triacylglycerol transfer activity by microsomal triglyceride transfer protein during evolution. Biochemistry, 46, 12263–12274.

THOMAS P. (1873): Rapport médical sur le Bou Frida. In Publication du gouvernementgénéral civil de l'Algérie. Edited by: Jourdan A. Algiers

WASS M.N., KELLEY L.A., STERNBERG M.J. (2010): 3DLigandSite: predicting ligand-binding ssites using similar structures. Nucleic Acids Research, 38, 469-473.
YAROV-YAROVOY V., SCHONBRUN J., BAKER D. (2006): Multipass membrane protein structure prediction using Rosetta. Proteins, 62, 1010–1025.

Received 26 June 2017; accepted for publication 14 August 2017