ejpmr, 2021,8(2), 354-359

SJIF Impact Factor 6.222

EUROPEAN JOURNAL OF PHARMACEUTICAL AND MEDICAL RESEARCH

www.ejpmr.com

Research Article

ISSN 2394-3211

EJPMR

IN-SILICO STUDY: PROTEIN STRUCTURE PREDICTION, COMPARISON AND COMPARATIVE STRUCTURAL ANALYSIS BY ALIGNMENT OF SEVERE ACUTE RESPIRATORY SYNDROME CORONAVIRUS 2 NUCLEOCAPSID PROTEIN (SARS-COV-19 N PROTEIN) AND DIFFERENT ORGANISM'S HOST WITH ACE2 PROTEIN

Krushna K. Dishware [1]*, Dr. Faiyaz K. Shaikh[1] and Dr. Sanjay N. Harke [2]

[1]Dept. of Biotechnology and Bioinformatics, MGM IBT, MGMU, Aurangabad (431001).
[2]Dept. of Bioinformatics, MGM's Institute of Biosciences and Technology, MGM Campus, Aurangabad (431001).

***Corresponding Author: Krushna K. Dishware**

Dept. of Biotechnology and Bioinformatics, MGM IBT, MGMU, Aurangabad (431001).

Article Received on 02/12/2020 **Article Revised on 22/12/2020** **Article Accepted on 12/01/2021**

ABSTRACT

Human beings are presently experiencing a serious global public health issue that is Coronavirus disease 19 (COVID-19). The World Health Organization (WHO) on March 11, 2020, has declared the novel Coronavirus disease 19 (COVID-19) outbreak a global pandemic. Coronavirus disease 19 (COVID-19) has a single - stranded RNA genome. There are seven strains of human Coronavirus (CoVs) ,out of those three strains are highly pathogenic (SARS-CoV, MERS-CoV, 2019-nCoV), which causes an endemic of severe Coronavirus (CoV) disease. The most important four structural proteins of COVID-19 are S (spike), E (envelope), M (membrane), and N (nucleocapsid) proteins. The nucleocapsid protein is one of the core components of the SARS-CoV-2. Nucleocapsid protein alone is able to form the capsid. In this present work, we have predicted tertiary structure of human, Bats (Pteropus vampyrus), Mus musculus with ACE2 protein with the help of Swiss model (automated protein structure homology-modelling server). Protein structure alignment and comparison of different organism's host with ACE2 protein and SARS-CoV-2 Nucleocapsid protein using PyMol (3-Dimensional structure visualization Tool). We also performed Superimposition of proteins, comparative structural analysis. Calculate their RMSD (Root Mean Square Deviation) and RMS score.

KEYWORDS: COVID-19, SARS-CoV-2 Nucleocapsid protein, Structure prediction, Structure alignment and comparison, comparative structural analysis.

INTRODUCTION

COVID-19

The disease was initially found in Wuhan, china and expanded to nearly 200 countries. The USA,UK and Spain are the most affected countries due to the coronavirus disease (covid-19).[16] The coronavirus disease (covid-19) is declared as a global epidemic by the world health organisation (WHO).[3] Severe acute respiratory syndrome coronavirus is the new virus declared by the International committee on taxonomy of viruses on 11 February, 2020.[3,8] This is the largest RNA virus family classified into four genera i.e. alpha, beta, gamma and delta. The SARS-CoV-2 belongs to beta-genus.[9] A novel coronavirus SARS-CoV-2 genome is about 82% identical to the SARS-CoV.[9] The nucleocapsid protein and envelope protein of SARS Coronavirus (SARS-CoV) and SARS Coronavirus 2 (SARS-CoV-2) share 89.6% and 96% sequence identities.[9] The symptoms of covid-19 disease mainly include cough, difficulty in breathing, fever, sore throat.[6] The targets of coronavirus are structural protein, nonstructural protein, RNA-dependent RNA polymerase and the angiotensin-converting enzyme II (ACE2). The four structural proteins are Spike(S), Nucleocapsid(N), Envelope(E), and Membrane(M).[1]

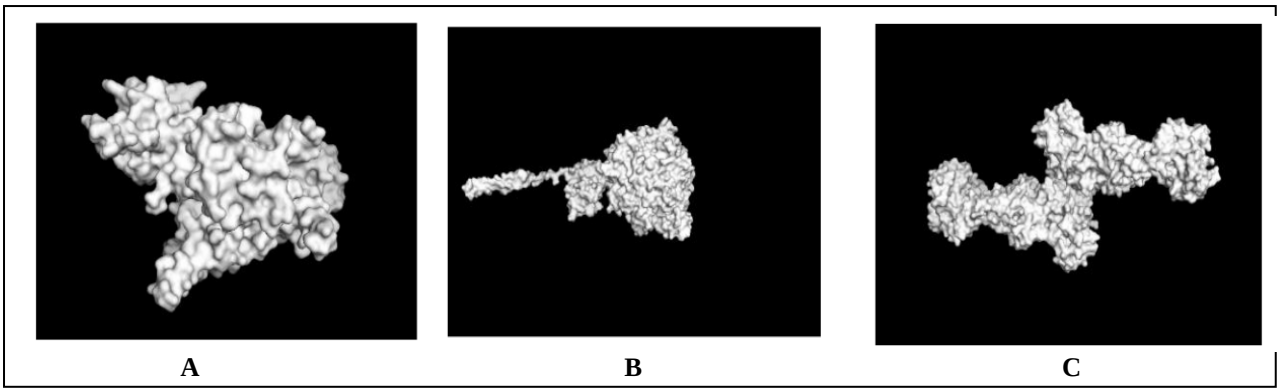


Figure 1: **Showing 3-Dimensional Structure of (A) Homosapiens ACE2 protein, (B) Bats (Pteropus vampyrus) protein, (C) Mus musculus ACE2 protein visualized in PyMol is reported in surface representation (white color).**

NUCLEOCAPSID PROTEIN

It is a viral protein coat that surrounds the genome. This genome is either DNA or RNA. Viral nucleocapsids contain major constituents of nucleocapsid protein.[2] In Coronavirus disease 2019 (COVID-19), the nucleocapsid protein alone is able to form the capsid. The main function of nucleocapsid protein is the only coat that protects the genome from the external world. Nucleocapsid protein is also one of the most expressed viral proteins. Understanding the function and properties of the nucleocapsid protein is most important to a biologist to understand the biological processes of the virus and develop effective tools or vaccines to control the infection. The N- protein is encoded by the ninth ORF of severe acute respiratory syndrome coronavirus (SARS-CoV). The nucleocapsid protein is a 46-kDa protein composed of 422 amino acids.[2] The two different features of nucleocapsid protein are being capable of recognising the genomic RNA and self-associating into an oligomer to form the capsid. The primary function of the nucleocapsid protein is only to package the genomic RNA in a protective covering.[4] Finally, the nucleocapsid protein is called the most expressed protein of the severe acute respiratory syndrome coronavirus (SARS-CoV). Hence, any known information generated from the analysis or study of nucleocapsid protein, either in-vivo or in-silico will definitely help to maximize our understanding of the biological processes of severe acute respiratory syndrome coronavirus.[2]

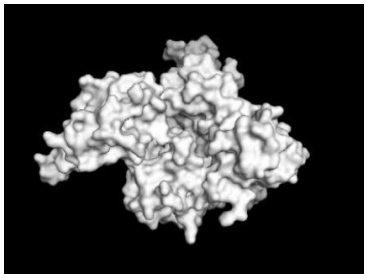


Figure 2: 3-Dimensional Structure of SARS-CoV-2 Nucleocapsid protein visualized in PyMol reported in surface representation (white color).

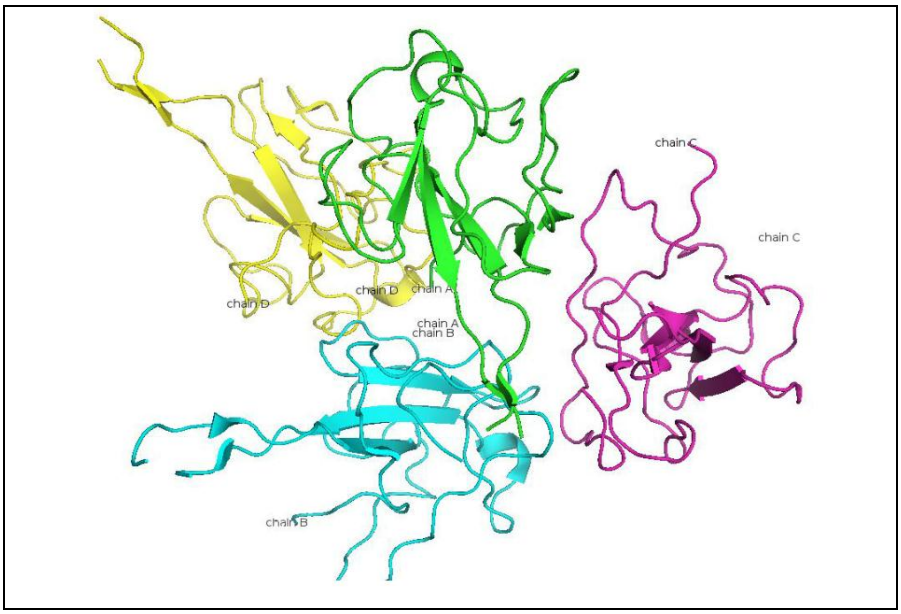
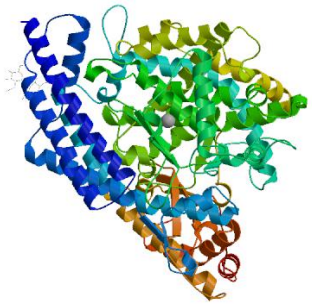
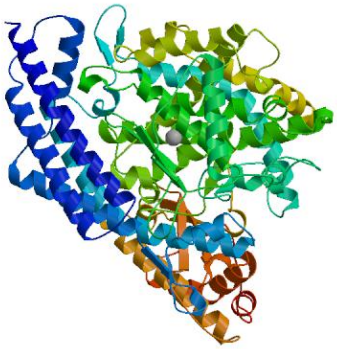
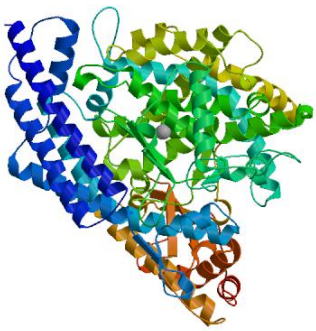


Figure 3: SARS-CoV-2 Nucleocapsid protein structure visualized in PyMol is reported in cartoon representation [chain A(green color), chain B(cyan color), chain C (magentas color), chain D(yellow)].

PROTEIN STRUCTURE PREDICTION

It is a fundamental area of computational biology.[7,12] There are three computational methods to predict 3 Dimensional structures of protein are Homology modeling, threading and ab initio.[13] Homology modeling and threading are knowledge-based methods,they predict protein 3D structures based on knowledge of existing protein structural information available in databases.[13] Homology modeling builds an atomic model using amino acid sequence of protein based on an experimentally determined structure i.e. closely related at the sequence level. This modeling method predicts protein structure based on sequence homology with known structure present in databases for Ex. Protein Data Bank (PDB).[12]

Table 1: Shows Organism name with its NCBI Accession ID, Predicted 3D structure and its Q-mean score.

SR.NO	Organism Name	NCBI Accession ID	Predicted 3D Protein Structure	Q-mean score
1	**Human**	BAB40370.1		-0.92
2	**Bats (Pteropus vampyrus)**	XP_011361275.1		-1.03
3	**Mus musculus**	NP_001123985.1		-0.88

PROTEIN STRUCTURE ALIGNMENT

The comparison and alignment of protein structure has come to be a fundamental and widely used task in computational biology.[12] There are three main steps needed for comparing two protein structure first step is the detection of their common similarities, second step is the alignment of the structure based on such similarities and the last step is a statistical measure of the similarity considering the first two steps.[13] Structure comparison refers to the analysis of the similarities and differences between two or more structures. Structure alignment refers to establishing which amino acid residues are equivalent between them. A method may align the central of a protein structure at very high accuracy (i.e very low RMSD score) and very low coverage that is except loop regions while a another method approach may prefer to increases the coverage (i.e. include the loop regions in the alignment) to the loss of accuracy (i.e. increasing the RMSD score).[5,7,9]

Table 2: Shows Structure Alignment in between SARS-CoV-2 Nucleocapsid protein and Different Organism host with ACE2 protein and its RMS score.

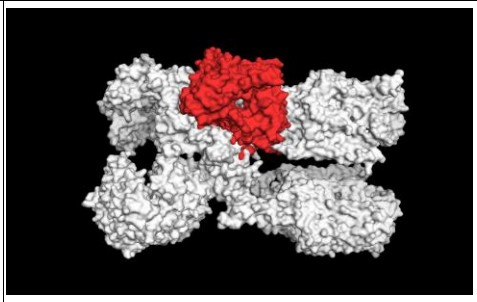
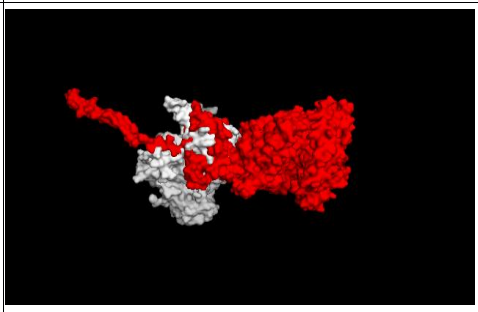
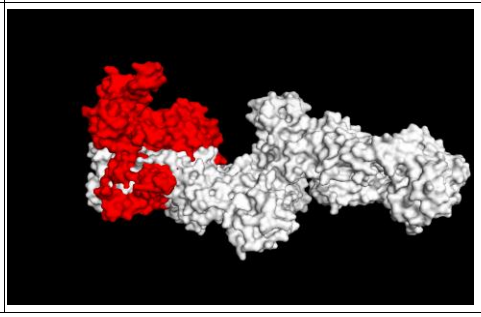
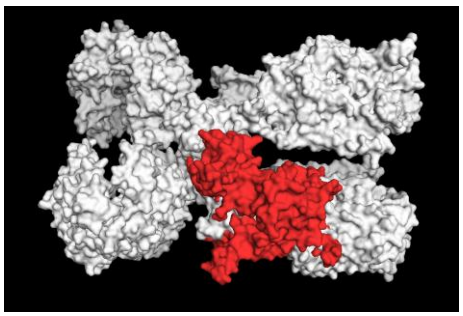
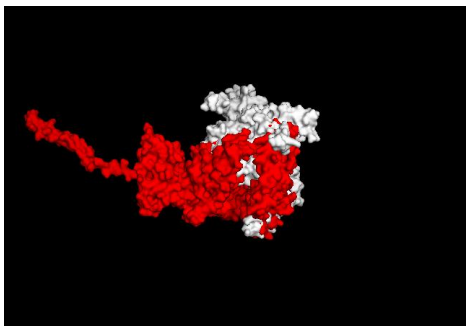
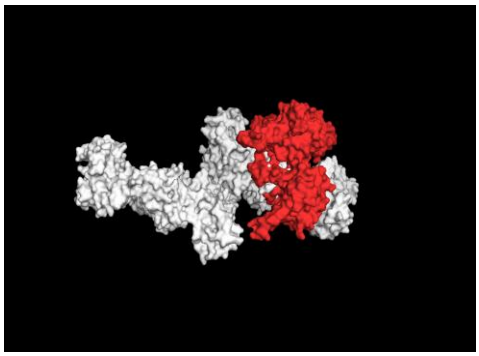
SR. NO	OGRANISM NAME	STRUCTURE	RMS SCORE
1	ACE2 Human		RMS = 2.520 (58 to 58 atoms)
2	ACE2 Bats (Pteropus vampyrus)		RMS = 2.444 (60 to 60 atoms)
3	ACE2 Mus musculus		RMS = 22.723 (1176 to 1176 atoms)

Table 3: Shows Structure Alignment in between SARS-CoV-2 Nucleocapsid protein and Different Organism host with ACE2 protein its RMSD score.

SR.NO	OGRANISM NAME	STRUCTURE	RMSD SCORE
1	ACE2 Human		8.491559 over 64 residues
2	ACE2 Bats (Pteropus vampyrus)		6.326575 over 64 residues

3	ACE2 Mus musculus		6.603576 over 72 residues

MATERIALS AND METHODS

Data collection

The Amino acid sequences of Homosapien, Bats (Pteropus vampyrus) and Mus Musulus (NCBI Accession ID:- BAB40370.1, AHI85757.1 and NP_001123985.1) with ACE2 protein were downloaded in fasta file format from the NCBI (National center of Biotechnology Information) resources https://www.ncbi.nlm.nih.gov/. The crystal/tertiary structure of the SARS-CoV-2 Nucleocapsid protein was taken from the RCSB (Protein Data Bank, PDB ID: 6m3m) Shown in figure 2. https://www.rcsb.org.

Tertiary Structure prediction

We have predicted tertiary structure of Homosapien, Bats (Pteropus vampyrus) and Mus Musculus with ACE2 protein by using homology modeling method. The structure of Homosapien, Bats (Pteropus vampyrus) and Mus Musculus with ACE2 protein is predicted by using 'SWISS model' (https://swissmodel.expasy.org/) which is an online tool provided by Expasy server (Expert Protein Analysis System). Three Dimensional structures of Homosapien, Bats (Pteropus vampyrus) and Mus Musculus with ACE2 protein are predicted (Shown in figure 1). This Tool also provides the Q-Mean (Qualitative Model Energy Analysis) and Z-score of the predicted structure. The Q-Mean score is near to zero indicates the good Quality of predicted structure (Table 1).

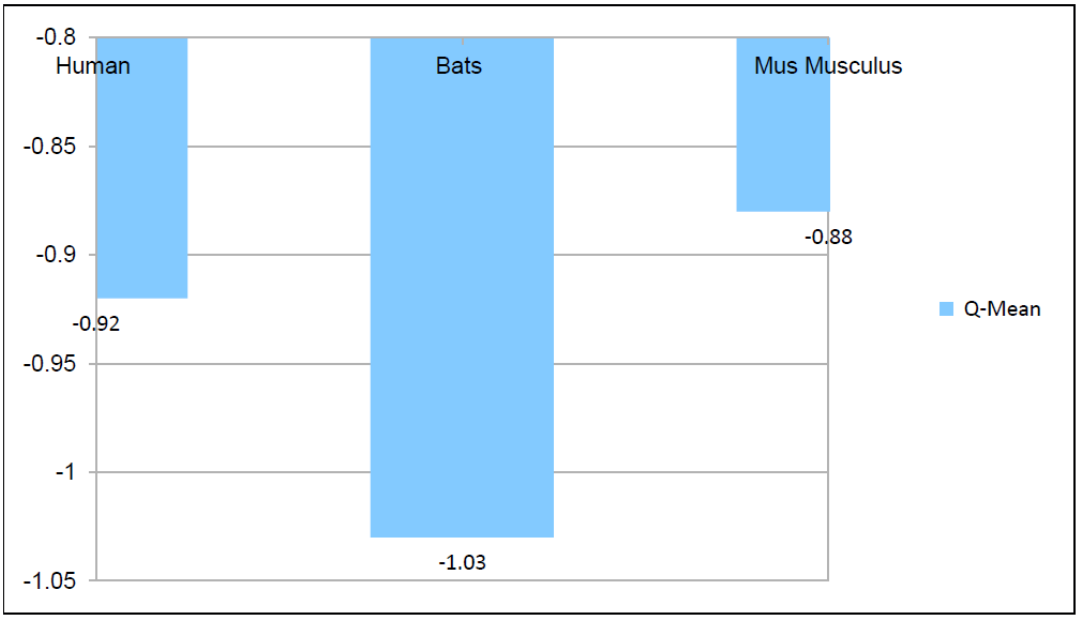
Structure Alignment

Structure Alignment of Homosapien, Bats (Pteropus vampyrus) and Mus Musculus with ACE2 protein and SARS CoV-2 Nucleocapsid protein by using "cealign" command in PyMol Visualization Tool.[5] The PyMol Tool cealign command gives an RMSD score (Show in Table 3).[10] The three Dimentional structure of Homosapien, Bats (Pteropus vampyrus) and Mus Musculus with ACE2 protein and SARS COV-2 Nucleocapsid protein were superimposed using "super" command available in PyMol its gives the RMS score (Show Table 2).[5,11]

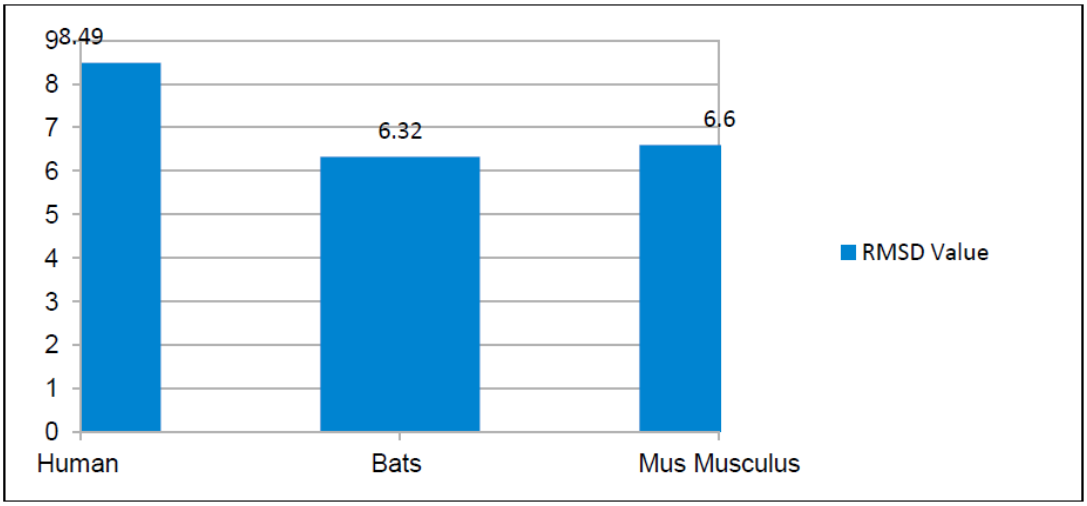
RESULT

In this study, we predicted tertiary structure of Homosapien, Bats (Pteropus vampyrus) and Mus Musculus with ACE2 protein using Homology modeling method. For predicting tertiary structure from amino acid sequences we use the "SWISS model" online Tool provided by Expasy server. Q-mean score which is "-0.92,-1.39,-0.88" near to zero shows good quality of predicted structure. Structural similarity between two proteins usually shows Low RMSD score which is shown in table 3. RMSD score is near to zero shows better the Structure similarity of two protein. RMSD Values of Bats and Mus Musculus with ACE2 protein show slight differences. Structural comparison of Homosapien, Bats (Pteropus vampyrus) and Mus Musculus with ACE2 protein and SARS CoV-2 Nucleocapsid protein as shown by graph 2.

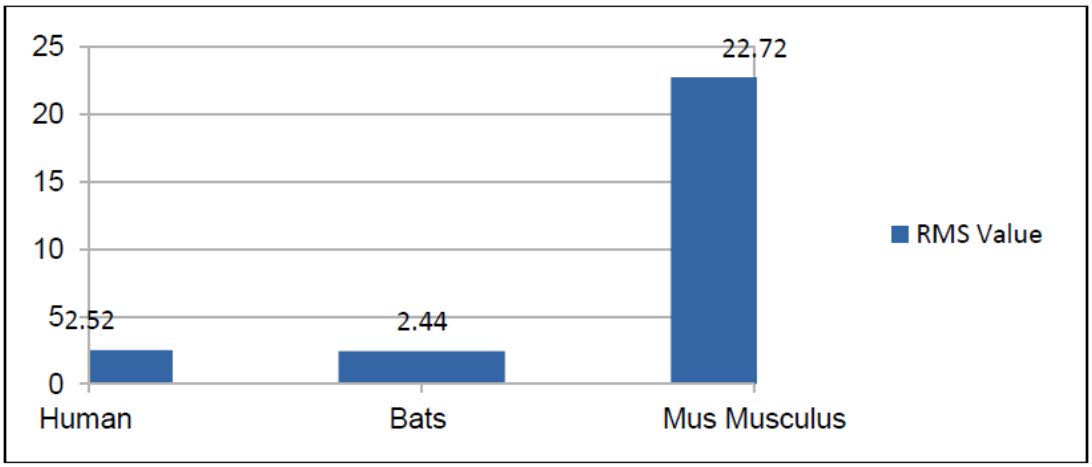
Super command is more effective than the "cealign" command, because we perform superimposes of protein structure with a minimum sequence similarity. The interaction between SARS CoV-2 Nucleocapsid protein and Homosapien, Bats (Pteropus vampyrus) and Mus Musculus with ACE2 protein are highly homologous, Bats (Pteropus vampyrus) with ACE2 protein having RMS score "2.44" shows high accuracy, Mus musculus with ACE2 protein is having loss of accuracy that shows in table 2. Human and Bats with ACE2 protein show slight differences in RMS Value. The performing superimposition of SARS CoV-2 Nucleocapsid protein and Homosapien, Bats (Pteropus vampyrus) and Mus musculus with ACE2 protein shows comparative structural analysis. The binding ability of proteins encoded by different organism hosts with ACE2 protein was analyzed.



Graph 1: Shows the Q-mean score.



Graph 2: Shows RMSD score between SARS-CoV-2 Nucleocapsid protein and Different Organism's Host with ACE2 protein.



Graph 3: Shows RMS score between SARS-CoV-2 Nucleocapsid protein and Different Organism's Host with ACE2 protein.

CONCLUSION

The structure alignment and superimposition of two proteins was performed which show high homology, their alignment of ACE2 Bats and SARS-CoV-2 Nucleocapsid protein is quite high. The study of ACE2 protein encoded with Different Organism hosts would be helpful to obtain more genetic and functional information about SARS-CoV-2. The potential susceptibility analyses of Different organism hosts would be of great importance for controlling the virus spread, treating viral disease and protecting life.

REFERENCES

1. Manisha Prajapat, Phulen Sarma, Nishant Shekhar, Pramod Avti, Sheweta Sinha, Hardeep Kaur, et al. Drug targets for coronavirus: A systematic review, 2020 Jan-Feb; 52(1): 56–65. Published online 2020 Mar 11. doi: 10.4103/ijp.IJP_115_20 https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7074424/.
2. Milan Surjit and Suniol K. Lal. The Nucleocapsid Protein of the SARS Coronavirus: Structure, Function and Therapeutic Potential, 2009 Jul 22; 129–151. Published online 2009 Jul 22. doi: 10.1007/978-3-642-03683-5_9 https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7176212/.
3. Ambrish Kumar Srivastava, Abhishek Kumar, et al. On the Inhibition of COVID-19 Protease by Indian Plants: An In Silico Investigation. [Submitted on 5 Apr 2020] https://arxiv.org/abs/2004.03411.
4. McBride R, van Zyl M and Fielding BC. The coronavirus nucleocapsid is a multifunctional protein. Viruses, 2014 Aug 7; 6(8): 2991-3018. doi: 10.3390/v6082991. PMID: 25105276; PMCID: PMC4147684.
5. Ivan Mercurio, Vincenzo Tragni, et al. Protein structure analysis of the interactions between SARS CoV-2 spike protein and the human ACE2 receptor: from conformational changes to novel neutralizing antibodies. Doi: https://doi.org/10.1101/2020.04.17.046185.
6. Rajesh Ghosh, Ayon chakraborty, et al. Evaluation of green tea polyphenols as novel coronavirus (SARS CoV-2) main protease (Mpro) inhibitors – an *in silico* docking and molecular dynamics simulation study. 2020 Jun 22. DOI: 10.1080/07391102.2020.1779818. PMID: 32568613.
7. Subodh et al. structure, pathway prediction for functional elements of e2 glycoprotein precursor in sars-cov-2 to analyze it's role in covid-19 disease, WJPMR, 2020; 6(8): 246-250.
8. Ying Cao, Yeping Sun, et al. Analysis of ACE2 Gene-Encoded Proteins Across Mammalian Species, 2020 Jul 3. DOI:10.3389/fvets.2020.00457 PMID: 32719819.
9. Grijesh Jaiswal, et al. In-silico design of a potential inhibitor of SARS-Cov-2 S protein. https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0240004.
10. Fred E.Cohen and Michael J E. Sternberg. On the prediction of protein structure: The significance of the root-mean-square deviation, 5 April 1980; 138(2): 321-333.
11. DeLano WL. The PyMOL Molecular Graphics System, Version 1.1. Schr{ö}dinger LLC. 2002. https://doi.org/10.1038/hr.2014.17.
12. Jenny Gu and Philip E. Bourne. Structural Bioinformatics, 2nd Edition.
13. Jin Xiong, Essential Bioinformatics Book.