International Journal of Advanced Research in Biological Sciences
ISSN: 2348-8069
www.ijarbs.com
(A Peer Reviewed, Referred, Indexed and Open Access Journal)
DOI: 10.22192/ijarbs Coden: IJARQG (USA) Volume 9, Issue 5 -2022

Research Article

Special Issue on Potential Applications of Bioinformatics in Biological Sciences - PABBS 2022

DOI: http://dx.doi.org/10.22192/ijarbs.2022.09.05.01.016

Comparative Structural Analysis of CFTRABCC7in Human and Rabbit

M.Vinoth[1*],Dr. R. Priya[2], Mahendran Radha[3]
Department of Bioinformatics, Vels institute of Science and Technology & Advanced Studies, Chennai.
*Corresponding e-mail: *vinoth.sls@velsuniv.ac.in*

Abstract

Mutations in the gene encoding the cystic fibrosis transmembrane conductance regulator cause cystic fibrosis, an autosomal recessive illness (CFTR).Cystic fibrosis (CF) is an incurable, chronic disease, which causes severe damages to respiratory and digestive tracts. It is the most common genetically inherited disease among Caucasians. For decades, preventing cystic fibrosis (CF) patients from developing lung illness has been an aim.With the discovery of the CF transmembrane conductance regulator (CFTR) gene and other breakthroughs in airway epithelial biology, a much better picture of illness has emerged. Pathophysiology has been discovered. The aim is to analyse CFTR ABCC7 a homologousgene responsible for causing genetic disease Cystic fibrosis, in order to lead research on this disease.

Keywords: Comparative genomics, Cystic fibrosis, CF, Structural Analysis, Motif Prediction, Primary and Secondary Structure, Tertiary Structure, Sequence Alignment.

Introduction

Cystic fibrosis -- An inherited life-threatening disorder that damages the lungs and digestive system. Cystic fibrosis affects the cells that produce mucus, sweat and digestive juices. It causes these fluids to become thick and sticky. They then plug uptubes, ducts and passageways. Symptoms vary and can include cough, repeated lung infections ,inability to gain weight and fatty stools. Treatments may ease symptoms and reduce complications. Newborn screening helps with early diagnosis.

Cystic fibrosis is an autosomal recessive genetic disorder that affects ion transport in exocrine glands. Inadequate ion transport causes dehydration and the production of thick secretions

in organs such as the lungs, sinuses, pancreas, intestines, hepatobiliary tree, and vas deferens. Although cystic fibrosis usually is diagnosed through a sweat chloride test, medical imaging is used to monitor pathologic changes caused by thedisease. This disease is caused by defects in CF genes, the so-called mutations incystic fibrosis transmembrane conductance regulator (CFTR) gene population. At present over 100,000 people suffer from this disease worldwide.
In this article, the comparative analysis of structure of CFTRABCC7 in Human and Rabbit using online tools and databases.

Materials and Methods

The following resources have been used to conduct analytical studies.

National Center for Biotechnology Information:

The National Center for Biotechnology Information advances science and healthby providing access to biomedical and genomic information.

Uniprotkb:

The UniProt Knowledgebase (UniProtKB) is the central hub for the collection of functional information on proteins, with accurate, consistent and rich annotation. In addition to capturing the core data mandatory for each UniProtKB entry (mainly, the amino acid sequence, protein name or description, taxonomic data and citationin formation),as much annotation information as possible is added.

Pairwise Alignment:

Pairwise Sequence Alignment is used to identify regions of similarity that may indicate functional, structural and/or evolutionary relationships between two biological sequences (protein or nucleic acid).

Protparam:

ProtParam (References / Documentation) is a tool which allows the computation of various physical and chemical parameters for a given protein stored in Swiss-Prot or TrEMBL or for a user entered protein sequence. The computed parameters include the molecular weight, theoretical pi, amino acid composition, atomic composition, extinction coefficient, estimated half-life, instability index, aliphatic index and grand average of hydropathicity(GRAVY)

GORIV:

The GOR method (Garnier-Osguthorpe-Robson) is an information theory-based method for the prediction of secondary structures in proteins. Like Chou-Fasman, theGOR method is based on probability parameters derived from empirical studies of known protein tertiarys tructures solvedby X-ray crystallography.

SWISS MODEL:

SWISS-MODEL is a fully automated protein structure homology-modellingserver,accessibleviatheExPASywebserver,orfromtheprogramDeepView(SwissPdb-Viewer). The purpose of this server is to make protein modeling accessible to all life science researchers worldwide.

BLAST:

Basic Local Alignment Search Tool finds regions of similarity between biological sequences. The program compares nucleotide or protein sequences to sequence databases and calculates the statistical significance.

GenomeNet:

GenomeNet is a Japanese network of database and computational services for genome research and related research areas in biomedical sciences, operated by the Kyoto University Bioinformatics Center. It provides information about kegg pathway, motif search, analysis of sequence, genome and chemical, and also has various bioinformatics tools.

Methods:

NCBI-GENE:

Using the NCBI database, homologous genes CFTR ABCC7 responsible for CysticFibrosis ; sequences are retrieved from the Uniprot database and then used for furtherpredictions.

PAIRWISEALIGNMENT(matcherEMBOSS):

Pairwise sequence alignment of Homo sapien and Oryctolagus cuniculus -- CFTRABCC7shows 95.8% sequencesimilarity.

##

Program: matcher

#Rundate:Tue27Apr202110:57:42#
Commandline: matcher
-auto

-stdout

#-asequence emboss_matcher-I20210427-110025-0629-66005465-p1m.asequence#-bsequence emboss_matcher-I20210427-110025-0629-66005465-p1m.bsequence#-datafile EBLOSUM62
-gapopen 14

-gapextend 4

-alternatives 1# -aformat3 pair#-sprotein1
-sprotein2

Align_format: pair#Report_file:stdout
##

#=================================
=====#
#Aligned_sequences: 2

1: CFTR_HUMAN

2: CFTR_RABIT

Matrix: EBLOSUM62

Gap_penalty: 14

Extend_penalty: 4#
Length: 1481

Identity: 1366/1481 (92.2%)# Similarity: 1419/1481 (95.8%)# Gaps: 1/1481 (0.1%)
Score: 6980#
#

#=================================
=====

PRIMARYANDSECONDARYSTRUCTURETOOLS:

Using **Protparam** tool, the primary structure information are predicted, thenumber of amino acids in a human and rabbit protein sequence is **1480** and **1481** respectively. The molecular weight is **168141.57** and **168042.42** theoretical valueis **8.91**and **9.01** for both.

Leucine has a high amino acid composition with a neutral charge of **12.4%** and 12.6%**.**

The consequence of Gor IV is that the human alpha helix is 44.19% which is higher than the 37.64% random coil and is also higher in the mouse alpha helix than random coil having percentage of 43.55% and 36.19%, So when protein is introduced to the drug it may be curable.

(HUMAN)

Number of amino acids:
148Molecularweight:168141.57
Theoretical pI: 8.91Aminoacidcomposition:

Ala (A)	83	5.6%
Arg (R)	78	5.3%
Asn (N)	54	3.6%
Asp (D)	58	3.9%
Cys (C)	18	1.2%
Gln (Q)	67	4.5%
Glu (E)	93	6.3%
Gly (G)	84	5.7%
His (H)	25	1.7%
Ile(I)	119	8.0%
Leu (L)	183	12.4%
Lys(K)	92	6.2%
Met (M)	37	2.5%
Phe (F)	85	5.7%
Pro (P)	45	3.0%
Ser (S)	123	8.3%
Thr (T)	83	5.6%
Trp(W)	23	1.6%
Tyr(Y)	40	2.7%
Val(V)	90	6.1%
Pyl (O)	0	0.0%
Sec (U)	0	0.0%

(B) 0 0.0%

(Z) 0 0.0%

(X) 0 0.0%

Totalnumberofnegativelychargedresidues(Asp+Glu):151Totalnumberofpositivelychargedresidues(Arg+Lys):170

Atomic composition:

Carbon	C	7640	
Hydrogen	H		12106
Nitrogen	N	2000	
Oxygen	O	2150	
Sulfur	S	55	

Formula: C7640H12106N2000O2150S55

Totalnumberofatoms:23951

(RABBIT)

Number of amino acids: 1481

Molecular weight: 168042.42

Theoretical pI: 9.01

Amino acid composition:Ala (A)84 5.7%
Arg (R)80 5.4%

Asn (N) 53 3.6%

Asp (D) 59 4.0%

Cys (C) 17 1.1%

Gln (Q) 63 4.3%
Glu (E) 90 6.1%

Gly (G) 84 5.7%

His (H) 24 1.6%

Ile(I)118 8.0%
Leu (L) 187 12.6%

Lys(K)91 6.1%

Met (M)40 2.7%

Phe (F)82 5.5%

Pro (P)43 2.9%

Ser (S) 134 9.0%

Thr (T)82 5.5%

Trp(W)23 1.6%

Tyr(Y)42 2.8%

Val(V)85 5.7%

Pyl (O)0 0.0%

Sec (U)0 0.0%

(B) 0 0.0%

(Z) 0 0.0%

(X) 0 0.0%

Totalnumberofnegativelychargedresidues(Asp+Glu):149Totalnumberofpositivelychargedresidues(Arg+Lys):171

Atomic composition:

Carbon	C	7623	
Hydrogen	H		12097
Nitrogen	N	1999	
Oxygen	O	2154	
Sulfur	S	57	

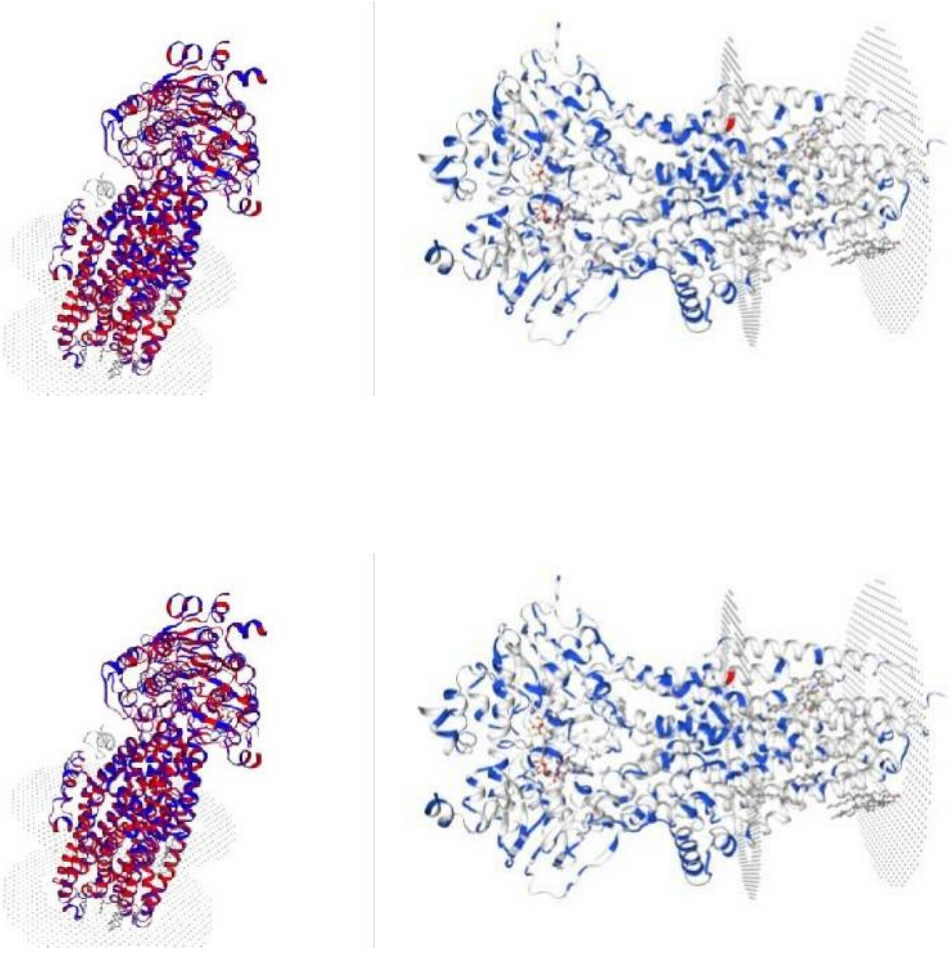
Formula: C7623H12097N1999O2154S57

Totalnumberofatoms:23930

Tertiary structure

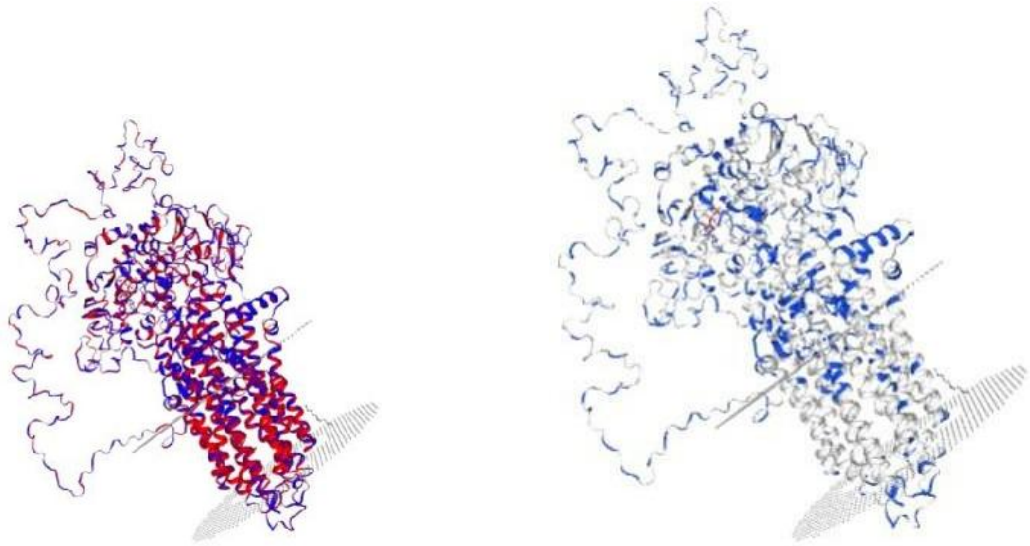
ThetertiaryproteinstructureisdevelopedusingSWISSMODELandthestructureis differentiated by the color of the polar and hydrophobic region of the given proteinsequence is mentioned in.

(HUMAN)



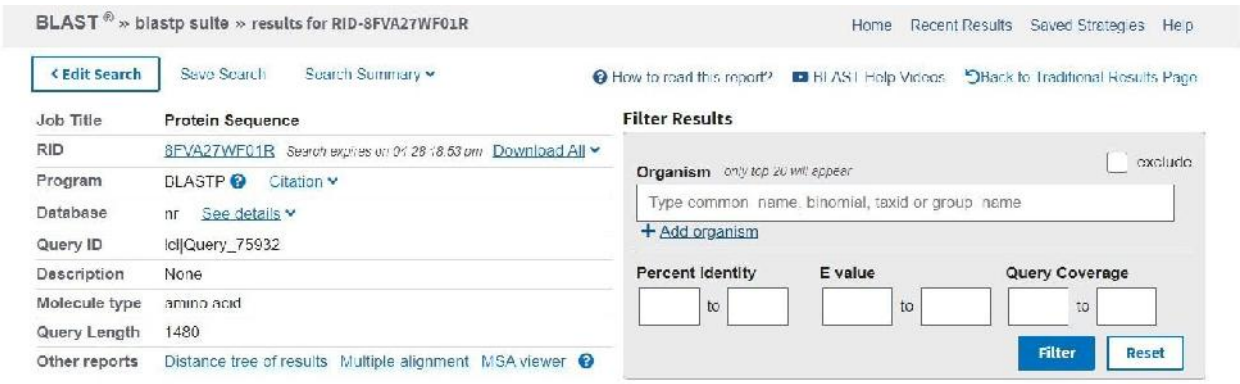
(RABBIT)

SIMILAR SEQUENCEALIGNMENTS

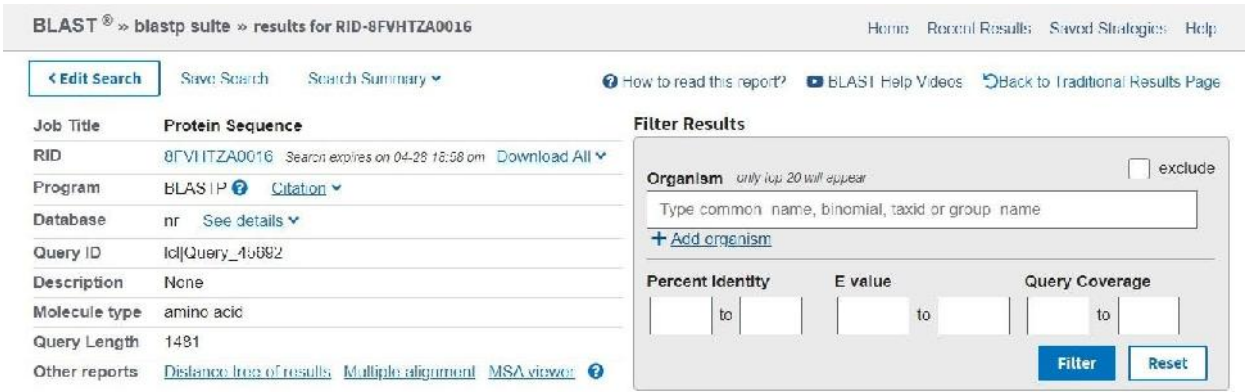


BLASTPprogramssearchproteindatabasesusingaproteinquery.ItpredictssignificantAlignments for protein sequences of Humanand rabbit.

(HUMAN)



(RABBIT)



MOTIF PREDICTION:

By motif analysis **CFTR** ABCC7 pattern belongs to PF14396, Cystic fibrosis TMconductance regulator (CFTR), regulator domain family, and the human **CFTR** ABCC7family highly lie between 639-849 and rabbit PF14396, (CFTR), regulator domain familylie between 639-849 region. And other such motifs of Human and chimpanzee lies inABC transporter transmembrane region, RecF/RecN/SMC N terminal domain, AAAATPasedomain, Double-GTPase 2.

(HUMAN)

Pfam (17 motifs)

Pfam	Position(Independent E-value)		Description
CFTR_R	639..849(1.2e-95)	Detail	PF14396, Cystic fibrosis TM conductance regulator (CFTR), regulator domain
ABC_membrane	82..350(5.2e-42) 862..1147(7.7e-50)	Detail	PF00664, ABC transporter transmembrane region
ABC_tran	441..575(7.4e-23) 1227..1374(4.3e-33)	Detail	PF00005, ABC transporter
AAA_21	1241..1322(0.041) 1346..1386(0.16)	Detail	PF13304, AAA domain, putative AbiEii toxin, Type IV TA system
SMC_N	1345..1418(0.025)	Detail	PF02463, RecF/RecN/SMC N terminal domain
AAA_29	1233..1265(0.0012)	Detail	PF13555, P-loop containing region of AAA domain
RsgA_GTPase	441..484(0.048)	Detail	PF03193, RsgA GTPase
DUF87	1238..1262(0.074)	Detail	PF01935, Helicase HerA, central domain
T2SSE	1243..1264(0.11)	Detail	PF00437, Type II/IV secretion system protein
AAA_16	1239..1265(0.3)	Detail	PF13191, AAA ATPase domain
MMR_HSR1	1239..1259(0.26)	Detail	PF01926, 50S ribosome-binding GTPase
DO-GTPase2	1239..1284(0.32)	Detail	PF19993, Double-GTPase 2
Zeta_toxin	449..479(0.37)	Detail	PF06414, Zeta toxin
MeaB	1224..1266(0.013)	Detail	PF03308, Methylmalonyl Co-A mutase-associated GTPase MeaB
AAA_30	1239..1323(0.28)	Detail	PF13604, AAA domain
NTPase_1	1240..1289(0.31)	Detail	PF03266, NTPase
Intein_splicing	369..516(0.25)	Detail	PF14890, Intein splicing domain

(RABBIT)

Pfam (16 motifs)

Pfam	Position(Independent E-value)		Description
CFTR_R	639..849(3.2e-97)	Detail	PF14396, Cystic fibrosis TM conductance regulator (CFTR), regulator domain
ABC_membrane	82..349(1.7e-41) 863..1147(6e-50)	Detail	PF00664, ABC transporter transmembrane region
ABC_tran	441..575(7.8e-24) 1228..1375(3e-33)	Detail	PF00005, ABC transporter
AAA_21	1242..1323(0.042) 1346..1387(0.091)	Detail	PF13304, AAA domain, putative AbiEii toxin, Type IV TA system
SMC_N	1346..1418(0.036)	Detail	PF02463, RecF/RecN/SMC N terminal domain
AAA_29	1234..1266(0.0012)	Detail	PF13555, P-loop containing region of AAA domain
RsgA_GTPase	441..484(0.067)	Detail	PF03193, RsgA GTPase
T2SSE	1200..1265(0.055)	Detail	PF00437, Type II/IV secretion system protein
DUF87	1239..1263(0.074)	Detail	PF01935, Helicase HerA, central domain
AAA_16	1240..1266(0.3)	Detail	PF13191, AAA ATPase domain
DO-GTPase2	1240..1285(0.32)	Detail	PF19993, Double-GTPase 2
MeaB	1224..1267(0.011)	Detail	PF03308, Methylmalonyl Co-A mutase-associated GTPase MeaB
MMR_HSR1	1240..1260(0.26)	Detail	PF01926, 50S ribosome-binding GTPase
Zeta_toxin	450..479(0.21)	Detail	PF06414, Zeta toxin
AAA_30	1240..1324(0.29)	Detail	PF13604, AAA domain
NTPase_1	1241..1290(0.31)	Detail	PF03266, NTPase

Results

According to the predictions, the structural evaluation of homologous human andrabbit genes is closely similar as by structure, sequence and motif prediction. Thestructural information predicted from the tools gives rise to the assumption that theprotein can be truly activated by the drug because the high percentage of amino acid composition is Leucine with neutral, hydrophobic and nonpolar properties and also the alpha helix percentage is higher than the random coil. By resulting this it allowsscientists to design mice to test a hypothesis or to generate rabbit models of human disease by using techniques such as homologous recombination, CRISPR/Cas9 and random transgenesis, etc.

Conclusion

Although I conclude that the above findings mean that the structural evaluation of homologous genes of different species is largely identical. Someone may further analyse the functional relationship between them in order to research the structural and functional analysis of CFTRABCC7 gene from homo sapien and rabbit.

References

1. Heidi A. Ernst, et al, (2018), A comparative structural analysis of the surface properties of asco-laccases.AnneScott,(2013),cysticfibrosis8 4(5):493-513
2. BlankaKlimova,etal,(2017),CysticFibrosis Revisited-aReviewStudy,13(2):102-109.
3. Welsh, M.; Ramsey, B.W.; Accurso, F.J.; Cutting, G.R. Cystic fibrosis. In The Metabolic and Molecular Basis of Inherited Disease, 8th ed.; Scriver, C.R., Beaudet, A.L., Sly, W.S., Valle, D., Childs, B., Vogelstein, B., Eds.; McGraw Hill: New York, NY, USA, 2001; pp. 5121–5188.
4. Dodge, J.A.; Lewis, P.A. Cystic fibrosis is no longer an important cause of childhood death in the UK. Arch. Dis. Child. 2005, 90, 547.
5. Corriveau, S.; Sykes, J.; Stephenson, A.L. Cystic fibrosis survival: The changing epidemiology. Curr. Opin. Pulm. Med. 2018, 24, 574–578.
6. Dodge JA, Lewis PA, Stanton M, Wilsher J. Cystic fibrosis mortality and survival in the UK: 1947-2003. *Eur Respir J* 2007;29:522-6.
7. Rosenstein BJ, Cutting GR, for the Cystic Fibrosis Foundation Consensus Panel. The diagnosis of cystic fibrosis: a consensus statement. *J Pediatr* 1998;132:589-95.
8. Middleton PG, Geddes DM, Alton EWFW. Protocols for in vivo measurement of the ion transport defects in cystic fibrosis nasal epithelium. *Eur Respir J* 1994;7:2050-6.
9. Yung MW, Gould J, Upton GJ. Nasal polyposis in children with cystic fibrosis: a long-term follow-up study. *Ann Otol Rhinol Laryngol* 2002;111:1081-6.
10. Deterding R, Retsch-Bogart G, Milgram L, Gibson R, Daines C, Zeitlin PL, et al. Safety and tolerability of denufosol tetrasodium inhalation solution, a novel P2Y2 receptor agonist: results of a phase 1/phase 2 multicenter study in mild to moderate cystic fibrosis. *Pediatr Pulmonol* 2005;39:339-48.
11. Daviskas E, Anderson SD. Hyperosmolar agents and clearance of mucus in the diseased airway. *J Aerosol Med* 2006;19:100-9.
12. Davies JC, Alton EW. Airway gene therapy. *Adv Genet* 2005;54:291-314.
13. Yung MW, Gould J, Upton GJ. Nasal polyposis in children with cystic fibrosis: a long-term follow-up study. *Ann Otol Rhinol Laryngol* 2002;111:1081-6.

14. Collins FS. Cystic fibrosis: molecular biology and therapeutic implications. *Science*. 1992; **256**(5058): 7 74- 779.
15. Kerem B, Rommens JM, Buchanan JA, et al. Identification of the cystic fibrosis gene: genetic analysis. *Science*. 1989; **245**(4922): 1073- 1080.
16. Vinoth, M., and Natarajan, B., "Isolation, characterization, antibiological and anticancer activity of Tetrahydrobenzochromone glycosides from leaf of Melochia corchorifolia." Rasayan Journal of Chemistry, 14(2), pp. 768-777, 2021.
17. Vinoth, M., and Natarajan, B., "Isolation, structural elucidation of flavonolignan derivative from Melochia corchorifolia - Evaluation of their biological, anticancer activity, and molecular docking studies." Rasayan Journal of Chemistry, 14(3), pp. 166-1679, 2021.
18. Vinoth, M., and Natarajan, B., and ShanmugaSundaram, C., "Characterization and evaluation ethyl acetate extract of Melochia corchorifolia leaf- anticancer antibiological and molecular docking studies on breast cancer estrogen receptor." Indian Journal of Chemical Technology, 28(2), pp. 188-196, 2021.
19. Vinoth, M., and Natarajan, B., "Evaluation of phytochemical, anthelmintic activity and Antidiabetic activity-Melochia corchorifolia ethanol extract." International Journal of Research in Pharmaceutical Sciences, 12(1), pp. 1-6, 2021.

Access this Article in Online	
	Website: www.ijarbs.com
Quick Response Code	Subject: Bioinformatics
DOI:10.22192/ijarbs.2022.09.05.01.016	

How to cite this article:
M.Vinoth, Dr. R. Priya, Mahendran Radha. (2022). Comparative Structural Analysis of CFTRABCC7in Human and Rabbit. Int. J. Adv. Res. Biol. Sci. 9(5): Special Issue 1: 162-172.
DOI: http://dx.doi.org/10.22192/ijarbs.2022.09.05.01.016