In Silico Biomodelling and Docking Studies of Claudin 1: A Rational Approach of Drug Design for Enteropathogenic *E.coli* Infections

Arpitha B M[1], Narasimha Sharma[1], Sinosh Skariyachan*[1]

[1]Department of Biotechnology, Dayananda Sagar College of Engineering, Bangalore, India
*Corresponding author: sinoshskariya@gmail.com
Phone: +91-9739654015, Fax: +91-80-26635929

ABSTRACT

Claudins are family of proteins of tight junctions establishing the paracellular barrier that controls the flow of molecules in the intercellular space between the cells of an epithelium. Claudin 1 protein plays significant role in Enteropathogenic *Escherichia coli* infection which is an important agent of infectious diarrhea, especially in pediatric populations. The crystal structure of Claudin 1 is yet unknown to the scientific public, hence a 3D structure is very essential for structural studies, protein – ligand interaction and designing of novel agonists against the infection. In this study we modelled a 3D structure of Claudin 1 by X-ray crystal structure of Apct Transporter of *Methanocaldococcus jannaschii* (PDB ID: 3GI9, Chain C) used as the template. Our study found that Claudin 1 predominantly consists of α helix. The RMSD value of modelled protein was found to be 2.0 A^o and steriochemical validation shows 88. 9% residues are in allowed region of Ramachandran plot. Further validation was done by various empirical force fields. Overall quality factor of the model identified to be 66.83 and error values of individual residues are negligible. The modeled protein was submitted to Protein Model Database and can be downloaded with the ID: PM0076543. Molecular docking studies with selected ligands were carried out and concluded that Cathepsin L and Celecoxib were the best ligands of choice with E_{total}– 162.3, -154.3 respectively. Our study concluded that these inhibitors could be used as potential drug candidates against Enteropathogenic *E.coli* Infections

KEY WORDS: EPEC, Claudin 1, Homology modeling, Validation, Cathepsin L, Molecular docking

INTRODUCTION

Enteropathogenic and Enterohemorrhagic *Escherichia coli* are important causal agent of infectious diarrhea, particularly amongst pediatric populations. While enteropathogenic *E. coli* is a significant health threat in developing countries, enterohemorrhagic *E. coli* causes sporadic, sometimes deadly outbreaks of hemorrhagic colitis, with a serious complication, hemolytic uremic syndrome, occurring in a proportion of cases. [1,2,3]. Infection of humans with enterohemorrhagic *Escherichia coli* (EHEC) serotype O157:H7 causes outbreaks of hemorrhagic colitis and hemolytic uremic syndrome and it is related to production of Shiga-like toxins, an important cause of acute kidney failure. The related pathogen enteropathogenic *E. coli* (EPEC) is an important cause of prolonged watery diarrhea in infants in underdeveloped countries. [4,5,6].

Tight junctions (TJ) are the most apical cell-cell contacts and are important for barrier function in epithelial and endothelial cells. A number of integral membrane proteins associated with the TJ have been identified during recent years, including occludin, junctional adhesion molecule, and the claudin family consisting of at least 24 members. [7,8,9,10] Claudins and occludins build up the functional units responsible for the tight sealing of the cells in epithelial sheets, whereas TJ proteins, such as Tight Junction Protein 1 (TJP1), are responsible for linking Claudins and occludin to the actin cytoskeleton.[11] Claudins comprise four transmembrane domains along with two extracellular loops and two cytoplasmic domains (Fig.1). Recent advances revealed that Claudins are directly involved in intercellular sealing of simple as well as stratified epithelia in vertebrates. [12,13,14]. During the course of infection, enteropathogenic and Enterohemorrhagic *Escherichia coli* subvert the host cell signaling machinery and hijack the actin cytoskeleton to tighten their interaction with the gut epithelium, while avoiding phagocytosis by professional phagocytes. [15,16,17,18]

Comparative modelling predicts the three-dimensional structure of a given protein sequence (target) based primarily on its alignment to one or more proteins of known structure (templates). The prediction process consists of fold assignment, target-template alignment, model building, and model evaluation.[19] Docking is a method which predicts the preferred orientation of one molecule to a second when bound to each other to form a

stable complex. [20]. A variety of experimental and computational techniques can be used to identify possible protein binding partners of protein. The prediction of putative protein– ligand interaction geometries by using computational docking methods is of increasing importance in the field of structure based drug designing [21]. In this study we have modelled a detailed 3D structure of Claudin 1 and docking studies were carried out to design and optimize novel therapeutic agents for Enteropathogenic *E.coli* infections

MATERIALS AND METHODS:

Sequence retrieval of Claudin 1

The protein sequence of Human Claudin 1- Accession: O95832 was retrieved from UNIPROT [22] database and the FASTA sequence is used for our studies

>sp|O95832|CLD1_HUMAN Claudin-1 OS=Homo sapiens GN=CLDN1 PE=1 SV=1
MANAGLQLLGFILAFLGWIGAIVSTALPQWRIYSYAGDNIVTAQAMYEGLWMSCVSQSTGQIQCKVFD
SLLNLSSTLQATRALMVVGILLGVIAIFVATVGMKCMKCLEDDEVQKMRMAVIGGAIFLLAGLAILVA
TAWYGNRIVQEFYDPMTPVNARYEFGQALFTGWAAASLCLLGGALLCCSCPRKTTSYPTPRPYPKPAP
SSGKDYV

Screening for best homologous templates

The best homologous proteins of Human Claudin 1 were selected based on the percentage of identity, similarity, expectation value and alignment scores using PSI-BLAST [23] algorithm across Proten Data Bank (PDB).

Multiple sequence alignment and phylogenetic characterization

The best homologous sequences were subjected to Multiple Sequence Alignment and Phylogenetic Characterization. MSA was performed by T-COFFEE [24] to analyze the conservation factor in the closely related proteins. The guide tree created from the best homologs was submitted to NJ Plot to interpret the evolutionary pattern of the homologous set.

Proteogenome analysis of Claudin1

The protein sequence of Claudin 1 was reverse translated by the Reverse Translate tool hosted in the ExPASy server. The nucleotide sequence of Claudin was subjected to Open Reading Frame analysis by the NCBI-ORF Finder. The locations of the functional genes present in the nucleotide sequence of claudin were analyzed by the GENSCAN [25]. The Secondary structure conformation of Claudin protein sequence was predicted by PSIPRED[26]. The domains and motifs, single peptides, internal repeats, intrinsic protein disorder and outlier homologues were predicted by the tool SMART[27]. The Post translational modifications of target protein sequence were predicted by dbPTM [28]. The hydrophobicity and topology prediction was done using TOP PRED[29]

In silico Comparative Modeling of Cludin-1 Protein

An *in silico* comparative modelling of the Claudin protein was performed by the MODELLER 9v7[30]. The best homolog identified earlier in the proteogenomic characterization was used to generate alignment, atom and the script files for modelling. The target and template were superimposed by DaliLite [31] server to analyze the RMSD value. The modelled protein was visualized by PyMOL[32].

Model refinement, validation and submission of modeled structure to PMDB

The modeled protein is validated by molecular dynamics and mechanics with the help of various empirical force fields such as ANOLEA [33], GROMOS [34] and VERIFY3D[35]. The parameters included the covalent bond distances and angles, steriochemical validation and atom nomenclature were validated using PROCHECK [36]. The statistics of non-bonded interactions between different atom types were detected and value of the error function was analyzed by ERRAT[37]. The modelled β-Enolase Protein was deposited to the Protein Model Data Base [38].

Selection of Potential Drug Candidates against Claudin1 protein

On the basis of literature studies and drug database survey, about 10 inhibitors were selected for docking process to identify the potential drug candidate against the Claudin 1 related disorders. The selected inhibitors are Aspirin[39], Cathepsin L[40], Sodium butyrate and Trichostatin[41] Indomethacin [42], Prostaglandin E2[43], Glucose[44], Cysteine[45], Celecoxib and Diclofenac[46]. The structural coordinates of these molecules were retrieved from NCBI PubChem [47] in SDF format and converted to PDB.

Molecular Docking

A rigid body docking was performed with HEX 6.1[48] by SP Fourier Transform, FFT steric scan, FFT final search and MM refinement. The clustering histogram with the scoring function was generated to analyze the binding energy of each selected conformations. The energy values in the clustering histogram of the docked complexes were analyzed. The ligand that produced the best E_{Total} , minimum energy, was selected as the best inhibitor. The docked complex is viewed and the interaction residues of amino acids with the ligands were analyzed by PyMOL

RESULTS AND DISCUSSION

Retrieval of protein sequence

The protein sequence of Human Claudin 1 consists of 211 amino acids. It plays a major role in tight junction-specific obliteration of the intercellular space, through calcium-independent cell-adhesion activity and also acts as a co-receptor for HCV entry into hepatic cells. It is expressed in liver and kidney, heart, brain, spleen, lung and testi**s.**

Screening of Best homologous

The PSI – BLAST results were analysed and the best protein hits were selected based on the percentage of identity, similarity and query coverage. The identity range was 29-50%, similarity range was 45-60% and E-value range was 2.6- 8.2. (Table-1)

MSA and phylogenetic characterization

The multiple sequence analysis was performed with T- COFFEE and the conservation present in the target protein was interpreted. The phylogenetic relationship between the selected templates and target was thus analysed.(Fig 2) The crystal structure of 'Apct Transporter bound to 7f11 monoclonal Fab fragment' of *Methanocaldococcus jannaschii* (PDB ID: 3GI9, Chain C) was found to have the closest relationship with the Human Claudin 1 protein. Hence the protein structure 3GI9 was selected as the best template for Human Claudin 1Protein for Homology Modelling.

Proteogenomic analysis

The ORF FINDER displayed all possible six frames of translation of the reverse translated Human Claudin 1 protein sequence. Out of the 6 frames of translation only +1, +2, -1, -2, and -3 frames possess the potential to translate to protein sequences. The +1 frame being the longest reading frame and it consist of many coding regions. The GENSCAN output showed that initial exon is present in the strand of 593 bps starting from 95 to 687. The molecular weight and isoelectric point of the protein were found to be 22743.8 Dalton and 8.41 respectively. The secondary structure of Claudin was predicted by PSI-PRED. It has been noticed that Claudin 1 consists majorly of helices. It was predicted that 56.4% were α helices, 9.95% were extended strand, 4.27% were β strands and 29.38% were random coils. The topology analysis revealed that protein consists of central transmembrane helical segment, internal helix cap, internal loop and an external loop. The hydrophobicity profile and the topology of Claudin protein were predicted by TOPPRED. The functional motifs predicted by SMART indicated a domain of Pfam ID: PF00822. The low complexity region is present between the 119 and 136 amino acids. The domain is expressed in many tissues but mainly by schwann cells as a component of myelin of the peripheral nervous system (PNS). It plays a role both in myelinisation and in cell proliferation. The proteins of this family are about 160 to 173 amino acid residues in size, and contain four transmembrane segments. PMP-22, EMP-1, -2 and -3 are highly similar.

Comparative modelling

A pair wise alignment between the target and temple is performed (Fig 3).The Modeller program was executed and a series of files were generated and the final modelled protein was written as 'Claudin.1B99990001' The superimposition was performed to analyse the backbone threading and fold recognition of target. (Fig.4) .There is 207 residues are aligned between the target and template with a Z score of 18.5. It has noticed that RMSD value is 2.0 and which indicate the backbone configuration of the protein is good. The modelled protein was visualized in PyMol. It was seen that the helices were predominant in the modelled protein and coils were present in small proportion. (Fig.5).

Validation of the modelled protein

The refinement of model was done by various empirical force fields such as ANOLEA, GROMOS and VERIFY-3D (Fig.7).The modelled protein is validated by Ramachandran Plot generated by PROCHECK. The Ramchandran plot accounts for 88.9% of the residues in the allowed region, 8.9% of the residues in the

additionally allowed regions, about 1.7% in the generously allowed and disallowed regions each.(Fig.6). The overall quality of the model is identified to be 66.832.(Fig.8).

Submission of the modelled structure in to Protein model database

The modelled protein was then submitted to Protein Model database with the unique ID PM0076543. This structure is available in the database and can be downloaded for further structural studies.

Docking between of Claudin with best inhibitors

The docking was performed effectively with all the selected ligands. (Table II) The best ligand was analysed based on the E_{Total} value given by the clustering histogram. Cathepsin L was analysed as the best ligand with E_{Total} -162.3 (Fig.9) Celecoxib also showed a significant binding affinity to the protein with energy minima of -154.3. The main interacting residues of protein with Cathepsin L are ARG158, THY159 and MET102.The docking study clearly shows that Cathepsin L and Celecoxib have strong inhibitory activity against claudin 1 and these could be used as drugs of choice against Enteropathogenic *E.coli* Infections.

CONCLUSION

Claudin-1 is a member of the Claudin protein family that participates in the formation of tight junctions between adjacent cells. The hallmark of EPEC/EHEC infections is induction of attaching and effacing (A/E) lesions that damage intestinal epithelial cells. Thus, agents capable of enhancing or maintaining the epithelial barrier, through preservation of the tight junction or via modulating intracellular signalling pathways that compete with those activated by bacteria, may be of value in reversing the effects of EHEC infection.

The structure of Claudin 1 is not yet known to public as there is no evidence in crystallographic data in PDB and other structural databases. Thus there is a need for the three dimensional structure of Claudin 1 for study of complete structural bioinformatics, protein – ligand interaction and designing of new agonists against the infection. The initial studies were mainly focused on proteogenomic aspects of Claudin 1. The structure of Claudin 1 is modelled based on the template crystal structure of 'Apct Transporter protein of *Methanocaldococcus jannaschii.* The model is validated by various empirical force fields. The Ramchandran plot accounts for 88.9% of the residues in the allowed region, 8.9% of the residues in the additionally allowed regions, about 1.7% in the generously allowed and disallowed regions each. The overall quality of the model is identified to be 66.832. The secondary structure of Claudin 1 mainly consists of α helices and turns. The validity of the modelled protein was so significant and it is submitted to Protein Modelling Database with unique ID, PM0076543. The suitable ligands against Claudin were shortlisted based on the survey of various literatures. The docking study indicated that Cathepsin L and Celecoxib were good inhibitors against Claudin 1 and these could be used as potential drugs of choice against Enteropathogenic *E.coli* Infections

REFERENCES

[1] Peter J. M. Ceponis, Derek M. McKay, Joyce C. Y. Ching, Perpetual Pereira, Philip M. Sherman, Mar, (2003), Enterohemorrhagic *Escherichia coli* O157:H7 Disrupts Stat1-Mediated Gamma Interferon Signal Transduction in Epithelial Cells, *Infect. Immun*., 2003, 71, 1396–1404.

[2] Smalley K.S, Brafford P, Haass N.K, Brandner J.M, Brown E, et al., , Up regulated expression of zonula occludens protein-1 in human melanoma associates with N-cadherin and contributes to invasion and adhesion, *Am J Pathol* 2005,166: 1541–1554

[3] Donnenberg M. S and T. S. Whittam, Pathogenesis and evolution of virulence in enteropathogenic and Enterohemorrhagic *Escherichia coli, J. Clin. Investig*. 2001 107:539–548.

[4] Punita Dhawan, Amar B. Singh, Natasha G. Deane,YiRan No, Sheng-Ru Shiou,Carl Schmidt,John Neff, M. Kay Washington, and R. DanielBeauchamp,. Claudin-1 regulates cellular transformation and metastatic behavior in colon cancer , *J Clin Invest*, 2005, 115,1765-1776

[5] Moon, R.T., Bowerman, B., Boutros, M., and Perrimon,N. The promise and perils of Wnt signaling through beta-catenin,. *Science.* 2002, 296:1644–1646.

[6] Wong, N.A.C.S., and Pignatelli, M.. Betacatenin— a linchpin in colorectal carcinogenesis, *Am. J. Pathol*. 2002, 160:389–401.

[7] Johnson-Henry, K., J. L. Wallace, N. S. Basappa, R. Soni, G. K. P. Wu, andP. M. Sherman.,Inhibition of attaching and effacing lesion formation following enteropathogenic *Escherichia coli* and Shiga toxin-producing E.coli infection. *Infect. Immun*. 2001, 69:7152–7158.

[8] Tsukita S. H, Furuse M, Occludin and claudins in tight-junction strands: leading or supporting players Trends, *Cell Biol* 1999,9: 268–273.

[9] Teresa Lopardo., Nadia Lo Iacono., Barbara Marinari, Maria L. Giustizieri, Daniel G. Cyr, Giorgio Merlo, Francesca Crosti, Antonio Costanzo, Luisa Guerrini .Claudin-1 Is a p63 Target Gene with a Crucial Role in Epithelial Development , 2008 *PLoS ONE* 2008, 3: 3, 1-8.

[10] Hecht G, Microbes and microbial toxins: paradigms for microbial mucosal interactions. VII. Enteropathogenic *Escherichia coli*: physiological alterations from an extracellular position, *Am. J. Physiol*. 2001, 281:G1-G7.

[11] Ching J. C, N. L. Jones, P. J. M. Ceponis, M. A. Karmali, and P. M. Sherman, *Escherichia coli* Shiga-like toxins induce apoptosis and cleavage of poly (ADP-ribose) polymerase via in vitro activation of caspases, 2002, *Infect. Immun.* 70:4669–4677.

[12] Smalley KS, Brafford P, Haass NK, Brandner JM, Brown E, Up regulated expression of zonula occludens protein-1 in human melanoma associates with N-cadherin and contributes to invasion and adhesion. *Am J Pathol* 2005,166: 1541–1554

[13] Anderson J. M, Molecular structure of tight junctions and their role in epithelial transport. 2001, *News Physiol Sci*, 13:126–130.

[14] Emmanuelle Caron, Valerie F Crepin, Nandi Simpson, Stuart Knutton, Junkal Garmendia and Gad Frankel.. Subversion of actin dynamics by EPEC and EHEC, *Curr Opin Microbiol.* 2006,9 :40-45.

[15] Johnson-Henry, K. J. L. Wallace, N. S. Basappa, R. Soni, G. K. P. Wu, and P. M. Sherman, Inhibition of attaching and effacing lesion formation following enteropathogenic *Escherichia coli* and Shiga toxin-producing *E.coli* infection, *Infect. Immun*. 2001,69:7152–7158
[16] Michael S. Donnenberg and Thomas S. Whittam, Enteropathogenic and Enterohemorrhagic *Escherichia coli* Virulence Gene Regulation, 2007, *Infect. Immun* 4199–4210.
[17] Pupo G.M, Karaolis D. K. R, Lan R.T, and Reeves P.R, Evolutionary relationships among pathogenic and nonpathogenic *Escherichia coli* strains inferred from multi locus enzyme electrophoresis and mdh sequence studies, *Infect. Immun.* 1997, 65:2685–2692.

[18] Horne C, Vallance B. A, Deng W, Finlay B. B, Current progress in enteropathogenic and Enterohemorrhagic *Escherichia coli* vaccines, *Expert Rev Vaccines.* 2002, 4 :483-93.
[19] Martí-Renom M. A, Stuart A. C, Fiser A, Sánchez R, Melo F, Sali A, Comparative protein structure modelling of genes and genomes, *Annu Rev Biophys Biomol Struct*, 2000, 291-325.
[20] Martin Zachairias, Protein–protein docking with a reduced protein model accounting for side-chain flexibility, *Protein Sci*, 2003, 1271–1282
[21] Lyskov S. and Gray J. J, The Rosetta Dock server for local protein-protein docking, *Nucleic Acids Res.,* 2008, 36: 233–238
[22] Brigitte Boeckmann, Amos Bairoch,Rolf Apweiler, Marie-Claude Blatter, Anne Estreicher, Elisabeth Gasteiger; The SWISS-PROT protein knowledgebase and its supplement TrEMBL, *Nucleic Acids Res,* 2003, 31: 365-370
[23] Altschul S F, Gish W, Miller W, Myers E W and Lipman D J; Basic local alignment search tool. *J. Mol. Biol.* 1990,215:403–410
[24] C. Notredame, D. Higgins, J. Heringa, T-Coffee: A novel method for multiple sequence alignments".*J of Mol Bio*, 2000,,302, 205-217,
[25] Burge, C. and Karlin, S. Prediction of complete gene structures in human genomic DNA. *J. Mol. Biol.* 1997, 268, 78-94.
[26] McGuffin LJ, Bryson K, Jones.The PSIPRED protein structure prediction server, *Bioinformatics*. 2000,16:404-5.
[27] Jörg Schultz,Frank Milpetz,Peer Bork, Chris P. Ponting SMART, a simple modular architecture research tool: Identification of signaling domains. *Proc. Natl. Acad. Sci.* 1998 ,95:5857-5864
[28] Tzong-Yi Lee1,5, Justin Bo-Kai Hsu1, Wen-Chi Chang1,6, Ting-Yuan Wang1,Po-Chiang Hsu2 and Hsien-Da Huang, A comprehensive resource for integrating and displaying protein post-translational modifications, *BMC Res*, 2009, J.2:111
[29] Claros MG, von Heijne G. TopPred II: Improved software for membrane protein structure predictions. *Comput Appl Biosci*. 1994, 10:685-6.
[30] Sali A, Marti-Renom M A, Stuart A C, Fiser A, Sanchez R, Melo F,Comparative protein structure modeling of genes and genomes. *Annu Rev Biophys Biomol Struct* 2000, 29: 291-325
[31] L. Holm, S. Kääriäinen, P. Rosenström, and A. Schenkel, Searching protein structure databases with DaliLite v.3,*Bioinformatics;* 2008, 24: 2780–2781
[32] Daniel Seeliger and Bert L. de Groot, Ligand docking and binding site analysis with PyMOL and Autodock/Vina, *J Comput Aided Mol Des.* 2010, 24 : 417–422
[33] Melo F, Devos D, Depiereux E, Feytmans E.ANOLEA: a www server to assess protein structures. *Proc Int Conf Intell Syst Mol Biol*. 1997, 5:187-90.
[34] Cao Z, Lin Z, Wang J, Liu H..Refining the description of peptide backbone conformations improves protein simulations using the GROMOS 53A6 force field. *J Comput Chem*. 2009, 31:1-23
[35] Eisenberg D, Lüthy R, Bowie JU, VERIFY3D: assessment of protein models with three-dimensional profiles. *Methods Enzymol.* 1997, 277:396-404.
[36] Laskowski, R.A. PROCHECK: a program to check the stereiochemical quality of protein structures. *J. Appl. Cryst.,* 1993, 26:283–291.
[37] Colovos C, Yeates TO. Verification of protein structures: patterns of nonbonded atomic interactions. *Protein Sci*: 1993, 12: 1511-9

[38] Tiziana Castrignanò, Paolo D'Onorio De Meo, Domenico Cozzetto1, Ivano Giuseppe Talamo1 and Anna Tramontano, The PMDB Protein Model Database, *Nucleic Acids Res*: 2006, 34:10.1093/nar/gkj105
[39] Oshima T, Miwa H, Joh T, Aspirin induces gastric epithelial barrier dysfunction by activating p38 MAPK via claudin-7. *Am J Physiol Cell Physiol*. 2008, 295:800-6
[40] François Boudreau,1 , Carine R. Lussier, Sébastien Mongrain, Mathieu Darsigny, Loss of cathepsin L activity promotes claudin-1 overexpression and intestinal neoplasia, *FASEB Journal*. 2007;21:3853-3865
[41] Krishnan M, Singh AB, Smith JJ, Sharma A, Chen X, Eschrich S, Yeatman TJ, Beauchamp RD, Dhawan P, HDAC inhibitors regulate claudin-1 expression in colon cancer cells through modulation of mRNA stability,Oncogene. 2010, 29: 305-12.
[42] Mima S, Tsutsumi S, Ushijima H, Takeda M, Fukuda I, Yokomizo K, Suzuki K, Sano K, Nakanishi T, Induction of claudin-4 by nonsteroidal anti-inflammatory drugs and its contribution to their chemopreventive effect, *Cancer Res*. 2005,65:1868-76.
[43] Marcelo N Tanaka , Bruno L Diaz , Wanderley de Souza Jose A Morgado-Diaz Prostaglandin E2-EP1 and EP2 receptor signaling promotes apical junctional complex disassembly of Caco-2 human colorectal cancer cells.*BMC Cell Biology*, 2008, 9:63-68
[44] Marta Villarroela, Marta García-Ramíreza, b, Lidia Corralizaa, b, Cristina Hernándeza, b and Rafael Simóa, Effects of high glucose concentration on the barrier function and the expression of tight junction proteins in human retinal pigment epithelial cells, *Exp Eye Res.* 2009, 89:913-20.
[45] Kovalenko OV, Yang XH, Hemler ME A novel cysteine cross-linking method reveals a direct association between claudin-1 and tetraspanin CD9.Mol Cell Proteomics. 2007, 6: 1855-67.
[46] Blanca L. Valle, Patrice J. Morin, Claudins in Cancer Biology, *Current Topics in Membranes*, 2010, 65: 293-333
[47] Wang Y, Bolton E, Dracheva S, Karapetyan K, Shoemaker BA, Suzek TO, Wang J, Xiao J, Zhang J, Bryant S H. An overview of the PubChem BioAssay resource. *Nucleic Acids Res*: 2010, 255-66
[48] Ritchie D W, Evaluation of protein docking predictions using Hex 3.1 in CAPRI rounds 1 and 2, *Proteins*. 2003; 98-106

Table 1: Best Homologous Template Structures for Claudin 1 protein

PDB ID	Chain	Length	Percentage of Identity	Percentage of Similarity	E-value	Organism
3D6L	A	137	29	53	2.6	*Campylobacter jejuni*
1VFG	A	390	36	50	5.1	*Aquifex aeolicus*
2W0A	A	455	42	60	5.9	*Mycobacterium tuberculosis*
1Y5L	C	225	38	58	7.9	*Escherichia coli*
3GI9	C	444	50	58	8.2	*Methanocaldococcus jannaschii*

Table 2 Docking binding energies of selected inhibitors towards modeled Claudin 1

Sl.No.	Pub chem ID	Inhibitor	Binding energy
1	CID:2244	Aspirin	-81.2
2	CID:16760361	Cathepsin L	-162.3
3	CID:2662	Celecoxib	-154.3
4	CID:594	Cysteine	-50
5	CID:3033	Diclofenac	-124.5
6	CID:24749	Glucose	-111.1
7	CID:3715	Indomethacin	-123.4
8	CID:5280360	Prostaglandin E2	-132.6
9	CID:5222465	Sodium butyrate	-116.8
10	CID:444732	Trichostatin A	-132

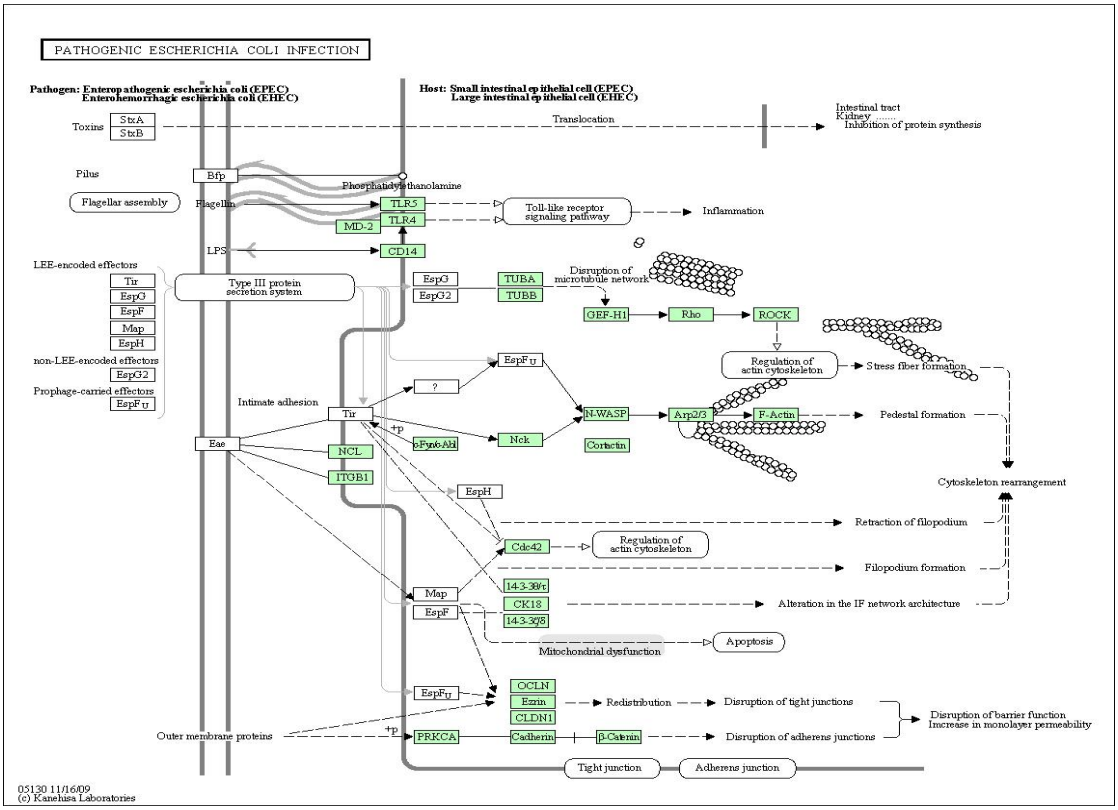

Fig 1 KEGG Molecular pathway showing the role of Claudin 1 protein in Enteropathogenic *E.coli* infections

Fig2 Phylogenetic Characterization of the best homologs of Claudin 1

```
gi|256032630|pdb|3GI9|C      MELKNKKLSLWEAVSMAVGVMIGASIFSIFGVGAKIAGRNLPETFILSGI
sp|O95832|CLD1_HUMAN         ----------------------------------------------MANAGL
                                                                            :  ::: 

gi|256032630|pdb|3GI9|C      YALLVAYSYTKLGAKIVSNAGPIAFIHKAIGDNIITG-ALSILLWMSYVI
sp|O95832|CLD1_HUMAN         QLLGFILAFLGWIGAIVSTALPQWRIYSYAGDNIVTAQAMYEGLWMSCVS
                               *  .   ::     . ***.* *    *:.   ****:*. **    **** *

gi|256032630|pdb|3GI9|C      SIALFAKGFAGYFLPLINAPINTFNIAITEIGIVAFFTALNFFGSKAVGR
sp|O95832|CLD1_HUMAN         -------------------------QSTGQIQCKVFDSLLNLSSTLQATR
                                                      :   ::    .* : **: .:   . *

gi|256032630|pdb|3GI9|C      AEFFIVLVKLLILGLFIFAGLITIHPSYVIPDLAPSAVSGMIFASAIFFL
sp|O95832|CLD1_HUMAN         ---ALMVVGILLGVIAIFVATVGMKCMKCLEDDEVQKMRMAVIGGAIFLL
                                  :::* :*:   : **.. : ::     : *    . :    ::..***:*

gi|256032630|pdb|3GI9|C      SYMGFGVITNASEHIENPKKNVPRAIFISILIVMFVYVGVAISAIGNLPI
sp|O95832|CLD1_HUMAN         AGLAILVAT----------------AWYGNRIVQEFYDPMTPVNAR-----
                             : :.: * *                * : . ::   *     ..:.*

gi|256032630|pdb|3GI9|C      DELIKASENALAVAAKPFLGNLGFLLISIGALFSISSAMNATIYGGANVA
sp|O95832|CLD1_HUMAN         YEFGQALFTGWAAASLCLLG----------GALLCCSCPRKTTSYP-TPRP
                              *: :*   .. *.*:   :**          ***:. *.. ::* *   :   .

gi|256032630|pdb|3GI9|C      YSLAKDGELPEFFERKVWFKSTEGLYITSALGVLFALLFNMEGVASITSA
sp|O95832|CLD1_HUMAN         YPKPAPSSGKDYV-------------------------------------
                             *. .  ..  ::.
```

Fig 3 Sequence alignment between target Claudin and template

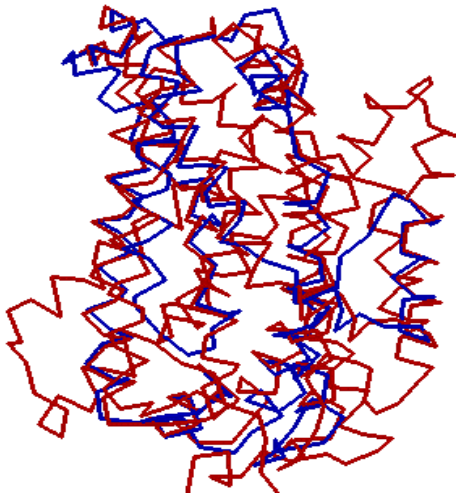

Fig 4. Superimposed structure of target and template

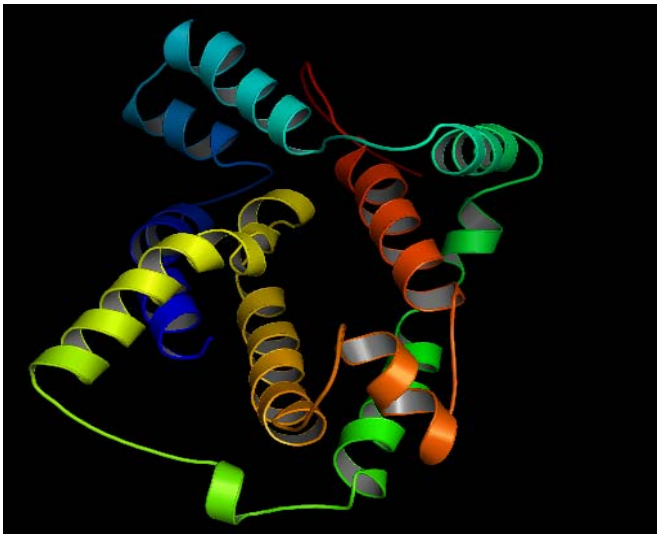

Figure 5 : Modelled Claudin protein visualized by PyMol

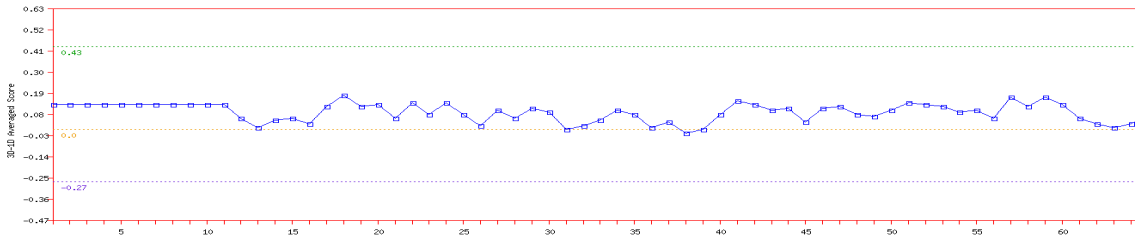

Fig 6 VERIFY 3D plot of modelled protein

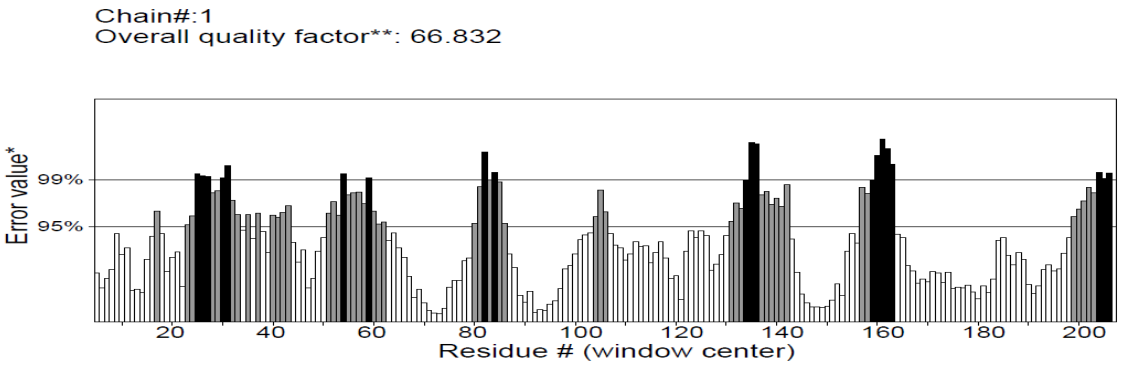

Fig 8 Quality factor analysis of Claudin 1 by ERRAT

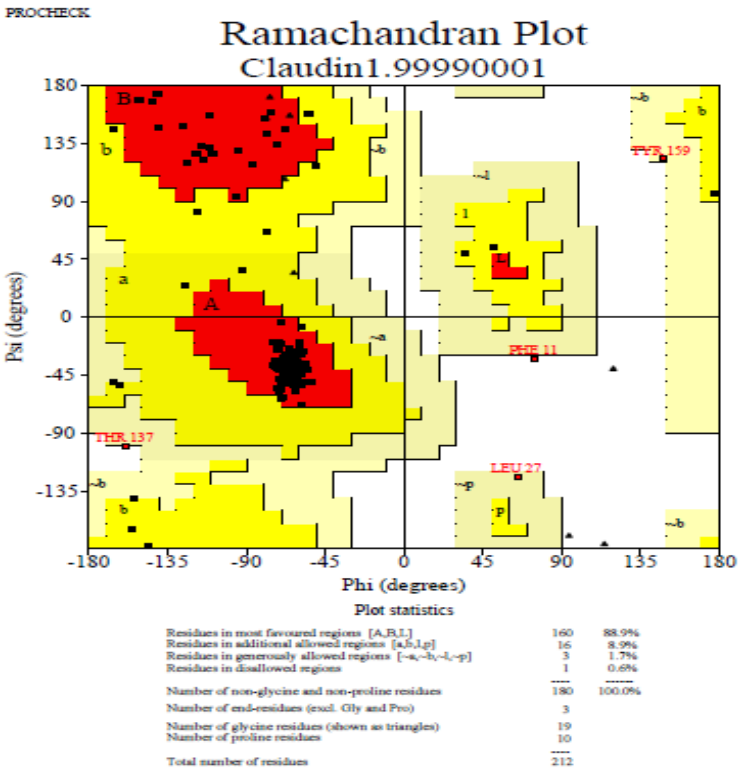

Fig 7. Ramchandran plot of Modeled protein generated by PROCHECK

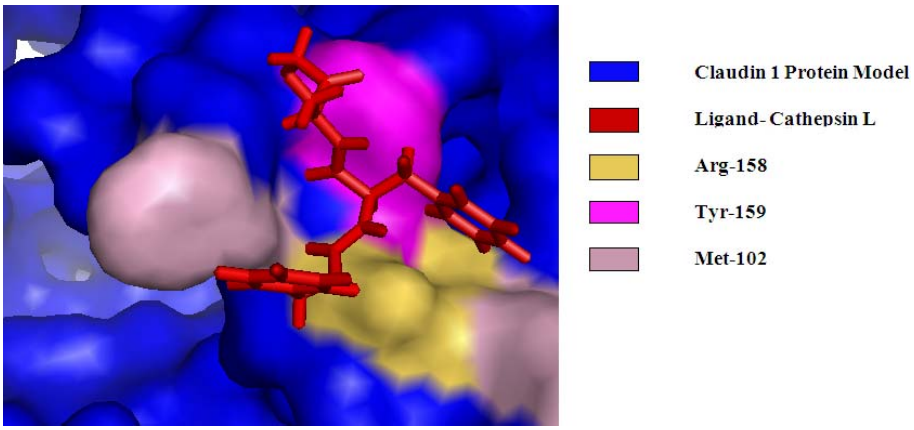

Fig 9 Docked complex showing the Cathepsin L to modelled Claudin