Insilico Drug Designing and Pharmacokinetics Studies on the Gene, Macc1 (Metastasis Associated Gene in Colon Cancer 1), Responsible for Colon Cancer

S. Maithreyee [1], Prabha.V[2], Nithya.G[3] and Balaji Munivelan[4]

1. *Research Scholar, PG & Research ,Dept.of Biochemistry ,DKM College for Women,Vellore.*
2. *Head of Department , PG & Research ,Dept.of Biochemistry ,DKM College for Women,Vellore.*
3. *Asst.Professor, PG & Research ,Dept.of Biochemistry ,DKM College for Women,Vellore*

CEO & Seniour Bioinformatician, ABS Geno-informatics , Chennai

Abstract: *Globally more than 1 million people get colorectal cancer every year resulting in about 715,000 deaths as of 2010 up from 490,000 in 1990. As of 2012, it is the second most common cause of cancer in women (9.2% of diagnoses) and the third most common in men (10%). In this research, we focused on the potential mutated gene, MACC1 - Metastasis-Associated Gene in Colon Cancer 1. We examined the potential mutated sites which are responsible for colon cancer. We modeled the 3D structure of both the normal and the mutated protein and validated it. The potential existing drug and the anti-oxidant agent were combined using advanced cheminformatics protocols. The designed drug was validated using pharmacokinteics techniques (ADME and Toxicity). Finally the designed drug and the control drug were docked using automated drug docking server. Based on the drug binding scores, we deduced that the designed drug has good binding affinities with the target modeled protein. Hence, we finally conclude that the regions of MACC1 act as therapeutic agent for Colon cancer.*

Keywords: MACC1 Colorectal Cancer and Protein Modeling

INTRODUCTION

Worldwide more than 1 million people are affected by colorectal cancer every year bringing about approximately 715,000 deaths as of 2010 up from 490,000 in 1990 [1]. Colorectal cancer (CRC) is also called as bowel cancer, colon cancer, or rectal cancer. It is the development of cancer from the colon or rectum. Signs and symptoms comprise of blood in the stool, change in bowel movements, weight loss, and feeling tired all the time. Worldwide, colorectal cancer is the third most common type of cancer. It is more common in developed countries, where more than 65% of cases are found. It is more common in men than women. Epigenetic alterations are very frequent in colon cancer. Only 1 or 2 oncogene mutations and 1 to 5 tumor suppressor mutations occur in an average colon cancer, with about 60 further "passenger" mutations [2].

METHODOLOGY

Bioinformatics

- **Target selection:** Based on the various clinical and molecular genetics literature studies, MACC1 gene was chosen which is directly involved in Human Colon Cancer. OMIM (https://pubmed.ncbi.nlm.nih.gov/19098908/) and Cancer index (http://www.cancerindex.org/geneweb/MACC1.htm). **Sequence retrieval system:** The gene coded protein sequence was retrieved from uniprot database in *FASTA* format. *Uniprot* proteomics database (https://www.uniprot.org/uniprot/Q6ZN28).
- **Protein sequence analysis:** The protein sequence analysis studies were done using *Scan Prosite (*https://prosite.expasy.org/*)* database in order to identify the active sites, motifs and domain of the MACC1. **Protein 3D structure prediction:** The amino acids sequence of

MACC1 protein was converted in to 3D structure using an automated homology modelling server called *CPH 3.0 model* server (http://www.cbs.dtu.dk/services/CPHmodels/).

Cheminformatics

- Drug compound selection: The existing drug compound of colon cancer was retrieved from *NCBI Pubchem* Compound database (https://pubchem.ncbi.nlm.nih.gov/) in order perform drug designing studies. The retrieved chemical 2D structure was converted into 3D structure using Discovery studio software. **Drug designing and validation**: The existing drug molecule and antioxidant molecule were combined using Molinspiration software (https://www.molinspiration.com/) and the drug likeness score properties were validated.

Pharmacokinetics

- ADMET prediction: the complete ADMET properties studies were done using an advanced pharmacokinetics server Swiss ADMET (http://www.swissadme.ch/).

Molecular drug docking

- The designed chemical compounds were introduced with the modelled protein colon cancer target MACC1 3D structure using an automated molecular protein –drug docking server called *PatchDock* (https://bioinfo3d.cs.tau.ac.il/PatchDock/). The docking results were validated using discovery studio software in order to view the H-bond interaction and drug – protein bind affinities.

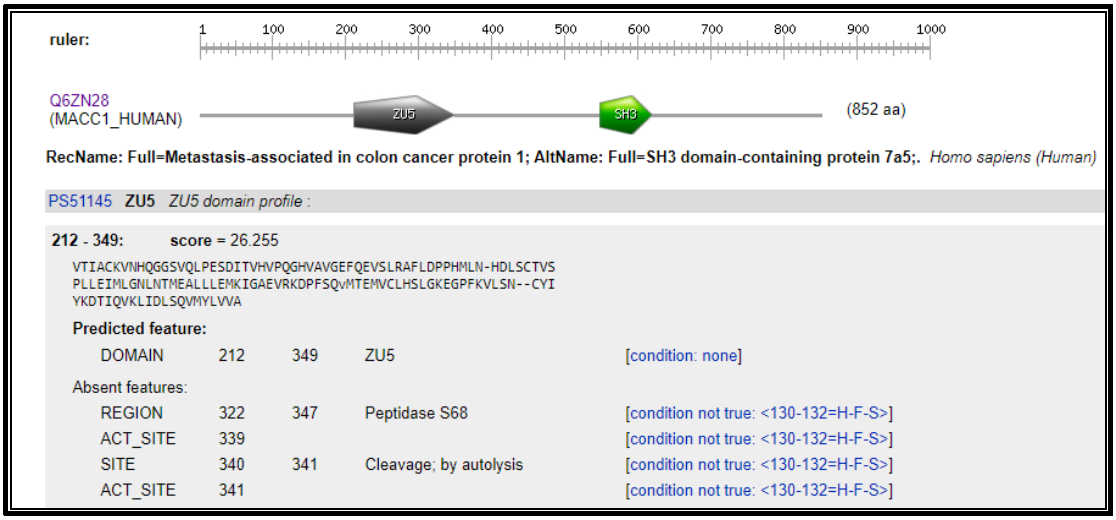
RESULTS AND DISCUSSION

Table 1: Target Gene Summary

Gene Name	Protein Name	Chr Loc	Gene ID	OMIM ID	Nucleotide Length	Protein Length	Uniprot ID
MACC1	Metastasis-Associated Gene In Colon Cancer 1	7	346389	612646	2559 nt	852 aa	Q6ZN28

The above table gives the molecular protein information of *MACC1- Metastasis-Associated Gene in Colon Cancer 1*.

Fig: 2 Motif and Domain prediction



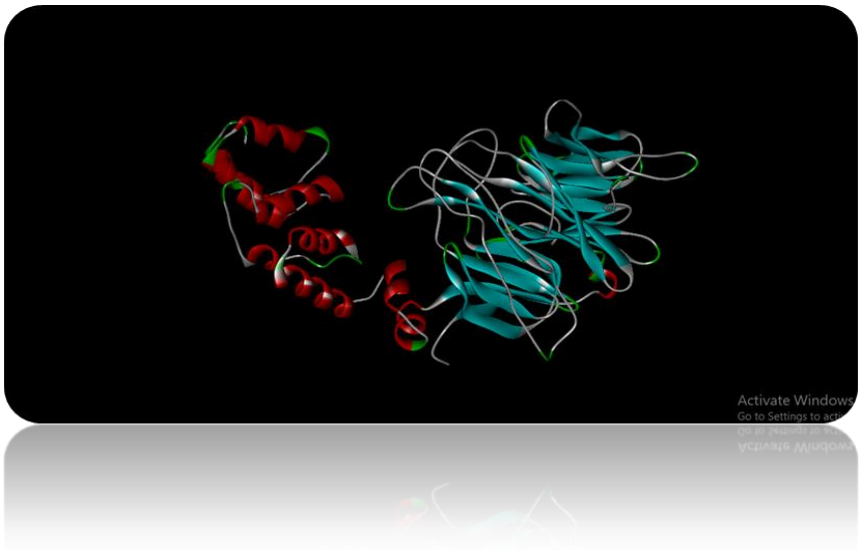
The above picture shows the Scan Prosite result page. MACC1 protein sequence represents the active amino acids with domain range.

In this research investigation, the colon cancer target protein MACC1 is present in the 7th chromosome and the length of its gene sequence and protein sequence is 2559nt and 852 aa respectively. In the primary step of the investigation, motif and domain analysis were performed. Fig.2 shows the functional part of the amino acids in MACC1. The identified functional part of the MACC1 protein sequence was analyzed using scan prosite tool. The results clearly elucidate the domain and motifs present in the sequence. The predicted function of the domain range 212 to 349 is ZU5 and that of 549 to 619 is SH3. The amino acids present in the active sites 339, 340 and 341 are Peptidase S68, Cleavage and autolysis respectively.

ZU5 domain is created by a compact β-sheet-rich core with an unstable number of surface loops and helices [3] and [4]. The β-strand-rich core consists of two antiparallel sheets interacting with each other in parallel. According to our research finding, MACC1 protein matches with the domain signature of ZU5 domain.

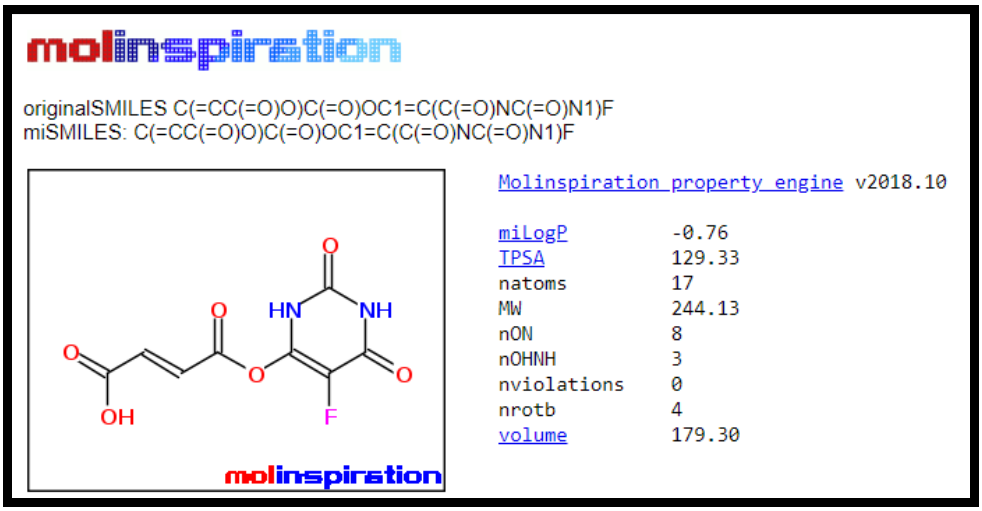
Src homology 3 (SH3) domain is a small protein domain of approximately 60 amino-acid residues first found out as a homology sequence in the non-catalytic part of various cytoplasmic protein tyrosine kinases (e.g. Src, Abl, Lck) [5] [6] [7]. It has been identified in a large variety of other intracellular or membrane-associated proteins. The SH3 domain has a characteristic fold which consists of five or six β-strands positioned as two tightly packed anti-parallel β sheets.. Here, CPH model server was used which is a web-server predicting protein 3D-structure by use of single template homology modelling. The server uses a hybrid of the scoring functions of CPHmodels-2.0 and a novel remote homology-modelling algorithm. MACC1 sequence was modelled using the fast CPHmodels-2.0 profile-profile scoring function suitable for close homology modelling. Based on CPH model server technique [8] [9], MACC1 was applied to this server to model the 3D structure. The predicted structure was viewed with the molecular visualization tool, Discovery studio software. Fig.3 to 7 shows the 3D structure of MACC1 protein. Each figure explains the intra-molecular structural properties of the secondary structure, atoms model, N and C terminals, hydrophobicity and the properties of amino acid residues. These findings play a significant role in structure-based drug designing.

Fig: 3 Three dimensional structure of *MACC1-protein*



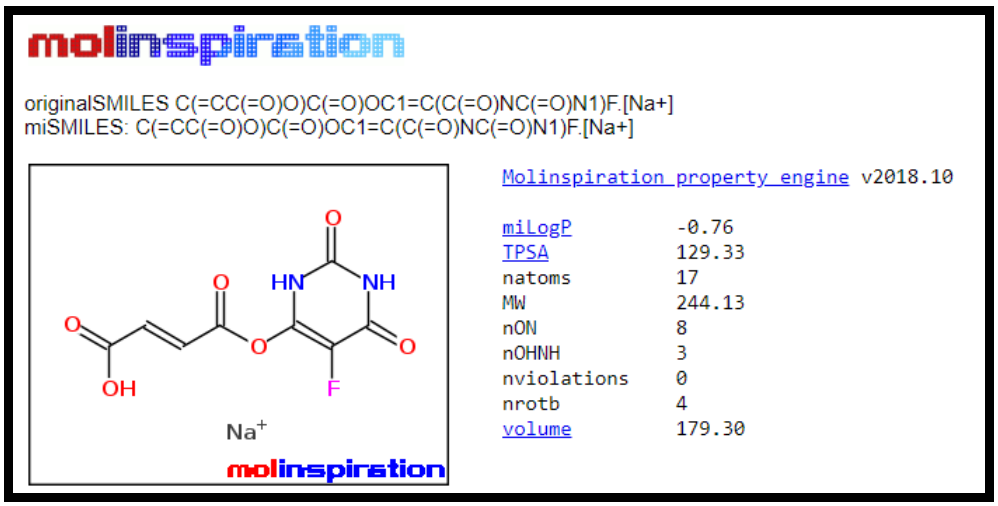
The above picture shows the 3D structure of MACC1 protein in secondary structure model color using discovery studio software. Colour indication: red –helix, blue-sheets, green-turns and white-coiled region.

Fig: 10 MOLINSPIRATION
Calculation of the Molecular Properties of *de novo* Drug 1



The above picture shows the validation studies of the designed *de novo* drug 1 (5-Fluorouracil + Fumaric Acid) using Molinspiration software (Calculation of Molecular Properties)

Fig : 13 Molinspiration – calculation of molecular properties



The above picture shows the validation studies of the designed *de novo* drug 2 (5-Fluorouracil + Fumaric Acid + Sodium) using Molinspiration software (Calculation of Molecular Properties) *ACC1-protein*

Fig.16: Pharmacokinetics studies

ADMET *de novo drug*

Physicochemical Properties	
Formula	C8H5FN2O6
Molecular weight	244.13 g/mol
Num. heavy atoms	17
Num. arom. heavy atoms	6
Fraction Csp3	0.00
Num. rotatable bonds	4
Num. H-bond acceptors	7
Num. H-bond donors	3
Molar Refractivity	50.05
TPSA	129.32 Å²
Lipophilicity	
Log $P_{o/w}$ (iLOGP)	0.48
Log $P_{o/w}$ (XLOGP3)	-0.93
Log $P_{o/w}$ (WLOGP)	-0.83
Log $P_{o/w}$ (MLOGP)	-0.07
Log $P_{o/w}$ (SILICOS-IT)	0.98
Consensus Log $P_{o/w}$	-0.07
Water Solubility	
Log *S* (ESOL)	-0.76
Solubility	4.19e+01 mg/ml ; 1.72e-01 mol/l
Class	Very soluble
Log *S* (Ali)	-1.30
Solubility	1.22e+01 mg/ml ; 4.99e-02 mol/l
Class	Very soluble
Log *S* (SILICOS-IT)	-0.91
Solubility	3.03e+01 mg/ml ; 1.24e-01 mol/l
Class	Soluble
Pharmacokinetics	
GI absorption	High
BBB permeant	No
P-gp substrate	No
CYP1A2 inhibitor	No
CYP2C19 inhibitor	No
CYP2C9 inhibitor	No
CYP2D6 inhibitor	No

CYP3A4 inhibitor	No
Log K_p (skin permeation)	-8.45 cm/s
Druglikeness	
Lipinski	Yes; 0 violation
Ghose	No; 1 violation: WLOGP<-0.4
Veber	Yes
Egan	Yes
Muegge	Yes
Bioavailability Score	0.56
Medicinal Chemistry	
PAINS	0 alert
Brenk	1 alert: michael_acceptor_1
Leadlikeness	No; 1 violation: MW<250
Synthetic accessibility	2.51

The above image shows the output page of the SWISS ADMET server. The above values clearly show that the designed drug compound 1 (5-Fluorouracil + Fumaric Acid) obeys the standard ADMET profiles.

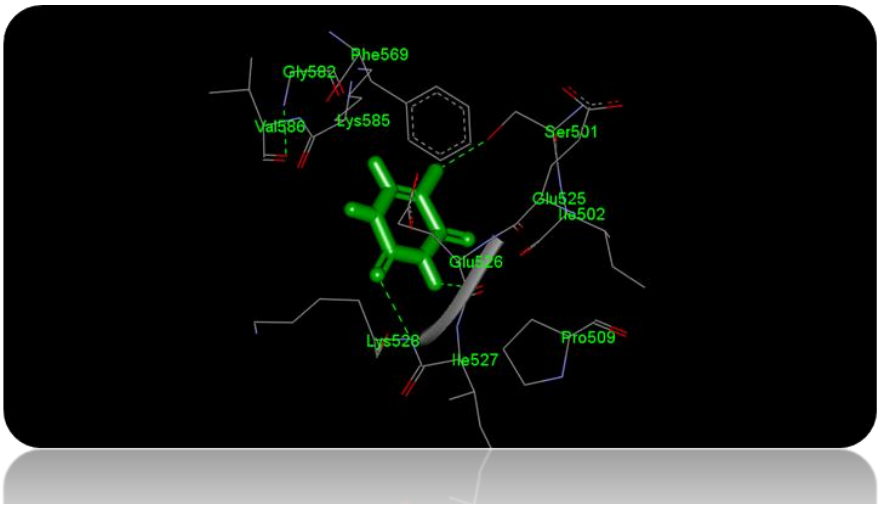
Fig: 17. Pharmacokinetics studies
ADMET *de novo drug 2*

Physicochemical Properties	
Formula	C8H5FN2NaO6
Molecular weight	267.12 g/mol
Num. heavy atoms	18
Num. arom. heavy atoms	6
Fraction Csp3	0.00
Num. rotatable bonds	4
Num. H-bond acceptors	7
Num. H-bond donors	3
Molar Refractivity	50.05
TPSA	129.32 Å²
Lipophilicity	
Log $P_{o/w}$ (iLOGP)	-8.95
Log $P_{o/w}$ (XLOGP3)	-0.93
Log $P_{o/w}$ (WLOGP)	-0.83
Log $P_{o/w}$ (MLOGP)	-0.07
Log $P_{o/w}$ (SILICOS-IT)	0.98
Consensus Log $P_{o/w}$	-1.96

Water Solubility	
Log *S* (ESOL)	-0.89
Solubility	3.42e+01 mg/ml ; 1.28e-01 mol/l
Class	Very soluble
Log *S* (Ali)	-1.30
Solubility	1.33e+01 mg/ml ; 4.99e-02 mol/l
Class	Very soluble
Log *S* (SILICOS-IT)	-0.91
Solubility	3.32e+01 mg/ml ; 1.24e-01 mol/l
Class	Soluble
Pharmacokinetics	
GI absorption	High
BBB permeant	No
P-gp substrate	Yes
CYP1A2 inhibitor	No
CYP2C19 inhibitor	No
CYP2C9 inhibitor	No
CYP2D6 inhibitor	No
CYP3A4 inhibitor	No
Log K_p (skin permeation)	-8.59 cm/s
Druglikeness	
Lipinski	Yes; 0 violation
Ghose	No; 1 violation: WLOGP<-0.4
Veber	Yes
Egan	Yes
Muegge	Yes
Bioavailability Score	0.55
Medicinal Chemistry	
PAINS	0 alert
Brenk	1 alert: michael_acceptor_1
Leadlikeness	Yes
Synthetic accessibility	2.56

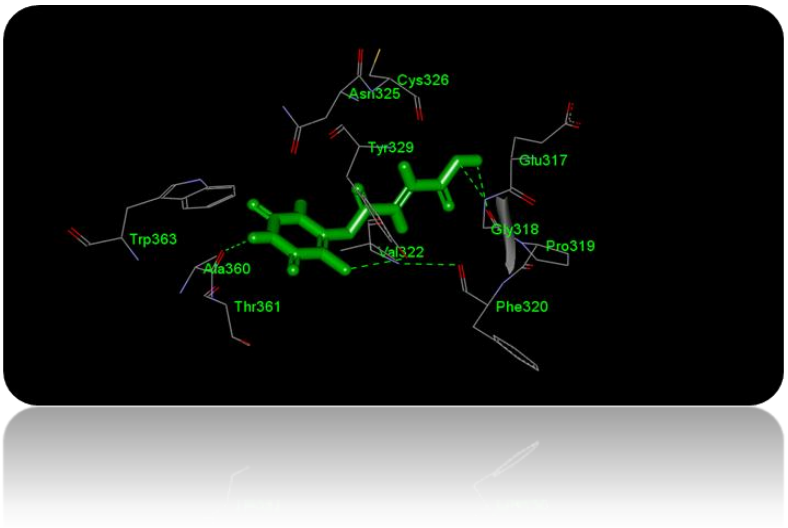
The above image shows the output page of the SWISS ADMET server. The above values clearly show that the designed drug compound 2 (5-Fluorouracil + Fumaric Acid + Sodium) obeys the standard ADMET profiles.

Molecular Drug Docking
Fig: 20. 3D Structure of Drug: H- Bond interactions



The above picture shows the existing drug (5-Fluorouracil) represented in green colour with drug binding cavities with respective amino acids of MACC1 protein.

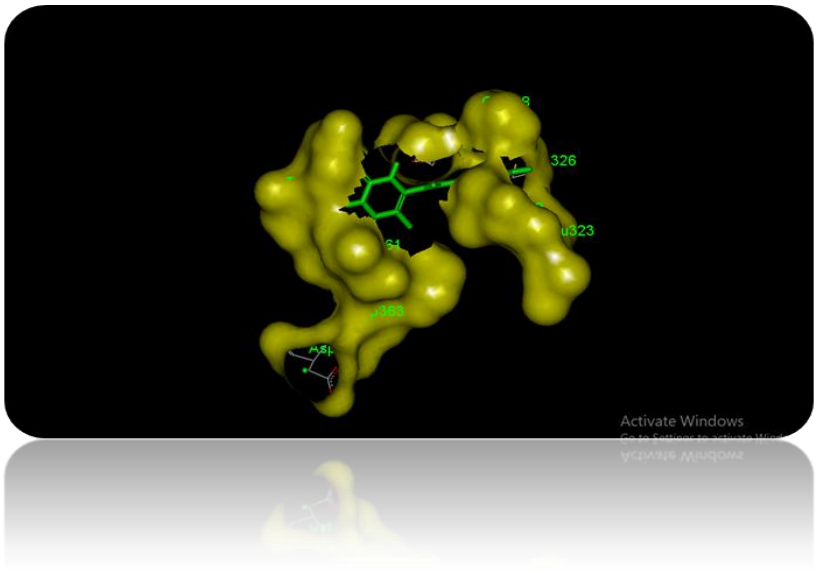
Molecular Drug Docking
Fig : 23 3D Structure of Drug : H- Bond interactions



The above picture shows the binding of the *de novo* drug 1 (5-Fluorouracil + Fumaric Acid), represented in green colour with drug binding cavities with respective amino acids of MACC1 protein.

Molecular Drug Docking

Fig.27: 3D Structure of Drug : H- Bond interactions



The above picture shows the binding of the *de novo* drug 2 (5-Fluorouracil + Fumaric Acid + Sodium) with the drug binding cavities with respective amino acids of MACC1 protein in surface model view.

Molecular Drug Docking

A new and highly efficient algorithm was identified for docking two molecules [10] [11]. While here, the results obtained by applying the algorithm to the docking of two protein molecules were demonstrated, the algorithm can be applied to receptor–drug cases as well. In molecular docking studies, the designed compound, that is, 5-fluorouracil combined with fumaric acid and 5-fluorouracil combined with Fumaric Acid + sodium ion (*de novo* drug 1 and *de novo* drug 2 respectively) was introduced to MACC1 protein. In this study, they observed the H-bond interacting amino acids with binding affinity scores. They simultaneously docked the existing drug, 5-fluorouracil with MACC1 protein. Table 2 shows the drug-docking result scores. Fig.18, 19 and 20 show the Patchdock result scores of MACC1 protein with 5-fluorouracil (existing drug) and the 3D structure of H-bond interactions with the respective drug-binding amino acid pockets (SER :501,LYS:528,GLY:582,GLY:584,GLU:526). Similarly, Fig.21, 22 and 23 show the Patchdock result scores of MACC1 protein with 5-fluorouracil + Fumaric acid (*de novo* drug 1) and the 3D structure of H-bond interactions with the respective drug-binding amino acid pockets (GLY:318,VAL:322,PHE:326,TYP:329,THR:361,ALA:359,TRP:363,ASP:359,GLY:318,ALA:360). Fig. 24,25 and 26 show the Patchdock result scores of MACC1 protein with 5-fluorouracil + sodium ion (*de novo* drug 2) and the 3D structure of H-bond interactions (ASN:325,LEU:323,TYR:329,TYR:345,TYR:361,ALA:359) with the respective drug-binding amino acid pockets.

The atomic contact energy between the *de novo* drug 1 with MACC1 protein shows -140.66. Similarly, that for the *de novo* drug 2 with MACC1 protein shows -146.18. Whereas, the atomic contact energy for the existing drug molecule, 5-flurorouracil with MACC1 protein is -90.31. Hence, interestingly, in the study the designed drug candidate is well bound with MACC1 protein. According to theory, negative value indicates more binding affinities between the target and the drug. Here, it was proved that our designed drug molecule is an efficient drug candidate for MACC1 protein which is responsible for colon cancer. Ultimately, the designed *de novo* drug candidate 1 has potential inhibitory properties in reducing the cancer cell growth, whereas, the designed *de novo* candidate 2 helps in both reducing the inhibitory and anti-inflammatory properties.

Conclusion

In conclusion, based on the *Insilico* results, it was reported that the designed novel drug candidates would act as efficient molecules for treating colon cancer. Complete pharmacokinetics profiling studies were carried out for the designed novel drug molecules. Overall, the drug docking results clearly explain the potential binding affinities between MACC1 protein and the designed candidates. This molecule would increase the inhibitory effect and reduce the inflammatory properties and hence reduce the growth of the cancer cells in colon cancer.

REFERENCES

[1] *Cunningham D, Atkin W, Lenz HJ, Lynch HT, Minsky B, Nordlinger B, Starling N (March 2010). "Colorectal cancer". Lancet. 375 (9719): 1030–1047.*

[2] *Vogelstein B, Papadopoulos N, Velculescu VE, Zhou S, Diaz LA, Kinzler KW (March 2013). "Cancer genome landscapes". Science. 339 (6127): 1546–1558.*

[3] *Yasunaga M. Ipsaro J.J. Mondragon A. Structurally similar but functionally diverse ZU5 domains in human erythrocyte ankyrin, J. Mol. Biol. 417:336-350(2012).*

[4] *Wang C. Yu C. Ye F. Wei Z. Zhang M, Structure of the ZU5-ZU5-UPA-DD tandem of ankyrin-B reveals interaction surfaces necessary for ankyrin function. Proc. Natl. Acad. Sci. U.S.A. 109:4822-4827(2012).*

[5] *Musacchio A. Gibson T. Lehto V.P. Saraste M, SH3--an abundant protein domain in search of a function. FEBS Lett. 307:55-61(1992).*

[6] *Pawson T. Schlessingert J. SH2 and SH3 domains Curr. Biol. 3:434-442(1993).*

[7] *Pawson T, Protein modules and signalling networks. Nature 373:573-580(1995).*

[8] *Nielsen M., Lundegaard C., Lund O., Petersen . CPHmodels-3.0 - Remote homology modeling using structure guided sequence profiles TN. Nucleic Acids Research, 2010, Vol. 38.*

[9] *O. Lund, M. Nielsen, C. Lundegaard, P. WornCPHmodels 2.0: X3M a Computer Program to Extract 3D Models. Abstract at the CASP5 conference A102, 2002.*

[10] *Duhovny D, Nussinov R, Wolfson HJ. Efficient Unbound Docking of Rigid Molecules. In Gusfield et al., Ed. Proceedings of the 2'nd Workshop on Algorithms in Bioinformatics(WABI) Rome, Italy, Lecture Notes in Computer Science 2452, pp. 185-200, Springer Verlag, 2002 .*

[11] *Schneidman-Duhovny D, Inbar Y, Nussinov R, Wolfson HJ. PatchDock and SymmDock: servers for rigid and symmetric docking. Nucl. Acids. Res. 33: W363-367, 2005.*