

ISSN: 0976-3031

Available Online at <http://www.recentscientific.com>

CODEN: IJRSFP (USA)

International Journal of Recent Scientific Research  
Vol. 15, Issue, 06, pp.4814-4820, June, 2024

International Journal of  
Recent Scientific  
Research

DOI: 10.24327/IJRSR

## RESEARCH ARTICLE

# IN SILICO MUTAGENESIS FOR IMPROVING ACTIVITY OF CYTOSINE DEAMINASE OF BIFIDOBACTERIUM BREVE: APPROACH TOWARDS EFFECTIVE PRO-DRUG GENE THERAPY IN COLORECTAL CANCER

Sanchita Dolui and Dr. Debdoot Gupta

Faculty, Department of Microbiology, The University of Burdwan. Burdwan-713104, West-Bengal

DOI: <http://dx.doi.org/10.24327/ijrsr.20241506.0904>

### ARTICLE INFO

#### Article History:

Received 18<sup>th</sup> May, 2024

Received in revised form 26<sup>th</sup> May, 2024

Accepted 16<sup>th</sup> June, 2024

Published online 28<sup>th</sup> June, 2024

#### Keywords:

Cancer, *Bifidobacterium breve*, In silico mutagenesis.

### ABSTRACT

Solitary Colorectal cancer (CRC) is the third most type of cancer that leading causes of morbidity and mortality worldwide. Among the recent therapeutic strategies the application of probiotics for treating cancer cell is most promising one. *Bifidobacterium breve* is a probiotic that can be applied in treating tumor cell (for its anaerobic nature) with least adverse effect. Prodrug/enzyme therapy is one of its mechanisms of therapeutic potentiality which contains 5FC and cytosine deaminase (CD). Generally, *Bifidobacterium* with transgene of *E coli* CD is recent choice for therapy which can be avoided by increasing its own CD activity through in silico mutagenesis. In the present work, attempts were made to predict 3D structure of CD and increasing its activity using in-silico mutagenesis. The physicochemical parameters of CD of *Bifidobacterium breve* found through PROTPARAM tool show that it is thermally stable and hydrophilic. After prediction of its structure through SWISS-MODEL and verification through SAVES it was found to be most reliable and stable. PrankWeb predicts 21 active sites and searching for favorable and stabilizing mutations using CUPSAT for position 151 and 89 showed various values of  $\Delta\Delta G$  in kcal/mol among which the one with highest value was selected. Mutant proteins were modeled accordingly and each of the modeled protein was docked with 5FC and the binding energy was checked. Mutants L89R and S151I had binding energy of -5.2 kcal/mol which is less than that of wild type CD (-4.6 kcal/mol) showing better interaction with 5FC. This in silico strategy can be validated in wet lab for improving the therapeutic potentiality of *Bifidobacterium breve*.

Copyright© The author(s) 2024, This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution and reproduction in any medium, provided the original work is properly cited.

### INTRODUCTION

Among non-communicable disease, Colon cancer is one of the leading causes of morbidity and mortality worldwide (1). Colon cancer remains a major cause of death worldwide with predictions that by 2030, the number of cancer-related deaths will reach nearly 20 million (2). Cancer cells are mutated self cells that escape from normal growth regulatory mechanisms (1). The techniques available to treat cancer are chemotherapy, surgery and radiation, chemotherapy is being the main choice of treatment (1). As healthy cells are damaged during chemotherapy and cancer cells become drug resistant, it is very difficult to treat cancer in a way that does not cause further damage (3). Therefore there is an urgent need for specific targeted therapies of cancer cells that may allow cancer treatment or be used as adjuvants to reduce conventional therapeutic doses of anticancer drugs (1). Suicide gene therapy is a therapeutic approach, also known as gene directed enzyme prodrug therapy, in which therapeutic transgenes have the ability to convert non toxic prodrug that penetrates the tumor into a cytotoxic drug or can express a toxic gene expression

product. There are systems that have been extensively investigated:- Cytosine deaminase (CD) gene of *E.coli* which converts the pro drug 5-fluorocytosine (5-FC) to 5-fluorouracil (5-FU) (4). Several studies have shown that microbes are capable of curing the deadly disease. Several studies have proven that the probiotics have anticancer effect (5). Probiotics are the beneficial bacteria and have fewer side effects in human health. The most commonly used probiotics are *Lactobacillus* sp and *Bifidobacterium* sp (5). Recently it has been shown that the tumor microenvironment is hypoxic and anaerobic bacteria are found to be grown in hypoxic condition; so these bacteria may be used as delivery agent for drug application. For example, *Clostridium* sp and *Bifidobacterium* sp can selectively germinate and grow in hypoxic areas of solid tumors after intravenous injection (6). *Bifidobacterium breve* are non pathogenic and anaerobic, Gram (+) bacteria within the normal bacterial flora found in the lower small intestinal and large intestine (6). Since *Bifidobacterium breve* is a probiotic, its application for gene/prodrug therapy does not cause host side effects. Cytosine deaminase (CD)/5FC is one type of gene/prodrug therapy that can be performed through

\*Corresponding author: Dr. Debdoot Gupta

Faculty, Department of Microbiology, The University of Burdwan. Burdwan-713104, West-Bengal

*Bifidobacterium breve*. (7). For gaining details of the activity of cytosine deaminase (CD) in this context, its structure should be explored because there is no experimentally derived structure stored in PDB database. Though *Bifidobacterium breve* has its own Cytosine deaminase gene, various experiments are conducted by taking transformed *Bifidobacterium breve* which has been transfected with the cytosine deaminase (CD) gene of *E.coli* by using a suitable expression plasmid (8). This may be due to the less active cytosine deaminase (CD) gene of *Bifidobacterium breve*. If the activity of cytosine deaminase (CD) can be improved through mutational analysis, the need of transfecting with *E.coli* cytosine deaminase gene will be avoided. Mutagenesis through wet lab method is a time consuming as well as laborious method, so in silico mutational analysis may be a good choice for improving the activity of the enzyme. The aim of this project is to perform in silico mutagenesis of cytosine deaminase (CD) of *Bifidobacterium breve* and examine its reflection in the cDNA level for improving enzyme activity that will more efficiently bind with 5FC to produce toxic drug from non-toxic prodrug which may direct towards better anticancer activity of this probiotic.

## MATERIALS AND METHODS

### Retrieval of the target protein Sequence and its concerned substrate

The amino acid sequence of the target protein molecule, Cytosine deaminase (CD) of *Bifidobacterium breve* was retrieved from the National Center for Biotechnology Information database (NCBI) (<http://www.ncbi.nlm.nih.gov>) (9). The structure of the substrate of CD, 5-fluorocytosine was derived from the PubChem (<https://pubchem.ncbi.nlm.nih.gov>) (10).

### Orthologous analysis

Checking the presence of similar sequences of proteins in other organisms was done through BlastP v.2.2.17 (<https://blast.ncbi.nlm.nih.gov/Blast.cgi?PAGE=Proteins>) (11).

### Multiple sequence alignment

Clustal Omega v.1.2.4 program was used for multiple sequence alignment of the protein (<https://www.ebi.ac.uk/Tools/msa/clustalo/>) (12) and for making the phylogenetic tree.

### Determine the physicochemical properties of Protein

The ProtParam server(<https://web.expasy.org/protparam/>) (13) was used for the determination of the physicochemical characteristics of the examined sequence including the molecular weight (MW), the theoretical isoelectric point (pI), the extinction coefficient, the total number of positive and negative residues, the instability index, the aliphatic index, the amino acid composition (%), N terminal of the sequence and the Grand Average hydropathy (GRAVY).

### The Secondary Structure prediction

The protein's secondary structure was predicted by NetSurfP v.2.0 (<https://services.healthtech.dtu.dk/services/NetSurfP-2.0/>) (14).

### Protein Structure prediction and verification

The protein structure was predicted by SWISS-MODEL (<https://swissmodel.expasy.org/>) (15) and was verified by SAVES v.6.0 (<https://saves.mbi.ucla.edu/>) (16-20).

## Active site prediction of protein

Active sites of the protein were predicted by PrankWeb v.2.4 (<https://prankweb.cz/>) (21).

## Prediction of favorable and stabilizing mutations

CUPSAT (Cologne University Protein Stability Analysis Tool) was used for the prediction of protein stability changes upon point mutations (single amino acid substitution) (<https://cupsat.brenda-enzymes.org/>) (22). The most stable and favorable mutations obtained by CUPSAT for active site amino acid positions is introduced into wild type protein sequence manually and the mutated variants were modeled using SWISS-MODEL.

## The Molecular Docking analysis

he binding energy of protein and ligand was calculated using PyRx v.1.0 (<https://pyrx.sourceforge.io/>) (23).

## RESULT AND DISCUSSION

### Preliminary analysis taking the sequence of protein sequence and ligand structure

In silico analysis of proteins provide a means of rapidly analyzing the sequence and structural aspects of growing number of proteins in post-genomic era [24, 25]. Here the protein sequence with length 453 was downloaded for further in silico characterization. Searching for its homologue in different other species shows its presence in some other species (*Clostridia bacterium*, *Lachnospiraceae bacterium*, *Coprococcus eutactus*, *Coprococcus aceti*) when the cut-off percentage was fixed to 60% in Fig.1. MSA by taking those sequences produces a phylogenetic tree which was shown in Fig.2. In phylogenetic tree, the closest neighbor of CD of *Bifidobacterium breve* is *uncultured bacterium* and second closest neighbors are *Clostridia bacterium*, *Mediterraneibacter faecis*.

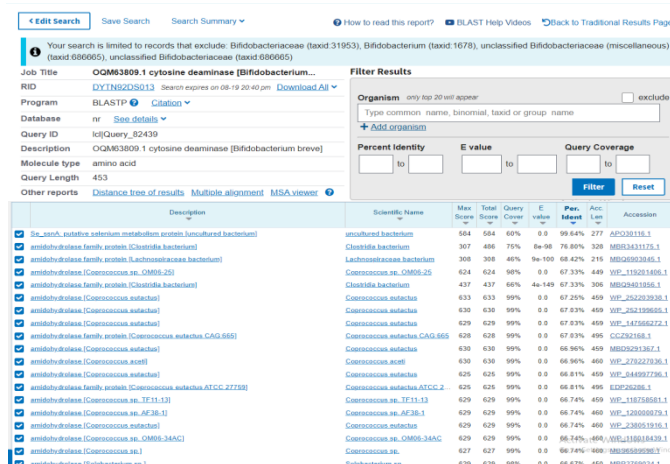


Fig.1 BLASTP Output

### Analysis of physicochemical properties

The physicochemical properties of the protein found through PROTPARAM tool was tabulated in the Table 1. Molecular weight was critical property of all macromolecule. The cytosine deaminase (CD) has 49857.24Da that's mean this molecular weight can move through the SDS page. The theoretical pI of cytosine deaminase (CD) was 4.78 that means at which the net charge of cytosine deaminase (CD) is zero. It was used for separating various protein molecules.

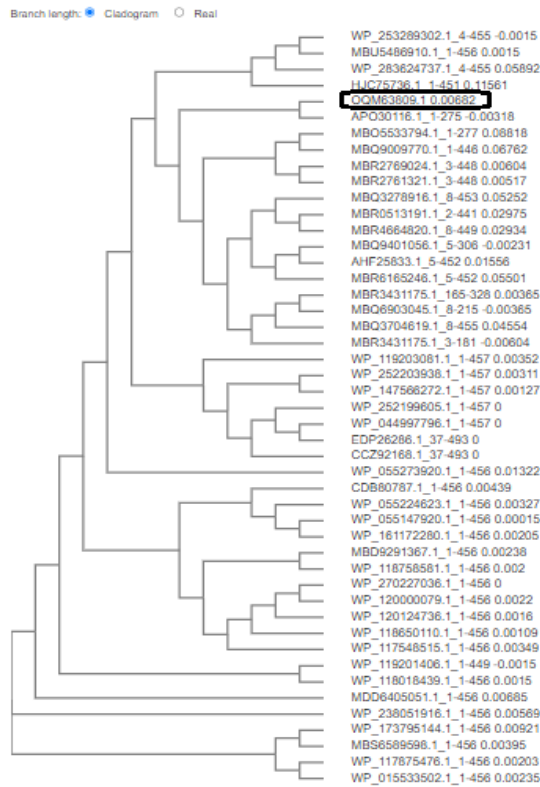


Fig.2 Phylogenetic Tree

Table 1 Different physicochemical properties of Cytosine deaminase (CD) of *Bifidobacterium breve*

Parameters	Value
Number of amino acids	453
Molecular weight	49857.24
Theoretical pI	4.78
Half-life	30 hours (mammalian reticulocytes, in vitro). >20 hours (yeast, in vivo). >10 hours (Escherichia coli, in vivo).
Instability index	30.28
Aliphatic index	77.77
Grand average of hydropath (GRAVY)	-0.237

Based on the results obtained, the cytosine deaminase (CD) with instability index being equal to 30.28 is a stable protein, more than 10 hours in the half-life (26). On the other hand, the aliphatic and hydrophobic indexes of the enzyme confirm its thermal stability and water solubility. The Aliphatic Index (AI) which has been defined as the relative volume of a protein occupied by the aliphatic side chains has been regarded as a positive factor for the increase in the thermal stability of the globular proteins (26). So, high AI of the cytosine deaminase (CD) which is equal to 77.77 indicates these proteins are thermally stable as well as they contain high amount of hydrophobic amino acids. GRAVY of cytosine deaminase (CD) is -0.237 that's mean it was hydrophilic. Moreover, the very low Grand Average hydropath (GRAVY) index of this enzyme refers to its simple interaction with water (27).

### Secondary structure prediction of the protein

The secondary structure of the protein was found to have combination of helix, strand and coil (Fig.3).

Below is a graphical representation of 453 residue predictions across 1 sequence. Running time was 105 seconds (105 seconds per sequence). Hover your mouse over a sequence position to see all outputs.

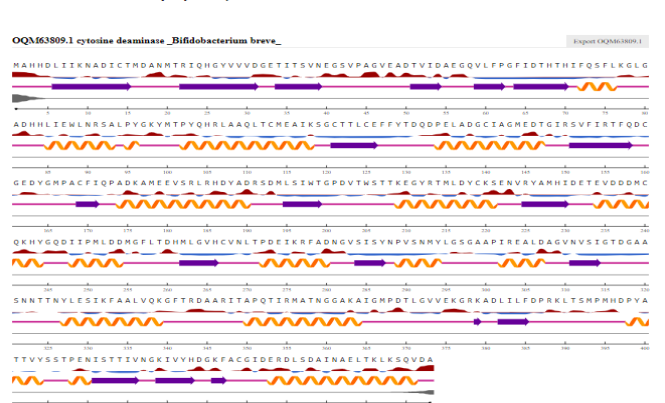


Fig.3 Secondary structure

### Protein structure prediction and verification

The structure of cytosine deaminase (CD) was not present in PDB which is needed for further in silico analysis. So, the 3D modeling was done in the SWISS MODEL server that created three models and the best model was selected. The SWISS-MODEL web server this enzyme based on similar structure with AlphaFold DB model of A0A413C6B sequence Identity 80.31%.



Fig.4 3D molecular structure of Cytosine deaminase (CD)

The modeled structure was visualized through Pymol server shown in Fig.4: The structure of cytosine deaminase (CD) was put in the SAVES server for verifying the structure. It was a validated protein as verified by SAVES software and shown in Fig.5.

ERRAT value of cytosine deaminase (CD) is 93.7079 means that cytosine deaminase (CD) is a high quality protein. ERRAT plot shows error values for residues. The Y-axis represents the error value and the X-axis represents the amino acid residues of the protein model. The ERRAT plot was shown in Fig.6.

VERIFY\_3D assesses the compatibility of the predicted 3D model based on the local environment of each amino acid (17). This program produces an averaged 3D-1D score for each residue, and the structure with more than 80% of residues with a score over 0.1 is considered reliable. Cytosine deaminase (CD) has 82.78% of the residues and averaged 3D-1D score

was 0.1. So, I can conclude that the cytosine deaminase (CD) is reliable. VERIFY\_3D output was shown in Fig.7.

In WHATCHECK, Green colour denotes highly reliable validated, yellow colour denotes not much reliable/mildly issued, Red colour denotes not validated. So, I can conclude that the cytosine deaminase (CD) is shown some of highly reliable regions, a short number of regions shows much reliable and low number regions shows not validated<sup>3</sup>. WHATCHECK output was shown in Fig.8.

The PROCHECK tool requires modeled protein file as an input and generates the Ramachandran plot. The Ramachandran plot predicts the structural stereochemical property. In Ramachandran plot, 90.7% amino acids were found in most favoured regions that make this protein stable. PROCHECK output and Ramachandran plot was shown in Fig.9 and Fig.10. The structure was pass by the ERRAT, VERIFY3D, WHATCHECK and PROCHECK.

**UCLA-DOE LAB — SAVES v6.0**



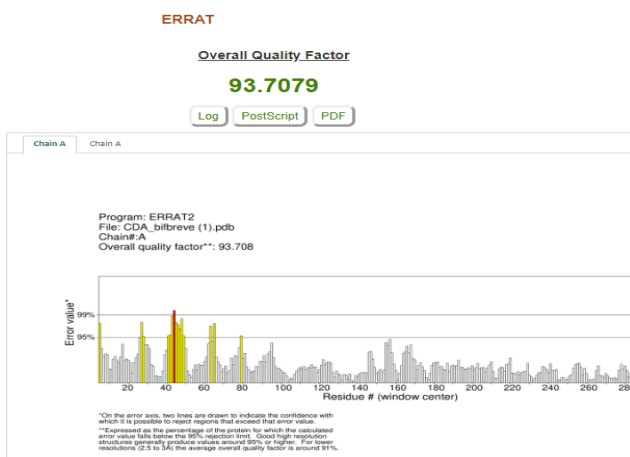
Job 1367778 has been created

New Job

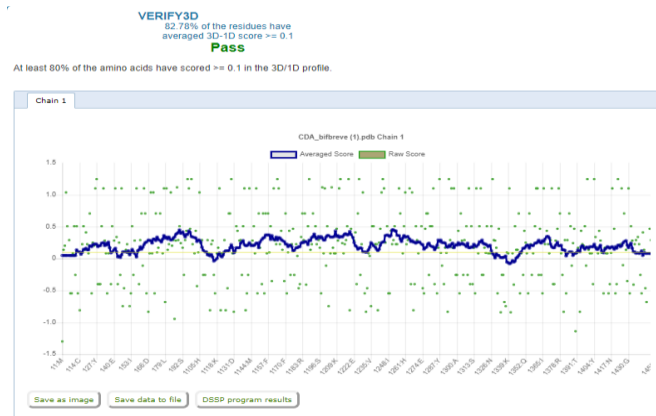
job #1367778: CDA\_bifbreve (1).pdb [job link] [3D Viewer]

<p><b>ERRAT</b></p> <p>Analyzes the statistics of non-bonded interactions between different atom types and plots the value of the error function versus position of a 9-residue sliding window, calculated by a comparison with statistics from highly refined structures.</p> <p>Start</p>	<p><b>Verify3D</b></p> <p>Determines the compatibility of an atomic model (3D) with its own amino acid sequence (1D) by assigning a structural class based on its location and environment (alpha, beta, loop, polar, nonpolar etc) and comparing the results to good structures.</p> <p>Start</p>	<p><b>PROVE</b></p> <p>Temporarily down at the moment</p>
<p><b>WHATCHECK</b></p> <p>Derived from a subset of protein verification tools from the WHATIF program (Wend, 1990), this does extensive checking of many stereochemical parameters of the residues in the model.</p> <p>Start</p>	<p><b>PROCHECK</b></p> <p>Checks the stereochemical quality of a protein structure by analyzing residue-by-residue geometry and overall structure geometry.</p> <p>Start</p>	<p><b>OPEN</b></p> <p>We are open to suggestions for a 6th program to operate in this window. If you know of a program that we could run locally on our server that would be most useful, please let us know; email holton at mbi dot ucla dot edu with your suggestion</p>

**Fig.5 SAVES output.**



**Fig.6 ERRAT output.**



**Fig.7 VERIFY3D Output**

Headings

**WHATCHECK**

- #1 Error: Missing unit cell information
- #2 Error: Missing symmetry information
- #3 Note: No rounded coordinates detected
- #4 Note: Valine nomenclature OK
- #5 Note: Threonine nomenclature OK
- #6 Note: Isoleucine nomenclature OK
- #7 Note: Leucine nomenclature OK
- #8 Warning: Arginine nomenclature problem
- #9 Warning: Tyrosine conversion problem
- #10 Warning: Phenylalanine conversion problem
- #11 Warning: Aspartic acid convention problem
- #12 Warning: Glutamic acid convention problem
- #13 Note: L-α-D conformation checked OK
- #14 Note: Chain names are unique
- #15 Note: All bond lengths OK
- #16 Note: No missing atoms detected
- #17 Note: All bond lengths OK
- #18 Note: Normal bond length variability
- #19 Warning: Directionality in bond lengths
- #20 Warning: Unusual bond angles
- #21 Warning: Normal bond angle variability
- #22 Error: Side chain planarity problems
- #23 Note: Atoms connected to aromatic rings OK
- #24 Note: PRO puckering amplitude OK
- #25 Warning: Strange PRO puckering phases
- #26 Warning: Torsion angle evaluation shows unusual conformations
- #27 Note: Ramachandran Z-score OK
- #28 Note: Omega angle restraint OK
- #29 Note: chi-1/chi-2 angle correlation Z-score OK
- #30 Note: Ramachandran plot
- #31 Note: Secondary structure
- #32 Error: Abnormally short interatomic distances
- #33 Warning: Abnormal packing environment for some residues
- #34 Note: No series of residues with bad packing environment
- #35 Note: Structural average packing environment OK
- #36 Note: Backbone oxygen evaluation OK
- #37 Note: Rotamers checked OK
- #38 Warning: Unusual backbone conformations
- #39 Error: Backbone conformation Z-score very low
- #40 Note: Symmetry related water molecules check not performed
- #41 Error: Hydrophobic E-factor error
- #42 Error: HIS\_ASN\_OLN side chain flips
- #43 Note: Histidine type assignments
- #44 Warning: Buried unsatisfied hydrogen bond donors
- #45 Warning: Buried unsatisfied hydrogen bond acceptors
- #46 Note: Overall summary report

**Fig.8 WHATCHECK output**

Out of 9 evaluations

- Errors: 3
- Warning: 4
- Pass: 2

The evaluations are the '\*' (Warning) and '\*\*' (Error) in the summary. The categories on the left do not always correspond in number due to PROCHECK output documents.

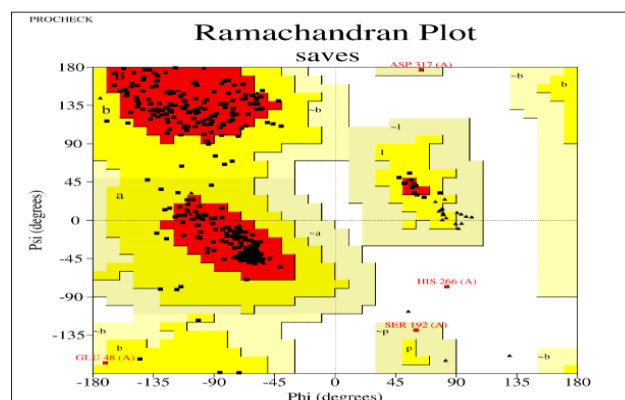
Summary	
Ramachandran plot	Warning
All Ramachandrans	Warning
Chi1-chi2 plots	Pass
Main-chain params	
Side-chain params	Error
Residue properties	Warning
Bond len/angle	Warning
M/c bond lengths	
M/c bond angles	
Planar groups	Warning
Program output	

Summary file

```

----- P R O C H E C K S U M M A R Y -----
/var/www/SAVES/jobs/1367778/saves.pdb 1.5 453 residues
Ramachandran plot: 98.7% core 0.3% allow 0.6% gener 0.3% disall
All Ramachandrans: 13 labelled residues (out of 453)
Chi1-chi2 plots: 4 labelled residues (out of 263)
Side-chain params: 5 better 9 worse
Residue properties: Max.deviation: 5.7 Morris et al class: 3 3 2
3 c1a-peptides bond len/angle: 6.1
G-factors: chi2: 0.11 covar1: 0.40 Overall: 0.45
Planar groups: 92.6% within limits 7.4% highlighted 3 off graph
* May be worth investigating further. ** Worth investigating further.
    
```

**Fig.9 PROCHECK output**



Plot statistics		
Residues in most favoured regions [A,B,L]	361	90.7%
Residues in additional allowed regions [a,b,l,p]	33	8.3%
Residues in generously allowed regions [-a,-b,-l,-p]	3	0.8%
Residues in disallowed regions	1	0.3%
-----		
Number of non-glycine and non-proline residues	398	100.0%
Number of end-residues (excl. Gly and Pro)	2	
Number of glycine residues (shown as triangles)	35	
Number of proline residues	18	
-----		
Total number of residues	453	

Based on an analysis of 118 structures of resolution of at least 2.0 Angstroms and R-factor no greater than 20%, a good quality model would be expected to have over 90% in the most favoured regions.

Fig.10 Ramachandran Plot

### Active sites prediction

Amino acids in the active site of a protein were important for binding to the substrates. The protein-substrate binding was responsible for the stability of a protein. Active sites of a protein were determined by PrankWeb and 21 active sites were found. The obtained result was shown in Table-2.

Table 2 Active sites of Cytosine deaminase (CD)

Residue No.	Amino acids
70	Histidine
73	Glutamine
85	Leucine
86	Isoleucine
88	Tryptophan
89	Leucine
94	Leucine
129	Tryptophan
151	Serine
164	Tyrptophan
166	Methionine
202	Aspartic acid
203	Valine
205	Tyrptophan
229	Histidine
232	Glutamic acid
236	Aspartic acid
266	Histidine
295	Leucine
317	Aspartic acid
321	Serine

### Prediction of favorable and stabilizing mutations

CUPSAT server was used for performing mutagenesis of a protein. CUPSAT mutational analysis results provide the amino acid residues with the favorable and stabilizing mutation with predicted  $\Delta\Delta G$  (kcal/mol). It contains information about mutation site, its structural features and comprehensive information about changes in protein stability for 19 possible substitutions for a specific amino acid position. A total of 19 stable and favorable mutations were found for cytosine deaminase (CD) of *Bifidobacterium breve* which could be used for improving activity of this protein (Table 3); two positions (H70, M166) were there where no stable substitutions were found. The favorable and stabilizing mutation with highest  $\Delta\Delta G$  (kcal/mol) among all possible mutations were selected shown in Fig.11. All mutated structures were modeled through SWISS-MODEL.

Table 3 CUPSAT output

Position of mutated residue	Wild type residue	Mutated residue	$\Delta\Delta G$ (kcal/mol)
70	Histidine	-	-
73	Glutamine	Tryptophan	8.57
85	Leucine	Phenylalane ne	0.38
86	Isoleucine	Arginine	0.11
88	Tryptophan	Phenylalane ne	0.51
89	Leucine	Arginine	0.54
94	Leucine	Isoleucine	0.72
129	Tryptophan	Threonine	0.49
151	Serine	Isoleucine	6.95
164	Tyrptophan	Proline	5.94
166	Methionine	-	-
202	Aspartic acid	Lysine	14.08
203	Valine	Histidine	0.1
205	Tyrptophan	Aspartic acid	0.26
229	Histidine	Cysteine	9.45
232	Glutamic acid	Leucine	2.91
236	Aspartic acid	Isoleucine	2.77
266	Histidine	Glycine	5.12
295	Leucine	Lysine	1.08
317	Aspartic acid	Methionine	3.46
321	Serine	Aspartic acid	3.75

Mutation Site	Structural Features
Protein: CDBB.PDB	SS element: Sheet
Chain: A	Solvent accessibility: 3.41%
Wild type AA: SER	Torsion angles: -138.8°, 142.3°
Residue ID: 151	

Amino acid	Overall Stability	Torsion*	Predicted $\Delta\Delta G$ (kcal/mol)
GLY	Stabilising	Unfavourable	0.32
ALA	Stabilising	Unfavourable	0.53
VAL	Stabilising	Favourable	6.73
LEU	Stabilising	Unfavourable	6.74
ILE	Stabilising	Favourable	6.96
MET	Stabilising	Favourable	1.07
PRO	Destabilising	Unfavourable	-4.54
TRP	Stabilising	Favourable	3.52
THR	Destabilising	Favourable	-0.07
PHE	Stabilising	Favourable	5.03
GLN	Destabilising	Favourable	-4.21
LYS	Destabilising	Unfavourable	-4.79
TYR	Stabilising	Favourable	5.12
ASN	Destabilising	Unfavourable	-1.15
CYS	Stabilising	Unfavourable	10.14
GLU	Destabilising	Unfavourable	-1.37
ASP	Destabilising	Unfavourable	-6.3
ARG	Destabilising	Favourable	-0.13
HIS	Destabilising	Favourable	2.17

Fig.11 CUPSAT output

### Prediction of the mutated protein structure and its interaction with the ligand

After prediction of the mutated structure of cytosine deaminase (CD) of *Bifidobacterium breve* its interaction with substrate 5FC was compared with that of wild type CD of *Bifidobacterium breve* and CD of *E coli* in terms of binding energy. Binding energy of any receptor ligand complex shows the structural stability of this complex. Binding energy of the protein ligand complex before and after mutation was shown in Table 4. The table shows positions 89, 151 is most promising for in silico mutagenesis and the mutated protein has better interaction than the wild type protein and that of CD protein of *E coli*. Detailed docking analysis for the interaction study also shows that the mutated CD of *Bifidobacterium breve* has more

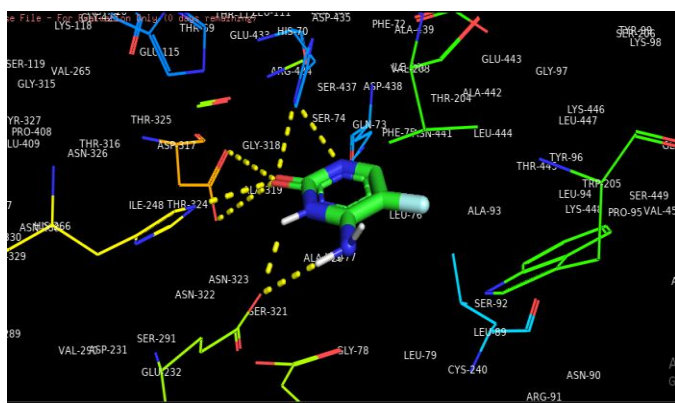
hydrogen bonds than that of wild type CD protein which reflects it better interaction with 5FC (Fig 12 and 14). The better interaction means the activity of CD of *Bifidobacterium breve* is enhanced.

**Table4** Docking result

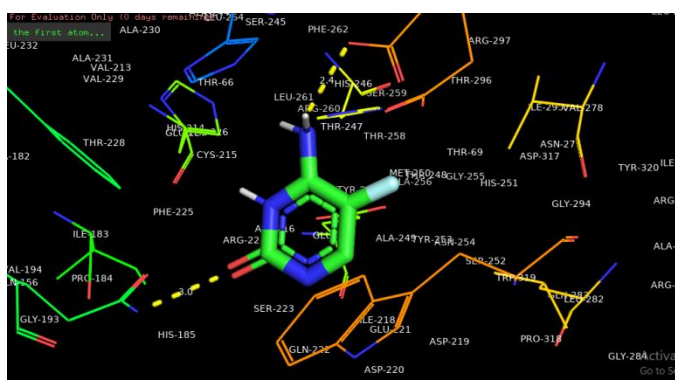
Substrate-ligand complex	Binding energy(Kcal/mol)
Wild type Cytosine deaminase (CD) of <i>Bifidobacterium breve</i> + 5FC	-4.6
Wild type Cytosine deaminase (CD) of <i>E.coli</i> + 5FC	-5.2
Mutation at 89,151 positions of the Cytosine deaminase (CD) of <i>Bifidobacterium breve</i> + 5FC	-5.2

### Conclusion and Future aspect

*Bifidobacterium breve* is a probiotic which is being used for treatment of colorectal cancer. As it is anaerobic in nature it can grow in the tumor microenvironment which is hypoxic in nature. If it is applied in cancer patient, it selectively enter into the tumor cells and initiate there the prodrug/enzyme therapy by producing cytosine deaminase protein which converts the prodrug 5FC to 5FU which ultimately kills the tumor cells (approach known as suicidal gene therapy).

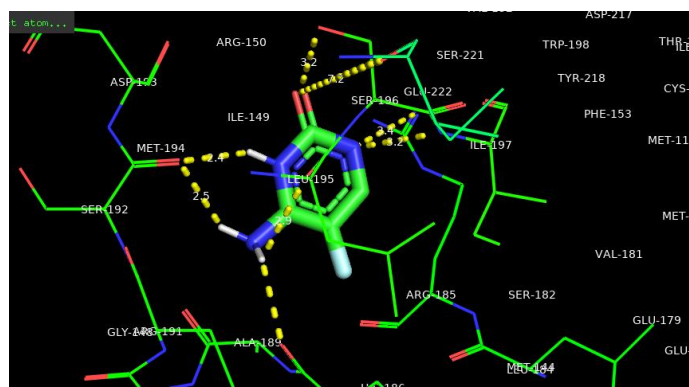


**Fig.12** Dock complex of wild type cytosine deaminase (CD) of *Bifidobacterium breve* with 5FC



**Fig.13** Cytosine deaminase (CD) of *E.coli*. with 5FC

Nowadays transfected *Bifidobacterium breve* with transformed CD gene is used for this purpose though *Bifidobacterium breve* has its own copy of CD gene. In this study in silico mutagenesis study on the hot spot region of the predicted structure of CD protein of *Bifidobacterium breve* was performed for improving the activity of this enzyme which can be checked by their docking score with 5FC.



**Fig.14** Dock complex of mutated Cytosine deaminase of *Bifidobacterium breve* with 5FC

It reveals that positions 89,151 of CD protein is a promising site for in silico mutagenesis and the mutated protein has better binding affinity with the substrate(-5.2 Kcal/mol). The better affinity means the activity of CD of *Bifidobacterium breve* is enhanced. So depending on the mutated protein sequence suitable cDNA construct can be made and used to transfect the bacteria for the production of desired mutated protein. In future this procedure can be verified through wet lab method which can be considered as a positive step towards future cancer treatment strategy.

### References

- Ferlay, J., Soerjomataram, I., Ervik, M., Dikshit, R., Eser, S., Mathers, C., & Bray, F. (2013). GLOBOCAN 2012 v1. 0, cancer incidence and mortality worldwide. *Iarc Cancerbase, 11*.
- Siegel RL, Miller KD, Fedewa SA, Ahnen DJ, Meester RG, Barzi A, Jemal A. Colorectal cancer statistics, 2017. *CA: a cancer journal for clinicians.* 2017 May 6;67(3):177-93.c
- Mills, H., Acquah, R., Tang, N., Cheung, L., Klenk, S., Glassen, R., & Van, T. N. (2022). The use of bacteria in cancer treatment: a review from the perspective of cellular microbiology. *Emergency Medicine International, 2022*.
- Zarogoulidis, P., Darwiche, K., Sakkas, A., Yarmus, L., Huang, H., Li, Q., & Malecki, M. (2013). Suicide gene therapy for cancer—current strategies. *Journal of genetic syndrome & gene therapy, 4*.
- Lee, J. E., Lee, J., Kim, J. H., Cho, N., Lee, S. H., Park, S. B., ... & Yoo, H. M. (2019). Characterization of the anti-cancer activity of the probiotic bacterium *Lactobacillus fermentum* using 2D vs. 3D culture in colorectal cancer cells. *Biomolecules, 9*(10), 557.
- Yazawa, K., Fujimori, M., Amano, J., Kano, Y., & Taniguchi, S. I. (2000). *Bifidobacterium longum* as a delivery system for cancer gene therapy: selective localization and growth in hypoxic tumors. *Cancer gene therapy, 7*(2), 269-274.
- Wei, H., Chen, L., Lian, G., Yang, J., Li, F., Zou, Y., & Yin, Y. (2018). Antitumor mechanisms of bifidobacteria. *Oncology Letters, 16*(1), 3-8.
- Nakamura, T., Sasaki, T., Fujimori, M., Yazawa, K., Kano, Y., Amano, J., & Taniguchi, S. I. (2002). Cloned cytosine deaminase gene expression of *Bifidobacterium longum* and application to enzyme/pro-drug therapy of hypoxic solid

- tumors. *Bioscience, biotechnology, and biochemistry*, 66(11), 2362-2366.
9. Schoch CL, Ciuffo S, Domrachev M, Hotton CL, Kannan S, Khovanskaya R, Leipe D, McVeigh R, O'Neill K, Robbertse B, Sharma S. NCBI Taxonomy: a comprehensive update on curation, resources and tools. *Database*. 2020; 2020:baaa062.
  10. Kim S, Chen J, Cheng T, Gindulyte A, He J, He S, Li Q, Shoemaker BA, Thiessen PA, Yu B, Zaslavsky L. PubChem 2023 update. *Nucleic acids research*. 2023 Jan 6; 51 (D1):D1373-80.
  11. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. *Journal of molecular biology*. 1990 Oct 5; 215 (3):403-10.
  12. Madeira F, Pearce M, Tivey AR, Basutkar P, Lee J, Edbali O, Madhusoodanan N, Kolesnikov A, Lopez R. Search and sequence analysis tools services from EMBL-EBI in 2022. *Nucleic acids research*. 2022 Jul 5;50 (W1):W276-9
  13. Gasteiger E, Hoogland C, Gattiker A, Duvaud SE, Wilkins MR, Appel RD, Bairoch A. Protein identification and analysis tools on the ExPASy server. Humana press; 2005.
  14. Klausen MS, Jespersen MC, Nielsen H, Jensen KK, Jurtz VI, Soenderby CK, Sommer MO, Winther O, Nielsen M, Petersen B, Marcatili P. NetSurfP-2.0: Improved prediction of protein structural features by integrated deep learning. *Proteins: Structure, Function, and Bioinformatics*. 2019 Jun; 87(6):520-7.
  15. Waterhouse A, Bertoni M, Bienert S, Studer G, Tauriello G, Gumienny R, Heer FT, de Beer TA, Rempfer C, Bordoli L, Lepore R. SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic acids research*. 2018 Jul 2; 46(W1):W296-303.
  16. Colovos C, Yeates TO. Verification of protein structures: patterns of nonbonded atomic interactions. *Protein science*. 1993 Sep; 2(9):1511-9.
  17. Bowie JU, Lüthy R, Eisenberg D. A method to identify protein sequences that fold into a known three-dimensional structure. *Science*. 1991 Jul 12; 253(5016):164-70.
  18. Pontius J, Richelle J, Wodak SJ. Deviations from standard atomic volumes as a quality measure for protein crystal structures. *Journal of molecular biology*. 1996 Nov 22; 264(1):121-36.
  19. MacArthur MW, Laskowski RA, Thornton JM. Knowledge-based validation of protein structure coordinates derived by X-ray crystallography and NMR spectroscopy. *Current Opinion in Structural Biology*. 1994 Oct 1; 4(5):731-7.
  20. Hooft RW, Vriend G, Sander C, Abola EE. Errors in protein structures. *Nature*. 1996 May 23; 381(6580):272-.
  21. Jendele L, Krivak R, Skoda P, Novotny M, Hoksza D. PrankWeb: a web server for ligand binding site prediction and visualization. *Nucleic acids research*. 2019 Jul 2; 47(W1):W345-9.
  22. Parthiban V, Gromiha MM, Schomburg D. CUPSAT: prediction of protein stability upon point mutations. *Nucleic acids research*. 2006 Jul 1; 34(suppl\_2):W239-42.
  23. Pawar RP, Rohane SH. Role of autodock vina in PyRx molecular docking.
  24. Michalovich D, Overington J, Fagan R. Protein sequence analysis in silico: application of structure-based bioinformatics to genomic initiatives. *Current Opinion in Pharmacology*. 2002 Oct 1; 2(5):574-80.
  25. Kanehisa M, Bork P. Bioinformatics in the post-sequence era. *Nature genetics*. 2003 Mar; 33(3):305-10.
  26. Ikai A. Thermostability and aliphatic index of globular proteins. *The Journal of Biochemistry*. 1980 Oct 1;88(6):1895-8.
  27. Kyte J, Doolittle RF. A simple method for displaying the hydropathic character of a protein. *Journal of molecular biology*. 1982 May 5; 157(1):105-32.
  28. Du X, Li Y, Xia YL, Ai SM, Liang J, Sang P, Ji XL, Liu SQ. Insights into protein-ligand interactions: mechanisms, models, and methods. *International journal of molecular sciences*. 2016 Jan 26; 17(2):144.

#### How to cite this article:

Sanchita Dolui and Debdoot Gupta. (2024). In silico mutagenesis for improving activity of cytosine deaminase of bifidobacterium breve: approach towards effective pro-drug gene therapy in colorectal cancer. *Int J Recent Sci Res*.15 (06), pp.4814-4820.

\*\*\*\*\*