PLOS ONE

RESEARCH ARTICLE

Atomic Insight into the Altered O^6-Methylguanine-DNA Methyltransferase Protein Architecture in Gastric Cancer

Naveed Anjum Chikan[1,2], Shoiab Bukhari[2], Nadeem Shabir[3], Asif Amin[2], Sheikh Shafi[4], Raies Ahmad Qadri[2], Trupti Navin Chandra Patel[1]*

1 Division of Medical Biotechnology, School of Bioscience and Technology, VIT University, Vellore, Tamil Nadu, 632014, India, **2** Departments of Biotechnology, University of Kashmir, Srinagar, Kashmir, 190006, India, **3** Department of Animal Biotechnology, College of Veterinary Sciences, Anand Agricultural University, Anand, Gujarat, India, 388 001, **4** Department of Clinical Biochemistry, Sher-i- Kashmir Institute of Medical Sciences, Srinagar, Kashmir, 190011, India

* tnpatel@vit.ac.in

Citation: Chikan NA, Bukhari S, Shabir N, Amin A, Shafi S, Qadri RA, et al. (2015) Atomic Insight into the Altered O^6-Methylguanine-DNA Methyltransferase Protein Architecture in Gastric Cancer. PLoS ONE 10(5): e0127741. doi:10.1371/journal.pone.0127741

Academic Editor: Reiner Albert Veitia, Institut Jacques Monod, FRANCE

Received: December 16, 2014

Accepted: April 19, 2015

Published: May 26, 2015

Data Availability Statement: All relevant data are within the paper and its Supporting Information files.

Funding: The research was funded by VIT University research associate grant and the funding agency had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

Abstract

O^6-methylguanine-DNA methyltransferase (MGMT) is one of the major DNA repair protein that counteracts the alkalyting agent-induced DNA damage by replacing O^6-methylguanine (mutagenic lesion) back to guanine, eventually suppressing the mismatch errors and double strand crosslinks. Exonic alterations in the form of nucleotide polymorphism may result in altered protein structure that in turn can lead to the loss of function. In the present study, we focused on the population feared for high exposure to alkylating agents owing to their typical and specialized dietary habits. To this end, gastric cancer patients pooled out from the population were selected for the mutational screening of a specific error prone region of MGMT gene. We found that nearly 40% of the studied neoplastic samples harbored missense mutation at codon^{151} resulting into Serine to Isoleucine variation. This variation resulted in bringing about the structural disorder, subsequently ensuing into a major stoichiometric variance in recognition domain, substrate binding and selectivity loop of the active site of the MGMT protein, as observed under virtual microscope of molecular dynamics simulation (MDS). The atomic insight into MGMT protein by computational approach showed a significant change in the intra molecular hydrogen bond pattern, thus leading to the observed structural anomalies. To further examine the mutational implications on regulatory plugs of MGMT that holds the protein in a DNA-Binding position, a MDS based analysis was carried out on, all known physically interacting amino acids essentially clustered into groups based on their position and function. The results generated by physical-functional clustering of protein indicated that the identified mutation in the vicinity of the active site of MGMT protein causes the local and global destabilization of a protein by either eliminating the stabilizing salt bridges in cluster C3, C4, and C5 or by locally destabilizing the "protein stabilizing hing" mapped on C3-C4 cluster, preceding the active site.

Introduction

Although declining, the malady of gastric cancer, according to GOLOBOCON 2012 is still the third leading cause of cancer deaths worldwide [1, 2]. In the pathogenesis of this disease, various genetic and molecular alterations take place leading to the malignant transformation of gastric mucosa [3]. This transformation is a multi step process that entails the abnormalities in important cellular functions such as DNA repair, adhesion, signal transduction, cell differentiation and others [4,5]. Alkylating carcinogens like N-Nitrosodimethylamine, Methyl Nitrosourea (NMU), N-methyl-N'-nitro-N-nitroguanidine etc. lead to formation of O^6-Methylguanine, a DNA adduct whose presence leads to induction of mutations (G:C—A:T transition) and results in development of cancer [6–10]. *MGMT* is the enzyme responsible for the repairing O^6-methylguanine adducts [11–13]. MGMT is a suicidal enzyme that removes a methyl group from the O^6-position in guanine and transfers it to its own cystine residue at codon 145 in the protein, thus inactivating itself while repairing guanine [14]. Under the exposure of NMU, MGMT-defective mice have been seen to develop cancer [15], while as transgenic mice carrying extra copies of the foreign MGMT gene were less prone to the disease [16]. The, genetic polymorphism of this enzyme has proven to be a potential risk factor for cancer [17–22]. This study thus focuses on mutational profiling of error prone region of Exon 5 of MGMT which encodes for the active site of the protein, viz active site surrounded by domains responsible for holding onto DNA [13]. The representative population of gastric cancer patients that has been selected for this study presents a unique cohort essentially being highly exposed to dietary alkylating agents [6, 23–28].

The use of *Insilico* techniques to understand the effect of polymorphism on protein structure and dynamics has been in practice and a plethora of work has been done in this regard [29–32]. The computer aided prediction methods using evolutionary and structure based prediction gives an insight into the damaging capability of the polymorphism [33]. The molecular dynamics can be used to observe the conformational changes the polymorphism can inflict in the protein. These conformational changes in the three dimensional structure of protein can affect the physiological affinities and various biochemical pathway interactions. To examine the effect of mutation at evolutionary as well as atomic level, Insilico predictions using different servers as well as MDS of the Wildtype (wt) and Mutant (Mu) MGMT protein was carried out. For MDS protein trajectories and atomic interaction analysis, gromacs inbuilt tools were used. Principle component analysis (PCA) was conducted to estimate the flexibility of both structures. Free energy landscapes (FEL) of native and Mu MGMT were also studied to comprehend the effect of mutation.

Results and Discussion

The Exon 5 segment of MGMT gene, was successfully amplified from all samples. Amplicons after sequencing showed a transversion mutation in codon 151AGC, the sequences of which have been submitted to GeneBank bearing accession numbers KM000795 and KM000796. The in silico tools to study the possible damaging effect of the mutation were selected meticulously, so as each factor is looked into and double checked by other tool which uses different algorithm. The details of the servers that are used in our study are described in S1 Table, where there algorithm, working and criteria for prediction is given. Selected server predicts the mutation to be damaging. The MDS simulation trajectories for 30ns run for wt and mutant protein were analyzed extensively using gromacs inbuilt tools. S1 Fig shows a nsSNP at codon 151 that leads to a missense mutation from Ser to Ile, otherwise in its wild-type form helps in Protein-DNA interactions [34–36]. As shown in Fig 1, wtMGMT (PDBID:1T39) SER 151, besides

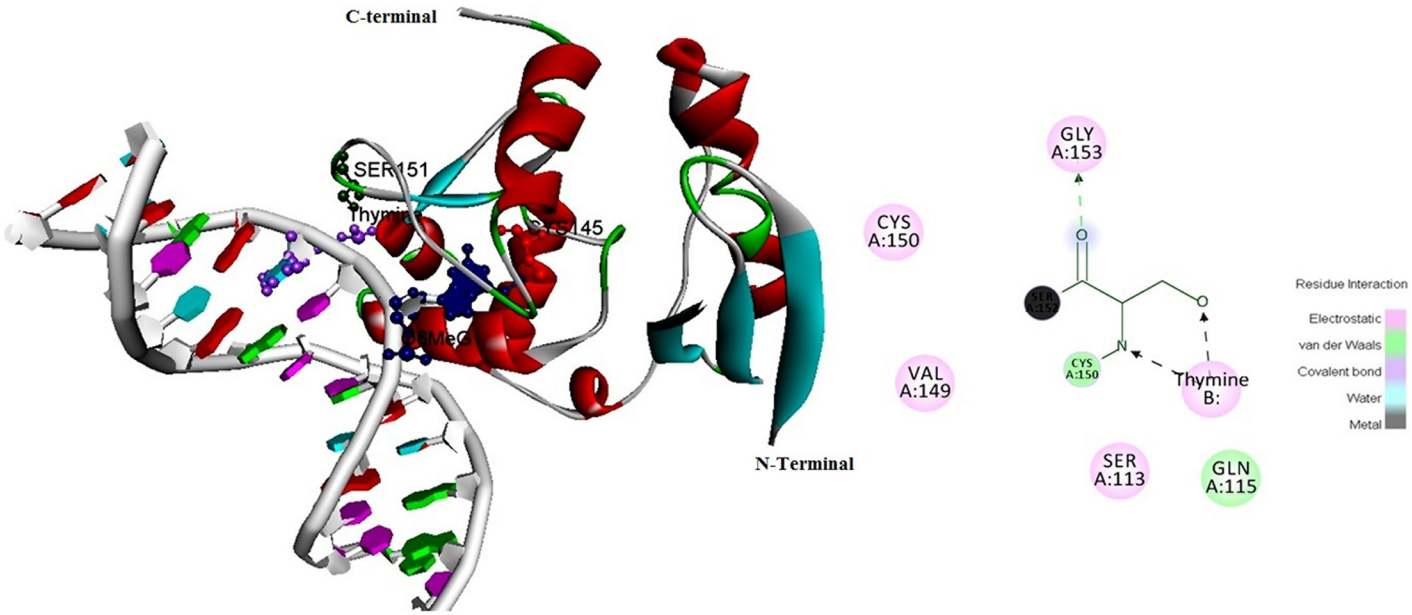


Fig 1. Representative picture of modeled wild type MGMT protein docked to minor groove of DNA depicts a stable interaction between Ser151 of enzyme with thymine base, attaining stability with the help of two hydrogen bonds shown in the figure as dotted arrows.

doi:10.1371/journal.pone.0127741.g001

making normal electrostatic interaction with thymine also formed two hydrogen bonds with it via amide nitrogen.

Fig 2 shows the snapshots of both wt and Mu structures at different time intervals, stipulating the synopsis of the effect of mutation on structural dynamics of MGMT. From snap shots, the Mu structure other than revealing expanded conformation, also formed helical conformation at amino acid number 87 to 90, giving an idea that the mutation does not favor the structural compactness of the protein, which inturn leads to its compromised and aberrated

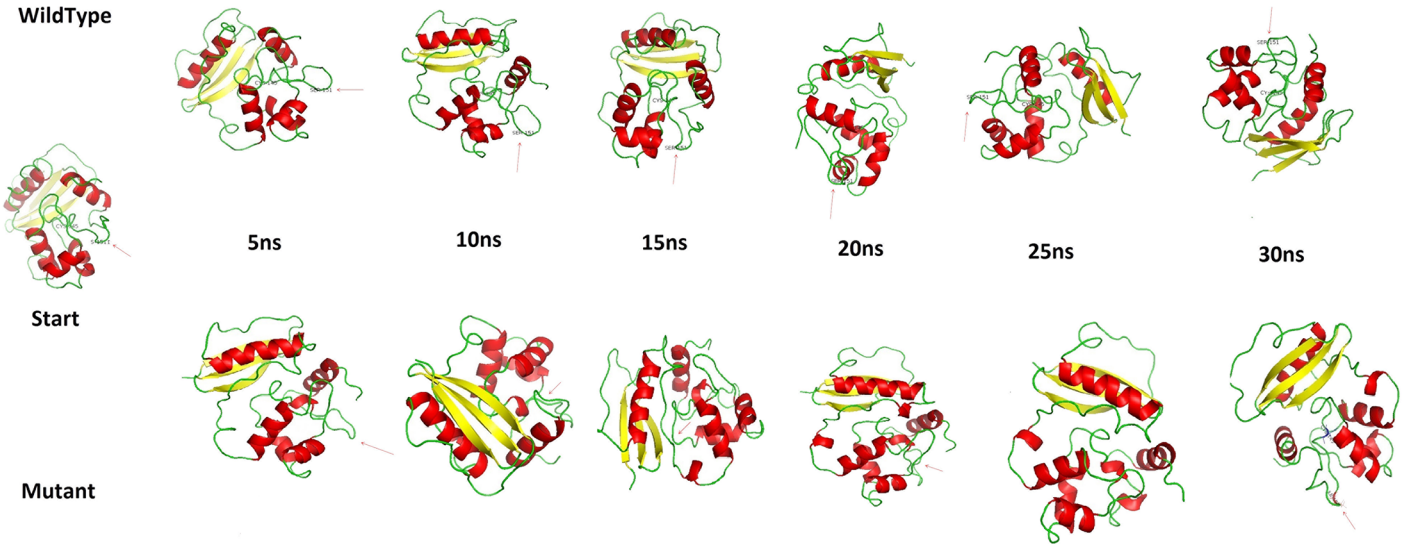


Fig 2. Illustrative representation of drastic conformational variability that a mutant structure (arrow specifies the site of an mutated residue) undergoes when compared wild type protein structure. The snapshots were retrieved at every 5 ns interval along the 30 ns simulation.

doi:10.1371/journal.pone.0127741.g002

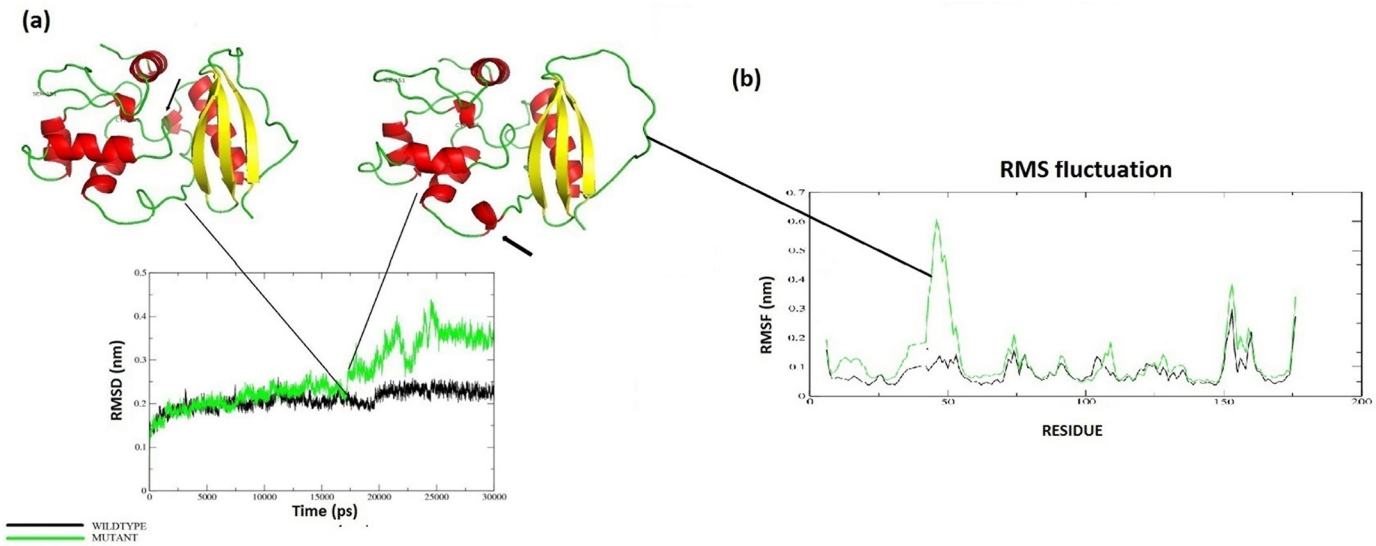


Fig 3. (a) Protein RMSDs for wt and Mu MGMT structures at 300 K. wt is shown in black and Mu in green. (Insets A and B) shows the relative structures at the point of RMSD jump. (b) MGMT residue RMSF along the MDS and the arrow pointing out to the region showing maximum fluctuation. (c) The RMSD vs. Atomic units. Plot showing highly unstable Mu curve in red.

doi:10.1371/journal.pone.0127741.g003

conformation having a considerable structural shift that is pivotal in causing defunct protein function [37]. After visual analysis, g_rms tool was used to calculate the RMSD for protein atoms, using the starting structure as a reference. The mutant structure showed abrupt elevation in RMSD at around 17 ns. On observing the anomaly at the structure level, we found that helical and loop content of the mutant structure varied (Fig 3A). The RMSD from the average over time is referred as RMSF, g_rmsf was used to calculate the atomic standard deviation and on observation, the Mu structure showed higher flexibility. The RMSF of both structures showed a slight change at residue 151, but is varying considerably in a protein loop region of 27 to 53 (Fig 3B), which might be the resultant of an intermolecular long range tertiary interaction variation. In r_rmsf tool the option-oq was used to convert the RMSF value into B Factor values and implicit them on the average structure (blue representing the most stable and red most fluctuating). The comparative B factor projection (S2A Fig) on wt and Mu MGMT primarily indicates fluctuations variations within the average structure, giving us an insight into the change in the fluctuating pattern between the two structures. The coloring pattern is default ranging between blue to red. A significant change in fluctuation observed in Mu structure besides which the average secondary structure layout (S2B Fig) differed considerably which again implies that the Mu can be disadvantageous to DNA repair.

To analyze the shape of the protein at each given time, g_ gyrate tool was used, which calculates the Radius of Gyration of a group of atoms along the x-, y- and z-axis, as a function of time. Our results demonstrate the major deviation in Radii of gyration in Mu structure, passed after 17 ns run (S3 Fig). While as it is known that the MGMT structure does not vary to the great extent when compared to MGMT-bound-DNA structure, indicative of stable bound structure via close association of recognition residues (Ala126, Ala127, Ala129, Gly131 and Gly132), and Ser93, Thr95, Gln115, Asn123, and Ser151, interacting with the phosphate backbone of DNA[36] however since the radi of gyration was recorded to be increased due to the mutation and therefore suggesting the expanded overturned protein structure presumably awkwardly shifts the Arginine finger (intrahelical positioned Arg128) from its position, which

is responsible for promoting the flipping of nucleotide into the MGMT active site, thus could impair the diligence needed for removing O^6-methylguanine adduct from DNA

Further as we know that each amino acid has its own hydrophobicity-value, the original wild-type residue and newly introduced mutant residue differ in this property. To evaluate this, we used g_sas tool which computes hydrophobic, hydrophilic and total SASA of the protein over time. The mutant structure has greater SASA which correlates with our earlier finding of increased Rg in mutant structure (S4 Fig). To check the effect of the Mu on the MGMT structure docked with DNA PDB ID:1T39 [38], we used Discovery studio to color and compute the hydrophobicity according to kyte-dolittle scale (S5 Fig). The wt hydrophobicity and five residue running average hydrophobicity were -0.8 and 0.94 respectively, whereas the corresponding values for Mu residue were considerably higher at 4.5 and 2, thus showing that Mu residue is more hydrophobic than the wt residue. The indexed deviation in the values of mutant protein hydrophobicity relative to the wt protein could profoundly affect the stiochiochemistry of hydrogen bond formation between the enzyme and DNA, as is evident from S5 Fig. Subsequently, the unfavorable Enzyme-DNA docking can lead to non-responsiveness of the enzyme with respect to its cooperative functionality.

To further the understanding of mutation on protein dynamics, we divided important amino acids involved in physical interaction with DNA and Mg^+ ion into clusters (Fig 4) depending upon their position and contributions in DNA docking, base flipping and DNA repair [36]. Cluster1 contained five amino-acids involving in DNA docking viz. SER93, PHE94, THR95, ASN123 and LYS125. Cluster 2 contained single amino acid ARG 135 also involved in DNA docking. Cluster 3 contained three amino-acids TYR114, GLN115 and SER151 where TYR 114 is involved in base flipping required for DNA repair and the other two have roles in DNA docking. Cluster 4, besides containing cluster 3 amino-acids, contained CYS145 which is an active site of MGMT, responsible for DNA repair. Cluster 5 consist three amino acids (CYS24, HIS29 and HIS85) all of which interact with Mg^+ ion. g_rama tool was used to generate phi/psi dihedral combinations of selected clusters and was used to compute the angles as a function of time. Their contour plot was generated using energy minima to understand their respective mobility (S6 Fig). All the selected clusters were affected by the mutation from

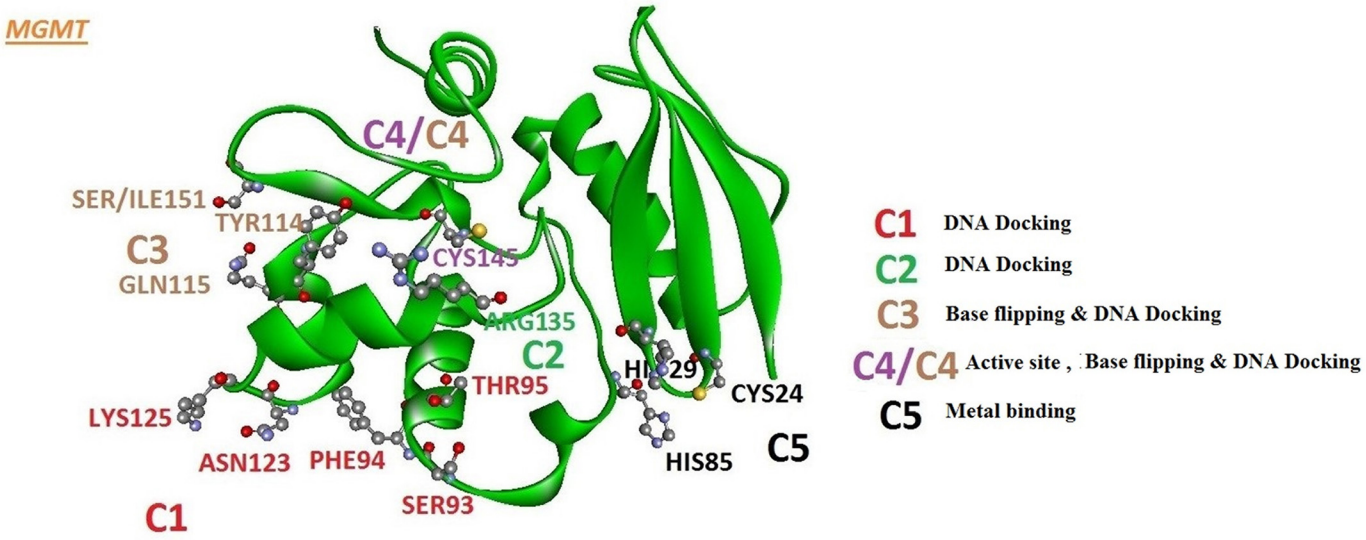


Fig 4. MGMT protein shown with the domains assembled into functionally important clusters that surround the active site.

doi:10.1371/journal.pone.0127741.g004

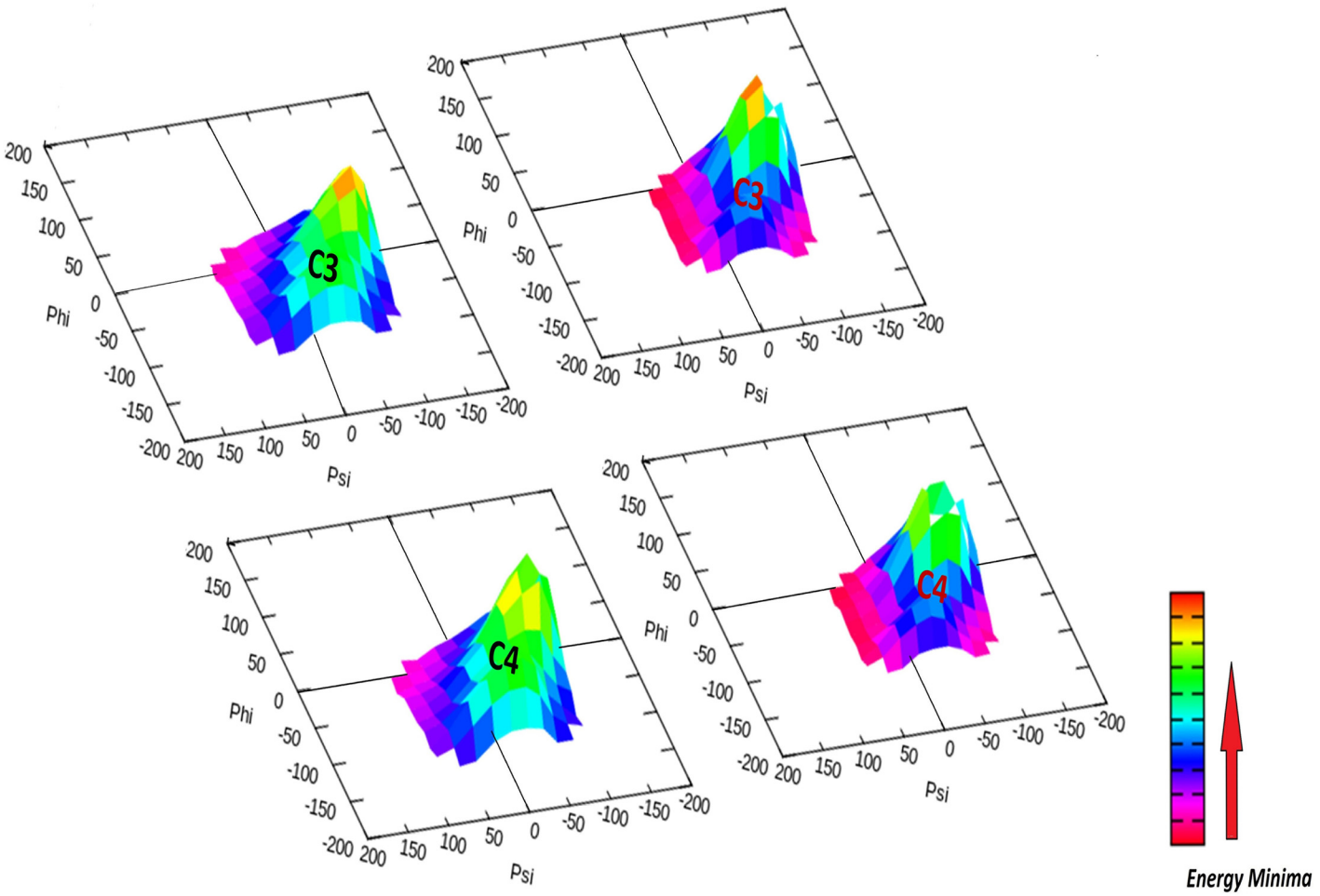


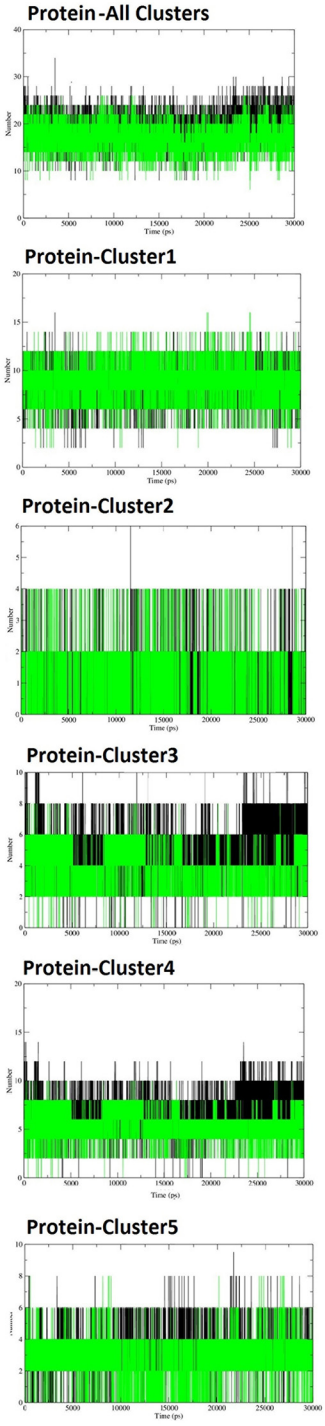
Fig 5. Three dimensional time dependent Phi/Psi distribution showing the change in the Gibbs free energy by the mutation in the Clusters 3 and 4 (black = wt, Red = Mu).

doi:10.1371/journal.pone.0127741.g005

SER151 to ILE151. To understand the effect, particularly on cluster 3 and 4, the Psi /phi distributions pertaining to the labeled energy minima were plotted (Fig 5). The difference in the peak region of energy minima can be observed in the corresponding wt and Mu clusters, giving distinctive impression of possible imparity in DNA repair.

For deeper understanding of the structural variation observed till now, we looked into intra hydrogen bond formation of the selected clusters using g_hband tool, the results of which have been shown in Fig 6. All the clusters selected for this analysis show the decrease in average number of hydrogen bonds per frame in mutant structure expect Cluster 1. The increase in the number of average hydrogen bonds per frame in Cluster 1 is slender in comparison to the variations we observe. The total decrease in the average hydrogen bond formation per frame is in co-relation with increased RMSF and Rg in mutant structure. The result generated by this analysis is conclusively entailing the anomaly observed till now with change in intra hydrogen bond pattern.

To understand the effect of this mutation on global correlated motions in atomic simulations, PCA, a mathematical technique that is efficient in characterizing the general folding and non-folding features of protein, was used. The technique identifies dominant motions in the protein by extracting principal modes involved in the motion involved in the molecule. The



Structure	Selection	Atoms Involved	Possible H-Bonds	Av. H-Bonds (per Frame)
Wildtype	Protein-All Clusters	(2546) (233)	53028	19.324
	Protein-Cluster1	(2546) (80)	49572	8.732
	Protein-Cluster2	(2546) (48)	49140	1.581
	Protein-Cluster3	(2546) (49)	49032	5.235
	Protein-Cluster4	(2546) (59)	49248	6.833
	Protein-Cluster5	(2546) (45)	49140	3.684
Mutant	Protein-All Clusters	(2554) (241)	52352	17.173
	Protein-Cluster1	(2554) (80)	49235	8.861
	Protein-Cluster2	(2554) (48)	48805	1.373
	Protein-Cluster3	(2554) (57)	48695	4.170
	Protein-Cluster4	(2554) (68)	48912	5.587
	Protein-Cluster5	(2554) (45)	48805	2.958

Fig 6. Intra protein hydrogen bond profile of wt and Mu MGMT protein over time at 300K.

doi:10.1371/journal.pone.0127741.g006

principal components of protein motion were computed as the eigenvectors (Ev) of the mass weighted covariance matrix of protein atoms. The calculation of these values was carried out using essential dynamics (ED) method according to standard protocol [39] available within the GROMACS software package. Two of the first eight Ev's that account to more that 85% motion

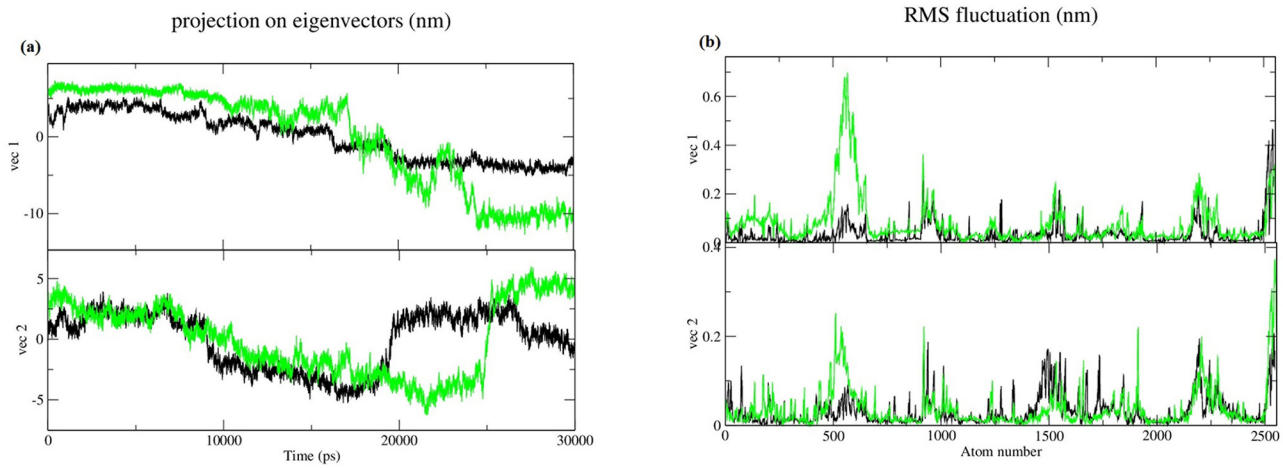


Fig 7. (a) The motion of principle EV overlaid sequentially. Black and green colour represents wt and Mu respectively (b) RMSF of all atoms of both vectors.

doi:10.1371/journal.pone.0127741.g007

of overall system were selected for analysis, the projection over time and RMSF fluctuation of which is depicted in Fig 7. Both the Ev's were combined into one single trajectory; the combination produced a common set of Principal Component (PC) eigenvectors for wt and Mu MGMT, making direct comparison possible among different systems. The trajectories were obtained using g-covar and g-anaeig of gromacs utilities. In Fig 8(A) the projections, PC 1 vs. PC 2, of both structures are projected (black wt/ red Mu), the cluster obtained from wt structure is stable, where as the projection of first two PC of mutant covers a large area. To further analyze the PC projections, their free energy surfaces were plotted (Fig 8B) which revealed that the stability of wt over the run is uniform over time as compared to Mu based on the energy minima basins formed by both. The structures with minimum energy were retrieved from the free energy land scape at different points of time. The structures on the right side of each projection in Fig 8(C) of PC are from the start of simulation to the left one from the near end of simulation. This analysis was crucial in elucidating the compromised free energy landscape of Mu structure, an observation that besides corroborating with our previous results, has conclusively implied a drastic conformational change in Mu structure.

Conclusion

Incongruities of DNA repair and cancer etiology are synonyms in a way that it is the occurrence of the mutation that has been widely accepted as the basis of cancer. A mutation in a DNA repair protein that could impair its function (S7 Fig) can create pretumorigenic environment and can assist in cancer progression at any stage. MGMT being one of the important DNA repair protein has an essential role in maintaining genomic stability by removing O^6Methyleguanine adducts. Thus, a significant genetic polymorphism in this protein will have an effect on cancer development and its progression. As none of the studies till date has reported mutational analysis of MGMT using MDS, it has primarily prompted us to look into the possibility of MGMT being mutated in a classified population where consumption of foods containing higher levels of N-nitroso compounds is common and gastric cancer is prevalent.

The use of molecular dynamics to study the effect of novel mutation at codon ^{151} has given us an insight into architecture of both the structures at an atomic level over a run of 30 ns period. The effect of the mutation was not only limited to its vicinity, but also impinged on overall structure including secondary elements at different locations of the protein. The structural

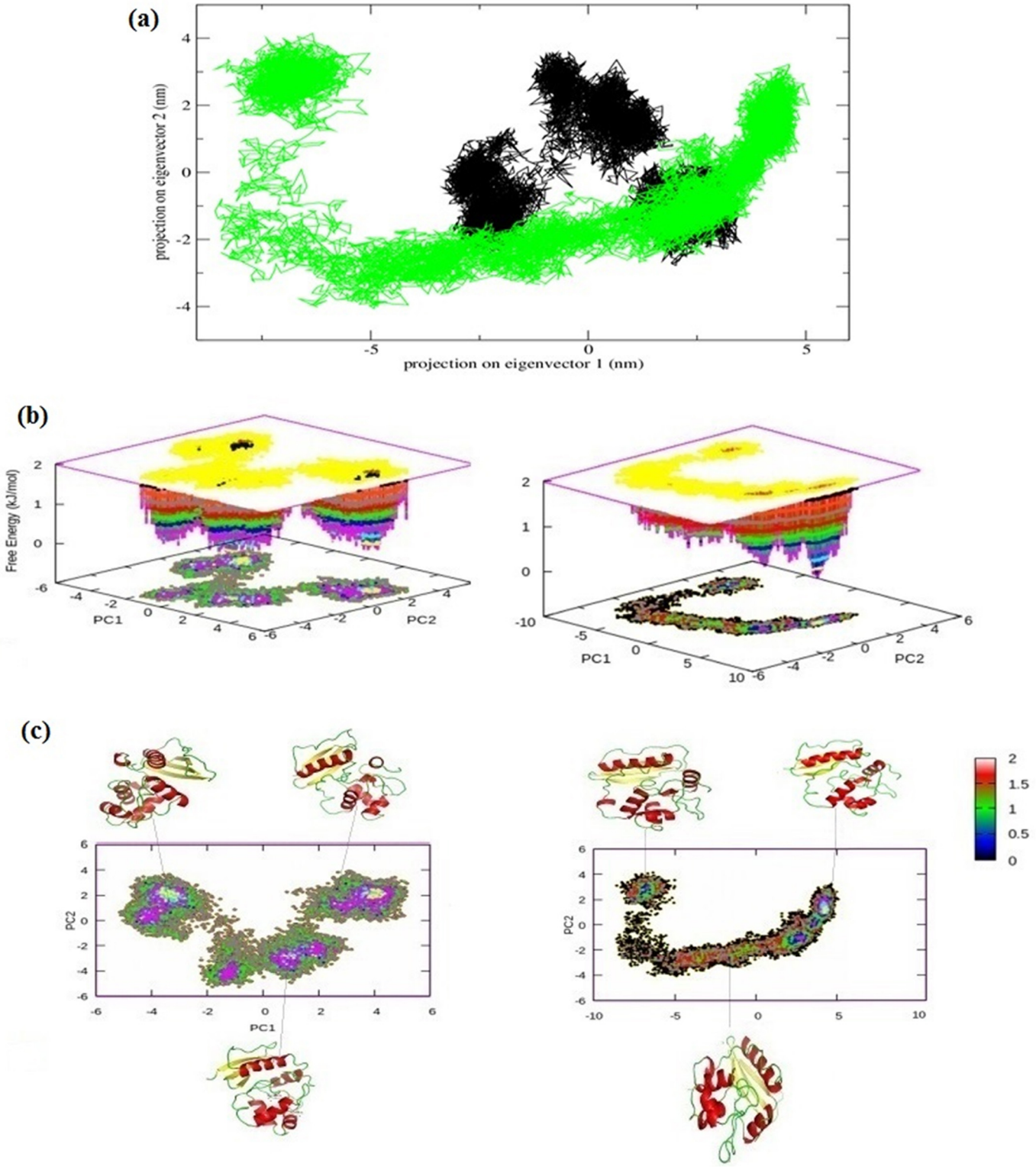


Fig 8. (a) Two dimensional representation of the motion of both structures along the first two principal eigenvectors, Black and green represent the wt and Mu MGMT respectively. (b) FEL of both the motions generated separately. (c) Separate two dimensional representations of PCA of both wt and Mu MGMT with the inset of three most stable structures at different point of time.

doi:10.1371/journal.pone.0127741.g008

transitions observed in secondary elements, promotes the collapse of structural architecture of Mu MGMT protein. The FEL obtained by quasiharmonic analysis (PCA) also concluded that the mutation considerably affects the stability of the MGMT over time, a factor that can hamper the normal stoichiometric fashion of DNA repair by MGMT.

The explored mutation in exon 5 appears to be associated with driver mutation, which seems to affect DNA/protein interaction, an important factor that could affect DNA docking, base flipping and ultimately repair mechanism, which if impaired, could also result in genome wide increase in O^6 methyl guanine adducts leading to increased genomic instability.

Materials and Methods

Ethical statement

The protocols/experiments involving the use of human specimens were duly examined and approved by University Human Ethics Committee (UHEC), VIT University, Vellore (UHEC--VIT/2011).

Patients and tissue collection

A total of 30 patients diagnosed with gastric carcinoma admitted to Sheri-Kashmir Institute of Medical sciences (SKIMS), Srinagar were considered for the study. Patients undergoing surgery as the primary treatment at different stages of the disease were recruited for the study with their consent. The characteristics of the studied patients are listed in S2 Table.

Tumor samples $5mm^3$ were excised from surgically resected specimens within the tumor mass, excluding the margin. Adjacent non-neoplastic samples of similar dimension were taken from the resection margin, approximately 10mm from the macroscopic tumor edge and subsequently confirmed as benign by routine histopathology at SKIMS. A total of 30 tumor and 30 normal tissue samples were collected, and stored at -80°C until analysis.

DNA extraction and Polymerase Chain Reaction

DNA was extracted from $2mm^3$ tissue samples using DNA extraction kit (Hi Pura Mammalian Genomic DNA Isolation Kit—HiMedia). Concentration and quality of DNA was measured by routine spectrophotometeric analysis. Amplification of the Exon 5 regions of the MGMT exon, was carried out in gradient minicycler (Eppendorf) in a 25μl reaction mix containing 1 μl (400ng/μl) genomic DNA, DNA polymerase {1X PCR buffer (200mM Tris HCl, 200mM KCl, 50 mM, (NH4)2 SO4) supplied with 25mM Mgcl2,Fermentas}, Nuclease free water and 1 μl of forward (5'- GCCCGTGCAGGTACGGTCTT-3') and reverse (5'- AGCTCCCGCTCCCTTG AGCC-3') primers each. The annealing temperature was optimized at 65.5°C. To facilitate Polymerase Chain Reaction (PCR) product analysis for mutation, PCR product sequencing was carried out.

SNP Damage Prediction. The damage prediction of the polymorphisim was carried out using SIFT [40], Polyphen-2[41], PhD-SNP [42], MutPred [43], SNAP [44], SNPs & Go[45] and PoPMuSiC [46].

Molecular dynamics simulation

MDS studies were performed by Gromacs 4.5.3 package [47]. For wt MGMT, the PDB structure 1QNT [48] was used as a starting structure for MDS. Accelrys Discovery Studio [49] was used to make the single point mutation on the wild type structure. Both, wt and Mu MGMT were applied with GROMOS96 43a1 force field and then placed in a model of a pre-equilibrated water bath and counter-ions were added to achieve a neutral box using the "genion" tool

that comes along with gromacs package. Solvent molecules were restrained to the original position with a force constrain of 100Kcal/mol for 5000 steps before being subjected to energy minimization for 5000 iteration. For regulating the temperature inside the box, Berendsen temperature coupling method [50] was used. Electrostatic interactions were computed using the Particle Mesh Ewald method [51]. Ionizing state of the residues, pressure and other parameters were set in the standard range. Non-bonded pair list was updated after every 10 steps and conformations were stored every 2 pico seconds (ps). Position restraint simulation for 500 ps was implemented to allow solvent molecules to enter the cavity region of structure. Finally, system was subjected to MDS for 30 nano seconds (ns). Root mean Square Deviation (RMSD), Root Mean Square Fluctuation (RMSF), Solvent Accessible Surface Area (SASA), Radius of gyration (Rg) and PCA were carried out by using inbuilt gromacs tools. g_hbond was used to calculate the number of distinct hydrogen bonds formed by specific residues to other amino acids within the protein during simulations (NH bond). g_sham was used extensively to obtain free energy landscape. Graphs were plotted using Grace GUI toolkit 5.1.22 version while as free energy landscapes were plotted using gnuplot 4.6.0 version. All visualizations were carried out using Pymol, Ligplus, VMD [52] and graphs were plotted using Grace Program [53] and GNUPlot. Trajectories were analyzed using the inbuilt tool in the GROMACS distribution.

Supporting Information

S1 Fig. a) A representative chromatogram of MGMT exon 5 showing the single base pair, G>T at position 151 as indicated by an arrow in the neoplastic chromatogram. b) Alignment of exon 5 sequence that was amplified from neoplastic and non-neoplastic tissue (adjacent normal) with that of wild type (Reference-sequence acquired from NCBI) was translated and the SNP mapped was shown to change of Serine into Isoleucine.
(TIF)

S2 Fig. (a) Average Tertiary structures colored according to Bfactor values (b) Average Secondary structure representation of both structures.
(TIF)

S3 Fig. (a) Radii of gyration of wt and Mu MGMT shown separately. (b) Rg of all atoms of wt and Mu MGMT versus time at 300K.
wt is represted by Black and Mu by Green.
(TIF)

S4 Fig. Solvent Accessible Surface Area of wt (black) and Mu (green) MGMT over time at 300K.
(TIF)

S5 Fig. Variation In color (Surface of protein according to kyte-doolittle scale) at mutated region and the graphical representation of variation in kyte-doolittle scale in single amino acid and five average running hydrophobicity of both wt and Mu MGMT.
(TIF)

S6 Fig. Time dependent Ramachandran Contour plot of all the selected clusters over time, each line showing the transition of 1.
(TIF)

S7 Fig. Pictorial Representation of GC:AT Transition by impaired MGMT.
(TIF)

S1 Table. Polymorphism Prediction using different servers.
(DOCX)

S2 Table. Characteristics of Study Subjects.
(DOCX)

Acknowledgments

The authors would like to acknowledge Dr. Daniele Granata for his kind inputs on free energy landscapes.

Author Contributions

Conceived and designed the experiments: NAC TNCP. Performed the experiments: NAC. Analyzed the data: NAC SB. Contributed reagents/materials/analysis tools: TNCP RAQ SS NS. Wrote the paper: NAC SB AA NS.

References

1. Ferlay, J., Soerjomataram, I., Dikshit, R., Eser, S., Mathers, C., Rebelo, M., et al. (2014). Cancer incidence and mortality worldwide: Sources, methods and major patterns in GLOBOCAN 2012. Int. J. Cancer
2. Parkin D. M., Bray F., Ferlay J., & Pisani P. (2005). Global cancer statistics, 2002. CA: a cancer journal for clinicians, 55(2), 74–108.
3. Zheng L., Wang L., Ajani J., & Xie K. (2004). Molecular basis of gastric cancer development and progression. Gastric cancer, 7(2), 61–77. PMID: 15224192
4. Fiocca R., Luinetti O., Villani L., Mastracci L., Quilici P., Grillo F., et al. (2000). Molecular mechanisms involved in the pathogenesis of gastric carcinoma: interactions between genetic alterations, cellular phenotype and cancer histotype. Hepato-gastroenterology, 48(42), 1523–1530.
5. Elrifai W., & Powell S. (2002). Molecular biology of gastric cancer*. Seminars in Radiation Oncology, 12(2), 128–140. PMID: 11979414
6. Chikan N. A., Shabir N., Shaffi S., Mir M. R., & Patel T. N. (2012). N-Nitrosodimethylamine in the Kashmiri Diet and Possible Roles in the High Incidence of Gastrointestinal Cancers. Asian Pacific Journal of Cancer Prevention, 13, 1077–1079. PMID: 22631641
7. Sekiguchi M., Nakabeppu Y., Sakumi K., & Tuzuki T. (1996). DNA-repair methyltransferase as a molecular device for preventing mutation and cancer. Journal of cancer research and clinical oncology, 122 (4), 199–206. PMID: 8601571
8. Engelbergs J., Thomale J., & Rajewsky M. F. (2000). Role of DNA repair in carcinogen-induced *ras* mutation. Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis, 450(1), 139–153.
9. Ito T., Nakamura T., Maki H., & Sekiguchi M. (1994). Roles of transcription and repair in alkylation mutagenesis. Mutation Research/DNA Repair, 314(3), 273–285. PMID: 7513059
10. Warren J. J., Forsberg L. J., & Beese L. S. (2006). The structural basis for the mutagenicity of O6-methyl-guanine lesions. Proceedings of the National Academy of Sciences, 103(52), 19701–19706. PMID: 17179038
11. Montesano R., Becker R., Hall J., Likhachev A., Lu S. H., Umbenhauer D., et al. (1985). Repair of DNA alkylation adducts in mammalian cells. Biochimie, 67(9), 919–928. PMID: 3910113
12. Pegg A. E. (1990). Mammalian O6-alkylguanine-DNA alkyltransferase: regulation and importance in response to alkylating carcinogenic and therapeutic agents. Cancer Research, 50(19), 6119–6129. PMID: 2205376
13. Edara S., Kanugula S., & Pegg A. E. (1999). Expression of the inactive C145A mutant human O6-alkylguanine-DNA alkyltransferase in E. coli increases cell killing and mutations by N-methyl-N′-nitro-N-nitrosoguanidine. Carcinogenesis, 20(1), 103–108. PMID: 9934856
14. Pegg A. E., & Byers T. L. (1992). Repair of DNA containing O6-alkylguanine.The FASEB journal, 6(6), 2302–2310. PMID: 1544541
15. Sakumi K., Shiraishi A., Shimizu S., Tsuzuki T., Ishikawa T., & Sekiguchi M. (1997). Methylnitrosourea-induced tumorigenesis in MGMT gene knockout mice. Cancer research, 57(12), 2415–2418. PMID: 9192819

16. Nakatsuru Y., Matsukuma S., Sekiguchi M., & Ishikawa T. (1991). Characterization of O ^{6}-methylguanine-DNA methyltransferase in transgenic mice introduced with the E. coli ada gene. Mutation Research/DNA Repair, 254(3), 225–230. PMID: 2052012

17. Kaur T. B., Travaline J. M., Gaughan J. P., Richie J. P., Stellman S. D., & Lazarus P. (2000). Role of polymorphisms in codons 143 and 160 of the O6-alkylguanine DNA alkyltransferase gene in lung cancer risk. Cancer Epidemiology Biomarkers & Prevention, 9(3), 339–342.

18. Egyhazi S., Ma S., Smoczynski K., Hansson J., Platz A., & Ringborg U. (2002). Novel O6-methylguanine-DNA methyltransferase SNPs: A frequency comparison of patients with familial melanoma and healthy individuals in Sweden. Human mutation, 20(5), 408–409. PMID: 12402349

19. Li C., Liu J., Li A., Qian L., Wang X., Wei Q., et al. (2005). Exon 3 polymorphisms and haplotypes of O^{6}*methylguanine-DNA methyltransferase* and risk of bladder cancer in southern China: A case–control analysis. Cancer letters, 227(1), 49–57. PMID: 15885889

20. Cohet C., Borel S., Nyberg F., Mukeria A., Brüske-Hohlfeld I., Constantinescu V., et al., (2004). Exon 5 Polymorphisms in the O6-Alkylguanine DNA Alkyltransferase Gene and Lung Cancer Risk in Non–Smokers Exposed to Second-Hand Smoke. Cancer Epidemiology Biomarkers & Prevention, 13(2), 320–323.

21. Ford B. N., Ruttan C. C., Kyle V. L., Brackley M. E., & Glickman B. W. (2000). Identification of single nucleotide polymorphisms in human DNA repair genes. Carcinogenesis, 21(11), 1977–1981. PMID: 11062157

22. Wang L., Liu H., Zhang Z., Spitz M. R., & Wei Q. (2006). Association of genetic variants of O6-methylguanine-DNA methyltransferase with risk of lung cancer in non-Hispanic Whites. Cancer Epidemiology Biomarkers & Prevention, 15(12), 2364–2369.

23. Yang M., Coles B. F., Caporaso N. E., Choi Y., Lang N. P., & Kadlubar F. F. (2004). Lack of association between Caucasian lung cancer risk and *O*^{6}-methylguanine-DNA methyltransferase-codon 178 genetic polymorphism. Lung Cancer, 44(3), 281–286. PMID: 15140540

24. Dhar G. M., Shah G. N., & Naheed B. (1993). Epidemiological trend in the distribution of cancer in Kashmir Valley. Journal of epidemiology and community health, 47(4), 290–292. PMID: 8228764

25. Khuroo M. S., Zargar S. A., Mahajan R., & Banday M. A. (1992). High incidence of oesophageal and gastric cancer in Kashmir in a population with special personal and dietary habits. Gut, 33(1), 11–15. PMID: 1740265

26. Kumar R., Mende P., Tricker A. R., Siddiqi M., & Preussmann R. (1990). *N*-Nitroso compounds and their precursors in Brassica oleracea.Cancer letters, 54(1), 61–65.

27. Siddiqi M., Tricker A. R., & Preussmann R. (1988). The occurrence of preformed *N*-nitroso compounds in food samples from a high risk area of esophageal cancer in Kashmir, India. Cancer letters, 39(1), 37–43. PMID: 3345506

28. Siddiqi M., Tricker A. R., & Preussmann R. (1988). Formation of *N*-nitroso compounds under simulated gastric conditions from Kashmir foodstuffs.Cancer letters, 39(3), 259–265. PMID: 3359419

29. Priyadharshini C. J., & George P. D. C. (2014). Single amino acid polymorphism in aldehyde dehydrogenase gene superfamily. Frontiers in bioscience (Landmark edition), 20, 335–376.

30. George P. D. C., Rajith B., Chakraborty C., NagaSundaram N., Ali S. K., & Zhu H. (2014). Structural signature of the G719S-T790M double mutation in the EGFR kinase domain and its response to inhibitors. Scientific reports, 4.

31. George D. C. P., Chakraborty C., Haneef S. S., NagaSundaram N., Chen L., & Zhu H. (2014). Evolution-and structure-based computational strategy reveals the impact of deleterious missense mutations on MODY 2 (maturity-onset diabetes of the young, type 2). Theranostics, 4(4), 366. doi: 10.7150/thno.7473 PMID: 24578721

32. Bukhari S., Mokhdomi T. A., Chikan N. A., Amin A., Qazi H., Wani S. H., et al. (2015). Affinity proteomics led identification of vimentin as a potential biomarker in colon cancers: insights from serological screening and computational modelling. Molecular BioSystems, 11(1), 159–169. doi: 10.1039/c4mb00506f PMID: 25319351

33. Doss C., Chakraborty C., Abishek N. M. P., & Kumar D. T. (2014). Application of Evolutionary Based in Silico Methods to Predict the Impact of Single Amino Acid Substitutions in Vitelliform Macular Dystrophy. ADVANCES IN PROTEIN CHEMISTRY AND STRUCTURAL BIOLOGY, VOL 94, 94, 177–267. doi: 10.1016/B978-0-12-800168-4.00006-8 PMID: 24629188

34. Duguid E. M., Rice P. A., & He C. (2005). The structure of the human AGT protein bound to DNA and its implications for damage detection. Journal of molecular biology, 350(4), 657–666. PMID: 15964013

35. Daniels D. S., & Tainer J. A. (2000). Conserved structural motifs governing the stoichiometric repair of alkylated DNA by *O*^{6}-alkylguanine-DNA alkyltransferase. Mutation Research/DNA Repair, 460(3), 151–163.

36. Daniels D. S., Mol C. D., Arvai A. S., Kanugula S., Pegg A. E., & Tainer J. A. (2000). Active and alkylated human AGT structures: a novel zinc site, inhibitor and extrahelical base binding. The EMBO journal, 19(7), 1719–1730. PMID: 10747039

37. Petsko G. A., & Ringe D. (2004). Protein structure and function. New Science Press.

38. Daniels D. S., Woo T. T., Luu K. X., Noll D. M., Clarke N. D., Pegg A. E., et al. (2004). DNA binding and nucleotide flipping by the human DNA repair protein AGT. Nature structural & molecular biology, 11(8), 714–720.

39. Amadei A., Linssen A. B. M., De Groot B. L., Van Aalten D. M. F., & Berendsen H. J. C. (1996). An efficient method for sampling the essential subspace of proteins. Journal of Biomolecular Structure and Dynamics, 13(4), 615–625. PMID: 8906882

40. Kumar P., Henikoff S., & Ng P. C. (2009). Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. Nature protocols, 4(7), 1073–1081. doi: 10.1038/nprot.2009.86 PMID: 19561590

41. Adzhubei I. A., Schmidt S., Peshkin L., Ramensky V. E., Gerasimova A., Bork P., et al. (2010). A method and server for predicting damaging missense mutations. Nature methods, 7(4), 248–249. doi: 10.1038/nmeth0410-248 PMID: 20354512

42. Capriotti E., Calabrese R., & Casadio R. (2006). Predicting the insurgence of human genetic diseases associated to single point protein mutations with support vector machines and evolutionary information. Bioinformatics, 22(22), 2729–2734. PMID: 16895930

43. Li B., Krishnan V. G., Mort M. E., Xin F., Kamati K. K., Cooper D. N., et al. (2009). Automated inference of molecular mechanisms of disease from amino acid substitutions. Bioinformatics, 25(21), 2744–2750. doi: 10.1093/bioinformatics/btp528 PMID: 19734154

44. Bromberg Y., & Rost B. (2007). SNAP: predict effect of non-synonymous polymorphisms on function. Nucleic acids research, 35(11), 3823–3835. PMID: 17526529

45. Calabrese R., Capriotti E., Fariselli P., Martelli P. L., & Casadio R. (2009). Functional annotations improve the predictive score of human disease-related mutations in proteins. Human mutation, 30(8), 1237–1244. doi: 10.1002/humu.21047 PMID: 19514061

46. Gonnelli G., Rooman M., & Dehouck Y. (2012). Structure-based mutant stability predictions on proteins of unknown structure. Journal of biotechnology, 161(3), 287–293. doi: 10.1016/j.jbiotec.2012.06.020 PMID: 22782143

47. Hess B., Kutzner C., Van Der Spoel D., & Lindahl E. (2008). GROMACS 4: Algorithms for highly efficient, load-balanced, and scalable molecular simulation. Journal of chemical theory and computation, 4(3), 435–447.

48. Wibley J. E., Pegg A. E., & Moody P. C. (2000). Crystal structure of the human O6-alkylguanine-DNA alkyltransferase. Nucleic acids research, 28(2), 393–401. PMID: 10606635

49. Accelrys Software Inc., Discovery Studio Modeling Environment, Release 3.5, San Diego: Accelrys Software Inc., 2012.

50. Berendsen H.J.C., Postma J.P.M., van Gunsteren W.F, Dinola ,A., Haak J.R. (1984). Molecular dynamics with coupling to an external bath. J. Chem. Phys. 81, 3684–3690

51. Cheatham T.E., Miller J.L., Fox T., Darden T.A., Kollman P.A. (1995). Molecular dynamics simulations on solvated biomolecular systems: the particle mesh ewald method leads to stable trajectories of DNA, RNA, and proteins, J. Am. Chem. Soc. 14, 4193–4194.

52. Humphrey W., Dalke A., & Schulten K. (1996). VMD: visual molecular dynamics. Journal of molecular graphics, 14(1), 33–38. PMID: 8744570

53. Turner P. J. (2005). XMGRACE, Version 5.1. 19. Center for Coastal and Land-Margin Research, Oregon Graduate Institute of Science and Technology, Beaverton, OR.