Academic Sciences International Journal of Pharmacy and Pharmaceutical Sciences

ISSN- 0975-1491 Vol 5, Issue 4, 2013

Research Article

MOLECULAR DYNAMICS SIMULATIONS OF NATIVE AND MUTANT GLUCOSAMINYL (N-ACETYL) TRANSFERASE 1 PROVE THEIR BIOCHEMICAL VARIATIONS

L. INBATHAMIZH[1] AND E. PADMINI*[2]

[1]Research and Development Centre, Bharathiar University, Coimbatore 641046, Tamilnadu, India, *[2]Department of Biochemistry, Bharathi Women's College, Chennai 600108, Tamilnadu, India. Email: inba66in@yahoo.com

Received: 01 Aug 2013, Revised and Accepted: 02 Sep 2013

ABSTRACT

Objective: With the emergence of deadly diseases, the need for alternative and effective therapies enhances. Target-based therapy seems to be more potent than chemotherapy to treat cancer. Glucosaminyl (N-acetyl) transferase 1 (GCNT1) is a major enzyme involved in cell adhesion, glycosylation and branching at the surface and in the secretions of prostate cancer causing glycoproteins. Its mutant with Adenine/Guanine polymorphism in GCNT1 is associated with the increased risk of prostate cancer. Both the native and mutant GCNT1s pose to be potent drug targets for prostate cancer. The present study is framed with an objective to confirm through molecular dynamics simulations, the biochemical variations that exist between these two target proteins due to mutation.

Method: *In silico* methods are adopted to study the molecular dynamics of these proteins along with their biochemistry and stability.

Results: The results confirm significant differences between the two proteins in all the analyzed aspects.

Conclusion: Besides suggesting the effect of mutation on protein conformation and further its association with the progress of the disease, the study provides an insight into the drug targets for prostate cancer therapy.

Keywords: Glucosaminyl (N-acetyl) transferase 1, Prostate cancer, Drug targets, Molecular dynamics simulations, *In silico* methods, Mutation.

INTRODUCTION

Cancer is responsible for many deaths (1 in 8) worldwide with metastasis being the major cause of death from cancer. Effective management, treatment and cure of cancer are undoubtedly crucial. The control of cancer may benefit from the potential that resides in alternative therapies [1]. Targeted drug therapy which seeks to understand the basic chemistry and physiology of the disease with drugs interacting with specific key molecules in cancer cells and causing fewer side effects on healthy cells would be more effective [2]. With the tremendous explosion in the large amounts of available biological information and completion of human genome sequencing, it would be possible to identify and work with drug targets employing a series of *in silico* techniques. Identification and molecular studies of useful drug targets that can elicit the desired effect through interaction with the drug ligands have been the paradigm for the drug development process in the era of molecular medicine.

GCNT1(Enzyme Commission number EC=2.4.1.102), also called core 2 (beta-1,6-N-acetyl glucosaminyl transferase) with UniProt ID: Q02742, seems to be an important metabolic enzyme involved in cancer especially in prostate cancer to greater extent and also to significant extent in breast cancer and colon cancer. It forms critical branches in O-glycans catalyzing the conversion of Uridine 5' diphosphate (UDP)-N-acetyl-D-glucosamine and beta-D-galactosyl-1,3-N-acetyl-D-galactosaminyl-R to UDP and beta-D-galactosyl-1,3-(N-acetyl-beta-D-glucosaminyl-1,6)-N-acetyl-D-galactosaminyl-R where R is an alkyl group. GCNT1 is responsible for glycosylation and branching at the surface and in the secretions of cancer causing glycoproteins [3]. It is one of the best tumor markers for cancer [4]. Expression of GCNT1 is positively correlated with the progression of prostate cancer in humans. GCNT1 cells adhere more strongly to prostate stromal cells, type IV collagen and laminin. This increased adhesion may be a cause for aggressive tumor formation by prostate cancer cells rich in GCNT1 [5].

GCNT catalyzes the addition of complex oligosaccharide side chains to glycoproteins, regulating the expression and function of several proteins involved in tumor metastasis [6]. Elevated levels of GCNT substrates are correlated with various markers of tumor progression, including positive node status, large tumor size and high nuclear grade. Cell-surface expression of beta 1-6 branched oligosaccharides in metastatic tumor cells is specifically associated with increased invasion of human amnion basement membranes *in vitro* [7]. It has been shown that overexpression of GCNT1 results in increased phosphorylation of mitogen-activated protein kinase and c-fos promoter activation [8].

GCNT1 is an inverting, metal ion-independent family 14 glycosyltransferase that catalyzes the formation of the core 2 O-glycan [9]. Beta 1,6 N-acetyl glucosaminyl transferase functions in both core 2 and core 4 O-glycan branch formation [10]. The redundancy in beta 1,6 N-acetyl glucosaminyl transferase capable of forming core 2 O-glycans is important for understanding the mechanisms leading to specific changes in core 2 branching during cell development and malignant transformation.

Statistical studies demonstrate that GCNT1 is an independent predictor for progressed pathological stage and for Prostate-specific membrane antigen (PSMA) relapse [11]. GCNT1, specifically detected in cancer tissues of 63% of patients, is closely correlated with the vessel invasion, as well as depth of tumor invasion [12]. Earlier studies have indicated that the levels of this enzyme are elevated in human prostate cancer PC3 cell lines. Reducing its activity in PC3 cells is associated with reduced expression of matrix metalloproteinases (MMP-2 and MMP-9) and vascular endothelial growth factor (VEGF), resulting in inhibition of invasion and angiogenesis, implicating it as a therapeutic target in the regulation of malignant properties and treatment of cancer [13].

The structure of a protein could change in various ways due to the biochemical differences of the amino acid variant whether acidic, basic, or hydrophobic, and by the location of the variant in the protein sequence [14]. Sequence analysis has shown that Adenine (A) / Guanine (G) polymorphism in GCNT1 with mutation of isoleucine to valine at 152 in the protein, is associated with the increased risk of prostate cancer [15]. It is found that while men with the GG genotype have a 3.60-fold increased risk, those with the AG genotype have a 1.58-fold compared with those with the AA genotype. The G allele is found to have a gene dosage effect for prostate cancer risk and is more frequently observed in the prostate cancer group. The molecular study of this mutant type of GCNT1 seems to be important in investigating its association with the enhanced risk of the disease.

One of the essential tools in the study of biomacromolecules includes the evaluation of molecular dynamics simulations. These

biocomputational techniques provide extensive information on the fluctuations and conformational changes of proteins based on the calculation of the time dependent behavior of the molecular system. Thus, this study exploits these techniques to substantiate the conformational changes brought about in GCNT1 upon mutation which may find a valuable application in target-based therapy for prostate cancer.

MATERIALS AND METHODS

Bioinformatics tools

Modeller 9v1

Modeller is used for homology or comparative modeling of protein three-dimensional structures by aligning the sequence to be modeled with known related structures [16].

Biochemistry-online analysis tool

A tool for the calculation of biochemical parameters of protein from its sequence [17].

I-Mutant v 2.0

A support vector machine-based web server for the automatic prediction of protein stability changes upon single point mutation starting from the protein structure or from the protein sequence [18]. The tool is trained on a data set derived from ProTherm that is presently the most comprehensive database of experimental data on protein mutations.

GROMACS (GROningen MAchine for Chemical Simulations) 3.2v

GROMACS is an open source software package originally developed by the department of Biophysical Chemistry, Groningen University, for simulating the forces and movements of atoms in molecular systems over time (molecular dynamics) in biomolecules [19].

RasMol version 2.7.3

RasMol is a program for molecular graphics visualization originally developed by Roger Sayle. It is used primarily for the depiction and exploration of biological macromolecule structures [20,21].

VMD (Visual Molecular Dynamics)

A molecular visualization program for displaying, animating and analyzing large biomolecular systems using three dimensional (3D) graphics and built-in scripting [22].

Methodology

Single point mutation and modeling of mutant protein

In order to study the variations in the structure and properties of GCNT1 due to mutation, single point mutation was created manually. Isoleucine at position 152 in the sequence of native GCNT1 (nGCNT1) was replaced by valine based on the association of the mutation with the increased risk of prostate cancer. The mutant GCNT1 (mGCNT1) thus created was modeled by Modeller9v1 as done for nGCNT1 [23]. The modeled structure of mGCNT1 was visualized and compared with that of nGCNT1 using RasMol version 2.7.3.

Biochemistry analysis

For details regarding variations in the biochemical parameters between nGCNT1 and mGCNT1, Biochemistry-online analysis tool was utilized, with the protein sequence in one-letter coding fed as the input.

Validation of stability changes in protein on mutation

Stability changes in GCNT1 upon mutation in terms of free energy change at 37 ^{o}C and pH 7 were validated using I-Mutant v 2.0. The protein sequence in raw format and one letter code with the position of mutation and the new residue were the required inputs.

Molecular dynamics simulations

GROMACS program was invoked for the molecular dynamics studies of both nGCNT1 and mGCNT1. The PDB files of the proteins were processed with pdb2gmx command to generate the topology and add necessary hydrogen atoms. The proteins were centered in suitably sized boxes for simulation using "editconf" command. A more realistic simulation was created by solvating the empty space in the boxes with the Simple Point Charge (SPC) 216, a 3-point water model using "genbox" command. The energy minimization and molecular dynamics position restraint were set up using "grompp" command. Position Restraint was performed using "mdrun" command. Again "grompp" and "mdrun" commands were used to set up and perform molecular dynamics simulations. VMD was used to visualize the output structures as .gro files.

The structural changes the proteins had undergone in a period of time of 0-10 pico seconds (ps) at 37 ^{o}C were analyzed. The energy changes in the system were calculated using g_energy. The function g_helix computed all kinds of helix properties.

RESULTS AND DISCUSSION

Biochemistry and composition of nGCNT1 and mGCNT1 varied as presented in Table 1a and Table 1b. Length of the two proteins remained same as there was no addition or deletion of amino acids. But their Brutto formula and molecular weight were different due to the change in the number of C and H atoms with the mutation of isoleucine to valine. This was reflected in the composition of nGCNT1 with larger quantity and percentage of isoleucine and lesser quantity and percentage of valine while *vice-versa* in mGCNT1. The charge under pH 7 and the isoelectric point of the proteins showed no difference as both the amino acids involved in the mutation were non-polar and neutral.

Table 1a: Biochemistry of nGCNT1 and mGCNT1 by analysis tool

Parameter/Protein	nGCNT1	mGCNT1
Length	428 amino acids	428 amino acids
Brutto formula	$C_{2259}H_{3519}N_{595}O_{626}S_{24}$	$C_{2258}H_{3517}N_{595}O_{626}S_{24}$
Molecular weight (with isotopes 13C 15N D)	56114.82 Da	56097.82 Da
Isoelectric point pI	9.29	9.29
Charge under pH 7	7.21	7.21

Table 1b: Composition of nGCNT1 and mGCNT1 by analysis tool

Amino acid	Quantity		%	
	nGCNT1	mGCNT1	nGCNT1	mGCNT1
A	23	23	5.38	5.38
C	9	9	2.11	2.11
G	15	15	3.51	3.51
V	36	37	8.42	8.65
L	45	45	10.52	10.52
I	22	21	5.15	4.91
M	15	15	3.51	3.51
S	30	30	7.01	7.01
T	17	17	3.98	3.98
F	19	19	4.44	4.44
Y	21	21	4.91	4.91
W	8	8	1.87	1.87
D	22	22	5.15	5.15
E	28	28	6.55	6.55
N	17	17	3.98	3.98
Q	12	12	2.81	2.81
H	13	13	3.04	3.04
K	32	32	7.48	7.48
R	24	24	5.61	5.61
P	20	20	4.68	4.68

Shaded regions indicate changes between nGCNT1 and mGCNT1. Amino acids are represented in single letter code.

Differences were also observed in the modeled structures of nGCNT1 and mGCNT1 as illustrated in Fig. 1. The number of atoms, total bonds and H-bonds in nGCNT1 exceeded that in mGCNT1 while the number of turns in mGCNT1 was more than that in nGCNT1. The

number of groups, helices and strands remained the same in both nGCNT1 and mGCNT1.

To which extent mutations affected the stability had been one of the basic questions in protein structural studies. Prediction of protein stability free energy change (DDG) upon single point mutation might help the annotation process. The negative value of DDG affirmed the decrease in the stability of GCNT1 on mutation as shown in the I-Mutant outputs in Fig. 2.

Native **Mutant**

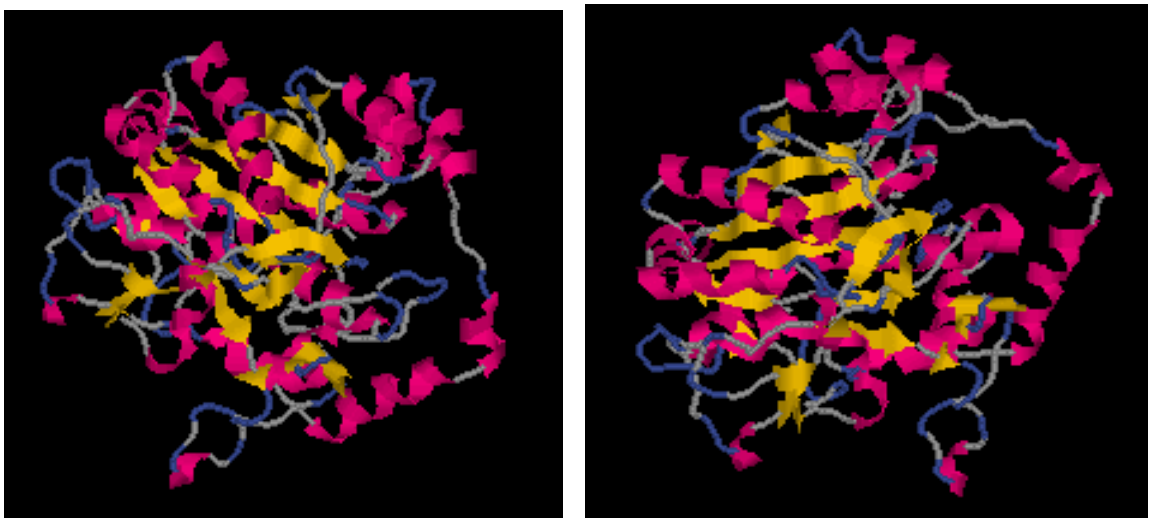
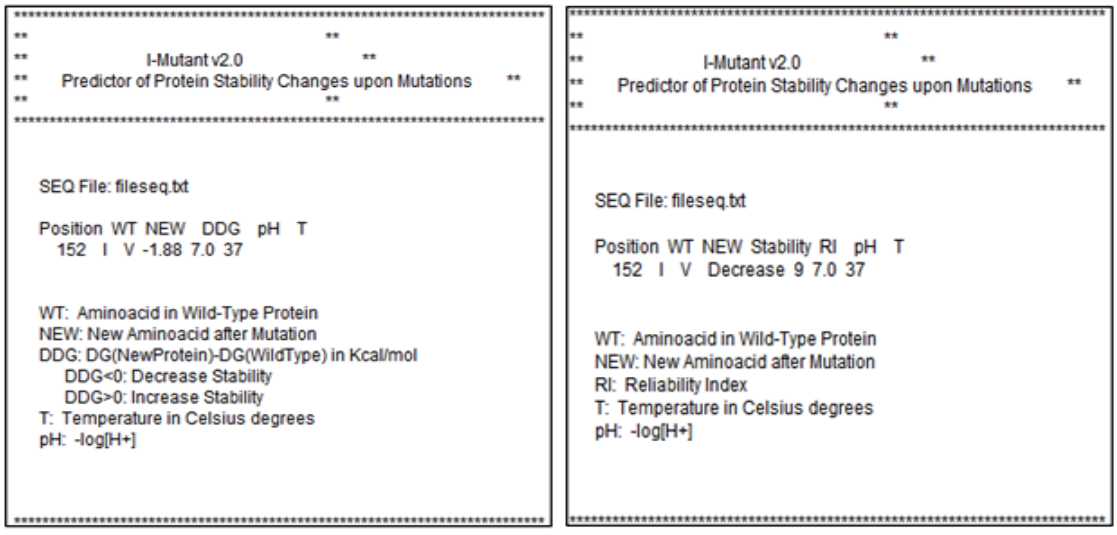


Fig. 1: 3D cartoon models of nGCNT1 and mGCNT1 as deciphered by Rasmol. Native model represents 428 groups, 3504 atoms and 3592 bonds with 21 helices, 19 strands, 39 turns and 262 H-bonds. Mutant model represents 428 groups, 3503 atoms and 3591 bonds with 21 helices, 19 strands, 40 turns and 259 H-bonds.



(a) **(b)**

Fig. 2: Stability changes of GCNT upon mutation. Negative value of DDG (a) affirms the decrease in the stability (b).

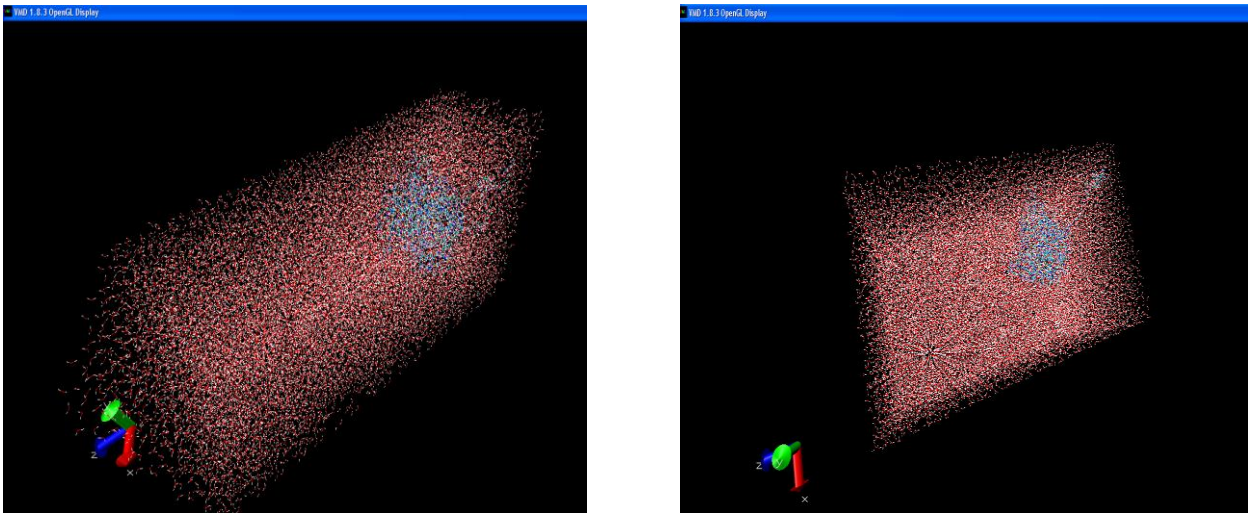
Molecular Dynamics studies probed the relationship between the molecular structure, movement and function of GCNT1. Molecular dynamics simulations were performed with GROMACS to obtain the structural information of nGCNT1 and mGCNT1 at 37 ^{o}C in aqueous environment.

The topology of the proteins was generated with the detailed description of all the atom types and charges on the atoms as well as the bonds, angles, and dihedrals. The "pdb2gmx" command converted the .pdb file into .gro file.

For simulation to take a reasonable amount of time the size of the system was restricted by placing the molecule in a box by "editconf". Only forces within this box were calculated. The dimensions were set to the diameter of the system. Generally, the SPC water model performed the best in water box simulations. The native and mutant protein models of GCNT1 in their respective simulated solvent boxes were visualized by VMD as depicted by Fig. 3.

Energy minimization was performed to remove overlapping atoms, to reduce steric clashes in the starting structure, thereby lowering forces and therefore potential energy in the system. The file em.mdp contained the parameters required for the different steps in preparing and running the energy minimization simulation.

Position restraining further prepared the system by holding the target molecule stable but allowing the water to settle around it. The file pr.mdp contained the parameters required for the different steps in preparing and running the position restraint simulation. The file md.mdp contained the parameters required for the different steps in preparing and running the molecular dynamics simulation.



(a) nGCNT1 **(b) mGCNT1**

Fig. 3: VMD models of (a)nGCNT1 (b)mGCNT1 entrapped in simulated water boxes

Energy components were extracted from an energy file by g_energy. An .xvg file was created which had the energy changes for 0-10 ps. This program read the energy file, which was produced during the simulation. Graph plotted for the obtained energy values suggested a few changes in the energy level during the time interval. Plotting the graph for the obtained values indicated a decrease in energy from the initial structure to the simulated structure in nGCNT1 whereas in mGCNT1, energy increased as shown in Fig. 4a. Also, the energy of nGCNT1 seemed to be lesser than that of mGCNT1 both at the beginning and at the end of the simulation. This reflected the decrease in the stability of mGCNT1 as validated by I-Mutant.

Analysis of helical properties upon simulation suggested differences in the secondary structure helix of GCNT1 and its mutant form. Initially, the peptide was checked for the longest helical part. This was determined by H-bonds and Phi/Psi angles. That bit was fitted to an ideal alpha helix around the Z-axis and centered on the origin for computing the helical properties. Fluctuations were observed in all the analyses in both nGCNT1 and mGCNT1 during the simulations between 0 to 10 ps.

Total helix length obtained from the file len-ahx.xvg was the average rise times the number of helical residues. Helix length decreased in both the protein from 0 to 10 ps as shown in Fig. 4b, though the value was more for nGCNT1 than mGCNT1.

Helix radius given by the file radius.xvg was the RMSD in two dimensions for all alpha C atoms. For an ideal helix the radius was 0.23 nm. Radius of helix of nGCNT1 structure both at the beginning and the end of simulation was same as that of an ideal helix while for mGCNT1; both the stages showed deviations with more deviation towards end as plotted in Fig. 4c.

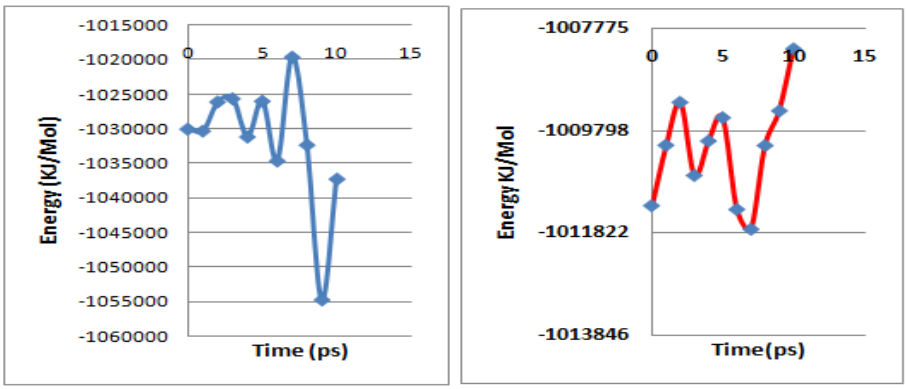
RMSD from ideal helix was calculated for the alpha C atoms only and obtained by the file rms-ahx.xvg. While the initial and final values remained the same for nGCNT1, the final value was more than the initial in case of mGCNT1 and the values were greater than that of nGCNT1 as depicted in Fig. 4d, indicating greater deviation of mGCNT1's secondary structure from that of an ideal helix.

RMSD was a measure of the structural change in the protein structures during molecular dynamics simulations. It was also used as a quantifier of the stability of the system. RMSD of the initial and simulated structures of the proteins were computed by g_confrms after least-squares fitting of the second structure on the first one. The value was found to be 0.155733 Å for the native protein and 0.159724 Å for the mutant protein. The value being more for the mutant suggested the decrease in its stability during the simulations.

The early view of proteins as relatively rigid structures had been replaced by dynamic models in which the internal motions and resulting conformational changes played an essential role in their function [24]. The results of the molecular dynamics study showed different paths of deviation in nGCNT1 and mGCNT1 reflecting the changes in the protein conformation upon mutation and justifying the biochemical variations between the two.

Single nucleotide polymorphisms (SNPs) had been found to have a profound effect on the structure and function of the resulting proteins. It was estimated that approximately 25% of SNPs in the human population might be deleterious to protein function [25]. Another study had reported that 80% of the disease associated SNPs resulted in protein destabilization [26]. SNP of A to G and the subsequent mutation of isoleucine to valine in the sequence was responsible for the observed changes in the structure and properties of GCNT1.

(a)



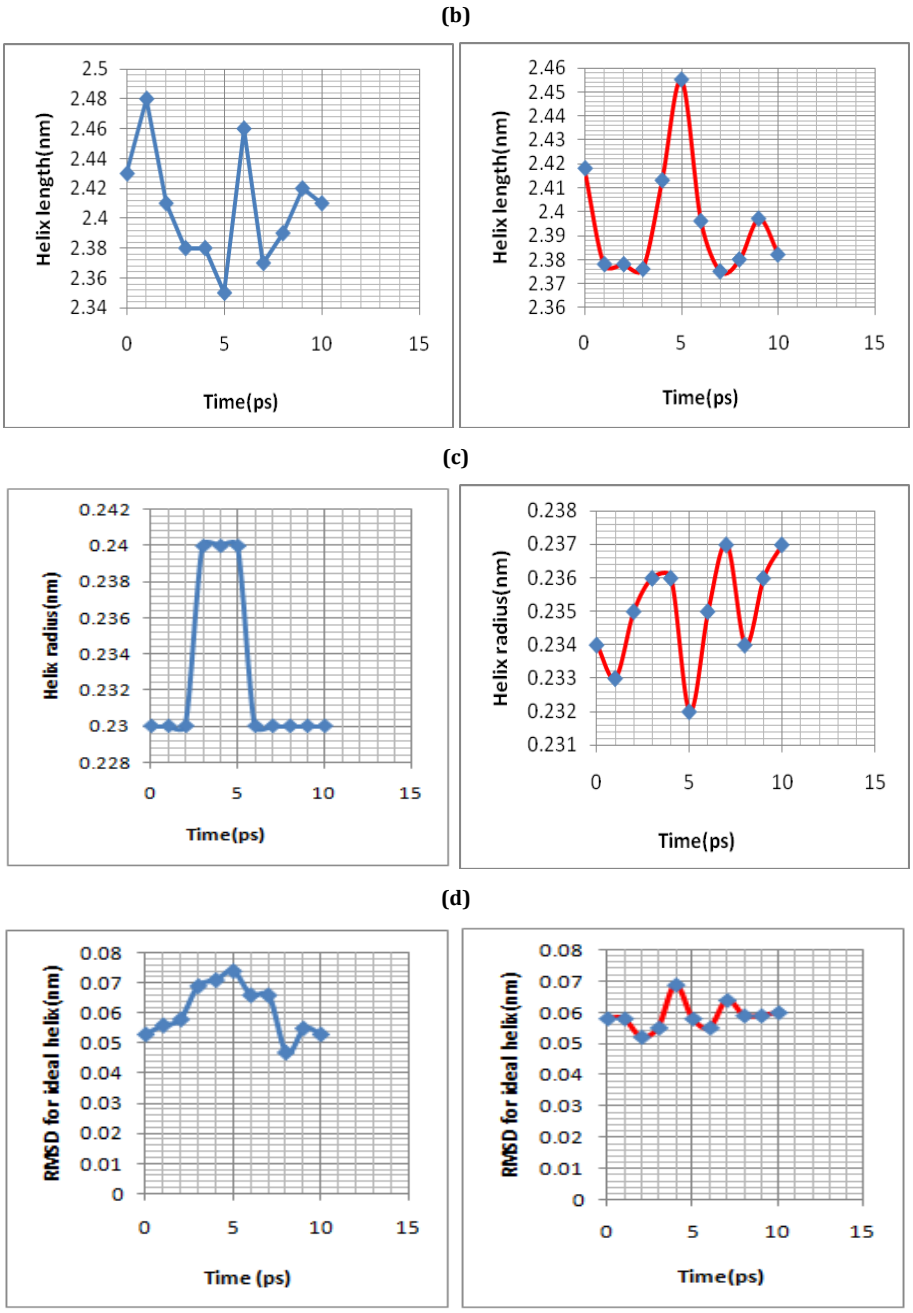


Fig. 4: Changes in (a)Energy (b)Helix length (c)Helix radius (d)RMSD from ideal helix of nGCNT1 (blue) and mGCNT1(red) during molecular dynamics simulations.

CONCLUSION

Deleterious mutations in proteins modified the physicochemical properties of amino acids that affected their stability and dynamic behaviour. Further, these mutations would disrupt the binding sites of the proteins and prohibit their interaction with ligands. Comparison of the simulation results proved the biochemical variations that existed between the native GCNT1 and mutant GCNT1, and suggested that the transition of GCNT1 structure upon mutation might play an essential role in increasing the risk of prostate cancer. Further, the molecular dynamics simulation study of GCNT1 at various conditions would help to understand the association of molecular changes in GCNT1 with the progression of the disease. Besides emphasizing the impact of mutation on protein conformation and biochemistry, this study revealed the implications these *in silico* molecular simulations had to support the observed changes.

ACKNOWLEDGEMENT

The authors thank Meenakshi College for Women, Chennai and Helixinfosystems, Chennai for their support in carrying out the work.

REFERENCES

1. Rumki Nath, Saswati Roy, Biplab De, Dutta Choudhury M. Anticancer and antioxidant activity of croton: A review. Int J Pharm Pharm Sci 2013; 5 Suppl 2:63-70.
2. Sonika Redhu, Anil Jindal. Molecular Modelling: A new scaffold for Drug Design. Int J Pharm Pharm Sci 2013; 5 Suppl 1:5-8.
3. Brockhausen I. Pathways of O-glycan biosynthesis in cancer cells. Biochem Biophys Acta 1999; 1473 :67-95.
4. Nakayama J, Shimizu F, Katsuyama T. Glycosyltransferase genes as tumor marker. Rinsho Byori 2002; Suppl 123:142-148.

5. Hagisawa S, Ohyama C, Takahashi T, Endoh M, Moriya T, Nakayama J, et al. Expression of core 2 {beta} 1, 6-N-acetylglucosaminyltransferase facilitates prostate cancer progression. Glycobiology 2005; 15 (10):1016–1024.
6. Summar F Siddiqui, John Pawelek, Tamara Handerson, Chen-Yong Lin, Robert B Dickson, David L Rimm, et al. Coexpression of ß1,6-N-Acetyl glucosaminyl transferase Glycoprotein Substrates Defines Aggressive Breast Cancers with Poor Outcome. Cancer Epidemiol Biomarkers Prev 2005; 14 :2517-2523.
7. Yagel S, Feinmesser R, Waghorne C, Lala PK, Breitman ML, Dennis JW. Evidence that beta 1-6 branched Asn-linked oligosaccharides on metastatic tumor cells facilitate invasion of basement membranes. Int J Cancer 1989; 44 :685-690.
8. Koya D, Dennis JW, Warren CE, Takahara N, Schoen FJ, Nishio Y, et al. 1999. Overexpression of core 2 N-acetylglycosaminyltransferase enhances cytokine actions and induces hypertrophic myocardium in transgenic mice. FASEB J 1999; 13 (15):2329-2337.
9. John E Pak, Pascal Arnoux, Sihong Zhou, Prashanth Sivarajah, Malathy Satkunarajah, Xuekun Xing et al. X-ray Crystal Structure of Leukocyte Type Core 2 ß1,6-N-Acetylglucosaminyltransferase. J Biol Chem 2006; 281 (36):26693-26701.
10. Schwientek T, Nomoto M, Levery SB, Merkx G, van Kessel AG, Bennett EP, et al. Control of O-glycan branch formation. Molecular cloning of human cDNA encoding a novel beta1,6-N-acetylglucosaminyltransferase forming core 2 and core 4. J Biol Chem 1999; 274 (8):4504-4512.
11. Ohyama C, Kanto S, Kato K, Nakano O, Arai Y, Kato T, et al. Natural killer cells attack tumor cells expressing high levels of sialyl Lewis x oligosaccharides. Proc Natl Acad Sci 2002; 99 (21):13789–13794.
12. Shimodaira K, Nakayama J, Nakamura N, Hasebe O, Katsuyama T, Fukuda M. Carcinoma-associated expression of core 2 beta-1,6-N-acetylglucosaminyltransferase gene in human colorectal cancer: role of O-glycans in tumor progression. Cancer Res 1997; 57 (23):5201-5206.
13. Lynch TP, Ferrer CM, Jackson SR, Shahriari KS, Vosseller K, Reginato MJ. Critical role of O-Linked β-N-acetylglucosamine transferase in prostate cancer invasion, angiogenesis, and metastasis. J Biol Chem 2012; 287 (14):11070-11081.
14. Doss CGP, NagaSundaram N. Investigating the Structural Impacts of I64T and P311S Mutations in APE1-DNA Complex: A Molecular Dynamics Approach. PLoS one 2012; 7 (2):e31677.
15. Wang L, Mitoma J, Tsuchiya N, Narita S, Horikawa Y, Habuchi T, et al. An A/G polymorphism of core 2 branching enzyme gene is associated with prostate cancer. Biochem Biophys Res Commun 2005; 331 (4):958-963.
16. Eswar N, Marti-Renom MA, Webb B, Madhusudhan MS, Eramian D, Shen M, et al. Comparative protein structure modeling with Modeller. Current Protocols in Bioinformatics 2006; Suppl 15:5.6.1-5.6.30. URL: www.salilab.org/modeller
17. Klimenkov Vitaly. Biochemistry - online; 2003. URL: http://vitalonic.narod.ru/ biochem/
18. Capriotti E, Fariselli P, Casadio R. I-Mutant2.0: predicting stability changes upon mutation from the protein sequence or structure. Nucleic Acids Res 2005; 33 :W306-W310. URL: http://gpcr.biocomp.unibo.it/cgi/predictors/I-Mutant2.0/I-Mutant2.0.cgi.
19. Lindahl E, Hess B, van der Spoel D. GROMACS: A package for molecular simulation and trajectory analysis. J Mol Mod 2001; 7:306-317. URL: http://www.gromacs.org
20. Roger Sayle, James Milner-White E. RasMol: Biomolecular graphics for all. Trends in Biochemical Sciences 1995; 20 (9):374-376.
21. Herbert J Bernstein. Recent changes to RasMol, recombining the variants. Trends in Biochemical Sciences 2000; 25 (9):453-455.
22. Cho, Hyun Sun, Dashdorj, Naranbaatar, Schotte, Friedrich, et al. Protein structural dynamics in solution unveiled via 100-ps time-resolved x-ray scattering. Proceedings of the National Academy of Sciences of the United States of America 2010; 107 :7281-7286. URL: http://www.ks.uiuc.edu
23. Inbathamizh L, Padmini E. In silico Studies On the Inhibitory Effects of Calcitriol and 5, 5'-Dithiobis-2-Nitrobenzoic Acid on Human Glucosaminyl N-Acetyl Transferase 1 Activity. Asian J Exp Biol Sci 2012a; 3 (1):14-21.
24. Martin Karplus, Andrew McCammon. Molecular dynamics simulations of biomolecules. Nat Struct Mol Biol 2002; 9 (9):646–652.
25. Yue P, Moult J. Identification and analysis of deleterious human SNPs. J Mol Biol 2006; 356 :1263–1274.
26. Wang Z, Moult J. SNPs, protein structure, and disease. Hum Mutat 2001; 17 (4):263–270.