ORIGINAL ARTICLE

In silico analysis of detrimental mutations in ADD domain of chromatin remodeling protein ATRX that cause ATR-X syndrome: X-linked disorder

P. Chandrasekaran · C. George Priya Doss ·
J. Nisha · Rao Sethumadhavan · V. Shanthi ·
K. Ramanathan · R. Rajasekaran

Abstract Many biological functions involve specific interactions of proteins. Mutations in ATRX gene can change the sequence and structure of a protein thereby impairing its function. Thus, the dysfunction of chromatin remodeling protein ATRX as a result of amino acid substitution in ADD domain often underlies the human disease, ATR-X syndrome. In general, it is mainly caused by amino acid substitution that interfered with the interactions of interest at the interface level. Hence, the study of protein–protein interactions and the interface that mediate the interactions stand important for understanding of biological function. In this work, we address the loss of function in chromatin remodeling protein ATRX due to loss of structural stability that affect the functional activity in mutant ADD domain by 17 missense mutations viz., G175E, N179S, P190A, P190L, P190S, L192F, V194I, C200S, Q219P, C220R, C220Y, W222S, C243F, R246C, R246L, G249C, and G249D. Furthermore, the loss of binding affinity of ADD domain with their interacting partner namely histone H3-peptide were investigated by (1) computing the RMSD (root mean square deviation) for the ADD domain of native with all the 17 mutants, (2) computing intra-molecular interactions in ADD domain of native with all the mutants, (3) computing binding affinity of native and mutant structures of ADD domain with histone H3-peptide through docking studies, and (4) cross validating the loss of function on binding affinity through inter-molecular interactions and normal mode analysis. Finally, as from our computational result, we concluded that all the parameters mentioned above used for studying ADD domain of mutant structures showed the decreased potential values compared to native structure that underlie ATR-X syndrome.

1 Introduction

The ability of proteins to form specific, stable complexes with other proteins is fundamental to many biological processes (Routledge et al. 2009). Mutations in human genes can change the sequence and structure of a protein, impair its function, and could lead to disease (Steward et al. 2003). Wide ranges of highly debilitating diseases are associated with the failure of proteins to maintain their native structures (Goldman et al. 1997). The alpha thalassemia/mental retardation X-linked (ATRX) protein belongs to the Switch 2, sucrose non-fermenting 2 (SWII/SNF2) ATP-dependent helicase families that exhibit chromatin remodeling activity (Iwase et al. 2011; Baumann and De La Fuente 2009; De La Fuente et al. 2011). The ATRX gene is located on the X chromosome at Xq13. It is 300 kb large gene spanning and contains 36 exons (Dhayalan et al. 2011; Wong et al. 2010). The ATRX protein has a helicase/ATPase

P. Chandrasekaran · R. Sethumadhavan · V. Shanthi ·
K. Ramanathan · R. Rajasekaran (✉)
Bioinformatics Division, School of Biosciences and Technology,
Vellore Institute of Technology, Vellore 632014,
Tamil Nadu, India
e-mail: rrajasekaran@vit.ac.in

C. G. P. Doss
Medical Biotechnology Division, School of Biosciences
and Technology, Vellore Institute of Technology,
Vellore 632014, Tamil Nadu, India

J. Nisha
Department of Biotechnology, Holy Cross College,
Trichy 620002, Tamil Nadu, India

domain at its carboxyl terminus and this domain is characteristic of the catalytic subunits of chromatin remodeling complexes. At its amino terminus, the ATRX protein has a domain composed of a Plant Homeo Domain (PHD) and a GATA-like zinc-finger motif, known as the ATRX-DNMT3A-DNMT3L (ADD) domain because this configuration of Zinc fingers has been found in the DNA-methyltransferases DNMT3A and DNMT3L (Valadez-Graham et al. 2012; Xue et al. 2003). Mutations within the coding region of the ATRX gene cause ATR-X syndrome in human patients, an X-linked genetic disease, which is characterized by variable combinations of severe mental retardation, characteristic dysmorphic facial features, alpha thalassemia, seizures, urogenital abnormalities, and sex reversal (Baumann and De La Fuente 2009; Berube et al. 2008; Baker et al. 2008). ATRX is a large protein (2,492 residues) in which almost all non-truncating mutations associated with ATR-X syndrome fall within 97 conserved residues of the ADD domain or 733 conserved residues of the Snf2 domain. Such pathogenic mutations are not seen in the remaining poorly conserved, structurally disordered 1,662 residues of ATRX. It seems likely that mutations in these regions are not observed in ATR-X syndrome because they act as neutral polymorphisms rather than lethal mutations (Mitson et al. 2011). The association of ATRX mutations with a reduction in alpha globin synthesis in alpha thalassemia patients suggests that the protein plays a role in the regulation of alpha globin gene expression (Wong et al. 2010). Genome-wide analysis has shown that in euchromatin the predominant targets of ATRX are sequences containing VNTRs (variable number of tandem repeats). Many of these are G and C rich with high proportion of CpG dinucleotides. These observations explain why ATRX mutations affect the α-globin cluster but not the β-globin cluster that cause α-thalassaemia. The α-cluster lies in a GC-rich subtelomeric region containing a high density of CpG islands and G-rich TRs (tandem repeats) whereas the β-globin cluster has none of these features (Law et al. 2010). The ADD domain of the ATRX protein is a hot spot for mutations causing the ATR-X syndrome (Dhayalan et al. 2011). The disease causing mutations fall into two groups viz., the majority affect buried residues thereby affecting the structural integrity of the ADD domain and the another group affects a cluster of surface residues which are likely to perturb a potential protein interaction site. The effects of individual point mutations on the folding state and the stability of ADD domain correlate well with the levels of mutant ATRX protein in patients, providing insights into the molecular pathophysiology of ATR-X syndrome (Argentaro et al. 2007).

In the present study, we computationally performed the sequence and structure-based mutagenesis to explore the function of ADD domain of ATRX by investigating its potential interaction with histone H3-peptide through inter–intra molecular interactions, docking studies and normal mode analysis.

2 Materials and methods

2.1 Data sets

The protein sequence and variants (single amino acid polymorphisms/missense mutations/point mutations/nsSNP) for ADD domain were obtained from the Swissprot database (Yip et al. 2008) available at http://www.expasy.ch/sprot/ to find out the detrimental point mutants. The 3D Cartesian coordinates were obtained from Protein Data Bank for in silico mutation modeling and docking studies based on detrimental point mutations.

2.2 Prediction of protein stability upon single amino acid substitution using support vector machine-based tool, I-Mutant 2.0

We predicted nsSNP causing protein stability change using I-Mutant 2.0 (Capriotti et al. 2005), a support vector machine (SVM)-based tool for the automatic prediction of protein stability change upon single amino acid substitution available at http://gpcr.biocomp.unibo.it/cgi/predictors/I-Mutant2.0/I-Mutant2.0.cgi. Its predictions were performed starting either from the protein structure or, more importantly, from the protein sequence. The output files show the predicted free energy change value or sign (DDG), which was calculated from the unfolding Gibbs free energy value of the mutated protein minus the unfolding Gibbs free energy value of the native type. Positive DDG values mean that the mutated protein possesses high stability and vice versa.

2.3 Analyzing the functional consequence of nsSNP by sequence homology-based method, SIFT

We used the program SIFT (Ng and Henikoff 2003), to specifically detect the deleterious single amino acid polymorphism, available at http://blocks.fhcrc.org/sift/SIFT.html. Sorting intolerant from tolerant (SIFT), a sequence homology-based tool, presumed that important amino acids would be conserved in the protein family. Hence, changes at well-conserved positions tend to be predicted as deleterious. We submitted the query in the form of protein sequences. The underlying principle of this program supposed that, SIFT take a query sequence and use multiple alignments to predict tolerated and deleterious substitutions for every position of the query sequence. The SIFT

predicted whether submitted SAP nsSNP affected the protein function based on sequence homology and amino acid properties. The cut-off value was a tolerance index of ≥ 0.05. The higher the tolerance index, the lesser the functional impact of a particular amino acid substitution. Thus, the calculated score represented the likelihood of mutability at the site of amino acid substitution (Ng and Henikoff 2001).

2.4 Simulation for functional change in point mutant by structure homology-based method, PolyPhen

Analyzing the damaged point mutations at the structural level was considered to be very important to understand the functional activity of the concerned protein. We used the server PolyPhen (Ramensky et al. 2002), available at http://coot.embl.de/PolyPhen/ for this purpose. The input options for the PolyPhen server were protein sequence or SWALL database ID or accession number together with sequence position of two amino acid variants. We submitted the query in the form of protein sequence with mutational position and two amino acid variants. It analyzed the impact of nsSNP by mapping the amino acid substitutions on the protein 3D structure to explore whether the amino acid substitution was likely to destroy the hydrophobic core of the protein, the electrostatic interaction and other important features of protein. It calculated position-specific independent (PSIC) scores for each of the two variants and then computed the PSIC score difference between them. The higher the PSIC score difference, the higher was the possible functional impact of a particular amino acid substitution.

2.5 Modeling single amino acid polymorphism location on protein structure to compute total energy and RMSD

Structure analysis was performed for computing the total energy and evaluating the structural deviation between native and mutant types by means of root mean square deviation (RMSD). We used the web resource Protein Data Bank (Berman et al. 2000) to identify the 3D structure of ADD domain (PDB ID: 3QLC) of ATRX protein. We confirmed the mutation position and the mutation residue in ADD domain. The mutation was performed using SWISSPDB viewer and the energy minimization for 3D structures was performed by NOMAD-Ref server (Lindahl et al. 2006). This server used Gromacs as default force field for energy minimization based on the methods of steepest descent, conjugate gradient and L-BFGS methods (Delarue and Dumas 2004). We used the iFold server for simulated annealing, which was based on discrete molecular dynamics and was one of the fastest strategies for simulating protein dynamics. This server was also efficient in sampling the vast conformation space of biomolecules in both length and time scales (Sharma et al. 2006; Mohammed Said et al. 2012). The total energy was computed for both the native and mutants of ADD domain by GROMOS96 (Schuler et al. 2001) implemented in DeepView. Divergence in mutant structure with native structure was due to mutations, deletions, and insertions (Han et al. 2006; Rajasekaran and Sethumadhavan 2010a) and the deviation between the two structures could alter the functional activity (Varfolomeev et al. 2002) which was evaluated by their RMSD values.

2.6 Computing intra- and inter-molecular interactions in ADD domain of ATRX protein

We used PIC (Protein Interactions Calculator) server (Tina et al. 2007) for computing intra- and inter-molecular interactions for both native and mutant structures available at: http://crick.mbu.iisc.ernet.in/~PIC. It accepted atomic coordinate set of a protein structure in the standard Protein Data Bank (PDB) format. Interactions within a protein structure and interactions between proteins in an assembly were essential considerations in understanding molecular basis of stability and functions of proteins and their complexes. There were several weak and strong interactions that render stability to a protein structure or an assembly. It computes various interactions such as interaction between apolar residues, disulphide bridges; hydrogen bond between main chain atoms; hydrogen bond between main chain and side chain atoms; hydrogen bond between two side chain atoms; interaction between oppositely charged amino acids (ionic interactions), aromatic–aromatic interactions, aromatic–sulfur interactions and cation–π interactions.

2.7 Docking between ADD domain and histone H3-peptide

Protein–peptide interactions were among the most prevalent and important interactions in the cell that play a key role in major cellular processes, including signal transduction, transcription regulation, cell localization, protein degradation, and immune response. We used Rosetta FlexPepDock web server (London et al. 2011) (http://flexpepdock.furmanlab.cs.huji.ac.il/), a high-resolution protocol for the refinement of protein–peptide complex structures implemented in the Rosetta modeling suite frame-work. Started from a coarse model of the interaction, FlexPepDock performed a Monte Carlo-Minimization-based approach to refine all the peptide's degrees of freedom (rigid body orientation, backbone and side chain flexibility) as well as the protein receptor side chain

conformations. The modeling of a protein–peptide inter-action was divided into several consecutive steps, each representing a smaller subproblem, in line with prevalent approaches for modeling and docking of globular proteins: (1) model the receptor structure, (2) predict potential binding sites on the receptor surface, (3) model the peptide backbone in the binding site, and (4) refine the peptide–protein complex to high resolution.

2.8 Exploring the flexibility of binding pocket by normal mode analysis

A quantitative measure of the atomic motions in proteins was obtained from the mean square fluctuations of the atoms relative to their average positions. These were related to the B-factor (Yuan et al. 2005). Analysis of B-factors, therefore, was likely to provide newer insights into protein dynamics, flexibility of amino acids, and protein stability (Parthasarathy and Murthy 2000). It was to be noted that, protein flexibility was important for protein function and for rational drug design (Carlson and McCammon 2000). Also, flexibility of certain amino acids in protein was useful for various types of interactions. Moreover, flexibility of amino acids in binding pocket was considered to be a significant parameter to understand the binding efficiency. In fact, loss of flexibility impaired the binding effect (Hinkle and Tobacman 2003) and vice versa. Hence, this was analyzed by the B-factor, which was computed from the mean square displacement $<R^2>$ of the

lowest-frequency normal mode using ElNémo server (Suhre and Sanejouand 2004).

3 Results and discussion

3.1 SNP dataset

The ADD domain and a total of 17 variants namely G175E, N179S, P190A, P190L, P190S, L192F, V194I, C200S, Q219P, C220R, C220Y, W222S, C243F, R246C, R246L, G249C, and G249D investigated in this work were retrieved from Swissprot database.

3.2 Identifying the detrimental missense mutations with well-recognized computational tools

In this study, we focused our attention to investigate the functional impact of amino acid substitution in ADD domain associated with ATR-X syndrome using existing computational methods I-Mutant 2.0, SIFT and PolyPhen. Over the past 5 years, computational approaches to in silico analysis of amino acid substitutions on 3D structure of protein to understand the structural and functional impact of mutant structure of protein have improved considerably. In this aspect, we have brought in silico model with two diverse approaches that would be helpful to experimental biologists as an alternative method in determining the functional nsSNPs in ATR-X syndrome. In this analysis,

Table 1 Functionally significant detrimental missense mutations by total energy, RMSD, I-Mutant2.0, SIFT, and PolyPhen

Variants	Total energy (KJ/mol)	RMSD (Å)	DDG	Tolerant index	PSIC SD
Native	−8,991.874	–	**–**	**–**	**–**
G175E	−8,546.884	0.29	**−0.63**	**0.00**	**0.999**
N179S	−8,363.465	0.75	**−1.46**	**0.01**	**0.193**
P190A	−8,571.004	0.51	**−3.11**	**0.01**	**1.000**
P190L	−8,516.848	0.61	**−1.33**	**0.01**	**1.000**
P190S	−8,540.866	0.93	**−2.91**	**0.00**	**1.000**
L192F	−8,549.327	1.11	**−0.67**	**0.00**	**1.000**
V194I	−8,606.086	1.15	**−0.83**	**0.00**	**0.991**
C200S	−8,614.933	1.12	**−3.27**	**0.00**	**1.000**
Q219P	−8,263.912	0.44	**−2.16**	**0.02**	**0.997**
C220R	−8,569.088	0.59	**−2.65**	**0.00**	**1.000**
C220Y	−8,562.291	0.81	**−2.60**	**0.00**	**1.000**
W222S	−8,496.043	0.98	**−2.14**	**0.00**	**1.000**
C243F	−8,521.195	0.69	**−0.84**	**0.00**	**1.000**
R246C	−8,217.157	0.78	**−0.53**	**0.00**	**1.000**
R246L	−8,268.191	0.93	**−0.94**	**0.00**	**0.997**
G249C	−8,580.952	0.52	**−1.66**	**0.00**	**1.000**
G249D	−8,530.296	0.68	**−1.83**	**0.01**	**1.000**

Letters in bold indicate less stable, deleterious and damaging by I-Mutant, SIFT and PolyPhen, respectively
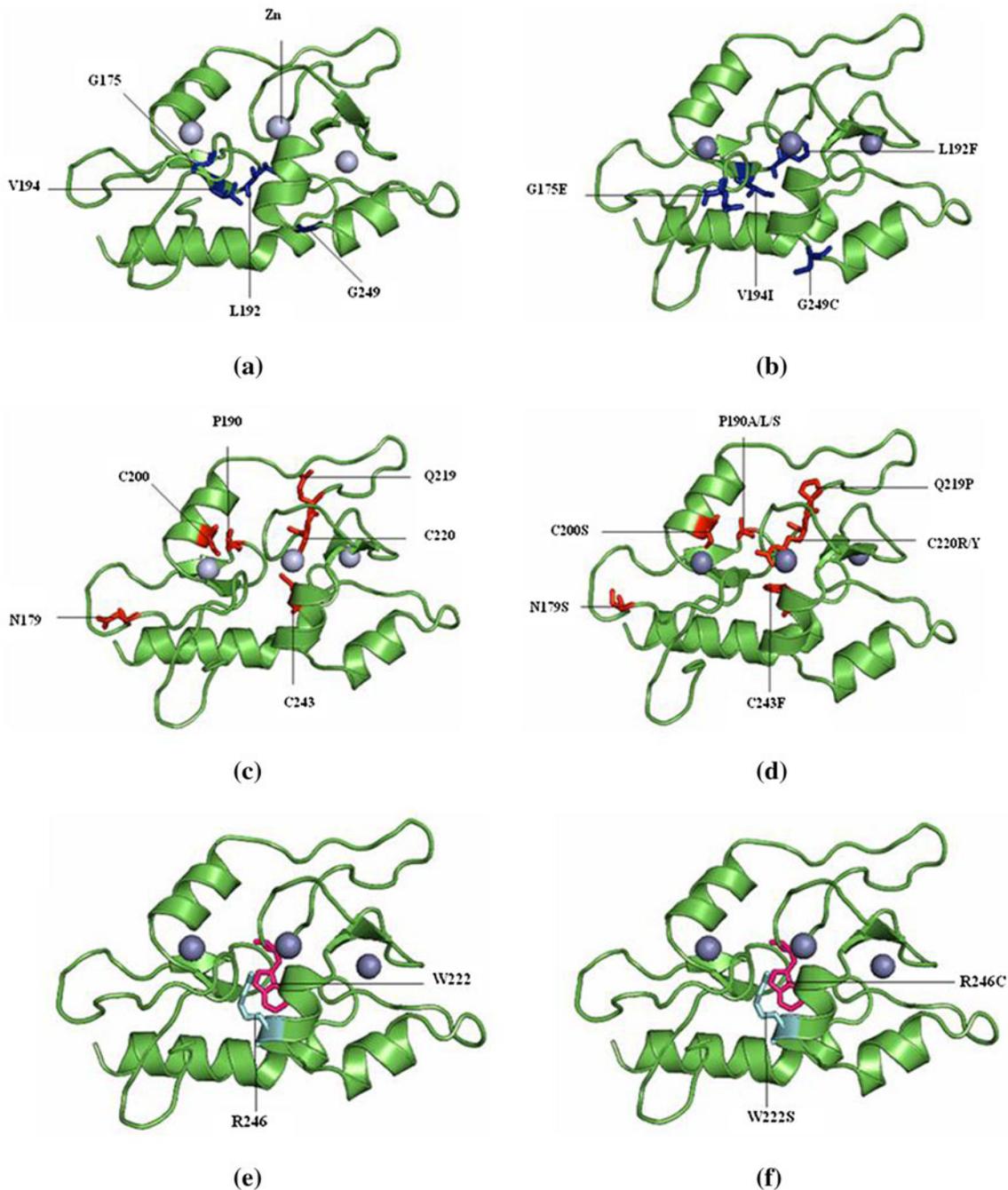
Fig. 1 Shows the native and mutated residues of ADD domain of ATRX protein represented as sticks. **a** Non-polar amino acid residues (*G175*, *L192*, *V194*, and *G249*) (*blue*) of native ADD domain and **b** their corresponding mutated amino acids (*E175*, *F192*, *P194*, and *G249C*) (*blue*) at their respective position. **c** Polar amino acids (*N179*, *P190*, *C200*, *Q219*, *C220*, and *C243*) (*red*) of native ADD domain and **d** their corresponding mutated amino acids (*S179*, *A190*, *S200*, *P219*, *R220*, and *F243*) (*red*) at their respective position. **e** Aromatic and positively charged amino acid residues (*W222* (*magenta*) and R246 (*cyan*)) of native ADD domain, and **f** their corresponding mutated amino acid residues (*S222* (*magenta*) and *C246* (*cyan*)) at their respective position (color figure online)

we used sequence (SIFT) and structure-based methods (I-Mutant 2.0 and PolyPhen), the most common approaches used in SNP prediction. According to I-Mutant 2.0, more negative the free energy value (DDG value), less stable the given point mutation was likely to be. Out of 17 variants, 5 variants (C200S, P190A, P190S, C220R, and C220Y) which showed DDG values of −3.27, −3.11, −2.91, −2.65, and −2.60, respectively, were considered to be less stable and deleterious as listed in Table 1. The other 12 variants showed the DDG values ranging from −0.53 to
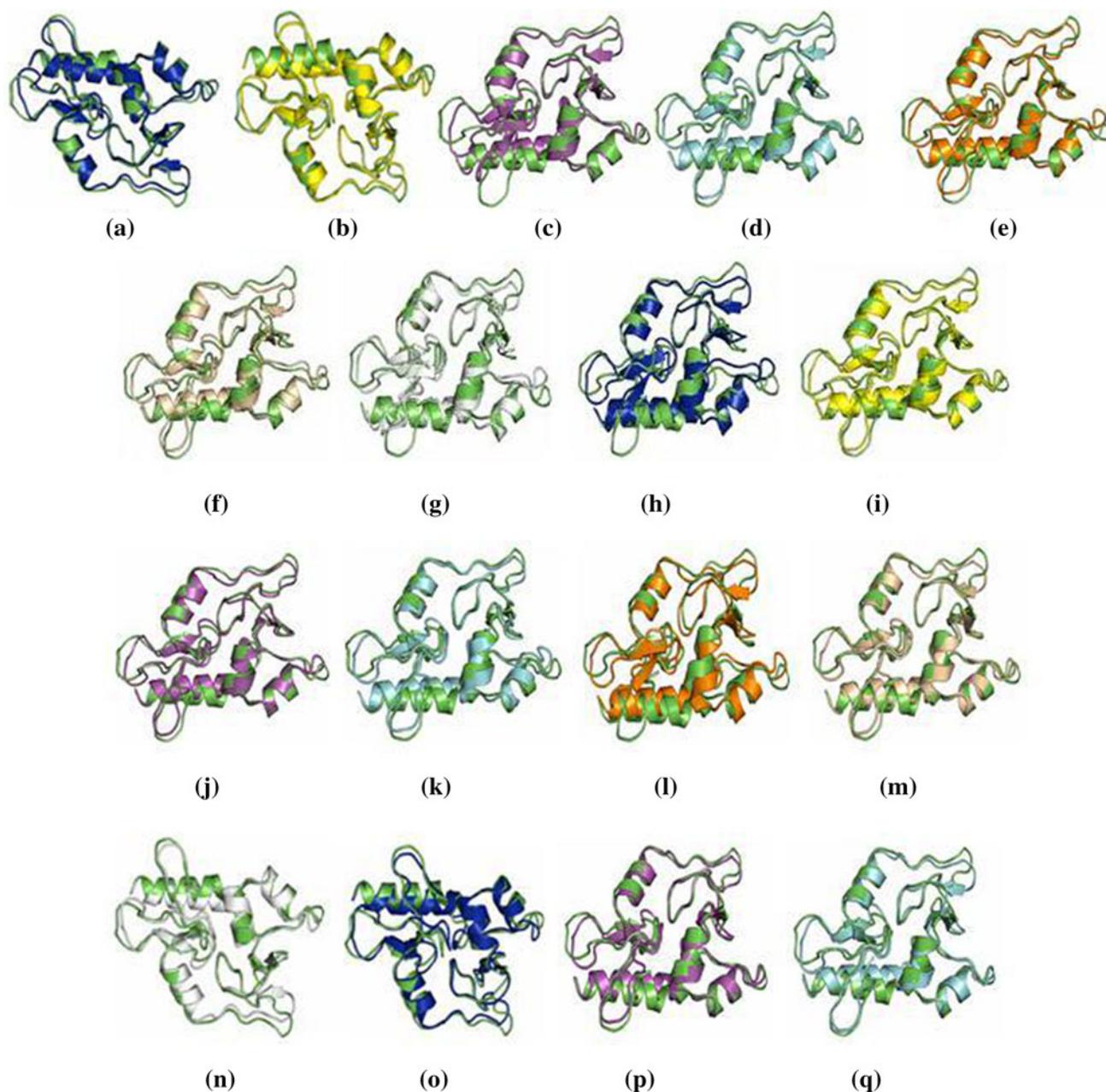
Fig. 2 Superimposed structure of native ADD domain (*green*) with all 17 mutant structures of ADD domain. **a** ADD domain mutant G175E (*blue*). **b** ADD domain mutant N179S (*yellow*). **c** ADD domain mutant P190A (*magenta*). **d** ADD domain mutant P190L (*cyan*). **e** ADD domain mutant P190S (*orange*). **f** ADD domain mutant L192F (*tint*). **g** ADD domain mutant V194I (*grey*). **h** ADD domain mutant C200S (*blue*). **i** ADD domain mutant Q219P (*yellow*). **j** ADD domain mutant C220R (*magenta*). **k** ADD domain mutant C220Y (*cyan*). **l** ADD domain mutant W222S (*orange*). **m** ADD domain mutant C243F (*tint*). **n** ADD domain mutant R246C (*grey*). **o** ADD domain mutant R246L (*blue*). **p** ADD domain mutant G249C (*magenta*). **q** ADD domain mutant G249D (*cyan*) (color figure online)

−2.16. Moreover, we categorized these 17 non-synonymous variants into the following groups such as non-polar, polar, aromatic and positively charged amino acids based on their physio-chemical properties as shown in Fig. 1. Of the 17 variants showed a negative $\Delta\Delta G$, two variants (C220Y and C243F) changed from polar amino acids to

aromatic amino acids, two variants (P190A and P190L) changed from polar to non-polar, two variants (G175E and G249D) changed from non-polar to negatively charged followed by six variants viz., L192F, G249C, C220R, R246L, R246C, W222S changed their amino acids from non-polar to aromatic, non-polar to polar, polar to

positively charged, positively charged to non-polar, positively charged to polar, and aromatic amino acid to polar amino acids, respectively. In addition, four variants (N179S, P190S, C200S, and Q219P) retained its polar amino acids and one more variant V194I retained its non-polar amino acid. Indeed, by considering only amino acid substitutions based on physico-chemical properties, we could not be able to identify the detrimental effect. Rather, by considering the sequence conservation along with the above said properties could have more advantages and reliable to find out the detrimental effect of missense mutations.

The conservation level of a particular position in a protein was determined using a sequence homology-based tool, SIFT. Protein sequences of 17 variants were submitted independently to SIFT program to check its tolerance index. Higher the tolerance index, lesser the functional impact, a particular amino acid substitution was likely to have and vice versa. Out of 17 variants, 12 variants, G175E, P190S, L192F, V194I, C200S, C220R, C220Y, W222S, C243F, R246C, R246L, and G249C had the tolerance index score of 0.00, the other five variants, N179S, P190A, P190L, G249D had the tolerance index score of 0.01 followed by one variant Q219P with tolerance index score of 0.02 as outlined in Table 1. A lower score 0.00–0.05 indicated that the nsSNPs were more damaging to protein function. Out of 17 variants, 12 variants mentioned above had the less tolerance index score and were more deleterious. Such scores enabled the quantitative comparison and ranking of SNPs in the order of their biological significance, and were useful for biologist to decide which SNPs of a gene they should first look at.

The structural levels of alteration have been determined by applying PolyPhen program. Protein sequence with mutational position and amino acid variants associated to 17 single point mutants investigated in this work had been submitted as input to the PolyPhen server and the results as shown in Table 1. A PSIC score difference of 0.00 and above was considered to be damaging. A total of 17 variants were considered to be damaging by PolyPhen and these 17 variants exhibited a PSIC score difference between 0.193 and 1.000. It was interesting to observe that both sequence and structure-based computational methods were predicted as these 17 variants were less stable, deleterious and damaging by I-Mutant 2.0, SIFT and PolyPhen, respectively. Therefore, we considered all these 17 variants as detrimental based on our results obtained by well-known computational tools. Further, we focused our attention on these 17 variants being mapped into the ADD domain to investigate the structural analysis by comparing total energy, RMSD, inter–intra molecular interactions, docking studies and normal mode analysis of mutant type with native type structure.

Table 2 Intra-molecular interactions in ADD domain of ATRX protein

Variants	H–I	MC–MC	MC–SC	SC–SC	I–I	A–A	A–S	C–π	Total
Native	79	168	121	103	8	6	3	3	491
G175E	74	166	106	103	8	6	3	3	469
N179S	73	160	113	88	7	6	2	2	451
P190A	72	166	100	98	7	6	3	2	454
P190L	73	163	100	99	6	6	4	2	453
P190S	71	158	121	92	7	5	5	1	460
L192F	69	168	107	94	7	6	2	3	456
V194I	73	166	106	96	7	6	3	3	460
C200S	72	168	101	84	6	7	7	3	448
Q219P	75	165	100	81	7	6	3	1	438
C220R	73	160	118	101	8	6	2	2	470
C220Y	77	165	117	98	7	6	5	2	477
W222S	68	165	104	103	7	4	2	3	456
C243F	79	163	113	71	8	7	5	3	449
R246C	76	166	99	94	7	6	3	2	453
R246L	72	161	114	89	8	6	3	2	455
G249C	72	162	107	103	6	6	3	2	461
G249D	74	163	107	93	7	6	2	2	454

H–I Hydrophobic interaction within 5 Å, *MC–MC* main chain–main chain hydrogen bonds, *MC–SC* main chain–side chain hydrogen bonds, *SC–SC* side chain–side chain hydrogen bonds, *I–I* ionic interactions within 6 Å, *A–A* aromatic–aromatic interactions within 4.5 and 7 Å, *A–S* aromatic–sulfur interactions within 5.3 Å, *C–π* cation–pi interaction within 6 Å

3.3 Computing total energy and RMSD by modeling of mutant structures

Non-synonymous single nucleotide polymorphism (nsSNP) of genes introduced amino acid changes to proteins, and played an important role in providing genetic functional diversity. To understand the structural characteristics of nsSNPs, we have mapped a set of nsSNPs derived from the Swissprot database to the structural surface of ADD domain (PDB ID: 3QLN). Mutation analysis was performed based on the results obtained from I-Mutant 2.0, SIFT, PolyPhen. Mutation at specified position was performed by SWISSPDB viewer independently to get mutant structures. Energy minimization for all the mutant models and their native structures was achieved using the Nomad-Ref server, followed by simulated annealing using iFOLD server. The employment of energy minimization preserved native geometry of 3D structure of protein. In many cases, an initial structure obtained by homology modeling method, or by amino acid substitution had certain sites in which atoms were closely positioned, valent bonds were much extended, or side chains exist in unusual conformations. For regularization of this non-equilibrium spatial structures the method of simulated annealing was employed (Sharma et al. 2006). The total energy for all the mutant and native structures after minimization are listed in Table 1. The total energy for the native protein was −8,991.874 kJ/mol. Change in total energy due to

Table 3 Inter-molecular interactions of ADD domain with histone H3-peptide

Variants	Rosetta energy score (Kcal/mol)	H–I	MC–MC	MC–SC	SC–SC	I–I	Total
Native	−88.709	3	8	9	6	3	29
G175E	−83.215	3	7	8	1	3	22
N179S	−67.477	3	8	5	3	3	22
P190A	−84.918	3	6	5	1	3	18
P190L	−80.009	3	7	9	1	3	23
P190S	−75.744	3	8	4	1	3	18
L192F	−57.269	3	2	2	2	4	13
V194I	−41.790	2	2	3	1	4	12
C200S	−62.502	2	3	7	8	4	24
Q219P	−41.915	3	7	3	3	3	19
C220R	−49.535	3	7	7	4	3	24
C220Y	−74.014	–	3	8	3	4	18
W222S	−66.674	2	5	3	6	3	19
C243F	−71.016	3	7	5	2	3	20
R246C	−76.931	2	6	6	4	3	21
R246L	−73.323	3	7	6	1	3	20
G249C	−83.215	3	5	3	2	4	17
G249D	−69.991	2	3	5	6	3	19

H–I hydrophobic interaction within 5 Å, *MC–MC* main chain–main chain hydrogen bonds, *MC–SC* main chain–side chain hydrogen bonds, *SC–SC* side chain–side chain hydrogen bonds, *I–I* ionic interactions within 6 Å

mutation was noticeable in the 3QLN mutants ranging from −8,521.195 to −8,614.933 kJ/mol. Higher the total energy, lesser the stability of protein structure would be (Rajasekaran and Sethumadhavan 2010b). Since, all the 17 mutants had the total energy higher than the native structure; we considered these missense mutations to have deleterious effect based on structural stability.

RMSD was the measure of the deviation of the mutant structures from their native structure conformations. Higher the RMSD value, the more the deviation between the native type and mutant type structures. Among all the 17 mutants, mutant V194I exhibited a high RMSD value of 1.15 Å followed by two mutants C200S and L192F with 1.12 and 1.11 Å RMSD score, respectively. It was interesting to observe that these three mutants were also predicted as more deleterious and damaging according to SIFT and PolyPhen, respectively. So the sequence and structure-based analysis of ADD domain was useful to understand the structural and functional impact of protein. The remaining 14 mutants had the RMSD value ranging between 0.29 and 0.98 Å as outlined in Table 1. Assuming that the minimized native structure had ideal mutual orientation of functional groups and the RMSD values characterized the effect of amino acid substitution on spatial 3D structure architecture (atom deviations from initial structure) (Sharma et al. 2006). Structural changes, in turn, affected functional activity due to amino acid substitution disturbing the binding efficiency of ADD domain with its interacting

partner leading to ATR-X syndrome. Figure 2 illustrates the superimposed structure of native with all the 17 mutant structures of ADD domain. It was to be noted that, 8 variants, namely N179S, C220R, C220Y, C243F, R246C, R246L, G249C, and G249D (Argentaro et al. 2007; Badens et al. 2006) which were considered as detrimental missense mutations by various sequence and structure-based computational methods were well supported with experimental studies performed elsewhere.

3.4 Computing intra-molecular interactions in ADD domain of ATRX protein

Interactions within a protein structure and interactions between proteins in an assembly were essential considerations in understanding molecular basis of stability and functions of proteins and their complexes. There were several weak and strong intra-molecular interactions that rendered stability to a protein structure. Hence these intra-molecular interactions have been computed by PIC server to further substantiate the stability of protein structure. Based on our analysis, we found a total number of 491 intra-molecular interactions in native ADD domain which includes 79 hydrophobic, 392 hydrogen bonding (i.e., 168 main chain–main chains, 121 main chain–side chain, 103 side chain–side chain), 8 ionic, 6 aromatic–aromatic, 3 aromatic–sulfur, and 3 cation–π interactions. Based on our in silico mutagenesis studies, we found that all the 17
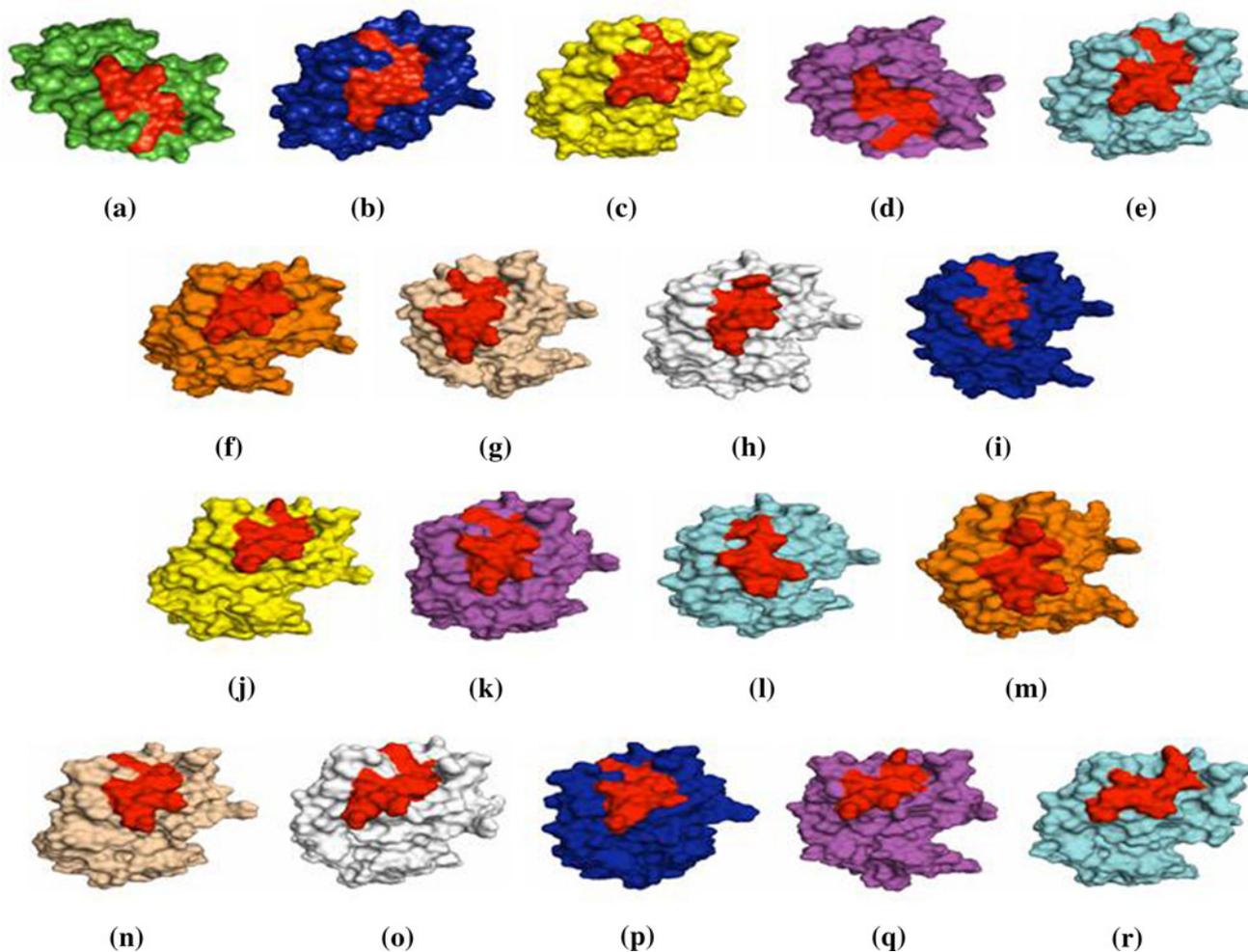
Fig. 3 Docked complex structure of ADD domain of native ATRX and all the 17 mutant structures with histone H3-peptide (*red*). **a** Native ADD domain (*green*). **b** ADD domain mutant G175E (*blue*). **c** ADD domain mutant N179S (*yellow*). **d** ADD domain mutant P190A (*magenta*). **e** ADD domain mutant P190L (*cyan*). **f** ADD domain mutant P190S (*orange*). **g** ADD domain mutant L192F (*tint*). **h** ADD domain mutant V194I (*grey*). **i** ADD domain mutant C200S (*blue*). **j** ADD domain mutant Q219P (*yellow*). **k** ADD domain mutant C220R (*magenta*). **l** ADD domain mutant C220Y (*cyan*). **m** ADD domain mutant W222S (*orange*). **n** ADD domain mutant C243F (*tint*). **o** ADD domain mutant R246C (*grey*). **p** ADD domain mutant R246L (*blue*). **q** ADD domain mutant G249C (*magenta*). **r** ADD domain mutant G249D (*cyan*) (color figure online)

mutant structures encompass the intra-molecular interactions between the ranges of 438–477 as could be seen from Table 2. From this result, we confirmed that all 17 mutant structures having decreased stability due to reduction in intra-molecular interaction compared to native type ADD domain. Moreover, we also analyzed the inter-molecular interactions between the ADD domain and histone H3-peptide for further course of investigation.

3.5 Rationale of binding efficiency for native and mutant structures of ADD domain with its interacting partner's histone H3-peptide

Large-scale identification of protein–protein interactions in functional complexes represented an efficient route to elucidate the regulatory rules of cellular functions (Du et al. 2009; Manisk Kumar and Krishna 2013). In this present study, we identified the potential interactions between the ADD domain of ATRX and histone H3-peptide to understand more about how the deleterious mutations in ADD domain affect the binding affinity with H3-peptide, lead to ATR-X syndrome. Hence, we performed protein–peptide docking by FlexPepDock server for identifying the binding affinity between them in terms of Rosetta energy score. In this regard, we selected the PDB ID: 3QLC which had ADD domain complex with H3-peptide. To understand the binding affinity of H3-peptide with both native and mutants of ADD domain, we unbound the H3-peptide from ADD domain from the PDBID: 3QLC (chain A contain ADD domain and Chain C contain Histone H3-Peptide). We

Table 4 Comparison of normalized mean square displacement of H3-peptide binding amino acids in native and mutants of ATRX

Binding residues	Native	G175E	N179S	P190A	P190L	P190S	L192F	V194I	C200S	Q219P
Normalized mean square displacement $\langle R^2 \rangle$										
Asp(212)	0.0466	0.0424[a]	0.0462[a]	0.0445[a]	0.0407[a]	0.0434[a]	0.0332[a]	0.0381[a]	0.0403[a]	0.0357[a]
Met(216)	0.0299	0.0274[a]	0.0271[a]	0.0292[a]	0.0257[a]	0.0252[a]	0.0212[a]	0.0240[a]	0.0262[a]	0.0229[a]
Asp(217)	0.0220	0.0202[a]	0.0190[a]	0.0210[a]	0.0199[a]	0.0187[a]	0.0175[a]	0.0172[a]	0.0182[a]	0.0184[a]
Glu(218)	0.0251	0.0216[a]	0.0227[a]	0.0225[a]	0.0228[a]	0.0240[a]	0.0214[a]	0.0208[a]	0.0193[a]	0.0207[a]
Gly(226)	0.0212	0.0119[a]	0.0138[a]	0.0119[a]	0.0156[a]	0.0174[a]	0.0160[a]	0.0163[a]	0.0100[a]	0.0155[a]
Gly(227)	0.0225	0.0193[a]	0.0221[a]	0.0201[a]	0.0228	0.0246	0.0219[a]	0.0221[a]	0.0148[a]	0.0211[a]
Asn(228)	0.0181	0.0145[a]	0.0150[a]	0.0150[a]	0.0171[a]	0.0177[a]	0.0156[a]	0.0154[a]	0.0070[a]	0.0168[a]
Leu(229)	0.0124	0.0098*	0.0091[a]	0.0109[a]	0.0109[a]	0.0115[a]	0.0097[a]	0.0109[a]	0.0077[a]	0.0106[a]
Ile(230)	0.0071	0.0064[a]	0.0072	0.0076	0.0076	0.0070[a]	0.0054[a]	0.0074	0.0074	0.0062[a]
Cys(231)	0.0080	0.0088	0.0076[a]	0.0099	0.0082	0.0069[a]	0.0084	0.0089	0.0101	0.0068[a]
Asp(233)	0.0143	0.0122[a]	0.0109[a]	0.0125[a]	0.0109[a]	0.0120[a]	0.0157	0.0171	0.0200	0.0116[a]
Arg(250)	0.0322	0.0277[a]	0.0307[a]	0.0280[a]	0.0316[a]	0.0325	0.0281[a]	0.0376	0.0360	0.0292[a]
Ser(254)	0.0295	0.0314	0.0225[a]	0.0224[a]	0.0326	0.0267[a]	0.0205[a]	0.0347	0.0391	0.0274[a]
Trp(263)	0.0139	0.0135[a]	0.0129[a]	0.0125[a]	0.0147	0.0146	0.0105[a]	0.0064[a]	0.0038[a]	0.0148

C220R	C220Y	W222S	C243F	R246C	R246L	G249C	G249D
Normalized mean square displacement $\langle R^2 \rangle$							
0.0355[a]	0.0421[a]	0.0391[a]	0.0342[a]	0.0406[a]	0.0480	0.0412[a]	0.0341[a]
0.0214[a]	0.0242[a]	0.0268[a]	0.0222[a]	0.0252[a]	0.0275[a]	0.0252[a]	0.0242[a]
0.0161[a]	0.0184[a]	0.0202[a]	0.0178[a]	0.0206[a]	0.0197[a]	0.0188[a]	0.0194[a]
0.0193[a]	0.0230[a]	0.0214[a]	0.0218[a]	0.0228[a]	0.0244[a]	0.0221[a]	0.0193[a]
0.0136[a]	0.0181[a]	0.0122[a]	0.0167[a]	0.0142[a]	0.0168[a]	0.0155[a]	0.0131[a]
0.0181[a]	0.0216[a]	0.0163[a]	0.0192[a]	0.0209[a]	0.0244	0.0208[a]	0.0147[a]
0.0161[a]	0.0185	0.0085[a]	0.0158[a]	0.0157[a]	0.0160[a]	0.0136[a]	0.0069[a]
0.0097[a]	0.0133	0.0075[a]	0.0109[a]	0.0099[a]	0.0105[a]	0.0084[a]	0.0086[a]
0.0052[a]	0.0065[a]	0.0042[a]	0.0048[a]	0.0061[a]	0.0065[a]	0.0051[a]	0.0053[a]
0.0059[a]	0.0063[a]	0.0098	0.0076[a]	**0.0080**	0.0074[a]	0.0081	0.0109
0.0128[a]	0.0116[a]	0.0154	0.0142[aa]	0.0139[a]	0.0115[a]	0.0111[a]	0.0136[a]
0.0288[a]	0.0338	0.0170[a]	0.0293[a]	0.0306[a]	0.0313[a]	0.0274[a]	0.0172[a]
0.0248[a]	0.0270[a]	0.0185[a]	0.0256[a]	0.0336	0.0284[a]	**0.0295**	0.0344
0.0143	0.0169	0.0106[a]	0.0130[a]	0.0114[a]	0.0136[a]	0.0128[a]	0.0092[a]

Bold number indicates amino acids with identical flexibility of both native and mutants

[a] Indicates amino acids with decreased flexibility of mutants than native

performed the point mutation by SWISSPDB viewer for the 17 variants using the PDB ID: 3QLC (Chain A). Energy minimization for all the mutant models and their native structures was achieved using the Nomad-Ref server and followed by simulated annealing using iFOLD server to get optimized structures. Subsequently, docking was performed for both the native type and all the 17 mutant types of ADD domain with H3-peptide.

Table 3 shows the Rosetta energy score between histone H3-peptide and native ADD domain that was found to be −88.709 kcal/mol whereas with 17 mutants, the Rosetta energy score was found between the ranges −41.790 and −84.918 kcal/mol as outlined in Table 3. The binding affinity between all the 17 mutant structures of ADD domain

and histone H3-peptide had been found to be less as compared to native complex. Figure 3 shows the docked complex of native and all the 17 mutant structures with histone H3-peptide. Moreover, we computationally found that two mutants, V194I and Q219P displayed more pronounced reduction in binding to histone H3-peptide, compared to other mutants with Rosetta energy score of −41.790 and −41.915 kcal/mol, respectively, as depicted in Fig. 3h, j.

The atomic resolution exploration of protein–peptide interactions found essential for all biological functions because it allowed the comprehensive knowledge of the physical basis of affinity, as well as the understanding of molecular recognition at intermolecular interface level (Goldman et al. 1997; Talavera et al. 2011; Moreira et al.

2007; Jingyu et al. 2013). So, we further evaluated these 17 variants of ADD domain that affected by binding affinity with histone H3-peptide through interactions of interest that disrupted in the interacting residues at the interface level. It is noted that, the major intermolecular interactions are noticed such as hydrophobic interactions, inter-molecular hydrogen bonding (i.e., main chain/main chain interactions, main chain/side chain, side chain/side chain interactions) and ionic interactions. In this analysis, we found 29 inter-molecular interactions mediated by 3 hydrophobic interactions, 23 hydrogen bonding namely 8 main chain–main chains, 9 main chain–side chains and 6 side chain–side chain followed by 3 ionic interactions between the native ADD domain and histone H3-peptide. Whereas, all the 17 mutants structure of ADD domain with histone H3-peptide have inter-molecular interactions between the ranges of 12–24 as could be seen from Table 3. Hence, reductions in the inter-molecular interactions between all the 17 mutants structure of ADD domain with histone H3-peptide affect its binding efficiency.

3.6 Identifying the number of amino acids with decreased flexibility in the interacting residues of ADD domain

We selected the PDB ID 3QLC, the complex of ADD domain (chain A) with histone H3-peptide (chain C). It could be seen from Table 4 that, 14 amino acids namely Asp(212), Met(216), Asp(217), Glu(218), Gly(226), Gly(227), Asn(228), Leu(229), Ile(230), Cys(231), Asp(233), Arg(250), Ser(254), and Trp(263) were identified as binding residues for ADD domain with histone H3-peptide by PDB sum (Laskowski 2001). For the further course of investigation, we performed normal mode analysis to compare the flexibility of binding amino acids of both native and mutants of ADD domain to confirm the detrimental effect of all the 17 mutants by using the program Elnemo. Table 4 depicts the flexibility of amino acids of both native type and mutant types by means of normalized mean square displacement $<R^2>$. We considered the $<R^2>$ of binding amino acids of native ADD domain as referred to compare the binding amino acids of 17 mutants. Hence we further sorted out these data into three different ranges of flexibility. One was the $<R^2>$ of binding amino acids of mutants which was exactly the same as $<R^2>$ of the binding amino acids of native named as 'identical flexibility.' The other was the $<R^2>$ of binding amino acids of mutants which was higher than $<R^2>$ of the binding amino acids of native named as 'increased flexibility.' And the last was the $<R^2>$ of binding amino acids of mutants which was lesser than $<R^2>$ of binding amino acids of native named as 'decreased flexibility.' The three different ranges of flexibility of binding amino acids of 17 mutants could be

Table 5 Binding amino acids of mutants with different ranges of flexibility based on $< R^2>$

Mutants	$<R^2>$		
	A	B	C
G175E	0	2	12
N179S	0	1	13
P190A	0	2	12
P190L	0	5	9
P190S	0	3	11
L192F	0	2	12
V194I	0	5	9
C200S	0	5	9
Q219P	0	1	13
C220R	0	1	13
C220Y	0	4	10
W222S	0	2	12
C243F	0	0	14
R246C	1	1	12
R246L	0	2	12
G249C	1	1	12
G249D	0	2	12

A amino acids with identical flexibility of both native and mutants, *B* amino acids with increased flexibility of mutants than native, *C* amino acids with decreased flexibility of mutants than native

seen from Table 5. In case of ATRX syndrome most of the mutations occurred in the binding regions of ADD domain as mentioned above. Based on the results obtained from 17 mutant structures, we identified a loss of flexibility in mutant structures by comparison with native crystal structure of ADD domain. From this result, we confirmed that, out of 238 binding residues in all 17 mutant structures, 2 binding residues having identical flexibility, 39 binding residues having increased flexibility, and 197 binding residues having decreased flexibility. This result could clearly exemplified that decreased flexibility in the binding residues of 17 mutant structures of ADD domain changed the conformational state of interfacial residues and could affect the evolutionary molecular recognition of histone H3-peptide according to "conformational selection" method (Keskin et al. 2005). Interestingly, two variants L192F and Q219P had more pronounced effect on ADD domain by computational tools (I-Mutant 2.0, SIFT, PolyPhen) and structure-based methods (total energy, RMSD, inter–intra molecular interactions, Rosetta energy score and normal mode analysis) leads to cause ATR-X syndrome.

4 Conclusions

In this study, we computationally analyzed the 17 detrimental missense mutations in sequence and structure-based

level to understand the structural and functional effect of these missense mutations in ADD domain of chromatin remodeling protein ATRX. From this work, we concluded that all the 17 mutants have been considered to be less stable, deleterious and damaging by I-Mutant2.0, SIFT and PolyPhen programs, respectively. The RMSD between the native type and the mutant type of ADD domain had also been found in the range of 0.29–1.15 Å. Moreover, docking was performed between histone H3-peptide with native and mutant structures of ADD domain which showed the Rosetta energy score between −41.790 and −84.918 kcal/mol. Our computational results suggest that the normal function of ADD domain required its high-affinity binding to histone H3-peptide and mutations in ADD domain which directly affected the binding of histone H3-peptide that underlie ATR-X syndrome. Further, we evaluated that the decreased binding affinity of mutants ADD domain with histone H3-peptide were due to pronounced reduction in inter–intra molecular interactions. The less binding affinity had been observed on these 17 mutants with histone H3-peptide as the majority of binding amino acids of those 17 mutants had 'decreased flexibility' by normal mode analysis which could be the cause for ATR-X syndrome. Identified detrimental missense mutations by sequence (SIFT) and structure-based (I-Mutant 2.0, PolyPhen) computational tools had good agreement with structure-based analysis (total energy, RMSD, inter–intra molecular interactions, docking studies and normal mode analysis). The overall scope and novelty of our work was (1) to consider computationally a suitable protocol for identifying the detrimental missense mutation before wet lab experimentation and, (2) to provide optimal path for further clinical and experimental studies to characterize this mutants of ADD domain in depth.

References

Argentaro A, Yang JC, Chapman L, Kowalczyk MS, Gibbons RJ, Higgs DR, Neuhaus D, Rhodes D (2007) Structural consequences of disease-causing mutations in the ATRX-DNMT3-DNMT3L (ADD) domain of the chromatin-associated protein ATRX. Proc Natl Acad Sci USA 104:11939–11944

Badens C, Martini N, Courrier S, DesPortes V, Touraine R, Levy N, Edery P (2006) ATRX syndrome in a girl with a heterozygous mutation in the ATRX Zn finger domain and a totally skewed X-inactivation pattern. Am J Med Genet 140:2212–2215

Baker LA, Allis CD, Wang GG (2008) PHD fingers in human diseases: disorders arising from misinterpreting epigenetic marks. Mutat Res 647:3–12

Baumann C, De La Fuente R (2009) ATRX marks the inactive X chromosome (Xi) in somatic cells and during imprinted X chromosome inactivation in trophoblast stem cells. Chromosoma 118:209–222

Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN, Bourne PE (2000) The Protein Data Bank. Nucleic Acids Res 28:235–242

Berube NG, Healy J, Medina CF, Wu S, Hodgson T, Jagla M, Picketts DJ (2008) Patient mutations alter ATRX targeting to PML nuclear bodies. Eur J Hum Genet 16:192–201

Capriotti E, Fariselli P, Casadio R (2005) I-Mutant2.0: predicting stability changes upon mutation from the protein sequence or structure. Nucleic Acids Res 33:306–310

Carlson HA, McCammon JA (2000) Accommodating protein flexibility in computational drug design. Mol Pharmacol 57:213–218

De La Fuente R, Baumann C, Viveiros MM (2011) Role of ATRX in chromatin structure and function: implications for chromosome instability and human disease. Reproduction 142:221–234

Delarue M, Dumas P (2004) On the use of low-frequency normal modes to enforce collective movements in refining macromolecular structural models. Proc Natl Acad Sci USA 101:6957–6962

Dhayalan A, Tamas R, Bock I, Tattermusch A, Dimitrova E, Kudithipudi S, Ragozin S, Jeltsch A (2011) The ATRX-ADD domain binds to H3 tail peptides and reads the combined methylation state of K4 and K9. Hum Mol Genet 20:2195–2203

Du Y, Zhou J, Fan J, Shen Z, Chen X (2009) Streamline proteomic approach for characterizing protein–protein interaction network in a RAD52 protein complex. J Proteome Res 8:2211–2217

Goldman ER, Dall Acqua W, Braden BC, Mariuzza RA (1997) Analysis of binding interactions in an idiotope-antiidiotope protein–protein complex by double mutant cycles. Biochemistry 1997(36):49–56

Han JH, Kerrison N, Chothia C, Teichmann SA (2006) Divergence of inter domain geometry in two-domain proteins. Structure 14:935–945

Hinkle A, Tobacman LS (2003) Folding and function of the troponin tail domain. Effects of cardiomyopathic troponin T mutations. J Biol Chem 278:506–513

Iwase S, Xiang B, Ghosh S, Ren T, Lewis PW, Cochrane JC, Allis CD, Picketts DJ, Patel DJ, Li H, Shi Y (2011) ATRX ADD domain links an atypical histone methylation recognition mechanism to human mental-retardation syndrome. Nat Struct Mol Biol 18:769–776

Jingyu H, Wei Z, Yi-Ping PC (2013) Dynamically predicting protein functions from semantic associations of protein. Netw Model Anal Health Inform Bioinforma. doi:10.1007/s13721-013-0024-z

Keskin O, Ma B, Rogale K, Gunasekaran K, Nussinov R (2005) Protein-protein interactions: organization, cooperativity and mapping in a bottom-up systems biology approach. Phys Biol 2:24–35

Laskowski RA (2001) PDBsum: summaries and analyses of PDB structures. Nucleic Acids Res 29:221–222

Law MJ, Lower KM, Voon HP, Hughes JR, Garrick D, Viprakasit V, Mitson M, De Gobbi M, Marra M, Morris A, Abbott A, Wilder SP, Taylor S, Santos GM, Cross J, Ayyub H, Jones S, Ragoussis J, Rhodes D, Dunham I, Higgs DR, Gibbons RJ (2010) ATR-X syndrome protein targets tandem repeats and influences allele-specific expression in a size-dependent manner. Cell 143: 367–378

Lindahl E, Azuara C, Koehl P, Delarue M (2006) NOMAD-Ref: visualization, deformation and refinement of macromolecular structures based on all-atom normal mode analysis. Nucleic Acids Res 34:52–56

London N, Raveh B, Cohen E, Fathi G, Schueler-Furman O (2011) Rosetta FlexPepDock web server-high resolution modeling of peptide-protein interactions. Nucleic Acids Res 39:249–253

Manisk Kumar M, Krishna M (2013) Modeling and simulation analysis of propyl-thiouracil (PTU), an anti-thyroid drug on thyroid peroxidase (TPO), thyroid stimulating hormone receptor (TSHR), and sodium iodide (NIS) symporter based on systems

biology approach. Netw Model Anal Health Inform Bioinforma 2:45–57

Mitson M, Kelley LA, Sternberg MJ, Higgs DR, Gibbons RJ (2011) Functional significance of mutations in the Snf2 domain of ATRX. Hum Mol Genet 20:2603–2610

Mohammed Said AR, Mohammed Azmi AB, Rosni A, Ahamad Tajudin K (2012) A hybrid harmony search algorithm for ab initio protein tertiary structure prediction. Netw Model Anal Health Inform Bioinforma. doi:10.1007/s13721-012-0013-7

Moreira IS, Fernandes PA, Ramos MJ (2007) Hot spots—a review of the protein–protein interface determinant amino-acid residues. Proteins 2007(68):803–812

Ng PC, Henikoff S (2001) Predicting deleterious amino acid substitutions. Genome Res 11:863–874

Ng PC, Henikoff S (2003) SIFT: predicting amino acid changes that affect protein function. Nucleic Acids Res 31:3812–3814

Parthasarathy S, Murthy MR (2000) Protein thermal stability: insights from atomic displacement parameters (B values). Protein Eng 13:9–13

Rajasekaran R, Sethumadhavan R (2010a) Exploring the cause of drug resistance by the detrimental missense mutations in KIT receptor: computational approach. Amino Acids 39:651–660

Rajasekaran R, Sethumadhavan R (2010b) In silico identification of significant detrimental missense mutations of EGFR and their effect with 4-anilinoquinazoline-based drugs. Appl Biochem Biotechnol 160:1723–1733

Ramensky V, Bork P, Sunyaev S (2002) Human non-synonymous SNPs: server and survey. Nucleic Acids Res 30:3894–3900

Routledge KE, Tartaglia GG, Platt GW, Vendruscolo M, Radford SE (2009) Competition between intramolecular and intermolecular interactions in an amyloid-forming protein. J Mol Biol 389:776–786

Schuler LD, Daura X, Van Gunsteren WF (2001) An improved GROMOS96 force field for aliphatic hydrocarbons in the condensed phase. J Comput Chem 22:1205–1218

Sharma S, Ding F, Nie H, Watson D, Unnithan A, Lopp J, Pozefsky D, Dokholyan NV (2006) iFold: a platform for interactive folding simulations of proteins. Bioinformatics 22:2693–2694

Steward RE, MacArthur MW, Laskowski RA, Thornton JM (2003) Molecular basis of inherited diseases: a structural perspective. Trends Genet 19:505–513

Suhre K, Sanejouand YH (2004) ElNemo: a normal mode web server for protein movement analysis and the generation of templates for molecular replacement. Nucleic Acids Res 32:610–614

Talavera D, Robertson DL, Lovell SC (2011) Characterization of protein–protein interaction interfaces from a single species. PLoS One 6

Tina KG, Bhadra R, Srinivasan N (2007) PIC: protein interactions calculator. Nucleic Acids Res 35:473–476

Valadez-Graham V, Yoshioka Y, Velazquez O, Kawamori A, Vazquez M, Neumann A, Yamaguchi M, Zurita M (2012) XNP/dATRX interacts with DREF in the chromatin to regulate gene expression. Nucleic Acids Res 40:1460–1474

Varfolomeev SD, Uporov IV, Fedorov EV (2002) Bioinformatics and molecular modeling in chemical enzymology. Active sites of hydrolases. Biochemistry 67:1099–1108

Wong LH, McGhie JD, Sim M, Anderson MA, Ahn S, Hannan RD, George AJ, Morgan KA, Mann JR, Choo KH (2010) ATRX interacts with H3.3 in maintaining telomere structural integrity in pluripotent embryonic stem cells. Genome Res 20:351–360

Xue Y, Gibbons R, Yan Z, Yang D, McDowell TL, Sechi S, Qin J, Zhou S, Higgs D, Wang W (2003) The ATRX syndrome protein forms a chromatin-remodeling complex with Daxx and localizes in promyelocytic leukemia nuclear bodies. Proc Natl Acad Sci USA 100:10635–10640

Yip YL, Famiglietti M, Gos A, Duek PD, David FP, Gateau A, Bairoch A (2008) Annotating single amino acid polymorphisms in the UniProt/Swiss-Prot knowledgebase. Hum Mutat 29:361–366

Yuan Z, Bailey TL, Teasdale RD (2005) Prediction of protein B-factor profiles. Proteins 58:905–912