Cell Biochem Biophys (2016) 74:109–127
DOI 10.1007/s12013-016-0722-0

ORIGINAL PAPER

In Silico Analysis of FMR1 Gene Missense SNPs

Akin Tekcan[1]

Received: 22 October 2015 / Accepted: 27 January 2016 / Published online: 15 February 2016

Abstract The *FMR1* gene, a member of the fragile X-related gene family, is responsible for fragile X syndrome (FXS). Missense single-nucleotide polymorphisms (SNPs) are responsible for many complex diseases. The effect of *FMR1* gene missense SNPs is unknown. The aim of this study, using in silico techniques, was to analyze all known missense mutations that can affect the functionality of the *FMR1* gene, leading to mental retardation (MR) and FXS. Data on the human FMR1 gene were collected from the Ensembl database (release 81), National Centre for Biological Information dbSNP Short Genetic Variations database, 1000 Genomes Browser, and NHLBI Exome Sequencing Project Exome Variant Server. In silico analysis was then performed. One hundred-twenty different missense SNPs of the *FMR1* gene were determined. Of these, 11.66 % of the *FMR1* gene missense SNPs were in highly conserved domains, and 83.33 % were in domains with high variety. The results of the in silico prediction analysis showed that 31.66 % of the FMR1 gene SNPs were disease related and that 50 % of SNPs had a pathogenic effect. The results of the structural and functional analysis revealed that although the R138Q mutation did not seem to have a damaging effect on the protein, the G266E and I304N SNPs appeared to disturb the interaction between the domains and affect the function of the protein. This is the first study to analyze all missense SNPs of the *FMR1* gene. The results indicate the applicability of a bioinformatics approach to FXS and other FMR1-related diseases. I think that the analysis of FMR1 gene missense SNPs using bioinformatics methods would help diagnosis of FXS and other FMR1-related diseases.

Keywords FMR1 · Fragile X syndrome · Missense SNP · In silico analysis

Introduction

The fragile X mental retardation 1 (*FMR1*) gene is located on Xq27.3 [1]. It is a member of the fragile X-related gene family, which synthesizes the fragile X-related proteins 1 and 2 (FMRP 1 and 2) [2, 3]. These RNA-binding proteins (RBPs) are highly expressed in the nervous and reproductive systems [1–4]. RBPs contain various structural motifs, such as RNA recognition motif (RRM) and dsRNA binding domain. RBPs play key roles in post-transcriptional modification, RNA transport, and the control of protein synthesis [5]. They also play a role in mRNA transcription, splicing, turnover, polyadenylation, transport, translational control, nuclear export, editing, and RNA degradation [1, 6]. RBP-induced control of translational mechanisms is essential for neurological functions [7], and a lack of the FMR1 protein can cause neurological disorders, such as intellectual disability (ID) and mental retardation (MR) [8]. FMRP, FMR1, and FMR2 proteins have three RBP motifs (two hnRNP K-homology [KH1 and KH2] motifs and one arginine-glycine-glycine [RGG] motif) [1, 6, 9].

Fragile X syndrome (FXS)(MIM #300624) is caused by altered expression of the FMR1 gene and the loss of the FMRPs [1]. The loss of FMRPs is also responsible for premature ovarian failure (MIM #311360) [10] and fragile

Electronic supplementary material The online version of this article (doi:10.1007/s12013-016-0722-0) contains supplementary material, which is available to authorized users.

✉ Akin Tekcan
akintekcan@hotmail.com

[1] School of Health, Ahi Evran University, Kirsehir, Turkey

X tremor/ataxia syndrome (MIM #300623). FXS is a classic example of the loss of translational control. FXS is the leading monogenic cause of inherited ID and autism spectrum disorder (ASD) [11, 12]. In FXS, the level of FMRP is lower than normal, and FXS occurs due to excess FMR1 transcripts forming aggregates. To prevent FXS, the expression level of the *FMR1* gene must be controlled [1].

The 5′ untranslated region (UTR) of the *FMR1* gene is polymorphic, with normal alleles ranging from 5 to 44 CGG repeats, gray zone alleles containing 45–54 repeats, premutation alleles containing 55–200 repeats, and full mutation alleles containing 200 repeats or more. Expansion to more than 200 repeats results in hypermethylation and silencing of the FMR1 gene product [1, 10, 12, 13]. Small deletions and single-nucleotide polymorphisms (SNPs) in critical domains of the FMR1 gene may cause FXS-like phenotypes. It has been suggested that patients with a clinical FXS-like phenotype who present with developmental delays but do not have the full *FMR1* mutation should be routinely tested for further mutations of the *FMR1* coding region [14]. The prevalence of other mutations in the *FMR1* coding region is unknown [13, 14]. G266E [1, 12, 14], I304N [12], S27X [12], and R138Q [13, 15] mutations in the *FMR1* gene were reported to cause FXS. According to several studies, *FMR1* gene missense SNPs should be scanned [12–16].

There are 12,817 FMR1-related variations reported in the Ensembl (release 81) database. These include frameshift, missense, synonymous, and intron variants. Of the 12,817 *FMR1*-related variations, 1251 are classified as missense mutations, with altered amino acid sequences, and there are 120 different missense SNPs [17]. The missense SNPs are categorized according to whether they exert a pathogenic effect or whether they are deleterious and tolerated. These classifications are based on the results of bioinformatics-based in silico analyses. In silico techniques are important to determine the effects of mutations on the structure and function of proteins [18], and a number of studies have used them to analyze missense mutations in different genes [18–23]. The aim of the present study, using in silico techniques, was to analyze all known missense SNPs that can affect the functionality of the *FMR1* gene leading to MR and FXS. To the best of our knowledge, this is the first in silico analysis of *FMR1* gene missense mutations related to MR and FXS.

Materials and Methods

Data Set Used for Mutation Annotation

Human FMR1 gene information data were collected from the Ensembl release 81 [17], National Centre for Biological Information (NCBI) dbSNP Short Genetic Variations database [24], 1000 Genomes Browser [25], and NHLBI Exome Sequencing Project (ESP) Exome Variant Server [26]. These databases revealed 120 different missense SNPs or nonsynonymous mutations in the *FMR1* gene. The amino acid sequences of the FMR1 protein were retrieved from the Uniprot database (accession number Q06787) [27]. Information on the expression profiles of the FMR1 protein was retrieved from the RCSB Protein Data Bank (PDB) [28] and Multi-Omics Profiling Expression Database [29].

Multiple Sequence Alignment (MSA)

Ten sequence homologs were defined with ExPASy Basic Local Alignment Search Tool (BLAST) [30] and aligned to analyze their level of evolutionary conservation [31]. Multiple sequence alignment (MSA) of the sequences was performed using ClustalW [32], T-COFFEE [33], and Muscle [34]. WebLogo was used for graphical representation of the aligned amino acid sequences [35]. In the determination and visualization of the domains of the FMR1 protein family, the Pfam database was utilized [36]. The variety of the residues in each domain of the FMR1 was analyzed using the ConSurf server [37].

Molecular Interface Analysis and Predictions of Protein–Protein Interactions

Molecular interface analysis of the FMR1 protein (PDB codes #2bkd and #2qnd) was performed with PDBePISA (Protein, Interfaces, Structure, and Assemblies) [38]. The ConSurf tool was used to display the cysteines of the FMR1 protein and its disulfide, sulfur, and selenium bonds [37]. The structure of the protein was visualized by the Java Viewer for Chemical Structures in 3D (Jmol) [39].

The Search Tool for the Retrieval of Interacting Genes/Proteins (STRING) online database was used to retrieve functional partners of the FMR1 protein. In the STRING analysis, data were extracted on neighborhoods, gene fusion, co-occurrence, coexpression, experiments, database, and text mining. Predictions with a confidence score higher than 0.7 were included in this study.

Prediction of the Pathogenic Effects of SNPs

Sorting Tolerant From Intolerant (SIFT), Protein Analysis Through Evolutionary Relationships (PANTHER), and Polymorphism Phenotyping (PolyPhen) analysis tools were used to determine the pathogenic effects of the *FMR1* gene missense SNPs. PolyPhen analyzes the potential effect of amino acid changes according to homology and assigns a score to each SNP of between 0 and 1. The SIFT tool

classifies SNPs as tolerated (≥0.05) and/or deleterious (<0.05), the PolyPhen tool classifies them as benign, probably damaging, or possibly damaging, and PANTHER classifies them as neutral or likely to be disease causing.

Prediction of Disease-Related SNPs

In the bioinformatics analysis of disease-related *FMR1* gene SNPs, two algorithms, the Predictor of human Deleterious Single-Nucleotide Polymorphisms (PhD-SNP) and Predicting Human Disease-related Mutations in Proteins with Functional Annotations (SNPs&GO), were used. PhD-SNP, which predicts human deleterious SNPs, utilizes support vector machines (SVMs). The results obtained using both algorithms, with the mutations presented as neutral or disease causing and assigned scores of between 0 and 1. When the analysis was performed, other data, such as the UNIPROT number, mutation position, wild-type residue, and substituting residue, were also entered.

Prediction of Protein Stability Changes Upon SNPs from the Protein Sequence

I-Mutant2.0 is an SVM-based web server for the automatic prediction of protein stability changes upon single-site mutations. The tool was trained on a dataset derived from ProTherm, which is the most comprehensive database of experimental data on protein mutations. The input parameters were a pH of 7 and temperature of 25 °C. The reliability index and prediction scores (0 = unreliable and 10 = reliable) were obtained using I-Mutant2.0.

Visualization and Evaluation of the Structural Effects of the SNPs

To visually depict the structural effects of FMR1 gene SNPs, the HOPE web tool was used (http://www.cmbi.ru.nl/hope/). HOPE collects structural information from various sources, including the 3D protein structure, sequence annotations in UniProt, and predictions from DAS-servers. HOPE combines this information to analyze the effect of a specific mutation on the protein structure. HOPE is an online web service where the user can submit a sequence and mutation.

Results

MSA

The MSA analysis of the FMR1 protein using the Pfam database revealed five domains. The domains were Agenet (PF05641), KH1, KH2 (PF00013, protein domains), FXMRP1 C Core (PF12235, RGG motif), and FXMR C2 (PF16098) (Table 1). One hundred-twenty different missense SNPs or nonsynonymous mutations were determined in the *FMR1* gene. Of these *FMR1* gene missense SNPs, 11.66 % were in highly conserved regions, such as the FXMR C2 domain, and 83.33 % were in nonconserved regions, such as KH1, KH2, and FXMRP1 domains. Although the KH1, KH2, and FXMRP1 C Core domains were present in a wide variety of regions, the FXMR C2 domain was found only in a highly conserved region. Two hundred-fourteen mutated residues (not only missense mutations) were identified as functional (62.14 %) and structurally (37.85 %) important. Of these 214, 66.82 % were in nonconserved domains, especially Agenet and KH (Table 1). The MSA results, ConSurf alignments, and WebLogo files are provided in supplementary material 1, 2, and 3, respectively.

Result of Molecular Interfaces and Protein–Protein Interactions

The results of the interface analysis of the bond type and surface areas of the PDB#2bkd and #2qnd 3D structures revealed 16 H-bonds, one salt bridge, and two covalent bonds. The residues forming a hydrogen/disulfide bond, salt bridge, or covalent link are shown Table 2. In the PDB#2bkd and #2qnd structures of the FMR1 protein (Uniprot ID Q06787), there were 35 interface residues. The accessible surface areas ranged from 11.54 angstrom to 183.99 angstrom. Five sulfur atoms, three cysteines, no disulfide bonds, and two methionines (no selenomethionine) were determined in the protein (Fig. 1). The 3D structure and interfaces of the N-terminal domain of the FMRP are displayed in Fig. 2.

Figure 3 shows the results of the protein associations using the STRING database. The STRING analysis revealed 10 interactive objects (GRM5, DLG3, CYFIP1, CYFIP2, ZNF385B, EIF2C2, EIF2C4, EIF2C1, EIF2C3, and MFS4 proteins) for FMR1 protein in protein–protein interaction software. Among these, CYFIP1 and CYFIP2 belonged to the cytoplasmic FMR1-interacting protein family, and EIF2C2, EIF2C4, EIF2C1, and EIF2C3 were eukaryotic translation initiation factors. Thus, the results revealed a high degree of association between the FMR1 protein and the translation initiation process.

Pathogenic Effects of the SNPs

The pathogenic effects of the missense SNPs were predicted using three bioinformatics algorithms (SIFT, PolyPhen, and PANTHER). The SIFT program predicts whether an amino acid substitution caused by a missense

Table 1 The residues variety at each domain of FMR1 protein

Domains	Mutations	CI	B/E	F/S	Mutations	CI	B/E	F/S	Mutations	CI	B/E	F/S
Agenet	D63D,M,A,Q,E	6.4	E		A82A,C,G	9.8	B		A101N,Y,H,A,S,T,C,P	8.6	B	
	E64D,E,L	9.9	E	F	K83M,L,V,R,K,I,T,G	5.2	E		T102S,A,T,P	8.6	E	
	V65 L,I,V	9.9	B	S	V84I,V,L	9.9	B	S	Y103F,C,L,Y	8.6	B	
	E66E,L	9.9	E	F	R85R,K,N,Q	9.8	E	F	N104N,S,K,T	9.8	E	F
	V67I,V,A	9.9	B	S	M86 V,R,M	9.9	B	S	E105E,D	9.9	E	F
	Y68L,Y,F	4.1	B		I87L,S,M,C,T,V,I	6.4	B		I106I,V	9.9	B	S
	S69S,A,T	9.9	B	S	K88N,K	9.9	E	F	V107I,V	9.9	B	S
	R70E,K,S,R	9.7	E	F	G89G,M	9.9	B	S	T108S,N,A,P,D,T	8.6	B	
	A71T,P,A,S	8.6	E		E90E,A,D,G	7.5	E		I109F,I,T,G,A,H,S,V,L,Y	4.2	B	
	N72 N,T	9.9	E	F	F91I,F	9.8	B		E110G,D,E	8.7	E	F
	E73D,E,N	6.3	E		Y92Y,L,S,H,Q,C,F	7.5	B		R111K,R	9.9	E	F
	K74G,N,K,S,L,H,Q	6.4	E		V93A,I,V	9.8	B	S	L112I,L	9.7	B	
	E75E	9.9	E	F	I94I,V,L	9.8	B	S	R113R,W	9.9	E	F
	P76P,V,Q,A	9.8	E	F	E95K,E,V	9.9	B	S	S114S,A,P,T,Y,L,M,V	7.5	E	
	C77W,Y,S,C,R	8.6	B		Y96X,Y	9.9	B	S	V115I,V,T,A,L,K	8.6	E	
	C78C,G,E	9.8	B		A97M,A,Q,L,S,V,P	8.6	B		N116N	9.9	E	F
	W79W	9.9	B	S	A98S,A,G,D,V	9.8	B		P117Q,S,L,K,R,P,T	8.6	E	
	W80W	9.9	B	S	C99C,I,L,W	8.6	B		N118N,S	9.9	E	
	L81L,K,E,M,R,P	6.3	B		D100D,N,E	7.5	E		K119C,T,P,Q,S,K	8.6	E	
KH1	Q222E,Q,V	9.8	E	F	I241I	9.9	B	S	C260S,A,G,C	8.7	B	
	F223V,F	9.9	B	S	Q242Q,M	9.9	E	F	T261A,N,T	9.9	E	F
	I224I,G,T,A,S,K,V,R,Q,M,N	2.1	E		Q243N,Q,A,T	9.7	E	F	F262F	9.9	B	S
	V225V	9.9	B	S	A244A	9.9	B	S	H263R,H,K	9.8	E	F
	R226A,S,R	9.9	E	F	R245R	9.9	E	F	I264I	9.9	B	S
	E227T,P,D,E,S,L	6.4	E		K246K,L,R	9.8	E	F	Y265Y,S,R,T,F	7.5	E	
	D228D,E,H	9.9	E	F	V247L,V,T	9.8	E	F	G266G	9.9	B	S
	L229L	9.9	B	S	P248S,A,E,P,D	5.3	E		E267D,E,Q,N	9.8	E	F
	M230I,R,M	9.9	B	S	G249G,N	9.9	E	F	D268S,N,E,D,T	7.5	E	
	G231G	9.9	B	S	V250I,V	9.8	B	S	Q269E,Q,V,D,R,A,H,K,S,P,T	5.2	E	
	L232P,L	9.9	B	S	T251T,V,S,L	9.8	B		D270D,E	7.5	E	
	A233A	9.9	B	S	A252N,S,A,G,T	8.7	B		A271A,S	9.9	B	S
	I234I	9.9	B	S	I253I	9.9	B	S	V272V,C	9.9	B	S
	G235G	9.9	B	S	D254D,E	9.7	E	F	K273R,Q,K,S	6.4	E	
	T236S,A,T	9.7	B		L255L	9.9	B	S	K274V,R,E,M,Q,L,I,T,A,K	5.2	E	
	H237R,H	9.9	B	S	D256G,D,V,N,Q,E	8.6	E		A275A	9.9	B	S

Table 1 continued

Domains	Mutations	CI	B/E	F/S	Mutations	CI	B/E	F/S	Mutations	CI	B/E	F/S
	G238G	9.9	B	S	E257E	9.9	E	F	R276R	9.9	E	F
	A239T,A,S	9.9	B	S	D258N,E,Q,D,K,S,A,C	7.5	E		S277R,V,N,L,E,C,T,G,S,A	5.2	E	
	N240N	9.9	E	F	T259T,G,S	7.6	E					
KH2	V285L,Y,V,S,A,T,P,I,F	6.4	B		G298G	9.9	E	F	S311S	9.9	E	F
	I286L,M,F,I,V	8.6	B		K299K,R	9.9	E	F	G312G,C	9.9	E	F
	Q287K,E,Q,H,D,P	8.6	E		N300N,S	8.7	E	F	V313V	9.9	B	S
	V288V,I	9.9	B	S	G301G	9.9	B	S	V314F,V,I	9.9	B	S
	P289P	9.9	B	S	K302K,R	9.9	E	F	R315R,L,K	9.9	E	F
	R290R	9.9	E	F	L303I,V,L,A	8.7	B		V316L,I,V	9.9	B	S
	N291I,S,A,D,V,N,L,M	7.5	E		I304I	9.9	B	S	R317K,V,C,R	9.7	E	F
	L292L,F	9.9	B	S	Q305Q	9.9	E	F	I318F,I,V	9.8	B	
	V293I,V,L	9.8	B		E306E	9.9	E	F	E319A,E	9.9	E	F
	G294G	9.9	B	S	I307I,V,M	9.8	B		A320A,G,P	8.7	E	
	K295K	9.9	E	F	V308V	9.9	B	S	E321E,G,D,I	8.7	E	
	V296V	9.9	B	S	D309N,D	9.9	E	F	N322D,G,L,S,N	9.7	E	F
	I297I	9.9	B	S	K310K	9.9	E	F	–	–		
FXMRP1 C Core	H420H	9.9	B	S	Y461F,I,Y,S,A	6.4	E		T502A,N,S,I,T,P	8.6	E	
	L421L,M,I,V	9.7	B		V462Q,M,L,N,V,A,H,S,I,T,G	3.1	E		E503A,E,S,D,T	7.4	E	
	N422N,S,A,V	8.6	E		T463A,L,S,V,F,T	6.3	E		S504D,R,C,E,S	9.7	E	F
	Y423H,Y	9.9	B	S	D464E,D	7.5	E		D505A,E,L,N,V,R,D	4.1	E	
	L424L	9.9	B	S	D465E,S,Y,N,D,G	4.2	E		H506R,V,N,G,T,F,K,H,A	6.3	E	
	K425T,K,N,Q	9.8	E	F	G466A,L,S,N,G	6.4	E		R507P,R,D,K,N,E	5.2	E	
	E426EE,V,D	9.9	E	F	Q467S,A,P,G,T,N,Q,M,D,V	3.1	E		D508R,D,T,E,A	5.8	E	
	V427L,A,M,V	9.8	B		G468P,D,T,G,A,S	6.3	E		E509E,D,T	8.7	E	
	D428H,E,Q,D	9.8	E	F	M469V,Y,L,M,P,T,I,S,A	3.1	E		L510P,R,Q,A,L	8.6	B	
	Q429K,E,A,Q	9.8	E	F	G470P,T,G,S,L,E,A	5.2	E		S511T,D,R,P,K,S	9.7	E	F
	L430L,M,V	9.8	E	F	R471R,G,Q,H	8.5	E		D512P,D	9.9	E	F
	R431R,G,L,Y,H	9.7	E	F	G472G,T,D,V,N	5.2	E		W513W,F,R	8.6	B	
	L432L,M,Q,I	5.3	E		S473G,T,R,A,Q,S	6.3	E		S514S,L,I,G	9.7	B	
	E433A,E,Q,K,D	9.7	E	F	R474K,R,G	8.7	E	F	L515G,T,F,I,S,A,V,L	6.2	E	
	R434R,K,N,H	9.8	E	F	P475I,T,R,P,Q,N	7.5	E		A516G,A,N,L,S	8.7	B	
	L435S,L,M	8.7	B		Y476F,G,M,Y	7.5	E		P517Q,A,N,V,F,G,P	4.1	E	
	Q436H,E,Q	9.9	E	F	R477L,Y,N,R,S,A	5.2	E		T518A,E,Q,D,T	7.5	E	
	I437M,I,G	9.9	B	S	N478A,N,S,G,P,R	5.3	E		E519D,E,K,S	7.4	E	
	D438M,A,D	9.9	E	F	R479G,R	9.8	E	F	E520D,R,V,L,H,E,Q	5.2	E	
	E439Q,E,S,D	9.8	E	F	G480G	9.9	E	F	E521A,G,T,E,N,L,D,R	4.1	E	

Table 1 continued

Domains	Mutations	CI	B/E	F/S	Mutations	CI	B/E	F/S	Mutations	CI	B/E	F/S
	Q440S,Q,E	9.9	E	F	H481A,H,R	9.8	E	F	R522S,K,R,T,G	6.4	E	
	L441L,C	9.9	B	S	G482G,A,S	8.6	E		E523G,N,L,Y,E,M,D,V	3.1	E	
	R442R,D	9.9	E	F	R483G,R,H	9.8	E	F	S524R,V,N,Q,C,T,G,S	2.1	E	
	Q443R,A,Q,N,S	8.6	E		R484K,Q,G,R	9.7	E	F	F525R,G,F,Y,L,Q	2.1	B	
	I444I,F,T,M,A,Q,L	8.6	B		G485A,P,T,G	6.4	E		L526L,Q,R,K,S,H,P,G,T	4.1	E	
	G445G,T,H,V,R,Q,E,L	7.5	E		P486P,T,G,S,A,H,R,Y,N,Q	5.2	E		R527R,G,E,K	7.5	E	
	A446L,M,E,Q,A,G,V	7.5	E		G487A,H,S,N,V,T,G	5.2	E		R528A,K,S,V,R	8.5	E	
	S447T,G,V,I,S,A	7.4	E		Y488A,H,S,F,P,G,T,E,Y,N	4.1	E		G529S,I,G,T,P,V,R,D	4.1	E	
	S448V,Y,L,N,F,I,P,T,G,H,S	4.2	E		T489Q,A,N,S,T,R	4.2	E		D530G,R,D,L,E	7.5	E	
	R449S,P,T,G,M,D,R,V	6.4	E		S490Y,S,F,D	9.8	E	F	G531I,T,G,R,D,E,S	4.1	E	
	P450S,H,A,G,T,P,R	5.2	E		G491V,R,G,A,S	8.6	E		R532R,P,G,S,K,E	6.3	E	
	P451G,P,S,H,A,V,Y,L,Q	3.1	E		T492P,R,T,S,Y,N,A	7.6	E		R533K,E,R,D	6.3	E	
	P452Q,E,N,D,A,S,G,T,P	3.1	E		N493W,Q,N,K,S,G,T	8.6	E		R534E,H,K,N,R,G	7.5	E	
	N453A,S,I,F,T,G,P,N,R	5.2	E		S494G,H,A,L,S	9.8	E	F	G535A,S,I,P,G,E,L,N,V	2.1	E	
	R454V,D,R,Q,E,P,T,G,S	7.5	E		E495E,S,N,D,R,G	6.3	E		G536P,G,A,H,K,S,Q,N	3.1	E	
	T455V,Q,E,M,N,T,G,P,A,S	2.1	E		A496A,H,F,P,G,Q,M,L,R,D	5.2	B		G537T,G,S,K,A,H,R,V,L,Q	3.1	E	
	D456A,S,I,P,G,T,E,L,N,V,D	2.1	E		S497N,S,H,A	9.8	E	F	G538L,N,R,S,K,A,H,P,G,T	1.1	E	
	K457K,A,D,R	7.5	E		N498A,Y,N,R,T	8.7	E		R539D,R,N,L,S,E	7.4	E	
	E458D,R,T,G,E,K,N	3.1	E		A499L,N,A,P,T,F,V	6.4	E		G540N,S,G,D	7.3	E	
	K459E,K,G,T,R,P	5.2	E		S500L,S,N	9.9	E	F	Q541Q,S,Y,V,R,C	5.2	E	
	S460D,R,N,P,G,T,S	4.1	E		E501E,M,N	9.8	E	F	G542G,T,S,L,N	6.3	E	
FXMR C2	G549A,S,I,G,N,Y,V	2.1	E		S577F,I,T,S	7.4	E		L605P,L,Q	7.4	E	
	F550K,S,P,G,F,L,Y,N	2.1	E		L578S,L,F	8.5	B		R606K,R	9.8	E	F
	K551V,N,Y,G,P,K,S,H	3.1	E		Q579L,H,Q	9.7	E		T607V,T,G,P,M,A,S	4.1	E	
	G552P,G,L,K,S,N	3.1	E		I580I,V,L	9.7	B		G608V,M,L,N,F,I,P,G,T	1.1	E	
	N553P,G,T,K,S,A,D,N,E	4.1	E		R581N,C,R	9.7	E		K609T,R,K	9.7	E	
	D554L,E,M,Q,D,G	3.1	E		V582L,G,I,V	5.2	B		D610E,D	4.1	E	
	D555K,A,Q,E,D	5.2	E		D583D	9.9	E	F	R611P,R	9.8	E	F
	H556L,Q,M,V,K,H,G,F,I	3.1	E		C584Y,S,G,C	4.1	E		N612V,P,G,A,S,N	4.1	E	
	S557H,E,Q,S,L,T,P	5.2	E		N585S,N	9.8	E	F	Q613L,Q,M,P,I,V	5.1	E	
	R558W,Q,S,L,V,F,R	5.1	E		N586N	9.9	E	F	K614K	9.9	E	F
	T559K,S,A,H,P,G,T,N,Q	1.1	E		E587E	9.9	E	F	K615K	9.9	E	F
	D560E,Y,L,D,G,T	6.2	E		R588R	9.9	E	F	E616D,E	8.5	E	
	N561A,S,I,P,T,M,L,N,V,D	3.1	E		S589S,A,T	9.7	E	F	K617N,K	9.8	E	F
	R562D,R,T,E,K,S	6.3	E		V590R,V	9.8	E		P618Q,A,S,T,P	5.2	E	
	P563D,L,Q,P,T,S,K,A	2.1	E		H591V,Q,H	9.6	E		D619D,E	8.5	E	

Table 1 continued

Domains	Mutations	CI	B/E	F/S	Mutations	CI	B/E	F/S	Mutations	CI	B/E	F/S
	R564E,A,H,S,R,P	7.4	E		T592H,A,T	8.5	E		S620A,N,S,V,I,T,G	8.5	E	
	N565F,P,H,N,S,Y	6.3	E		K593A,S,K	8.5	E		V621V,C,T,A,Q,L	2.1	E	
	P566V,R,L,I,T,P,A,S	3.1	E		T594F,P,R,T,A,S	5.2	E		D622D,A,E	8.5	E	
	R567M,S,R	9.7	E		L595G,R,A,S,L	6.2	E		G623N,S,A,G,P,V	2.1	E	
	E568D,G,E,A,L	5.2	E		Q596L,K,H,E,Q,G,P	3.1	E		Q624T,P,Q,S	4.1	E	
	A569S,A,P,T,G	1.1	E		N597G,R,P,Q,N,S	5.2	E		Q625T,R,P,H,Q,A	6.3	E	
	K570E,A,K,I	8.5	E		T598S,A,G,T,P,V	4.1	E		P626A,S,L,V,P	4.1	E	
	G571T,G,P,A,K,L,S	2.1	E		S599Q,A,S,F,G,P,D	5.1	E		L627V,L	8.5	E	
	R572E,K,R	9.6	E		S600P,G,V,S,Y,N	2.1	E		V628V	9.9	E	F
	T573P,G,T,A,V,L,N,M,Q	2.1	E		E601G,C,R,S,E	7.4	E		N629N	9.9	E	F
	T574S,H,E,A,Q,T	4.1	E		G602P,G,S,A,H	5.2	E		G630G	9.9	E	F
	D575E,A,L,D	8.5	E		S603P,T,S,H,D,Y,N,Q	5.1	E		V631I,V	9.8	E	F
	G576G,D,N,A,E	7.4	E		R604H,S,R,G	7.5	E		P632P	9.8	E	F

CI confidence interval calculated with Bayesian method, *B/E* burried/exposed residue, *F/S* functional/structural residue

SNP affects the function of a protein and exerts a pathogenic effect. SIFT predictions are based on the degree of conservation of amino acid residues in sequence alignments derived from closely related sequences in Position-Specific Iterated BLAST. SIFT can be applied to naturally occurring missense SNPs [40]. PANTHER is a protein (and gene) classification system that was designed to facilitate high-throughput analyses. To predict the likely functional consequence of a mutation, PANTHER assigns scores to hidden Markov models [41]. PolyPhen is a prediction method, which enables the analysis of all human nonsynonymous SNPs publicly available via the dbSNP database [42]. Of the 120 SNPs analyzed, 61 (50.83 %) SNPs were deleterious, and 59 (49.16 %) SNPs were tolerated, according to the SIFT analysis. According to the PolyPhen analysis, 54 (45 %) SNPs were benign, and 55 (55 %) SNPs were probably damaging and/or possible damaging. Based on the results of the PANTHER analysis, 50 (41.6 %) SNPs were neutral, and 55 (45.83 %) SNPs were disease causing. Fifteen SNPs (V345I, A390T, V600 M, E331Q, E331 K, P334S, P339L, R344G, S362G, S387F, E391 K, V383G, V600E, V607A, and N608I) were not classified using the PANTHER tool. The results of all three bioinformatics analysis tools were compatible with each other (Table 3).

Disease-Related SNPs

PhD-SNP and SNP&GO, which are SVM-based classifiers, were used in the in silico analysis of disease-related *FMR1* gene SNPs. The SVM method (SVM-Profile) classifies mutations into disease causing and neutral, using a vector of two elements derived from the sequence profile as inputs [43]. SNPs&GO is a web server for the prediction of human disease-related single-point protein mutations. SNPs&GO is one of the best scoring classifiers available for predicting whether a mutation at the protein level is disease related [44]. The prediction of the pathogenic effect of FMR1 SNPs using SVMs included SNPs predicted to be deleterious by at least two of the tools (SIFT, PolyPhen, and PANTHER). Eighty-nine (74.16 %) SNPs and 31 (25.83 %) SNPs were determined, respectively, as neutral and disease-related SNPs in PhD-SNP and SNP&GO analyses (Table 3). Based on the results of the SNP&GO analysis, 82 (68.33 %) SNPs were neutral, and 38 (31.66 %) SNPs were disease related. According to the results of the PhD-SNP and SNP&GO analyses, 13 (C99R, E75V, A153P, R180L, G231D, A233D, A235V, E257V, L279P, I304N, N401T, R410H, and R421Q) of the *FMR1* SNPs were strongly disease related (score $\geq$0.800) (Table 3). The results of both bioinformatics analyses were compatible with each other.

Table 2 Molecular interfacing analysis of FMR1 protein #2bkd and #2qnd 3D structures

Residues	HSDC	BSA (Å^2)	ASA (Å^2)	Residues	HSDC	BSA (Å^2)	ASA (Å^2)
ASP 22		10.15	78.25	VAL 308	H	16.93	16.93
VAL 23	C	14.77	68.57	ASP 309	H	59.12	99.22
GLU 25	C	43.58	183.99	GLY 312		15.96	49.21
ASP 26		15.19	99.25	VAL 313	H	11.66	12.40
SER 27	H	5.49	29.63	VAL 314	H	37.29	55.53
GLN 41		0.29	108.95	ARG 315	H S	52.33	110.74
PHE 44		0.94	39.56	VAL 316	H	26.97	26.97
PHE 126		4.53	122.72	ARG 317	S	45.24	80.34
ARG 290	H	89.35	151.66	ILE 318	H	56.85	63.75
ASN 291		36.52	98.21	GLU 319		39.66	68.97
LEU 292		0.00	12.47	ALA 320	H	96.90	105.94
VAL 293		15.72	19.24	GLU 321	H	33.05	64.11
GLY 294		31.21	54.14	ASN 322	H	81.97	87.38
ILE 297		42.02	47.70	GLU 323	H	30.13	157.00
ASN 300	H	109.51	133.08	LYS 324		8.29	168.93
GLY 301		9.20	18.25	ASN 325	H	33.49	66.10
ILE 304		9.21	13.06	VAL 333		1.34	11.54
GLN 305	H	64.80	92.76	–	–	–	–

HSDC bond types hydrogen/disulfide/salt/covalent, *BSA* burried surface area, *ASA* accessible surface area, *Å^2* angstrom

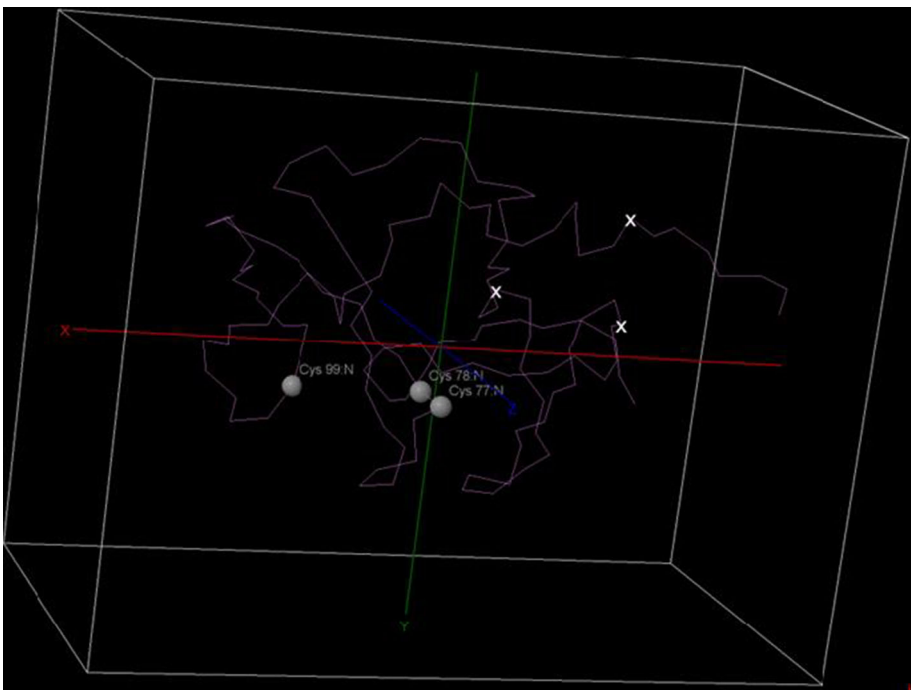

Fig. 1 Display of the cysteines on the structure of FMR1 protein N-terminal domain with ConSurf Tool. X, Y, and Z axes were shown with *red*, *green,* and *blue*, respectively. The modified histidinol residues of the protein were specified with X (as a *white*) (Color figure online)

SNP-Induced Changes in Protein Stability

In the present study, the *FMR1* SNPs were analyzed using the I-Mutant2.0 bioinformatics analysis tool [45]. In mutation analyses, the prediction of protein stability is a very important parameter. Many protein design and analysis techniques incorporate point mutations with increased or decreased stability. The I-Mutant2.0 identified 89 (74.16 %) SNPs that decreased protein stability. Besides, it was found that 12 of 13 strongly disease-related mutations which identified in PhD-SNP and SNP&GO analyses decreased protein stability (Table 3).

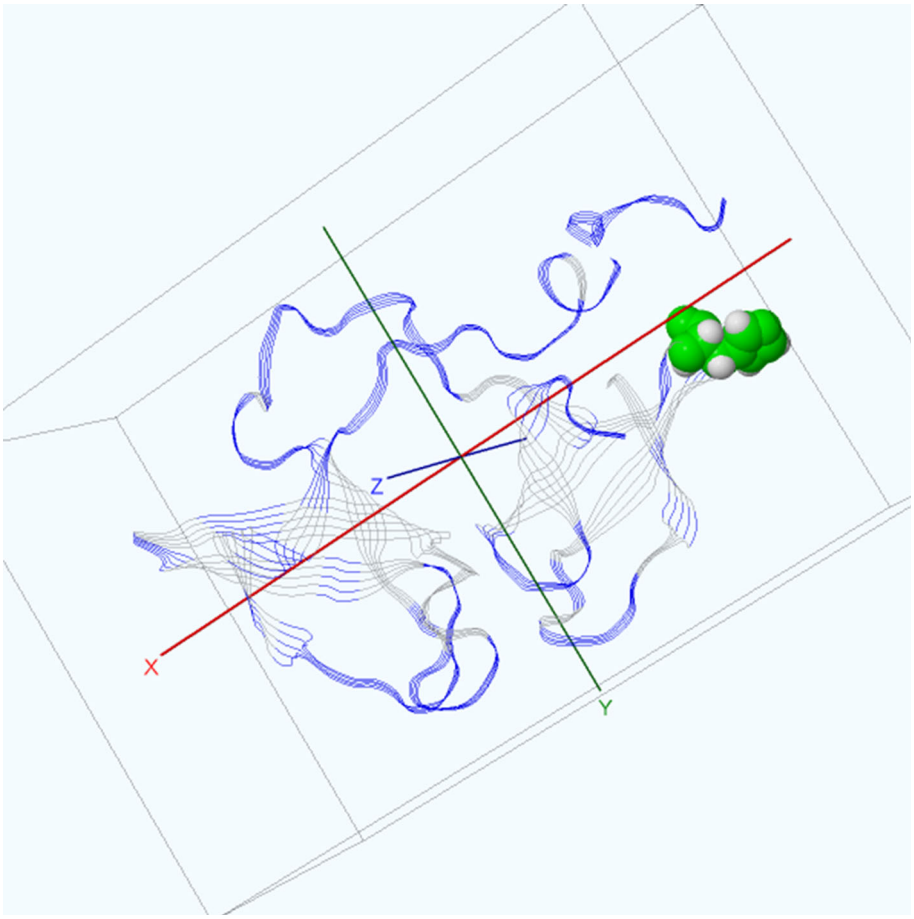

Fig. 2 The visualization of the FMR1 protein motifs (*blue*). The interfaces of the N-terminal domain (with *green*) of protein were displayed. The *gray lines*, X (*red*), Y (*green*), and Z (*blue*) axes shows the spatial position of the molecule (Color figure online)

SNP-Induced Changes in the Structure and Function of Proteins

HOPE presents the data on the effects of a mutation in a format that can be easily understood by those without a bioinformatics background. HOPE is an easy-to-use webserver, which analyzes the structural effects of a mutation [46]. When a protein sequence and mutation are input, HOPE collects and combines available information from a series of webservers and databases and produces a mutation report complete with results, figures, and animations. Using information from the Protein Data Bank, HOPE analyzed the structural effect of three mutations (R138Q [PDB#4ova], G266E [PDB#2qnd], and I304N [PDB#2qnd]).

The R138Q mutation resulted in amino acid changes from arginine to glutamine, and the mutant residue was smaller than that of the wild-type residue, potentially leading to the loss of protein–protein interactions. The wild-type residue was positively charged, and the mutant residue was neutral. Loss of the charge of the wild-type residue can result in the loss of interactions with other molecules or residues. This mutation matches some clinical features seen in, for example, developmental delays. The severity of the effects of this variant is not clear. The wild-type residue was strongly conserved, but some other residue types were observed at the same position. Homologous proteins were present at the same position as the mutant residue, and this mutation possibly did not have a damaging effect on the protein, as shown in Table 3. The mutant residue was located near a highly conserved position (Fig. 4a1–a3).

The G266E mutation resulted in amino acid changes from glycine to glutamic acid. The mutant residue was larger than the wild-type residue. The wild-type residue was buried in the core of the protein. The wild-type residue was neutral, and the mutant residue was negatively charged, possibly leading to protein-folding problems. The wild-type residue was more hydrophobic than the mutant residue. The wild-type residue was a glycine, the most flexible of all residues. This flexibility might be necessary for the protein's function. Mutation of this glycine abolished the protein's function. The wild-type residue was not conserved at this position. Homologous proteins existed with that other residue type than with the wild-type residue in the protein sequence. The other residue type was

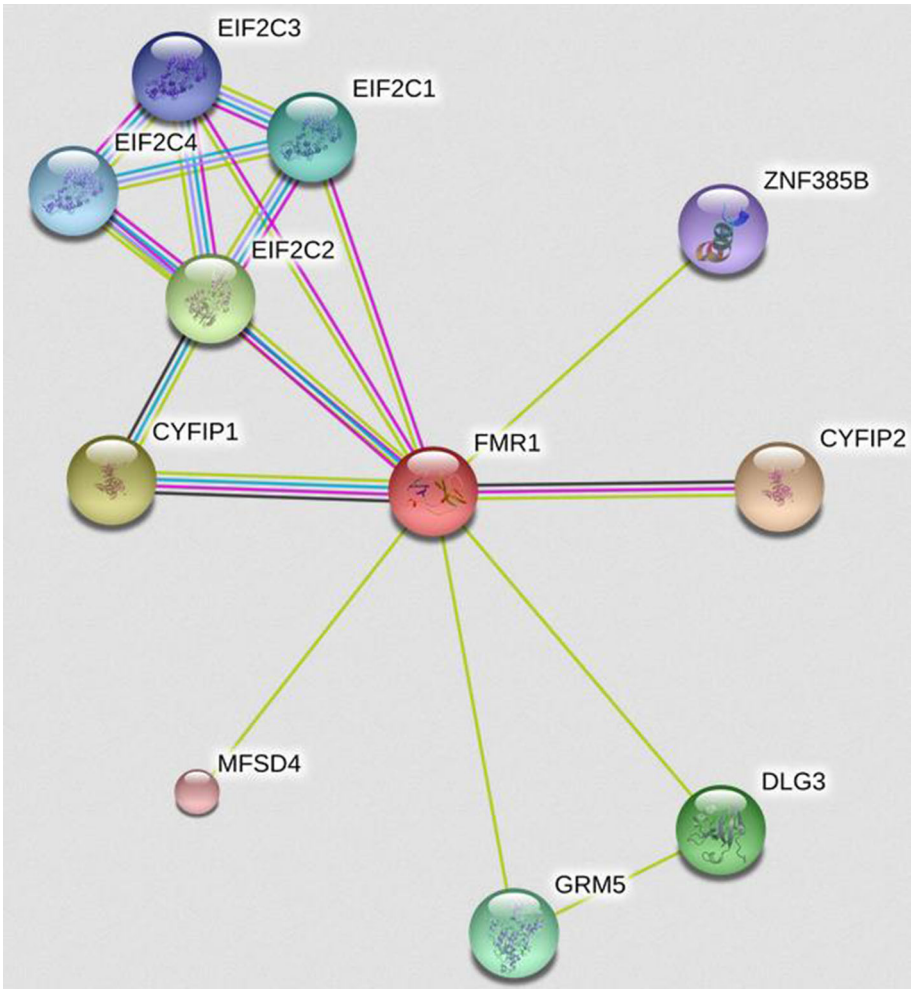

Fig. 3 STRING database generated protein interaction network of FMR1 protein

dissimilar to the mutant residue. Therefore, the mutation was possibly damaging, as shown in Table 3. The mutated residue was located in a domain that is important for binding of other molecules and in contact with residues in a domain that is also important for binding. The mutation might disturb the interaction between these two domains and therefore affect the function of the protein (Fig. 4b1–b3).

The I304 N mutation resulted in amino acid changes from isoleucine to asparagine in I304N. In common with the G266E mutation, the mutant residue was larger than the wild-type residue. The wild-type residue was buried in the core of the protein. Based on the PDB-file and the PISA-assembly, this residue showed multimer contact. As the PISA database contains protein assemblies that are highly likely to be biologically relevant, this results strongly suggested that the residue was in contact with other proteins. The increase in the size of the mutated residue at this position may disturb multimeric interactions. The residue was less hydrophobic as a result of the mutation. As hydrophobicity can be important for multimerization, this mutation could affect multimer interactions. The I304N mutation is the same as that seen in FXS. The wild-type residue was not conserved at this position. Homologous proteins existed with that other residue type than with the wild-type residue in this protein sequence. The other residue type was dissimilar to the mutant residue. Therefore, the mutation was possibly damaging, as shown in Table 3. The mutant residue was located near a highly conserved position. The mutated residue was located in a domain that is important for binding of other molecules and in contact with residues in a domain that is important for binding. Thus, this mutation might disturb the interaction between these two domains and, as such, affect the function of the protein (Fig. 4c1–c3).

Discussion

FXS syndrome is one of the most prevalent genetic causes of MR. Although FXS is caused by a genetic mutation in a single gene, it produces a phenotypically complex disorder,

Table 3 Results of the prediction and protein stability analysis methods obtained by analysis of known missense mutations and SNPs of FMR1 gene

Mutation	Type	AA	SIFT (score)	PolyPhen (score)	PANTHER (score)	PhD-SNP (score)	SNPs&GO (score)	I-Mutant2.0 (score)
rs201498256	SNP	Y68F	Tolerated (0.53)	Benign (0.276)	Neutral (0.181)	Neutral (0.107)	Neutral (0.040)	Increase (0)
rs112763380	SNP	C99R	Deleterious (0.01)	Probably damaging (0.939)	Disease (0.790)	Disease (0.898)	Disease (0.916)	Decrease (5)
rs201580891	SNP	K119N	Tolerated (0.05)	Possibly damaging (0.895)	Neutral (0.418)	Neutral (0.380)	Neutral (0.326)	Decrease (3)
rs200163413	SNP	R138Q	Tolerated (0.17)	Benign (0.358)	Disease (0.780)	Neutral (0.281)	Neutral (0.454)	Decrease (8)
rs29281	SNP	A145S	Tolerated (1)	Benign (0)	Neutral (0.160)	Neutral (0.025)	Neutral (0.025)	Decrease (5)
rs398123678	SNP	H147R	Deleterious (0)	Probably damaging (0.991)	Disease (0.631)	Disease (0.663)	Disease (0.824)	Increase (0)
rs201326944	SNP	S217W	Deleterious (0)	Probably damaging (0.992)	Disease (0.930)	Disease (0.649)	Disease (0.632)	Increase (5)
rs200721607	SNP	A239T	Deleterious (0.05)	Possibly damaging (0.646)	Neutral (0.432)	Neutral (0.382)	Neutral (0.442)	Decrease (7)
rs139029212	SNP	K273R	Tolerated (0.49)	Benign (0.036)	Neutral (0.147)	Neutral (0.156)	Neutral (0.051)	Decrease (5)
rs143161663	SNP	S277G	Tolerated (0.37)	Benign (0.001)	Neutral (0.228)	Neutral (0.149)	Neutral (0.068)	Decrease (6)
rs121434622	SNP	I304N	Deleterious (0)	Probably damaging (0.998)	Disease (0.891)	Disease (0.878)	Disease (0.891)	Decrease (5)
rs147478734	SNP	V345I	Tolerated (0.29)	Benign (0.004)	Unclassified	Neutral (0.034)	Neutral (0.014)	Decrease (3)
rs2187601	SNP	A390T	Tolerated (0.4)	Benign (0.001)	Unclassified	Neutral (0.335)	Neutral (0.157)	Decrease (7)
rs182830086	SNP	A478S	Tolerated (0.12)	Possibly damaging (0.641)	Neutral (0.176)	Neutral (0.156)	Neutral (0.134)	Increase (4)
rs144392181	SNP	E500D	Tolerated (0.35)	Benign (0.012)	Disease (0.558)	Neutral (0.273)	Neutral (0.229)	Increase (3)
rs143889976	SNP	A434V	Tolerated (0.09)	Benign (0.024)	Disease (0.752)	Neutral (0.364)	Disease (0.537)	Decrease (6)
rs367795320	SNP	L505P	Tolerated (0.26)	Benign (0.012)	Disease (0.588)	Neutral (0.461)	Neutral (0.388)	Decrease (4)
rs145953697	SNP	R512Q	Tolerated (0.06)	Possibly damaging (0.491)	Neutral (0.488)	Neutral (0.471)	Neutral (0.392)	Increase (1)
rs371495007	SNP	E458A	Tolerated (0.45)	Benign (0)	Neutral (0.186)	Neutral (0.338)	Neutral (0.136)	Decrease (8)
rs139801134	SNP	T463M	Tolerated (0.41)	Benign (0.002)	Neutral (0.369)	Neutral (0.310)	Neutral (0.228)	Decrease (4)
rs186789410	SNP	D533V	Deleterious (0)	Benign (0.255)	Disease (0.659)	Neutral (0.472)	Disease (0.624)	Increase (2)
rs369023505	SNP	P417Q	Tolerated (0.29)	Benign (0)	Disease (0.895)	Disease (0.850)	Disease (0.863)	Decrease (9)
rs372019441	SNP	R537Q	Tolerated (0.33)	Benign (0.056)	Neutral (0.242)	Neutral (0.346)	Neutral (0.184)	Increase (4)
rs376588908	SNP	R541C	Deleterious (0.01)	Probably damaging (0.961)	Disease (0.580)	Neutral (0.366)	Neutral (0.381)	Increase (2)
rs369585221	SNP	I475T	Tolerated (0.44)	Benign (0)	Neutral (0.123)	Neutral (0.261)	Neutral (0.125)	Decrease (3)

Table 3 continued

Mutation	Type	AA	SIFT (score)	PolyPhen (score)	PANTHER (score)	PhD-SNP (score)	SNPs&GO (score)	I-Mutant2.0 (score)
rs587780338	SNP	C563Y	Tolerated (0.23)	Benign (0.003)	Disease (0.727)	Neutral (0.381)	Neutral (0.304)	Decrease (6)
rs201041299	SNP	S579I	Deleterious (0.01)	Benign (0.054)	Neutral (0.403)	Disease (0.558)	Neutral (0.313)	Decrease (2)
rs45540244	SNP	T529I	Deleterious (0)	Probably damaging (0.926)	Disease (0.598)	Neutral (0.337)	Neutral (0.299)	Decrease (1)
rs376406395	SNP	V600M	Tolerated (0.07)	Benign (0.027)	Unclassified	Neutral (0.243)	Neutral (0.093)	Increase (4)
rs372396040	SNP	S537L	Deleterious (0)	Possibly damaging (0.514)	Neutral (0.278)	Neutral (0.295)	Neutral (0.167)	Increase (5)
COSM311247	SNP	E25V	Deleterious (0)	Probably damaging (0.913)	Disease (0.762)	Disease (0.632)	Disease (0.751)	Increase (2)
COSM1466194	SNP	A31T	Tolerated (0.25)	Benign (0.016)	Neutral (0.168)	Neutral (0.393)	Neutral (0.221)	Decrease (7)
COSM3233433	SNP	P38A	Tolerated (0.13)	Benign (0.006)	Neutral (0.314)	Neutral (0.341)	Neutral (0.244)	Decrease (8)
COSM1490574	SNP	I59L	Tolerated (0.2)	Benign (0.005)	Disease (0.443)	Neutral (0.556)	Neutral (0.468)	Decrease (7)
COSM373883	SNP	E75V	Deleterious (0)	Probably damaging (0.997)	Disease (0.876)	Disease (0.760)	Disease (0.853)	Decrease (2)
COSM3843791	SNP	E95K	Tolerated (0.08)	Benign (0.037)	Disease (0.431)	Neutral (0.793)	Disease (0.721)	Decrease (7)
COSM1116459	SNP	R113K	Deleterious (0)	Probably damaging (0.984)	Disease (0.404)	Neutral (0.576)	Disease (0.562)	Decrease (8)
COSM4156642	SNP	K119N	Tolerated (0.05)	Possibly damaging (0.895)	Neutral (0.418)	Neutral (0.380)	Neutral (0.326)	Decrease (3)
COSM1715720	SNP	P120S	Tolerated (0.07)	Benign (0.097)	Neutral (0.203)	Neutral (0.377)	Neutral (0.292)	Decrease (9)
COSM1466196	SNP	A121V	Tolerated (1)	Benign (0.105)	Neutral (0.108)	Neutral (0.088)	Neutral (0.047)	Increase (0)
COSM3973350	SNP	F126S	Deleterious (0)	Probably damaging (0.999)	Neutral (0.757)	Neutral (0.730)	Neutral (0.830)	Decrease (9)
COSM123202	SNP	I129M	Deleterious (0.05)	Benign (0.051)	Neutral (0.325)	Neutral (0.270)	Neutral (0.225)	Decrease (8)
COSM260968	SNP	R138Q	Tolerated (0.17)	Benign (0.358)	Neutral	Disease	Neutral	Decrease (8)
COSM1556914	SNP	C141F	Deleterious (0)	Possibly damaging (0.535)	Disease (0.780)	Disease (0.281)	Disease (0.454)	Decrease (1)
COSM1556913	SNP	K148N	Deleterious (0.02)	Possibly damaging (0.614)	Disease (0.650)	Neutral (0.487)	Disease (0.570)	Increase (5)
COSM3363851	SNP	A153P	Deleterious (0)	Probably damaging (0.991)	Disease (0.732)	Disease (0.821)	Disease (0.820)	Decrease (2)
COSM756016	SNP	R180L	Deleterious (0.01)	Possibly damaging (0.644)	Disease (0.724)	Disease (0.798)	Disease (0.798)	Decrease (5)
COSM3939836	SNP	M183I	Tolerated (0.23)	Possibly damaging (0.646)	Neutral (0.124)	Neutral (0.309)	Neutral (0.154)	Decrease (7)
COSM3379546	SNP	R190Q	Deleterious (0)	Probably damaging (0.998)	Disease (0.764)	Disease (0.648)	Disease (0.715)	Decrease (6)
COSM161113	SNP	R193H	Deleterious (0.01)	Benign (0.445)	Disease (0.704)	Disease (0.554)	Disease (0.555)	Decrease (8)
COSM1756430	SNP	T194I	Deleterious (0)	Possibly damaging (0.773)	Disease (0.652)	Disease (0.644)	Disease (0.598)	Decrease (5)
COSM4107275	SNP	L196V	Tolerated (0.22)	Benign (0.421)	Neutral (0.234)	Neutral (0.135)	Neutral (0.101)	Decrease (8)

Table 3 continued

Mutation	Type	AA	SIFT (score)	PolyPhen (score)	PANTHER (score)	PhD-SNP (score)	SNPs&GO (score)	I-Mutant2.0 (score)
COSM1682943	SNP	S197Y	Deleterious (0.03)	Benign (0.254)	Neutral (0.316)	Neutral (0.364)	Neutral (0.137)	Increase (3)
COSM1569654	SNP	I199V	Deleterious (0.03)	Benign (0.209)	Neutral (0.201)	Neutral (0.090)	Neutral (0.047)	Decrease (6)
COSM1176797	SNP	E204K	Deleterious (0)	Probably damaging (0.95)	Disease (0.734)	Neutral (0.384)	Disease (0.609)	Decrease (4)
COSM1556912	SNP	S206C	Tolerated (0.06)	Possibly damaging (0.879)	Disease (0.550)	Neutral (0.102)	Neutral (0.060)	Increase (3)
COSM1625590	SNP	S206G	Deleterious (0.03)	Benign (0.205)	Neutral (0.298)	Neutral (0.097)	Neutral (0.044)	Decrease (5)
COSM1466197	SNP	M230L	Tolerated (0.05)	Possibly damaging (0.835)	Neutral (0.414)	Disease (0.721)	Disease (0.707)	Decrease (7)
COSM1116460	SNP	G231D	Deleterious (0)	Probably damaging (1)	Disease (0.876)	Disease (0.889)	Disease (0.886)	Decrease (8)
COSM4156644	SNP	A233D	Deleterious (0)	Probably damaging (0.996)	Disease (0.833)	Disease (0.901)	Disease (0.886)	Decrease (6)
COSM1726869	SNP	G235V	Deleterious (0)	Probably damaging (1)	Disease (0.899)	Disease (0.800)	Disease (0.830)	Decrease (1)
COSM1116461	SNP	V250A	Deleterious (0)	Probably damaging (0.954)	Disease (0.626)	Neutral (0.428)	Disease (0.523)	Decrease (9)
COSM367831	SNP	L255I	Deleterious (0.02)	Probably damaging (0.994)	Disease (0.551)	Disease (0.586)	Disease (0.520)	Decrease (7)
COSM1556911	SNP	E257V	Deleterious (0)	Probably damaging (0.996)	Disease (0.870)	Disease (0.817)	Disease (0.902)	Increase (2)
COSM1116462	SNP	T261A	Deleterious (0.01)	Probably damaging (0.996)	Disease (0.576)	Disease (0.558)	Disease (0.621)	Decrease (7)
COSM1466198	SNP	G266E	Deleterious (0)	Probably damaging (1)	Disease (0.879)	Disease (0.722)	Disease (0.761)	Decrease (2)
COSM488058	SNP	E267D	Deleterious (0.02)	Possibly damaging (0.723)	Disease (0.606)	Disease (0.570)	Disease (0.661)	Decrease (7)
COSM1116463	SNP	S277R	Tolerated (0.21)	Benign (0.034)	Neutral (0.442)	Neutral (0.374)	Neutral (0.315)	Increase (5)
COSM1116464	SNP	L279P	Deleterious (0)	Probably damaging (0.998)	Disease (0.912)	Disease (0.878)	Disease (0.878)	Decrease (9)
COSM1116465	SNP	E280K	Deleterious (0)	Probably damaging (1)	Disease (0.734)	Disease (0.811)	Disease (0.844)	Decrease (8)
COSM3964840	SNP	V285L	Tolerated (0.34)	Benign (0.053)	Neutral (0.188)	Neutral (0.285)	Neutral (0.177)	Decrease (7)
COSM456904	SNP	Q287E	Deleterious (0.03)	Benign (0.086)	Disease (0.369)	Neutral (0.617)	Disease (0.626)	Decrease (1)
COSM1116466	SNP	E319K	Deleterious (0)	Probably damaging (0.966)	Disease (0.479)	Neutral (0.553)	Disease (0.519)	Decrease (9)
COSM1116467	SNP	E321D	Tolerated (1)	Benign (0.086)	Neutral (0.321)	Neutral (0.051)	Neutral (0.020)	Decrease (8)
COSM1756431	SNP	E330Q	Deleterious (0.04)	Possibly damaging (0.726)	Neutral (0.367)	Neutral (0.171)	Neutral (0.141)	Decrease (4)
COSM1556910	SNP	E331Q	Tolerated (0.35)	Benign (0.361)	Unclassified	Neutral (0.070)	Neutral (0.039)	Decrease (4)
COSM260969	SNP	E331K	Tolerated (0.34)	Benign (0.201)	Unclassified	Neutral (0.107)	Neutral (0.076)	Decrease (5)
COSM756014	SNP	P334S	Tolerated (0.27)	Benign (0.234)	Unclassified	Neutral (0.259)	Neutral (0.104)	Decrease (8)
COSM1580267	SNP	P339L	Tolerated (0.1)	Benign (0.067)	Unclassified	Neutral (0.300)	Neutral (0.098)	Decrease (2)

Table 3 continued

Mutation	Type	AA	SIFT (score)	PolyPhen (score)	PANTHER (score)	PhD-SNP (score)	SNPs&GO (score)	I-Mutant2.0 (score)
COSM1556909	SNP	R344G	Tolerated (0.36)	Benign (0.204)	Unclassified	Neutral (0.281)	Neutral (0.182)	Decrease (8)
COSM1116468	SNP	S362G	Tolerated (0.21)	Benign (0.021)	Unclassified	Neutral (0.227)	Neutral (0.124)	Decrease (7)
COSM3558878	SNP	S387F	Deleterious (0.01)	Benign (0.035)	Unclassified	Neutral (0.233)	Neutral (0.126)	Increase (7)
COSM161114	SNP	E391K	Tolerated (0.66)	Benign (0.04)	Unclassified	Neutral (0.220)	Neutral (0.152)	Decrease (7)
COSM4156646	SNP	V383G	Deleterious (0)	Probably damaging (0.992)	Unclassified	Neutral (0.496)	Neutral (0.253)	Decrease (10)
COSM1116469	SNP	N401T	Tolerated (0.21)	Possibly damaging (0.712)	Disease (0.866)	Disease (0.876)	Disease (0.918)	Decrease (8)
COSM266551	SNP	R410H	Tolerated (0.07)	Possibly damaging (0.505)	Disease (0.929)	Disease (0.918)	Disease (0.918)	Decrease (9)
COSM70802	SNP	R421Q	Deleterious (0.02)	Possibly damaging (0.743)	Disease (0.828)	Disease (0.818)	Disease (0.840)	Decrease (7)
COSM756013	SNP	R433P	Deleterious (0.04)	Benign (0.042)	Disease (0.855)	Disease (0.740)	Disease (0.759)	Increase (6)
COSM3558880	SNP	E437K	Tolerated (0.19)	Benign (0.011)	Disease (0.866)	Disease (0.780)	Disease (0.867)	Decrease (8)
COSM1116471	SNP	P454L	Tolerated (0.08)	Probably damaging (1)	Neutral (0.498)	Disease (0.610)	Disease (0.516)	Decrease (7)
COSM1116472	SNP	G461S	Tolerated (0.34)	Benign (0.05)	Neutral (0.113)	Neutral (0.476)	Neutral (0.288)	Decrease (7)
COSM1252632	SNP	R462G	Tolerated (0.28)	Probably damaging (0.999)	Neutral (0.314)	Neutral (0.341)	Neutral (0.234)	Decrease (9)
COSM205745	SNP	R463C	Deleterious (0.01)	Probably damaging (1)	Disease (0.530)	Neutral (0.440)	Neutral (0.347)	Decrease (7)
COSM3913517	SNP	S479F	Deleterious (0)	Possibly damaging (0.861)	Disease (0.892)	Disease (0.608)	Disease (0.790)	Increase (0)
COSM456905	SNP	E488K	Tolerated (0.08)	Probably damaging (0.999)	Neutral (0.210)	Neutral (0.493)	Neutral (0.385)	Decrease (4)
COSM3372112	SNP	S493L	Tolerated (0.1)	Benign (0.159)	Disease (0.677)	Neutral (0.363)	Disease (0.502)	Increase (6)
COSM4107281	SNP	E499K	Deleterious (0.04)	Possibly damaging (0.759)	Neutral (0.300)	Neutral (0.202)	Neutral (0.143)	Decrease (7)
COSM260970	SNP	H506R	Deleterious (0.01)	Probably damaging (0.979)	Neutral (0.161)	Neutral (0.051)	Neutral (0.014)	Decrease (5)
COSM1116473	SNP	E438D	Deleterious (0)	Possibly damaging (0.771)	Disease (0.615)	Neutral (0.499)	Disease (0.517)	Increase (6)
COSM260971	SNP	T440M	Tolerated (0.33)	Probably damaging (0.978)	Disease (0.914)	Neutral (0.476)	Disease (0.723)	Decrease (3)
COSM1116474	SNP	R511W	Deleterious (0.02)	Possibly damaging (0.88)	Disease (0.885)	Neutral (0.484)	Neutral (0.482)	Increase (4)
COSM3390469	SNP	R512W	Deleterious (0)	Probably damaging (0.975)	Disease (0.900)	Neutral (0.461)	Disease (0.587)	Decrease (1)
COSM1116475	SNP	G517E	Deleterious (0.04)	Benign (0.038)	Neutral (0.128)	Neutral (0.302)	Neutral (0.102)	Decrease (0)
COSM1556907	SNP	H535L	Tolerated (0.33)	Benign (0.009)	Neutral (0.220)	Neutral (0.287)	Neutral (0.132)	Increase (4)
COSM1116476	SNP	R537Q	Tolerated (0.33)	Benign (0.056)	Neutral (0.242)	Neutral (0.346)	Neutral (0.184)	Increase (4)
COSM4107282	SNP	R541H	Deleterious (0.03)	Probably damaging (0.944)	Neutral	Neutral	Neutral	Increase (1)

Table 3 continued

Mutation	Type	AA	SIFT (score)	PolyPhen (score)	PANTHER (score)	PhD-SNP (score)	SNPs&GO (score)	I-Mutant2.0 (score)
COSM1490575	SNP	R543G	Deleterious (0.02)	Possibly damaging (0.863)	Disease (0.513)	Disease (0.588)	Disease (0.682)	Increase (4)
COSM1116477	SNP	P487L	Tolerated (0.07)	Benign (0.001)	Disease (0.545)	Neutral (0.457)	Neutral (0.476)	Decrease (1)
COSM611324	SNP	R583L	Deleterious (0.01)	Possibly damaging (0.611)	Neutral (0.167)	Disease (0.825)	Neutral (0.294)	Decrease (6)
COSM205746	SNP	R585C	Deleterious (0)	Probably damaging (0.999)	Disease (0.529)	Disease (0.784)	Neutral (0.464)	Decrease (3)
COSM1331191	SNP	R585H	Tolerated (0.1)	Probably damaging (0.999)	Neutral (0.254)	Disease (0.770)	Neutral (0.311)	Decrease (9)
COSM1116478	SNP	T586M	Tolerated (0.1)	Benign (0.058)	Neutral (0.408)	Disease (0.726)	Neutral (0.377)	Increase (2)
COSM4107283	SNP	R517Q	Deleterious (0)	Possibly damaging (0.781)	Neutral (0.184)	Neutral (0.345)	Neutral (0.131)	Decrease (7)
COSM1116479	SNP	R569C	Deleterious (0)	Probably damaging (0.998)	Neutral (0.459)	Neutral (0.419)	Neutral (0.160)	Decrease (4)
COSM85250	SNP	E595D	Tolerated (0.45)	Benign (0.065)	Neutral (0.313)	Disease (0.528)	Neutral (0.182)	Decrease (7)
COSM3363853	SNP	P597T	Tolerated (1)	Benign (0.002)	Neutral (0.381)	Neutral (0.245)	Neutral (0.201)	Increase (4)
COSM756011	SNP	V600E	Tolerated (0.38)	Benign (0.131)	Unclassified	Neutral (0.351)	Neutral (0.196)	Increase (8)
COSM1116480	SNP	V607A	Deleterious (0)	Probably damaging (0.997)	Unclassified	Neutral (0.144)	Neutral (0.096)	Decrease (9)
COSM3780367	SNP	N608I	Deleterious (0)	Probably damaging (1)	Unclassified	Neutral (0.161)	Neutral (0.097)	Decrease (4)
COSM1556908	sSubs	G515M	Deleterious (0)	Probably damaging(0.976)	Neutral (0.475)	Neutral (0.209)	Neutral (0.169)	Decrease (7)

with neurological and psychiatric features. In the majority of cases, the mutation consists of an expansion of a CGG trinucleotide repeat within the 5′ UTR of the *FMR1* gene [47]. The behavioral overlap between FXS and ASD is so common that approximately 72 % of patients with FXS were reported to exhibit ASD symptoms in different studies [48–50]. At the cellular level, FXS is associated with immature dendritic spine morphology. FMRP is an essential protein for synaptic development and plasticity, because it is a key negative regulator, which can down-regulate or upregulate mRNA synthesis and synaptic protein synthesis. Using Fmr1-knockout animal models, researchers demonstrated increased translation and protein synthesis in the hippocampus [16, 48, 49]. In our previous report, we showed that FMR1 gene premutations led to premature ovarian failure in women [10]. Other studies reported that some missense mutations, such as R138Q, G266E, I304 N, G482S, and R534H, in exons, introns, and the 3′ UTR region of the *FMR1* gene caused FXS [1, 12–16]. None of the previous studies investigated missense mutations of the *FMR1* gene. Therefore, the effects of missense mutations of the *FMR1* gene in FXS were unknown. The present study is the first attempt to analyze all missense SNPs of the *FMR1* gene. The results suggest that bioinformatics approaches can reveal important information about missense mutations in FXS- and other *FMR1* gene-related diseases, such as premature ovarian failure.

Using different sequencing methods, the number of identified missense mutations in the human genome has accumulated in databases. Processing the vast amount of data that exists on genetic variants requires in silico analysis tools [51]. According to some studies, the determination of genetic variants will have important consequences for future therapies and personalized medicine [13, 22, 40, 51]. Multiple sequence alignments are widely used in bioinformatics analyses [32]. These provide information on phylogenetic trees, structure prediction, and critical residues [34]. In the present study, the majority of *FMR1* gene missense mutations occurred in domains showing high

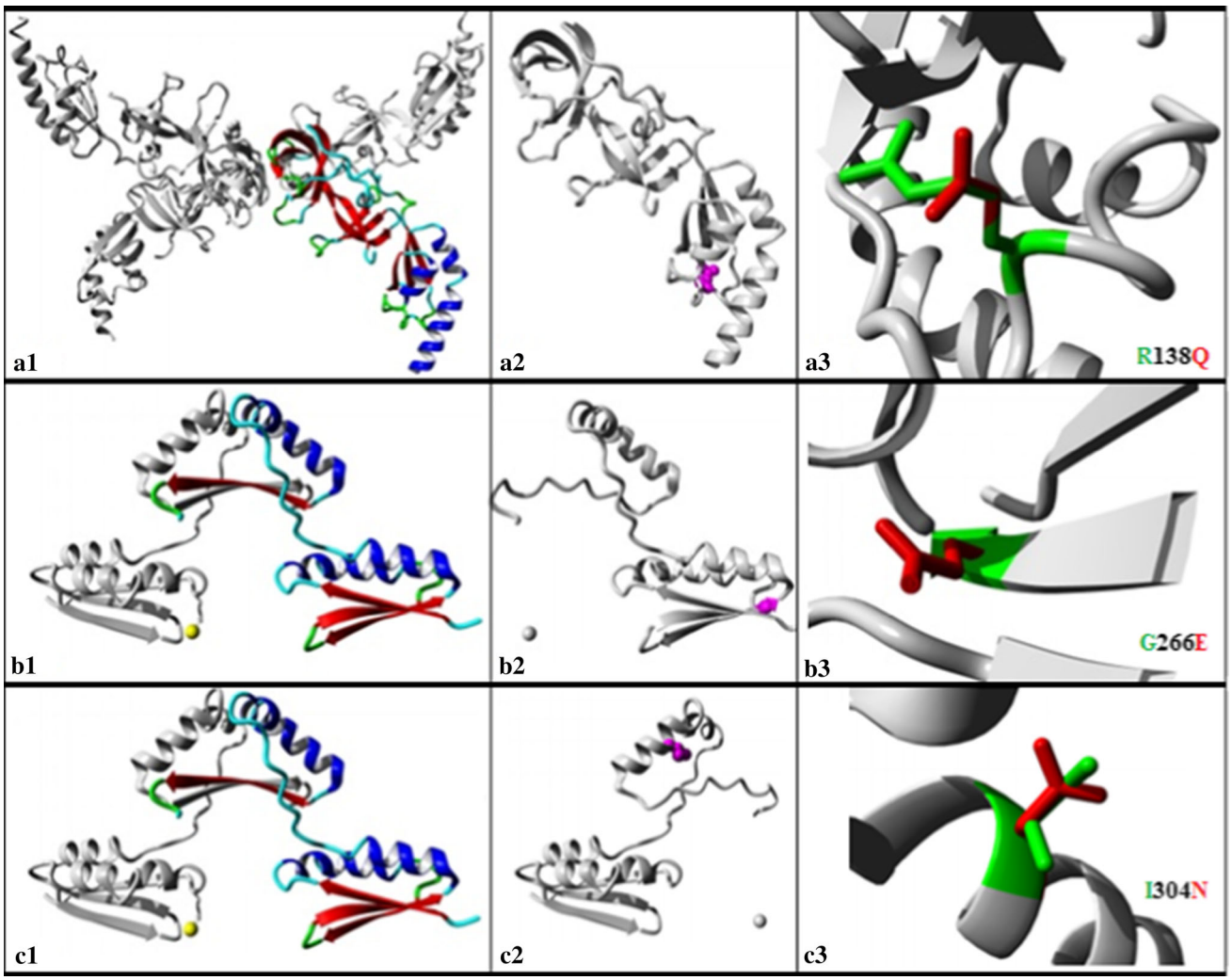

Fig. 4 The visualization of 3D structural information of R138Q (PDB#4ova), G266E (PDB#2qnd), and I304N (PDB#2qnd) SNPs. **a–c** Represents the R138Q, G266E, and I304N mutations, respectively. **a.1**, **b.1, c.1** The protein is colored by element; α-helix: blue, β-strand: *red*, turn: *green*, 3/10 helix: *yellow*, random coil: *cyan*, and other molecules: *gray*. **a.2**, **b.2, c.2** The protein is colored *gray*. The side chain of mutated residue is colored *magenta*. **a.3**, **b.3, c.3** The protein is colored *gray*. The side chains both wild-type and mutant residue are shown and colored *green* and *red*, respectively (Color figure online)

variety, especially the Agenet and KH domains. In bioinformatics approaches, macromolecular interfaces are classified as "biologically relevant" or "insignificant" (crystal packing) according to a scoring system. The calculation of the type of interface depends on the interface area and atomic composition, hydrophobicity, charge, and topological complementarity of the residue [38, 52]. FMRP performs its tasks through interactions with several protein partners. The majority of these protein partners are connected to the amino terminus of the protein by an independently folded domain, termed the N-terminal domain of FMR. Thus, missense SNPs that are located in this region likely have damaging effects (Fig. 2) [53].

In the present study, the results of the interface analysis of the PDB#2bkd and PDB#2qnd 3D structures revealed 35 interface residues. Four of these were missense mutations (F126S, I304 N, E319 K, and E321D) (Tables 2, 3). These mutations can damage protein formation. These mutations can be produced experimentally using site-directed mutagenesis and similar techniques. However, this is time-consuming and often requires the use of computational prediction methods to select the best possible combinations [54]. Thermodynamic data on proteins are essential for understanding the mechanism of protein folding and stability and for designing stable mutants. The results of analyses of thermodynamic data, together with sequence and structural information, can provide a valuable resource for developing algorithms to elucidate the mechanism of protein folding and stability and to predict mutation-induced changes in protein stability [45]. In the present study, the I-Mutant2.0 web server was used to analyze protein stability changes caused by a single-point mutation.

The results showed that the free energy of 74.16 % of the *FMR1* gene SNPs was decreased (I-Mutant2.0). To the best of our knowledge, this is the first study to analyze the effects of FMR1 gene SNPs on protein stability.

Prediction approaches can provide the analysis of numerous SNPs in a short time [51]. In this study, the results of the prediction methods revealed that 31.66 % of *FMR1* gene SNPs were disease related (PhD-SNP, SNP&GO) and that 50 % had a pathogenic effect (SIFT, PolyPhen, and PANTHER). Although there were some discrepancies in the prediction analysis, on the whole, the results of the mutation analyses overlapped. To eliminate the discrepancies in the prediction analysis and identify which SNPs were harmful, an algorithm was used, and the SNPs that can be considered disease related or pathogenic at were identified using at least two of the bioinformatics tools. Using the prediction methods and structural analyses, the results showed that R138Q, G266E, and I304N SNPs of the FMR1 gene led to a loss of protein function. Other studies reported that these SNPs were associated with FXS [1, 12, 14–16]. The FMRP shows both biochemical and genetic interactions with components of the miRNA pathway, suggesting that it may be involved in translational suppression [3]. A previous study found G266E and I304 N missense mutations in FXS patients with a normal number CGG repeats [12]. These mutations might disturb the interaction between domains that are important for binding of other molecules and that are in contact with residues, thereby affecting the function of the protein. In accordance with the findings of this study, the results of a molecular docking analysis showed that *FMR1* I304N SNP affected two binding sites located on the KH2 domain and concluded that this might lead to a loss of protein function [55]. Another study reported that the number of CGG repeats was not increased in a male patient with development delay who had a missense mutation of the FMR1 gene, R138Q, in a highly conserved region [56]. Handt et al. stated that undetected mutations of the *FMR1* gene might account for FXS-like phenotypes and MR phenotypes [13]. In the present study, the STRING analysis of protein–protein interactions determined that the FMR1 protein was closely associated with cytoplasmic FMRP interacting protein 1 and 2 protein (CYFIP1 and CYFI), which are candidates for ID, autism, and FXS [57]. The findings indicate that missense mutations that damage interactions between proteins and their domains may cause FXS-like phenotypes.

Conclusions

In this study, missense mutations of the *FMR1* gene were identified, and their sequences, functions, thermodynamics, and structural characteristics were evaluated using 14 bioinformatics methods. In silico approaches allow large numbers of mutations to be analyzed at the same time and simulations of the predicted effects of missense mutations at the protein level. The present study revealed the effects of all missense mutations of the *FMR1* gene. It showed that approximately 30–50 % of all *FMR1* gene missense SNPs are associated with diseases and that these mutations disrupt protein structure and function. The structural analysis demonstrated that R138Q, G266E, and I304 N missense SNPs were closely associated with FXS-like phenotypes. An analysis of *FMR1* gene missense SNPs could help diagnose FXS and MR in patients with a normal number of CGG repeats.

References

1. Chen, E., & Joseph, S. (2015). Fragile X mental retardation protein: a paradigm for translational control by RNA-binding proteins. *Biochimie, 114*, 147–154.
2. Laggerbauer, B., Ostareck, D., Keidel, E. M., Ostareck-Lederer, A., & Fischer, U. (2001). Evidence that fragile X mental retardation protein is a negative regulator of translation. *Hum. Mol. Genet., 10*(4), 329–338.
3. Jin, P., Zarnescu, D. C., Ceman, S., Nakamoto, M., Mowrey, J., Jongens, T. A., et al. (2004). Biochemical and genetic interaction between the fragile X mental retardation protein and the microRNA pathway. *Nat. Neurosci., 7*(2), 113–117.
4. Hinds, H. L., Ashley, C. T., Sutcliffe, J. S., Nelson, D. L., Warren, S. T., Housman, D. E., & Schalling, M. (1993). Tissue specific expression of FMR-1 provides evidence for a functional role in fragile X syndrome. *Nat. Genet., 3*, 36–43.
5. Ascano, M., Mukherjee, N., Bandaru, P., Miller, J. B., Nusbaum, J. D., Corcoran, D. L., et al. (2012). FMRP targets distinct mRNA sequence elements to regulate protein expression. *Nature, 492*, 382–386.
6. Zhou, H., Mangelsdorf, M., Liu, J., Zhu, L., & Wu, J. Y. (2014). RNA-binding proteins in neurological diseases. *Sci China Life Sci, 57*(4), 432–444.
7. Darnell, J. C., & Richter, J. D. (2012). Cytoplasmic RNA-binding proteins and the control of complex brain function. *Cold Spring Harb. Perspect. Biol., 4*(8), 012344.
8. Bhakar, A. L., Dolen, G., & Bear, M. F. (2012). The pathophysiology of fragile X (and what it teaches us about synapses). *Annu. Rev. Neurosci., 35*, 417–443.
9. Valverde, R., Pozdnyakova, I., Kajander, T., Venkatraman, J., & Regan, L. (2007). Fragile X mental retardation syndrome: structure of the KH1-KH2 domains of fragile X mental retardation protein. *Structure, 15*(9), 1090–1098.
10. Tural, S., Tekcan, A., Kara, N., Elbistan, M., Güven, D., & Ali Tasdemir, H. (2015). FMR1 gene mutation screening by TP-PCR in patients with premature ovarian failure and fragile-X. *Gynecol. Endocrinol., 31*(3), 191–195.
11. Greco, C. M., Berman, R. F., Martin, R. M., Tassone, F., Schwartz, P. H., Chang, A., et al. (2006). Neuropathology of fragile X-associated tremor/ataxia syndrome (FXTAS). *Brain, 129*, 243–255.
12. Luo, S., Huang, W., Xia, Q., Du, Q., Wu, L., & Duan, R. (2015). Mutational analyses of the FMR1 gene in Chinese pediatric population of fragile x suspects: low tolerance for-point mutation. *J. Child Neurol., 30*(6), 803–806.
13. Handt, M., Epplen, A., Hoffjan, S., Mese, K., Epplen, J. T., & Dekomien, G. (2014). Point mutation frequency in the FMR1

gene as revealed by fragile X syndrome screening. *Mol. Cell. Probes, 28*(5–6), 279–283.
14. Myrick, L. K., Nakamoto-Kinoshita, M., Lindor, N. M., Kirmani, S., Cheng, X., & Warren, S. T. (2014). Fragile X syndrome due to a missense mutation. *Eur. J. Hum. Genet., 22*(10), 1185–1189.
15. Wang, T., Bray, S. M., & Warren, S. T. (2012). New perspectives on the biology of fragile X syndrome. *Curr. Opin. Genet. Dev., 22*(3), 256–263.
16. Myrick, L. K., Deng, P. Y., Hashimoto, H., Oh, Y. M., Cho, Y., Poidevin, M. J., et al. (2015). Independent role for presynaptic FMRP revealed by an FMR1 missense mutation associated with intellectual disability and seizures. *Proc Natl Acad Sci USA, 112*(4), 949–956.
17. Cunningham, F., Amode, M. R., Barrell, D., Beal, K., Billis, K., Brent, S., et al. (2015). Ensembl 2015. *Nucleic Acids Res, 43*((Database issue)), 662–669.
18. Yilmaz, A. (2014). Bioinformatic Analysis of GJB2 Gene Missense Mutations. *Cell Biochem. Biophys., 71*, 1623–1642.
19. Cardona, F., Tormos-Pérez, M., & Pérez-Tur, J. (2014). Structural and functional in silico analysis of LRRK2 missense substitutions. *Mol. Biol. Rep., 41*(4), 2529–2542.
20. Divanshu, G., Lekshmi, M., & Shanthi, V. (2014). In silico studies of deleterious non-synonymous single nucleotide polymorphisms (nsSNPs) of *NRL* gene. *Netw. Model. Anal. Health Inform. Bioinform., 3*(59), 1–7.
21. Doss, C. G., Chakraborty, C., Chen, L., & Zhu, H. (2014). Integrating in silico prediction methods, molecular docking, and molecular dynamics simulation to predict the impact of ALK missense mutations in structural perspective. *Biomed. Res. Int., 19*(13), 1–14.
22. Li, B., Seligman, C., Thusberg, J., Miller, J. L., Auer, J., Whirl-Carrillo, M., et al. (2014). In silico comparative characterization of pharmacogenomic missense variants. *BMC Genom. Suppl., 4*, 4.
23. Raza, S. I., Muhammad, D., Jan, A., Ali, R. H., Hassan, M., Ahmad, W., & Rashid, S. (2014). In silico analysis of missense mutations in LPAR6 reveals abnormal phospholipid signaling pathway leading to hypotrichosis. *PLoS One, 9*(8), 104756.
24. Sherry, S. T., Ward, M. H., Kholodov, M., Baker, J., Phan, L., Smigielski, E. M., & Sirotkin, K. (2001). dbSNP: the NCBI database of genetic variation. *Nucleic Acids Res., 29*(1), 308–311.
25. 1000 Genomes Project Consortium, Abecasis, G. R., Auton, A., Brooks, L. D., DePristo, M. A., Durbin, R. M., et al. (2012). An integrated map of genetic variation from 1,092 human genomes. *Nature, 491*(7422), 56–65.
26. http://evs.gs.washington.edu/EVS/.
27. UniProt Consortium (2015) UniProt: a hub for protein information. Nucleic Acids Res 43(Database issue):204-12.
28. Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., et al. (2000). The Protein Data Bank. *Nucleic Acids Res., 28*(1), 235–242.
29. Montague, E., Stanberry, L., Higdon, R., Janko, I., Lee, E., Anderson, N., et al. (2014). MOPED 2.5–an integrated multi-omics resource: multi-omics profiling expression database now includes transcriptomics data. *OMICS, 18*(6), 335–343.
30. http://web.expasy.org/blast/.
31. Artimo, P., Jonnalagedda, M., Arnold, K., Baratin, D., Csardi, G., de Castro, E., et al. (2012). ExPASy: SIB bioinformatics resource portal. *Nucleic Acids Res, 40*(Web Server issue), 597–603.
32. Larkin, M. A., Blackshields, G., Brown, N. P., Chenna, R., McGettigan, P. A., McWilliam, H., et al. (2007). Clustal W and Clustal X version 2.0. *Bioinformatics, 23*(21), 2947–2948.
33. Notredame, C., Higgins, D. G., & Heringa, J. (2000). T-coffee: T-COFFEE: a novel method for fast and accurate multiple sequence alignment. *J. Mol. Biol., 302*(1), 205–217.
34. Edgar, R. C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res., 32*(5), 1792–1797.
35. Crooks, G. E., Hon, G., Chandonia, J. M., & Brenner, S. E. (2004). WebLogo: a sequence logo generator. *Genome Res., 14*(6), 1188–1190.
36. Finn, R. D., Bateman, A., Clements, J., Coggill, P., Eberhardt, R. Y., Eddy, S. R., et al. (2014). Pfam: the protein families database. *Nucleic Acids Res, 42*(Database issue), 222–230.
37. Celniker, G., Nimrod, G., Ashkenazy, H., Glaser, F., Martz, E., Mayrose, I., et al. (2013). ConSurf: using evolutionary data to raise testable hypotheses about protein function. *Isr. J. Chem., 53*, 199–206.
38. Krissinel, E., & Henrick, K. (2007). Inference of macromolecular assemblies from crystalline state. *J. Mol. Biol., 372*(3), 774–797.
39. http://www.jmol.org/.
40. Kumar, P., Henikoff, S., & Ng, P. C. (2009). Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. *Nat. Protoc., 4*(7), 1073–1081.
41. Mi, H., Lazareva-Ulitsky, B., Loo, R., Kejariwal, A., Vandergriff, J., Rabkin, S., et al. (2005). The PANTHER database of protein families, subfamilies, functions and pathways. *Nucleic Acids Res, 33*(Database issue), 284–288.
42. Ramensky, V., Bork, P., & Sunyaev, S. (2002). Human non-synonymous SNPs: server and survey. *Nucleic Acids Res., 30*(17), 3894–3900.
43. 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.
44. 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., 30*(8), 1237–1244.
45. Bava, K. A., Gromiha, M. M., Uedaira, H., Kitajima, K., & Sarai, A. (2004). ProTherm, version 4.0: thermodynamic database for proteins and mutants. *Nucleic Acids Res, 32*(Database issue), 120–121.
46. Venselaar, H., Te, Beek T. A., Kuipers, R. K., Hekkelman, M. L., & Vriend, G. (2010). Protein structure analysis of mutations causing inheritable diseases. An e-Science approach with life scientist friendly interfaces. *BMC Bioinform., 11*, 548.
47. Kaufmann, W. E., Cohen, S., Sun, H. T., & Ho, G. (2002). Molecular phenotype of fragile X syndrome: FMRP, FXRPs, and protein targets. *Microsc Res Tech., 57*(3), 135–144.
48. Kazdoba, T. M., Leach, P. T., Silverman, J. L., & Crawley, J. N. (2014). Modeling fragile X syndrome in the Fmr1 knockout mouse. *Intractable Rare Dis. Res., 3*(4), 118–133.
49. Lozano, R., Rosero, C. A., & Hagerman, R. J. (2014). Fragile X spectrum disorders. *Intractable Rare Dis. Res., 3*(4), 134–146.
50. Rosti, R. O., Sadek, A. A., Vaux, K. K., & Gleeson, J. G. (2014). The genetic landscape of autism spectrum disorders. *Dev. Med. Child Neurol., 56*(1), 12–18.
51. Kumar, A., Rajendran, V., Sethumadhavan, R., Shukla, P., Tiwari, S., & Purohit, R. (2014). Computational SNP analysis: current approaches and future prospects. *Cell Biochem. Biophys., 68*(2), 233–239.
52. Henrick, K., & Thornton, J. (1998). PQS: a protein quaternary structure file server. *Trends Biochem. Sci., 23*, 358–361.
53. Ramos, A., Hollingworth, D., Adinolfi, S., Castets, M., Kelly, G., Frenkiel, T. A., et al. (2006). The structure of the N-terminal domain of the fragile X mental retardation protein: a platform for protein-protein interaction. *Structure, 14*(1), 21–31.
54. Kumar, M. D., Bava, K. A., Gromiha, M. M., Prabakaran, P., Kitajima, K., Uedaira, H., & Sarai, A. (2006). ProTherm and

ProNIT: thermodynamic databases for proteins and protein-nucleic acid interactions. *Nucleic Acids Res., 1*(34), 204–206.
55. Marino, D., Achsel, T., Lacoux, C., Falconi, M., & Bagni, C. (2014). Molecular dynamics simulations show how the FMRP Ile304Asn mutation destabilizes the KH2 domain structure and affects its function. *J. Biomol. Struct. Dyn., 32*(3), 337–350.
56. Collins, S. C., Bray, S. M., Suhl, J. A., Cutler, D. J., Coffee, B., Zwick, M. E., & Wrren, S. T. (2010). Identification of novel FMR1 variants by massively parallel sequencing in developmentally delayed males. *Am. J. Med. Genet. A, 152*, 2512–2520.
57. Abekhoukh, S., & Bardoni, B. (2014). CYFIP family proteins between autism and intellectual disability: links with Fragile X syndrome. *Front. Cell. Neurosci., 8*, 81.