Available online at www.sciencedirect.com
ScienceDirect
Genomics 90 (2007) 447–452
GENOMICS
www.elsevier.com/locate/ygeno

Identification and in silico analysis of functional SNPs of the *BRCA1* gene

R. Rajasekaran, C. Sudandiradoss, C. George Priya Doss, Rao Sethumadhavan *

School of Biotechnology, Chemical, and Biomedical Engineering, Vellore Institute of Technology University, Vellore 632014, Tamil Nadu, India

Received 8 June 2007; accepted 8 July 2007

Abstract

Single-nucleotide polymorphisms (SNPs) play a major role in the understanding of the genetic basis of many complex human diseases. Also, the genetics of human phenotype variation could be understood by knowing the functions of these SNPs. It is still a major challenge to identify the functional SNPs in a disease-related gene. In this work, we have analyzed the genetic variation that can alter the expression and the function of the *BRCA1* gene using computational methods. Of the total 477 SNPs, 65 were found to be nonsynonymous (ns) SNPs. Among the 14 SNPs in the untranslated region, 4 were found in the 5′ and 10 were found in the 3′ untranslated region (UTR). It was found that 16.9% of the nsSNPs were damaging, by both the SIFT and the PolyPhen servers. The UTR Resource tool suggested that 2 of 4 SNPs in the 5′ UTR and 3 of 10 SNPs in the 3′ UTR might change the protein expression levels. We identified major mutations from proline to serine at positions 1776 and 1812 of the native protein of the *BRCA1* gene. From a comparison of the stabilizing residues of the native and mutant proteins, we propose that an nsSNP (rs1800751) could be an important candidate for the breast cancer caused by the *BRCA1* gene.

Keywords: *BRCA1* gene; Breast cancer; SIFT; PolyPhen; FastSNP; Modeled structure

Single-nucleotide polymorphisms (SNPs) account for the more common form of human genetic variation. About 500,000 SNPs fall in the coding regions of the human genome [1]. Among these, the nonsynonymous SNPs (nsSNPs) cause changes in the amino acid residues. These are likely to be an important factor contributing to the functional diversity of the encoded proteins in the human population [2]. The nsSNPs affect gene regulation by altering DNA and transcriptional binding factors [3] and the maintenance of the structural integrity of cells and tissues [4]. Also, nsSNPs affect the functional roles of proteins in the signal transduction of visual, hormonal, and other stimulants [5,6].

Mutations in the cancer susceptibility gene *BRCA1* greatly increase the risk of breast and ovarian cancer [7]. At present most mutations in the *BRCA1* gene have been identified to be point mutations or small insertions and deletions. An important database, http://research.nhgri.nih.gov/bic/, gives a world of information on breast cancer. Although our literature survey showed that there is a wide choice of literature on the *BRCA1* gene related to breast cancer, there have been no computational studies undertaken for an in silico investigation of the nsSNP mutations in *BRCA1*. We undertook this work mainly to perform a computational analysis of the nsSNPs in the *BRCA1* gene, to identify the possible mutations and propose a modeled structure for the mutant protein. We report that the mutation from proline to serine at the residue position of 1812 in the native protein of *BRCA1* gene could be a candidate of major concern for the disease of breast cancer caused by the *BRCA1* gene.

Results and discussion

SNP dataset from dbSNP

The *BRCA1* gene investigated in this work was retrieved from the dbSNP database [8]. It contained a total of 477 SNPs, of which 65 were nsSNPs and 14 were in noncoding regions, which comprise 4 SNPs in the 5′ UTR and 10 SNPs in the 3′ UTR. The rest were in the intron region. We selected nonsynonymous coding SNPs and 5′ and 3′ UTR SNPs for our investigation. The distribution of nsSNPs in coding regions and SNPs in the UTRs is shown in Fig. 1. It can be seen from Fig. 1 that 13% of the total SNPs are nsSNPs, and 0.8 and 2% of the total SNPs are in the 5′ and 3′ UTRs. Further it can be seen that the number of nsSNPs in

* Corresponding author. Fax: +91 4162243092.
E-mail address: sethu_rao@bsnl.in (R. Sethumadhavan).

0888-7543/$ - see front matter © 2007 Elsevier Inc. All rights reserved.
doi:10.1016/j.ygeno.2007.07.004

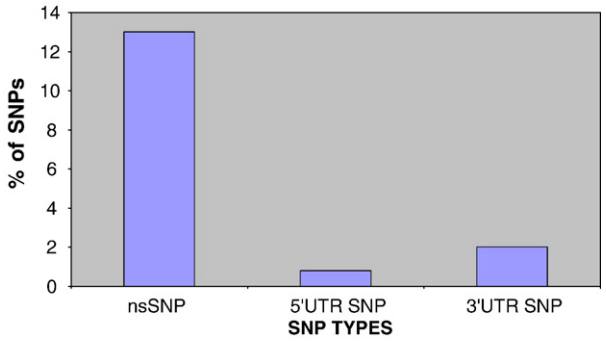

Fig. 1. Distribution of nonsynonymous, 5′ UTR, and 3′ UTR SNPs.

the coding region is much higher compared to the SNPs in the 5′ and 3′ untranslated regions.

Deleterious nsSNP found by the SIFT program

The conservation level of a particular position in a protein was determined by using a sequence homology-based tool, SIFT [9]. The protein sequences of 65 nsSNPs were submitted independently to the SIFT program to check its tolerance index. The higher the tolerance index, the less functional impact a particular amino acid substitution is likely to have, and vice versa. Among the 65 nsSNPs, 28 were found to be deleterious, having a tolerance index score of ≤0.05. The results are shown in Table 1.

Table 1
List of nsSNPs that were predicted to have functional significance by SIFT

SNP ID	Nucleotide change	Amino acid change	Tolerance index
rs16941	A/G	E1038G	0.03
rs799917	C/T	P871L	0.03
rs1799950	A/G	Q315R	0.02
rs1799966	A/G	S1613G	0.02
rs1799967	G/A	M1652I	0.01
rs1800707	G/T	K1406N	0.01
rs1800709	C/T	R800W	0.00
rs1800726	G/C	A1641P	0.02
rs1800744	G/T	S1512I	0.01
rs1800751	C/T	P1812S	0.00
rs1800757	C/T	P1776S	0.00
rs4986847	A/C	I925L	0.02
rs7502059	G/A	A794V	0.02
rs8176153	G/A	G275S	0.03
rs11658785	C/A	R256M	0.03
rs12946486	T/G	K1338Q	0.01
rs28897672	G/T	C61G	0.00
rs28897673	A/G	Y105C	0.04
rs28897674	A/C	S153R	0.03
rs28897696	A/C	A564E	0.01
rs28897698	G/T	V1809F	0.00
rs28897699	G/T	Q1857H	0.02
rs28897680	A/C	Q687P	0.03
rs28897681	G/T	D654Y	0.01
rs28897682	A/T	N769Y	0.00
rs28897683	A/C	T785K	0.00
rs28897684	A/G	F801G	0.04
rs28897689	A/G	R1306G	0.01

Table 2
List of nsSNPs that were predicted to be functionally significant by PolyPhen

SNP ID	Nucleotide change	Amino acid change	PSIC SD
rs1799950	A/G	Q315R	1.965
rs1800709	C/T	R800W	1.630
rs1800751	C/T	P1812S	2.396
rs28897672	G/T	C61G	2.700
rs28897673	A/G	Y105C	2.025
rs28897681	G/T	D654Y	2.003
rs28897682	A/T	N769Y	2.074
rs28897683	A/C	T785K	1.626
rs28897684	A/G	E801G	2.144
rs28897687	G/T	N1195K	1.653
rs28897689	A/G	R1306G	2.006
rs28897695	G/T	F518C	1.766
rs28897696	A/C	A564E	1.125
rs4986854	C/T	M484T	1.890

We observed that, of 28 deleterious nsSNPs, 7 showed a highly deleterious tolerance index score of 0.00, and 7 showed a tolerance index score of 0.01, followed by 6, 6, and 2 nsSNPs with a tolerance index of 0.02, 0.03, and 0.04, respectively. Six nsSNPs showed a nucleotide change of A → G, 5 nsSNPs A → C, 1 nsSNP A → T, 3 nsSNPs G → A, 6 nsSNPs G → T, 1 nsSNP G → C, 4 nsSNPs C → T, 1 nsSNP C → A, and 1 nsSNP T → G. A → G and G → T nucleotide changes occurred the maximum number of times and A → T, G → C, C → A, and T → G nucleotide changes occurred a minimum number of times, as can be seen from Table 1. The nucleotide change C → T accounted for the highest number of deleterious nsSNPs, with a SIFT tolerance index of 0.00. This was closely followed by the nucleotide change G → T, which showed a tolerance index of 0.01. Also, of the 7 nsSNPs that showed a SIFT tolerance index of 0.00, 3 of them changed to an aromatic amino acid in the mutant type from a nonaromatic amino acid in the native protein.

Damaged nsSNP found by the PolyPhen server

The structural levels of alteration were determined by applying the PolyPhen program [11]. Sixty-five protein sequences of nsSNPs investigated in this work were submitted as input to the PolyPhen server and the results are shown in Table 2. A position-specific independent count (PSIC) score difference of 1.1 and above is considered to be damaging. It can be seen that, of 65 nsSNPs, 14 were considered to be damaging. All 14 nsSNPs exhibited a PSIC score difference in the range 1.125 to 2.396.

Twelve nsSNPs that were observed to be deleterious by the SIFT program also were damaging according to PolyPhen. Hence, we could infer that the results obtained on the basis of sequence details (SIFT) were in good correlation with the results obtained for structural details (PolyPhen), as can be seen from Tables 1 and 2. It can be seen from Tables 1 and 2 that 4 nsSNPs (rs1800751, rs28897672, rs28897682, rs28897684) had a SIFT tolerance index of 0.00 and PSIC score difference ≥2.00. Hence the mutations occurring with these 4 nsSNPs would be of prime importance in the identification of breast cancer caused by the *BRCA1* gene, according to SIFT and PolyPhen results.

Functional SNPs in UTR found by the FastSNP server

Table 3 shows the list of SNPs in the 5′ untranslated region that are predicted to be functionally significant. We used the server FastSNP [16] for this purpose. According to this server, of 14 UTR SNPs of the *BRCA1* gene, 2 in the 5′ UTR, namely rs11658785 and rs8176153, were predicted to be damaging, with a risk ranking of 2–3 and 1–3, respectively. The nucleotide changes were A → C for SNP ID rs11658785 and A → G for SNP ID rs8176153. However, this server did not predict any functional significance for the 3′ UTR and hence, we used another server to check for any functional significance due to the 3′ UTR.

Functional SNPs in UTR found by the UTRscan server

Polymorphisms in the 3′ UTR affect gene expression by affecting the ribosomal translation of mRNA or by influencing the RNA half-life [25]. Table 4 shows the list of SNPs in the 3′ untranslated region that are predicted to be of functional significance. We used the UTRscan server for this purpose [19]. We analyzed the same 14 UTR SNPs in UTRscan that were analyzed by the FastSNP server. The UTRscan server finds patterns of regulatory region motifs from the UTR database and gives information about whether the matched pattern is damaged. We found that three UTR SNPs, namely, rs11655841, rs8176318, and rs8176317, were predicted to be damaging by this server, as can be seen from Table 4.

The 15-lipoxygenase (15-LOX) differentiation control element (15-LOX-DICE) controls 15-LOX synthesis, which catalyzes the degradation of lipids and is an important factor responsible for the degradation of mitochondria during reticulocyte maturation. This 15-LOX-DICE exists in the three 3′ UTR SNPs that were considered to be of functional significance and hence can be thought to be damaging to the *BRCA1* gene. We used two different servers to get the functional SNPs in both 5′ and 3′ UTRs. We observed that two SNPs in the 5′ UTR and three SNPs in the 3′ UTR were predicted to be of functional significance as per the FastSNP and UTRscan servers, respectively.

Modeling of mutant structure

Mapping the deleterious nsSNPs into protein structure information was performed through the Single Amino Acid Polymorphism database (SAAPdb) [21]. The available structure for the *BRCA1* gene has the PDB ID 1jnx.

Table 3
List of SNPs (UTR mRNA) predicted to be functionally significant by FastSNP

SNP ID	Nucleotide change	UTR position	Level of risk	Possible functional effect
rs11658785	A/C	5′ UTR	Low-medium (2–3)	Splicing regulation
rs8176153	A/G	5′ UTR	Very low-medium (1–3)	Promoter/regulatory region

Table 4
List of SNPs (UTR mRNA) predicted to be functionally significant by UTRscan

SNP ID	Nucleotide change	UTR position	Functional element change
rs11655841	C/G	3′ UTR	15-LOX-DICE → no pattern
rs8176318	G/T	3′ UTR	15-LOX-DICE → no pattern
rs8176317	A/G	3′ UTR	15-LOX-DICE → no pattern

According to this resource, the mutation occurred for 1jnx mainly at two SNP Ids, namely, rs1800751 and rs1800757. The mutations were at residue position 1812 (P → S) and at position 1776 (P → S). The proline-to-serine mutations for 1jnx at positions 1812 and 1776 were performed by the SWISSPDB viewer independently to get two modeled structures. Then energy minimizations were performed by the NOMAD-Ref server [22] for the native-type protein (PDB 1jnx) and the two mutant-type proteins 1jnx (P1812S) and 1jnx (P1776S).

It can be seen from Table 5 that total energy for the native-type structure (PDB 1jnx) and the two mutant-type structures 1jnx (P1812S) and 1jnx (P1776S) was −4521.023, −10975.899, and −10707.296 Kcal/mol, respectively. Table 5 also shows that the RMSD values between the native type (1jnx) and the mutant type 1jnx (P1812S) is 2.99 Å and between the native type 1jnx and the mutant type 1jnx (P1776) is 2.95 Å. The higher the RMSD value is, the more the deviation between the two structures is, which in turn changes their functional activity. Since the RMSD values are higher for two mutant-type structures compared to the native-type structure, 1jnx, these two nsSNPs could be believed to affect the structure of the proteins. These two nsSNPs were also shown to be deleterious according to the SIFT program and, one of the nsSNPs, i.e., rs1800751, was shown to be damaging according to the PolyPhen server.

It can be seen from Table 5 that the both the RMSD value and the total energy of the mutant 1jnx (P1776S) are slightly lesser than the RMSD and total energy of mutant 1jnx (P1812S). Therefore the PolyPhen server, which is based on structural details, does not predict rs1800757 to be damaging. This analysis portrays that mutant type P1812 with the SNP ID rs1800751 would be expected to be more deleterious and damaging compared to mutant type 1jnx (P1776S) with the SNP ID rs1800757. The superimposed structures of the native protein 1jnx with the two mutant-type proteins (P1812S and P1776S) are shown in Figs. 2A and 2B, respectively.

Computing stabilizing residues between native structure and mutant modeled structures

We used the SRide server [24] to identify the stabilizing residues of the native-type structure and mutant modeled structures. The results are shown in Table 6. Nine stabilizing residues were identified in the native-type 1jnx structure. Seven stabilizing residues were identified in mutant model 1jnx (P1812S). Ten stabilizing residues were identified in mutant model 1jnx (P1776S). Two stabilizing residues, namely Val (1714) and Glu(1713), were found to be common to both native-structure 1jnx and mutant model 1jnx (P1812S). The other stabilizing residues of the native structure were not present in

Table 5
RMSD and total energy of native-structure 1jnx and mutant models

Parameter	1jnx native-type structure	1jnx 1812 mutant (rs1800751) with 1jnx native-type structure	1jnx 1776 mutant (rs1800757) with 1jnx native-type structure
RMSD of entire structure		2.99 Å	2.95 Å
Total energy after energy minimization	−4521.023 Kcal/mol	−10,975.899 Kcal/mol	−10,707.296 Kcal/mol

mutant model 1jnx (P1812S). Four stabilizing residues, namely Val(1714), Ser(1715), Glu(1735), and Glu(1794), were found to be common to both native-structure 1jnx and mutant model-structure 1jnx (P1776S). The remaining stabilizing residues of the native structure were not seen in mutant model 1jnx (P1776S). This analysis revealed that a higher number of stabilizing residues in the mutant-type 1jnx (P1776S) matched with the native protein structure compared to the mutant-type 1jnx (P1812S). Therefore we predict that the mutation from proline to serine at residue position 1812 in the native-type protein will be more deleterious and the mutation in this SNP could be an important candidate for breast cancer caused by the *BRCA1* gene.

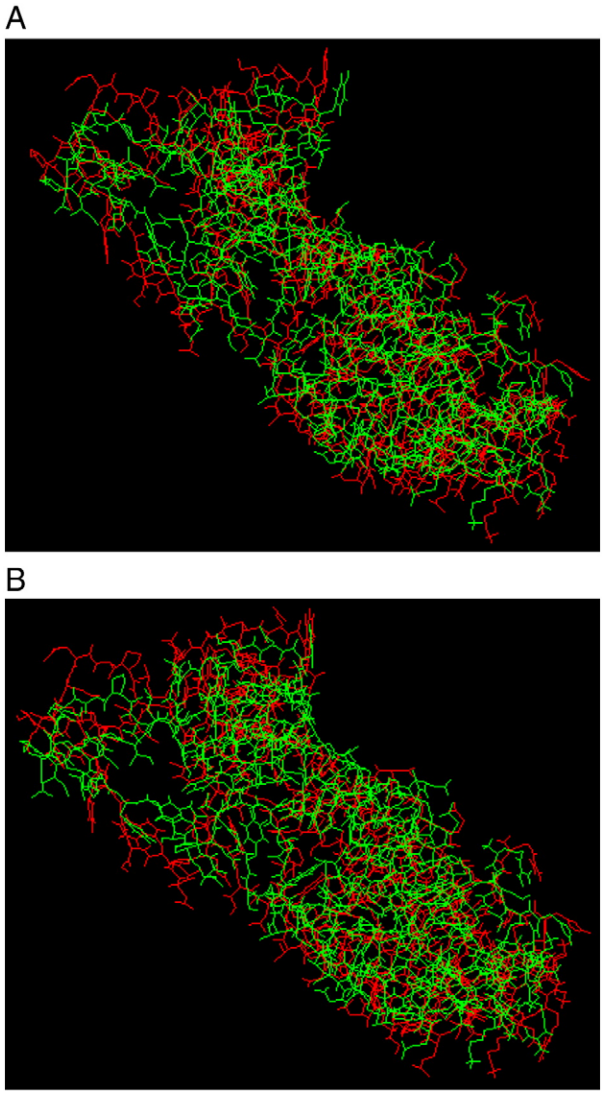

Fig. 2. (A) Superimposed structure of native protein 1jnx (red color) with mutant protein P1812S (green color). (B) Superimposed structure of native protein 1jnx (red color) with mutant protein P1776S (green color).

Conclusion

The breast cancer *BRCA1* gene was investigated in this work by evaluating the influence of functional SNPs through computation methods. Of a total of 477 SNPs in the *BRCA1* gene, 65 were found to be nonsynonymous and 4 and 10 SNPs were found to be in the 5′ and 3′ untranslated regions. Of 65 nsSNPs, 28 were found to be deleterious by SIFT and 14 were damaging as per the PolyPhen server. Twelve nsSNPs were found to be common in both the SIFT and the PolyPhen server. Two SNPs in the 5′ UTR and 3 SNPs in the 3′ UTR were found to be of functional significance. It was found that the major mutation in the native protein of the *BRCA1* gene was from proline to serine. Of two nsSNPs that had this mutation, we conclude that rs1800751 with a mutation of proline to serine at position 1812 in the native protein 1jnx could be the main target mutation for the breast cancer caused by the *BRCA1* gene.

Materials and methods

Datasets

The SNPs and their related protein sequence for the *BRCA1* gene were obtained from the dbSNP [8] (http://www.ncbi.nlm.nih.gov/SNP/) for our computational analysis.

Analysis of functional consequences of coding nsSNPs by sequence-homology-based method (SIFT)

We used the program SIFT [9] available at http://blocks.fhcrc.org/sift/SIFT.html to detect the deleterious coding nonsynonymous SNPs. SIFT is a sequence-

Table 6
Stabilizing residues in native and mutant models of 1jnx

Native-type structure 1jnx	Mutant model structure 1jnx (P1812S)	Mutant model structure 1jnx (P1776S)
Met(1689), Val(1713), **Val(1714)**, **Ser(1715)**, **Glu(1735)**, **Glu(1794)**, Val(1833), Trp(1837), Leu(1854)	Val(1687), Leu(1705), **Val(1714)**, **Glu(1735)**, Cys(1787), Val(1792), Gln(1857	Val(1687), Leu(1705), **Val(1714)**, **Ser(1715)**, **Glu(1735)**, Val(1736), Cys(1787), **Glu(1794)**, Gln(1811), Gln(1857

Residues shown in boldface are common to both native protein and mutant-type structures.

homology-based tool that presumes that important amino acids will be conserved in the protein family. Hence, changes at well-conserved positions tend to be predicted as deleterious [9]. We submitted the query in the form of SNP IDs or as protein sequences. The underlying principle of this program is that SIFT takes a query sequence and uses multiple alignment information to predict tolerated and deleterious substitutions for every position of the query sequence. SIFT is a multistep procedure that, given a protein sequence, (a) searches for similar sequences, (b) chooses closely related sequences that may share similar functions, (c) obtains the multiple alignment of the chosen sequences, and (d) calculates normalized probabilities for all possible substitutions at each position from the alignment. Substitutions at each position with normalized probabilities less than a chosen cutoff are predicted to be deleterious and those greater than or equal to the cutoff are predicted to be tolerated [10]. The cutoff value in the SIFT program is a tolerance index of ≥ 0.05. The higher the tolerance index, the less functional impact a particular amino acid substitution is likely to have.

Simulation for functional change in coding nsSNPs by structure-homology-based method (PolyPhen)

Analyzing the damaged coding nonsynonymous SNPs at the structural level is considered to be very important to understand the functional activity of the protein of concern. We used the server PolyPhen [11], which is available at http://coot.embl.de/PolyPhen/, for this purpose. Input options for the PolyPhen server are protein sequence or SWALL database ID or accession number together with sequence position with two amino acid variants. We submitted the query in the form of protein sequence with mutational position and two amino acid variants. Sequence-based characterization of the substitution site, profile analysis of homologous sequences, and mapping of substitution site to a known protein three-dimensional structure are the parameters taken into account by the PolyPhen server to calculate the score. It calculates PSIC scores for each of the two variants and then computes the PSIC score difference between them. The higher the PSIC score difference is, the higher is the functional impact a particular amino acid substitution is likely to have.

Functional significance of noncoding SNPs in regulatory untranslated regions

Recent studies show that SNPs have functional effects on protein structure by a single change in the amino acid [12,13] and on transcriptional regulation [14,15]. We used the Web server FastSNP [16] available at http://fastsnp.ibms.sinica.edu.tw for predicting the functional significance of the 5′ and 3′ UTRs of the *BRCA1* gene. The FastSNP server follows the decision tree principle with external Web service access to TFSearch, which predicts whether a noncoding SNP alters the transcription factor-binding site of a gene. The score will be given by this server on the basis of levels of risk with a ranking of 0, 1, 2, 3, 4, or 5. This signifies the levels of no, very low, low, medium, high, and very high effect, respectively.

Scanning of UTR SNPs in UTR site

The 5′ and 3′ UTRs are involved in various biological processes such as posttranscriptional regulatory pathways, stability, and translational efficiency [17,18]. We used the program UTRscan [19] available at http://www.ba.itb.cnr.it/BIG/UTRScan/, which allows one to search the user-submitted sequences for any of the patterns collected in the UTR site. UTRsite is a collection of functional sequence patterns located in 5′ or 3′ UTR sequences. Briefly, two or three sequences of each UTR SNP that have a different nucleotide at an SNP position are analyzed by UTRscan, which looks for UTR functional elements by searching through user-submitted sequence data for the patterns defined in the UTRsite and UTR databases. If different sequences for each UTR SNP are found to have different functional patterns, this UTR SNP is predicted to have functional significance. The Internet resources for UTR analysis are UTRdb and UTRsite. UTRdb contains experimentally proven biological activity of functional patterns of UTR sequence from eukaryotic mRNAs [20]. The UTRsite has the data collected from UTRdb and also is continuously enriched with new functional patterns.

Modeling nsSNP locations on protein structure and their RMSD difference

Structure analysis was performed for evaluating the structural stability of native and mutant protein. We used the Web resource SAAPdb [21] to identify the protein related to the *BRCA1* gene (PDB ID 1jnx). We also confirmed the mutation positions and the mutation residues from this server. The mutation was performed by using the SWISSPDB viewer and energy minimization for 3D structures was performed by the NOMAD-Ref server [22]. This server use Gromacs as the default force field for energy minimization based on the methods of steepest descent, conjugate gradient, and L-BFGS [23]. We used the conjugate gradient method for optimizing the 3D structures. The deviation between the two structures is evaluated by their RMSD values.

Computation of stabilizing residues

To check the stability of the native and mutant modeled structures, identification of the stabilizing residues is useful. We used the server SRide [24] for identifying the stabilizing residues in native protein and mutant models. Stabilizing residues were computed using parameters such as surrounding hydrophobicity, long-range order, stabilization center, and conservation score as described by Magyar et al. [24].

Acknowledgments

The authors thank the management of Vellore Institute of Technology University for providing the facilities to carry out this work and Professor Lazar Mathew, Dean, for his encouragement. The authors also thank the reviewers for making useful suggestions for the improvement of the manuscript.

References

[1] F.S. Collins, L.D. Brooks, A DNA polymorphism discovery resource for research on human genetic variations, Genomic Res. 8 (1998) 1229–1231.
[2] E.S. Lander, The new genomics: global views of biology, Science 274 (1996) 536–539.
[3] J. Barroso, M. Gurnell, V.E. Croeley, M. Agostini, J.W. Schwabe, M.A. Soos, et al., Dominant negative mutations in human PPAR gamma associated with severe insulin resistance diabetes mellitus and hypertension, Nature 402 (1999) 880–883.
[4] R. Thomas, R. McConnell, J. Whittacker, P. Kirkpatrick, J. Bradley, R. Stanford, Identification of mutations in the repeated part of the autosomal dominant polycystic kidney disease type 1 gene PKD1 by long range PCR, Am. J. Hum. Genet. 65 (1999) 39–49.
[5] T.P. Dryja, T.L. Mcgee, L.B. Halu, G.S. Conley, J.E. Olsson, E. Reichel, et al., Mutations within the rhodopsin gene in patients with autosomal dominant retinitis pigmentosa, N. Engl. J. Med. 323 (1990) 1302–1307.
[6] E.P. Smith, J. Boyd, G.R. Frank, M. Takahashi, R.M. Cohen, B. Specker, et al., Estrogen resistance caused by a mutation in the estrogen-receptor gene in a man, N. Engl. J. Med. 331 (1994) 1056–1061.
[7] Y. Miki, J. Swensen, D. Shattuck-Eidens, et al., A strong candidate for the breast and ovarian cancer susceptibility gene *BRCA1*, Science 266 (1994) 66–71.
[8] S.T. Sherry, M.H. Ward, M. Kholodov, J. Baker, L. Phan, M. Smigielski, K. Sirotkin, dbSNP: the NCBI database of genetic variation, Nucleic Acids Res. 29 (2001) 308–311.
[9] C.P. Ng, S. Henikoff, SIFT: predicting amino acid changes that affect protein function, Nucleic Acids Res. 31 (2003) 3812–3814.
[10] C.P. Ng, S. Henikoff, Predicting deleterious amino acid substitutions, Genome Res. 11 (2001) 863–874.
[11] V. Ramensky, P. Bork, S. Sunyaev, Human non-synonymous SNPs: server and survey, Nucleic Acids Res. 30 (2002) 3894–3900.
[12] M. Cargill, D. Altshuler, J. Ireland, P. Sklar, K. Ardlie, N. Patil, N. Shaw, C.R. Lane, E.P. Lim, N. Kalyanaraman, et al., Characterization of single-

nucleotide polymorphisms in coding regions of human genes, Nat. Genet. 22 (1999) 231–238.
[13] S. Sunyaev, V. Ramensky, P. Bork, Towards a structural basis of human non-synonymous single nucleotide polymorphisms, Trends Genet. 16 (2000) 198–200.
[14] L. Prokunina, M.E. Alarcn-Riquelme, Regulatory SNPs in complex diseases: their identification and functional validation, Expert Rev. Mol. Med. (2004) 1–15.
[15] L. Prokunina, C. Castillejo-Lopez, F. Oberg, I. Gunnarsson, L. Berg, V. Magnusson, A.J. Brookes, D. Tentler, H. Kristjansdottir, G. Grondal, et al., A regulatory polymorphism in PDCD1 is associated with susceptibility to systemic lupus erythematosus in humans, Nat. Genet. 32 (2002) 666–669.
[16] H.-Y. Yuan, J.-J. Chiou, W.-H. Tseng, C.-H. Liu, C.-K. Liu, Y.-J. Lin, H.-H. Wang, A. Yao, Y.-T. Chen, C.-N. Hsu, FASTSNP: an always up-to-date and extendable service for SNP function analysis and prioritization, Nucleic Acids Res. 34 (2006) W635–W641.
[17] N. Sonenberg, mRNA translation: influence of the 5′ and 3′ untranslated regions, Curr. Opin. Genet. Dev. 4 (1994) 310–315.
[18] R. Nowak, Mining treasures from 'junk DNA', Science 263 (1994) 608–610.
[19] G. Pesole, S. Liuni, Internet resources for the functional analysis of 5′ and 3′ untranslated regions of eukaryotic mRNA, Trends Genet. 15 (1999) 378.
[20] G. Pesole, S. Liuni, G. Grillo, F. Licciulli, F. Mignone, C. Gissi, C. Saccone, UTRdb and UTRsite: specialized databases of sequences and functional elements of 5′ and 3′ untranslated regions of eukaryotic mRNAs, Nucleic Acids Res. 30 (2002) 335–340.
[21] A. Cavallo, A.C. Martin, Mapping SNPs to protein sequence and structure data, Bioinformatics 8 (2005) 1443–1450.
[22] E. Lindahl, C. Azuara, P. Koehl, M. Delarue, NOMAD-Ref: visualization, deformation and refinement of macromolecular structures based on all-atom normal mode analysis, Nucleic Acids Res. 34 (2006) W52–W56.
[23] M. Delarue, P. Dumas, On the use of low-frequency normal modes to enforce collective movements in refining macromolecular structural models, Proc. Natl. Acad. Sci. U. S. A. 101 (2004) 6957–6962.
[24] C. Magyar, M.M. Gromiha, G. Pujadas, G.E. Tusnády, I. Simon, SRide: a server for identifying stabilizing residues in proteins, Nucleic Acids Res. 33 (2005) W303–W305.
[25] S. Van Deventer, Cytokine and cytokine receptor polymorphisms in infectious disease, Intensive Care Med. 26 (2000) S98–S102.