Spring 4-24-2013

In Silico Analysis Shows That Single Aminoacid Variations In Rhesus Macacque Fcγreceptor Affect Protein Stability And Binding Affinity To IgG1

Rashesh Sanghvi

IN SILICO ANALYSIS SHOWS THAT SINGLE AMINOACID VARIATIONS IN RHESUS MACACQUE FCγRECEPTOR AFFECT PROTEIN STABILITY AND BINDING AFFINITY TO IGG1

by

RASHESH SANGHVI

Under the Direction of Dr.Roberta Attanasio

ABSTRACT

Rhesus macaques are a widely used animal model of human diseases and related immune responses. Fc receptors (FcRs) mediate the interaction between antibody molecules and innate killing mechanisms, consequently eliminating the pathogen. In rhesus macaques, FcRs are highly polymorphic. To evaluate the potential influence of FcγR polymorphisms on the interaction with antibody molecules, we performed *in silico* analysis using SIFT, Provean, nsSNPAnalyzer, I-Mutant, MuSTAB and iPTREE-STAB web servers. V20G in FcγRI, I137K in FcγRII and I233V in FcγRIII were further analyzed structurally using FOLD-X, AMMP and Chimera to calculate changes in folding and interaction energy and for structure visualization. Results from our analysis suggest that the selected variations destabilize protein structure. Additionally, Q32R increases the binding affinity of FcγRI, whereas A131T decreases the binding affinity of FcγRII towards IgG1. Together, our results indicate that these substitutions might influence effector and regulatory mechanisms resulting from antibody/FcR interactions.

INDEX WORDS: FcγR, Rhesus Macaque, Single nucleotide polymorphism, *in silico* analysis, AIDS

IN SILICO ANALYSIS SHOWS THAT SINGLE AMINOACID VARIATIONS IN RHESUS

MACACQUE FCγRECEPTOR AFFECT PROTEIN STABILITY AND BINDING AFFINITY TO

IGG1

by

RASHESH SANGHVI

A Thesis Submitted in Partial Fulfillment of the Requirements for the Degree of

Master of Science

in the College of Arts and Sciences

Georgia State University

2013

IN SILICO ANALYSIS SHOWS THAT SINGLE AMINOACID VARIATIONS IN RHESUS

MACACQUE FCγRECEPTOR AFFECT PROTEIN STABILITY AND BINDING AFFINITY TO

IGG1

by

RASHESH SANGHVI

Committee Chair: Dr. Roberta Attanasio

Committee: Dr. Robert Harrison

Dr. Franco Scinicariello

Electronic Version Approved:

Office of Graduate Studies

College of Arts and Sciences

Georgia State University

May 2013

ACKNOWLEDGEMENTS

I would like to express my deepest gratitude to my advisor, Dr. Roberta Attanasio, for her continuous support and guidance throughout my Masters program. She has encouraged and supported all my academic decisions and provided me with opportunity to generate an idea for my thesis project. I thank Dr. Robert Harrison for supporting me and guiding me in the field of Bioinformatics. I am grateful to Dr. Franco Scinicariello for being in my thesis committee and supporting me.

I would like to thank Palak Gupta and Manali Rupji for supporting me during the duration of my Masters program.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1   INTRODUCTION

1.1   Rhesus Macaque as Animal Models of Human Diseases

Animal models are an indispensible part for biomedical research. They are of paramount importance for studying pathogenesis, host immune responses to viral infections as well as immunogenicity and protective efficacy of vaccines. Nonhuman primates, specifically rhesus macaque models have been used in large number of studies on human infectious diseases such as yellow fever, degenerative neurologic disease like prion disease, childhood disease such as polio and tropical diseases like hepatitis E and bartonellosis as well as newly emergent, sexually transmitted disease caused by immunodeficiency virus. Macaque models are also used in study of oncogenic and bioterrorism associated diseases [1].   Macaques serve as models for studying passive immunization and various therapeutic approaches.

Rhesus macaques (*Macaca mulatta*), are very closely related to humans anatomically, physiologically and phylogenetically [1,2]. The patho-physiological responses of humans and nonhuman primates to internal and external insults are remarkably similar[3]. Nonhuman primates play an important role in understanding significant infectious diseases such as acquired immunodeficiency syndrome (AIDS), hepatitis and malaria, their treatment and prevention. They are also used for studying chronic neurological degenerative disorders like Parkinson's and Alzheimer's diseases [4]. Simian immunodeficiency virus-infected macaques show a decrease in CD4+T cells and also develop opportunistic diseases, malignancies and diseases of the central nervous system as would be seen in humans with human immunodeficiency virus (HIV) [5]. Thus, rhesus macaques provide an efficient model for studies related to diseases caused by single etiologies like HIV and influenza virus. Rhesus macaques play an important role as animal model in understanding the host innate and adaptive immune responses elicited

by HIV vaccines, and are considered the best animal model currently available for AIDS research [6,7].

1.2   Molecules of Adaptive Immunity

Neutralizing antibodies play an important role in protective responses to HIV/SIV infection. Antibodies (Abs), also known as immunoglobulin (Igs), interact with their specific cellular membrane bound Fc receptors (FcRs) to give rise to effector functions that ultimately destroy the pathogen. The Ab/FcR interaction is also involved in immune responses operating during autoimmune diseases and cancer and can be manipulated for therapeutic applications.

FcRs are a family of membrane-bound glycoproteins mainly expressed on the surface of phagocytic effector cells like macrophages and neutrophils. They are also expressed on natural killer (NK) cells. FcRs interacts with the Fc region of the antibody, when the antibody binds the antigen forming immunocomplexes, setting into motion the cellular responses.

FcRs belong to the immunoglobulin super family. Chromosome mapping and cloning experiments indicate that FcR genes are the result of gene duplication of a common ancestor. FcγR interact with the Fc region of the IgG family of antibodies. The three basic classes of human FcγR are: FcγRI(CD64), FcγRII(CD32) and FcγRIII(CD16)[8]. All the FcγRs show overall similar intron-exon assembly consisting of a leader region, 2 c-like domains for FcγRII and FcγRIII and 3 for FcγRI, and the connecting transmembrane-cytoplasmic region[9]. Figure 1.1 shows the different types of human FcRs, their structure, the cell types expressing the various FcRs, and the binding affinity to the different antibody classes. The association of FcRs with different intracellular chains is cell type-dependent [10].
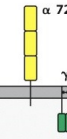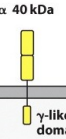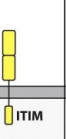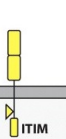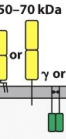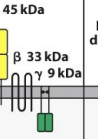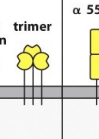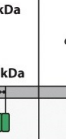
Receptor	FcγRI (CD64)	FcγRII-A (CD32)	FcγRII-B2 (CD32)	FcγRII-B1 (CD32)	FcγRIII (CD16)	FcεRI	FcεRII (CD23)	FcαRI (CD89)	Fcα/μR
Structure	α 72 kDa / γ / γ-like domain	α 40 kDa / γ-like domain	/ ITIM	/ ITIM	α 50–70 kDa / or / γ or ζ	α 45 kDa / β 33 kDa / γ 9 kDa	lectin trimer domain / N	α 55–75 kDa / γ 9 kDa	α 70 kDa
Binding / Order of affinity	IgG1 / $10^8\,M^{-1}$ / 1) IgG1=IgG3 / 2) IgG4 / 3) IgG2	IgG1 / $2 \times 10^6\,M^{-1}$ / 1) IgG1 / 2) IgG3=IgG2* / 3) IgG4	IgG1 / $2 \times 10^6\,M^{-1}$ / 1) IgG1=IgG3 / 2) IgG4 / 3) IgG2	IgG1 / $2 \times 10^6\,M^{-1}$ / 1) IgG1=IgG3 / 2) IgG4 / 3) IgG2	IgG1 / $5 \times 10^5\,M^{-1}$ / IgG1=IgG3	IgE / $10^{10}\,M^{-1}$	IgE / 2–$7 \times 10^7\,M^{-1}$ (trimer) / $2 \times 10^6\,M^{-1}$ (monomer)	IgA1, IgA2 / $10^7\,M^{-1}$ / IgA1=IgA2	IgA, IgM / $3 \times 10^9\,M^{-1}$ / 1) IgM / 2) IgA
Cell type	Macrophages Neutrophils Eosinophils Dendritic cells	Macrophages Neutrophils Eosinophils Platelets Langerhans cells	Macrophages Neutrophils Eosinophils	B cells Mast cells	NK cells Eosinophils Macrophages Neutrophils Mast cells	Mast cells Basophils	Eosinophils B cells	Macrophages Eosinophils† Neutrophils	Macrophages B cells
Effect of ligation	Uptake Stimulation Activation of respiratory burst Induction of killing	Uptake Granule release (eosinophils)	Uptake Inhibition of stimulation	No uptake Inhibition of stimulation	Induction of killing (NK cells)	Secretion of granules	Degranulation	Uptake Induction of killing	Uptake

Figure 1.1 Distinct receptors for the Fc region of the different immunoglobulin classes are expressed on different accessory cells. From: Janeway's Immunobiology, 8th edition (©Garland Science) [10]

In humans, there are 8 genes coding for FcγRs which are present as a gene cluster in chromosome 1. Their gene products show highly diversified binding affinity towards the IgG group. The diversity of FcγRs is further increased due to the presence of various single nucleotide polymorphisms in the extracellular c-like domains. FcγR-related mechanisms are also involved in autoimmune diseases and inflammation [11]. Therefore, it is important to characterize FcR polymorphisms and study the effects of these polymorphisms on the interaction of FcγR with Igs.

1.3 Interaction of Fc Receptors and Antibody

Antibody Fc-dependent mechanisms include antibody-dependent cell-mediated cytotoxicity (ADCC), FcR- dependent increase in uptake of antibody-antigen complex by antigen presenting cells [12,13], FcR- mediated enhancement of MHC class I-restricted cross-presentation [14] and modulation of immune responses by differential engagement of antibodies with activating and inhibitory Fc receptors [15]. FcγR-mediated immune responses result in removal of pathogens and virus-infected or cancer cells. They are also involved in hypersensitivity reactions [11,16,17]. FcγRs are present on various cells like natural killer cells, dendetric cells, mast

cells, monocytes, B-cells, eosinophils and neutrophils. The Fc portions of IgG-sensitized anti-gens are recognized by the FcγR present on the membranes of the cells, resulting in cell activation through the immune-receptor tyrosin-based activation/inhibitory motif. The cell response depends on the type of FcR, antibody isotype and cell type. The interaction of the FcγRIIa with immune complex (IC) leads to cell activation of neutrophils, which results in phagocytosis of the antigen, degranulation of the neutrophil and oxidative burst. The degranulation process in turn results in local inflammation. NK cells destroy the antibody coated cells by recognizing them using the NK-membrane bound FcγRIIIa (CD16). IC interaction with FcR on the B-cells leads to maturation, thus resulting in differentiation to plasma cells and memory B-cells. The plasma cells secrete antibodies specific for the pathogen. Differential engagement with either activating or inhibitory FcR leads to either MHC class I restricted or MHC class II restricted antigen presentation. For all the immune responses mediated by FcR, the interaction of the FcR with the Fc region of the Igs is essential.

Different classes of FcγRs show differential binding affinity towards the IgG subclasses. In order to understand this behavior, it is essential to investigate the 3D structure of the FcγR-IgG complex and determine the amino acids involved in the interaction. The FcγRI-IgG complex crystal structure is yet to be solved and hence exact interacting amino acids cannot be determined. Ile88, Trp90, Trp113, His134, Val158, Gly159 and Lys 161 at the interface of the FcγRIII and IgG1 are involved in hydrophobic interactions. Thr116, Lys120, His134, His 135, Val158 and Lys161 are involved in hydrogen bonds and salt bridges [18]. Similarly Trp90, Trp113, Pro117, Val119, Lys120, Lys128, Ser129, Phe132, Arg134, Leu135 and Tyr160 are involved in interaction of FcγRII with the IgG Fc ligand [19].

1.4   *In Silico* Analysis of Single Nucleotide Polymorphism

Single Nucleotide Polymorphisms (SNP) is commonly found across the genome. The non-synonymous (ns) mutations that result into amino acid residue changes, play an important

role in diversifying protein function [20]. nsSNPs result in amino acid variations that may cause changes in protein stability. The physiochemical properties of the mutant residues differ from the wild type, resulting in this change. The amino acid change may also disrupt salt bridges or hydrogen bonding affecting protein dynamics and thus stability. It may also disrupt the binding site and affect protein interaction by altering the specificity of the protein, blocking the active site or affecting the binding affinity [21]. Identifying possibly deleterious nsSNPs or those responsible for a specific phenotype is of a major concern and requires testing thousands of SNPs in the respective genes. Because it is not feasible to analyze all SNPs experimentally, there is obvious need for *in silico* analysis.

Computational methods are needed to analyze the effects of such large number of nsSNPs. Several *in silico* analysis tools have been developed based on machine learning to predict protein stability changes due to single amino acid substitutions. These tools identify the deleterious nsSNPs from the neutral ones based on sequence and structure approaches [22] and take into consideration the physicochemical properties like acidic, basic, or hydrophobic, the conservation of the amino acid at a specific position and protein structure. Some of these tools predict just the direction of the stability change which, in certain biological applications, may be more relevant than magnitude.

The 3D structure has not been analyzed and deposited in the Protein Data Bank (PDB) for all proteins, thus a sequence-based approach is more appealing than structural information based methods [23]. In this study, different sequence based algorithms like Sorting Intolerant from Tolerant (SIFT) and PROVEAN, I-Mutant 2. 0, nsSNPAnalyzer, MuSTAB and iPTREES-TAB were used for the prioritization of high-risk nsSNPs in FcomprisecγRs.

Since structural information for most proteins is not available, it is necessary to predict the 3D structure of the protein by comparative modeling. The resulting models can be used to predict the effects of the nsSNPs on protein stability and protein-protein interaction by mapping the mutations on the predicted structures. The mutated structures are predicted using homol-

ogy modeling and then are compared to the wild type structure to estimate the effects of the single amino acid substitution at the structural level.

Another method of analyzing the effect of the nsSNPs is by estimating changes in the protein stability. This can be investigated by measuring the difference of Gibbs free energy ($\Delta\Delta G$) between the wild type and mutated state of the protein ($\Delta\Delta G = \Delta G_{wild} - \Delta G_{mutated}$) where ΔG = Gibbs protein unfolding energy. The discrimination among deleterious or functionally effective nsSNPs from the neutral ones is based on the notion that protein stability perturbation should be above a certain threshold $\Delta\Delta G$ (± 1 kcal/mol) [24] for it to be functionally important.

1.5   Purpose of the Study

Various Studies have shown correlation between the polymorphisms and the function of the FcγRs. The polymorphism at position 158 of human FcγRIII, which codes for valine (V) or a phenylalanine (F) results in low or high susceptibility for antilymphocyte globulins (AGL), respectively[25]. **FcγRIIA** R131 polymorphism has less affinity towards IgG2 compared to H131 [26,27]. S48N polymorphism also affects the affinity towards IgG binding and causes recurrent infections [28]. V158/F158 polymorphism in **FcγRIIIA** causes difference in affinity for IgG1 and IgG3 and differential glycoslysation of **FcγRIIIB** causes difference in affinity for IgG1[29]. These allelic variants have shown to affect the susceptibility to various diseases like systemic lupus erythematosus [30] and inflammation following an organ transplant indicating the importance of these SNPs on the functional domain of FcγR. An I187T substitution in the transmembrane region of FcγRIIB affects the inhibitory function on B cells[27]. A study on clearance rates of RhD(Rhesus D)-positive RBCs showed faster clearance rates for FcγRIIA- H131 or FcγRIIA-V158 homozygous subjects [31]. A recent study on AIDS vaccine efficacy in rhesus macaques showed TLR7 variant M5 allele very likely destabilizes an RNA secondary structure predicted for the V5 allele and also affects the survival time post infection in SIV-infected rhesus macaques [32].

The research pertaining to AIDS is done with the assumption that the macaque FcγR and their interaction with the Fc receptors of the Igs represent that of humans. The sequence homology of the macaque FcγR I, FcγR II and FcγR III is 94-95%, 88-90% and 91.7% with those of humans. In addition, 11 distinct allelic sequences were identified indicating high polymorphism in macaque, which contained total 15 nsSNPs [33].   Thus it is required to study the impact of the polymorphisms in the macaque FcγR on their interaction with IgG subclasses.

1.6   Results

Our *in silico* analysis suggests that the presence of mutations (V20G in FcγRI, I137K in FcγRII and I233V in FcγRIII) can affect the structure and thus the function of the FcγR in the form of protein stability.   In addition, A131T in FcγRII affects the protein-protein interaction with IgG1.

2   EXPERIMENT

Based on the various *in silico* analyses carried out to predict the effects of single amino acid variations and to increase prediction accuracy, we combined prediction obtained from various *in silico* methods. We used *in silico* methods based on sequence and structure of the protein. The schematic diagram of the approach utilized is shown in Figure 2. 1.
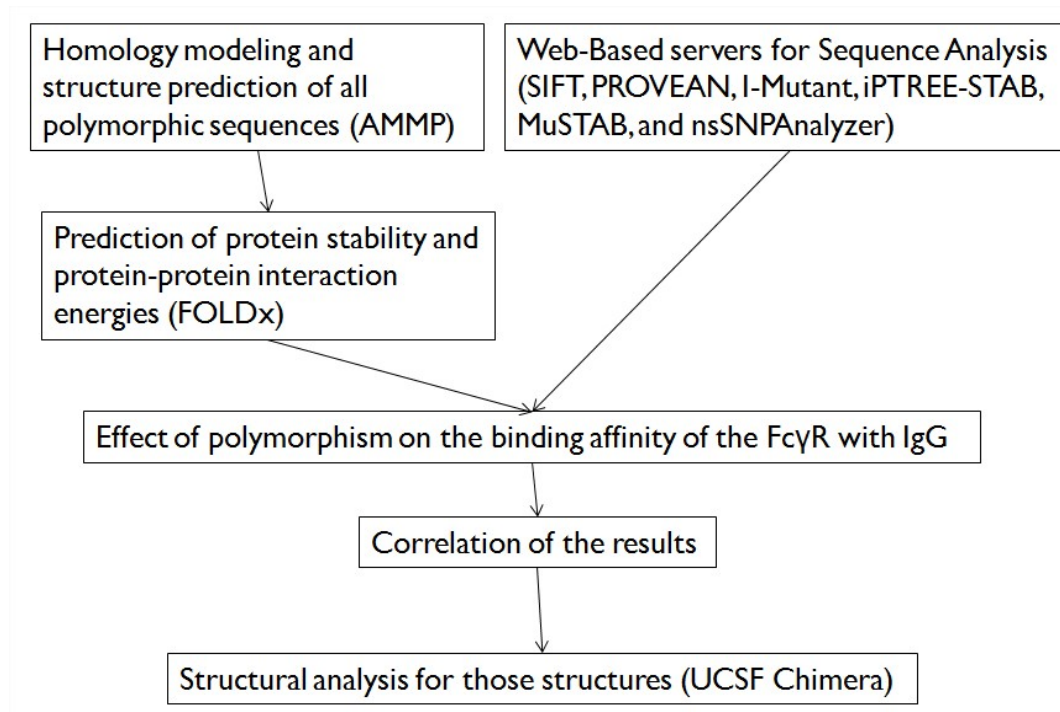
Figure 2.1   Schematic diagram of the approach used to study the effect of single amino acid variations on protein stability and binding affinity. The programs and software used are indicated in the boxes.

2.1   Data

2.1.1   Sequence

The protein sequences of the Indian rhesus macaque FcγRs and IgGs were obtained from the Protein database at National Centre for Biological Information (http://www. ncbi. nlm. nih. gov/protein). Three allelic sequences of FcγRI (CD64), five allelic sequences of FcγRII (CD32) and three allelic sequences of FcγRIII (CD16) having GenBank accession numbers: AEC03710.1- AEC03712.1, AEC03702.1- AEC03706.1

and AEC03707.1- AEC03709.1 respectively, were analyzed in this study [33]. In these sequences, 16 polymorphic sites were analyzed, three in FcγRI, nine in FcγRII and four in FcγRIII. The protein sequence of Indian rhesus macaque IgG1 sequenced in a previous study having GenBank accession numbers: AAQ57554. 1 [34] was used to form the complex with FcγRs. The 3D protein structures, 1T83 and 3RY6, were obtained from Protein Data Bank (PDB).

2.1.2 Homology Modeling

All sequences were used for sequence based homology modeling. They were submitted to an online server at http://bmcc3. cs. gsu. edu/, which uses the molecular mechanics program AMMP to predict the 3D structure of a protein based on sequence. The modeler uses two techniques for optimization of the predicted structure. The first is the genetic algorithm that builds a set of 20 structures varying from the start point and is optimized by conjugate gradient method. The second is the four-dimensional embedding. This method is useful in solving the three-dimensional problem in four-dimensions and then forcing the solution back into three dimensions.

3D structures of the FcγRI and FcγRIII predicted by AMMP structure prediction server were superimposed on the chain C of the reference Protein Data Bank structure (PDB) 1T83 and the predicted structures of FcγRII on the chain C of the reference Protein Data Bank structure (PDB) 3RY6. The predicted structure of Indian macaque IgG1 was superimposed on the IgG chains A and B of both the reference structures.

2.2 Web-Based Servers for Sequence Analysis

Many web servers are available to predict the effect of single amino acid variations on protein stability and protein binding efficiency. SIFT, I-Mutant2.0, I-Mutant3.0, MuStab, iPTree-stab and snpAnalyser were used in this study. The first sequence in all the FcγRs was as-

sumed to be the base sequence. Then the observed polymorphisms in other sequences were analyzed.

Sorting Intolerant From Tolerant (SIFT) predicts intolerant mutations using homology and requires only the sequence to obtain position specific information from alignments. SIFT scores the mutation at a specific position based on a position-specific scoring matrix (PSSM). SIFT also takes into consideration the physiochemical properties of each amino acid and predicts if the amino acid change is tolerant or intolerant, i.e whether the change severely affects the protein function. The output from SIFT includes a score from 0-1, where <=0. 05 is the threshold for tolerance [35].

Protein Variation Effect Analyzer (PROVEAN), a sequence based predictor, estimates the effect of protein sequence variation on protein function. It is based on a clustering method where BLAST hits with more than 75% global sequence identity are clustered together and top 30 such clusters form a supporting sequence set. A delta alignment scoring system is used, where the scores of each supporting sequence are averaged within and across clusters to generate the final PROVEAN score. A protein variant is said to be "deleterious" if the final score is below a certain threshold (default is -2. 5), or is predicted to be "neutral" if the score is above the threshold [36].

I-Mutant2.0 and I-Mutant3.0 is based on Support Vector Machine algorithm to predict the stability of the protein on single amino acid variations. It can predict protein stability changes by using protein sequence or structure. It has an overall accuracy of 77% when prediction is based on protein sequence. I-Mutant2.0 and I-Mutant3.0 predicts the DDG values as a regression estimator and also the sign of the stability change. I-Mutant3.0 furthermore classifies mutations into three categories: neutral mutation (-0.5<=DDG<=0.5), large decrease (<=-0.5) and large increase (>0.5).

MuSTAB is another SVM based web server that uses various biochemical features, structural features and various biological features to predict the effect of the amino acid change

on the protein stability and function. It uses amino acid composition, conformational parameters for alpha helix, beta sheet, and average buried area in folded state, polarity and number of codons for amino acid as a classifier to predict protein stability changes due to single amino acid variations. It has 84.59% accuracy, 70.29% sensitivity and 90.98% specificity [23].

IPTREE-STAB web server is based on an adaptive boosting algorithm and utilizes decision trees for accessing the protein stability and also predicts the DDG value. It provides a binary classification as stabilizing or destabilizing depending on the DDG value. While discriminating proteins due to single amino acid change, it also considers three residues adjacent to the polymorphism site in N and C terminal. iPTREESTAB achieves an overall accuracy of 82.1%, 75.3% sensitivity and 84.5% specificity [23,37] .

nsSNPAnalyzer is a random forest algorithm based tool. nsSNPAnalyzer uses the multiple sequence alignment information along with the structural information like solvent accessibility, secondary structure and polarity to distinguish between neutral and disease mutation. It also considers the similarity and dissimilarity between the wild type and the mutant residue [38].

2.3   Structural Analysis

The predicted structures were viewed in University of California, San Francisco (UCSF) chimera. It is a computationally intensive program for visualization of molecular models and it provides an interactive interface for the user for analyzing the models and model related data. It provides a platform for analyzing sequence alignments, generating homology models, molecular docking, viewing various density models, and also comparing different models by super imposition [39].

The mutant and wild type structures were superimposed and the effect of the non synonymous variation was observed in terms of steric hindrance due to the changes of the side chains and change in the charge of the amino acid. Then, the degree of change in the hydro-

phobicity or hydrophyllicity of the mutated amino acid and its effect on the interacting intra chain and inter chain molecules was analyzed.

2.4   FOLDX

FOLDx is a computer algorithm that computes the protein stability and protein-protein interaction values. It compares between the wild type and the mutant protein structures based on various energy interactions like van der Waals clashes, electrostatic interactions and hydrogen bonding. The FOLDx energy function includes the terms that have been determined as crucial by protein stability experiments, and the unfolding energy is calculated using the following equation:

$$DG = DG_{vdw} + DG_{solvH} + DG_{solvP} + DG_{wb} + DG_{Hbond} + DG_{el} + DG_{kon} + TDS_{mc} + TDS_{sc} + TDS_{tr}$$

where ΔG_{vdw} depicts the total van derWaals contributions of all atoms compared to that with solvenyt; ΔG_{solvH} and ΔG_{solvP} give the changes in the folded and unfolded state salvation energy; ΔG_{wb}, gives the stabilizing free energy generated due to formation of more than one hydrogen bond between water molecules and protein; ΔG_{Hbond} is the free energy difference between the formation of an intramolecular hydrogen bond compared to intermolecular hydrogen bond formation (with solvent); ΔG_{el} gives the energy by electrostatic bonds formed by charged groups; ΔG_{kon} reflects the effect of electrostatic interactions on the kon. ΔS_{mc} depicts the changes in entropy due to the back bone of the protein in the folded state; ΔS_{sc} depicts the entropic cost of side chain optimization in a particular protein conformation [40,41]

2.4.1   Optimizing Models

The modeled structures were subjected to an optimization procedure using the repairPDB function of FoldX. This corrects the poor torsion angles and van der Waals clashes if found in any residue. It checks if any other rotamer exists for the same residue that is more stable and avoids steric clashes in a specific position.

2.4.2   Energy Calculations

The difference in the energy of the folded and unfolded protein gives an estimate of the Gibbs free energy of folding. The effect of the single amino acid variations on the stability of the protein is analyzed by computing the difference between the Gibbs free energy of folding of the mutant structure and the wild type

$$DDG = DG_{Mutant} - DG_{WT}$$

where DG is the Gibbs free energy of folding.

FoldX calculates the Gibbs free energies of the two molecules A (DG_A) and B (DG_B) as well as the complex AB (DG_{AB}). The interaction energy can be estimated from the difference between these energies using the formula listed below:

$$DG_{binding} = DG_{AB} - (DG_A + DG_B)$$

The difference in the binding energies ($DG_{binding}$) of the mutant and wild type structures esti-mates the -effect of the single amino acid variation on the stability of the protein-protein interac-tion.

3   RESULTS

A total of 3 nsSNP from 3 allelic FcγRI sequences, 9 nsSNP from 5 allelic FcγRII sequences and 4 nsSNP from 3 allelic FcγRIII sequences obtained from the 9 Indian macaques were iden-tified.

3.1   Proximity of SNPs to IgG Binding Site

Figure 3.1 Alignment of rhesus macaque FcγRI allelic sequences (GenBank accession numbers: HQ423394-HQ423396) with a human FcγRIa (GenBank accession number BC020823.1).
The allelic sequences are labeled based on the animal number used to obtain the respective sequence [33]. The horizontal blue lines represent the start of the D1, D2 and D3 domains. TM-CO represents the trans-membrane and cytoplasmic region. The vertical red arrows indicate the SNP positions in the allelic se-quences of rhesus macaque CD64. The red boxes indicate the amino acids involved in binding with IgG1 as per FOLDx prediction.

FcγRI has three Ig like domains and all the three SNPs are present in the N-terminal domain (D1). Since crystal structure of human FcγRI-IgG complex has not been solved yet, no information regarding the amino acids involved in the interaction with IgG is currently available.

However, on the basis of the information related to other human FcR-IgG complex structures, it is reasonable to assume that most of the interacting amino acids lie within the second domain (D2).

Figure 3.2   Interaction of the amino acids at the FcγRI and Fc region of IgG1. Hydrogen bonds formed by the amino acids predicted to be involved in FcγRI binding to IgG1 are indicated. FcγRI is in green, chain A of IgG1 is in blue and chain B of IgG1 is in red.

Using the AnalyzeComplex method of FOLDx, we determined the amino acids present at the FcγRI-IgG1 interface. Another command of FOLDx, PrintNetworks, independently predicted the amino acids involved in intramolecular hydrogen bond formation and electrostatic interaction. Arg99, Lys127, Tyr130, Tyr145, His164 and Tyr173 were common amino acids predicted from both commands of FOLDx and were predicted to be involved in hydrogen bond formation with IgG1. Arg99, Asp100, Lys125, Lys127, Tyr130, Phe143, Phe144, Tyr145, Arg146, Lys170 and Arg172 were predicted to be involved in electrostatic interactions with IgG1. Since none of the SNPs are present in the FcR-IgG binding site, they would not affect the binding af-

finity of the rhesus FcγRI to IgG1. However, these SNPs have potential to affect protein stability

and may result in change of function or loss of function. Figure 3.2 indicates the hydrogen

bonds formed by amino acids predicted to be involved in FcγRI interaction with IgG1 by FOLDx.

The Gln at position 32 forms a hydrogen bond with His at position 99.
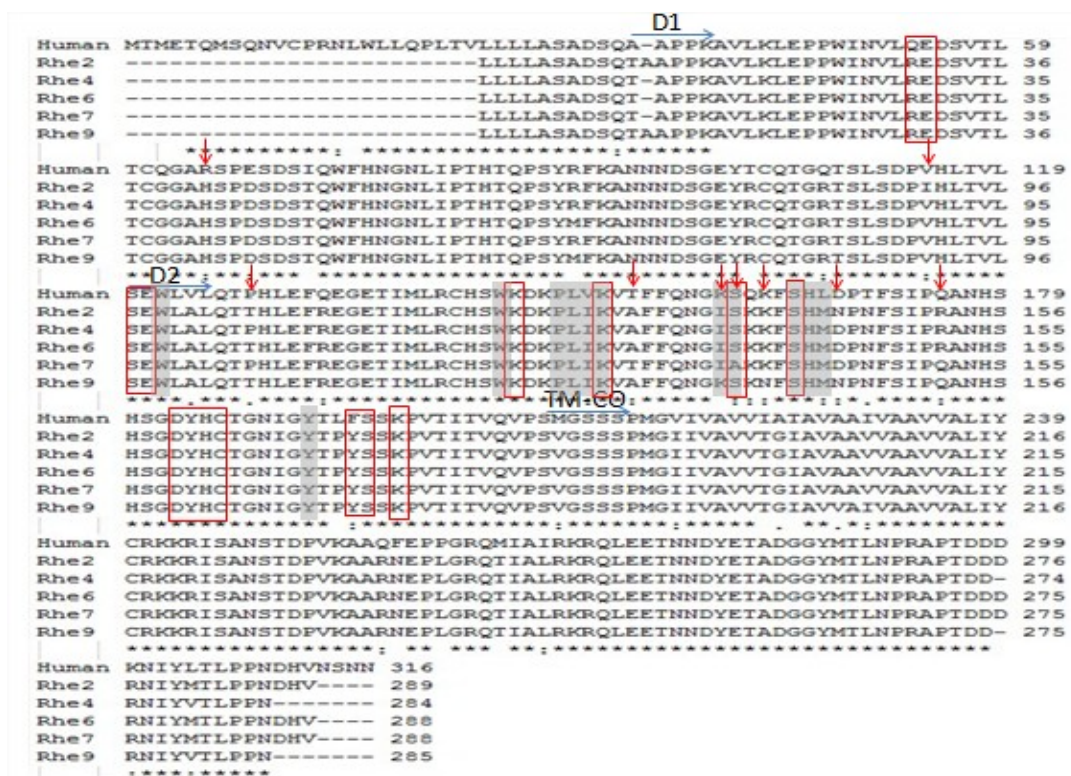
Figure 3.3 Sequence alignment of rhesus macaque FcγRII allelic sequences (GenBank accession numbers: HQ423389-HQ423393) with a human FcγRIIa (GenBank accession number AAH20823.1). The allelic sequences are labeled based on the animal number used to obtain the respective sequence [33]. The horizontal blue lines represent the start of the D1 and D2 domains.   TM-CO represents the transmembrane and cytoplasmic region.   The vertical red arrows indicate the SNP positions in the allelic sequences of rhesus macaque CD32.   The red boxes indicate the amino acids involved in binding with IgG1 as per FOLDx prediction.   The grey shaded regions represent the amino acids involved in human FcγRII-IgG1 binding.

FcγRII has two Ig like domains and two SNPs are present in the N-terminal domain (D1).

The remaining 7 SNPs are present in the C-terminal domain (D2). The amino acids involved in

human FcγRII-IgG1 interaction are Trp90, Trp113, Pro117, Val119, Lys120, Lys128, Ser129,

Phe132, Arg134, Leu135 and Tyr160 [19]. These are highlighted in Figure3.2 in grey shade.

Using the AnalyzeComplex method of FOLDx, we determined the amino acids present

at the rhesus FcγRII-IgG1 interface. Another command of FOLDx, PrintNetworks, independ-

ently predicted the amino acids involved in intramolecular hydrogen bond formation and electro-

static interaction. Arg30, Lys125, Lys129, Ser138, Ser142, His162, Cyc163 and Ser173 were

common amino acids predicted from both commands of FOLDx and were predicted to be in-

volved in hydrogen bond formation with IgG1. Arg30, Glu31, Glu98, Lys123, Lys125, Lys129,

Asp160, Tyr161, Tyr172 and Lys175 were predicted to be involved in electrostatic interactions

with IgG1.
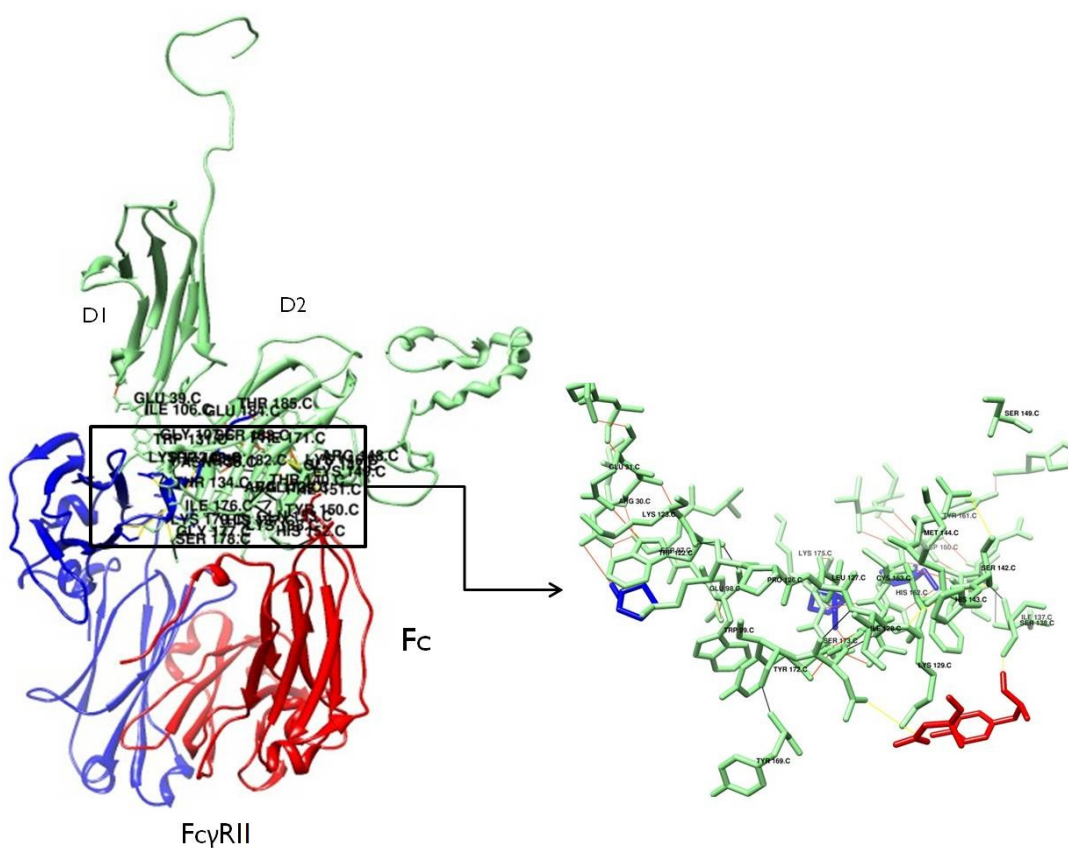
Figure 3.4 Interaction of the amino acids at the FcγRII and Fc region of IgG1. Hydrogen bonds formed by the amino acids predicted to be involved in FcγRII binding to IgG1 are indicated. FcγRII is in green, chain A of IgG1 is in blue and chain B of IgG1 is in red.

Figure 3.4 indicates the hydrogen bonds formed by the amino acids predicted from the

human FcγRII-IgG1 crystal structure analysis [19] and by FOLDx analysis, to be involved in

FcγRII interaction with IgG1. The hydrogen bonds formed by residues predicted by crystal

structure analysis are indicated in black. The ones formed by the residues predicted by FOLdx

analysis are indicated in orange and those formed by residues predicted by both the techniques are indicated in yellow.

FcγRIII has two Ig like domains and out of the four SNPS, one is present in the N-terminal domain (D1), another one is present in the C-terminal domain (D2) and the other two are within the cytoplasmic resion of FcγRIII. The amino acids involved in human FcγRIII-IgG1 interaction are Ile88, Trp90, Trp113, Thr116, Lys120, His134, His 135, Val158, Gly159 and Lys 161 [18], These are highlighted in Figure 3.3 in grey shade.

Figure 3.5 Sequence alignment of rhesus macaque FcγRIII allelic sequences (GenBank accession numbers: HQ423386-HQ423388) with a human FcγRIIIa (GenBank accession number CAA34753. 1). The allelic sequences are labeled based on the animal number used to obtain the respective sequence [33]. The horizontal blue lines represent the start of the D1 and D2 domains. TM-CO represents the transmembrane and cytoplasmic region.   The vertical red arrows indicate the SNP positions in the allelic sequences of rhesus macaque CD16. The red boxes indicate the amino acids involved in binding with IgG1 as per FOLDx prediction. The grey shaded regions represent the amino acids involved in human FcγRIII-IgG1 binding.

Using the AnalyzeComplex method of FOLDx, we determined the amino acids present at the rhesus FcγRIII-IgG1 interface. Another command of FOLDx, PrintNetworks, independently predicted the amino acids involved in intramolecular hydrogen bond formation and electro-

static interaction. Thr134, Lys138, Lys146, Arg148, Tyr150, Phe171, Ser178, Lys179, Ser182,

Ser183 and Thr185 were common amino acids predicted from both commands of FOLDx and

were predicted to be involved in hydrogen bond formation with IgG1. Glu39, Lys132, Lys138,

Lys146, Arg148, Lys149, Tyr150, Phe151, Arg173, Lys179 and Glu184 were predicted to be

involved in electrostatic interactions with IgG1.
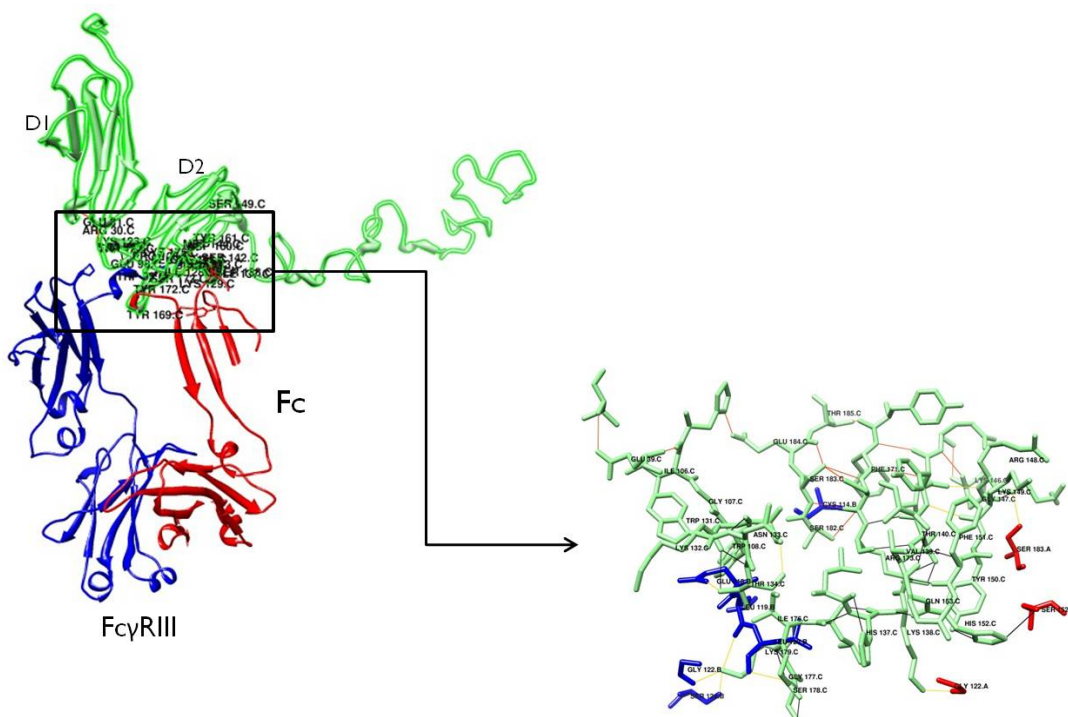
Figure 3.6 Interaction of the amino acids at the FcγRIII and Fc region of IgG1.   Hydrogen bonds formed by the amino acids predicted to be involved in FcγRIII binding to IgG1 are indicated.   FcγRIII is in green, chain A of IgG1 is in blue and chain B of IgG1 is in red.

Figure 3.6 indicates the hydrogen bonds formed by the amino acids predicted from the

human FcγRIII-IgG1 crystal structure analysis [18] and by FOLDx analysis, to be involved in

FcγRIII interaction with IgG1. The hydrogen bonds formed by residues predicted by crystal

structure analysis are indicated in black. The ones formed by the residues predicted by FOLdx

analysis are indicated in orange and those formed by residues predicted by both the techniques

are indicated in yellow.
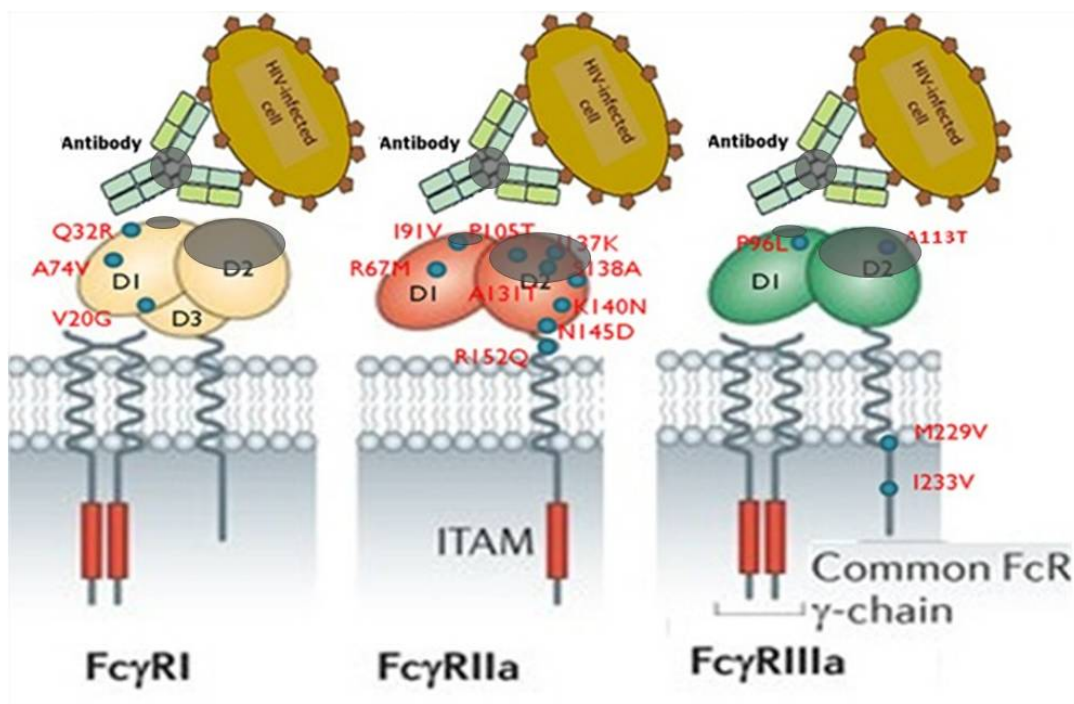
Figure 3.7 Position of Single amino acid variations on FcγRs with respect to the interaction site of the FcγRs with the Fc region of the antibodies. The grey shaded region indicates the binding sites to the antibody Fc region.

3.2   Sequence Based Predictions

SIFT, a sequence based predictor, estimates the effect of amino acid variations on protein function based on homology of the sequence and the physio-chemical properties of the amino acid residues. It also takes into account the conservation of the sequence through evolution. SIFT scores were classified as damaging (0.00–0.05) or tolerant (0.051–1.00). Of the 15 nsSNPs, V20G and V74A in FCγRI, R67M, P105T, K140N in FCγRII and V233I in FCγRIII showed the tolerance score of 0.00, 0.03, 0.04, 0.01, 0.02 and 0.03 respectively [35].

PROVEAN predicts the effect of the variant on the biological function of the protein based on sequence homology. PROVEAN scores are classified as "deleterious" if below a certain threshold (here -2.5) and "neutral" if above it. Of the 15 nsSNPs, V20G in FCγRI, P105T, T131A, K137I and K140N in FCγRII and L96P in FCγRIII were predicted "deleterious" with PROVEAN scores of -4.951, -4.714, -2.88,   -4.084, -2.783 and -5.504 respectively.

I-Mutant 3.0 is a SVM based tool used to predict the DDG stability values and the reliability index for the DDG value. It also classifies the prediction as largely destabilizing (DDG<-0.5 Kcal/mol) or largely stabilizing (DDG>0.5 Kcal/mol) or having a weak effect (-0.5<=DDG<=0.5 Kcal/mol). For this study we have considered a threshold of -1.5 Kcal/mol to consider a SNP to be destabilizing. Of the 15 nsSNPs, V20G and V74A in FCγRI and I137K in FCγRII were considered destabilizing with DDG values -2.19, -1.65 and -1.85 Kcal/mol respectively.

nsSNPAnalyzer predicts whether an nsSNP has a phenotypic effect. It considers the structural environment, conservation of the residue at that position by sequence homology and the difference in the physical properties of the mutant compared to the wild type amino acid. nsSNPAnalyzer classifies mutations as "disease" or "neutral" depending on its internal score. Of the 15 nsSNPs, V20G in FCγRI and I233V in FCγRIII were considered to be "disease" causing.

MuSTAB predicts the effect of the nsSNP on protein stability depending on various biochemical, empirical and biological features. Along with predicting whether the nsSNP increases or decreases the stability of the protein, it also gives the prediction confidence. For this study we have considered nsSNP to be destabilizing if the prediction confidence is above 95%. V20G in FCγRI, I137K in FCγRII and I233V in FCγRIII were predicted to decrease the protein stability considerably.

iPTREE-STAB is a decision tree-based, predictive tool that not only discriminates the nsSNP on the basis of increase or decrease in thermal stability but also predicts the DDG values. The nsSNP having DDG values > -2.5 kcal/mol were considered to be truly destabilizing. V20G in FCγRI and I137K in FCγRII were predicted to decrease the thermal stability of the protein.

The nsSNPs predicted to have a phenotypic effect according to SIFT or PROVEAN and destabilize the protein according to at least two of the tools from I-Mutant3. 0, nsSNPAnalyzer,

MuSTAB and iPTREE-STAB were used for further analysis. V20G in FCγRI, I137K in FCγRII and I233V in FCγRIII were further analyzed structurally to understand their effect.

Table 3.1 Sequence based analysis by SIFT

	Amino Acid Substitution	Straight Mutations		Reverse Mutations	
		Score	Prediction	Score	Prediction
FcγRI	G20V	1	TOLERATED	0	AFFECT PROTEIN FUNCTION
FcγRI	Q32R	0.06	TOLERATED	1	TOLERATED
FcγRI	A74V	1	TOLERATED	0.03	AFFECT PROTEIN FUNCTION
FcγRII	R67M	0.04	AFFECT PROTEIN FUNCTION	0.17	TOLERATED
FcγRII	I91V	1	TOLERATED	0.18	TOLERATED
FcγRII	T105P	1	TOLERATED	0.01	AFFECT PROTEIN FUNCTION
FcγRII	A131T	1	TOLERATED	0.21	TOLERATED
FcγRII	I137K	1	TOLERATED	0.07	TOLERATED
FcγRII	S138A	0.48	TOLERATED	1	TOLERATED
FcγRII	K140N	0.02	AFFECT PROTEIN FUNCTION	1	TOLERATED
FcγRII	N145D	0.9	TOLERATED	0.94	TOLERATED
FcγRII	R152Q	0.67	TOLERATED	0.27	TOLERATED
FcγRIII	P96L	1	TOLERATED	0.09	TOLERATED
FcγRIII	A113T	1	TOLERATED	0.47	TOLERATED
FcγRIII	M229V	1	TOLERATED	0.06	TOLERATED
FcγRIII	I233V	0.09	TOLERATED	0.03	AFFECT PROTEIN FUNCTION

Table 3.2 Sequence based analysis by I-Mutant for forward mutations.

	Straight Mutations	IMutant2.0			IMutant3.0 (Structure Based)			
	Amino Acid Substitution (WT Position MT)	DDG (Kcal/mol)	Stability	RI	DDG (Kcal/mol)	SVM 2 prediction	RI	RSA
FcγRI	G20V	-0.49	Decrease	2	-0.57	Decrease	8	34.3
FcγRI	Q32R	-0.84	Decrease	4	-0.35	Decrease	3	21.3
FcγRI	A74V	-0.35	Decrease	5	-0.11	Decrease	3	50.8
FcγRII	R67M	-1.58	Decrease	7	-0.15	Decrease	4	74.7
FcγRII	I91V	-0.02	Decrease	4	-0.33	Decrease	3	16.2
FcγRII	T105P	-0.42	Decrease	0	-1.06	Decrease	5	15.9
FcγRII	A131T	-0.59	Decrease	4	-1.02	Decrease	4	11.8
FcγRII	**I137K**	**-2.26**	**Decrease**	9	-1.05	Decrease	9	61.6
FcγRII	S138A	-0.73	Decrease	5	-0.14	Decrease	7	70

	Amino Acid	DDG	Stability	RI	DDG	Stability	RI	RSA
	K140N	0.07	Decrease	2	-0.6	Decrease	1	55.4
	N145D	-0.51	Decrease	3	-0.39	Increase	0	55.3
	R152Q	-1.7	Decrease	8	-0.62	Decrease	7	79.5
FcγRIII	P96L	-0.78	Decrease	6	-1.02	Decrease	8	26.1
	A113T	-0.88	Decrease	6	-1.23	Decrease	9	14.5
	M229V	-0.62	Decrease	7	-0.8	Decrease	7	40
	I233V	**-1.28**	**Decrease**	9	-0.77	Decrease	7	29.2

Table 3.3 Sequence based analysis by I-Mutant for reverse mutations.

	Reverse Mutations	IMutant2.0			IMutant3.0 (Structure Based)			
	Amino Acid Substitution (WT Position MT)	DDG (Kcal/mol)	Stability	RI	DDG (Kcal/mol)	SVM 2 Stability prediction	RI	RSA
FcγRI	**V20G**	**-3.95**	**Decrease**	**9**	-2.54	Decrease	9	44.2
	R32Q	-0.93	Decrease	8	-1.27	Decrease	9	38.4
	V74A	**-1.79**	**Decrease**	**9**	-1.53	Decrease	9	42.9
FcγRII	M67R	-0.88	Decrease	7	-0.34	Decrease	4	58.5
	V91I	-0.91	Decrease	4	-0.28	Decrease	4	21.5
	P105T	**-1.4**	**Decrease**	**7**	-1.11	Decrease	9	7
	T131A	**-3.15**	**Decrease**	**9**	-1.58	Decrease	10	16.6
	K137I	-0.03	Decrease	3	-0.21	Increase	4	70
	A138S	-0.39	Decrease	9	-0.48	Decrease	3	60.8
	N140K	**-2.03**	**Decrease**	**7**	-0.78	Decrease	8	49.9
	D145N	**-2.43**	**Decrease**	**8**	-0.64	Decrease	6	54.1
	Q152R	0.54	Decrease	0	0.04	Increase	2	73.9
FcγRIII	L96P	0.28	Decrease	1	-1.75	Decrease	8	23.5
	T113A	-0.7	Decrease	8	-1.78	Decrease	9	3.6
	V229M	**-3.48**	**Decrease**	**8**	-1.1	Decrease	8	37.7
	V233I	-1.23	Decrease	5	-0.96	Decrease	7	31.9

Table 3.4 Sequence based prediction tools for forward mutations.

	Amino Acid Substitution (WT Position MT)	Provean Score	Provean Prediction	SNP Analyser Prediction	MuSTAB Prediction (protein stability)	MuSTAB confidence	iPTREE-STAB Prediction	iPTREE-STAB Discrimination
FcγRI	G20V	4.732	Neutral	Neutral	INCREASED	26.61%	-1.5838 kcal/mol	positive (stabilizing)
	Q32R	-1.909	Neutral	Neutral	DECREASED	92.32%	-0.9691 kcal/mol	negative (destabilizing)
	A74V	1.025	Neutral	Neutral	INCREASED	29.11%	-1.3200 kcal/mol	negative (destabilizing)
FcγRII	R67M	-1.254	Neutral	Neutral	INCREASED	30.89%	-1.1536 kcal/mol	negative (destabilizing)
	I91V	0.723	Neutral	Neutral	DECREASED	92.50%	-2.5446 kcal/mol	negative (destabilizing)

	Amino Acid Substitution	Provean Score	Provean Prediction	SNPAnalyser Prediction	MuSTAB Prediction (protein stability)	MuSTAB confidence	iPTREE-STAB Prediction	iPTREE-STAB Discrimination
	T105P	4.714	Neutral	Neutral	DECREASED	83.04%	-0.1737 kcal/mol	positive (stabilizing)
	A131T	2.88	Neutral	Neutral	DECREASED	78.75%	-0.1405 kcal/mol	negative (destabilizing)
	I137K	**4.084**	**Neutral**	**Neutral**	**DECREASED**	**98.21%**	**-2.9075 kcal/mol**	**negative (destabilizing)**
	S138A	-0.75	Neutral	Neutral	DECREASED	79.64%	-0.5370 kcal/mol	negative (destabilizing)
	K140N	-2.783	Deleterious	Neutral	DECREASED	89.29%	-0.5850 kcal/mol	negative (destabilizing)
	N145D	-2.023	Neutral	Neutral	DECREASED	89.64%	-1.7270 kcal/mol	negative (destabilizing)
	R152Q	0.47	Neutral	Neutral	DECREASED	85.89%	-0.0586 kcal/mol	negative (destabilizing)
FcγRIII	P96L	5.554	Neutral	Neutral	INCREASED	26.61%	0.4325 kcal/mol	negative (destabilizing)
	A113T	1.787	Neutral	Neutral	DECREASED	81.61%	-0.8613 kcal/mol	negative (destabilizing)
	M229V	0.758	Neutral	Neutral	DECREASED	87.32%	-0.3474 kcal/mol	negative (destabilizing)
	I233V	**-0.705**	**Neutral**	**Disease**	**DECREASED**	**98.21%**	**-0.3474 kcal/mol**	**negative (destabilizing)**

Table 3.5 Sequence based prediction tools for reverse mutations.

	Amino Acid Substitution (WT Position MT)	Provean Score	Provean Prediction	SNPAnalyser Prediction	MuSTAB Prediction (protein stability)	MuSTAB confidence	iPTREE-STAB Prediction	iPTREE-STAB Discrimination
FcγRI	**V20G**	**-4.951**	**Deleterious**	**Disease**	**DECREASED**	**95.71%**	**-3.1872 kcal/mol**	**negative (destabilizing)**
	R32Q	1.965	Neutral	Neutral	DECREASED	93.04%	-0.9691 kcal/mol	negative (destabilizing)
	V74A	-1.165	Neutral	Neutral	DECREASED	94.64%	-1.8655 kcal/mol	negative (destabilizing)
FcγRII	M67R	1.254	Neutral	Neutral	DECREASED	79.64%	-1.1536 kcal/mol	negative (destabilizing)
	V91I	-0.723	Neutral	Neutral	DECREASED	92.50%	-2.5446 kcal/mol	negative (destabilizing)
	P105T	-4.714	Deleterious	Neutral	DECREASED	78.39%	-0.1737 kcal/mol	negative (destabilizing)
	T131A	-2.88	Deleterious	Neutral	DECREASED	92.50%	-0.1405 kcal/mol	negative (destabilizing)
	K137I	**-4.084**	**Deleterious**	Neutral	INCREASED	30.36%	-1.8950 kcal/mol	negative (destabilizing)
	A138S	0.85	Neutral	Neutral	DECREASED	88.93%	0.4520 kcal/mol	negative (destabilizing)
	N140K	2.816	Neutral	Neutral	DECREASED	79.64%	-1.5851 kcal/mol	negative (destabilizing)
	D145N	1.723	Neutral	Neutral	DECREASED	88.93%	-1.7270 kcal/mol	negative (destabilizing)
	Q152R	-0.47	Neutral	Neutral	INCREASED	25.18%	0.8900 kcal/mol	negative (destabilizing)
FcγRIII	L96P	-5.504	Deleterious	Neutral	DECREASED	87.86%	-1.0747 kcal/mol	negative (destabilizing)
	T113A	-1.97	Neutral	Neutral	DECREASED	91.07%	-0.8613 kcal/mol	negative (destabilizing)
	V229M	-0.725	Neutral	Neutral	DECREASED	94.11%	-0.3474 kcal/mol	negative (destabilizing)
	V233I	0.738	Neutral	Neutral	DECREASED	96.79%	-0.3474 kcal/mol	positive (stabilizing)

3.3 Structure Homology Modeling and FOLDx Analysis

Single amino acid mutations can drastically modify the protein structure stability. Thus, the modeling of a protein's 3D structure and the knowledge of its structural information is necessary for complete understanding of its functionality. The 3D structure of the 3 allelic se-

quences of FcγRI, 5 allelic sequences of FcγRII and 3 allelic sequences of FcγRIII were pre-
dicted via homology modeling using AMMP. Model 1 for each of the prediction was considered
to be the structure for the respective sequence. The available structure PDB ID 3RJD was the
reference for FcγRI. PDB ID 3RY4 was the reference structure for FcγRII except for sequence
HQ423391, which used PDB ID 3D5O as the reference. PDB ID 1T83 was the reference struc-
ture for FcγRIII except for sequence HQ423388, which used PDB ID 1E4J as the reference.
The 3D structures obtained for FcγRI and FcγRIII were superimposed over chain c of 1T83 and
those of FcγRII were superimposed over chain c of 3RY6.

The IgG1 sequences of Indian macaques were obtained from NCBI [34]. The 3D struc-
ture of the IgG1 was predicted via homology modeling using AMMP. The predicted structures
were superimposed over chains A and B of 1T83 and 3RY6. The FcγR and IgG1 predicted su-
perimposed structures were combined into single PDB file using UCSF Chimera.

3.4   FOLDx Analysis

The effects of the nsSNPs on protein stability and protein interaction were computed
with Foldx, which uses an empirical energy equation to calculate the Gibbs free energy DDG.
The empirical energy terms consider the location, type of a mutated residue [42]. FOLDx is a
structure based prediction tool. Two different analysis protocols were utilized to obtain maxi-
mum information over the effect of the single amino acid substitutions: 1) All the nsSNPs were
considered singularly and their effect on the protein stability and interaction potential was de-
termined, 2) The nsSNPs were considered according to the allelic sequences obtained from the
previous study. Initially, all the structures were minimized using the RepairPDB function of
Foldx to obtain a stable protein stability value. Then the structures for each single amino acid
variation were generated using the BuildModel feature of Foldx3.0. Finally, the effect of each
single amino acid variation on the interaction between the FcγRs and IgG1 was determined us-

ing the Analyze Complex feature. When the DDG > 0, then the mutation is considered destabi-

lizing and when DDG < 0, it is considered stabilizing.

3.4.1 Effect of Single nsSNP on Protein Stability and Protein Interaction

BuildModel feature of FOLDx3. 0 was used to generate structures for each single amino

acid variations. For this purpose, the amino acid residues at the mutation sites in the first allelic

sequence in each FcγR were considered as the wild type and the amino acid residues from

other sequences were considered to be mutant residues. To avoid any miscalculation in obtain-

ing the effect on the stability, reverse mutations were also analyzed, where the structures gen-

erated from the previous step by BuildModel were used as wild type structure and used as input

for the same procedure of BuildModel. The threshold of +- 1kcal/mol was considered to differ-

entiate between neutral and deleterious amino acid substitution, based on its effect on the pro-

tein stability [24]. The interaction energy between the chain c and chain a + chain b, which

forms the IgG1, was determined using the AnalyzeComplex feature of FOLDx3.0. Then, the

difference between the wild type structure and mutant structure interaction energies was deter-

mined. If the difference is greater than zero, the mutation decreases the affinity between the

two molecules. When the difference is less than zero, it increases the affinity. A threshold of +-

1kcal/mol was considered to analyze whether or not the amino acid substitution affects the pro-

tein interaction.

Table 3.6 SNPs affecting protein stability and protein interaction by FOLDx

	Amino Acid Substitu-tion	Straight Mutations		Reverse Mutations	
		DDG for pro-tein	Difference in Interaction Energy	DDG for pro-tein	Difference in Interaction Energy
FcγRI	G20V	-1.09	0	**1.13**	0
	Q32R	**1.01**	**-1.25**	-0.33	0.5
	A74V	0.22	0.14	-0.19	-0.14
FcγRII	R67M	**1.55**	0	-1.13	0
	I91V	**1.62**	0	-0.3	0
	T105P	0.94	0	**1.04**	0
	A131T	-0.57	**4.62**	0.53	-1.2

	I137K	-0.28	0.1	0.32	0.3
	S138A	-0.54	0.82	0.61	0.01
	K140N	**1.42**	0	-1.59	0.02
	N145D	-0.37	-0.09	0.34	0.15
	R152Q	0.13	0	-0.17	0
FcγRIII	P96L	0.25	0	-0.18	0
	A113T	**2.34**	0	-0.22	0
	M229V	0.73	0	-0.66	0
	I233V	0.11	0	0.37	0

V20G and Q32R in FcγRI were predicted to affect protein stability having DDG values of 1.13 kcal/mol and 1.01 kcal/mol. As the DDG value of Q32R is very near to the threshold value, it was considered to have neutral effect on protein stability. R67M, V91I, P105T and K140N in FcγRII were predicted to affect protein stability having DDG values 1.55, 1.62, 1.04 and 1.42 kcal/mol. Although the DDG value of P105T is near the threshold value, SIFT also predicted deleterious effects on the protein and was thus considered for further analysis. A113Tin FcγRIII was predicted to affect protein stability significantly, having DDG values 2.34 kcal/mol. Q32R in FcγRI is predicted to increase the affinity towards IgG1 and A131T in FcγRII is predicted to affect the protein interaction by greatly decreasing the affinity towards IgG1.

3.4.2   Effect of nsSNPs on Protein Stability and Protein Interaction as per Allelic Sequences

Only the DDG (change) values of FOLDx are trained to approach experimental values. Thus the DG values of the wild type or mutant structures do not provide any inference on their own. Again the first allelic sequence of each FcγR is considered as reference sequence and using the BuildModel feature of FOLDx3.0 the remaining allelic sequence structures were generated. This feature also provides the change in the Gibbs free energy (DDG) as compared to the wild type allelic sequence. The threshold of +- 1kcal/mol was considered to analyze whether or not the amino acid substitution affected the protein stability [24]. Utilizing the AnalyzeCom-

plex feature of FOLDx3.0, the interaction energy between the chain c and chain a + chain b, which forms the IgG1, was determined. The interaction energy of all the allelic sequences was compared. A similar threshold of +- 1kcal/mol was used to analyze the effect of amino acid substitution on the protein interaction.

Table 3.7 Protein stability and protein interaction of allelic sequences by FOLDx

	Allelic Sequences	Straight Mutations		Reverse Mutations	
		DDG for protein	Difference in Interaction Energy	DDG for protein	Difference in Interaction Energy
FcγRI	Rhe3_Rhe8	0.01	-0.31	**-2.17**	0.33
	Rhe3_Rhe9	-0.93	0.14	0.98	-0.09
	Rhe8_Rhe9	0.03	0.22	0.58	-1.82
FcγRII	Rhe2_Rhe4	1.18	-0.77	2.93	-0.6
	Rhe2_Rhe6	**2.9**	-0.38	-1.28	-0.84
	Rhe2_Rhe7	0.03	**4.64**	5.09	-3.45
	Rhe2_Rhe9	4.53	0.07	0.1	-0.22
	Rhe4_Rhe6	4.28	0	-0.93	0
	Rhe4_Rhe7	-0.01	1.23	1.66	-0.23
	Rhe4_Rhe9	4.48	-0.41	-0.8	1.83
	Rhe6_Rhe7	-1.4	6.05	**4.24**	0.07
	Rhe6_Rhe9	2.09	-0.84	0.62	1.67
	Rhe7_Rhe9	**4.68**	-2.3	-2.14	**1.18**
FcγRIII	Rhe3_Rhe4	**3.86**	0	-2.67	0
	Rhe3_Rhe9	0.09	0	1.51	0
	Rhe4_Rhe9	-3.79	0.01	**2.72**	0

FcγRI has three allelic sequences isolated from Rhe3, Rhe8 and Rhe9 [33] (GenBank accession numbers HQ423394-HQ423396). Rhe3 has G, Q and A at positions 20, 32 and 74 respectively, whereas Rhe8 has V, R and V and Rhe9 have V, Q and V at those positions respectively. Rhe8 is more stable than Rhe3 where the protein stability increases by 2.17 Kcal/mol. Rhe8 has more affinity for IgG1 compared to HQ423396.

FcγRII has five allelic sequences isolated from Rhe2, Rhe4, Rhe6, Rhe7 and Rhe9 [33] (GenBank accession numbers HQ423389-HQ423393). Rhe2 has R, I, T, A, I, S, K, N and R at

positions 67, 91, 105, 131, 137, 138, 140, 145 and 152. Rhe4 has V91, P105 and D145; Rhe6 has M67, V91 and D145; Rhe7 has V91, P105, T131, A138, D145 and Q152; Rhe9 has M67, V91, K137, N140 and Q152 compared to Rhe2. Rhe7 is the most stable allelic sequence and Rhe9 is the most unstable with a 4.68 Kcal/mol difference in stability. The order of sequence stability in decreasing order is Rhe7> Rhe4> Rhe2> Rhe6> Rhe9. The highest difference in affinity towards IgG1 is between allelic sequences Rhe6 and Rhe7 with a difference of 6.05 Kcal/mol in interaction energy, where Rhe6 has more affinity towards IgG1. Rhe2 also has more affinity towards IgG1 compared to Rhe7, with a difference of 4.64 Kcal/mol in interaction energy. It can be concluded that P105, T131, A138 and Q152 decrease the affinity for IgG1, from which only the positions 131 and 138 are in the binding site. Thus, a Thr at 131 and Ala at 138 affect the FcγRII-IgG1 interaction.

FcγRIII has three allelic sequences isolated from Rhe3, Rhe4 and Rhe9 [33] (GenBank accession numbers HQ423386-HQ423388). Rhe3 has P, A, M and I at positions 96, 113, 229 and 233. Rhe4 has L96, T113, V229 and V233; Rhe4 has L96 compared to Rhe3. Rhe9 is the most stable allelic sequence and Rhe4 is the least stable. Thus, a Leu at position 96 makes FcγRIII more stable. However, due to the other mutations in Rhe4 at positions 113, 229 and 233 it becomes more unstable. Results from single amino acid variation analysis with FoldX indicated that the Thr at position 113 destabilizes the protein compared to Ala. Thus, we can infer that the effect of Leu at 96 on the stability of the protein is not only masked by the Thr at 113, but also further destabilizes the protein.

3.5   Analysis of Structure Using UCSF Chimera

Only the single amino acid variations that were indicated as affecting protein stability or interaction energy by more than three analysis tools used were chosen to be analyzed structurally by UCSF Chimera. Thus, structural analysis of the mutant and wild type modeled structures for V_{20}/G_{20} in FcγRI, I_{137}/K_{137} in FcγRII and I_{233}/V_{233} in FcγRIII was carried out by superimposing

the structures and comparing the physical properties, solvent accessibility, hydrophobicity and charge density.

V_{20}/G_{20} mutation site in FcγRI indicated that the effect of the mutation may be due to lower hydrophobicity of Gly compared to Val in a hydrophobic patch and decrease protein stability. Gly at 20 has higher solvent accessibility and RSA compared to Val at 20. Val interacts with Leu at position 3 which may take part in stabilizing the protein while folding.

I_{137}/K_{137} mutation site in FcγRII indicated that the effect of the mutation may be due to a strong positive charge and the hydrophilic nature of Lys compared to Ile, which has no charge and is hydrophobic in nature. The hydrogen bond present between Lys137 and Asn135 may stabilize the protein as compared to Ile137. The strong positive charge of Lys137 leads to electrostatic interactions with Lys139, Lys140 and Asp160, further affecting the stability of the protein.

I_{233}/V_{233} mutation site in FcγRIII Ile233 causes additional VdWClashes with Asp253, which may result in decreased protein stability. Furthermore, both residues have similar hydrophobicity properties. FcγR is a membrane bound protein and this mutation site is at the c-terminal of the protein which forms the proximal cytoplasmic tail of the protein. Probably due to this reason the residue is highly conserved and may play an important role in the cell signaling pathways. In which, any change in the residue may affect protein functionality.
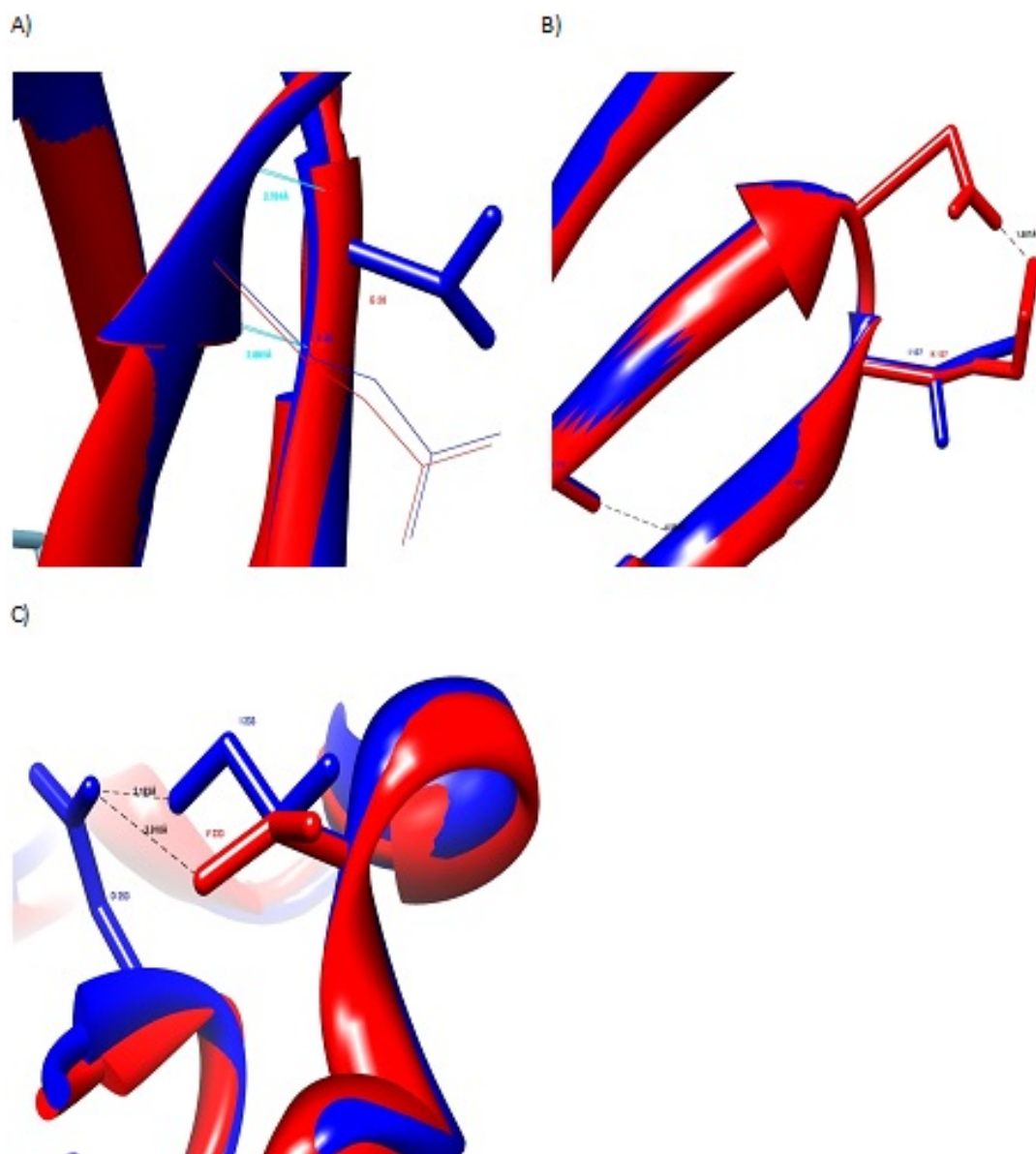
Figure 3.8 A) Superimposed structures of Gly and Val at position 20 in FcγRI. B) Super-imposed structures of Ile and Lys at position 137 in FcγRII. C) Superimposed structures of Ile and Val at position 233 in FcγRIII

A_{131}/T_{131} mutation in FcγRII decreases the affinity of FcγRII for IgGI as indicated by FoldX analysis. The mutant residue Thr is hydrophilic in nature and is involved in hydrogen bonding as opposed to the hydrophobic nature of Ala. The mutation site in present in a buried region with a very low RSA value.   Because of this, the introduction of a hydrophilic residue may affect protein functionality. Thr131 interacts with Ser183 of chain B by hydrogen bonding.

Thr131 also causes VdWClashes AsnB182 and SerB183, which may prevent interaction be-
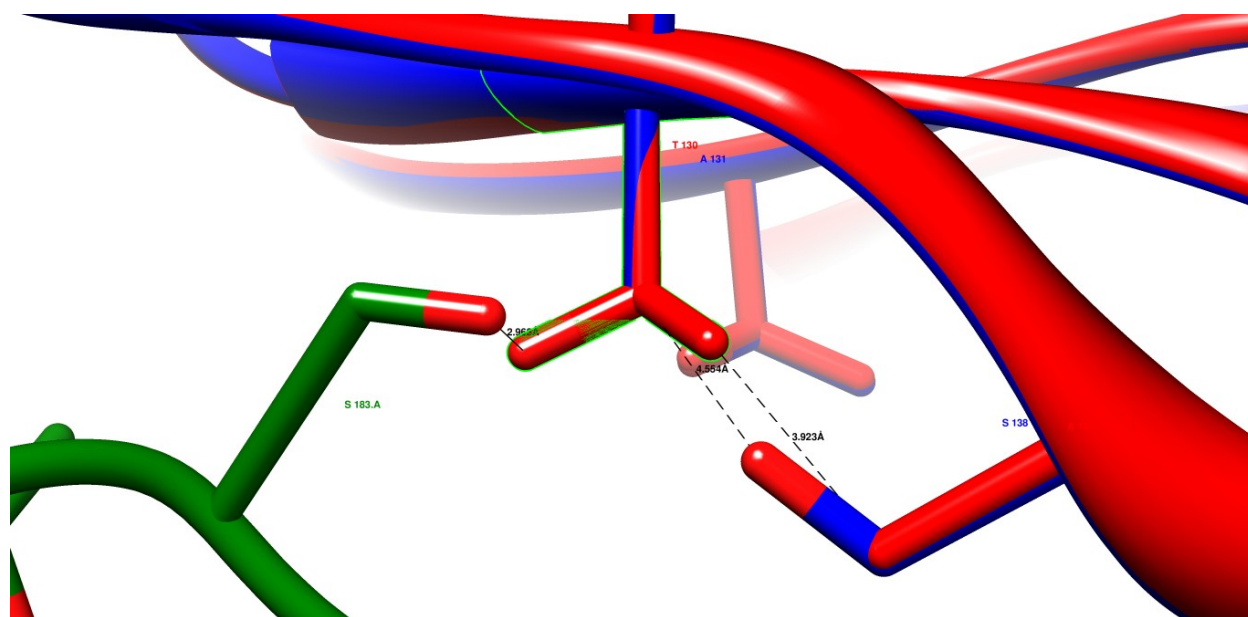
tween FcγRII and IgG1.

Figure 3.9   Superimposed structures of Ala (Blue) and Thr (Red) at position 131 in FcγRII, along with IgG1 (Green).

4   CONCLUSIONS

The number of SNPs reported in the SNP databases such as dbSNP at NCBI is increasing. Due to the large number of SNPs it becomes difficult to plan population based experiments to identify the SNP that can potentially affect the protein function. Various non synonymous SNPs are reported to be involved in disease development. Thus, it becomes necessary to identify potential diseases causing SNPs from the neutral ones. Many *in silico* methods provide information on the effect of the polymorphisms on protein structure and function. Various studies have been performed using *in silico* analysis approaches to predict the functional nsSNPs for genes such as G6PD, BARF and PTEN [1, 2]. The current *in silico* methods have wide range of performance and are chosen depending on different aspects of the proteins. Thus, a single method would be insufficient to provide an accurate prediction of functional SNPs and their effects. Therefore, to increase prediction accuracy, a combination of various methods based on evolutionary conservation, physio-chemical properties and protein structure are utilized. [43]

The nsSNPs in FcRs can affect their affinity towards the immunoglobulin and thus affect the response of the adaptive immune system. nsSNPs in FcγRs have been reported to affect the binding affinity towards the IgGs and clearance rates of the Rhesus D-positive red blood cells [31]. In our study, we investigated naturally present polymorphisms in FcγRs of rhesus macaque to identify functionally important ones. An *in silico* analysis approach was undertaken using SIFT, PROVEAN, nsSNPAnalyzer, I-Mutant, MuSTAB and iPTREE-STAB online web servers. SIFT and nsSNPAnalyzer are highly utilized pathogeneticity prediction methods [44]. Provean is a relatively new prediction method by JCVI [36]. According to SIFT and nsSNPAnalyzer V20G in FCγRI and I233V in FCγRIII were predicted to affect protein function and cause disease. Whereas, V20G in FCγRI and P105T and K140N in FCγRII were predicted by SIFT and PROVEAN to affect protein function. I-Mutant, MuSTAB and iPTREE-STAB predict the

affect of the SNP on the protein stability. MuSTAB only gives the direction of the stability change, whereas I-Mutant and IPTREE-STAB also predict the value of the protein stability change (DDG). V20G in FCγRI and I137K in FCγRII were predicted to decrease protein stability considerably. MuSTAB also predicted I233V in FCγRIII to decrease protein stability, which correlated with the SIFT and nsSNPAnalyzer predictions. FOLDx is also a protein stability predictor, but it utilized structure as compared to the sequence based prediction of the above three predictors. I-Mutant3.0 and FOLDx were reported to give best predictive performance [45].

Structures of the human FcγR are available in Protein Data Bank (PDB) and have been used to analyze the effect of polymorphisms [46,47]. Although the FcγRs of rhesus macaque are highly similar to their human counterparts, a 3D structure is essential for analyzing the impact of the SNPs.   Therefore, we predicted the 3D structure of the rhesus macaque FcγRs through homology modeling using the AMMP program. Depending on the highest sequence similarity according to AMMP profiling, the FcγRI modeled structure was based on PDB structure 3RJD, FcγRII was based on 3RY4 and FcγRII was based on 1T83.   The first sequence from each FcγR was considered to be the native sequence and their structures were predicted. The predicted structures were superimposed on to 1T83 for FcγRI and FcγRII and 3RY6 for FcγRIII. The PDB structures were utilized to superimpose the modeled structures of rhesus FcγR, as the BLAST search of the sequences gave the respective hits. 1T83 is a complex of human FcγRIII and IgG1 and 3RY6 is a complex of human FcγRII and IgG1. There is no significant difference in the native FcγR and FcγR in FcγR-IgG1 complex[47]. Using the FOLDx program, the mutant structures were modeled and analyzed. Each mutation was considered individually to study the inherent effect of the SNP. In addition, the allelic sequences were analyzed to investigate if the polymorphisms neutralize each other by occurring simultaneously as an act of preservation of function by nature.

FOLDx was used to analyze the effects of single amino acid variations on the structure and stability of the protein. Further, differences between the folding energies (kcal/mol) of the

wild type structure and the mutated modeled structures were analyzed. A threshold of +-1 kcal/mol was used to discriminate between functionally important and neutral SNPs. Studies have suggested that a +-1 kcal/mol threshold is not too strict and considers possible SNPs affecting protein stability [24]. A difference of 1 kcal/mol was observed between forward mutation and reverse mutation stability analysis over the conventional sign reversal of DDG, using FOLDx [48]. The amino acids predicted by FOLDx commands AnalyseComplex and PrintNetwork, for rhesus FcγRII and FcγRIII coincided to a certain extent with the amino acids known to be involved in human FcγRII and FcγRIII interaction with human IgG1 [18,19]. The crystal structures of the rhesus FcγRI, FcγRII and FcγRIII should be solved in order to confirm the involvement of the FOLDx predicted amino acids in interaction with rhesus IgG1.

Our results indicate that Val -> Gly at position 20 in FcγRI decreases protein stability having DDG values 1.13 kcal/mol, which may be due to the introduction of a hydrophilic amino acid in a hydrophobic pocket. Val -> Gly at position 20 in FcγRI was predicted deleterious by all the computational methods used in this study.

Polymorphism may not always manifest into a disease, it may lead to change in specificity instead if the mutations are present near the binding site. Thus, decrease in protein stability predicted by FOLDx can lead to change in specificity of FcγR towards IgGs. Our FOLDx analysis suggested Gln -> Arg at position 32 in FcγRI to decrease the stability by 1.01 kcal/mol. The decrease in stability may be caused due to replacement with a strong positive charge. The mutation site is also near the glycosylation site and may interact with 2-(Acetylamino)-2-Deoxy-A-D-Glucopyranose (NDG) molecule. The carbohydrate molecules are important for protein folding [49] and thus the folding energy may be affected due to the addition of a strong positive charged amino acid. Arg -> Met at position 67 in FcγRII is predicted to decrease the stability by 1.55 kcal/mol. The wild type amino acid, Arg, is hydrophilic and has a very strong positive charge, which is replaced by a hydrophobic, no charged amino acid. This may possibly disturb the electrostatic interactions. Ile -> Val at position 91 in FcγRII decreases the protein stability by

1.62 kcal/mol. This decrease could possibly result from the interference in hydrogen bond formation and surface accessibility. Pro -> Thr at position 105 in FcγRII may affect protein stability due to the replacement of the not hydrophobic amino acid with hydrophilic amino acid and decrease the protein stability by 1.04 kcal/mol. SIFT and PROVEAN also predicted the Pro -> Thr change at position 105 to be deleterious, which indicates that the change is in a conserved region. Lys-> Asn at position 140 in FcγRII decreases the protein stability by 1.42 kcal/mol, possibly due to the loss of strong positive charge of Lys. The PDB structure 3RY6 used to superimpose the FcγRII modeled structures had a GLYCEROL (GOL) molecule present at the hinge of the two domains of FcγRII. The mutation site Lys-> Asn at position 140 in FcγRII lies in the vicinity of the GOL molecule and the electrostatic interaction may be disturbed causing loss in stability. Ala-> Thr at position 113 in FcγRIII decreases protein stability significantly by 2.34 kcal/mol, which may be caused by the introduction of hydrophilic amino acid in a hydrophobic pocket, as the wild type residue is also hydrophobic. Further, if a single amino acid variation shows a change in protein stability (DDG values) or protein-protein interaction ($DDG_{binding}$ values), it should give comparable values with the sign reversal for the reverse mutation. This would indicate that the prediction of the effect of the single amino acid variation on the protein structure or protein-protein interaction is substantial. Val -> Gly at position 20 in FcγRI Arg -> Met at position 67 in FcγRII and Lys -> Asn at position 140 in FcγRII give comparable values with the sign reversal for the protein stability and protein-protein interaction analyses by FOLDx, which suggests that these polymorphisms should also be considered as a potential target for future experiments.

In conclusion, *in silico* analysis of the FcγRs indicates that V20G in FcγRI decreases protein stability and is not tolerated. In FcγRII, the I137K destabilizes the protein and A131T causes decrease in affinity of FcγRII towards IgG1. In FcγRIII, I233V and A113T destabilize the protein and may result in loss of function. Val at positions 229 and 233 in FcγRIII were found in the group of SIV infected rhesus macaque which were unable to make ant-SIV antibodies com-

pared to those having Met and Ile at those positions, respectively [50]. Thus for further analysis, these mutations should be given priority to obtain detailed information on their effects. In order to confirm the structure modeled in this study, the actual structures should be determined by X-Ray crystallography or Nuclear magnetic resonance (NMR) spectroscopy. In addition, binding studies should be carried out to investigate the effect of the mutations on the interaction with IgGs [28].

REFERENCES

1. Gardner MB, Luciw PA (2008) Macaque models of human infectious disease. ILAR J 49: 220-255.
2. Shedlock DJ, Silvestri G, Weiner DB (2009) Monkeying around with HIV vaccines: using rhesus macaques to define 'gatekeepers' for clinical trials. Nat Rev Immunol 9: 717-728.
3. Stone WH, Treichel RC, VandeBerg JL (1987) Genetic significance of some common primate models in biomedical research. Prog Clin Biol Res 229: 73-93.
4. Sibal LR, Samson KJ (2001) Nonhuman primates: a critical role in current disease research. ILAR J 42: 74-84.
5. Johnson RP (1996) Macaque models for AIDS vaccine development. Curr Opin Immunol 8: 554-560.
6. Herodin F, Thullier P, Garin D, Drouet M (2005) Nonhuman primates are relevant models for research in hematology, immunology and virology. Eur Cytokine Netw 16: 104-116.
7. Staprans SI, Feinberg MB, Shiver JW, Casimiro DR (2010) Role of nonhuman primates in the evaluation of candidate AIDS vaccines: an industry perspective. Curr Opin HIV AIDS 5: 377-385.
8. Jefferis R, Lund J (2002) Interaction sites on human IgG-Fc for FcgammaR: current models. Immunol Lett 82: 57-65.
9. van de Winkel JG, Capel PJ (1993) Human IgG Fc receptor heterogeneity: molecular aspects and clinical implications. Immunol Today 14: 215-221.
10. Murphy K (2012) Janeway's Immunobiology: Garland Science.
11. Clynes R, Takechi Y, Moroi Y, Houghton A, Ravetch JV (1998) Fc receptors are required in passive and active immunity to melanoma. Proc Natl Acad Sci U S A 95: 652-656.
12. Lanzavecchia A (1990) Receptor-mediated antigen uptake and its effect on antigen presentation to class II-restricted T lymphocytes. Annu Rev Immunol 8: 773-793.
13. Manca F, Fenoglio D, Li Pira G, Kunkl A, Celada F (1991) Effect of antigen/antibody ratio on macrophage uptake, processing, and presentation to T cells of antigen complexed with polyclonal antibodies. J Exp Med 173: 37-48.
14. Villinger F, Mayne AE, Bostik P, Mori K, Jensen PE, et al. (2003) Evidence for antibody-mediated enhancement of simian immunodeficiency virus (SIV) Gag antigen processing and cross presentation in SIV-infected rhesus macaques. J Virol 77: 10-24.
15. Brady LJ (2005) Antibody-mediated immunomodulation: a strategy to improve host responses against microbial antigens. Infect Immun 73: 671-678.
16. Sylvestre DL, Ravetch JV (1994) Fc receptors initiate the Arthus reaction: redefining the inflammatory cascade. Science 265: 1095-1098.
17. Oettgen HC, Martin TR, Wynshaw-Boris A, Deng C, Drazen JM, et al. (1994) Active anaphylaxis in IgE-deficient mice. Nature 370: 367-370.
18. Radaev S, Motyka S, Fridman WH, Sautes-Fridman C, Sun PD (2001) The structure of a human type III Fcgamma receptor in complex with Fc. J Biol Chem 276: 16469-16477.
19. Ramsland PA, Farrugia W, Bradford TM, Sardjono CT, Esparon S, et al. (2011) Structural basis for Fc gammaRIIa recognition of human IgG and formation of inflammatory signaling complexes. J Immunol 187: 3208-3217.

20. Alshatwi AA, Hasan TN, Syed NA, Shafi G, Grace BL (2012) Identification of functional SNPs in BARD1 gene and in silico analysis of damaging SNPs: based on data procured from dbSNP database. PLoS One 7: e43939.
21. Teng S, Srivastava AK, Schwartz CE, Alexov E, Wang L (2010) Structural assessment of the effects of amino acid substitutions on protein stability and protein protein interaction. Int J Comput Biol Drug Des 3: 334-349.
22. Doss CG, Nagasundaram N (2012) Investigating the structural impacts of I64T and P311S mutations in APE1-DNA complex: a molecular dynamics approach. PLoS One 7: e31677.
23. Teng S, Srivastava AK, Wang L (2010) Sequence feature-based prediction of protein stability changes upon amino acid substitutions. BMC Genomics 11 Suppl 2: S5.
24. Juritz E, Fornasari MS, Martelli PL, Fariselli P, Casadio R, et al. (2012) On the effect of protein conformation diversity in discriminating among neutral and disease related single amino acid substitutions. BMC Genomics 13 Suppl 4: S5.
25. Ternant D, Buchler M, Beneton M, Alvan G, Ohresser M, et al. (2008) Interindividual variability in the concentration-effect relationship of antilymphocyte globulins - a possible influence of FcgammaRIIIa genetic polymorphism. Br J Clin Pharmacol 65: 60-68.
26. Bruhns P, Iannascoli B, England P, Mancardi DA, Fernandez N, et al. (2009) Specificity and affinity of human Fcgamma receptors and their polymorphic variants for human IgG subclasses. Blood 113: 3716-3725.
27. Li X, Ptacek TS, Brown EE, Edberg JC (2009) Fcgamma receptors: structure, function and role as genetic risk factors in SLE. Genes Immun 10: 380-389.
28. de Haas M, Koene HR, Kleijer M, de Vries E, Simsek S, et al. (1996) A triallelic Fc gamma receptor type IIIA polymorphism influences the binding of human IgG by NK cell Fc gamma RIIIa. J Immunol 156: 2948-2955.
29. Binstadt BA, Geha RS, Bonilla FA (2003) IgG Fc receptor polymorphisms in human disease: implications for intravenous immunoglobulin therapy. J Allergy Clin Immunol 111: 697-703.
30. Manger K, Repp R, Jansen M, Geisselbrecht M, Wassmuth R, et al. (2002) Fcgamma receptor IIa, IIIa, and IIIb polymorphisms in German patients with systemic lupus erythematosus: association with clinical symptoms. Ann Rheum Dis 61: 786-792.
31. Miescher S, Spycher MO, Amstutz H, De Haas M, Kleijer M, et al. (2004) A single recombinant anti-RhD IgG prevents RhD immunization: association of RhD-positive red blood cell clearance rate with polymorphisms in the FcgammaRIIA and FcgammaIIIA genes. Blood 103: 4028-4035.
32. Siddiqui RA, Krawczak M, Platzer M, Sauermann U (2011) Association of TLR7 variants with AIDS-like disease and AIDS vaccine efficacy in rhesus macaques. PLoS One 6: e25474.
33. Nguyen DC, Scinicariello F, Attanasio R (2011) Characterization and allelic polymorphisms of rhesus macaque (Macaca mulatta) IgG Fc receptor genes. Immunogenetics 63: 351-362.
34. Scinicariello F, Engleman CN, Jayashankar L, McClure HM, Attanasio R (2004) Rhesus macaque antibody molecules: sequences and heterogeneity of alpha and gamma constant regions. Immunology 111: 66-74.
35. Ng PC, Henikoff S (2003) SIFT: Predicting amino acid changes that affect protein function. Nucleic Acids Res 31: 3812-3814.
36. Choi Y, Sims GE, Murphy S, Miller JR, Chan AP (2012) Predicting the functional effect of amino acid substitutions and indels. PLoS One 7: e46688.

37. Huang LT, Gromiha MM, Ho SY (2007) iPTREE-STAB: interpretable decision tree based method for predicting protein stability changes upon mutations. Bioinformatics 23: 1292-1293.
38. Bao L, Zhou M, Cui Y (2005) nsSNPAnalyzer: identifying disease-associated nonsynonymous single nucleotide polymorphisms. Nucleic Acids Res 33: W480-482.
39. Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, et al. (2004) UCSF Chimera--a visualization system for exploratory research and analysis. J Comput Chem 25: 1605-1612.
40. Schymkowitz J, Borg J, Stricher F, Nys R, Rousseau F, et al. (2005) The FoldX web server: an online force field. Nucleic Acids Res 33: W382-388.
41. Tokuriki N, Stricher F, Schymkowitz J, Serrano L, Tawfik DS (2007) The stability effects of protein mutations appear to be universally distributed. J Mol Biol 369: 1318-1332.
42. Tokuriki N, Stricher F, Serrano L, Tawfik DS (2008) How protein stability and new functions trade off. PLoS Comput Biol 4: e1000002.
43. George Priya Doss C, Rajith B (2012) A New Insight into Structural and Functional Impact of Single-Nucleotide Polymorphisms in PTEN Gene. Cell Biochem Biophys.
44. Thusberg J, Olatubosun A, Vihinen M (2011) Performance of mutation pathogenicity prediction methods on missense variants. Hum Mutat 32: 358-368.
45. Khan S, Vihinen M (2010) Performance of protein stability predictors. Hum Mutat 31: 675-684.
46. Oshima Y, Fujimura A (2002) Implication for how the single nucleotide polymorphism (SNP) of Fc receptor, Fc gamma RIIIa alters the interaction with anti-CD20 monoclonal antibody. Blood 99: 4649; author reply 4650.
47. Kim TD, Cho SE, Yang CH, Kim J (2001) Analysis of Fc gammaRIII and IgG Fc polymorphism reveals functional and evolutionary implications of protein-protein interaction. J Mol Evol 53: 1-9.
48. Christensen NJ, Kepp KP (2012) Accurate stabilities of laccase mutants predicted with a modified FoldX protocol. J Chem Inf Model 52: 3028-3042.
49. Ferrara C, Stuart F, Sondermann P, Brunker P, Umana P (2006) The carbohydrate at FcgammaRIIIa Asn-162. An element required for high affinity binding to non-fucosylated IgG glycoforms. J Biol Chem 281: 5032-5036.
50. Miller CJ, Genesca M, Abel K, Montefiori D, Forthal D, et al. (2007) Antiviral antibodies are necessary for control of simian immunodeficiency virus replication. J Virol 81: 5024-5035.