Protein structure modeling for variant pathogenicity prediction

Author: Sylt Schuurmans
Studentnumber: 333332
Study: Bio-informatica
Hanze hogeschool: Institute for Life Science and Technology
Hospital: Universitary Medical Center Groningen: Department of Genetics
Supervisor : Joeri van der Velde

June 4, 2019

Introduction

Around 1 in 17 people is affected by one of 7,000 known rare diseases. Most of these patients do not receive a diagnosis, which means they remain in uncertainty without a prognosis, are unable join specific patient support groups, and do not receive the most appropriate treatment. Next-generation sequencing (NGS) of DNA promises to establish a molecular diagnosis and help these patients but many challenges still stand in the way of maximum success. Recent years have seen great advances in computational tools that quickly reduce the amount of DNA variants to be interpreted by a human expert for potentially pathogenic effects []. Although algorithms can now safely remove around 95% of the harmless variants, this still leaves hundreds of variants to be investigated for a whole-exome sequenced patient, which is far too many for a quick and clear diagnosis. Current tools to predict variant pathogenicity rely on features such as evolutionary conservation, annotation of regulatory genomics elements or structural DNA features. These tools have already been optimized over many years and further significant improvements are not expected. Therefore there is still a great need for even more powerful variant prioritization tools. A refreshing alternative was presented by VIPUR [] which shows the potential of structural modelling of proteins to predict the actual effect of a specific variant on the function of that protein. This presents an exciting new opportunity to improve genome diagnostic variant prioritization. However, this predictor was (i) not integrated with the latest and greatest variant pathogenicity prediction approaches, (ii) was trained on relatively small number of variants, and (iii) did not result high quality software that was ready to be taken into routine diagnostic practice. To test this approach we will explore the potential pitfalls of protein modeling by evaluating the VIPUR pipeline and by examining a single protein with it variants.

Abbreviations

3D Three Dimensional
ACCP Solvent Accessible Surface Area
API Application Programming Interface
Bash Bourne Again Shell
CPU Central Processing Unit
CSV Comma Seperated Values
DNA Deoxyribonucleic Acid
DISC Death-Inducing Signaling Complex
FADD Fas Associated Death Domain protein
FasL Fas Ligand
FEM Fixed End Move
FHF Familial Hibernian Fever
GAVIN Gene-Aware Variant INterpretation
GRCh/hg Genome Reference Consortium Human Human Genome
HOPE Have yOur Protein Explained
LOMETS Local Meta-threading Server
MD Molecular Dynamics
MPI Message Parsing Interface
NCBI National Center for Biotechnology Information
NF-κB Nuclear Factor kappa-light-chain-enhancer of activated B cells
OpenGL Open Graphics Library
OS Operating System
OSF Open Science Framework
PDB Protein Data Bank
PM Pivot Movement
PSI-BLAST Position Specific Iterative BLAST
PSSM Position Specific Scoring Matrix
RCSB Research Collaboratory for Structural Bioinformatics
RNA Ribonucleic acid
RSMD Root Mean Square Deviation
SASA Solvent Accessible Surface Area
SCOP The Structural Classification of Proteins
SLURM Simple Linux Utility for Resource Management
SODD Silencer of Death Domain
SPVAA Simple Protein Variant Analysis Approach
TNF Tumor Necrosis Factor
TNFR1 Tumor Necrosis Factor Receptor Superfamily Member 1A TNFRSF1A Tumor Necrosis Factor Receptor Superfamily Member 1A
TRADD Tumor Necrosis Factor Receptor type 1-Associated DEATH Domain protein
TRAPS Tumor Necrosis Factor Associated Receptor-Associated Periodic Syndrome
VIPUR Variant Interpretation Using Rosetta
VTS VIPUR Training Set

Contents

List of Figures

List of Tables

1 Variant prediction in genome diagnostics and the addition of protein modeling

1.1 Mutations and its effects in the central dogma of molecular biology

Within the human genome mutations occur continuously by internal and external factors that substitute, remove, insert or alter the reading frame in a nucleotide sequence. Mutations are not without consequences and can be: beneficial, benign or in most cases pathogenic because they replace a nucleotide which serves a purpose at the specific position in a sequence. Alterations in sequences might lead to a difference in ribonucleic acid (RNA) transcription rates or differences in the RNA transcript that is formed from the deoxyribonucleic acid (DNA) which both can influence the cellular machinery. Mutations outside a gene could lead to lowered or heightened transcription of a protein, when a mutation resides inside a gene it could lead to proteins that are: unstable during or after formation, perform less optimal or are not functional [].

1.2 A general concept of structural levels within proteins and the effect of mutations

The formation of protein structures is classified in different levels, distinctions are made based on bindings and structures that arise from them. The order in which amino acids appear in a sequence is called the primary structure, in this level amino acids are only bound to each other by peptide bonds. Within a primary structure amino acids can form new peptide bonds between the N and C -terminus of an amino acid, with these bonds 3D structures are made called α helices and β-sheets that together make up the secondary structure. More alterations to a single amino acid sequence in the 3D can come from disulfide bridges, ion, hydrogen -bonds, hydrophobic and hydrophilic -interactions formed by the residues of the amino acids, together these bonds form the tertiary structure. By combining multiple tertiary structures the quaternary structure of a protein can be formed out of the mentioned bonds, bridges and interactions [].

Mutations within proteins can have different effects to protein structures, often single missense mutations often have minimal effect on the backbone of a protein [] but can result in destabilization of the structure when assembled or can disrupt the active site. Frameshift mutations often cause larger disruptions within the structure and often lead to proteins that are deformed or have early stop codons [].

1.3 Addition of structural data to diagnosis and treatment in healthcare

Acquiring information about DNA sequences highly relies on experimental sequencing methods and became cheaper over the years [] and found its use in diagnosing patients within the healthcare sector []. From the collected data by genome sequencing experiments most of the analysis is handled in-silico due to the quantities of data that is produced. Proteins often find their use in diagnosing diseases experimentally [], however in-silico it is often limited to information about conservation in the amino acid sequence which may lead to identical results as by analyzing DNA. Yet the 3D structure defines how a protein functions [] and by assessing structures of protein variants it becomes possible to determine the change in function and diagnose protein variants that were unclassifiable through finding conservation. Another advantage of the structural information is that it gives the possibility to develop treatment for diseases that are caused by a mutations. With experimental methods such as X-ray crystallography and nuclear magnetic resonance (NMR) more than 158000 structures [] have been completely revealed, however it is only a tiny fraction of the potential proteins possible without folds []. Making 3D structures is currently not common for diagnosis because it is relative expensive and is difficult to perform, some structures contain flexible regions which makes it hard to determine their exact position and can cause information loss about the structure [].

1.4 Protein modeling techniques

An alternative approach to determine structures is based on modeling the protein structure computationally from the amino acid sequence of the desired protein. A downside from computer generated models is that they do not follow the rules of physics and therefore not automatically fold into the correct confirmation. With the method homology modeling sequences of the requested protein are aligned to sequences of known experimental determined structures, based on these alignments a template is formed whereon structural fragments are built, it is not recommended to use this strategy if the sequence identity is less than 20% since there might not be any structural relation at that point []. Another approach is protein threading which relies on the observation of folds in previous determined experimental structures. Based on the occurrence of specific folds a probability is predicted that a certain residue in a protein might fold in that manner.

Strategies are continuously being improved and developed for proteins to determine the unknown structures, but all have the similar guidelines in avoiding steric hindrance [] and finding the lowest energies based on different scoring systems []. From the computer generated models many are less accurate than the experimental determined methods and are often compared to them for reference. However the computational models do not have follow the same laws of physics which bottleneck the current experimental methods in for example determining membrane proteins [].

1.5 A theoretical large scale implementation of structural protein variant assessment

With the wide spectrum of potential different proteins it can be difficult and maybe momentarily impossible to produce any form of universal protein assessment standardization that is able to determine if a mutation is harmful or not based on structural information. However a first step to solve such a complex problem would be by determining the correct approach, in this case it is assumed that a machine learning approach would be the best method for detecting patterns in structures and classifying the effect of structural changes. Because it has the ability to learn from structural mutations currently available, assuming that the current knowledge about structures and mutations is correct, and is able to develop new insights in how structural changes could affect proteins.

Since the problem is so complex it should be divided into smaller more feasible problems, beginning by separating the different protein classes, which for example can be done according to The Structural Classification of Proteins database (SCOP) []. A first discrimination between the proteins could be made based on protein type/fold class (membrane, globular, fibrous and disordered -proteins) because these differences already predetermine some functions and locations for certain proteins in a cell []. After formation of these classes each should have its own machine learning method applied so their features can be analyzed within context of where and how they function. The next set of discriminators is highly dependent on the variations in classes, but all have features in the end describing bonds, interactions and movement of complexes in protein structures. When for each of the main classes a method has been developed a meta classifier will determine based on certain aspects which method should be applied to determine the effect of mutation in a protein.

2 Monte Carlo method

2.1 Monte Carlo method

There are complex problems in a variety of research fields which could take up years or even centuries to compute with simple deterministic methods. For some problems there is an algorithm which makes it possible to cut down computation time significantly, but when no deterministic algorithm is available to speed up the process an empirical probabilistic method might be able to approximate the desired result. With the Monte Carlo method random samples are taken from the parameter space ,that describe a data set, and fed into a model which produces a potential outcome. By repeating the process more results are generated until at some point the data can display a pattern that describes the outcome. The result is a quantified probability which describes the chance that something might occur based on the quantity of occurrence generated by the model [].
The Monte Carlo methods can differ depending on the algorithm and application in which it is used, but in summary most implementations will follow a general pattern []:

0. Construct a model which is able to describe an outcome of the problem.
1. Define the space of which inputs can be used by the model to get an outcome (creating a parameter space).
2. Use the model to generate results based on random sampled input from the parameter space.
3. Order and determine which results are part of a certain outcome and draw conclusions on the generated statistical evidence.

2.2 The use of the Monte Carlo method and its pitfalls

The Monte Carlo method is widely used within various applications in different fields of science but it is limited in the type of problems it can solve and is suitable for; problems of which all the inputs are known but it is too inefficient to compute deterministically; situations that require uncertainty to be incorporated into the analysis and exploring parameters for a model that give a better impact than the current parameters. The mentioned type of problems it can solve all tend to rely on significant quantities of data which makes it a relative time consuming process for generating results. Meaning of the generated result is highly depended on the model and random sampling techniques which both contribute to an errors in the result [].

3 TRAPS disease and its proteins

3.1 Tumor Necrosis Factor Receptor Associated Syndrome

Tumor necrosis factor receptor-associated periodic syndrome (TRAPS) is classified as a rare disease (1 : 1,000,000) and was formerly known as Familial Hibernian fever (FHF) [], is a hereditary autosomal dominant disease which can cause recurring fevers with a duration from days up to several months. Symptoms during these fevers are: skin rash, swelling, inflammatory reactions across the whole body and pain in the abdomen, muscles and/or joints, a long term and lasting effect is the accumulation of amyloid within the kidneys and may result in other diseases []. TRAPS is known to be caused by mutations within the gene turmor necrosis factor receptor 1 (TNFRSFA1/TNRF1) (Section 3.2), the mutated proteins tend to get trapped in the cell and will be unable to reach the cell surface and therefore start activating a inflammatory response []. So far 158 mutations have been associated with the disease [], but more mutations have been identified in TNFRSF1A wherein some might be pathogenic (Sections 4.10, 4.11).

3.2 Tumor Necrosis Factor Receptor Super Family Member 1A

Tumor Necrosis Factor Receptor Super Family Member 1A (TNFRSF1A, TNFR1) is a gene located on chromosome 12 region 1 band 3 and sub-band 31. The gene produces a trans-membrane receptor consisting of 445 residues divided into 221 residue cytoplasmic section and a 171 extracellular part that consists of 4 conserved cysteine rich domains []. The receptor is ubiquitous across most cell surfaces ,but not on erythrocytes [], and can form two different types of unbound hexagonal clusters depending on the dimer formation []. When the structures are dimers the binding sites are exposed and make it possible for tumor necrosis factor (TNF) α and β (Section 3.3) to bind in trimeric form, with binding of TNF the dimers disconnect and three TNFR1s interact with the TNF trimer []. With the interaction of the TNF trimers with TNFR1 it can activate several pathways such as; the nuclear factor kappa-light-chain-enhancer of activated B cells (NF-κB), which enhances the transcription of various genes during inflammation, infection or other forms of external stress; also it is able to activate the extrinsic pathway of apoptosis after binding of TNF to TNFR1, by releasing the silencer of death domain (SODD) proteins release on the cytoplasmic site. Tumor Necrosis Factor Receptor type 1-Associated DEATH Domain protein (TRADD) [] will start to bind together with proteins that will form a complex which will attract Fas associated death domain (FADD) and after two hours[] if not inhibited. On binding of FADD initiator caspase 8 starts a cascade wherein caspase 3 is activated an will cleave aspartate out of proteins and thereby disrupting the metabolism[].

3.3 Tumor Necrosis Factor Alpha and Beta

The proteins TNF α and β are both pro-inflammatory cytokines that are produced as response to an infection or when a cell is damaged. Both are transcribed from their genes that reside in chromosome 6 in the p-arm at region 2 band 1 and sub-band 3. TNF α and β are 35% identical and 50% homologous to each other consisting out of 233 and 205 amino acid residues. Both are able to form a homotrimeric structures that can bind to the dimeric form TNFR1 (Section 3.2) to activate the extrinsic pathway[].

4 Materials and methods

4.1 VIPUR approach

VIPUR is a machine learning approach for predicting pathogenicity of proteins. The 106 features that were used for machine learning originate mainly (94%) from the Rosetta software suite (Section 4.3) applications; DDG monomer (Section 4.3.2), Relax (Section 4.3.1) and Rescore (Section 4.3.3), the remaining features were collected from PSI-BLAST (Section 4.5) and Probe (Section 4.6). All proteins in the VTS of which structures were known or had fragments available were collected from Modbase [] and SWISS-MODEL []. Proteins that did not have a structure within the databases were modeled with Modeller (Section 4.9 based on protein fragments that had the highest amino acid sequence identity to the protein. In some experimental determined structures duplicate chains, ligands, metals and non-standard amino acids were present, these inconsistencies are able to alter the features generated by software and could in some case hinder feature collection, therefor they were removed to make the data homogeneous. Structural mutations of proteins that are in the VTS were introduced by a script using PyMOL (Section 4.15) by default or PyRosetta (Section 4.4) if PyMOL was not available.

4.2 Simple protein variant analysis approach

Another approach for determining pathogenicity of a mutation is by assessing energy differences between a wild type and mutant protein residues inside its complex. Analyzing mutations from this perspective gives the ability to view a complex in whole and determine how residues cause perturbations in a complex. Missense mutations in monomers of complexes were made with Modeller (Section 4.9) and the backbone was refined with Rosetta's backrub application (Section 4.3.4), to lower the energy levels within side chains Rosetta relax (Section 4.3.1). This method shows similarities to that of VIPUR, was tested with TNFRSF1A (Section 3.2) and its ligands TNF α and β. This method keeps: duplicate chains ligands and metals within the structure, water is excluded since it can cause issues with Rosetta tools (Section **??**).

4.3 Rosetta

Rosetta is a software suite that has a variety of tools that are developed to aid in macro molecular and antibody ,analysis, design and modeling []. Both approaches rely on the Relax (Section 4.3.1) for minimizing side chains. VIPUR uses rescore (Section 4.3.3) to acquire information about protein structures.

Both methods rely on Relax to minimize energies in the side chains of the remodeled structures. With DDG monomer both rely on energy minimization's in the side chains of the protein structures and need to information on energy changes in

The scores generated for the machine learning within the VIPUR approach rely on results generated by Rosetta software and to apply this approach the steps are reproduced. Several strategies were employed for realizing mutated structures, the first strategy was to identify the whole structure of proteins

The initial structure of the protein was produced with the application abinitio relax. For the prediction the application requires an amino acid sequence to identify homologous sequences in a curated database. Homologous sequences within the database are found by the BLAST algorithm, when a

For the search of the sequences it uses the BLAST algorithm and to find homologous amino acid sequences which have protein structures.

requires an amino acid sequence and it takes an amino acid sequence as input and searches in a curated protein database BLAST for finding homologous sequences.

to align sequences with to acquire homologous sequences. The homologous With these sequences it finds structures related to the protein For the prediction of the initial structure of TNFR the application abinitio relax was used.

With this tool a sequence is inserted as input that is aligned to

Missense mutated proteins have an altered amino acid that can cause differences in interactions with other amino acids, which can influence the backbone or side chain positions of a protein and therefore affect the structure. Software that makes missense mutations in protein structures (Modeller, PyMOL, PyRosetta) tend to replace residues without optimizing, causing odd energy levels or steric hindrance to arise.
Rosetta software suite Version 3.10

4.3.1 Relax

Relax was the only application used by both methods which tried to minimize energies in local conformational search space [] within the mutated structures. From each minimization attempt the structure was saved and scores for certain properties were calculated and written into a single file. From this score file VIPUR collected all samples and made 83 features out of it, the detailed approach used the scores from a single structure for its assessment.

4.3.2 DDG Monomer

DDG monomer is meant to predict energetic stability of a point mutation in monomeric protein.The application was used by VIPUR to collect features related to energies and hydrogen, disfulfide, bonds and constraints differences between the wild type and a mutated protein. To execute the tool a script had to be ran that renumbers the wild type pdb file and it requires a "mutation file" that describes the change of a residue based on name and position changes to a different residue [].

4.3.3 Rescore

With this tool Rosetta scores can be calculated based on silent or PDB files proteins structures [] , the output is identical to that is written within the score files produced by Relax (Section 4.3.1).

4.3.4 Backrub

The backrub application is based on the Monte Carlo method (Section 2.2), and alters a protein by moving its backbone residues with a strategy called fix end move (FEM). With this strategy, groups of residues are selected at random from the structure, it can contain up to: four dihedral, two bond angles and two end points. Both ends of a group are fixated at their position in which a new angle α arises, within this angle residues are pivoted in their natural occurring maximum range of $\pm 10°$ []. With this method the backbones of newly introduced mutations were altered, for each attempt a new file was generated and a score was written to a score file, from which the lowest Rosetta scoring was selected to be further relaxed (Section 4.3.1). It was used on the mutated protein to relax the modified backbone structure.

4.4 PyRosetta

Is an application programming (API) which has Python bindings (Section 4.18) for the Rosetta software suite (Section 4.3), it founds its use in VIPUR when no PyMOL (Section 4.15) was available to mutate residues within a structure [].
Version 4

4.5 PSI-BLAST

Position specific iterative basic local alignment search tool (PSI-BLAST) focuses on distant relatives of proteins by making a profile of the sequence and querying it at a protein sequence database. With the generated results a new profile is constructed and queried again, these steps are repeated several times to determine which residues are found in relatives of the protein. The result is a position specific scoring matrix (PSSM) which describes the frequency of which residues are substituted by a specific

other residue, positive is more, negative is less common []. From the PSSMs sequences features were acquired for the VIPUR machine learning method.
Position-Specific Iterated BLAST 2.7.1+

4.6 Probe

Probe is able to evaluate atom packing for a single protein or interacting proteins by creating a probe, which is described as a sphere like object, that marks an area with dots when at least two non-covalent atoms are in contact with the probe at the same position []. VIPUR used this tool to calculate solvent accessible surface area (SASA or ACCP).
version 2.16.130520

4.7 Robetta prediction server

The web tool Robetta integrates several tools to form protein structures based on sequence alignments of previously discovered structures also known as homology modeling (Section 4.2). It requires an amino acid sequence, optionally constrains and fragments can be added to disallow movement of certain structures or add known fragments to avoid calculating pieces that are already known. With this information Robetta search with the help of sequence aligners for known fragments and tries to incorporate them into a single protein structure []. The used structures of TNFRSF1A (Section 3.3) were in complete and could therefore lack information regarding the structure when a mutation is introduced. To form a whole protein the, fragments of several known structures are joined by the Abinitio protocol within act on th with this web tool it was possible to predict the missing pieces of the protein

4.8 I-TASSER prediction server

The I-TASSER web server is a tool that is able to predict protein structures with a FASTA sequence. The first step it takes is finding structural templates which resemble the sequence by local meta-threading server (LOMETS). LOMETS starts with multiple sequence alignment of which several sequences will undergo protein threading by different programs to form structural templates. The templates are assessed based on the highest alignment Z-score, the program specific confidence score and sequence identity []. The known fragments of TNFRSF1A (Section 3.3) were given as a template to I-TASSER and modeled into a whole protein to make it possible to introduce mutations and predict pathogenicity of a variants.
Server version

4.9 Modeller

Modeller is software that is developed for homology modeling but it was used for its utilities which allowed to; complete protein data bank (PDB) structures with missing atoms; predict disulfide bonds that were missing and mutate protein residues [].
Version 9.21

4.10 GAVIN Machine Learning Data Table

Is a collection of nucleotide mutations from rare diseases used by the GAVIN [] machine learning approach. From this set the genes of TNFRSF1A (Section 3.3) with a missense mutation were filtered (Section 4.19) and written into a format which the variant effect predictor could (VEP) [] could read and translate from nucleotide to protein mutations. The classification of these variants was according to Clinvar significance values [].

4.11 GenomAD

The GenomAD database consists of unified data from large scale genome sequencing data projects and is based on genome reference consortium human genome build 37 human genome 19 (GRCh37/hg19). From this database missense mutations were collected for TNFRSF1A (Section 3.3), no classification was known for these mutations [].

4.12 Infevers

Is a website about hereditary auto immune diseases with for each disease a downloadable table about the known mutations and their classification. The table for TRAPS disease (Section 3.3) was used to collect missense mutations of TNFRSF1A gene [].

4.13 Research Collaboratory for Structural Bioinformatics

Research Collaboratory for Structural Bioinformatics (RCSB) is a database where whole or fragmented experimentally determined proteins structures that are published can be found and downloaded. The Fragments for modeling (Sections 4.3, 4.8) whole TNFRSF1A (Section 3.3) (1EXT []) and determining the differences in energy levels (Section 4.3.1) with TNF β (1TNR []) with the interaction site were acquired from this database [].

4.14 Uniprot

Knowledge from various omic domains about proteins has been linked together into single database called Uniprot which makes all information accessible at once, for TNFRSF1A (Section 3.3) the FASTA sequences were collected from Uniprot and for structures it redirected to (Section 4.13) [].

4.15 PyMOL

Visualization of 3D structures, making images of proteins, putting the known orientations of monomeres in position, replacing TNF β with TNF α and aligning the structures to measure the distance between X-ray crystal structures and the produced models of were done in PyMOL []. PyMOL had a different use in VIPUR where it was used in combination with Python (Section 4.18) to perform mutagenesis on the protein structures to introduce a missense mutations.
Version 2.2.3

4.16 HOPE

Have yOur Protein Explained (HOPE) is a web service that collects information of about a user specified missense mutation in a protein and comes from various sources. Uniprot (Section 4.14) is queried with BLAST to find homologous sequences and structures, other features that are found on Uniprot are active sites, domains and various other sequence features that help to identify the function of a region. From the BLAST results homology models are made with Yasara that are sent of to WHAT IF web services that calculate structural information about the protein. Before the formation of a report all information is put into a decision tree to asses mutational effects in contexts of: contacts, structural locations, non-structural features, previous variant information and amino acid properties to form an automated report []. *Version 1.1.1*

4.17 Bash

Unix like operating systems (OS) have a shell which allows users to interact with programs on a computer or with the computer itself based on commands submitted. The default shell for MacOS and also for several Linux distributions is the Bourne again shell (Bash) which was used to launch Python scripts

(Section 4.18) and submit jobs to the SLURM workload manager (Section 4.20).
Laptop Version GNU bash, version 3.2.57(1)-release (x86_64-apple-darwin18)
Server Version GNU bash, version 4.1.2(2)-release (x86_64-redhat-linux-gnu)

4.18 Python

Both VIPUR and the pipeline that minimizes backbone (Section 4.3.4) and side chain energies (Section 4.3.1) were written in Python due to its capabilities, ease of use and because modeller (Section 4.9) for MacOS relies on the system version of Python and does currently not support newer versions besides the one found within the OS of Mac. The mutations that were put together from the different tables (Sections 4.11, 4.12, 4.10) with R Section 4.19) were filtered by a Python script. To apply each mutation correctly on the proteins in the detailed method a script was written in which files were generated that described in a compact format on which chains and position a mutation resided.
Laptop version 2.7.15
Server version 2.7.11

4.19 R scripting language

With R the tables from GenomeAD, GAVIN and Infevers (Sections **??** 4.10 4.12) of TNFRSF1A missense mutations (Section 3.3) were merged together in a new comma seperated values file with their known classifications. Ordering and filtering the double mutations and removing double classifcations where done with Python (Section 4.18).
R scripting front-end version 3.5.2 (2018-12-20)

4.20 SLURM

For computational jobs where a laptop or desktop does not suffice because due to the lack computational resources a computer cluster could come to aid. These clusters consist out of several computers that execute resource intensive tasks, to manage these systems for many clients and to use these clusters optimal a workload manager mlike simple Linux utility resource management (SLURM), is installed. Jobs are submitted that request resources for execution and are scheduled on the systems queue which is ordered based on priorities, resource requirement and time.

4.21 MPI

Some tools from the Rosetta software suite (Sections 4.3) have the ability to use multiple central processing unit (CPU) cores from a single computer or from multiple computers. With a message parsing interface (MPI) it is possible for software to communicate between CPU cores on the same and on different computers to exchange information about processes and therefor solving solutions faster.
OpenMPI/1.8.8-GNU-4.9.3-2.25

5 Results

5.1 Revving the VIPUR approach to expand rare disease diagnostics

5.1.1 Preparatory steps for using the VIPUR approach

After the publication of VIPUR the tools, data and applications became available at the open science framework (OSF) [] which were downloaded and reviewed. All applications from the Rosetta software suite (Section 4.3) were pre-compiled without support for MPI (Section 4.21) and with that not the ability to benefit from multiple CPUs. The Rosetta software suite was rebuilt with MPI support in a slurm job where the compilation could benefit from multiple CPU cores.

5.1.2 VIPUR resolving system incompatibilities

Within the VIPUR pipeline residues were mutated to determine the effects of a structural mutation, by default missense mutations were inserted with PyMOL (Section 4.15), an alternative method integrated within the pipeline for situations wherein PyMOL was not accessible Pyrosetta (Section 4.4) could be used. Neither of these programs could be built or compiled because the lack of Open graphics library (OpenGL) for PyMOL and having the incorrect C++ and C libraries for PyRosetta. To bypass both programs and still be able to introduce mutations into PDB files Modeller (Secton 4.9) was introduced and built.

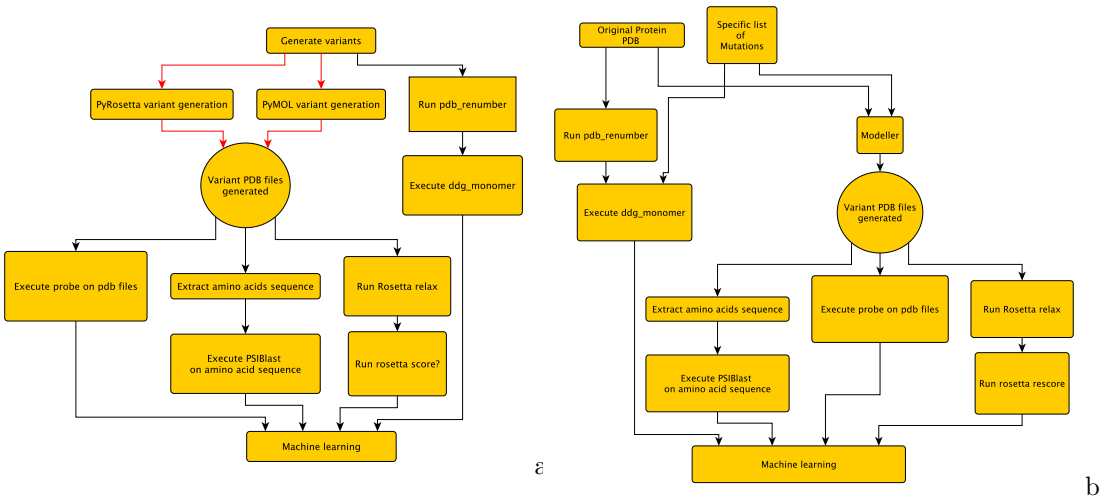
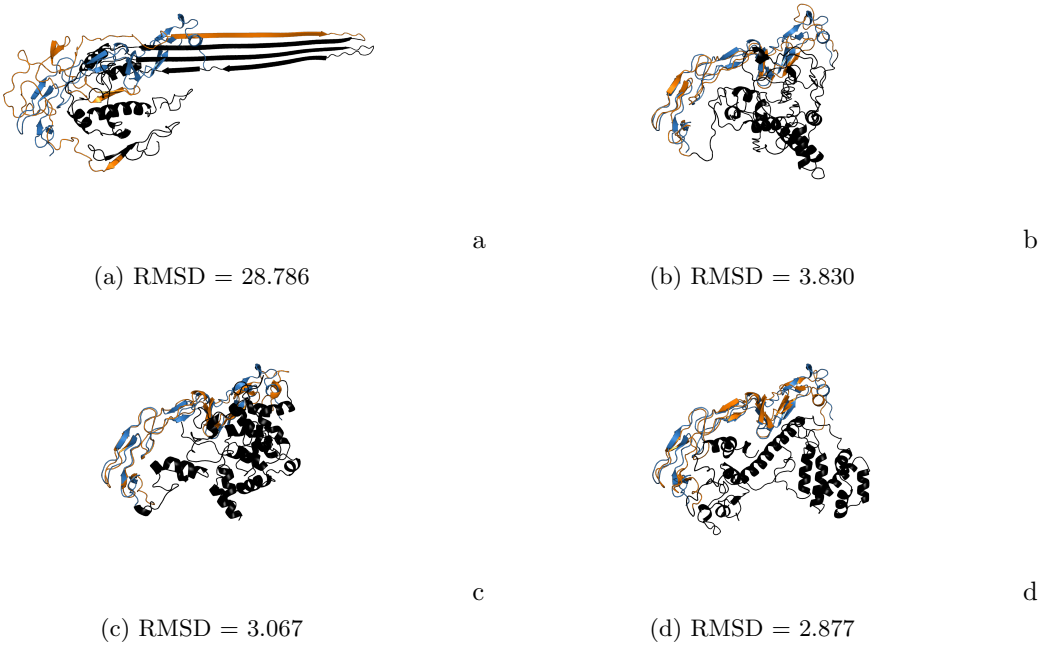


Figure 1: Both flowcharts illustrate the VIPUR pipeline wherein each block is a procedure the central circle is the purpose of the mutated applications and each arrow represents the path to it. Figure 1a has red arrows that indicate that both methods were incapable to produce the mutated PDB files. Within figure 1b the alternative method is proposed wherein PyMOL and PyRosetta (Sections 4.15, 4.4) is substituted by Modeller (Section 4.9) to acquire the mutated protein structures.

5.1.3 Expanding the VIPUR training set with data from TNFRSF1A by homology modeling and protein threading

Since the VTS did not have any features of TNFRSF1A (Section 3.2) the amino acid sequence was collected from Uniprot (Section 4.14) and the protein from RCSB (Section 4.13). The structures available of TNFRSF1A were incomplete, fragments for the TNF α and β binding site [] were available and its death domain that interacts with TRADD [] which plays a role in apoptosis (Section 3.2). To acquire a monomeric structure of TNFRSF1A two ab initio modeling web services I-TASSER and Robetta (Sections 4.8, 4.7) had been employed. Both were given the task to model the whole protein with and without a template to determine how well they could model a known structure and what it would form. Determination of which the best model was is based on the shortest distance ,defined root mean square deviation (RMSD), between a produced model compared to the X-ray crystallographic model of the TNFRSF1A binding site .



a

(a) RMSD = 28.786

b

(b) RMSD = 3.830

c

(c) RMSD = 3.067

d

(d) RMSD = 2.877

Figure 2: 3D structures of TNFRSF1A (2a, 2b: I-TASSER, 2c, 2d: Robetta) without (left: 2a, 2c) and with templates (right: 2b, 2d). The sky blue colored structure in each figure is an X-ray crystallographic model (1EXT) of the binding site of TNFRSF1A and the orange structure is the representation of that identical fragment in the model made by the web services.

5.2 Analyses of proteins variants TNFRSF1A

5.2.1 Requirements for determining structural and binding effects of protein variants

Protein variants can be assessed from multiple perspectives and together they can form a holistic view on how a protein works and how mutations affect its workings. However adding perspectives to the protein assessment makes it complex and requires expertise to determine its validity and contribution, therefore the analysis has been limited to basic structural information and also make the assessment inline with the VIPUR methods.

Various proteins consist of multiple chains that can be identical or different depending on their function [] and should be taken into account when assessing protein variants since one residue might alter the binding between chains and might alter the proteins formation. Different molecules and atoms that do not make up a protein but play a role in a pathway and function (ligands and co-receptors) are able to affect a proteins shape and can behave differently when a residue is mutated.

A different aspect that can change with mutations is the alteration in motions between structures which can allow or disallow certain movements to occur and with inhibit processes.

5.2.2 Single protein variant analysis approach and its tools

Before introducing mutations into a protein structure it is helpful to know if a mutation has been observed to avoid allocating resources to something that does not occur. Therefore three tables with observed TNFRSF1A mutations (Sections 4.10, 4.11, 4.12) have been combined with an R script (Section 4.19) into a single table consisting of two columns. The first column (split into three columns 1) contains strings that describes the: original residue, position and where it mutates to, the second column describes whether a formed mutation is harmful, with most mutations the effects have not been identified yet.

Original residue	Position in the protein sequence	New residue	Classification
Cys	44	Tyr	PATHOGENIC
Thr	44	Pro	PATHOGENIC
Thr	44	Ser	PATHOGENIC

Table 1: The format wherein mutations were filtered from the GAVIN, GenomAD and Infevers tables (Sections 4.10, 4.11, 4.12), describe whether a structural mutation is harmful or not. For many mutations it is unknown and other classifications are available, to view the whole table visit the supplementary.

For assessing variants in TNFRSF1A a structural fragment was used that contained TNF β (1TNR) [] and was made homotrimeric with PyMOL (Section 4.15) which results in six chains that emulate a bound TNFRSF1A with TNF β. The first column of the mutation table did not contain sufficient information to apply mutations correctly and within the PDB different numbering is used than in the amino acid sequence. To bundle the information and make it usable for introducing mutations a Python script (Section 4.18) has been written that combines the mutation table, PDB chains and the correct position within the sequence into a type of table which has sufficient information to mutate structures.

Iteration number	Filename	Chain	Residue index in chain	New residue
34	1tnr3_TNFA	R	0	TYR
34	1tnr3_TNFA	T	0	TYR
34	1tnr3_TNFA	S	0	TYR
35	1tnr3_TNFA	R	0	PRO
35	1tnr3_TNFA	T	0	PRO
35	1tnr3_TNFA	S	0	PRO
36	1tnr3_TNFA	R	0	SER
36	1tnr3_TNFA	T	0	SER
36	1tnr3_TNFA	S	0	SER

Table 2: The format that describes the mutations that should be made by Modeller (Section 4.9), with specifications of the: model, file. chain, residue index and the new residue. The whole table for TNFA and TNFB are visible within the supplementary.

To introduce mutations within PDB structures a Python script (Section 4.18) was written which used the generated mutation table (Table: 2) and a matching PDB structure, from the table; it acquires an iteration number which specifies if a mutation has to be stored in a single file or across multiple files; the filename serves as key which determine the PDB that should be used; letters specify chains, numbers are indices within the chains and the last column states the three letter residue where it should mutate to. When a structure is read in through the Python bindings of Modeller (Section 4.9) all non standard atoms and molecules are removed because Rosetta (Section 4.3) is not able to deal with those atoms. Just before mutagenesis takes place missing atoms are added to the structure that were difficult to determine with experimental methods(Section 1.3). After the insertion of all mutations a last guess is made where disulfide bonds are added between cysteines. With this process variants can be easily made, however the mutations do not present a correct protein because the mutated residues can put the protein in a high energy state.

In the attempt to make mutated structures behave more natural two tools from the Rosetta software suite (Section 4.3) have been used to minimize energies within protein structures. With the backrub application (Section 4.3.4) 1000 altered backbone models have been produced each with 10000 Monte Carlo moves (Sections 2.2), although it is known that the backbone changes are limited 1.2 replacing amino acids can make alterations in the backbone.

Relax **

We specifically chose to asses known mutations from infevers. **

5.3 Finding structural information and its mutations through HOPE

A final method to gain insight in protein structures and its mutations is with the webservice hope

6 Discussion

People with rare diseases are hard to diagnose

Prediction of pathogenicity in variants momentarily done based on sequence information and has been successful for certain groups of genes []. However pathogenticy of some genes with their variants cannot be classified by the currently used features for classification. Recently a method, called VIPUR, surfaced that incorporated sequence and protein data for classification of the pathogenicity from gene variants[].

In the attempt to reproduce the methods taken by the VIPUR approach on protein structures that are related to rare diseases it was realized that some questionable steps were being taken. With this approach all ligands were removed [] which changes the energies within and can therefor alter the structure [] and causes it to be analyzed from a single perspective instead of two when a bound ligand is also taken into account. Another step that was taken with VIPUR is that each structure is viewed as a monomer which is for some proteins not a problem, but for a complex that consists of multiple similar or a variety of different monomers makes it difficult to assess the effects.

To make predictions for new benign and pathogenic variants from TNFRSF1A, more information should be collected on how certain residues contribute to TNFRSF1A. More differences between interaction energies in mutated proteins could have been found be by adding molecular dynamics (MD) simulations of TNF α/β separately TNFRSF1A and combined with TNF docked into TNFRSF1A.

prediction of potential benign and pathogenic variants of TNFRSF1A isoforms should be included in the analysis to gain insight in which part of the proteins are highly important for the interactions and could result in a better prediction.

A significant contribution to gain more insight in how TNFRSF1A interacts with TNF α/β would have been the addition of molecular dynamic simulations; it shows how the proteins move on their own but also how the residues of the protein and the ligand interact with each other.

The VIPUR pipeline could not be executed because it was not possible to compile PyRosetta or PyMOL on the cluster.

however VIPUR has not been tested due to not having the correct software available and TNFRSF1A was not within the training data set of VIPUR.

Isoforms were not taken into account.

Within the publication of Probe is mentioned: "It requires both highly accurate structures and also the explicit inclusion of all hydrogen atoms and their van der Waals interactions." [].

The site of Probe mentions: "Meaningful analysis of molecular contact surfaces requires that ALL atoms are considered. Before using Probe, use the companion program Reduce to add hydrogens to the coordinate file." [], no mention of the Reduce software is mentioned in the VIPUR approach and therefor it is difficult to asses the meaning of previously acquired results in the VTS.

VIPUR is questionable because it has a limited amount of simulations. VIPUR uses PSI-blast to justify its results.

Describe the error made with apply mutations from different isoforms, see the table within results.

Only assessed a small piece of TNFRSF1A and did not even look at the class of proteins itself.

Good other suggestions for finding if the approach really means something is by using shap [].

The software written has a limited use currently and could be expanded to rapidly introduce mutations in multiple structures and chains at once.

It might have been useful to disable disulfide bridges if no cysteine residue is mutated because it is less likely that alterations are formed to disulfide bridges and otherwise they might be added without a reason.

Fragments were only available of 1EXT and 1TNR because it is a transmembrane protein which is difficult to make a structure from with X-ray crystallography. Maybe the first 30 residues were unnecessary because the might be signal peptides.

some steps have become questionable in structure of TNFRSF1A some questionable rare diseases some questionable training set some q to reproduce some of the results that were acquired with the VIPUR some questionable assu

Looking at the investigation t

With the resource at our disposal we were unable to reproduce any of the results that were produced by VIPUR for testing purposes, by

7 Conclusion

While VIPUR might be missing information to give a solid prediction about the pathogenicty of a protein variant, the detailed method used for determining the changes in energy levels could be a more reliable source for making predictions based of features.

of info that accurate we propsed another method for assessing protein structures within complex which may play a role in machine learning

Supplementary

Mutation	Classification
Ser4Phe	NA
Val6Met	NA
Pro7Thr	NA
Pro12Leu	NA
Glu14Lys	PATHOGENIC
Leu15Val	NA
Thr16Pro	PATHOGENIC
Thr16Ala	PATHOGENIC
Cys17Ser	PATHOGENIC
Val20Ala	Uncertain significance (VOUS)
Val21Ile	PATHOGENIC
Gly21Ala	NA
Val21Phe	PATHOGENIC
Val21Leu	PATHOGENIC
Arg24Trp	BENIGN
Gly26Ala	NA
Val27Asp	NA
Ile28Phe	NA
Arg28Trp	PATHOGENIC
Ala30Pro	PATHOGENIC
Ala30Gly	PATHOGENIC
Ala30Ser	PATHOGENIC
Ala30Thr	PATHOGENIC
Val31Gly	NA
Val33Leu	PATHOGENIC
Val33Met	PATHOGENIC
Arg34Thr	PATHOGENIC
Gly35Arg	NA
Ser38Pro	PATHOGENIC
Pro39Thr	PATHOGENIC
Asp41His	NA
Asp41Glu	NA
Thr43Ser	PATHOGENIC
Cys44Tyr	PATHOGENIC
Thr44Pro	PATHOGENIC
Thr44Ser	PATHOGENIC
Ser45Pro	PATHOGENIC
Ala48Thr	PATHOGENIC
Tyr49His	PATHOGENIC
Tyr49Cys	PATHOGENIC
Tyr49Asp	PATHOGENIC
Ala50Thr	POPULATION
Ile50Val	NA
His51Arg	PATHOGENIC
His51Tyr	PATHOGENIC
Ala51Gly	PATHOGENIC
Ala51Pro	PATHOGENIC
His51Gln	Likely pathogenic
Pro52Ala	PATHOGENIC
Asn54Asp	Uncertain significance (VOUS)
Ala54Thr	PATHOGENIC
Glu55Lys	PATHOGENIC
Ser56Thr	NA
Ser56Leu	NA
Ile57Ser	Uncertain significance (VOUS)
Cys58Ser	PATHOGENIC
Cys58Gly	PATHOGENIC
Cys58Phe	PATHOGENIC
Cys58Trp	Likely pathogenic
Cys58Tyr	PATHOGENIC
Cys58Arg	PATHOGENIC
Cys59Tyr	PATHOGENIC
Cys59Ser	NA
Cys59Arg	PATHOGENIC
Cys59Phe	PATHOGENIC
Cys62Tyr	PATHOGENIC
Cys62Gly	PATHOGENIC
Gly65Glu	PATHOGENIC
Pro66Arg	POPULATION
Ala66Thr	PATHOGENIC
Pro66Leu	BENIGN
Thr66Ile	Likely pathogenic
Tyr67Ser	PATHOGENIC
Tyr67Cys	PATHOGENIC
Pro68Leu	POPULATION
Leu68Phe	PATHOGENIC
Leu70Ser	POPULATION
Gly70Arg	PATHOGENIC
Thr70Met	PATHOGENIC
Asp71Glu	Likely pathogenic
Ser72Pro	PATHOGENIC
Cys72Gly	Pathogenic
Glu72Gly	POPULATION
Cys72Tyr	PATHOGENIC
Cys72Phe	PATHOGENIC
Cys72Arg	PATHOGENIC
Cys72Trp	PATHOGENIC
Cys72Ser	PATHOGENIC
Pro72Ser	POPULATION
Pro73Ser	NA
Cys73Arg	POPULATION
Cys74Arg	PATHOGENIC
Val74Gly	PATHOGENIC
Val74Met	PATHOGENIC
Val74Leu	PATHOGENIC
Cys74Tyr	PATHOGENIC
Cys75Arg	PATHOGENIC
Cys75Tyr	PATHOGENIC

Cys75Ser	PATHOGENIC
Thr75Ala	PATHOGENIC
Pro75Leu	NA
Pro75Arg	NA
Ala75Thr	POPULATION
Ala76Thr	PATHOGENIC
Glu76Asp	PATHOGENIC
Arg78Pro	PATHOGENIC
Thr79Met	PATHOGENIC
Ser79Pro	PATHOGENIC
Thr79Lys	Likely pathogenic
Phe80Ser	PATHOGENIC
Phe80Leu	PATHOGENIC
Thr80Ile	POPULATION
Phe80Val	PATHOGENIC
Cys81Ser	PATHOGENIC
Thr81Asn	PATHOGENIC
Val81Ala	POPULATION
Cys81Arg	PATHOGENIC
Cys81Trp	PATHOGENIC
Cys81Tyr	PATHOGENIC
Cys81Phe	PATHOGENIC
Arg82Gly	Uncertain significance (VOUS)
Cys82Trp	PATHOGENIC
Cys82Phe	PATHOGENIC
Cys82Tyr	PATHOGENIC
Gly83Asp	POPULATION
Cys84Arg	PATHOGENIC
Cys84Tyr	PATHOGENIC
Cys84Ser	PATHOGENIC
Glu85Asp	NA
Asn85Lys	PATHOGENIC
Asn85Ile	PATHOGENIC
His86Pro	PATHOGENIC
His86Tyr	PATHOGENIC
His86Leu	PATHOGENIC
Leu87Pro	PATHOGENIC
Leu87Phe	POPULATION
Lys87Glu	POPULATION
Gly87Ser	NA
Gln88Glu	POPULATION
Ser88Pro	PATHOGENIC
Phe89Ser	PATHOGENIC
Phe89Leu	PATHOGENIC
Phe89Val	PATHOGENIC
Cys90Ser	PATHOGENIC
Cys90Gly	PATHOGENIC
Cys90Tyr	PATHOGENIC
Ser90Ala	PATHOGENIC
Thr90Asn	PATHOGENIC
Thr90Ile	NA
Thr90Pro	NA
Arg90Trp	POPULATION
Ser90Pro	PATHOGENIC
Cys90Arg	PATHOGENIC
Tyr92His	POPULATION
Ser92Asn	POPULATION
Cys93Trp	PATHOGENIC
Cys93Ser	PATHOGENIC
Cys93Arg	PATHOGENIC
Asn94Lys	PATHOGENIC
Ser94Cys	PATHOGENIC
Ser94Gly	POPULATION
Asn94Ile	PATHOGENIC
His95Pro	PATHOGENIC
His95Leu	PATHOGENIC
Ala95Thr	POPULATION
Glu95Gly	POPULATION
His95Tyr	NA
Leu96Phe	POPULATION
Leu96Pro	PATHOGENIC
Cys96Tyr	PATHOGENIC
Arg97Gln	PATHOGENIC
Phe98Leu	PATHOGENIC
Phe98Cys	PATHOGENIC
His98Asn	NA
Phe98Ile	PATHOGENIC
Phe98Ser	PATHOGENIC
Cys99Arg	PATHOGENIC
Cys99Ser	PATHOGENIC
Cys99Tyr	PATHOGENIC
Cys99Gly	PATHOGENIC
Cys100Arg	PATHOGENIC
Ser101Asn	POPULATION
Pro102Ser	POPULATION
Asn102Asp	POPULATION
Cys102Arg	PATHOGENIC
Cys102Ser	PATHOGENIC
Cys102Trp	PATHOGENIC
Cys103Tyr	POPULATION
Ser103Cys	PATHOGENIC
Cys105Tyr	PATHOGENIC
Arg106Gln	PATHOGENIC
Ser106Pro	PATHOGENIC
Leu107Phe	POPULATION
Cys108Arg	PATHOGENIC
Cys108Tyr	PATHOGENIC
Thr109Ala	PATHOGENIC
Thr110Ile	POPULATION
Val111Leu	POPULATION
Gln111Lys	Uncertain significance (VOUS)
Arg112Pro	PATHOGENIC
Val112Leu	Likely pathogenic
His112Tyr	PATHOGENIC

Val112Met	NA
Leu113Phe	POPULATION
Thr114Ile	POPULATION
Ile114Ser	NA
Val115Ala	POPULATION
Ser115Pro	PATHOGENIC
Ser115Phe	NA
Cys116Tyr	PATHOGENIC
Cys116Phe	PATHOGENIC
Cys116Trp	PATHOGENIC
Cys117Tyr	PATHOGENIC
Cys117Arg	PATHOGENIC
Gly117Asp	POPULATION
Cys117Ser	Likely pathogenic
Cys118Tyr	PATHOGENIC
Thr118Ala	PATHOGENIC
Met118Val	POPULATION
Cys118Arg	PATHOGENIC
Val119Gly	NA
Val119Ala	NA
Arg121Trp	Likely pathogenic
Ser121Cys	POPULATION
Arg121Pro	NA
Arg121Gln	NA
Arg121Gly	NA
Asp122His	Likely pathogenic
Asp122Glu	Uncertain significance (VOUS)
Thr123Ile	NA
Val124Gly	NA
Val124Ala	NA
Val124Met	NA
Cys125Arg	Likely pathogenic
Cys125Trp	PATHOGENIC
Cys125Tyr	PATHOGENIC
Cys125Phe	PATHOGENIC
Gly126Asp	POPULATION
His126Arg	POPULATION
His126Tyr	POPULATION
Cys127Tyr	PATHOGENIC
Cys127Arg	PATHOGENIC
Phe129Leu	POPULATION
Asn130Lys	Likely pathogenic
Gln131Glu	NA
Arg133Trp	NA
Arg133Gln	NA
His134Pro	Uncertain significance (VOUS)
Tyr135His	NA
Glu135Val	POPULATION
Tyr135Cys	Likely pathogenic
Ser137Gly	NA
Glu138Gly	NA
Glu138Ala	Likely benign
Ser140Thr	POPULATION
Phe141Leu	PATHOGENIC
Phe141Ser	PATHOGENIC
Phe141Cys	PATHOGENIC
Phe141Ile	PATHOGENIC
Cys142Phe	POPULATION
Cys143Trp	Likely pathogenic
Cys143Arg	PATHOGENIC
Ser145Asn	POPULATION
Asn145Asp	NA
Asn145Ser	NA
Cys146Tyr	NA
Leu150Phe	POPULATION
Asn151Ser	NA
Gly152Ala	NA
Thr153Ile	POPULATION
Leu153Val	POPULATION
Val154Met	NA
Val154Leu	NA
His155Tyr	PATHOGENIC
Ile156Asn	PATHOGENIC
Leu156Phe	POPULATION
Val159Asp	PATHOGENIC
Arg160His	NA
Arg160Leu	NA
Arg160Ser	NA
Arg160Cys	NA
Lys161Arg	NA
Ser162Cys	NA
Ser162Ala	NA
Pro163Ser	NA
Asp164Asn	POPULATION
Glu164Lys	NA
Pro167Leu	NA
Pro167Ala	NA
Ser168Ala	NA
Pro169Thr	NA
His169Arg	NA
His169Tyr	NA
Pro169His	NA
His170Tyr	NA
Pro171Arg	NA
Phe172Leu	NA
Pro177Ser	NA
Glu178Val	NA
Glu178Lys	NA
Ala179Thr	NA
Ala179Ser	NA
Ala179Glu	NA
Ala179Val	NA
Leu182Val	NA
Ser183Thr	NA

Leu185Arg	NA
Cys185Phe	POPULATION
Pro186Ser	NA
Phe186Leu	BENIGN
Pro187His	NA
Pro187Leu	NA
Pro187Ala	NA
Ser188Asn	NA
Arg189Cys	NA
Glu190Lys	NA
Phe190Leu	NA
Phe190Val	NA
Gly191Asp	NA
Thr192Met	NA
Leu196Val	NA
Leu198Val	NA
Ile199Asn	PATHOGENIC
Ile199Thr	NA
Arg201Ser	NA
Arg201Cys	NA
Arg201Gly	NA
Arg201His	NA
Val202Asp	PATHOGENIC
His203Asn	NA
Thr205Asn	NA
Glu206Asp	NA
Asp207Asn	NA
Arg207Cys	NA
Arg207His	NA
His208Leu	NA
Phe209Tyr	NA
Gly214Arg	NA
Pro215Thr	NA
Trp215Cys	NA
Gly216Glu	NA
Pro217Leu	NA
Cys217Tyr	NA
Arg218Lys	NA
Arg219His	NA
Arg219Leu	NA
Arg219Cys	NA
Phe220Ser	NA
Phe220Leu	NA
Leu222Phe	NA
Phe222Ser	NA
Ser226Cys	NA
Ala226Thr	NA
Phe229Leu	NA
Ile230Val	NA
Gly231Val	NA
Met233Leu	NA
Arg235His	NA
Gln237Arg	NA
Arg238Gln	NA
Ser241Phe	NA
His242Arg	POPULATION
His242Tyr	POPULATION
Phe245Leu	POPULATION
Glu248Lys	NA
Glu248Asp	NA
Gly250Arg	NA
Leu251Phe	NA
Glu251Val	POPULATION
Pro253Thr	NA
Pro253Ala	NA
Glu253Lys	NA
Glu254Ala	NA
Glu255Gln	NA
Lys255Glu	NA
Gly256Glu	NA
Gly256Val	NA
Gly257Ala	NA
Gly257Arg	NA
Ala257Pro	NA
Ala257Val	NA
Gly259Arg	NA
Gly259Glu	NA
Leu259Ile	NA
Leu259Arg	NA
Gly260Val	NA
Thr262Ala	NA
Pro266Thr	NA
Pro266Ala	NA
Leu267Val	NA
Arg268Gln	NA
Trp269Arg	NA
Arg269Lys	BENIGN
Asn270Lys	NA
Asn270Asp	NA
Pro271Ser	NA
Pro271Ala	NA
Ser272Gly	NA
Phe273Leu	NA
Ser274Gly	NA
Gly274Arg	NA
Pro275Ser	NA
Gly277Glu	BENIGN
Thr280Ser	NA
Thr280Asn	NA
Pro281Arg	NA
Ser282Leu	NA
Arg284His	NA
Arg284Cys	NA
Arg284Leu	NA

Trp285Arg	NA
Phe285Leu	NA
Ser286Ile	NA
Pro287Ala	NA
Val288Leu	NA
Trp288Ser	NA
Val288Met	NA
Pro289Leu	NA
Ala289Thr	NA
Ser290Arg	NA
Phe293Leu	NA
Pro293Arg	NA
Trp295Cys	NA
Ser296Thr	NA
Arg298Gly	NA
Thr298Ala	NA
Pro301Ser	NA
Pro301His	NA
Gly302Ser	NA
Cys304Ser	NA
Pro305Arg	NA
Asn306Lys	NA
Pro310Leu	NA
Pro310Ser	NA
Arg311Cys	NA
Arg311His	NA
Arg312Lys	NA
Glu313Lys	NA
Glu313Gln	NA
Ala315Thr	NA
Pro317Ala	NA
Tyr318Cys	NA
Tyr318Phe	NA
Gly320Arg	NA
Gly320Glu	BENIGN
Asp322Glu	NA
Pro323Ser	NA
Ile324Val	NA
Ile324Asn	NA
Ala326Ser	NA
Leu329Phe	NA
Ala330Thr	NA
Ala330Val	NA
Pro333Ser	NA
Ile334Val	NA
Pro335His	NA
Pro335Arg	NA
Asn336Lys	NA
Asn336Asp	NA
Leu338Phe	NA
Lys340Arg	NA
Lys340Glu	NA
Glu342Asp	NA
His346Arg	NA
His346Gln	NA
Thr353Ala	NA
Thr353Pro	NA
Pro356Ser	NA
Pro356Ala	NA
Thr358Ala	NA
Glu364Val	NA
Pro368Ser	NA
Pro368Ala	NA
Ser381Asn	NA
Asp382His	NA
His383Asn	NA
Asp386Glu	NA
Leu388Arg	NA
Leu390Pro	NA
Arg394His	NA
Cys395Arg	NA
Cys395Trp	NA
Leu396Arg	NA
Glu398Lys	NA
Ser402Gly	NA
Ala405Gly	NA
Arg408Lys	NA
Arg409Trp	NA
Arg409Leu	NA
Pro412Leu	NA
Arg413Gly	NA
Arg413Gln	NA
Glu415Lys	NA
Ala416Gly	NA
Thr417Ala	NA
Glu419Gly	NA
Gly422Arg	NA
Arg423His	NA
Arg426His	NA
Gly432Val	NA
Glu435Gly	NA
Asp436Gly	NA
Glu438Lys	NA
Glu438Ala	NA
Glu438Gln	NA
Cys442Tyr	NA
Gly443Arg	NA
Gly443Val	NA
Pro448Leu	NA
Ser452Gly	NA
Leu453Pro	NA

Iteration number	Filename	Chain	Residue index in chain	New residue
34	1tnr3_TNFA	R	0	TYR
34	1tnr3_TNFA	T	0	TYR
34	1tnr3_TNFA	S	0	TYR
35	1tnr3_TNFA	R	0	PRO
35	1tnr3_TNFA	T	0	PRO
35	1tnr3_TNFA	S	0	PRO
36	1tnr3_TNFA	R	0	SER
36	1tnr3_TNFA	T	0	SER
36	1tnr3_TNFA	S	0	SER
37	1tnr3_TNFA	R	1	PRO
37	1tnr3_TNFA	T	1	PRO
37	1tnr3_TNFA	S	1	PRO
38	1tnr3_TNFA	R	4	THR
38	1tnr3_TNFA	T	4	THR
38	1tnr3_TNFA	S	4	THR
39	1tnr3_TNFA	R	5	HIS
39	1tnr3_TNFA	T	5	HIS
39	1tnr3_TNFA	S	5	HIS
40	1tnr3_TNFA	R	5	CYS
40	1tnr3_TNFA	T	5	CYS
40	1tnr3_TNFA	S	5	CYS
41	1tnr3_TNFA	R	5	ASP
41	1tnr3_TNFA	T	5	ASP
41	1tnr3_TNFA	S	5	ASP
42	1tnr3_TNFA	R	6	THR
42	1tnr3_TNFA	T	6	THR
42	1tnr3_TNFA	S	6	THR
43	1tnr3_TNFA	R	6	VAL
43	1tnr3_TNFA	T	6	VAL
43	1tnr3_TNFA	S	6	VAL
44	1tnr3_TNFA	R	7	ARG
44	1tnr3_TNFA	T	7	ARG
44	1tnr3_TNFA	S	7	ARG
45	1tnr3_TNFA	R	7	TYR
45	1tnr3_TNFA	T	7	TYR
45	1tnr3_TNFA	S	7	TYR
46	1tnr3_TNFA	R	7	GLY
46	1tnr3_TNFA	T	7	GLY
46	1tnr3_TNFA	S	7	GLY
47	1tnr3_TNFA	R	7	PRO
47	1tnr3_TNFA	T	7	PRO
47	1tnr3_TNFA	S	7	PRO
48	1tnr3_TNFA	R	7	GLN
48	1tnr3_TNFA	T	7	GLN
48	1tnr3_TNFA	S	7	GLN
49	1tnr3_TNFA	R	8	ALA
49	1tnr3_TNFA	T	8	ALA
49	1tnr3_TNFA	S	8	ALA
50	1tnr3_TNFA	R	10	ASP
50	1tnr3_TNFA	T	10	ASP
50	1tnr3_TNFA	S	10	ASP
51	1tnr3_TNFA	R	10	THR
51	1tnr3_TNFA	T	10	THR
51	1tnr3_TNFA	S	10	THR
52	1tnr3_TNFA	R	11	LYS
52	1tnr3_TNFA	T	11	LYS
52	1tnr3_TNFA	S	11	LYS
53	1tnr3_TNFA	R	12	THR
53	1tnr3_TNFA	T	12	THR
53	1tnr3_TNFA	S	12	THR
54	1tnr3_TNFA	R	12	LEU
54	1tnr3_TNFA	T	12	LEU
54	1tnr3_TNFA	S	12	LEU
55	1tnr3_TNFA	R	13	SER
55	1tnr3_TNFA	T	13	SER
55	1tnr3_TNFA	S	13	SER
56	1tnr3_TNFA	R	14	SER
56	1tnr3_TNFA	T	14	SER
56	1tnr3_TNFA	S	14	SER
57	1tnr3_TNFA	R	14	GLY
57	1tnr3_TNFA	T	14	GLY
57	1tnr3_TNFA	S	14	GLY
58	1tnr3_TNFA	R	14	PHE
58	1tnr3_TNFA	T	14	PHE
58	1tnr3_TNFA	S	14	PHE
59	1tnr3_TNFA	R	14	TRP
59	1tnr3_TNFA	T	14	TRP
59	1tnr3_TNFA	S	14	TRP
60	1tnr3_TNFA	R	14	TYR
60	1tnr3_TNFA	T	14	TYR
60	1tnr3_TNFA	S	14	TYR
61	1tnr3_TNFA	R	14	ARG
61	1tnr3_TNFA	T	14	ARG
61	1tnr3_TNFA	S	14	ARG
62	1tnr3_TNFA	R	15	TYR
62	1tnr3_TNFA	T	15	TYR
62	1tnr3_TNFA	S	15	TYR
63	1tnr3_TNFA	R	15	SER
63	1tnr3_TNFA	T	15	SER
63	1tnr3_TNFA	S	15	SER
64	1tnr3_TNFA	R	15	ARG
64	1tnr3_TNFA	T	15	ARG
64	1tnr3_TNFA	S	15	ARG
65	1tnr3_TNFA	R	15	PHE
65	1tnr3_TNFA	T	15	PHE
65	1tnr3_TNFA	S	15	PHE
66	1tnr3_TNFA	R	18	TYR
66	1tnr3_TNFA	T	18	TYR
66	1tnr3_TNFA	S	18	TYR
67	1tnr3_TNFA	R	18	GLY
67	1tnr3_TNFA	T	18	GLY
67	1tnr3_TNFA	S	18	GLY

68	1tnr3_TNFA	R	21	GLU
68	1tnr3_TNFA	T	21	GLU
68	1tnr3_TNFA	S	21	GLU
69	1tnr3_TNFA	R	22	ARG
69	1tnr3_TNFA	T	22	ARG
69	1tnr3_TNFA	S	22	ARG
70	1tnr3_TNFA	R	22	THR
70	1tnr3_TNFA	T	22	THR
70	1tnr3_TNFA	S	22	THR
71	1tnr3_TNFA	R	22	LEU
71	1tnr3_TNFA	T	22	LEU
71	1tnr3_TNFA	S	22	LEU
72	1tnr3_TNFA	R	22	ILE
72	1tnr3_TNFA	T	22	ILE
72	1tnr3_TNFA	S	22	ILE
73	1tnr3_TNFA	R	23	SER
73	1tnr3_TNFA	T	23	SER
73	1tnr3_TNFA	S	23	SER
74	1tnr3_TNFA	R	23	CYS
74	1tnr3_TNFA	T	23	CYS
74	1tnr3_TNFA	S	23	CYS
75	1tnr3_TNFA	R	24	LEU
75	1tnr3_TNFA	T	24	LEU
75	1tnr3_TNFA	S	24	LEU
76	1tnr3_TNFA	R	24	PHE
76	1tnr3_TNFA	T	24	PHE
76	1tnr3_TNFA	S	24	PHE
77	1tnr3_TNFA	R	26	SER
77	1tnr3_TNFA	T	26	SER
77	1tnr3_TNFA	S	26	SER
78	1tnr3_TNFA	R	26	ARG
78	1tnr3_TNFA	T	26	ARG
78	1tnr3_TNFA	S	26	ARG
79	1tnr3_TNFA	R	26	MET
79	1tnr3_TNFA	T	26	MET
79	1tnr3_TNFA	S	26	MET
80	1tnr3_TNFA	R	27	GLU
80	1tnr3_TNFA	T	27	GLU
80	1tnr3_TNFA	S	27	GLU
81	1tnr3_TNFA	R	28	PRO
81	1tnr3_TNFA	T	28	PRO
81	1tnr3_TNFA	S	28	PRO
82	1tnr3_TNFA	R	28	GLY
82	1tnr3_TNFA	T	28	GLY
82	1tnr3_TNFA	S	28	GLY
83	1tnr3_TNFA	R	28	GLY
83	1tnr3_TNFA	T	28	GLY
83	1tnr3_TNFA	S	28	GLY
84	1tnr3_TNFA	R	28	TYR
84	1tnr3_TNFA	T	28	TYR
84	1tnr3_TNFA	S	28	TYR
85	1tnr3_TNFA	R	28	PHE
85	1tnr3_TNFA	T	28	PHE
85	1tnr3_TNFA	S	28	PHE
86	1tnr3_TNFA	R	28	ARG
86	1tnr3_TNFA	T	28	ARG
86	1tnr3_TNFA	S	28	ARG
87	1tnr3_TNFA	R	28	TRP
87	1tnr3_TNFA	T	28	TRP
87	1tnr3_TNFA	S	28	TRP
88	1tnr3_TNFA	R	28	SER
88	1tnr3_TNFA	T	28	SER
88	1tnr3_TNFA	S	28	SER
89	1tnr3_TNFA	R	28	SER
89	1tnr3_TNFA	T	28	SER
89	1tnr3_TNFA	S	28	SER
90	1tnr3_TNFA	R	29	SER
90	1tnr3_TNFA	T	29	SER
90	1tnr3_TNFA	S	29	SER
91	1tnr3_TNFA	R	29	ARG
91	1tnr3_TNFA	T	29	ARG
91	1tnr3_TNFA	S	29	ARG
92	1tnr3_TNFA	R	30	ARG
92	1tnr3_TNFA	T	30	ARG
92	1tnr3_TNFA	S	30	ARG
93	1tnr3_TNFA	R	30	GLY
93	1tnr3_TNFA	T	30	GLY
93	1tnr3_TNFA	S	30	GLY
94	1tnr3_TNFA	R	30	MET
94	1tnr3_TNFA	T	30	MET
94	1tnr3_TNFA	S	30	MET
95	1tnr3_TNFA	R	30	LEU
95	1tnr3_TNFA	T	30	LEU
95	1tnr3_TNFA	S	30	LEU
96	1tnr3_TNFA	R	30	TYR
96	1tnr3_TNFA	T	30	TYR
96	1tnr3_TNFA	S	30	TYR
97	1tnr3_TNFA	R	31	ARG
97	1tnr3_TNFA	T	31	ARG
97	1tnr3_TNFA	S	31	ARG
98	1tnr3_TNFA	R	31	TYR
98	1tnr3_TNFA	T	31	TYR
98	1tnr3_TNFA	S	31	TYR
99	1tnr3_TNFA	R	31	SER
99	1tnr3_TNFA	T	31	SER
99	1tnr3_TNFA	S	31	SER
100	1tnr3_TNFA	R	31	ALA
100	1tnr3_TNFA	T	31	ALA
100	1tnr3_TNFA	S	31	ALA
101	1tnr3_TNFA	R	31	LEU
101	1tnr3_TNFA	T	31	LEU
101	1tnr3_TNFA	S	31	LEU
102	1tnr3_TNFA	R	31	ARG
102	1tnr3_TNFA	T	31	ARG

102	1tnr3_TNFA	S	31	ARG
103	1tnr3_TNFA	R	31	THR
103	1tnr3_TNFA	T	31	THR
103	1tnr3_TNFA	S	31	THR
104	1tnr3_TNFA	R	32	THR
104	1tnr3_TNFA	T	32	THR
104	1tnr3_TNFA	S	32	THR
105	1tnr3_TNFA	R	32	ASP
105	1tnr3_TNFA	T	32	ASP
105	1tnr3_TNFA	S	32	ASP
106	1tnr3_TNFA	R	34	PRO
106	1tnr3_TNFA	T	34	PRO
106	1tnr3_TNFA	S	34	PRO
107	1tnr3_TNFA	R	35	MET
107	1tnr3_TNFA	T	35	MET
107	1tnr3_TNFA	S	35	MET
108	1tnr3_TNFA	R	35	PRO
108	1tnr3_TNFA	T	35	PRO
108	1tnr3_TNFA	S	35	PRO
109	1tnr3_TNFA	R	35	LYS
109	1tnr3_TNFA	T	35	LYS
109	1tnr3_TNFA	S	35	LYS
110	1tnr3_TNFA	R	36	SER
110	1tnr3_TNFA	T	36	SER
110	1tnr3_TNFA	S	36	SER
111	1tnr3_TNFA	R	36	LEU
111	1tnr3_TNFA	T	36	LEU
111	1tnr3_TNFA	S	36	LEU
112	1tnr3_TNFA	R	36	ILE
112	1tnr3_TNFA	T	36	ILE
112	1tnr3_TNFA	S	36	ILE
113	1tnr3_TNFA	R	36	VAL
113	1tnr3_TNFA	T	36	VAL
113	1tnr3_TNFA	S	36	VAL
114	1tnr3_TNFA	R	37	SER
114	1tnr3_TNFA	T	37	SER
114	1tnr3_TNFA	S	37	SER
115	1tnr3_TNFA	R	37	ASN
115	1tnr3_TNFA	T	37	ASN
115	1tnr3_TNFA	S	37	ASN
116	1tnr3_TNFA	R	37	ALA
116	1tnr3_TNFA	T	37	ALA
116	1tnr3_TNFA	S	37	ALA
117	1tnr3_TNFA	R	37	ARG
117	1tnr3_TNFA	T	37	ARG
117	1tnr3_TNFA	S	37	ARG
118	1tnr3_TNFA	R	37	TRP
118	1tnr3_TNFA	T	37	TRP
118	1tnr3_TNFA	S	37	TRP
119	1tnr3_TNFA	R	37	TYR
119	1tnr3_TNFA	T	37	TYR
119	1tnr3_TNFA	S	37	TYR
120	1tnr3_TNFA	R	37	PHE
120	1tnr3_TNFA	T	37	PHE
120	1tnr3_TNFA	S	37	PHE
121	1tnr3_TNFA	R	38	GLY
121	1tnr3_TNFA	T	38	GLY
121	1tnr3_TNFA	S	38	GLY
122	1tnr3_TNFA	R	38	TRP
122	1tnr3_TNFA	T	38	TRP
122	1tnr3_TNFA	S	38	TRP
123	1tnr3_TNFA	R	38	PHE
123	1tnr3_TNFA	T	38	PHE
123	1tnr3_TNFA	S	38	PHE
124	1tnr3_TNFA	R	38	TYR
124	1tnr3_TNFA	T	38	TYR
124	1tnr3_TNFA	S	38	TYR
125	1tnr3_TNFA	R	39	ASP
125	1tnr3_TNFA	T	39	ASP
125	1tnr3_TNFA	S	39	ASP
126	1tnr3_TNFA	R	40	ARG
126	1tnr3_TNFA	T	40	ARG
126	1tnr3_TNFA	S	40	ARG
127	1tnr3_TNFA	R	40	TYR
127	1tnr3_TNFA	T	40	TYR
127	1tnr3_TNFA	S	40	TYR
128	1tnr3_TNFA	R	40	SER
128	1tnr3_TNFA	T	40	SER
128	1tnr3_TNFA	S	40	SER
129	1tnr3_TNFA	R	41	ASP
129	1tnr3_TNFA	T	41	ASP
129	1tnr3_TNFA	S	41	ASP
130	1tnr3_TNFA	R	41	LYS
130	1tnr3_TNFA	T	41	LYS
130	1tnr3_TNFA	S	41	LYS
131	1tnr3_TNFA	R	41	ILE
131	1tnr3_TNFA	T	41	ILE
131	1tnr3_TNFA	S	41	ILE
132	1tnr3_TNFA	R	42	PRO
132	1tnr3_TNFA	T	42	PRO
132	1tnr3_TNFA	S	42	PRO
133	1tnr3_TNFA	R	42	TYR
133	1tnr3_TNFA	T	42	TYR
133	1tnr3_TNFA	S	42	TYR
134	1tnr3_TNFA	R	42	LEU
134	1tnr3_TNFA	T	42	LEU
134	1tnr3_TNFA	S	42	LEU
135	1tnr3_TNFA	R	43	PRO
135	1tnr3_TNFA	T	43	PRO
135	1tnr3_TNFA	S	43	PRO
136	1tnr3_TNFA	R	43	PHE
136	1tnr3_TNFA	T	43	PHE
136	1tnr3_TNFA	S	43	PHE
137	1tnr3_TNFA	R	43	GLU

137	1tnr3_TNFA	T	43	GLU
137	1tnr3_TNFA	S	43	GLU
138	1tnr3_TNFA	R	43	SER
138	1tnr3_TNFA	T	43	SER
138	1tnr3_TNFA	S	43	SER
139	1tnr3_TNFA	R	44	GLU
139	1tnr3_TNFA	T	44	GLU
139	1tnr3_TNFA	S	44	GLU
140	1tnr3_TNFA	R	44	PRO
140	1tnr3_TNFA	T	44	PRO
140	1tnr3_TNFA	S	44	PRO
141	1tnr3_TNFA	R	45	SER
141	1tnr3_TNFA	T	45	SER
141	1tnr3_TNFA	S	45	SER
142	1tnr3_TNFA	R	45	LEU
142	1tnr3_TNFA	T	45	LEU
142	1tnr3_TNFA	S	45	LEU
143	1tnr3_TNFA	R	45	VAL
143	1tnr3_TNFA	T	45	VAL
143	1tnr3_TNFA	S	45	VAL
144	1tnr3_TNFA	R	46	SER
144	1tnr3_TNFA	T	46	SER
144	1tnr3_TNFA	S	46	SER
145	1tnr3_TNFA	R	46	GLY
145	1tnr3_TNFA	T	46	GLY
145	1tnr3_TNFA	S	46	GLY
146	1tnr3_TNFA	R	46	TYR
146	1tnr3_TNFA	T	46	TYR
146	1tnr3_TNFA	S	46	TYR
147	1tnr3_TNFA	R	46	ALA
147	1tnr3_TNFA	T	46	ALA
147	1tnr3_TNFA	S	46	ALA
148	1tnr3_TNFA	R	46	ASN
148	1tnr3_TNFA	T	46	ASN
148	1tnr3_TNFA	S	46	ASN
149	1tnr3_TNFA	R	46	ILE
149	1tnr3_TNFA	T	46	ILE
149	1tnr3_TNFA	S	46	ILE
150	1tnr3_TNFA	R	46	PRO
150	1tnr3_TNFA	T	46	PRO
150	1tnr3_TNFA	S	46	PRO
151	1tnr3_TNFA	R	46	TRP
151	1tnr3_TNFA	T	46	TRP
151	1tnr3_TNFA	S	46	TRP
152	1tnr3_TNFA	R	46	PRO
152	1tnr3_TNFA	T	46	PRO
152	1tnr3_TNFA	S	46	PRO
153	1tnr3_TNFA	R	46	ARG
153	1tnr3_TNFA	T	46	ARG
153	1tnr3_TNFA	S	46	ARG
154	1tnr3_TNFA	R	48	HIS
154	1tnr3_TNFA	T	48	HIS
154	1tnr3_TNFA	S	48	HIS
155	1tnr3_TNFA	R	48	ASN
155	1tnr3_TNFA	T	48	ASN
155	1tnr3_TNFA	S	48	ASN
156	1tnr3_TNFA	R	49	TRP
156	1tnr3_TNFA	T	49	TRP
156	1tnr3_TNFA	S	49	TRP
157	1tnr3_TNFA	R	49	SER
157	1tnr3_TNFA	T	49	SER
157	1tnr3_TNFA	S	49	SER
158	1tnr3_TNFA	R	49	ARG
158	1tnr3_TNFA	T	49	ARG
158	1tnr3_TNFA	S	49	ARG
159	1tnr3_TNFA	R	50	LYS
159	1tnr3_TNFA	T	50	LYS
159	1tnr3_TNFA	S	50	LYS
160	1tnr3_TNFA	R	50	CYS
160	1tnr3_TNFA	T	50	CYS
160	1tnr3_TNFA	S	50	CYS
161	1tnr3_TNFA	R	50	GLY
161	1tnr3_TNFA	T	50	GLY
161	1tnr3_TNFA	S	50	GLY
162	1tnr3_TNFA	R	50	ILE
162	1tnr3_TNFA	T	50	ILE
162	1tnr3_TNFA	S	50	ILE
163	1tnr3_TNFA	R	51	PRO
163	1tnr3_TNFA	T	51	PRO
163	1tnr3_TNFA	S	51	PRO
164	1tnr3_TNFA	R	51	LEU
164	1tnr3_TNFA	T	51	LEU
164	1tnr3_TNFA	S	51	LEU
165	1tnr3_TNFA	R	51	THR
165	1tnr3_TNFA	T	51	THR
165	1tnr3_TNFA	S	51	THR
166	1tnr3_TNFA	R	51	GLY
166	1tnr3_TNFA	T	51	GLY
166	1tnr3_TNFA	S	51	GLY
167	1tnr3_TNFA	R	51	TYR
167	1tnr3_TNFA	T	51	TYR
167	1tnr3_TNFA	S	51	TYR
168	1tnr3_TNFA	R	52	PHE
168	1tnr3_TNFA	T	52	PHE
168	1tnr3_TNFA	S	52	PHE
169	1tnr3_TNFA	R	52	PRO
169	1tnr3_TNFA	T	52	PRO
169	1tnr3_TNFA	S	52	PRO
170	1tnr3_TNFA	R	52	TYR
170	1tnr3_TNFA	T	52	TYR
170	1tnr3_TNFA	S	52	TYR
171	1tnr3_TNFA	R	53	GLN
171	1tnr3_TNFA	T	53	GLN
171	1tnr3_TNFA	S	53	GLN

172	1tnr3_TNFA	R	54	LEU
172	1tnr3_TNFA	T	54	LEU
172	1tnr3_TNFA	S	54	LEU
173	1tnr3_TNFA	R	54	CYS
173	1tnr3_TNFA	T	54	CYS
173	1tnr3_TNFA	S	54	CYS
174	1tnr3_TNFA	R	54	ASN
174	1tnr3_TNFA	T	54	ASN
174	1tnr3_TNFA	S	54	ASN
175	1tnr3_TNFA	R	54	ILE
175	1tnr3_TNFA	T	54	ILE
175	1tnr3_TNFA	S	54	ILE
176	1tnr3_TNFA	R	54	SER
176	1tnr3_TNFA	T	54	SER
176	1tnr3_TNFA	S	54	SER
177	1tnr3_TNFA	R	55	ARG
177	1tnr3_TNFA	T	55	ARG
177	1tnr3_TNFA	S	55	ARG
178	1tnr3_TNFA	R	55	SER
178	1tnr3_TNFA	T	55	SER
178	1tnr3_TNFA	S	55	SER
179	1tnr3_TNFA	R	55	TYR
179	1tnr3_TNFA	T	55	TYR
179	1tnr3_TNFA	S	55	TYR
180	1tnr3_TNFA	R	55	GLY
180	1tnr3_TNFA	T	55	GLY
180	1tnr3_TNFA	S	55	GLY
181	1tnr3_TNFA	R	56	ARG
181	1tnr3_TNFA	T	56	ARG
181	1tnr3_TNFA	S	56	ARG
182	1tnr3_TNFA	R	57	ASN
182	1tnr3_TNFA	T	57	ASN
182	1tnr3_TNFA	S	57	ASN
183	1tnr3_TNFA	R	58	SER
183	1tnr3_TNFA	T	58	SER
183	1tnr3_TNFA	S	58	SER
184	1tnr3_TNFA	R	58	ASP
184	1tnr3_TNFA	T	58	ASP
184	1tnr3_TNFA	S	58	ASP
185	1tnr3_TNFA	R	58	ARG
185	1tnr3_TNFA	T	58	ARG
185	1tnr3_TNFA	S	58	ARG
186	1tnr3_TNFA	R	58	SER
186	1tnr3_TNFA	T	58	SER
186	1tnr3_TNFA	S	58	SER
187	1tnr3_TNFA	R	58	TRP
187	1tnr3_TNFA	T	58	TRP
187	1tnr3_TNFA	S	58	TRP
188	1tnr3_TNFA	R	59	TYR
188	1tnr3_TNFA	T	59	TYR
188	1tnr3_TNFA	S	59	TYR
189	1tnr3_TNFA	R	59	CYS
189	1tnr3_TNFA	T	59	CYS
189	1tnr3_TNFA	S	59	CYS
190	1tnr3_TNFA	R	61	TYR
190	1tnr3_TNFA	T	61	TYR
190	1tnr3_TNFA	S	61	TYR
191	1tnr3_TNFA	R	62	GLN
191	1tnr3_TNFA	T	62	GLN
191	1tnr3_TNFA	S	62	GLN
192	1tnr3_TNFA	R	62	PRO
192	1tnr3_TNFA	T	62	PRO
192	1tnr3_TNFA	S	62	PRO
193	1tnr3_TNFA	R	63	PHE
193	1tnr3_TNFA	T	63	PHE
193	1tnr3_TNFA	S	63	PHE
194	1tnr3_TNFA	R	64	ARG
194	1tnr3_TNFA	T	64	ARG
194	1tnr3_TNFA	S	64	ARG
195	1tnr3_TNFA	R	64	TYR
195	1tnr3_TNFA	T	64	TYR
195	1tnr3_TNFA	S	64	TYR
196	1tnr3_TNFA	R	65	ALA
196	1tnr3_TNFA	T	65	ALA
196	1tnr3_TNFA	S	65	ALA
197	1tnr3_TNFA	R	66	ILE
197	1tnr3_TNFA	T	66	ILE
197	1tnr3_TNFA	S	66	ILE
198	1tnr3_TNFA	R	67	LEU
198	1tnr3_TNFA	T	67	LEU
198	1tnr3_TNFA	S	67	LEU
199	1tnr3_TNFA	R	67	LYS
199	1tnr3_TNFA	T	67	LYS
199	1tnr3_TNFA	S	67	LYS
200	1tnr3_TNFA	R	68	PRO
200	1tnr3_TNFA	T	68	PRO
200	1tnr3_TNFA	S	68	PRO
201	1tnr3_TNFA	R	68	LEU
201	1tnr3_TNFA	T	68	LEU
201	1tnr3_TNFA	S	68	LEU
202	1tnr3_TNFA	R	68	TYR
202	1tnr3_TNFA	T	68	TYR
202	1tnr3_TNFA	S	68	TYR
203	1tnr3_TNFA	R	68	MET
203	1tnr3_TNFA	T	68	MET
203	1tnr3_TNFA	S	68	MET
204	1tnr3_TNFA	R	69	PHE
204	1tnr3_TNFA	T	69	PHE
204	1tnr3_TNFA	S	69	PHE
205	1tnr3_TNFA	R	70	ILE
205	1tnr3_TNFA	T	70	ILE
205	1tnr3_TNFA	S	70	ILE
206	1tnr3_TNFA	R	70	SER
206	1tnr3_TNFA	T	70	SER

206	1tnr3_TNFA	S	70	SER
207	1tnr3_TNFA	R	71	ALA
207	1tnr3_TNFA	T	71	ALA
207	1tnr3_TNFA	S	71	ALA
208	1tnr3_TNFA	R	71	PRO
208	1tnr3_TNFA	T	71	PRO
208	1tnr3_TNFA	S	71	PRO
209	1tnr3_TNFA	R	71	PHE
209	1tnr3_TNFA	T	71	PHE
209	1tnr3_TNFA	S	71	PHE
210	1tnr3_TNFA	R	72	TYR
210	1tnr3_TNFA	T	72	TYR
210	1tnr3_TNFA	S	72	TYR
211	1tnr3_TNFA	R	72	PHE
211	1tnr3_TNFA	T	72	PHE
211	1tnr3_TNFA	S	72	PHE
212	1tnr3_TNFA	R	72	TRP
212	1tnr3_TNFA	T	72	TRP
212	1tnr3_TNFA	S	72	TRP
213	1tnr3_TNFA	R	73	TYR
213	1tnr3_TNFA	T	73	TYR
213	1tnr3_TNFA	S	73	TYR
214	1tnr3_TNFA	R	73	ARG
214	1tnr3_TNFA	T	73	ARG
214	1tnr3_TNFA	S	73	ARG
215	1tnr3_TNFA	R	73	ASP
215	1tnr3_TNFA	T	73	ASP
215	1tnr3_TNFA	S	73	ASP
216	1tnr3_TNFA	R	73	SER
216	1tnr3_TNFA	T	73	SER
216	1tnr3_TNFA	S	73	SER
217	1tnr3_TNFA	R	74	TYR
217	1tnr3_TNFA	T	74	TYR
217	1tnr3_TNFA	S	74	TYR
218	1tnr3_TNFA	R	74	ALA
218	1tnr3_TNFA	T	74	ALA
218	1tnr3_TNFA	S	74	ALA
219	1tnr3_TNFA	R	74	VAL
219	1tnr3_TNFA	T	74	VAL
219	1tnr3_TNFA	S	74	VAL
220	1tnr3_TNFA	R	74	ARG
220	1tnr3_TNFA	T	74	ARG
220	1tnr3_TNFA	S	74	ARG
221	1tnr3_TNFA	R	75	GLY
221	1tnr3_TNFA	T	75	GLY
221	1tnr3_TNFA	S	75	GLY
222	1tnr3_TNFA	R	75	ALA
222	1tnr3_TNFA	T	75	ALA
222	1tnr3_TNFA	S	75	ALA
223	1tnr3_TNFA	R	77	TRP
223	1tnr3_TNFA	T	77	TRP
223	1tnr3_TNFA	S	77	TRP
224	1tnr3_TNFA	R	77	CYS
224	1tnr3_TNFA	T	77	CYS
224	1tnr3_TNFA	S	77	CYS
225	1tnr3_TNFA	R	77	PRO
225	1tnr3_TNFA	T	77	PRO
225	1tnr3_TNFA	S	77	PRO
226	1tnr3_TNFA	R	77	GLN
226	1tnr3_TNFA	T	77	GLN
226	1tnr3_TNFA	S	77	GLN
227	1tnr3_TNFA	R	77	GLY
227	1tnr3_TNFA	T	77	GLY
227	1tnr3_TNFA	S	77	GLY
228	1tnr3_TNFA	R	78	HIS
228	1tnr3_TNFA	T	78	HIS
228	1tnr3_TNFA	S	78	HIS
229	1tnr3_TNFA	R	78	GLU
229	1tnr3_TNFA	T	78	GLU
229	1tnr3_TNFA	S	78	GLU
230	1tnr3_TNFA	R	79	ILE
230	1tnr3_TNFA	T	79	ILE
230	1tnr3_TNFA	S	79	ILE
231	1tnr3_TNFA	R	80	GLY
231	1tnr3_TNFA	T	80	GLY
231	1tnr3_TNFA	S	80	GLY
232	1tnr3_TNFA	R	80	ALA
232	1tnr3_TNFA	T	80	ALA
232	1tnr3_TNFA	S	80	ALA
233	1tnr3_TNFA	R	80	MET
233	1tnr3_TNFA	T	80	MET
233	1tnr3_TNFA	S	80	MET
234	1tnr3_TNFA	R	81	ARG
234	1tnr3_TNFA	T	81	ARG
234	1tnr3_TNFA	S	81	ARG
235	1tnr3_TNFA	R	81	TRP
235	1tnr3_TNFA	T	81	TRP
235	1tnr3_TNFA	S	81	TRP
236	1tnr3_TNFA	R	81	TYR
236	1tnr3_TNFA	T	81	TYR
236	1tnr3_TNFA	S	81	TYR
237	1tnr3_TNFA	R	81	PHE
237	1tnr3_TNFA	T	81	PHE
237	1tnr3_TNFA	S	81	PHE
238	1tnr3_TNFA	R	82	ASP
238	1tnr3_TNFA	T	82	ASP
238	1tnr3_TNFA	S	82	ASP
239	1tnr3_TNFA	R	82	ARG
239	1tnr3_TNFA	T	82	ARG
239	1tnr3_TNFA	S	82	ARG
240	1tnr3_TNFA	R	82	TYR
240	1tnr3_TNFA	T	82	TYR
240	1tnr3_TNFA	S	82	TYR
241	1tnr3_TNFA	R	83	TYR

241	1tnr3_TNFA	T	83	TYR
241	1tnr3_TNFA	S	83	TYR
242	1tnr3_TNFA	R	83	ARG
242	1tnr3_TNFA	T	83	ARG
242	1tnr3_TNFA	S	83	ARG
243	1tnr3_TNFA	R	85	LEU
243	1tnr3_TNFA	T	85	LEU
243	1tnr3_TNFA	S	85	LEU
244	1tnr3_TNFA	R	86	LYS
244	1tnr3_TNFA	T	86	LYS
244	1tnr3_TNFA	S	86	LYS
245	1tnr3_TNFA	R	87	GLU
245	1tnr3_TNFA	T	87	GLU
245	1tnr3_TNFA	S	87	GLU
246	1tnr3_TNFA	R	89	TRP
246	1tnr3_TNFA	T	89	TRP
246	1tnr3_TNFA	S	89	TRP
247	1tnr3_TNFA	R	89	GLN
247	1tnr3_TNFA	T	89	GLN
247	1tnr3_TNFA	S	89	GLN
248	1tnr3_TNFA	R	90	PRO
248	1tnr3_TNFA	T	90	PRO
248	1tnr3_TNFA	S	90	PRO
249	1tnr3_TNFA	R	91	HIS
249	1tnr3_TNFA	T	91	HIS
249	1tnr3_TNFA	S	91	HIS
250	1tnr3_TNFA	R	91	VAL
250	1tnr3_TNFA	T	91	VAL
250	1tnr3_TNFA	S	91	VAL
251	1tnr3_TNFA	R	91	CYS
251	1tnr3_TNFA	T	91	CYS
251	1tnr3_TNFA	S	91	CYS
252	1tnr3_TNFA	R	93	GLY
252	1tnr3_TNFA	T	93	GLY
252	1tnr3_TNFA	S	93	GLY
253	1tnr3_TNFA	R	94	GLY
253	1tnr3_TNFA	T	94	GLY
253	1tnr3_TNFA	S	94	GLY
254	1tnr3_TNFA	R	94	ALA
254	1tnr3_TNFA	T	94	ALA
254	1tnr3_TNFA	S	94	ALA
255	1tnr3_TNFA	R	96	THR
255	1tnr3_TNFA	T	96	THR
255	1tnr3_TNFA	S	96	THR
256	1tnr3_TNFA	R	97	LEU
256	1tnr3_TNFA	T	97	LEU
256	1tnr3_TNFA	S	97	LEU
257	1tnr3_TNFA	R	97	SER
257	1tnr3_TNFA	T	97	SER
257	1tnr3_TNFA	S	97	SER
258	1tnr3_TNFA	R	97	CYS
258	1tnr3_TNFA	T	97	CYS
258	1tnr3_TNFA	S	97	CYS
259	1tnr3_TNFA	R	97	ILE
259	1tnr3_TNFA	T	97	ILE
259	1tnr3_TNFA	S	97	ILE
260	1tnr3_TNFA	R	98	PHE
260	1tnr3_TNFA	T	98	PHE
260	1tnr3_TNFA	S	98	PHE
261	1tnr3_TNFA	R	99	TRP
261	1tnr3_TNFA	T	99	TRP
261	1tnr3_TNFA	S	99	TRP
262	1tnr3_TNFA	R	99	ARG
262	1tnr3_TNFA	T	99	ARG
262	1tnr3_TNFA	S	99	ARG
263	1tnr3_TNFA	R	101	ASN
263	1tnr3_TNFA	T	101	ASN
263	1tnr3_TNFA	S	101	ASN
264	1tnr3_TNFA	R	101	ASP
264	1tnr3_TNFA	T	101	ASP
264	1tnr3_TNFA	S	101	ASP
265	1tnr3_TNFA	R	101	SER
265	1tnr3_TNFA	T	101	SER
265	1tnr3_TNFA	S	101	SER
266	1tnr3_TNFA	R	102	TYR
266	1tnr3_TNFA	T	102	TYR
266	1tnr3_TNFA	S	102	TYR
267	1tnr3_TNFA	R	106	PHE
267	1tnr3_TNFA	T	106	PHE
267	1tnr3_TNFA	S	106	PHE
268	1tnr3_TNFA	R	107	SER
268	1tnr3_TNFA	T	107	SER
268	1tnr3_TNFA	S	107	SER
269	1tnr3_TNFA	R	108	ALA
269	1tnr3_TNFA	T	108	ALA
269	1tnr3_TNFA	S	108	ALA
270	1tnr3_TNFA	R	109	ILE
270	1tnr3_TNFA	T	109	ILE
270	1tnr3_TNFA	S	109	ILE
271	1tnr3_TNFA	R	109	VAL
271	1tnr3_TNFA	T	109	VAL
271	1tnr3_TNFA	S	109	VAL
272	1tnr3_TNFA	R	110	MET
272	1tnr3_TNFA	T	110	MET
272	1tnr3_TNFA	S	110	MET
273	1tnr3_TNFA	R	110	LEU
273	1tnr3_TNFA	T	110	LEU
273	1tnr3_TNFA	S	110	LEU
274	1tnr3_TNFA	R	111	TYR
274	1tnr3_TNFA	T	111	TYR
274	1tnr3_TNFA	S	111	TYR
275	1tnr3_TNFA	R	112	ASN
275	1tnr3_TNFA	T	112	ASN
275	1tnr3_TNFA	S	112	ASN

276	1tnr3_TNFA	R	112	PHE
276	1tnr3_TNFA	T	112	PHE
276	1tnr3_TNFA	S	112	PHE
277	1tnr3_TNFA	R	115	ASP
277	1tnr3_TNFA	T	115	ASP
277	1tnr3_TNFA	S	115	ASP
278	1tnr3_TNFA	R	116	HIS
278	1tnr3_TNFA	T	116	HIS
278	1tnr3_TNFA	S	116	HIS
279	1tnr3_TNFA	R	116	LEU
279	1tnr3_TNFA	T	116	LEU
279	1tnr3_TNFA	S	116	LEU
280	1tnr3_TNFA	R	116	SER
280	1tnr3_TNFA	T	116	SER
280	1tnr3_TNFA	S	116	SER
281	1tnr3_TNFA	R	116	CYS
281	1tnr3_TNFA	T	116	CYS
281	1tnr3_TNFA	S	116	CYS
282	1tnr3_TNFA	R	117	ARG
282	1tnr3_TNFA	T	117	ARG
282	1tnr3_TNFA	S	117	ARG
283	1tnr3_TNFA	R	118	CYS
283	1tnr3_TNFA	T	118	CYS
283	1tnr3_TNFA	S	118	CYS
284	1tnr3_TNFA	R	118	ALA
284	1tnr3_TNFA	T	118	ALA
284	1tnr3_TNFA	S	118	ALA
285	1tnr3_TNFA	R	119	SER
285	1tnr3_TNFA	T	119	SER
285	1tnr3_TNFA	S	119	SER
286	1tnr3_TNFA	R	120	ASN
286	1tnr3_TNFA	T	120	ASN
286	1tnr3_TNFA	S	120	ASN
287	1tnr3_TNFA	R	120	LYS
287	1tnr3_TNFA	T	120	LYS
287	1tnr3_TNFA	S	120	LYS
288	1tnr3_TNFA	R	123	LEU
288	1tnr3_TNFA	T	123	LEU
288	1tnr3_TNFA	S	123	LEU
289	1tnr3_TNFA	R	123	ALA
289	1tnr3_TNFA	T	123	ALA
289	1tnr3_TNFA	S	123	ALA
290	1tnr3_TNFA	R	124	ALA
290	1tnr3_TNFA	T	124	ALA
290	1tnr3_TNFA	S	124	ALA
291	1tnr3_TNFA	R	125	THR
291	1tnr3_TNFA	T	125	THR
291	1tnr3_TNFA	S	125	THR
292	1tnr3_TNFA	R	125	ARG
292	1tnr3_TNFA	T	125	ARG
292	1tnr3_TNFA	S	125	ARG
293	1tnr3_TNFA	R	125	TYR
293	1tnr3_TNFA	T	125	TYR
293	1tnr3_TNFA	S	125	TYR
294	1tnr3_TNFA	R	125	HIS
294	1tnr3_TNFA	T	125	HIS
294	1tnr3_TNFA	S	125	HIS
295	1tnr3_TNFA	R	126	TYR
295	1tnr3_TNFA	T	126	TYR
295	1tnr3_TNFA	S	126	TYR
296	1tnr3_TNFA	R	127	ARG
296	1tnr3_TNFA	T	127	ARG
296	1tnr3_TNFA	S	127	ARG
297	1tnr3_TNFA	R	128	LEU
297	1tnr3_TNFA	T	128	LEU
297	1tnr3_TNFA	S	128	LEU
298	1tnr3_TNFA	R	133	SER
298	1tnr3_TNFA	T	133	SER
298	1tnr3_TNFA	S	133	SER
299	1tnr3_TNFA	R	134	VAL
299	1tnr3_TNFA	T	134	VAL
299	1tnr3_TNFA	S	134	VAL
300	1tnr3_TNFA	R	134	LYS
300	1tnr3_TNFA	T	134	LYS
300	1tnr3_TNFA	S	134	LYS
301	1tnr3_TNFA	R	135	THR
301	1tnr3_TNFA	T	135	THR
301	1tnr3_TNFA	S	135	THR
302	1tnr3_TNFA	R	135	SER
302	1tnr3_TNFA	T	135	SER
302	1tnr3_TNFA	S	135	SER
303	1tnr3_TNFA	R	135	GLU
303	1tnr3_TNFA	T	135	GLU
303	1tnr3_TNFA	S	135	GLU
304	1tnr3_TNFA	R	135	VAL
304	1tnr3_TNFA	T	135	VAL
304	1tnr3_TNFA	S	135	VAL
305	1tnr3_TNFA	R	138	VAL
305	1tnr3_TNFA	T	138	VAL
305	1tnr3_TNFA	S	138	VAL

Iteration number	Filename	Chain	Residue index in chain	New residue
34	1tnr-3	R	0	YR
34	1tnr-3	T	0	YR
34	1tnr-3	S	0	YR
35	1tnr-3	R	0	RO
35	1tnr-3	T	0	RO
35	1tnr-3	S	0	RO
36	1tnr-3	R	0	ER
36	1tnr-3	T	0	ER
36	1tnr-3	S	0	ER
37	1tnr-3	R	1	RO
37	1tnr-3	T	1	RO
37	1tnr-3	S	1	RO
38	1tnr-3	R	4	HR
38	1tnr-3	T	4	HR
38	1tnr-3	S	4	HR
39	1tnr-3	R	5	IS
39	1tnr-3	T	5	IS
39	1tnr-3	S	5	IS
40	1tnr-3	R	5	YS
40	1tnr-3	T	5	YS
40	1tnr-3	S	5	YS
41	1tnr-3	R	5	SP
41	1tnr-3	T	5	SP
41	1tnr-3	S	5	SP
42	1tnr-3	R	6	HR
42	1tnr-3	T	6	HR
42	1tnr-3	S	6	HR
43	1tnr-3	R	6	AL
43	1tnr-3	T	6	AL
43	1tnr-3	S	6	AL
44	1tnr-3	R	7	RG
44	1tnr-3	T	7	RG
44	1tnr-3	S	7	RG
45	1tnr-3	R	7	YR
45	1tnr-3	T	7	YR
45	1tnr-3	S	7	YR
46	1tnr-3	R	7	LY
46	1tnr-3	T	7	LY
46	1tnr-3	S	7	LY
47	1tnr-3	R	7	RO
47	1tnr-3	T	7	RO
47	1tnr-3	S	7	RO
48	1tnr-3	R	7	LN
48	1tnr-3	T	7	LN
48	1tnr-3	S	7	LN
49	1tnr-3	R	8	LA
49	1tnr-3	T	8	LA
49	1tnr-3	S	8	LA
50	1tnr-3	R	10	ASP
50	1tnr-3	T	10	ASP
50	1tnr-3	S	10	ASP
51	1tnr-3	R	10	THR
51	1tnr-3	T	10	THR
51	1tnr-3	S	10	THR
52	1tnr-3	R	11	LYS
52	1tnr-3	T	11	LYS
52	1tnr-3	S	11	LYS
53	1tnr-3	R	12	THR
53	1tnr-3	T	12	THR
53	1tnr-3	S	12	THR
54	1tnr-3	R	12	LEU
54	1tnr-3	T	12	LEU
54	1tnr-3	S	12	LEU
55	1tnr-3	R	13	SER
55	1tnr-3	T	13	SER
55	1tnr-3	S	13	SER
56	1tnr-3	R	14	SER
56	1tnr-3	T	14	SER
56	1tnr-3	S	14	SER
57	1tnr-3	R	14	GLY
57	1tnr-3	T	14	GLY
57	1tnr-3	S	14	GLY
58	1tnr-3	R	14	PHE
58	1tnr-3	T	14	PHE
58	1tnr-3	S	14	PHE
59	1tnr-3	R	14	TRP
59	1tnr-3	T	14	TRP
59	1tnr-3	S	14	TRP
60	1tnr-3	R	14	TYR
60	1tnr-3	T	14	TYR
60	1tnr-3	S	14	TYR
61	1tnr-3	R	14	ARG
61	1tnr-3	T	14	ARG
61	1tnr-3	S	14	ARG
62	1tnr-3	R	15	TYR
62	1tnr-3	T	15	TYR
62	1tnr-3	S	15	TYR
63	1tnr-3	R	15	SER
63	1tnr-3	T	15	SER
63	1tnr-3	S	15	SER
64	1tnr-3	R	15	ARG
64	1tnr-3	T	15	ARG
64	1tnr-3	S	15	ARG
65	1tnr-3	R	15	PHE
65	1tnr-3	T	15	PHE
65	1tnr-3	S	15	PHE
66	1tnr-3	R	18	TYR
66	1tnr-3	T	18	TYR
66	1tnr-3	S	18	TYR
67	1tnr-3	R	18	GLY
67	1tnr-3	T	18	GLY
67	1tnr-3	S	18	GLY

68	1tnr-3	R	21	GLU
68	1tnr-3	T	21	GLU
68	1tnr-3	S	21	GLU
69	1tnr-3	R	22	ARG
69	1tnr-3	T	22	ARG
69	1tnr-3	S	22	ARG
70	1tnr-3	R	22	THR
70	1tnr-3	T	22	THR
70	1tnr-3	S	22	THR
71	1tnr-3	R	22	LEU
71	1tnr-3	T	22	LEU
71	1tnr-3	S	22	LEU
72	1tnr-3	R	22	ILE
72	1tnr-3	T	22	ILE
72	1tnr-3	S	22	ILE
73	1tnr-3	R	23	SER
73	1tnr-3	T	23	SER
73	1tnr-3	S	23	SER
74	1tnr-3	R	23	CYS
74	1tnr-3	T	23	CYS
74	1tnr-3	S	23	CYS
75	1tnr-3	R	24	LEU
75	1tnr-3	T	24	LEU
75	1tnr-3	S	24	LEU
76	1tnr-3	R	24	PHE
76	1tnr-3	T	24	PHE
76	1tnr-3	S	24	PHE
77	1tnr-3	R	26	SER
77	1tnr-3	T	26	SER
77	1tnr-3	S	26	SER
78	1tnr-3	R	26	ARG
78	1tnr-3	T	26	ARG
78	1tnr-3	S	26	ARG
79	1tnr-3	R	26	MET
79	1tnr-3	T	26	MET
79	1tnr-3	S	26	MET
80	1tnr-3	R	27	GLU
80	1tnr-3	T	27	GLU
80	1tnr-3	S	27	GLU
81	1tnr-3	R	28	PRO
81	1tnr-3	T	28	PRO
81	1tnr-3	S	28	PRO
82	1tnr-3	R	28	GLY
82	1tnr-3	T	28	GLY
82	1tnr-3	S	28	GLY
83	1tnr-3	R	28	GLY
83	1tnr-3	T	28	GLY
83	1tnr-3	S	28	GLY
84	1tnr-3	R	28	TYR
84	1tnr-3	T	28	TYR
84	1tnr-3	S	28	TYR
85	1tnr-3	R	28	PHE
85	1tnr-3	T	28	PHE
85	1tnr-3	S	28	PHE
86	1tnr-3	R	28	ARG
86	1tnr-3	T	28	ARG
86	1tnr-3	S	28	ARG
87	1tnr-3	R	28	TRP
87	1tnr-3	T	28	TRP
87	1tnr-3	S	28	TRP
88	1tnr-3	R	28	SER
88	1tnr-3	T	28	SER
88	1tnr-3	S	28	SER
89	1tnr-3	R	28	SER
89	1tnr-3	T	28	SER
89	1tnr-3	S	28	SER
90	1tnr-3	R	29	SER
90	1tnr-3	T	29	SER
90	1tnr-3	S	29	SER
91	1tnr-3	R	29	ARG
91	1tnr-3	T	29	ARG
91	1tnr-3	S	29	ARG
92	1tnr-3	R	30	ARG
92	1tnr-3	T	30	ARG
92	1tnr-3	S	30	ARG
93	1tnr-3	R	30	GLY
93	1tnr-3	T	30	GLY
93	1tnr-3	S	30	GLY
94	1tnr-3	R	30	MET
94	1tnr-3	T	30	MET
94	1tnr-3	S	30	MET
95	1tnr-3	R	30	LEU
95	1tnr-3	T	30	LEU
95	1tnr-3	S	30	LEU
96	1tnr-3	R	30	TYR
96	1tnr-3	T	30	TYR
96	1tnr-3	S	30	TYR
97	1tnr-3	R	31	ARG
97	1tnr-3	T	31	ARG
97	1tnr-3	S	31	ARG
98	1tnr-3	R	31	TYR
98	1tnr-3	T	31	TYR
98	1tnr-3	S	31	TYR
99	1tnr-3	R	31	SER
99	1tnr-3	T	31	SER
99	1tnr-3	S	31	SER
100	1tnr-3	R	31	ALA
100	1tnr-3	T	31	ALA
100	1tnr-3	S	31	ALA
101	1tnr-3	R	31	LEU
101	1tnr-3	T	31	LEU
101	1tnr-3	S	31	LEU
102	1tnr-3	R	31	ARG
102	1tnr-3	T	31	ARG

102	1tnr-3	S	31	ARG
103	1tnr-3	R	31	THR
103	1tnr-3	T	31	THR
103	1tnr-3	S	31	THR
104	1tnr-3	R	32	THR
104	1tnr-3	T	32	THR
104	1tnr-3	S	32	THR
105	1tnr-3	R	32	ASP
105	1tnr-3	T	32	ASP
105	1tnr-3	S	32	ASP
106	1tnr-3	R	34	PRO
106	1tnr-3	T	34	PRO
106	1tnr-3	S	34	PRO
107	1tnr-3	R	35	MET
107	1tnr-3	T	35	MET
107	1tnr-3	S	35	MET
108	1tnr-3	R	35	PRO
108	1tnr-3	T	35	PRO
108	1tnr-3	S	35	PRO
109	1tnr-3	R	35	LYS
109	1tnr-3	T	35	LYS
109	1tnr-3	S	35	LYS
110	1tnr-3	R	36	SER
110	1tnr-3	T	36	SER
110	1tnr-3	S	36	SER
111	1tnr-3	R	36	LEU
111	1tnr-3	T	36	LEU
111	1tnr-3	S	36	LEU
112	1tnr-3	R	36	ILE
112	1tnr-3	T	36	ILE
112	1tnr-3	S	36	ILE
113	1tnr-3	R	36	VAL
113	1tnr-3	T	36	VAL
113	1tnr-3	S	36	VAL
114	1tnr-3	R	37	SER
114	1tnr-3	T	37	SER
114	1tnr-3	S	37	SER
115	1tnr-3	R	37	ASN
115	1tnr-3	T	37	ASN
115	1tnr-3	S	37	ASN
116	1tnr-3	R	37	ALA
116	1tnr-3	T	37	ALA
116	1tnr-3	S	37	ALA
117	1tnr-3	R	37	ARG
117	1tnr-3	T	37	ARG
117	1tnr-3	S	37	ARG
118	1tnr-3	R	37	TRP
118	1tnr-3	T	37	TRP
118	1tnr-3	S	37	TRP
119	1tnr-3	R	37	TYR
119	1tnr-3	T	37	TYR
119	1tnr-3	S	37	TYR
120	1tnr-3	R	37	PHE
120	1tnr-3	T	37	PHE
120	1tnr-3	S	37	PHE
121	1tnr-3	R	38	GLY
121	1tnr-3	T	38	GLY
121	1tnr-3	S	38	GLY
122	1tnr-3	R	38	TRP
122	1tnr-3	T	38	TRP
122	1tnr-3	S	38	TRP
123	1tnr-3	R	38	PHE
123	1tnr-3	T	38	PHE
123	1tnr-3	S	38	PHE
124	1tnr-3	R	38	TYR
124	1tnr-3	T	38	TYR
124	1tnr-3	S	38	TYR
125	1tnr-3	R	39	ASP
125	1tnr-3	T	39	ASP
125	1tnr-3	S	39	ASP
126	1tnr-3	R	40	ARG
126	1tnr-3	T	40	ARG
126	1tnr-3	S	40	ARG
127	1tnr-3	R	40	TYR
127	1tnr-3	T	40	TYR
127	1tnr-3	S	40	TYR
128	1tnr-3	R	40	SER
128	1tnr-3	T	40	SER
128	1tnr-3	S	40	SER
129	1tnr-3	R	41	ASP
129	1tnr-3	T	41	ASP
129	1tnr-3	S	41	ASP
130	1tnr-3	R	41	LYS
130	1tnr-3	T	41	LYS
130	1tnr-3	S	41	LYS
131	1tnr-3	R	41	ILE
131	1tnr-3	T	41	ILE
131	1tnr-3	S	41	ILE
132	1tnr-3	R	42	PRO
132	1tnr-3	T	42	PRO
132	1tnr-3	S	42	PRO
133	1tnr-3	R	42	TYR
133	1tnr-3	T	42	TYR
133	1tnr-3	S	42	TYR
134	1tnr-3	R	42	LEU
134	1tnr-3	T	42	LEU
134	1tnr-3	S	42	LEU
135	1tnr-3	R	43	PRO
135	1tnr-3	T	43	PRO
135	1tnr-3	S	43	PRO
136	1tnr-3	R	43	PHE
136	1tnr-3	T	43	PHE
136	1tnr-3	S	43	PHE
137	1tnr-3	R	43	GLU

137	1tnr-3	T	43	GLU
137	1tnr-3	S	43	GLU
138	1tnr-3	R	43	SER
138	1tnr-3	T	43	SER
138	1tnr-3	S	43	SER
139	1tnr-3	R	44	GLU
139	1tnr-3	T	44	GLU
139	1tnr-3	S	44	GLU
140	1tnr-3	R	44	PRO
140	1tnr-3	T	44	PRO
140	1tnr-3	S	44	PRO
141	1tnr-3	R	45	SER
141	1tnr-3	T	45	SER
141	1tnr-3	S	45	SER
142	1tnr-3	R	45	LEU
142	1tnr-3	T	45	LEU
142	1tnr-3	S	45	LEU
143	1tnr-3	R	45	VAL
143	1tnr-3	T	45	VAL
143	1tnr-3	S	45	VAL
144	1tnr-3	R	46	SER
144	1tnr-3	T	46	SER
144	1tnr-3	S	46	SER
145	1tnr-3	R	46	GLY
145	1tnr-3	T	46	GLY
145	1tnr-3	S	46	GLY
146	1tnr-3	R	46	TYR
146	1tnr-3	T	46	TYR
146	1tnr-3	S	46	TYR
147	1tnr-3	R	46	ALA
147	1tnr-3	T	46	ALA
147	1tnr-3	S	46	ALA
148	1tnr-3	R	46	ASN
148	1tnr-3	T	46	ASN
148	1tnr-3	S	46	ASN
149	1tnr-3	R	46	ILE
149	1tnr-3	T	46	ILE
149	1tnr-3	S	46	ILE
150	1tnr-3	R	46	PRO
150	1tnr-3	T	46	PRO
150	1tnr-3	S	46	PRO
151	1tnr-3	R	46	TRP
151	1tnr-3	T	46	TRP
151	1tnr-3	S	46	TRP
152	1tnr-3	R	46	PRO
152	1tnr-3	T	46	PRO
152	1tnr-3	S	46	PRO
153	1tnr-3	R	46	ARG
153	1tnr-3	T	46	ARG
153	1tnr-3	S	46	ARG
154	1tnr-3	R	48	HIS
154	1tnr-3	T	48	HIS
154	1tnr-3	S	48	HIS
155	1tnr-3	R	48	ASN
155	1tnr-3	T	48	ASN
155	1tnr-3	S	48	ASN
156	1tnr-3	R	49	TRP
156	1tnr-3	T	49	TRP
156	1tnr-3	S	49	TRP
157	1tnr-3	R	49	SER
157	1tnr-3	T	49	SER
157	1tnr-3	S	49	SER
158	1tnr-3	R	49	ARG
158	1tnr-3	T	49	ARG
158	1tnr-3	S	49	ARG
159	1tnr-3	R	50	LYS
159	1tnr-3	T	50	LYS
159	1tnr-3	S	50	LYS
160	1tnr-3	R	50	CYS
160	1tnr-3	T	50	CYS
160	1tnr-3	S	50	CYS
161	1tnr-3	R	50	GLY
161	1tnr-3	T	50	GLY
161	1tnr-3	S	50	GLY
162	1tnr-3	R	50	ILE
162	1tnr-3	T	50	ILE
162	1tnr-3	S	50	ILE
163	1tnr-3	R	51	PRO
163	1tnr-3	T	51	PRO
163	1tnr-3	S	51	PRO
164	1tnr-3	R	51	LEU
164	1tnr-3	T	51	LEU
164	1tnr-3	S	51	LEU
165	1tnr-3	R	51	THR
165	1tnr-3	T	51	THR
165	1tnr-3	S	51	THR
166	1tnr-3	R	51	GLY
166	1tnr-3	T	51	GLY
166	1tnr-3	S	51	GLY
167	1tnr-3	R	51	TYR
167	1tnr-3	T	51	TYR
167	1tnr-3	S	51	TYR
168	1tnr-3	R	52	PHE
168	1tnr-3	T	52	PHE
168	1tnr-3	S	52	PHE
169	1tnr-3	R	52	PRO
169	1tnr-3	T	52	PRO
169	1tnr-3	S	52	PRO
170	1tnr-3	R	52	TYR
170	1tnr-3	T	52	TYR
170	1tnr-3	S	52	TYR
171	1tnr-3	R	53	GLN
171	1tnr-3	T	53	GLN
171	1tnr-3	S	53	GLN

172	1tnr-3	R	54	LEU
172	1tnr-3	T	54	LEU
172	1tnr-3	S	54	LEU
173	1tnr-3	R	54	CYS
173	1tnr-3	T	54	CYS
173	1tnr-3	S	54	CYS
174	1tnr-3	R	54	ASN
174	1tnr-3	T	54	ASN
174	1tnr-3	S	54	ASN
175	1tnr-3	R	54	ILE
175	1tnr-3	T	54	ILE
175	1tnr-3	S	54	ILE
176	1tnr-3	R	54	SER
176	1tnr-3	T	54	SER
176	1tnr-3	S	54	SER
177	1tnr-3	R	55	ARG
177	1tnr-3	T	55	ARG
177	1tnr-3	S	55	ARG
178	1tnr-3	R	55	SER
178	1tnr-3	T	55	SER
178	1tnr-3	S	55	SER
179	1tnr-3	R	55	TYR
179	1tnr-3	T	55	TYR
179	1tnr-3	S	55	TYR
180	1tnr-3	R	55	GLY
180	1tnr-3	T	55	GLY
180	1tnr-3	S	55	GLY
181	1tnr-3	R	56	ARG
181	1tnr-3	T	56	ARG
181	1tnr-3	S	56	ARG
182	1tnr-3	R	57	ASN
182	1tnr-3	T	57	ASN
182	1tnr-3	S	57	ASN
183	1tnr-3	R	58	SER
183	1tnr-3	T	58	SER
183	1tnr-3	S	58	SER
184	1tnr-3	R	58	ASP
184	1tnr-3	T	58	ASP
184	1tnr-3	S	58	ASP
185	1tnr-3	R	58	ARG
185	1tnr-3	T	58	ARG
185	1tnr-3	S	58	ARG
186	1tnr-3	R	58	SER
186	1tnr-3	T	58	SER
186	1tnr-3	S	58	SER
187	1tnr-3	R	58	TRP
187	1tnr-3	T	58	TRP
187	1tnr-3	S	58	TRP
188	1tnr-3	R	59	TYR
188	1tnr-3	T	59	TYR
188	1tnr-3	S	59	TYR
189	1tnr-3	R	59	CYS
189	1tnr-3	T	59	CYS
189	1tnr-3	S	59	CYS
190	1tnr-3	R	61	TYR
190	1tnr-3	T	61	TYR
190	1tnr-3	S	61	TYR
191	1tnr-3	R	62	GLN
191	1tnr-3	T	62	GLN
191	1tnr-3	S	62	GLN
192	1tnr-3	R	62	PRO
192	1tnr-3	T	62	PRO
192	1tnr-3	S	62	PRO
193	1tnr-3	R	63	PHE
193	1tnr-3	T	63	PHE
193	1tnr-3	S	63	PHE
194	1tnr-3	R	64	ARG
194	1tnr-3	T	64	ARG
194	1tnr-3	S	64	ARG
195	1tnr-3	R	64	TYR
195	1tnr-3	T	64	TYR
195	1tnr-3	S	64	TYR
196	1tnr-3	R	65	ALA
196	1tnr-3	T	65	ALA
196	1tnr-3	S	65	ALA
197	1tnr-3	R	66	ILE
197	1tnr-3	T	66	ILE
197	1tnr-3	S	66	ILE
198	1tnr-3	R	67	LEU
198	1tnr-3	T	67	LEU
198	1tnr-3	S	67	LEU
199	1tnr-3	R	67	LYS
199	1tnr-3	T	67	LYS
199	1tnr-3	S	67	LYS
200	1tnr-3	R	68	PRO
200	1tnr-3	T	68	PRO
200	1tnr-3	S	68	PRO
201	1tnr-3	R	68	LEU
201	1tnr-3	T	68	LEU
201	1tnr-3	S	68	LEU
202	1tnr-3	R	68	TYR
202	1tnr-3	T	68	TYR
202	1tnr-3	S	68	TYR
203	1tnr-3	R	68	MET
203	1tnr-3	T	68	MET
203	1tnr-3	S	68	MET
204	1tnr-3	R	69	PHE
204	1tnr-3	T	69	PHE
204	1tnr-3	S	69	PHE
205	1tnr-3	R	70	ILE
205	1tnr-3	T	70	ILE
205	1tnr-3	S	70	ILE
206	1tnr-3	R	70	SER
206	1tnr-3	T	70	SER

206	1tnr-3	S	70	SER
207	1tnr-3	R	71	ALA
207	1tnr-3	T	71	ALA
207	1tnr-3	S	71	ALA
208	1tnr-3	R	71	PRO
208	1tnr-3	T	71	PRO
208	1tnr-3	S	71	PRO
209	1tnr-3	R	71	PHE
209	1tnr-3	T	71	PHE
209	1tnr-3	S	71	PHE
210	1tnr-3	R	72	TYR
210	1tnr-3	T	72	TYR
210	1tnr-3	S	72	TYR
211	1tnr-3	R	72	PHE
211	1tnr-3	T	72	PHE
211	1tnr-3	S	72	PHE
212	1tnr-3	R	72	TRP
212	1tnr-3	T	72	TRP
212	1tnr-3	S	72	TRP
213	1tnr-3	R	73	TYR
213	1tnr-3	T	73	TYR
213	1tnr-3	S	73	TYR
214	1tnr-3	R	73	ARG
214	1tnr-3	T	73	ARG
214	1tnr-3	S	73	ARG
215	1tnr-3	R	73	ASP
215	1tnr-3	T	73	ASP
215	1tnr-3	S	73	ASP
216	1tnr-3	R	73	SER
216	1tnr-3	T	73	SER
216	1tnr-3	S	73	SER
217	1tnr-3	R	74	TYR
217	1tnr-3	T	74	TYR
217	1tnr-3	S	74	TYR
218	1tnr-3	R	74	ALA
218	1tnr-3	T	74	ALA
218	1tnr-3	S	74	ALA
219	1tnr-3	R	74	VAL
219	1tnr-3	T	74	VAL
219	1tnr-3	S	74	VAL
220	1tnr-3	R	74	ARG
220	1tnr-3	T	74	ARG
220	1tnr-3	S	74	ARG
221	1tnr-3	R	75	GLY
221	1tnr-3	T	75	GLY
221	1tnr-3	S	75	GLY
222	1tnr-3	R	75	ALA
222	1tnr-3	T	75	ALA
222	1tnr-3	S	75	ALA
223	1tnr-3	R	77	TRP
223	1tnr-3	T	77	TRP
223	1tnr-3	S	77	TRP
224	1tnr-3	R	77	CYS
224	1tnr-3	T	77	CYS
224	1tnr-3	S	77	CYS
225	1tnr-3	R	77	PRO
225	1tnr-3	T	77	PRO
225	1tnr-3	S	77	PRO
226	1tnr-3	R	77	GLN
226	1tnr-3	T	77	GLN
226	1tnr-3	S	77	GLN
227	1tnr-3	R	77	GLY
227	1tnr-3	T	77	GLY
227	1tnr-3	S	77	GLY
228	1tnr-3	R	78	HIS
228	1tnr-3	T	78	HIS
228	1tnr-3	S	78	HIS
229	1tnr-3	R	78	GLU
229	1tnr-3	T	78	GLU
229	1tnr-3	S	78	GLU
230	1tnr-3	R	79	ILE
230	1tnr-3	T	79	ILE
230	1tnr-3	S	79	ILE
231	1tnr-3	R	80	GLY
231	1tnr-3	T	80	GLY
231	1tnr-3	S	80	GLY
232	1tnr-3	R	80	ALA
232	1tnr-3	T	80	ALA
232	1tnr-3	S	80	ALA
233	1tnr-3	R	80	MET
233	1tnr-3	T	80	MET
233	1tnr-3	S	80	MET
234	1tnr-3	R	81	ARG
234	1tnr-3	T	81	ARG
234	1tnr-3	S	81	ARG
235	1tnr-3	R	81	TRP
235	1tnr-3	T	81	TRP
235	1tnr-3	S	81	TRP
236	1tnr-3	R	81	TYR
236	1tnr-3	T	81	TYR
236	1tnr-3	S	81	TYR
237	1tnr-3	R	81	PHE
237	1tnr-3	T	81	PHE
237	1tnr-3	S	81	PHE
238	1tnr-3	R	82	ASP
238	1tnr-3	T	82	ASP
238	1tnr-3	S	82	ASP
239	1tnr-3	R	82	ARG
239	1tnr-3	T	82	ARG
239	1tnr-3	S	82	ARG
240	1tnr-3	R	82	TYR
240	1tnr-3	T	82	TYR
240	1tnr-3	S	82	TYR
241	1tnr-3	R	83	TYR

241	1tnr-3	T	83	TYR
241	1tnr-3	S	83	TYR
242	1tnr-3	R	83	ARG
242	1tnr-3	T	83	ARG
242	1tnr-3	S	83	ARG
243	1tnr-3	R	85	LEU
243	1tnr-3	T	85	LEU
243	1tnr-3	S	85	LEU
244	1tnr-3	R	86	LYS
244	1tnr-3	T	86	LYS
244	1tnr-3	S	86	LYS
245	1tnr-3	R	87	GLU
245	1tnr-3	T	87	GLU
245	1tnr-3	S	87	GLU
246	1tnr-3	R	89	TRP
246	1tnr-3	T	89	TRP
246	1tnr-3	S	89	TRP
247	1tnr-3	R	89	GLN
247	1tnr-3	T	89	GLN
247	1tnr-3	S	89	GLN
248	1tnr-3	R	90	PRO
248	1tnr-3	T	90	PRO
248	1tnr-3	S	90	PRO
249	1tnr-3	R	91	HIS
249	1tnr-3	T	91	HIS
249	1tnr-3	S	91	HIS
250	1tnr-3	R	91	VAL
250	1tnr-3	T	91	VAL
250	1tnr-3	S	91	VAL
251	1tnr-3	R	91	CYS
251	1tnr-3	T	91	CYS
251	1tnr-3	S	91	CYS
252	1tnr-3	R	93	GLY
252	1tnr-3	T	93	GLY
252	1tnr-3	S	93	GLY
253	1tnr-3	R	94	GLY
253	1tnr-3	T	94	GLY
253	1tnr-3	S	94	GLY
254	1tnr-3	R	94	ALA
254	1tnr-3	T	94	ALA
254	1tnr-3	S	94	ALA
255	1tnr-3	R	96	THR
255	1tnr-3	T	96	THR
255	1tnr-3	S	96	THR
256	1tnr-3	R	97	LEU
256	1tnr-3	T	97	LEU
256	1tnr-3	S	97	LEU
257	1tnr-3	R	97	SER
257	1tnr-3	T	97	SER
257	1tnr-3	S	97	SER
258	1tnr-3	R	97	CYS
258	1tnr-3	T	97	CYS
258	1tnr-3	S	97	CYS
259	1tnr-3	R	97	ILE
259	1tnr-3	T	97	ILE
259	1tnr-3	S	97	ILE
260	1tnr-3	R	98	PHE
260	1tnr-3	T	98	PHE
260	1tnr-3	S	98	PHE
261	1tnr-3	R	99	TRP
261	1tnr-3	T	99	TRP
261	1tnr-3	S	99	TRP
262	1tnr-3	R	99	ARG
262	1tnr-3	T	99	ARG
262	1tnr-3	S	99	ARG
263	1tnr-3	R	101	ASN
263	1tnr-3	T	101	ASN
263	1tnr-3	S	101	ASN
264	1tnr-3	R	101	ASP
264	1tnr-3	T	101	ASP
264	1tnr-3	S	101	ASP
265	1tnr-3	R	101	SER
265	1tnr-3	T	101	SER
265	1tnr-3	S	101	SER
266	1tnr-3	R	102	TYR
266	1tnr-3	T	102	TYR
266	1tnr-3	S	102	TYR
267	1tnr-3	R	106	PHE
267	1tnr-3	T	106	PHE
267	1tnr-3	S	106	PHE
268	1tnr-3	R	107	SER
268	1tnr-3	T	107	SER
268	1tnr-3	S	107	SER
269	1tnr-3	R	108	ALA
269	1tnr-3	T	108	ALA
269	1tnr-3	S	108	ALA
270	1tnr-3	R	109	ILE
270	1tnr-3	T	109	ILE
270	1tnr-3	S	109	ILE
271	1tnr-3	R	109	VAL
271	1tnr-3	T	109	VAL
271	1tnr-3	S	109	VAL
272	1tnr-3	R	110	MET
272	1tnr-3	T	110	MET
272	1tnr-3	S	110	MET
273	1tnr-3	R	110	LEU
273	1tnr-3	T	110	LEU
273	1tnr-3	S	110	LEU
274	1tnr-3	R	111	TYR
274	1tnr-3	T	111	TYR
274	1tnr-3	S	111	TYR
275	1tnr-3	R	112	ASN
275	1tnr-3	T	112	ASN
275	1tnr-3	S	112	ASN

276	1tnr-3	R	112	PHE
276	1tnr-3	T	112	PHE
276	1tnr-3	S	112	PHE
277	1tnr-3	R	115	ASP
277	1tnr-3	T	115	ASP
277	1tnr-3	S	115	ASP
278	1tnr-3	R	116	HIS
278	1tnr-3	T	116	HIS
278	1tnr-3	S	116	HIS
279	1tnr-3	R	116	LEU
279	1tnr-3	T	116	LEU
279	1tnr-3	S	116	LEU
280	1tnr-3	R	116	SER
280	1tnr-3	T	116	SER
280	1tnr-3	S	116	SER
281	1tnr-3	R	116	CYS
281	1tnr-3	T	116	CYS
281	1tnr-3	S	116	CYS
282	1tnr-3	R	117	ARG
282	1tnr-3	T	117	ARG
282	1tnr-3	S	117	ARG
283	1tnr-3	R	118	CYS
283	1tnr-3	T	118	CYS
283	1tnr-3	S	118	CYS
284	1tnr-3	R	118	ALA
284	1tnr-3	T	118	ALA
284	1tnr-3	S	118	ALA
285	1tnr-3	R	119	SER
285	1tnr-3	T	119	SER
285	1tnr-3	S	119	SER
286	1tnr-3	R	120	ASN
286	1tnr-3	T	120	ASN
286	1tnr-3	S	120	ASN
287	1tnr-3	R	120	LYS
287	1tnr-3	T	120	LYS
287	1tnr-3	S	120	LYS
288	1tnr-3	R	123	LEU
288	1tnr-3	T	123	LEU
288	1tnr-3	S	123	LEU
289	1tnr-3	R	123	ALA
289	1tnr-3	T	123	ALA
289	1tnr-3	S	123	ALA
290	1tnr-3	R	124	ALA
290	1tnr-3	T	124	ALA
290	1tnr-3	S	124	ALA
291	1tnr-3	R	125	THR
291	1tnr-3	T	125	THR
291	1tnr-3	S	125	THR
292	1tnr-3	R	125	ARG
292	1tnr-3	T	125	ARG
292	1tnr-3	S	125	ARG
293	1tnr-3	R	125	TYR
293	1tnr-3	T	125	TYR
293	1tnr-3	S	125	TYR
294	1tnr-3	R	125	HIS
294	1tnr-3	T	125	HIS
294	1tnr-3	S	125	HIS
295	1tnr-3	R	126	TYR
295	1tnr-3	T	126	TYR
295	1tnr-3	S	126	TYR
296	1tnr-3	R	127	ARG
296	1tnr-3	T	127	ARG
296	1tnr-3	S	127	ARG
297	1tnr-3	R	128	LEU
297	1tnr-3	T	128	LEU
297	1tnr-3	S	128	LEU
298	1tnr-3	R	133	SER
298	1tnr-3	T	133	SER
298	1tnr-3	S	133	SER
299	1tnr-3	R	134	VAL
299	1tnr-3	T	134	VAL
299	1tnr-3	S	134	VAL
300	1tnr-3	R	134	LYS
300	1tnr-3	T	134	LYS
300	1tnr-3	S	134	LYS
301	1tnr-3	R	135	THR
301	1tnr-3	T	135	THR
301	1tnr-3	S	135	THR
302	1tnr-3	R	135	SER
302	1tnr-3	T	135	SER
302	1tnr-3	S	135	SER
303	1tnr-3	R	135	GLU
303	1tnr-3	T	135	GLU
303	1tnr-3	S	135	GLU
304	1tnr-3	R	135	VAL
304	1tnr-3	T	135	VAL
304	1tnr-3	S	135	VAL
305	1tnr-3	R	138	VAL
305	1tnr-3	T	138	VAL
305	1tnr-3	S	138	VAL

Method

The 3D-structure of your protein of interest is known. Information from this 3D-structure will be obtained using WHAT IF Web services, the UniProt database and the Reprof software.

The structural information was obtained from the analysis of PDB:

1EXT (http://www.rcsb.org/pdb/explore/explore.do?structureId=1EXT)
Annotations about this protein were obtained from UniProt entry

p19438 (http://www.uniprot.org/uniprot/p19438)
See the method page for more information.

Amino Acids

You are interested in the mutation of a Cysteine into a Glycine at position 62.

The figure below shows the schematic structures of the original (left) and the mutant (right) amino acid. The backbone, which is the same for each amino acid, is colored red. The side chain, unique for each amino acid, is colored black.

Each amino acid has its own specific size, charge, and hydrophobicity-value. The original wild-type residue and newly introduced mutant residue often differ in these properties.

The mutant residue is smaller than the wild-type residue.

The wild-type residue is more hydrophobic than the mutant residue.

The report will evaluate the effect of the mutation on the following features: Contacts made by the mutated residue, structural domains in which the residue is located, modifications on this residue and known variants for this residue. A feature will only be shown when information is available. A short conclusion based on just the amino acid properties is shown always. In case a 3D-structure/model is available you will also find images and animations in the report.

Contacts

The wild-type residue is annotated in UniProt to be involved in a cysteine bridge, which is important for stability of the protein. Only cysteines can make these type of bonds, the mutation causes loss of this interaction and will have a severe effect on the 3D-structure of the protein.

Together with loss of the cysteine bond, the differences between the old and new residue can cause destabilization of the structure.

Structure

The mutation is located within a stretch of residues that is repeated in the protein, this repeat is named TNFR-Cys 1. The mutation into another residue might disturb this repeat and consequently any function this repeat might have.

The mutation introduces a glycine at this position. Glycines are very flexible and can disturb the required rigidity of the protein at this position.

Variants

This mutation matches a previously described variant, with the following description: Familial hibernian fever (FHF) [MIM:142680].

See the ExPASy site about this variant:

VAR_019303 (http://www.expasy.org/cgi-bin/variant_pages/get-sprot-variant.pl?VAR_019303)
The variant is annotated with severity: DISEASE

The mutation is located in a region with known splice variants, described as:

Familial hibernian fever (FHF) [MIM:142680] Familial hibernian fever (FHF) [MIM:142680]

Conservation

Only this residue type was found at this position. Mutation of a 100% conserved residue is usually damaging for the protein.

The mutant and wild-type residue are not very similar. Based on this conservation information this mutation is probably damaging to the protein.

Domains

Interpro Domain	Gene Ontology Term	Broad Gene Ontology Term
Tnfr/Ngfr Cysteine-Rich Region IPR001368 (http://www.ebi.ac.uk/interpro/entry/IPR001368)	Protein Binding GO:0005515 (http://www.ebi.ac.uk/QuickGO/GTerm?id=GO:0005515)	Binding GO:0005488 (http://www.ebi.ac.uk/QuickGO/GTerm?id=GO:0005488) Molecular_Function GO:0003674 (http://www.ebi.ac.uk/QuickGO/GTerm?id=GO:0003674)
Tumor Necrosis Factor Receptor 1A, N-Terminal IPR033993 (http://www.ebi.ac.uk/interpro/entry/IPR033993)	None	None

The mutated residue is located in a domain that is important for binding of other molecules and in contact with residues in a domain that is also important for binding. The mutation might disturb the interaction between these two domains and as such affect the function of the protein.

The mutated residue is located in a domain that is important for binding of other molecules. The mutated residue is in contact with residues in another domain. It is possible that the mutation disturbs these contacts.

Amino Acid Properties

The wild-type and mutant amino acids differ in size.

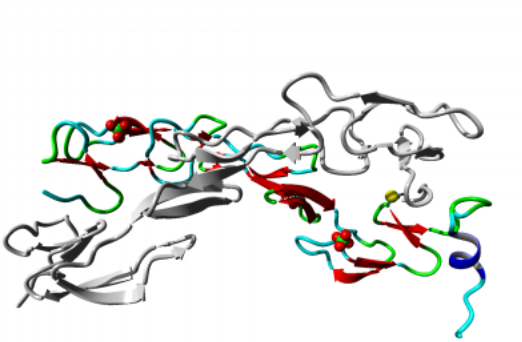
The mutant residue is smaller than the wild-type residue.

The mutation will cause an empty space in the core of the protein.

The hydrophobicity of the wild-type and mutant residue differs.

The mutation will cause loss of hydrophobic interactions in the core of the protein.

Images



Download (/hope/yasara/94214ee3-183d-441c-a1a1-c4a8e768cf26/16GLY_overview.png/)

Overview of the protein in ribbon-presentation. The protein is coloured by element; α-helix=blue, β-strand = red, turn=green, 3/10 helix=yellow and random coil=cyan. Other molecules in the complex are coloured grey when present.

Citation

Please use the following citation when referencing the results in your report:

Protein structure analysis of mutations causing inheritable diseases. An e-Science approach with life scientist friendly interfaces.

BMC Bioinformatics. 2010 Nov 8;11(1):548. DOI: 10.1186/1471-2105-11-548. (http://dx.doi.org/10.1186/1471-2105-11-548) PubMed: 21059217. (http://www.ncbi.nlm.nih.gov/pubmed/21059217)

Method

The 3D-structure of your protein of interest is known. Information from this 3D-structure will be obtained using WHAT IF Web services, the UniProt database and the Reprof software.

The structural information was obtained from the analysis of PDB:

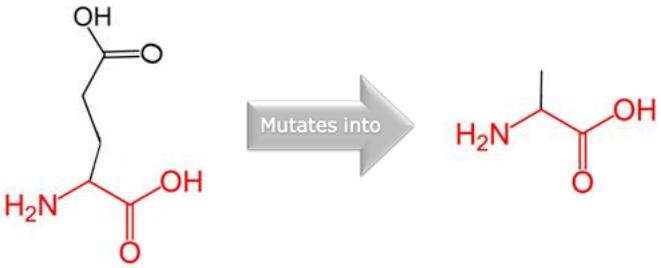
1EXT (http://www.rcsb.org/pdb/explore/explore.do?structureId=1EXT)
Annotations about this protein were obtained from UniProt entry

p19438 (http://www.uniprot.org/uniprot/p19438)
See the method page for more information.

Amino Acids

You are interested in the mutation of a Glutamic Acid into a Alanine at position 138.

The figure below shows the schematic structures of the original (left) and the mutant (right) amino acid. The backbone, which is the same for each amino acid, is colored red. The side chain, unique for each amino acid, is colored black.



Each amino acid has its own specific size, charge, and hydrophobicity-value. The original wild-type residue and newly introduced mutant residue often differ in these properties.

The mutant residue is smaller than the wild-type residue.

The wild-type residue charge was NEGATIVE, the mutant residue charge is NEUTRAL.

The mutant residue is more hydrophobic than the wild-type residue.

The report will evaluate the effect of the mutation on the following features: Contacts made by the mutated residue, structural domains in which the residue is located, modifications on this residue and known variants for this residue. A feature will only be shown when information is available. A short conclusion based on just the amino acid properties is shown always. In case a 3D-structure/model is available you will also find images and animations in the report.

Contacts

The wild-type residue forms a salt bridge with:

- Proline at position 368
- Leucine at position 390

The difference in charge will disturb the ionic interaction made by the original, wild-type residue.

Structure

The mutation is located within a stretch of residues that is repeated in the protein, this repeat is named TNFR-Cys 3. The mutation into another residue might disturb this repeat and consequently any function this repeat might have.

Conservation

The wild-type residue occurs often at this position in the sequence, but other residues have also been observed here.

Your mutant residue is among the other residue types that have been observed at this position in homologous sequences. This means that this mutation can occur at this position and is probably not damaging to the protein.

Domains

Interpro Domain	Gene Ontology Term	Broad Gene Ontology Term

Hope Version 1.1.1

Interpro Domain	Gene Ontology Term	Broad Gene Ontology Term
Tumour Necrosis Factor Receptor 1A IPR020419 (http://www.ebi.ac.uk/interpro/entry/IPR020419)	Tumor Necrosis Factor-Activated Receptor Activity GO:0005031 (http://www.ebi.ac.uk/QuickGO/GTerm?id=GO:0005031)	Molecular Transducer Activity GO:0060089 (http://www.ebi.ac.uk/QuickGO/GTerm?id=GO:0060089) Molecular_Function GO:0003674 (http://www.ebi.ac.uk/QuickGO/GTerm?id=GO:0003674)
Tnfr/Ngfr Cysteine-Rich Region IPR001368 (http://www.ebi.ac.uk/interpro/entry/IPR001368)	Protein Binding GO:0005515 (http://www.ebi.ac.uk/QuickGO/GTerm?id=GO:0005515)	Binding GO:0005488 (http://www.ebi.ac.uk/QuickGO/GTerm?id=GO:0005488) Molecular_Function GO:0003674 (http://www.ebi.ac.uk/QuickGO/GTerm?id=GO:0003674)
Tumor Necrosis Factor Receptor 1A, N-Terminal IPR033993 (http://www.ebi.ac.uk/interpro/entry/IPR033993)	None	None

The mutated residue is located in a domain that is important for binding of other molecules and in contact with residues in a domain that is also important for binding. The mutation might disturb the interaction between these two domains and as such affect the function of the protein.

The mutated residue is located in a domain that is important for binding of other molecules and in contact with residues in a domain that is important for the activity of the protein. The mutation might affect this interaction and thereby disturb signal transfer from binding domain to the activity domain.

The mutated residue is located in a domain that is important for binding of other molecules. The mutated residue is in contact with residues in another domain. It is possible that the mutation disturbs these contacts.

Amino Acid Properties

There is a difference in charge between the wild-type and mutant amino acid.

The charge of the wild-type residue is lost by this mutation. This can cause loss of interactions with other molecules.

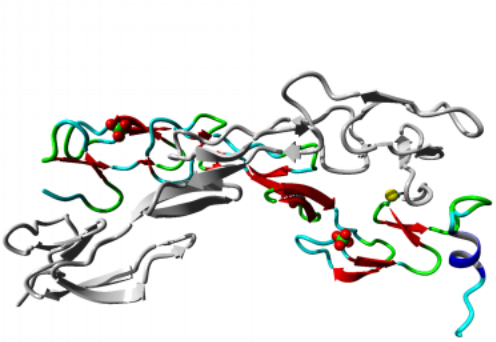
The wild-type and mutant amino acids differ in size.

The mutant residue is smaller than the wild-type residue.

This will cause a possible loss of external interactions.

The hydrophobicity of the wild-type and mutant residue differs.

Images



Download (/hope/yasara/5ecdbfa0-57b7-4570-b5d0-0def9401cf33/25ALA_overview.png/)

Overview of the protein in ribbon-presentation. The protein is coloured by element; α-helix=blue, β-strand = red, turn=green, 3/10 helix=yellow and random coil=cyan. Other molecules in the complex are coloured grey when present.

Citation

Please use the following citation when referencing the results in your report:

Protein structure analysis of mutations causing inheritable diseases. An e-Science approach with life scientist friendly interfaces.

BMC Bioinformatics. 2010 Nov 8;11(1):548. DOI: 10.1186/1471-2105-11-548. (http://dx.doi.org/10.1186/1471-2105-11-548) PubMed: 21059217. (http://www.ncbi.nlm.nih.gov/pubmed/21059217)

Method

The 3D-structure of your protein of interest is known. Information from this 3D-structure will be obtained using WHAT IF Web services, the UniProt database and the Reprof software.

The structural information was obtained from the analysis of PDB:

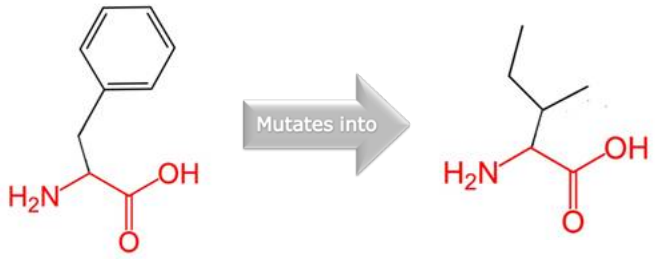
1EXT (http://www.rcsb.org/pdb/explore/explore.do?structureId=1EXT)
Annotations about this protein were obtained from UniProt entry

p19438 (http://www.uniprot.org/uniprot/p19438)
See the method page for more information.

Amino Acids

You are interested in the mutation of a Phenylalanine into a Isoleucine at position 141.

The figure below shows the schematic structures of the original (left) and the mutant (right) amino acid. The backbone, which is the same for each amino acid, is colored red. The side chain, unique for each amino acid, is colored black.



Each amino acid has its own specific size, charge, and hydrophobicity-value. The original wild-type residue and newly introduced mutant residue often differ in these properties.

The mutant residue is smaller than the wild-type residue.

The report will evaluate the effect of the mutation on the following features: Contacts made by the mutated residue, structural domains in which the residue is located, modifications on this residue and known variants for this residue. A feature will only be shown when information is available. A short conclusion based on just the amino acid properties is shown always. In case a 3D-structure/model is available you will also find images and animations in the report.

Structure

The mutation is located within a stretch of residues that is repeated in the protein, this repeat is named TNFR-Cys 3. The mutation into another residue might disturb this repeat and consequently any function this repeat might have.

Conservation

The wild-type residue is very conserved, but a few other residue types have been observed at this position too.

Your mutant residue was among the residues at this position observed in other sequences. This means that homologous proteins exist with the same residue type as your mutant at this position and this mutation is possibly not damaging to the protein.

Domains

Interpro Domain	Gene Ontology Term	Broad Gene Ontology Term
Tumour Necrosis Factor Receptor 1A IPR020419 (http://www.ebi.ac.uk/interpro/entry/IPR020419)	Tumor Necrosis Factor-Activated Receptor Activity GO:0005031 (http://www.ebi.ac.uk/QuickGO/GTerm?id=GO:0005031)	Molecular Transducer Activity GO:0060089 (http://www.ebi.ac.uk/QuickGO/GTerm?id=GO:0060089) Molecular_Function GO:0003674 (http://www.ebi.ac.uk/QuickGO/GTerm?id=GO:0003674)
Tnfr/Ngfr Cysteine-Rich Region IPR001368 (http://www.ebi.ac.uk/interpro/entry/IPR001368)	Protein Binding GO:0005515 (http://www.ebi.ac.uk/QuickGO/GTerm?id=GO:0005515)	Binding GO:0005488 (http://www.ebi.ac.uk/QuickGO/GTerm?id=GO:0005488) Molecular_Function GO:0003674 (http://www.ebi.ac.uk/QuickGO/GTerm?id=GO:0003674)

Hope Version 1.1.1

Interpro Domain	Gene Ontology Term	Broad Gene Ontology Term
Tumor Necrosis Factor Receptor 1A, N-Terminal IPR033993 (http://www.ebi.ac.uk/interpro/entry/IPR033993)	None	None

The mutated residue is located in a domain that is important for binding of other molecules and in contact with residues in a domain that is also important for binding. The mutation might disturb the interaction between these two domains and as such affect the function of the protein.

The mutated residue is located in a domain that is important for binding of other molecules and in contact with residues in a domain that is important for the activity of the protein. The mutation might affect this interaction and thereby disturb signal transfer from binding domain to the activity domain.

The mutated residue is located in a domain that is important for binding of other molecules. The mutated residue is in contact with residues in another domain. It is possible that the mutation disturbs these contacts.

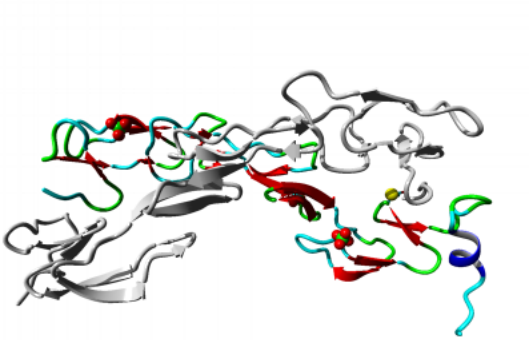
Amino Acid Properties

The wild-type and mutant amino acids differ in size.

The mutant residue is smaller than the wild-type residue.

The mutation will cause an empty space in the core of the protein.

Images



Download (/hope/yasara/2b880c31-81d1-4c55-bb0c-741ce23f4d97/25ILE_overview.png/)

Overview of the protein in ribbon-presentation. The protein is coloured by element; α-helix=blue, β-strand = red, turn=green, 3/10 helix=yellow and random coil=cyan. Other molecules in the complex are coloured grey when present.

Citation

Please use the following citation when referencing the results in your report:

Protein structure analysis of mutations causing inheritable diseases. An e-Science approach with life scientist friendly interfaces.

BMC Bioinformatics. 2010 Nov 8;11(1):548. DOI: 10.1186/1471-2105-11-548. (http://dx.doi.org/10.1186/1471-2105-11-548) PubMed: 21059217. (http://www.ncbi.nlm.nih.gov/pubmed/21059217)