EUROPEAN JOURNAL OF BIOMEDICAL AND PHARMACEUTICAL SCIENCES
http://www.ejbps.com

ISSN 2349-8870
Volume: 10
Issue: 10
91-101
Year: 2023

IN SILICO APPROACHES FOR ASSESSING THE IMPACT OF DELETERIOUS NSSNPS ON PROTEIN FUNCTION AND DRUG RESPONSE: BRIDGING PHARMACOGENOMICS AND PROTEIN STRUCTURE ANALYSIS

Sathishkumar A.[1*], Saisudha S.[2], Mohanapriya K.[1], Balasubramaniyam P. N.[3] and Ahamed Mueen M.[3]

[1]Assistant Professor, Department of Pharmaceutical Chemistry, Swamy Vivekananda College of Pharmacy, Elayampalayam, Tamil Nadu, India.
[2]Assistant Professor, Department of Pharmaceutics, Swamy Vivekananda College of Pharmacy, Elayampalayam, Tamil Nadu, India.
[3]Assistant Professor, Department of Pharmacognosy, Swamy Vivekananda College of Pharmacy, Elayampalayam, Tamil Nadu, India.

***Corresponding Author: Sathishkumar A.**
Assistant Professor, Department of Pharmaceutical Chemistry, Swamy Vivekananda College of Pharmacy, Elayampalayam, Tamil Nadu, India.

Article Received on 01/08/2023 **Article Revised on 22/08/2023** **Article Accepted on 12/09/2023**

ABSTRACT
Genetic variants allowing amino acid substitutions in proteins are termed nonsynonymous single-nucleotide polymorphisms (nsSNPs). It is possible for nsSNPs to have a large impact on human health due to their effect on protein structure and function. In the same way, polymorphisms in genes that encode drug targets play a profound role in drug response, either by directly affecting drug target function or by interfering with drug-target interaction. Pharmacogenomics explains how polymorphisms can have a significant impact on drug response, including the ability to produce potentially fatal adverse outcomes as well as equally severe lacks of therapeutic responses. During this review, deleterious nsSNPs are discussed both at the Protein structure and function are determined by a number of factors, including their effects on protein-protein interactions. Structure and properties of amino acids based on physicochemical properties are also used in order to predict how nsSNPs will affect function. Additionally, Pharmacogenomics in silico, which provides understanding of the phenotypic characteristics of diseases and individuals susceptibilities, is discussed. Using Our knowledge of these factors will soon allow us to determine the most effective treatment and drug therapy. In conclusion, this essay discusses how in silico approaches can be used to assess and develop drugs that are safe and effective.

KEYWORDS: Drug responses, drug-target proteins, pharmacogenomics, not-synonymous nucleotide polymorphisms, personalized medicine.

INTRODUCTION
In principle, pharmacogenomics offers A complete human genome sequence and a growing understanding of genetic polymorphisms have resulted in personalized medicines with reduced side effects.[1,2] ADRs result in thousands of avoidable deaths every year due to Millions of people have experienced severe adverse reactions in their lives.[3] It takes a long time and great effort to discover drugs, however, for a variety of reasons. One of these reasons is the individual's different reaction to drugs and unpredictable side effects.[2,4] Polymedication can result in dangerous drug interactions or adverse drug reactions (most commonly in geriatric patients) for certain psychiatric disorders, cardiopulmonary diseases, and cancer.[1,5] At present, it is not possible to predict how a medication will affect a particular individual. In drug-metabolizing enzymes and transporters, genetic polymorphisms can affect drug absorption, distribution, metabolism, elimination, and toxicity. By doing so, patients are more likely to be susceptible to treatment and have adverse reactions if drugs are administered.[2–6] The study of pharmacogenomics reveals, an understanding of how genetic polymorphisms are associated with medication response to custom-tailor drugs and therapies to match each person's genetic response to medications.[2–6] Pharmacogenomics can improve drug efficacy in several ways, including discovering new drug targets, overcoming drug resistance, and optimizing ADMET.[2–6]

To increase Oral bioavailability, efficacy, and potency of drugs, methods are used to identify hits, optimize them, and improve their affinity, selectivity, Stability, efficacy, and potency. The process of drug development begins As soon as a compound is identified that meets these requirements. The federal government approves the

marketing of products after performing molecular testing and animal trials. A poor ADMET profile is responsible for the failure of many new chemical entities during drug development. Computational advances have made it possible to predict ADMET properties concurrently with or before experiments, allowing large numbers of compounds to be evaluated before synthesized and analyzed.

About one in 1650 base pairs in the human genome is a single-nucleotide polymorphism (SNP), which constitutes the most common genetic variant. SNPs play a significant role in many diseases, so the identification of these genes and the understanding of how they relate to pathologic conditions will shed light on disease susceptibility and make targeted treatments more effective.

Non-synonymous SNPs (nsSNPs) are the most identifiable type of SNP, which affects the amino-acid sequence of the protein containing their SNP. A small percentage (0.1%) of these SNPs occur within coding regions of genes. Inheritance of susceptibility to disease is likely to be influenced by nonsynonymous SNPs that alter protein structure and function. DNA samples from general populations and the HapMap project have been used to identify most nsSNPs, which have been shown to affect disease and drug interaction.

Half of all genetic changes associated with human diseases are caused by nsSNPs. In light of the phenotypic consequences of these polymorphisms, they are considered deleterious nsSNPs. It is important to note that not all nsSNPs are linked to disease. A major challenge in medical genetics is to identify the phenotypes associated with nsSNPs, which are known as tolerant nsSNPs.[1,5,12,13] One of the most important problems today is to determine whether nsSNPs are disease-related. By distinguishing deleterious and tolerant nsSNPs, disease susceptibility can be assessed, as well as a genetic basis for human disease and pathogenesis.

Functional binding sites are often altered by nsSNPs because they disrupt protein structure and stability. Drug-receptor or drug-enzyme interactions may be affected by structural changes in receptors or active target enzyme sites. It is known that genetic variations in the genes encoding enzymes and transporters involved in drug metabolism can affect the ADMET profile of the drug, resulting in an increased amount of unmetabolized drugs located near the metabolic pathway. Among nsSNPs, some are located in voids or pockets (type P)[15–18], some are on convex surfaces (type S), and some are completely buried in protein structures (type I). In general, pathogenic nsSNPs are usually type P, while type I ones are rarely found. A number of studies have shown that 70% of disease-associated nsSNPs are associated with highly conserved residues within proteins. By considering physical differences in amino acid exchanges, evolution, and structural features, they are able to affect Functional, structural, and folding properties of proteins.[7,9,14,19–24]

An nsSNP with deleterious function needs to be identified. Identifying a single nsSNP is not sufficient to identify a disease or drug response connected to a target protein. In genome–phenotype analyses, groupings of genetic variants are particularly useful since they share a linkage disequilibrium. To identify loci in linkage disequilibrium across the entire genome, in silico methodologies are developed by integrating sets of nsSNPs. Using these approaches, we are more likely to identify polymorphisms of drug targets, enzymes that metabolize drugs, transporters for drugs, and genes that affect drug response, as well as genes and pathways that contribute to chronic disease etiology and pathogenesis.[7,23,25–29]

nsSNP can influence protein function

There are many ways in which nsSNPs affect cellular functions. It is true that nsSNPs can affect the function of proteins by altering their stability, protein-protein interactions, and other features as described below.

Active-site characteristics

A protein's ability to perform its function may be compromised if an amino acid in its active-site region is altered, thus leading to a disease phenotype. Catalytic residues can be critical sites, as can residues involved in binding ligands or interacting with partner molecules. These cases can be caused by loss of function or gain of function, altered binding specificities, or altered affinity of proteins, while the protein's stability remains unchanged. The biochemistry will, however, be affected by an alteration in the conformation of the active sites. A nsSNP may alter the characteristics of catalytic groups even if it occurs close to the active site.[30–32]

While Adapting a biochemical reaction is not the same as cancelling it, but altering its kinetics and physiological environment is the same thing of the cell (pH, temperature), leading to protein malfunctions. The Snyder-Robinson syndrome (SRS) is caused by three missense mutations in the spermine synthase gene (SMS), which encodes a dimeric protein (G56S, V132G, I150T). It is at position G56 that a dimer binding interface is located, at position V132 it is located, and at position I150 it is more distant from the dimer interface. There is almost no doubt that SRS results from mutations at site I150, even if there is no direct effect of its binding. There are mutations that can cause disease at site V132, or mutations that can be harmless. The hydrogen-bond effect and monomer stability are quite tolerable at site G56. It is expected that such mutations will cause disease since the SMS depends on dimer formation for its function.[30,31]

Accessibility of solvents

In several studies, nsSNPs that disrupt protein hydrophobic cores have been shown to affect protein function. It is important for protein stability that conserved residues are maintained hydrophobically with other residues in a protein core. If the original and mutated residues differ in volume, mass, and hydrophobicity, the nsSNP is more likely to be deleterious. This may cause protein aggregation, resulting in deleterious phenotypes, if a hydrophobic residue is introduced. Core hydrogen bonds are formed by hydrogen side-chains protein, a hydrophilic residue will destabilize folding processes. A specific nsSNP may have an impact on protein solubility, or change interactions with other molecules, there are different solvents or interfaces that can be used to interact with proteins (proteins with other proteins or proteins with DNA). A potential change to nsSNPs in highly conserved or charged surface patches should therefore be considered from a biological perspective. A nsSNP effect can then be modified by drug binding in cases like these.[9,10,27] Most nsSNPs are found at conserved sites within proteins, and only a few are completely buried inside proteins. It is noted that 88% of disease-related nsSNPs reside in surfaces pockets or interior voids of proteins, and do not follow a similar pattern.

Aggregation, folding, and flexibility of proteins

An amino-acid side-chain's intramolecular interactions and burial inside a molecule dictate the way proteins fold. As a result of protein folding, in this structure, helices, strands, turns, and coils comprise the main structural domains in an energetically favorable manner through the protein's secondary structure. In order for the tertiary structure of proteins to be stable, hydrogen bonds are formed between the polar groups of the main chains and hydrogen bonds between the side chains. It is possible for properties of residues at critical folding positions to change due to a 38 nsSNP that changes their biophysical properties. As a result, alterations in the folding parameters are possible. Mutants are often less stable even if they are correctly folded, or even if they are structurally stable, the protein's function may be compromised in some cases. In addition to disulfide bonds[42] and van der Waals[43], electrostatic interactions[44] may also be affected by the nsSNP[40], due to the disruption of the most important conformational stabilizing forces. Protein flexibility can also be altered by nsSNPs.[30,45]

It is common for proteins to undergo slight conformational changes when they perform a particular function. Mutations can alter protein flexibility, which governs protein conformational changes required for protein function. Conformational flexibility has a significant influence on protein aggregation. Inactivation results from breaking up and destabilizing native protein structures, revealing aggregation-prone regions.[46] Protein activation results from disruption and opening of native protein structures. Protein molecules are irreversibly interacting during aggregation. Even minor alterations such as nsSNPs may affect the solubility and aggregation propensity of proteins because proteins' sequences are a major factor in their aggregation propensity, at least in part. There are more alternate hydrophilic/hydrophobic stretches in disease-related proteins than in average proteins and more foldability and isoelectric point as compared to average proteins. The molecular mechanism of protein aggregation linked to neurodegenerative disorders like Alzheimer's disease is implicated in Protein aggregation is triggered by mutations. In a variant protein, the required quantities of secretion are not achieved at the cell membrane during cystic fibrosis because it can't fold correctly in the endoplasmic reticulum. Mutations can also lead to disorders such as emphysema that occur when proteins are improperly transported.

Amino acid substitutions

It has been estimated that half of the mutated sites interact strongly with amino acids; contact energies of residues are similar, except for allosteric residues. Remainders with special characteristics can cause a different class of mutations because of their unique characteristics. Protein secondary structures commonly contain glycine and proline residues, and their substitution leads to destabilization or a refolding of the protein. There are 39 most commonly modified amino acids in secondary structures, including arginine and glycine. There is an increased likelihood of disease-related nsSNPs in patients with tryptophan, tyrosine, and cysteine residues.[37,49]

Interactions between proteins, proteins and DNA, and proteins and membranes

Interactions between proteins and DNA are involved in most cellular processes, including transcription, signal transduction, and transport. In many cases, changes in amino acids at the interface can influence binding, affecting phenotypes and increasing disease susceptibility.[50]

Residue substitutions affect binding recognition and specificity when they are used in networking processes. Small side-chains replaced by bulky side-chains in a narrow binding pocket may completely or partially block the entrance of a partner group to the binding pocket. A large proportion of residues in proteins with a limited number of interactions are involved in their function because proteins with a limited number of interactions evolve slowly. The conservation of nsSNPs associated with disease is therefore more probable. There are often binding hot spots at protein–protein interfaces, which are polar or charged residues surrounding water, blocking hydrophobic residues that play the biggest role in binding. Several mutations that alter the electrostatic properties of protein surfaces can result in alterations in the folding or stability of the protein, as well as alterations in the affinity and specificity of its binding to its partners or ligands and, ultimately, the functionality

of the protein. As well, nsSNPs located near protein–DNA interfaces can affect DNA regulation54. Finally, nsSNPs located near protein-membrane interfaces can affect signal-processing across membranes, channels and pumps, and cellular adhesion.

Myelin protein zero monomers can break down when they are no longer able to interact with one another, causing Charcot-Marie-Tooth neuropathy. Allosteric regulation may also occur when Binding occurs between proteins. Mutations in the binding interface of pantothenate kinase can lead to inherited pantothenate kinase-associated neurodegeneration. Proteins can also be altered through mutations, causing them to interact differently. Mutations in the crystallin family of genes have altered the affinity of human crystallins for partners. These erroneous interactions result in congenital cataracts.

Localization within the cell

Proteins function and interact with their biological partners in a specific environment provided by subcellular localization. If a protein is located in the wrong subcellular location, it can damage DNA transcription factors can be interfered with by other proteins there. Altering protein expression will disrupt the normal cell cycle and lead to disease. A protein subcellular delocalization causes the disease in 1% of cases. There is evidence that some nsSNPs within the signaling peptide have the potential to cause a different subcellular localization for the native protein than the nsSNP within the signaling peptide would provide. Families with Deficits in factor X and hypoparathyroidism often have nsSNPs in their signal peptides.

PTMs (post-translation modifications)

After translation, PTMs modify protein side chains covalently. Amino acids are citrullinated, which alters their chemical properties or adds functional groups (phosphorylation) to them. PTMs alter the structure, localization, and function of proteins by creating or disrupting covalent bonds, a process critical to almost all cell signalling and cellular plasticity and dynamics. In some cases, nsSNPs cause the destabilization of proteins, changes in their interactions, or changes in their catalytic properties when substituting residues that cannot be modified. A number of cancer60 and immune deficiency61 phenotypes have been associated with detrimental nsSNPs in PTM sites.

Therefore, any change in these features caused by an nsSNP is likely to have an effect on the structure and stability of the protein, which in turn may affect the protein's function, increasing disease susceptibility. It is therefore crucial to improve Predicting functional nsSNP effects to facilitate drug discovery and development and their consequences.

Identifying functional nsSNPs and predicting their effects

There have been many nsSNPs identified in proteins that cause diseases. In silico methods are being developed to an approach to predicting how a set of proteins will react to deleterious mutations on databases combined with biochemical data on nsSNPs. An evolutionary sequence conservation analysis, Amino acid structure and physicochemical properties are the most common methods behind the development of these techniques.

Methods can be developed using compounds with physicochemical properties

Size, flexibility, and polarity of side chains influence protein folding, stability, function, and protein–protein interactions. Due to these distinct amino-acid properties, a substitution can be predicted for its impact and deleted mutations can be identified based on their compatibility. In addition, nsSNPs can alter the physicochemical properties of the original residue, disrupt the interacting interface, or affect protein stability and dynamics.[62]

Methods based on structure

Protein structure databases are used to match input sequences with protein structures. A number of predictive structure-based methods can model substitutions Since they do not require detailed information at the atomic level, they use the structure of a homologous protein rather than the input sequence's exact structure. In addition to solvent accessibility, secondary structure, active sites, and free energy differences between amino acids, these methods are used for investigating a number of structural properties. Substituting the nsSNP site may stabilize, enhance affinity, maintain residue-residue contacts, and enhance flexibility of the receptor ligand complex. Mutations and disease phenotypes must be fully understood with structural information. In order to determine if a specific nsSNP will When comparing a protein's three-dimensional (3D) structure with its normal folding or structure, mapping it can have a significant impact on the protein's normal folding or structure performed. Whether the amino-acid replacement destroys essential contacts that maintain structure, destabilizes electrostatic interactions, or interacts with ligands can be determined by mapping the nsSNP into the known three-dimensional (3D) structure [9]. In order to predict functional nsSNP effects on structure, in silico approaches have been useful due to the difficulty in eliciting 3D structures experimentally.

Using homology-based sequence analysis

Phylogenetic trees and sequence homology are used as evolutionary properties in sequence homology-based methods. Based on multiple sequence alignments, these methods determine the probability of substitutions and differentiate between tolerant mutations and disease-causing mutations. In natural selection, amino acids that are critical to protein function, stability, and interaction are retained among species. It is therefore expected that

highly conserved residues will evolve under strong selective pressure among homologous proteins.[7] The representative structure of such evolutionary conserved residues will be mapped. Using these methods, active sites and functional interfaces of proteins can be identified by extracting functionally important residues. It has been found that disease-related mutations are often targeted at Residues with high conservation and functional importance, including those that are found in enzyme active centres and those which play a key role in maintaining protein stability, especially those located in buried positions. There has been a finding Positions that are evolutionarily conserved are more prone to deleterious nsSNPs based on disease-associated mutation data and phylogenetic data across multiple species. Genes have a selective effect on organisms, neutral effects on organisms, and deleterious effects on organisms, which can all be repaired during evolution. Mutations are predicted to have an effect on protein function in about 26%–32% of cases.[34]

Sequential or structural features are commonly used to predict functional nsSNPs. The measurement of sequence similarity does not directly correlate with protein function, so methods based on sequence homology can result in incorrect annotations. It has been demonstrated that structural analyses are more accurate. However, studies based solely on structural comparisons have also produced ambiguous results.[9–12] Some proteins catalyze chemical reactions that are distinct from each other, but others catalyze similar reactions that are the same. If more than 70% of a protein's sequence is identical, sequence analysis can provide an accurate functional indication. In the case of a structural homology and a direct comparison of active sites, there is a greater probability of functional assignment by using structure. Binding and/or catalysis of a protein are commonly associated with a particular 3-D arrangement of residues. Understanding protein 3D structure provides valuable insights into how the protein functions and interacts. Only 1% of sequences have been experimentally analyzed for structure. Numerous studies have demonstrated that combining multiple sequence alignments with structural information of proteins can predict Protein function and deleterious nsSNPs.[9–12,21,34]

There are In silico approaches are becoming more prevalent being implemented into web servers to distinguish between deleterious nsSNPs that cause protein disorders and neutral polymorphisms that don't affect phenotypes. Nine, thirteen, seventeen, twenty-three,[27–29], [64–68] are the most representative.

1. In order to identify phenotypic changes associated with amino acid substitutions, Using sequence homology, SIFT sorts intolerants from tolerants.
2. Multiple sequence alignment is performed by Polymorphism Phenotyping (PolyPhen). Based on sequence analysis, the substitution site was characterized, two amino-acid variants were calculated for the position-specific independent count profile, and structural parameters and contacts were calculated. In this model, sequence information, evolutionary properties, and 3D structural information are used to predict whether an nsSNP will affect protein function, and it is optimized when structural information is available. This web server provides annotations for more than 11,000 nsSNPs. As PolyPhen maps amino-acid substitutions to proteins' 3D structures, it analyzes whether the substitution might destroy a protein's hydrophobic core, electrostatic interactions, ligand interactions, or other important features. The study analyses several structural parameters as well as several contact parameters in order to determine whether the substitution may destroy these features.

3. Based on the structure and sequence analysis of nsSNPs, SNPs3D assigns molecular functional effects. Models of gene-pathway-disease interaction are provided at the molecular level as well as various disease/gene relationships. Using this server, you can identify the genes that are likely to contribute to a specific disease, correlate the sets of candidate genes, and analyze how nsSNPs might affect the function of normal proteins.
4. The SNPeffect 4.0 tool analyzes the human proteome for disease variants and polymorphisms by means of molecular characterization and annotation. This technique improves the accuracy of the prediction of the functional role of local nsSNPs by combining sequence-based and structure-based bioinformatics tools.
5. By mapping nsSNPs onto protein sequences, functional pathways, and comparative protein structure models, the large-scale annotation of coding nsSNPs (LS-SNP) predicts where nsSNPs are causing the effect. SNPs can be detected within genes, haplotypes, and pathways using the results.
6. As part of Structure SNP (StSNP), nsSNPs are mapped and modeled on protein structures and linked to metabolic pathways in a web server. By examining possible disease-related pathways resulting from a particular nsSNP and comparing them with current molecular structure data, diseases can be linked to the current molecular structure information.
7. In AUTO-MUTE, nsSNPs are analyzed based on their disease potential using knowledge-based computational mutagenesis.

In these methods, the variant taken into account is different in terms of its nature and properties. Evolutionary information is sometimes used to support them. Various parameters relate to protein structure and/or function. Prediction performance would be enhanced by knowing the 3D structure of proteins. Methods that rely solely on sequence-level information often perform poorly, while others that incorporate structural and functional information perform well. Several studies have compared amino acid residue types, protein structures, and environments of amino acid substitution and examined whether they affect prediction performance. The best method could not be determined based on all of the evaluation measures.

Discovering new drugs in modern times

It is important that after identifying disease-associated nsSNPs, functional analyses are performed, the molecular mechanisms of disease phenotyping are understood, and drugs that target these mutants are developed. It has been shown that certain nsSNPs can affect the effectiveness and disposition of drugs. This type of nsSNP can have an impact on drug response and whether drugs cause adverse reactions if they occur within candidate drug-target proteins.

An ADMET profile, target potency, and selectivity determine the success of a drug. Historically, drugs were discovered by synthesizing compounds and screening them against a battery of in vivo biological screens, followed by further testing for ADMET properties of the candidates. It takes years of rigorous preclinical testing for new compounds to make it to clinical trials, but only about 8%.[70] According to estimates, 30% of active compounds do not show efficacy, and 50% are withdrawn from the pipeline because of toxicity.[71] In order to increase the success of clinical trials, ADMET properties must be evaluated early. Several factors prevent early determination of these factors, such as a lack of high-throughput in vitro assays, excessive amounts of compound needed for in vivo studies, and inability to predict some human toxicity by animal models. With the use of efficient technologies such as high-throughput screening, virtual screening, de novo design, ADMET screening in silico, structure-based drug design, and ADMET screening in silico, bioinformatics, genomics, proteomics, and combinatorial chemistry have revolutionized drug discovery in the 21st century.

To selectively treat diseases without producing side effects, drugs are developed targeting specific genes and molecular mechanisms in order to use a target-based approach to drug discovery. Simulating a chemical compound in silico and designing chemicals that may interfere with it is done using in silico techniques. Modern drug design can be accomplished with the help of these tools at any stage of the process.

1. Identification of targets. A potential drug target must be identified, along with its role in the disease, before drug discovery can begin. A drug target is a molecule involved in a signaling pathway that is specific to a particular disease condition, such as a gene or protein. Inhibiting an active region with a drug or promoting a molecule altered by a disease is one way of inhibiting or enhancing it with a drug. By using bioinformatics methods, such as homology-based analyses, ligand-based analyses, structures-based analyses, and HTS analyses, specific molecular targets and patient groups are identified.
2. Validation of targets. In order to demonstrate the benefit of a drug target, it must first show that it will improve or eliminate a particular phenotype. A wide range of approaches may be used for in silico characterization, including genetic pathway mapping, protein-protein interactions, disease locus mapping, and subcellular localization prediction.[74]
3. By identifying drugs from hit to lead, we ensure that they possess desirable ADMET properties, higher potency, physiochemical/metabolic properties suggestive of reasonable in vivo pharmacokinetics, and desirable adverse activities. Hit-to-lead analysis can be conducted using high-throughput biochemical, cellular, and natural product assays, structure-based design, and virtual HTS (vHTS).
4. The process of lead optimization involves refining a confirmed hit's chemical structure to improve its drug properties. Target affinity and selectivity are optimized in lead structures. Structure-based ADMET is currently aided by docking techniques.[75]
5. Whenever an active substance is to be used as a human medicine, preclinical studies with and without animals are required to limit risks. Pharmaceuticals, toxicologies, preformulation, formulation analyses, and pharmacokinetics all play a role in preclinical testing.[75]
6. Finding treatments that work in humans is made faster and safer through clinical trials. A defined period of time is set for collecting health data on patients with predetermined characteristics. Preventive, screening, diagnostic, and treatment trials are categorized as prevention, screening, diagnostic, treatment, and quality of life trials.
7. An NDA must be submitted and FDA approval must be obtained. Moreover, the purpose of the NDA is to demonstrate that the medicine is safe and effective for humans by offering information from previous years of research and making manufacturing and labeling proposals.

In silico Drug discovery

A variety of nsSNPs affect drug-target protein functions including G-protein-coupled receptors, enzymes, ion channels, and functions associated with detoxification pathways, which contribute to a variety of responses to therapeutic agents. Polymorphisms of various genes can be identified using in silico methods such as SIFT and PolyPhen. Computational approaches can be used to investigate potential therapeutic targets for diseases caused by nsSNPs. It is also possible to identify future targets by studying current therapeutic targets of drugs. With the use of bioinformatics tools, it is possible to identify new therapeutic targets and discover novel drugs. The goal of this is to improve the efficacy and safety profiles of drugs, both during the drug discovery and development process. In silico approaches such as ligand- and structure-based drug design, de novo design, and homology modeling can be employed for the identification of new selective, effective, and safe compounds to the development of drug candidates for clinical trials, depending on the availability of information about the drug target and potential drug.[76,77]

Designing from scratch and screening virtually

By using computer algorithms, new ligands are selected based on their structures predicted to have biological

activity on a given target. The ligand-based approach assumes that similar structures have similar biological activities.[78,79] It is divided into structural- and ligand-based methods. It is adequate to use these methods when the structure of a protein is not known since they operate only on known active ligands. Protein-ligand interactions and binding affinities can be predicted using docking, a computational tool for structure-based drug design. The docking patterns identified will be used to identify suitable conformations, and the scores will determine the affinity of those conformations. An important docking program is AutoDock, which can be found on pages 81 and 82. In Budapest, Hungary, Virtua Drug Ltd provides DockingServer, a complete web server.

Methods of design from scratch

A genetic algorithm is used to enhance the specified properties of a molecule to create new molecules with drug-like properties. The de novo design method, which is complementary to HTS83, creates chemical structures from scratch by assembling molecular fragments instead of searching in large collections of physically available compounds for bioactive molecules. A de novo design approach typically involves assembling molecules from fragments, evaluating their biological activity using computational functions, and optimizing the resulting compound. A de novo design method can either be based on the 3D structure of a ligand-binding cavity (ligand-based method) or on some comparison between candidate compounds and known reference ligands (ligand-based method). Docking programs can be used to build combinatorial libraries of molecules and perform virtual screenings in order to synthesize combinatorial libraries in drug discovery. The pharmacophore is a spatial arrangement of the features necessary for a molecule to interact with a specific drug target receptor in a specific manner, as an alternative to docking methods and as a complementary modeling method to 3D structures at the atomic level. In the PharmMapper webserver, potential drug targets are identified by annotating all target information from TargetBank, BindingDB, DrugBank, and potential drug target databases (namely PharmTargetDB) with the help of an annotated pharmacophore database. There are databases that contain information about genes associated with diseases, genetic variations, drugs, and 3D structures of pharmaceuticals.

Simulation of homology

Based on a related homologous protein's experimental 3D structure and its amino-acid sequence, an atomic-resolution model of the target protein is constructed. Sequences that fall below a 20% sequence identity can have very different structures when they have no similarity in their structure.[87,88] Protein structures appear to be more conserved than protein sequences among homologues.

The DNA microarray

This method measures thousands of genes in parallel, allowing for identification of potential targets for therapeutic intervention by generating clues to how they function. A change in gene expression can also be monitored following treatment with drugs.[89]

vHTS

As part of the vHTS application, small molecule compounds are compared against databases of protein targets to determine which compounds bind very strongly. An extract of a specific compound from the database can be used for further testing if there is a "hit" with it.[88,90]

Optimizing drug leads

Modifying the sequence and secondary structure of the potential drug is the purpose of this process. The exploration of related compounds to the lead candidate can be enhanced using software tools. Based on the available data on known ligands and the 3D structure of the target, various ligand- and structure-based methods are used to estimate the activity.[78,79]

Biological activity and bioavailability of drugs (Profile of ADMET)

Drugs are characterized by these characteristics. To understand how the ADMET drug profile is influenced by molecular features, quantitative structure–activity relationships (QSARs) have been traditionally used. A complex biological process can be modeled using statistical approaches based on experimental data and statistical approaches. It is possible to complement QSAR studies with structure-based methods as a result of the availability of the 3D structures of ADMET proteins. There have been several recent studies that demonstrate that in silico ADMET predictions can greatly reduce the number of compounds that need to be synthesized to achieve a drug's desired biochemical and/or physicochemical profiles.[91,92]

In drug discovery, genomic advances have led to a shift away from analyzing single-target interactions and toward analyzing multi-target interactions. It has been demonstrated that in silico methods offer the advantage of delivering new drug candidates more quickly and at a lower cost, according to a number of researches studies.[72,88,91,92] In the pharmaceutical industry and other industries, such as the consumer goods industry, there is great potential for the use of non-animal alternatives to ensure the safety of drugs.

Future developments and conclusions

Protein sequences and functions are primarily affected by nsSNPs. While nsSNPs constitute a small fraction of coding SNPs (8-12%), their impact on gene function is profound. The nsSNPs associated with disease predisposition compose about 30% of the whole. It is known to cause diseases most commonly when nsSNPs affect catalytic and regulatory sites and ligand-binding

areas in proteins with 3D structures. As a result of structural mutations, buried amino-acid residues, hydrogen bonds, salt bridges, and S–S bonds change sizes and charges in the protein core. Protein aggregation and aberrant folding occur as a result of these changes.

Drug efficacy and disposition are known to be affected by certain nsSNPs. The presence of these nsSNPs could influence drug response and whether they cause adverse reactions because these polymorphisms can be found in candidate drug-target proteins. Pharmacog-enomic studies comparing gene expression between normal and affected individuals can be performed to gain a better understanding of drug response, which can be attributed to multiple genes rather than a single-gene mutation. A pharmacogenomic screening of selected polymorphisms in clinical trial subjects would also reduce the size, speed, and cost of clinical trials because the subjects may have genetic profiles that would result in adverse reactions to a drug under evaluation or ineffective responses. Drug action mechanisms and novel therapeutic targets are already being identified by in silico pharmacogenomics. The genetics of chronic pain is a powerful tool for identifying relationships between the two. Among the results is the discovery that inward rectifying potassium channels are able to modulate analgesic responses to multiple drugs, which have been linked to opioid-induced hyperalgesia. Drug therapy benefitted from pharmacogenomics in the sense that it could be used to develop individualized medicines, develop new drugs, and regulate them. This challenge is not easy and is a complex one, as pharmacogenomics itself needs to deal with individual variations in drug therapy. However, we are only just getting started.

Functional nsSNPs could therefore be used to optimize drug therapy by improving effectiveness, diagnostics, and reducing toxicity. The most challenging aspect of future drug development will be determining genetic-environmental relationships, ethnicity, and inheritance patterns. Medications can be prescribed based on the genetic profile of each patient group if the goals of pharmacogenomics are satisfied, so that each group receives treatment based on its genetic profile, achieving maximum therapeutic benefit with minimal, tolerable adverse effects through appropriate medication targeting the right target at the right dose.

REFERENCES

1. Shastry BS. Pharmacogenetics and the concept of individualized medicine. Pharmacogenomics J, 2006; 6: 16–21.
2. Ma Q, Lu AY. Pharmacogenetics, pharmacogenomics, and individual- ized medicine. Pharmacol Rev, 2011; 63(2): 437–459.
3. Pharmacovigilance: ensuring the safe use of medicines. In: WHO Policy Perspectives on Medicines. Geneva: World Health Organization; 2004. Available from: http: //apps.who.int/medicinedocs/pdf/s6164e/s6164e. pdf. Accessed.
4. Suvarna BS. Role of pharmacogenomics in the drug development. Kathmandu Univ Med J (KUMJ), 2009; 7(2–26): 172–176.
5. Shastry BS. Role of SNPs and haplotypes in human disease and drug development. In: Ozkan M, Heller MJ, Ferrari M, editors. Micro/ Nano Technology in Genomics and Proteomics. New York: Springer, 2006; 2: 447–458.
6. Pang GS, Wang J, Wang Z, Lee CG. Predicting potentially functional SNPs in drug-response genes. Pharmacogenomics, 2009; 10(4): 639–653.
7. Ramensky V, Bork P, Sunyaev S. Human non-synonymous SNPs: server and survey. Nucleic Acids Res, 2002; 30: 3894–3900.
8. González-Castejón M, Marín F, Soler-Rivas C, Reglero G, Visioli F, Rodríguez-Casado A. Functional non-synonymous polymorphisms prediction methods: current approaches and future developments. Curr Med Chem, 2011; 18: 5095–5103.
9. Sunyaev S, Ramensky V, Koch I, Lathe W 3rd, Kondrashov AS, Bork P. Prediction of deleterious human alleles. Hum Mol Genet, 2001; 10(6): 591–597.
10. Wang Z, Moult J. SNPs, protein structure, and disease. Hum Mutat, 2001; 7: 263–270.
11. Yue P, Moult J. Identification and analysis of deleterious human SNPs. J Mol Biol, 2006; 356(5): 1263–1274.
12. Ng PC, Henikoff S. Predicting the effects of amino-acid substitutions on protein function. Annu Rev Genomics Hum Genet, 2006; 7: 61–80.
13. Masso M, Vaisman II. Knowledge-based computational mutagenesis for predicting the disease potential of human non-synonymous single-nucleotide polymorphisms. J Theor Biol, 2010; 266(4): 560–568.
14. Dimmic MW, Sunyaev S, Bustamante CD. Inferring SNP function using evolutionary, structural, and computational methods. Pac Symp Biocomput, 2005: 382–384.
15. Ung MU, Lu B, McCammon JA. E230Q Mutation of the catalytic subunit of cAMP-dependent protein kinase affects local structure and the binding of peptide inhibitor. Biopolymers, 2006; 81: 428–439.
16. Rignall TR, Baker JO, McCarter SL, et al. Effect of single-active- site cleft mutation on product specificity in a thermostable bacterial cellulase. Appl Biochem Biotechnol, 2002; 98–100: 383–394.
17. Erdin S, Ward RM, Venner E, Lichtarge O. Evolutionary trace annota- tion of protein function in the structural proteome. J Mol Biol, 2010; 396(5): 1451–1473.
18. Dean PM, Lewis RA, editors. Molecular Diversity in Drug Design. Dordrecht: Kluwer Academic Publishers; 1999.
19. Bao L, Zhou M, Cui Y. nsSNP analyzer: identifying disease-associated nonsynonymous single nucleotide

polymorphisms. Nucleic Acids Res, 2005: W480–W482.
20. Stitziel NO, Tseng YY, Pervouchine D, Goddeau D, Kasif S, Liang J. Structural location of disease-associated single-nucleotide polymorphisms. J Mol Biol, 2003; 327(5): 1021–1030.
21. Sunyaev S, Ramensky V, Bork P. Towards a structural basis of human non-synonymous single-nucleotide polymorphisms. Trends Genet, 2000; 16(5): 198–200.
22. Sunyaev SR, Lathe III WC, Ramensky VE, Bork P. SNP frequencies in human genes: an excess of rare alleles and differing modes of selection. Trends Genet, 2000; 16(8): 335–337.
23. Yue P, Melamud E, Moult J. SNPs3D: candidate gene and SNP selection for association studies. BMC Bioinformatics, 2006; 7: 166.
24. Cheng TM, Lu YE, Vendruscolo M, Lio P, Blundell TL. Prediction by graph theoretic measures of structural effects in proteins arising from non-synonymous single-nucleotide polymorphisms. PLoS Comput Biol, 2008; 4: e1000135.
25. Frederic MY, Lalande M, Boileau C, et al. UMD-predictor, a new prediction tool for nucleotide substitution pathogenicity-application to four genes: FBN1, FBN2, TGFBR1, and TGFBR2. Hum Mutat, 2009; 30: 952–959.
26. Bromberg Y, Rost B. SNAP: predict effect of non-synonymous poly-morphisms on function. Nucleic Acids Res, 2007; 35(11): 3823–3835.
27. Ryan M, Diekhans M, Lien S, Liu Y, Karchin R. LS-SNP/PDB: anno-tated non-synonymous SNPs mapped to Protein Data Bank structures. Bioinformatics, 2009; 25(11): 1431–1432.
28. Reumers J, Maurer-Stroh S, Schymkowitz J, Rousseau F. SNPeffect v2.0: a new step in investigating the molecular phenotypic effects of human non-synonymous SNPs. Bioinformatics, 2006; 22(17): 2183–2185.
29. Wang P, Dai M, Xuan W, et al. SNP function portal: a web database for exploring the function implication of SNP alleles. Bioinformatics, 2006; 22(14): E523–E529.
30. Zhang Z, Teng S, Wang L, Schwartz CE, Alexov E. Computational analysis of missense mutations causing Snyder-Robinson syndrome. Hum Mutat. 2010; 31(9): 1043–1049.
31. Zhang Z, Norris J, Schwartz C, Alexov E. In-silico and in vitro inves-tigations of the mutability of disease-causing missense mutation sites in spermine synthase. PLoS One. 2011; 6(5): Article ID e20373.
32. Alexov E. Numerical calculations of the pH of maximal protein stability: the effect of the sequence composition and three-dimensional structure. Eur J Biochem. 2004; 271(1): 173–185.
33. Loladze VV, Ermolenko DN, Makhatadze GI. Thermodynamic con-sequences of burial of polar and non-polar amino-acid residues in the protein interior. J Mol Biol. 2002; 320: 343–357.
34. Chasman D, Adams RM. Predicting the functional consequences of non-synonymous single-nucleotide polymorphisms: structure-based assessment of amino-acid variation. J Mol Biol. 2001; 307: 683–706.
35. Teng S, Michonova-Alexova E, Alexov E. Approaches and resources for prediction of the effects of non-synonymous single-nucleotide poly-morphism on protein function and interactions. Curr Pharm Biotechnol. 2008; 9(2): 123–133.
36. Dill KA. Dominant forces in protein folding. Biochemistry. 1990; 29(31): 7133–7155.
37. Dobson RJ, Munroe PB, Caulfield MJ, Saqi MAS. Predicting deleteri-ous nsSNPs: an analysis of sequence and structural attributes. BMC Bioinformatics. 2006; 7(1): 217.
38. Pace CN, Fu H, Fryar KL, et al. Contribution of hydrophobic interactions to protein stability. J Mol Biol. 2011; 408(3): 514–528.
39. Khan S; Vihinen M. Spectrum of disease-causing mutations in protein secondary structures. BMC Struct Biol. 2007; 7(1): 1–18.
40. Thomas PJ, Qu BH, Pedersen PL. Defective protein folding as a basis of human disease. Trends Biochem Sci. 1995; 20: 456–459.
41. Dill KA, Ozkan SB, Weikl TR, Chodera JD, Voelz VA. The protein folding problem: when will it be solved? Curr Opin Struct Biol. 2007; 17(3): 342–346.
42. Eriksson AE, Baase WA, Zhang XJ, et al. Response of a protein structure to cavity creating mutations and its relation to the hydrophobic effect. Science. 1992; 255: 178–183.
43. Betz SF. Disulfide bonds and the stability of globular proteins. Protein Sci. 1993; 2(10): 1551–1558.
44. Horovitz A, Serrano L, Avron B, Bycroft M, Fersht AR. Strength and cooperativity of contributions of surface salt bridges to protein stability. J Mol Biol. 1990; 216: 103–144.
45. Karplus M, Kuriyan J. Molecular dynamics and protein function. Proc Natl Acad Sci U S A. 2005; 102(19): 6679–6685.
46. Valerio M, Colosimo A, Conti F, et al. Early events in protein aggregation: molecular flexibility and hydrophobicity/charge interaction in amyloid peptides as studied by molecular dynamics simulations. Proteins. 2005; 58(1): 110–118.
47. Wong P, Fritz A, Frishman D. Designability, aggregation propensity and duplication of disease-associated proteins. Protein Eng Des Sel. 2005; 18: 503–508.
48. Keage HA, Carare RO, Friedland RP, et al. Population studies of spo-radic cerebral amyloid angiopathy and dementia: a systematic review. BMC Neurol. 2009; 9: 3.
49. Vitkup D, Sander C, Church G. The amino-acid mutational spectrum of human genetic disease. Genome Biol. 2003; 4(11): R72.

50. Gong S, Blundell TL. Structural and functional restraints on the occur-rence of single-amino-acid variations in human proteins. PLoS ONE. 2010; 5(2): e9186.
51. Goh KI, Cusick ME, Valle D, Childs B, Vidal M, Barabasi AL. The human disease network. Proc Natl Acad Sci U S A. 2007; 104(21): 8685–8690.
52. Teng S, Madej T, Panchenko A, Alexov E. Modeling effects of human single-nucleotide polymorphisms on protein-protein interactions. Biophys J. 2009; 96(6): 2178–2188.
53. Ma B, Elkayam T, Wolfson H, Nussinov R. Protein-protein inter-actions: structurally conserved residues distinguish between bind-ing sites and exposed protein surfaces. Proc Natl Acad Sci U S A. 2003; 100: 5772–5777.
54. Wright JD, Lim C. Mechanism of DNA-binding loss upon single-point mutation in p53. J Biosci. 2007; 32(5): 827–839.
55. Kwa LG, Wegmann D, Brugger B, Wieland FT, Wanner G, Braun P. Mutation of a single-residue, β-glutamate-20, alters protein-lipid inter-actions of light harvesting complex II. Mol Microbiol. 2008; 67(1): 63–77.
56. Hanemann CO, D'Urso D, Gabreëls-Festen AA, Müller HW. Mutation-dependent alteration in cellular distribution of peripheral myelin pro-tein 22 in nerve biopsies from Charcot-Marie-Tooth type 1A. Brain. 2000; 123(5): 1001–1006.
57. Laurila K, Vihinen M. Prediction of disease-related mutations affecting protein localization. BMC Genomics. 2009; 10: 122.
58. Moosawi F, Mohabatkar H. Computer-assisted analysis of subcellular localization signals and post-translational modifications of human prion proteins. J Biomedical Sci Eng. 2009; 2: 70–75.
59. Ryu GM, Song P, Kim KW, Oh KS, Park KJ, Kim JH. Genome-wide analysis to predict protein sequence variations that change phosphorylation sites or their corresponding kinases. Nucleic Acids Res. 2009; 37(4): 1297–1307.
60. Radivojac P, Baenziger PH, Kann MG, Mort ME, Hahn MW, Mooney SD. Gain and loss of phosphorylation sites in human cancer. Bioinformatics. 2008; 24(16): i241–i247.
61. Vogt G, Vogt B, Chuzhanova N, Julenius K, Cooper DN, Casanova JL. Gain-of-glycosylation mutations. Curr Opin Genet Dev. 2007; 17: 245–251.
62. Grantham R. Amino-acid difference formula to help explain protein evolution. Science. 1974; 185: 862–864.
63. Bao L, Cui Y. Prediction of the phenotypic effects of non-synonymous single-nucleotide polymorphisms using structural and evolutionary information. Bioinformatics. 2005; 21(10): 2185–2190.
64. Lichtarge O, Sowa ME. Evolutionary predictions of binding surfaces and interactions. Curr Opin Struct Biol. 2002; 12(1): 21–27.
65. Ben-Shimon A, Eisenstein M. Computational mapping of anchoring spots on protein surfaces. J Mol Biol. 2010; 402(1); 259–277.
66. Adzhubei IA, Schmidt S, Peshkin L, et al. A method and server for predict-ing damaging missense mutations. Nat Methods. 2010; 7(4): 248–249.
67. Flanagan SE, Patch AM, Ellard S. Using SIFT and PolyPhen to predict loss-of-function and gain-of-function mutations. Genet Test Mol Bio-markers. 2010; 14(4): 533–537. Erratum in: Genet Test Mol Biomarkers. 2010; 14(5): 730.
68. De Baets G, van Durme J, Reumers J, et al. SNPeffect 4.0: on-line prediction of molecular and structural effects of protein-coding variants. Nucleic Acids Res. 2012; 40: D935–D939.
69. Gleeson MP, Hersey A, Montanari D, Overington J. Probing the links between in vitro potency, ADMET and physicochemical parameters. Nat Rev Drug Discov. 2011; 10: 197–208.
70. Marchetti S, Schellens JH. The impact of FDA and EMEA guide-lines on drug development in relation to Phase 0 trials. Br J Cancer. 2007; 97: 577–581.
71. Wang J, Urban L. The impact of early ADME profiling on drug discov-ery and development strategy. Drug Discovery World. 2004; 5: 73–86.
72. Kermani BG. Pharmacogenomics. In: Markel S, León D, editors. Silico Technologies in Drug Target Identification and Validation. Vol 6 of Drug Discovery Series. Boca Raton, FL: CRC Press, Taylor & Francis Group, 2006: 323–343.
73. Hughes JP, Rees S, Kalindjian SB, Philpott KL. Principles of early drug discovery. Br J Pharmacol. 2011; 162(6): 1239–1249.
74. Goodnow RA. Hit and lead identification: integrated technology-based approaches. Drug Discov Today Technol. 2006; 3: 367–375.
75. Olejniczak K, Guenzel P, Bass R. Preclinical testing strategies. Drug Inf J. 2001; 35: 321–336.
76. Bajorath J. Computational analysis of ligand relationships within target families. Curr Opin Chem Biol. 2008; 12: 352–358.
77. Klebe G. Virtual ligand screening: strategies, perspectives and limitations. Drug Discov Today. 2006; 11: 580–594.
78. Wilson GL, Lill MA. Integrating structure-based and ligand-based approaches for computational drug design. Future Med Chem. 2011; 3(6): 735–750.
79. Sukumar N, Das S. Current trends in virtual high throughput screening using ligand-based and structure-based methods. Comb Chem High Throughput Screen. 2011; 14(10): 872–888.
80. Tresadern G, Bemporad D. Modeling approaches for ligand-based 3D similarity. Future Med Chem. 2010; 2(10): 1547–1561.
81. Cheng T, Li Q, Zhou Z, Wang Y, Bryant SH. Structure-based virtual screening for drug discovery: a problem-centric review. AAPS J. 2012; 14(1): 133–141.
82. Meng XY, Zhang HX, Mezei M, Cui M. Molecular docking: a powerful approach for structure-based

drug discovery. Curr Comput Aided Drug Des. 2011; 7(2): 146–157.
83. Hartenfeller M, Schneider G. De novo drug design. Methods Mol Biol. 2011; 672: 299–323.
84. Boda K, Johnson AP. Molecular complexity analysis of de novo designed ligands. J Med Chem. 2006; 49: 5869–5879.
85. Wolber G, Seidel T, Bendix F, Langer T. Molecule-pharmacophore superpositioning and pattern matching in computational drug design. Drug Discov Today. 2008; 13(1–2): 23–29.
86. Steindl M, Schuster D, Wolber G, Laggner C, Langer C. High throughput structure-based pharmacophore modeling as a basis for successful paral-lel virtual screening. J Comput Aided Mol Des. 2006; 20: 703–715.
87. Wieman H, Tøndel K, Anderssen E, Drabløs F. Homology-based modeling of targets for rational drug design. Mini Rev Med Chem. 2004; 4: 793–804
88. Satyajit D, Sovan S, Kapil S. Computer aided drug design – a new approach in drug design and discovery. Int J Pharm Sci Rev Res. 2010; 4(3): 146.
89. Liu Y, Verducci JS. Medicinal chemistr y and dr ug design. In: Ekinci D, editor. Microarray Analysis in Drug Discovery and Bio-marker Identification. InTech; 2012.
90. Subramaniam S, Mehrotra M, Gupta D. Virtual high throughput screening (vHTS) – A perspective. Bioinformation. 2008; 3(1): 14–17.
91. Gleeson MP, Hersey A, Hannongbua S. In-silico ADME models: a general assessment of their utility in drug discovery applications. Curr Top Med Chem. 2011; 11: 358–381.
92. Moroy G, Martiny VY, Vayer P, Villoutreix BO, Miteva MA. Towards in silico structure-based ADMET prediction in drug discovery. Drug Discov Today. 2012; 17(1–2): 44–55.