ELSEVIER

Journal of Molecular Graphics and Modelling 20 (2001) 199–218

Journal of Molecular Graphics and Modelling

www.elsevier.com/locate/JMGM

Prediction of potential toxicity and side effect protein targets of a small molecule by a ligand–protein inverse docking approach

Y.Z. Chen*, C.Y. Ung

Department of Computational Science, National University of Singapore, Blk S17, Level 7, 3 Science Drive 2, Singapore 117543, Singapore

Received 11 November 2000; received in revised form 18 April 2001; accepted 19 April 2001

Abstract

Determination of potential drug toxicity and side effect in early stages of drug development is important in reducing the cost and time of drug discovery. In this work, we explore a computer method for predicting potential toxicity and side effect protein targets of a small molecule. A ligand–protein inverse docking approach is used for computer-automated search of a protein cavity database to identify protein targets. This database is developed from protein 3D structures in the protein data bank (PDB). Docking is conducted by a procedure involving multiple conformer shape-matching alignment of a molecule to a cavity followed by molecular-mechanics torsion optimization and energy minimization on both the molecule and the protein residues at the binding region. Potential protein targets are selected by evaluation of molecular mechanics energy and, while applicable, further analysis of its binding competitiveness against other ligands that bind to the same receptor site in at least one PDB entry. Our results on several drugs show that 83% of the experimentally known toxicity and side effect targets for these drugs are predicted. The computer search successfully predicted 38 and missed five experimentally confirmed or implicated protein targets with available structure and in which binding involves no covalent bond. There are additional 30 predicted targets yet to be validated experimentally. Application of this computer approach can potentially facilitate the prediction of toxicity and side effect of a drug or drug lead. © 2001 Elsevier Science Inc. All rights reserved.

Keywords: ADME/TOX; Adverse drug reactions; Computer-aided drug design; Drug safety evaluation; Drug target; Docking; Scoring function; Protein receptors; Toxicity prediction; Aspirin; Gentamicin; Ibuprofen; Indinavir; Neomycin; Penicillin; Tamoxifen; Vitamin C

1. Introduction

Drug safety evaluation is an important issue in new drug discovery [1]. One of the objectives of animal tests and clinical trials is the assessment of toxicity and side effect of a drug candidate. These tests and trials have consumed a large percentage of the time and money spent on drug development [2–4]. Given a low success rate of drug candidates, detection of potential toxicity and side effect in early stages of drug development can potentially save money and time by focusing resources on drug leads and candidates likely to be safe to patients. This has lead to efforts towards the development of experimental techniques for molecular analysis and high-throughput screening of toxicological effects as an early assessment tool [5–9]. Computer programs have also been developed for prediction of toxicity and metabolism using statistically derived structure-toxicity/metabolism relationships [10–12]. While showing promising potential, these techniques also have their limits [9]. Hence, introduction of alternative methods that complement these techniques may provide additional tools to facilitate evaluation of drug safety profiles.

In this work, we explore the use of a computer method for finding toxicity and side effect protein targets of a small molecule. Knowledge of these targets combined with that of proteomics [13] and pharmacokinetic profile [14] can facilitate the assessment of potential toxicity and side effect of a drug candidate. In addition, it also gives the molecular mechanism of predicted effect. Our method is based on a ligand–protein inverse docking strategy such that a small molecule is attempted to dock to known ligand-binding pockets of each of a group of proteins associated with potential toxicity and side effect. A protein is considered as a candidate of toxicity and side effect target of a small molecule if that molecule can be docked into the protein. Here the term "inverse" is used because the method is used for finding proteins that will fit with a specific organic compound, rather than finding ligands that fit with a specific receptor protein.

The algorithm for docking a small molecule to a protein is based on an approach similar to that used in ligand–protein flexible docking [15]. In this approach, single or multiple

* Corresponding author. Tel.: +65-874-6877; fax: +65-774-6756.
E-mail address: yzchen@cz3.nus.edu.sg (Y.Z. Chen).

1093-3263/01/$ – see front matter © 2001 Elsevier Science Inc. All rights reserved.
PII: S1093-3263(01)00109-7

small molecules in single or multiple conformations are attempted to dock into a receptor site so as to find potential ligands. Testing results on a number of docking algorithms showed that they are capable of finding ligands and binding conformations at a receptor site close to experimentally determined structures [16–27]. Because of their capability in identifying ligands and binding conformations, these algorithms are expected to be equally applicable to an inverse docking process for finding multiple protein targets to which a small molecule can bind or weakly bind.

In contrast to other computer toxicity prediction methods [10–12], the method introduced in this work is unique in that it is a receptor-based approach. The potential toxicity and side effect of a drug molecule is probed by the search of its potential protein targets related to toxicity and side effects. These potential targets are selected from the following two criteria: the first is whether or not this molecule can be docked into each of these proteins, and the second is whether the binding of the docked molecule to its receptor is strong enough as evaluated by a molecular mechanics ligand–protein interaction energy. The inverse docking method also differs from other ligand–protein docking methods. The former is related to the docking of one small molecule to multiple proteins in a database, while the later is related to the docking of one or multiple small molecules to a single protein target [15–27]. The later is not readily applicable in identification of potential protein targets of a small molecule.

The proposed inverse docking strategy requires a sufficient number of potential toxicity and side effect related proteins of known 3D structure. There are 1040 entries of 38 types of proteins in the protein data bank (PDB) that fall into this category. Of these entries, there are 274 entries for human and mammalian protein. At present, the number of entries in the PDB increases at a rate of well over 100 per month. Approximately 17% of the entries in the PDB have unique sequence [28]. Introduction of high-throughput methods is expected to enable structural determination of 10 000 proteins with unique sequence within 5 years [29]. Thus, the number of proteins with known 3D structure is approaching a meaningful level to cover a diverse set of potential toxicity and side effect targets.

All small molecules appear to bind to cavities of proteins and nucleic acids. Thus, to facilitate computer-automated inverse docking search for finding protein targets of a small molecule, a protein cavity database has been developed from relevant protein entries in the PDB [30]. This database contains models of the individual cavity in each protein. A cavity model is a cluster of overlapping spheres that fill-up that cavity [31]. Docking to a cavity of protein can be conducted by matching ligand atom positions with the center of these spheres. This matching directs the initial positioning and orientation of the ligand in the actual cavity of the host protein. Optimization and scoring of the docked structure then follows [15–27]. The 3D structures of the proteins in the cavity database are needed in the optimization and scoring of the docked structures. Therefore, these are also included in the database.

An automated inverse docking procedure INVDOCK is conducted to search this database to identify protein targets of each of a selected number of clinical drugs. A drug is flexibly docked into each cavity by a procedure involving multiple conformer shape-matching alignment of the molecule to the cavity followed by molecular-mechanics torsion optimization and energy minimization on both the ligand and the binding region of the receptor. Potential protein targets are selected based on a new scoring scheme that performs binding competitive analysis in addition to the evaluation of molecular mechanics ligand–protein interaction energy. This scoring scheme has been tested on a number of ligand–protein complexes from PDB [30] and the computed root mean square deviation for most of these complexes is comparable to that from other docking studies. The identified toxicity and side effect protein targets of these drugs are then compared with available experimental findings so as to evaluate the applicability of the inverse docking procedure.

2. Methods

2.1. Selection of toxicity and side effect related proteins

Proteins associated with potential toxicity and side effect are selected based on available information from the medical biochemistry literature [32]. It is known that deficiency and inhibition of some proteins important in normal cellular function may result in toxicity or side effects. Examples of these proteins are those involved in key cellular metabolism processes such as glycolytic pathway, amino acid and nucleotide metabolism, urea cycle, citric acid cycle and oxidative phosphorylation in mitochondria [32,33]. Some drugs can be transformed into toxic agents by metabolizing enzymes such as cytochrome P450 and glutathione *S*-transferase [34,35]. Adverse reactions to drugs may also result from immune mechanism [36]. In addition, drug interaction with some proteins involved in signaling pathways may also produce toxic as well as anticancer effect [37]. Table 1 gives a list of proteins that are known to be associated with potential toxicity and side effect. Information about the physiological function, ligand/substrate, effect of deficiency/inhibition, and the number of corresponding PDB entries for these selected proteins [38–76] is also given in Table 1. There are a total of 1040 PDB entries corresponding to 38 types of proteins.

2.2. Protein cavity database

Each cavity entry in this database is derived from the corresponding toxicity and side effect related PDB entry by a procedure described below: Ligands and waters in

Table 1
Known toxicity and side effect-causing protein targets of drugs

Protein (number of total PDB entry)	Physiological function	Ligand (substrate (S), product (P), activator (A), inhibitor (I))	Site of action	Effect of deficiency/inhibition	References
Hexokinase (7)	Glycolysis	Glucose (S), ATP (S), glucose 6-P (P/I), ADP (P)	Cytosol	Hemolytic anemia	[38]
Phosphoglucose isomerase (6)	Glycolysis	Glucose (S), ATP (S), glucose 6-P (P/I), ADP (P)	Cytosol	Hemolytic anemia	[38]
Phosphofructokinase (5)	Glycolysis	Glucose (S), ATP (S), glucose 6-P (P/I), ADP (P)	Cytosol	Hemolytic anemia	[38]
Triose phosphate isomerase (7)	Glycolysis	Dihydroxyacetone phosphate (S/P), glyceraldehyde 3-P (S/P),	Cytosol	Hemolytic anemia	[38]
Phosphoglycerate kinase (7)	Glycolysis	1,3-bisphosphoglycerate (S), ADP (S), 3-phosphoglycerate (P), ATP (P)	Cytosol	Hemolytic anemia	[38]
Pyruvate kinase (12)	Glycolysis	Phosphoenolpyruvate (S), ADP (S), pyruvate (P), ATP (P)	Cytosol	Hemolytic anemia	[38]
UDP-galactose 4-epimerase (16)	Galactose metabolism	UDP-galactose (S), glucose (P), UDP (P)	Cytosol	Galactosemia	[39]
Glycogen phosphorylase (42)	Glycogenolysis	Glycogen (S), glucose 1-phosphate (P)	Liver	Hypoglycemia	[40]
Aconitase (14)	Citric acid cycle	Glycogen (S), glucose 1-phosphate (P)	Mitochondria	Neurotoxicity	[41]
Fumarase (6)	Citric acid cycle	Succinate (S), FAD (S), fumarate (P), FADH2 (P)	Mitochondria	Neurotoxicity induced by enzyme inhibition and deficiency	[42]
Cytochrome oxidase (3)	Oxidative phosphorylation	Azide (I), cyanide (I)	Mitochondria	Inhibition of electron transport may induce muscle weakness, fatigability, nausea, and aggravation of latent myocardial angina	[43]
Cytochrome C (115)	Oxidative phosphorylation		Mitochondria	Increased sensitivity to cell death signals triggered by TNF-α	[44]
ATP synthase (11)	ATP synthesis	ADP (S), ATP (P)	Mitochondria	Inhibition of electron transport may induce muscle weakness, fatigability, nausea, and aggravation of latent myocardial angina	[43]
Superoxide dismutase (67)	Free radical scavenger	Superoxide (S), NADPH (S), H_2O_2 (P), NADP (P), gentamicin (I)	Cytosol	Genotoxicity, nephrotoxicity	[45,65]
Catalase (24)	Free radical scavenger	H_2O_2 radical (S), water (P)	Mitochondria	Genotoxicity	[46]
Peroxidase (155)	Free radical scavenger	H_2O_2 (S)	Mitochondria	Oxidative stress	[47]
Acyl-CoA dehydrogenase (12)	Fatty acid oxidation	Acyl-CoA (S), *trans*-enoyl-CoA (P)	Mitochondria	Hypoglycemia, hyperammonemia, tissue fatty change, hypoketonemia	[48]
Phenylalanine hydroxylase (9)	Degradation of phenylalanine	Phenylalanine (S), tyrosine (P), tetrahydrobiopterin (S), quinonoid dihydrobiopterin (P)	Cytosol	Phenylketonuria	[49]
Arginase (11)	Urea cycle	Arginine (S), ornithine (P), urea (P)	Cytosol	Hyperammonemia	[32]
Ornithine transcarbomoylase (5)	Urea cycle	Ornithine (S), carbamoyl phosphate (S), citrulline (P)	Mitochondria	Hyperammonemia	[32]
Carbamoyl phosphate synthetase (9)	Urea cycle	Ammonia (S), carbon dioxide (S), carbamoyl phosphate (P)	Mitochondria	Hyperammonemia	[32]
Hypoxanthine–guanine phosphoribosyl transferase (7)	Nucleotide biosynthesis	Hypoxanthine (S), PRPP (S), guanine (S), inosine (P), guanylate (P)	Mitochondria	Hyperuricemia	[32]
Carbonic anhydrase (144)	pH regulation in blood and kidney	Carbon dioxide (S), bicarbonate ion (P)	Blood and kidney	Nephrotoxicity	[50]

Table 1 (*Continued*)

Protein (number of total PDB entry)	Physiological function	Ligand (substrate (S), product (P), activator (A), inhibitor (I))	Site of action	Effect of deficiency/inhibition	References
HLA (79)	Antigen recognition	Peptide (L)	Cell membrane	Allergy	[51]
Alcohol dehydrogenase (46)	Detoxification of alcohol	Alcohol (S), aldehyde (P)	Cytosol	Hepatoxicity	[52]
Glutamate dehydrogenase (10)	Amino acid degradation	Glutamate (S), NAD (S), ammonia (P), NADH (P)	Mitochondria	Nephrotoxicity	[53]
Glutathione synthetase (3)	Anti-oxidation	Gamma-Glu–Cys (S), glycine (S), ATP (S)	Cytosol	Genotoxicity	[54]
Glutathione *S*-transferase (89)	Anti-oxidation	Glutathione (S), leukotriene A4(LTA4) (S)	Cytosol	Genotoxicity	[55]
Catechol *O*-methyltransferase (1)	Catechol metabolism	Norepinephrine (S), vanillyl mandelic acid (P), neomycin (A)	Mitochondria	Increased catecholaminergic enzyme inhibition-induced neurotransmission. Activation causes neurotoxicity related to accumulation of dopamine in neural systems	[56,74]
Fatty acid-binding protein (26)	Lipid uptake	Fatty acid (L)	Blood stream	Altered distribution of lipid profile and adipogenesis	[57]
Acetylcholinesterase (41)	Neurotransmission	Acetylcholine (S)	Nerones	Cholinergic toxicity	[58]
Aldehyde dehydrogenase (15)	Alcohol metabilism	Acetaldehyde (S), acetyl-CoA (P)	Liver	Mitochondrial toxicity	[59]
Insulin (58)	Hormone that reduces glucose level	Insulin receptor (receptor)	Pancreas, blood stream	Drug-induced hyperglycemia	[60]
Retinoblastoma protein (Rb) (1)	Tumor suppression	Retinoblastoma-binding protein P46 (receptor)	Nucleus	Enzyme inhibition-induced malignant cell proliferation	[32]
Bcl-X protein (2)	Regulation of apoptosis	Voltage-dependent anion channel (VDAC) (binding protein), BAD protein (binding protein)	Cytosol	Inhibition-induced apoptosis (programmed cell death)	[32]
p53 Protein (22)	Regulation of apoptosis	Acetylation by CREB-binding protein (CRP) (protein binding), DNA (gene expression), TATA box-binding protein (TBP) (protein binding)	Cytosol, nucleus	Inhibition promoted malignant cell proliferation	[32]
Extracellular regulated kinase ERK (4)	Activation of AP-1 by Ha-Ras in ERK signaling pathway	Phosphorylation by MEK 1 protein (protein binding)	Cytosol	Activation by drug leading to stress	[32]
C-JUN N-terminal kinase (JNK) (1)	Regulation of JNK signaling pathway	Protein–protein interactions	Cytosol	Activation by drug leading to stress	[32]
G protein (15)	Signal transduction	GTP (S), GDP (P)	Cytosol	Inhibition causing large efflux of electrolytes and water, uncoupling of signaling cascade from hormone receptors	[32]
ICE-like cysteine protease	Apoptosis modulation	Peptides (S), aspirin (A)	Cytosol	Up-regulation activates apoptosis in cells	[61]
Butyrylcholinesterase	Serves as scavenger of toxins to protect acetylcholine-binding proteins	Aspirin (S)	Cytosol	Activation by some drugs such as aspirin and cocaine induce local anesthetics response	[62]
Gastric glycoproteins	Protection of gastric mucosal layer	Aspirin (I)	Mucosal layer of stomach	Inhibition causes drug-induced mucosal damage	[63]
Caspase	Regulation of apoptosis	Aspirin (A), gentamicin (A)	Cytosol	Activation causes induction of apoptosis in cells	[64]

N-acetyl-beta-D-glucosaminidase	Maintenance of renal function	*N*-acetyl-beta-D-glucosamine (NAG) (S)	Kidney	Nephrotoxicity	[65]
Gamma-glutamyl-transferase	Maintenance of liver and renal function	L-Gamma-glutamyl-*p*-nitroanilide (S)	Liver, kidney	Nephrotoxicity	[65]
Alkaline phosphatase	Protection against endotoxin	5-Phosphopolynucleotide (S), gentamicin (A)	Universal	Activation causes nephrotoxicity	[66]
Leucine aminopeptidase	An exopeptidase hydrolyzing the peptide bond adjacent to a free amino group	Peptides (S), gentamicin (A)	Kidney	Activation causes nephrotoxicity	[66]
Basolateral transporter (glucose transporter 1/GLUT1)	Glucose transport across plasma membrane of cells	Glucose (L), gentamicin (I)	Plasma membrane	Inhibition causes attenuation of stress response	[67]
Cathepsin B and L	Lysosomal proteases	Peptides (S), gentamicin (I)	Lysosomes	Inhibition causes reduction of protein catabolism in lysosomes	[68]
Na(+), K(+)-ATPases	Ion transport in renal	Gentamicin (I)	Plasma membrane	Inhibition causes nephrotoxicity	[69]
Lysosomal phospholipases A and C	Phospholipid hydrolysis	Phospholipid (S), gentamicin (I)	Lysosome	Inhibition causes nephrotoxicity in accumulation of phospholipid in lysosomes of kidney cortex	[70]
Tumor necrosis factor	A pleiotropic inflammatory cytokine	TNF-binding protein (protein binding), interferon beta (protein binding), ibuprofen (I)	Cytosol	Activation causes hypotension and metabolic acidosis	[71]
Multidrug transporter proteins MDR1 and MRP1	Transport of small ligands	ATP (S), ibuprofen (I)		Inefficiency of drug absorption across intestine	[72]
P-glycoproteins	Protection of intestine mucosa layer	ATP (S), indinavir (I), saquinavir (I), ritonavir (I)	Intestinal mucosal layer	Inhibition causes lower oral bioavailability of drug to targeted tissues	[73]
Tyrosine hydroxylase	Biosynthesis of neurotransmitter catecholamines	Tyrosine (S), tetrahydrobiopterin (S), neomycin (A)	Cytosol	Neurotoxicity related to accumulation of dopamine in neural systems	[74]
Tryptophane hydroxylase	Regulation of biosynthesis of neurotransmitors	Tryptophan (S), tetrahydrobiopterin (S), neomycin (A)	Cytosol	Neurotoxicity related to accumulation of dopamine in neural systems	[74]
Monoamine oxidase (MAO)	Regulation of biosynthesis of neurotransmitors	Monoamine (S), neomycin (A)	Cytosol	Neurotoxicity related to accumulation of dopamine in neural systems	[74]
NMDR receptor	Neural signal transmission	Neomycin (A)	Membrane-bound	Ototoxicity and excitotoxicity of the central nervous system	[75]
Phosphatidylinositol-specific phospholipase C	Signal transduction	Phosphatidylinositol (S), gentamicin (I)	Cytosol	Inhibition causes nephrotoxicity	[76]
P-glycoproteins	Protection of intestine mucosa layer	ATP (S), indinavir (I), saquinavir (I), ritonavir (I)	Intestinal mucosal layer	Inhibition causes lower oral bioavailability of drug to targeted tissues	[73]
Tyrosine hydroxylase	Biosynthesis of neurotransmitter catecholamines	Tyrosine (S), tetrahydrobiopterin (S), neomycin (A)	Cytosol	Neurotoxicity related to accumulation of dopamine in neural systems	[74]
Tryptophane hydroxylase	Regulation of biosynthesis of neurotransmitors	Tryptophan (S), tetrahydrobiopterin (S), neomycin (A)	Cytosol	Neurotoxicity related to accumulation of dopamine in neural systems	[74]
Monoamine oxidase (MAO)	Regulation of biosynthesis of neurotransmitors	Monoamine (S), neomycin (A)	Cytosol	Neurotoxicity related to accumulation of dopamine in neural systems	[74]
NMDR receptor	Neural signal transmission	Neomycin (A)	Membrane-bound	Ototoxicity and excitotoxicity of the central nervous system	[75]
Phosphatidylinositol-specific phospholipase C	Signal transduction	Phosphatidylinositol (S), gentamicin (I)	Cytosol	Inhibition causes nephrotoxicity	[76]

a PDB protein structure are first removed. The surface of this protein, as defined by Richards [77], is then generated. van der Waals surface of a solvent-accessible atom is generated using the respective parameter from the AMBER force field [78]. The in-ward facing surface covering the interface of van der Waals surfaces is computed by using a probe sphere of radius 1.4 Å. The whole protein surface is then coated and filled by a cluster of spheres by the method proposed by Kuntz et al. [31]. Further modification is made in our own implementation of the algorithm. For proteins complexed with one or more ligands, additional spheres of radius 1.6 Å are added to the position of each ligand atom. This is to ensure that known ligand-binding sites are adequately represented. The radius of 1.6 Å is chosen because it is the smallest van der Waals radius of the non-hydrogen atoms of a typical ligand. Moreover, spheres within 0.5 Å of each other are combined so as to reduce the size of a cluster. Each sphere is further checked for the extent of its surrounding space covered by protein atoms. The surrounding space is defined as a region within 15 Å of the center of the sphere. A sphere is considered as "covered" if more than 50% of the direction around the sphere is covered by protein atoms. The remaining spheres are defined as "non-covered".

In a cavity, there is usually a high concentration of "covered" spheres. One can thus extract a cavity from "covered" spheres highly concentrated in a particular region. For such a purpose, the "covered" spheres are divided into separate groups based on spatial separation of nearest neighbor spheres. A member in a group is close to ($<$3 Å) at least another member in the group. Each group of these "covered" spheres is thus considered as a cavity. In addition, "non-covered" spheres close to a "covered" sphere in a group are included in that group. This is to ensure that a cavity model sufficiently covers the cavity surface region where a tail of a ligand might be located. The total number of cavity entries of proteins of all species associated with potential toxicity and side effect is 2353.

2.3. Inverse docking procedure

An inverse docking procedure INVDOCK has been developed for automated search of every entry in a protein cavity database to find protein targets of a small molecule [30]. In the present study, the search is limited to protein entries relevant to potential toxicity and side effect. It is difficult to conduct cross-the-board analysis of the possible binding sites for a large number of cavity entries. Hence, all parts of each cavity are subject to docking. Docking to sites inside each cavity starts from known ligand-binding sites, followed by more interior sections and then remaining part. To save CPU time, the program proceeds to the next cavity entry when first successful dock is obtained without further searching other better binding modes within each cavity. All cavity entries are subjected to search unless the related protein has been identified as a target.

Although the optimum binding-mode is not specifically sought if the first successful dock is obtained, this may only affect the binding quality rather than the quantity. The INVDOCK search proceeds such that, unless the molecule is successfully docked in a particular position/orientation, it continues until all possible positions and orientations in a cavity are exhausted. Such a procedure ensures that all targets identified by an optimum search will be selected. Although the docked structure might not be optimum, it is less likely to be too far away from the optimal binding mode. This is because the prioritized search strategy is set-up such that it searches likely binding sites first followed by the more interior area of a cavity and then the rest. Our study indicates that this approach seems to dock a ligand to a site reasonably close to the experimentally observed positions in most of the ligand–protein structures studied [31].

Our own implemented algorithm for flexible ligand docking is similar to the multistep strategy approach proposed by Wang et al. [15]. Ligand conformers are sampled at a resolution similar to that of Wang et al. Conformers with intramolecular energy within 30 kcal/mol of the lowest energy are first selected. These conformers are further screened to select the lowest energy conformer and those with at least one torsion differing by 60° or more with each other. The force field used to compute intramolecular energy is described below. Docking of each sampled conformer to a cavity is by the following steps: First the ligand is aligned within the selected site by matching the position of each ligand atom with the center of spheres. Because of the relatively low-resolution nature of ligand conformation sampling, a certain degree of structural clash is allowed at this stage. Ligand atoms are allowed to be as close as 2.0 Å to a protein atom. Such a steric clash is to be released by a structural optimization and energy minimization procedure described below. This procedure also helps to improve the docking configuration.

A molecular-mechanics conformation optimization is conducted by a limited torsion space sampling of rotatable bonds in the ligand and those in the side-chain of the receptor amino acid residues at the binding site. Each rotatable bond is sampled in the range of $\pm 15°$. This is followed by 50 iterations of steepest decent Cartesian coordinate energy minimization on all ligand and protein atoms at the binding site so as to further optimize the ligand–protein complex. The selection of 50 iterations is based on the consideration of the balance between CPU time usage and sufficient optimization of ligand structure. Although higher number of iterations is more desirable, the cumulative CPU cost for searching thousands of proteins might be too high.

In both torsion optimization and energy minimization, AMBER force fields [78] are used for covalent bond, bond angle, torsion, and non-bonded van der Waals and electrostatic interactions. Morse potential [79], which is a function of donor–acceptor distance, is used to represent hydrogen bond terms [80,81]. This potential has been shown to give reasonable description of hydrogen bond energy and

dynamics in biomolecules [80,81]. The energy function is:

$$V = \frac{1}{2}\sum_{\text{bonds}} K_r(R - R_{\text{eq}})^2 + \frac{1}{2}\sum_{\text{angles}} K_\theta(\theta - \theta_{\text{eq}})^2 + \frac{1}{2}\sum_{\text{torsions}} V_n[1 - \cos(n(\phi - \phi_{\text{eq}}))] + \sum_{\text{H bonds}} [V_0(1 - \mathrm{e}^{-a(r-r^0)})^2 - V_0] + \sum_{\text{non bonded}} \left[\frac{A_{ij}}{r_{ij}^{12}} - \frac{B_{ij}}{r_{ij}^{6}} + \frac{q_i q_j}{\varepsilon_r r_{ij}}\right]$$

In this function, R, θ, and ϕ denotes bond length, angle and torsion angle, respectively; R_{eq}, θ_{eq} and ϕ_{eq} are taken as equilibrium bond length, angle and torsion angle, respectively, and their values are from the original PDB structure and the structure of the drug respectively; K_r and K_θ are covalent and bond angle bending force constant, respectively; V_n and n are torsion parameters; r is hydrogen bond donor–acceptor distance, and V_0, a and r_0 are hydrogen bond potential parameters.

In some cases, protein conformation is affected by ligand-binding. This binding-induced conformation change has been considered in a docking study that explores methods such as side-chain conformational sampling and protein ensemble generation [82]. Limited side-chain conformation sampling is considered in INVDOCK along with additional energy minimization for all atoms at the binding site. However, no attempt is made to explicitly sample ensemble of protein conformations. It can be seen from Table 1 that some proteins in PDB have multiple entries associated with different binding ligand or different mutants. Moreover, there are multiple conformations in most NMR structures. To a certain extent, INVDOCK search of these multiple forms of structures/conformations may serve as a partial sampling of the conformation for these proteins.

2.4. Scoring

Docked structures are selected by evaluation of a ligand–protein interaction energy function ΔE_{LP} composed of the same hydrogen bond and non-bonded terms as those used for structure optimization. We have analyzed ΔE_{LP} of 34 PDB ligand–protein complexes. These include a variety of proteins such as aspartyl proteases, serine proteases, metalloproteases, glucose-binding proteins and others. Ligand-binding in each of these selected structures involves hydrogen bond and non-bonded interactions only. Our analysis of these ligand–protein complexes shows that the computed ΔE_{LP} roughly follows a linear relationship with the number of ligand atoms. It is generally below $\Delta E_{\text{Threshold}} = -\alpha N$ kcal/mol, where N is the number of ligand atoms and a is a constant ~1.0. Hence, this statistically derived energy value can be used empirically as a threshold for screening likely binders. ΔE_{LP} can be required to be lower than $\Delta E_{\text{Threshold}}$ when selecting successfully docked structures.

Ligand-binding is competitive in nature. A drug is less likely to be effective if it binds to its receptor non-competitively against natural ligands and, to some extent, other drugs that bind to the same receptor site. This binding competitiveness may be partially taken into consideration for cavities known as ligand-bound in at least one PDB entry. Ligands in PDB structures are known binders and they can thus be considered as "competitors" of a molecule that bind to the same receptor site. In addition to evaluation against $\Delta E_{\text{Threshold}}$, the computed ΔE_{LP} is also required to be roughly similar in value to the ligand–protein interaction energy of the corresponding PDB ligands that bind to in the same cavity in this or other relevant PDB entries.

3. Results and discussion

The clinical agents studied in this work include aspirin, gentamicin, ibuprofen, indinavir, neomycin, penicillin G, 4H-tamoxifen, and Vitamin C. These drugs are selected because they have been subjects of extensive investigation including the probing of toxicity and side effect protein targets of these drugs. However, these drugs are selected before the literature of the toxicity or side effects was searched.

The 3D structure of each of these clinical drugs as displayed in Figs. 1 and 2, is from either the ACD3D database of MDL [83] or Woodcock's molecular models database [84]. These structures are selected because they were model-built. Computer 3D model building programs typically generate structures that are virtually identical (with root mean square deviation of a few angstroms) to the corresponding X-ray crystal structure in 38–57% of the cases [85]. A flexible docking algorithm can be better tested on these model-built structures as they show more conformational diversity than crystal structures. Moreover, given that the number of available model-built molecular structures is much more than that of X-ray crystal structures, model-built structures might more likely be used in an inverse docking study. The INVDOCK search of toxicity and side effect targets for each of the selected set of clinical agents is restricted to human and mammalian proteins in this work, as they are more relevant to toxicity and side effect in human. The average CPU time is 12 days on a 250 MHz SGI R10000 Octane workstation for searching all the 1425 cavity entries of human and mammalian proteins associated with potential toxicity and side effect.

The relatively slower computing speed arises partly because of the following two differences between our program and those of other docking studies: (1) our program searches the whole cavity, while in most docking studies a predefined binding site is searched. The later is often much smaller than the former in many cases; (2) our optimization algorithm includes that for the protein atoms at the binding site, which requires substantially more CPU time than those

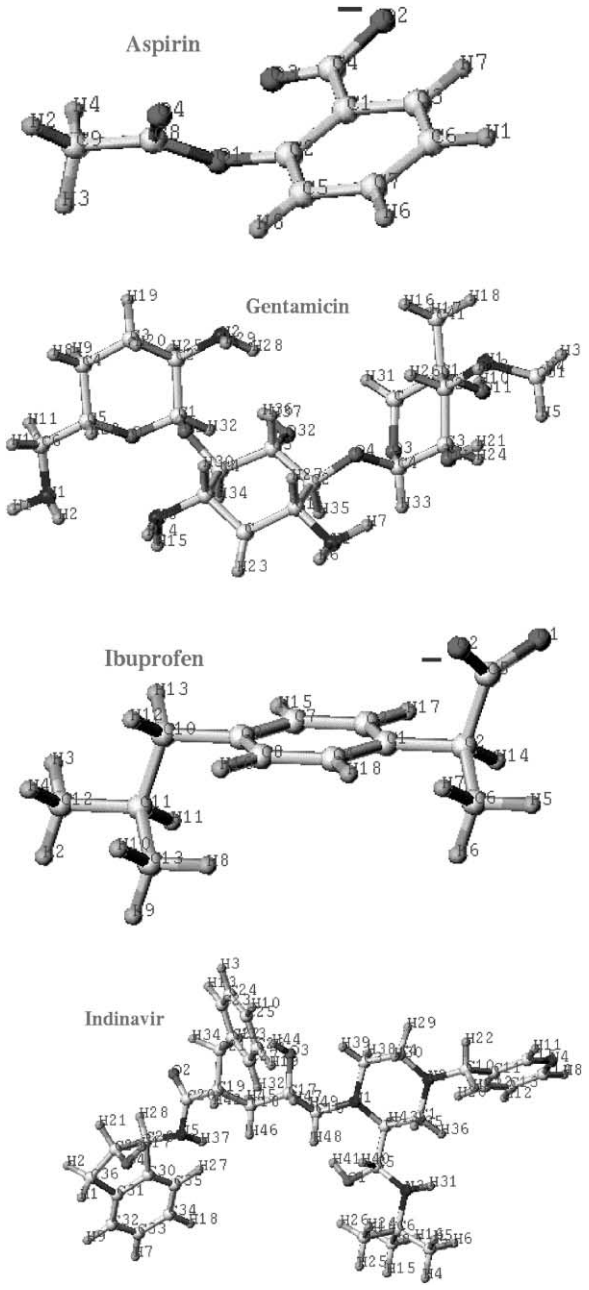

Fig. 1. Structure of therapeutic agents aspirin, gentamicin, ibuprofen and indinavir studied in this work.

in most docking studies in which the protein is held rigid. Further work is in progress in improving the computing speed of INVDOCK.

3.1. Potential toxicity and side effect protein targets of aspirin

Aspirin is widely used for their analgesic and anti-inflammatory properties. Aspirin overdose is one of the leading causes of accidental death in children. Aspirin induces toxic effects through a variety of ways [86] that include central stimulation of the respiratory center that results in hyperventilation, uncoupling of oxidative phosphorylation in mitochondria, interruption of glucose and fatty acid metabolism, alteration in capillary integrity, and alteration of platelet function and bleeding time.

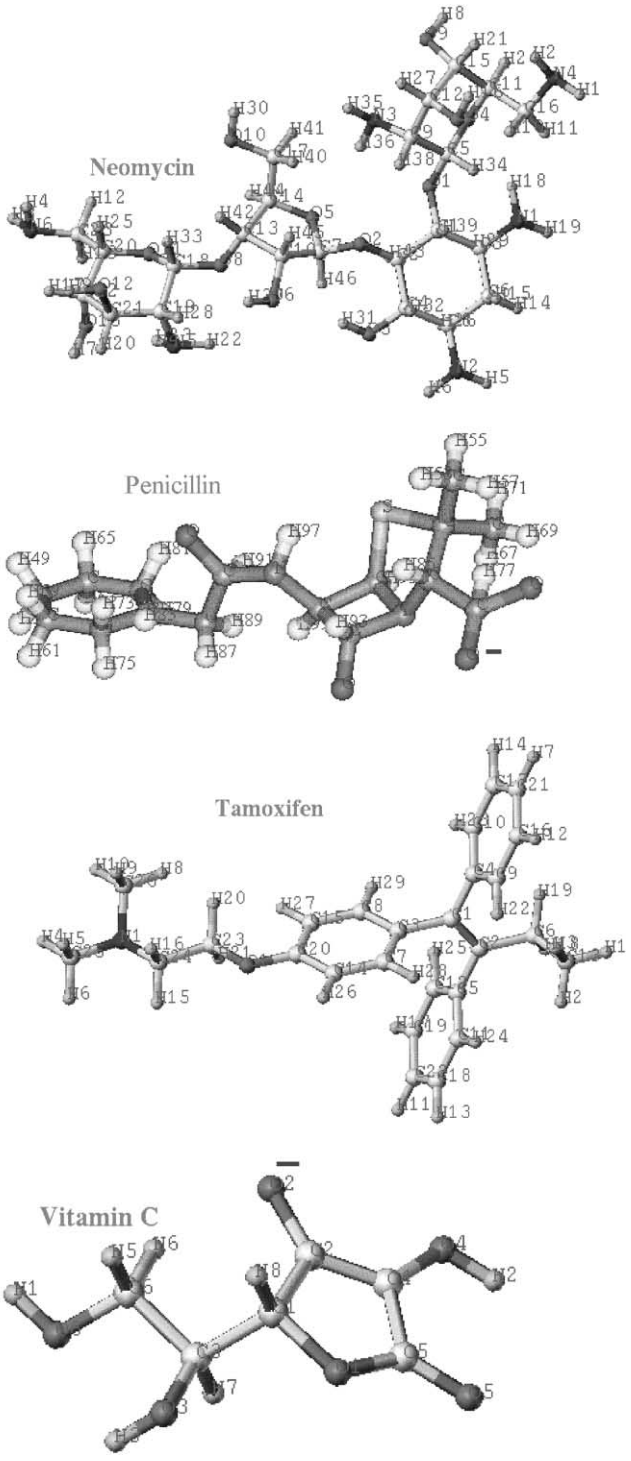

Fig. 2. Structure of therapeutic agents neomycin, penicillin G, 4H-tamoxifen and Vitamin C studied in this work.

Predicted toxicity and side effect targets from INVDOCK search is listed in Table 2. Also included in Table 2 is the computed ligand–protein interaction energy for each drug–

Table 2
Human and mammalian toxicity and side effect protein targets of aspirin identified from INVDOCK search

PDB	Protein	Experimental finding	Target status	Toxicity/side effect	E^a (kcal/mol)	Reference
1a42	Carbonic anhydrase II	Activation of enzyme activity that may lead to increase in plasma bicarbonate concentration	Implicated	Metabolic alkalosis (hypoventilation)	−37.6	[90]
1a6a	HLA-DR3	Change in HLA level	Implicated	Aspirin-induced asthma	−34.8	[91]
1a7c	Plasminogen activator inhibitor	Tissue-dependent response of protein	Implicated	Hypertension, thrombolysis	−27.8	[92]
1d6n	Hypoxanthine–guanine phosphoribosyltransferase			Excess uric acid in serum	−27.7	
1hdy	Alcohol dehydrogenase	Inhibition of activity	Confirmed	Increased blood alcohol level	−41.2	[87]
1hiq	Insulin	Tissue insensitivity to insulin	Implicated	Impaired glucose metabolism in insulin-sensitive cells	−41.8	[93]
1hmr	Fatty acid-binding protein	Increased binding capacity and protein content	Implicated	Effect on peroxisomal beta oxidation activity	−28.2	[94]
1mch	Immunoglobulin lambda light chain	Ig reactivities	Implicated	Allergic reaction to aspirin	−35.3	[95]
1pah	Phenylalanine hydroxylase			Phenylketonurea	−28.2	
2ant	Antithrombin	Irreversible acetylation of antithrombin	Confirmed	Blood coagulation, thrombolysis	−35.8	[89]
2hdh	L-3-Hydroxyacyl-CoA dehydrogenase	Reversible inhibition of enzyme activity	Confirmed	Effect on Reye's syndrome patients	−30.1	[88]

[a] E is the computed ligand–protein interaction energy for each aspirin–protein complex.

Table 3
Human and mammalian toxicity and side effect protein targets of gentamicin identified from INVDOCK search

PDB	Protein	Experimental finding	Target status	Toxicity/side effect	E^a (kcal/mol)	Reference
1a27	17-Beta-hydroxysteroid-dehydrogenase	Reduction of enzyme activity	Implicated	Caused alteration in the phosphatase activities of testes and accessory sex organs and a decline in the sperm count	−64.5	[99]
1ajr	Aspartate aminotransferase	Decreased activity	Implicated	Nephrotoxicity	−46.3	[97]
1azm	Carbonic anhydrase I	Inhibition	Confirmed	Nephrotoxicity	−49.4	[96]
1ch6	Glutamate dehydrogenase	Inhibitory effect	Confirmed	Nephrotoxicity	−57.3	[53]
4ald	Fructose-bisphosphate aldolase	Decreased enzyme activity	Implicated	Nephrotoxicity	−57.4	[98]
1hdy	Alcohol dehydrogenase			Hepatotoxicity	−53.2	
1pah	Phenylalanine hydroxylase			Phenylketonurea	−53.9	

[a] E is the computed ligand–protein interaction energy for each gentamicin–protein complex.

protein complex. This molecular mechanics-based energy might be useful to assess the binding enthalpy of a drug and to compare with that of other ligands that bind to the same receptor site. INVDOCK identified 11 potential targets. Three of these targets have been confirmed by experiments. Six others have been implicated by experiments. The activity of some of these proteins is changed by aspirin, which may be indicative of direct binding. Also, the level of other proteins has been changed by aspirin binding. Ligand-binding is known, in some cases, to regulate protein level. Therefore, it remains to be seen whether these six targets can be confirmed experimentally. The other two targets given in Table 2 are without relevant experimental data to either validate or invalidate them.

Experiment showed that aspirin increases alcohol level in blood via inhibition to alcohol dehydrogenase [87]. Our result indicates that aspirin binds to this protein and thus is consistent with this experimental finding. Aspirin is known to exert effects on Reye's syndrome patients via reversible inhibition of the activity of L-3-hydroxyacyl-CoA dehydrogenase [88]. It also acetylates antithrombin [89], which may lead to blood coagulation and thrombolysis. Both proteins are identified as a target by INVDOCK search.

Aspirin has been found to affect the activity of carbonic anhydrase II [90], which leads to an increased plasma bicarbonate concentration that might cause metabolic alkalosis. Aspirin is known to alter the HLA level and thus induce asthma [91]. Plasminogen activator inhibitor has been found to respond to aspirin [92], which might result in aspirin-induced hypertension and thrombolysis. Interaction of aspirin with insulin [93] may induce impairment of glucose metabolism in insulin-sensitive cells. Aspirin had been shown to increase binding capacity and content of fatty acid binding protein [94] thereby inducing peroximal beta oxidation. Reactivity of aspirin with immunoglobulin has been shown to cause allergy reaction [95].

The two other predicted targets that lack experimental validation/invalidation are hypoxanthine–guanine phosphoribosyltransferase and phenylalanine hydroxylase. Further investigation is needed to determine whether or not aspirin can bind to these proteins.

3.2. Potential toxicity/side effect protein targets of gentamicin

Gentamicin is an antibiotic and it is known to induce the following toxicity and side effects [86]: (a) ototoxicity to vestibular and cochlear cells; (b) nephrotoxicity due to proximal tubular damage and acute tubular necrosis; (c) competitive neuromuscular blockage if given rapidly intravenously together with other neuromuscular blocking drugs. The INVDOCK predicted toxicity and side effect targets of this drug are listed in Table 3, along with the computed ligand–protein interaction energy for each drug–protein complex. A total of seven potential targets are predicted by INVDOCK. Two of these proteins have been confirmed as a gentamicin target experimentally. The other two have been implicated by experiments. The rest are without available experimental finding to substantiate them.

Inhibition of carbonic anhydrase by gentamicin [96] is known to cause nephrotoxicity since carbonic anhydrase is a key enzyme that regulates the bicarbonate level in the blood. Loss of carbonic anhydrase activity results in pH regulation problems in the blood. In addition, gentamicin exerts inhibitory effect on glutamate dehydrogenase activity and thus glutamate synthesis, which leads to nephrotoxicity [53].

Three other INVDOCK identified toxicity targets of gentamicin are also implicated by experiments. For instance, the activity of two enzymes aspartate aminotransferase [97] and fructose-bisphosphate aldolase [98] has been found to decrease by gentamicin, which might lead to nephrotoxicity. The reduction of aspartate aminotransferase activity occurs at a higher dose of gentamicin [97], which seems to indicate weak interactions. Gentamicin is also known to cause an alteration in the phosphatase activities of testes and accessory sex organs and a decline in the sperm count [99]. The INVDOCK identified target responsible for this effect is 17-beta-hydroxysteroid-dehydrogenase [99]. Gentamicin causes reduction of the activity of this enzyme and thus affects steroidogenesis [99], which might contribute to its effect on accessory sex organs and on the decline of sperm count.

Table 4
Human and mammalian toxicity and side effect protein targets of ibuprofen identified from INVDOCK search

PDB	Protein	Experimental finding	Target status	Toxicity/side effect	E^a (kcal/mol)	Reference
1vid	Catechol *O*-methyltransferase			Hypertension	−33.8	
2ans	Adipocyte lipid-binding protein			Altered lipid distribution	−35.2	
1bo6	Estrogen sulfotransferase	Inhibitor	Confirmed	Sexual dysfunction	−33.8	[101]
4erk	Extracellular regulated kinase 2	Enhanced kinase activity	Implicated	Enhanced mitogenic effect	−32.9	[102]
5rla	Arginase	Decreased enzyme activity	Implicated	Hyperammonemia	−32.8	[103]

[a] E is the computed ligand–protein interaction energy for each ibuprofen–protein complex.

No report has been found to either implicate or invalidate the other two INVDOCK identified targets alcohol dehydrogenase and phenylalanine hydroxylase. Further investigation is needed to determine whether or not gentamicin can bind to these proteins.

3.3. Potential toxicity/side effect-causing protein targets of ibuprofen

Ibuprofen is a non-steroidal anti-inflammatory drug (NSAID) widely used for the relief of pain and inflammation. NSAIDs produce their therapeutic and main toxic effect by inhibiting cyclooxygenase, which in turn reduces the production of prostagladins [100]. Table 4 gives the list of INVDOCK predicted toxicity and side effect targets for this drug, along with the computed ligand–protein interaction energy for each drug–protein complex. Five targets are predicted in which one has been confirmed and two have been implicated by experiments. The other two predicted targets are yet to be validated or invalidated by experiment.

This drug is known to inhibit estrogen sulfotransferase leading to sexual dysfunction [101]. Estrogen sulfotransferase plays an important role in steroidogenesis involved in the production of steroidal sexual hormones. Inhibition of this enzyme by gentamicin is shown to exhibit adverse side effects particularly sexual dysfunction and disruption of hormone action [101]. INVDOCK identification of this protein as a toxicity target is consistent with results from this experimental study.

The two INVDOCK predicted targets that have been implicated by experiments are extracellular regulated kinase 2 (ERK2) [102] and arginase [103]. ERK2 plays an important role in intracellular signal transduction process that regulates cell proliferation. Ibuprofen has been shown to enhance the activity of ERK2 [102] that leads to the induction of DNA synthesis, which subsequently promote mitogenic effect of malignant proliferating cells. Study conducted by Mukhopadhyay et al. [103] indicates that ibuprofen decreases arginase activity. Arginase is a key enzyme in the urea cycle and thus reduction of the activity of this enzyme will subsequently lead to hyperammonemia.

The other two predicted targets are adipocyte lipid-binding protein and catechol *O*-methyltransferase. No experimental report has been found to either implicate or invalidate them. Further investigation is therefore needed to determine whether or not ibuprofen can bind to these proteins.

3.4. Potential toxicity/side effect-causing protein targets of indinavir

Indinavir is an inhibitor of HIV-1 protease and its clinical and toxicity effects had been studied in some detail. The predicted toxicity and side effect targets identified by INVDOCK search are listed in Table 5, along with the computed ligand–protein interaction energy for each drug–protein complex. INVDOCK produces six potential targets. One target has been confirmed and another three have been implicated by experiments. There are two other predicted targets yet to be validated or invalidated by experiment.

Table 5
Human and mammalian toxicity and side effect protein targets of indinavir identified from INVDOCK search

PDB	Protein	Experimental finding	Target status	Toxicity/side effect	E^a (kcal/mol)	Reference
1znj	Insulin	Drug-induced hyperglycemia	Implicated	Hyperglycemia and diabetes mellitus and lipodystropy	−72.6	[60]
2nll	Retinoic acid receptor	Changes in lipid metabolism and increase retinoic acid receptor signaling	Implicated	Increased plasma lipids and altered body fat distribution	−89.8	[104]
2pah	Phenylalanine hydroxylase			Phenylketonurea	−68.9	
1dmx	Murine carbonic anhydrase V	Increase of diuresis and changes in the urine pH	Implicated	Drug-induced urolithiasis	−70.9	[105]
1icm	Intestinal fatty acid-binding protein	Dose-responsive inhibition of adipogenesis	Confirmed	Reduced triglyceride accumulation	−70.1	[57]
2giw	Cytochrome C			Disorder of muscle and nerve cells	−72.5	

[a] E is the computed ligand–protein interaction energy for each indinavir–protein complex.

Some AIDS patients receiving therapy of HIV protease inhibitors such as indinavir have been reported to show a syndrome characterized by lipodystrophy [57]. Indinavir inhibits the activity of intestinal fatty acid binding protein in dose-dependent manner, resulting in inhibition of adipogenesis as measured by reduced triglyceride accumulation [57]. This protein is identified as a toxicity target by INVDOCK, which is consistent with experimental finding.

The predicted targets that have been implicated by experiments are insulin [60], retinoic acid receptor [104], and carbonic anhydrase [105]. Hyperglycemia and new-onset diabetes mellitus have been reported to occur in HIV-infected patients treated with protease inhibitors. Administration of HIV protease inhibitors such as indinavir has been found to increase the risk of hyperglycemia and diabetes mellitus [60]. In these cases, the increased plasma lipids and altered body fat distribution resembles the effect of Vitamin A overdose. This seems to indicate that indinavir and retinoids may exert their effects through similar mechanism. Interaction of indinavir with retinoic acid receptor has been implicated by the observation of altered lipid metabolism and enhanced retinoic acid receptor signaling process [104]. Drug-induced urolithiasis has been seen in 1.6% of the urinary calculi patients in France [105]. Reduction in carbonic anhydrase activity by indinavir likely is responsible for the increase in diuresis and for the changes in urine pH leading to urolithiasis [105].

The two other predicted targets are phenylalanine hydroxylase and cytochrome C. There are however no known experimental reports to either implicate or invalidate this. Further investigation is therefore needed to determine whether or not indinavir can bind to these proteins.

3.5. Potential toxicity/side effect-causing protein targets of neomycin

Neomycin is an aminoglycoside antibiotic known to induce the following toxic effects and side effects [86]: (a) ototoxicity to vestibular and cochlear cells; (b) nephrotoxicity due to proximal tubular damage and acute tubular necrosis; (c) competitive neuromuscular blockage if given rapidly intravenously together with other neuromuscular blocking drugs. Potential toxicity and side effect targets of this drug identified from INVDOCK search are listed in Table 6, along with the computed ligand–protein interaction energy for each drug–protein complex. The total number of INVDOCK predicted potential targets is 14. Only one of these has been confirmed by experiment. None-the-less, six of the predicted targets have been implicated by experiments. There are seven predicted targets that lack available experimental finding to substantiate them.

The experimentally confirmed target is carbonic anhydrase I [50]. Inhibition of the activity of carbonic anhydrase by neomycin is in a dose-dependent manner [50]. Since carbonic anhydrase is a key enzyme that regulate the bicarbonate level in the blood, loss of the activity of this enzyme results in pH regulation problems in the blood resulting in nephrotoxicity.

Two of the six experimentally implicated targets are malate dehydrogenase [106] and glutamate dehydrogenase [53]. Neomycin has been shown to increase the enzyme level of malate dehydrogenase [106] and decrease the enzymatic activity of glutamate dehydrogenase [53]. Both of these enzymes play important roles in amino acid metabolism and the metabolite products are then channeled to the urea

Table 6
Human and mammalian toxicity and side effect protein targets of neomycin identified from INVDOCK search

PDB	Protein	Experimental finding	Target status	Toxicity/side effect	E^a (kcal/mol)	Reference
1a6a	HLA-DR3			Allergy	−66.4	
1am6	Carbonic anhydrase	Inhibitor	Confirmed	Nephrotoxicity	−67.8	[50]
1d6n	Hypoxanthine–guanine phosphoribosyltransferase			Excess uric acid in body fluid	−89.5	
1dda	Alcohol dehydrogenase			Hepatotoxicity	−72.9	
1jnk	C-JUN N-terminal kinase	Activated the JNK pathway that associates with stress, injury, and apoptosis	Implicated	Caused hair cell death	−80.8	[107]
1pah	Phenylalanine hydroxylase			Phenylketonurea	−67.9	
1urn	U1A spliceosomal protein	Effect on affinity of aminoglycosides to various RNA molecules	Implicated	Interference with mRNA splicing processes	−65.3	[108]
2hgs	Glutathione synthetase			Genotoxicity	−87.5	
1acd	Adipocyte lipid-binding protein			Altered lipid distribution	−68.5	
1ch6	Glutamate dehydrogenase	Decreased enzymatic activity	Implicated	Nephrotoxicity	−91.9	[53]
1hlu	Beta-actin	Increased beta-actin level	Implicated	Hair cell damage	−88.5	[109]
2giw	Cytochrome C			Disorder of muscle and nerve cells	−65.3	
1c50	Glycogen phosphorylase	Stimulated enzyme activity	Implicated	Decreased liver glycogen content	−77.3	[110]
5mdh	Cytoplasmic malate dehydrogenase	Increased enzyme level	Implicated	Nephrotoxicity	−89.9	[106]

[a] E is the computed ligand–protein interaction energy for each neomycin–protein complex.

Table 7
Human and mammalian toxicity and side effect protein targets of penicillin G identified from INVDOCK search

PDB	Protein	Experimental finding	Target status	Toxicity/side effect	E[a] (kcal/mol)	Reference
10gs	Glutathione *S*-transferase	Non-competitive inhibitor	Confirmed	Hepatotoxicity	−58.4	[111]
1a6a	HLA-DR3	Binding to HLA molecule	Confirmed	Allergic inflammatory response	−46.4	[51]
1azm	Carbonic anhydrase I			Nephrotoxicity	−46.5	
1d6n	Hypoxanthine–guanine phosphoribosyltransferase			Excess uric acid in body fluid	−46.8	
1dda	Alcohol dehydrogenase			Hepatotoxicity	−46.1	
1lyy	Lysozyme	Increased urinary excretion of lysozyme	Implicated	Proteinuria	−54.0	[113]
1mcc	Immunoglobulin lambda light chain	Induced urinary excretion of free immunoglobulin	Implicated	Proteinuria	−45.6	[113]
1vid	Catechol *O*-methyltransferase	Metabolized by COMT	Confirmed	Competitive interference with normal COMT activity.	−45.4	[112]
1pah	Phenylalanine hydroxylase			Phenylketonurea	−53.4	
3ifb	Intestinal fatty acid-binding protein			Altered lipid distribution	−45.3	
1oth	Ornithine transcarbamoylase			Hyperammonemia	−55.4	
1ba9	Superoxide dismutase			Genotoxicity	−39.3	
4rla	Arginase			Hyperammonemia	−39.8	

[a] E is the computed ligand–protein interaction energy for each penicillin–protein complex.

cycle. Altered level or activity of each of these enzymes causes imbalance of the metabolism process thereby inducing nephrotoxicity.

The other four experimentally implicated targets are C-JUN N-terminal kinase (JNK) [107], U1A spliceosomal protein [108], beta-actin [109], and glycogen phosphorylase [110]. JNK is associated with stress, injury, and apoptosis activated in hair cells after trauma. Neomycin activates the JNK pathway, which has been shown to cause death of hair cells [107]. A number of aminoglycosides have been reported to interact and interfere with the function of various RNA molecules such as 16S rRNA. INVDOCK search has identified one of such complexes, U1A spliceosome that consists of both RNA and proteins [108]. Interaction of neomycin with this complex may interfere with mRNA splicing process. The protein level of beta-actin has been shown to increase by neomycin [109], which causes hair cell damage since beta-actin is an important structural component of the hair. Neomycin stimulates the activity of glycogen phosphorylase and thus causes a decrease in liver glycogen content [110], which affects glycogen synthesis and glycogen storage in the liver.

The other seven predicted potential targets are HLA-DR3, hypoxanthine–guanine phosphoribosyltransferase, phenylalanine hydroxylase, glutathione synthetase, lipid binding protein, and cytochrome C. No experimental report has been found to either implicate or invalidate these targets. Further investigation is therefore needed to determine whether or not neomycin can bind to these proteins.

3.6. Potential toxicity/side effect-causing protein targets of penicillin G

Penicillin G is a widely used beta-lactam antibiotic. It is known to cause renal dysfunction such as proteinuria. Potential toxicity targets identified from the INVDOCK search are listed in Table 7, along with the computed ligand–protein interaction energy for each drug–proteincomplex. INVDOCK predicts a total of 13 potential targets, three of which have been confirmed and another two have been implicated by experiments. There are however eight targets without available experimental data to validate or invalidate them. The three experimentally confirmed targets are glutathione *S*-transferase, HLA-DR3 and catechol *O*-methyltransferase. Penicillin has been found to act as a non-competitive inhibitor of glutathione *S*-transferase [111]. Since glutathione *S*-transferase plays important role in detoxification processes in the liver, inhibition of glutathione *S*-transferase may cause hepatotoxicity. Experiment also indicated that penicillin induces allergy via binding to HLA-DR3 [51]. Moreover, experiments have shown that penicillin inhibits the activity of catechol *O*-methyltransferase (COMT) [112] via competitive action. Since COMT plays important role in catechol metabolism and detoxification processes, inhibition of this enzyme may interfere with normal liver function.

The two experimentally implicated targets are immunoglobulin and lysozyme. It has been observed that treatment with penicillin results in urinary excretion of free immunoglobulin and an increased urinary excretion of lysozyme [113], which might be related to the effect of penicillin on proteinuria. The eight other predicted targets are carbonic anhydrase, hypoxanthine–guanine phosphoribosyltransferase, alcohol dehydrogenase, phenylalanine hydroxylase, fatty acid binding protein, ornithine transcarbamoylase, superoxide dismutase, and arginase. No experimental report has been found to either implicate or invalidate these targets. Further investigation is therefore needed to determine whether or not penicillin G can bind to these proteins.

Table 8
Human and mammalian toxicity and side effect protein targets of 4H-tamoxifen identified from INVDOCK search

PDB	Protein	Experimental finding	Target status	Toxicity/side effect	E[a] (kcal/mol)	Reference
10gs	Glutathione *S*-transferase	Suppressed enzyme activity	Confirmed	Genotoxicity and carcinogenicity	−50.1	[55]
1d6n	Hypoxanthine–guanine phosphoribosyltransferase			Excess uric acid in serum	−41.1	
1dda	Alcohol dehydrogenase	Inhibition	Confirmed	Enhanced ethanol's sedative effect	−43.8	[52]
1ugb	Carbonic anhydrase II			Nephrotoxicity	−40.4	
1a18	Adipocyte lipid-binding protein			Altered lipid distribution	−40.8	
1d3v	Arginase			Hyperammonemia	−40.5	

[a] E is the computed ligand–protein interaction energy for each tamoxifen–protein complex.

3.7. Potential toxicity/side effect-causing protein targets of 4H-tamoxifen

Tamoxifen is an anticancer drug widely used for treatment of breast cancer [114] and it has been approved as the first cancer preventive drug. Tamoxifen metabolite 4H-tamoxifen is believed to be the major contributor to the anti-oestrogenic effects of tamoxifen inside human body [114]. Hence 4H-tamoxifen is investigated in this study. Potential toxicity/side effect protein targets of 4H-tamoxifen identified by INVDOCK are given in Table 8, along with the computed ligand–protein interaction energy for each drug–protein complex. INVDOCK identified six potential targets, two of which have been confirmed by experiments. But the rest lacks experimental finding to validate or invalidate them.

Tamoxifen is known to cause genotoxicity by suppressing the activity of glutathione *S*-transferase [55]. The genotoxicity and carcinogenicity of tamoxifen have been attributed to metabolic activation of tamoxifen to an electrophile. The identification of glutathione *S*-transferase by INVDOCK as a toxicity target is consistent with experiment. In addition, tamoxifen has been found to enhance sedative effect of ethanol by inhibiting alcohol dehydrogenase [52]. The enzymatic study indicates an adverse metabolic influence by tamoxifen on hepatic metabolism of ethanol-derived acetaldehyde to the potentiation of the seductive effect of ethanol [52].

The other four predicted targets are hypoxanthine–guanine phosphoribosyltransferase, carbonic anhydrase, lipid binding protein, and arginase. We have not found a report in the literature that links tamoxifen to each of these proteins. There is also no report that indicates each of these proteins is not a target of tamoxifen or its analogs. Further investigation is therefore needed to determine whether or not 4H-tamoxifen can bind to these proteins.

3.8. Potential toxicity/side effect-causing protein targets of Vitamin C

Vitamin C (ascorbic acid) is well known for improving scurvy by stimulating synthesis of collagen. It is also an important anti-oxidant in normal cellular functions. This compound is non-toxic for most individuals, even in large doses [115]. Large doses (500 mg to several grams daily) can in a few individuals cause diarrhea and abdominal cramping [115]. It generally has no noticeable effect on renal calcium oxalate stones, systemic conditioning, uricosuria, Vitamin B_{12} destruction, mutagenicity, and iron overload [116,117].

INVDOCK search identifies several protein targets related to toxicity and they are listed in Table 9 along with relevant experimental findings. Moreover, the computed ligand–protein interaction energy for each drug–protein complex is also included in the Table 9. There are five potential targets predicted, two of which have been confirmed by experiments and the rest lack the available experimental data to either validate or invalidate them.

The two experimentally confirmed targets are alpha-amylase [118] and phenylalanine hydroxylase [119]. Vitamin C has been evaluated as an inhibitor of malt, bacterial, fungal, pancreatic and salivary alpha-amylases [118]. As a key enzyme in starch breakdown in pancreas, inhibition of alpha-amylase by excessive Vitamin C may interfere with starch hydrolysis process. It has been found that purified rat liver phenylalanine hydroxylase is inactivated in vitro by Vitamin C. Phenylalanine hydroxylase is an important

Table 9
Human and mammalian toxicity and side effect protein targets of Vitamin C identified from INVDOCK search

PDB	Protein	Experimental finding	Target status	Toxicity/side effect	E[a] (kcal/mol)	Reference
1a6a	HLA-DR3			Allergy	−55.6	
1bsi	Alpha-amylase	Inhibitor	Confirmed	Interfere starch hydrolysis	−45.6	[118]
1d6n	Hypoxanthine–guanine phosphoribosyltransferase			Excess uric acid in serum	−39.5	
1pah	Phenylalanine hydroxylase	Enzyme inactivation	Confirmed	May cause phenylketonurea	−31.8	[119]
1urn	U1A spliceosomal protein			Interfereence with mRNA processing	−33.1	

[a] E is the computed ligand–protein interaction energy for each Vitamin C–protein complex.

Table 10
Comparison of INVDOCK identified human and mammalian toxicity/side effect protein targets with those found in the available experimental data

Compound	Number of experimentally confirmed or implicated targets	Number of targets predicted by INVDOCK	Number of targets missed by INVDOCK	Number of targets without structure or involving covalent bond	Number of INVDOCK predicted targets without experimental finding
Aspirin	15	9	2	4	2
Gentamicin	17	5	2	10	2
Ibuprofen	5	3	0	2	2
Indinavir	6	4	0	2	2
Neomycin	14	7	1	6	6
Penicillin G	7	6	0	1	8
Tamoxifen	2	2	0	0	4
Vitamin C	2	2	0	0	3
Total	68	38	5	25	29

enzyme in phenyalanine metabolism. The inactivation of this enzyme has been known to cause phenylketonurea.

No experimental data has been found to either implicate or invalidate the other three INVDOCK identified targets HLA-DR3, hypoxanthine–guanine phosphoribosyltransferase, and U1A spliceosomal. Further investigation is therefore needed to determine whether or not Vitamin C can bind to these proteins.

3.9. Evaluation of INVDOCK performance

As shown in Table 10, there are a total of 68 protein targets for the eight therapeutic agents that have been confirmed or implicated by available experimental studies. INVDOCK predicted 38 of these. In addition, there are 29 INVDOCK predicted targets that lack experimental validation. Because of a limited scope of experimental study of toxicity targets of these drugs, it is not expected that all toxicity targets have been determined experimentally. Thus, it is difficult to give a complete assessment about which of the predicted targets are false and how many targets are missed without the knowledge of all the targets of these drugs. The evaluation given here should therefore be considered as a preliminary assessment based on currently available experimental data.

It is noted that some of the known toxicity targets of these drugs are not predicted by INVDOCK. Table 11 shows that there are 30 such INVDOCK missed targets. Of these missed targets, 22 are either without available ligand-bound structure of human or mammalian origin, or whose structure is incomplete (e.g. contains only an irrelevant section/domain of the protein). As this work only searches ligand-bound structures of human or mammalian origin, these 22 targets are beyond the capability of our inverse docking algorithm. These targets are: ICE-like cysteineproteases [61], butyrylcholinesterase [62], and glycoproteins [63] for aspirin; caspases [64], *N*-acetyl-beta-D-glucosaminidase [65], gamma-glutamyl-transferase [65], superoxide dismutase [65], alkaline phosphatase [66], basolateral transporters [67], cathepsin B and L [68], Na(+) and K(+)-ATPases [69], and lysosomal phospholipases A and C [70] for gentamicin; TNF for ibuprofen [71]; multidrug transporter proteins MDR1 and MRP1 [72], and *P*-glycoproteins [73] for indinavir; tyrosine hydroxylase [74], tryptophane hydroxylase [74], MAO [74], NMDA receptor [75], Mg(2+)–Na(+)–K(+)-ATPases [50], and phosphatidylinositol-specific phospholipase C [76] for neomycin.

In addition, there are 3 other INVDOCK missed targets to which the binding drug is linked by a covalent bond.

Table 11
Statistics of experimentally determined toxicity/side effect protein targets not predicted by INVDOCK[a]

Compound	Number of experimentally determined toxicity/side effect protein targets not predicted by INVDOCK	Number of protein targets without relevant structure in the database	Number of protein targets in which binding involving a covalent bond	Number of protein targets with available structure in the database, in which binding involving hydrogen bond and non-bonded interactions only
Aspirin	6	3	1	2
Gentamicin	12	10	0	2
Ibuprofen	2	1	1	0
Indinavir	2	2	0	0
Neomycin	7	6	0	1
Penicillin G	1	0	1	0
Tamoxifen	0	0	0	0
Vitamin C	0	0	0	0
Total	30	22	3	5

[a] Relevant structure means a complete and ligand-bound structure of human or mammalian origin. Prediction of a protein target without a relevant structure in our database or to which a drug is linked by a covalent bond is beyond the scope of the ligand–protein inverse docking approach presented in this work.

As our docking algorithm and force field are designed for ligand–protein binding involving intermolecular hydrogen bond and non-bonded interactions only, INVDOCK may not be suitable for the prediction of targets involving a covalent bond. Therefore, the miss of these three targets is not unexpected. These targets are hemoglobin for aspirin [120], hepatic proteins for ibuprofen [121], and penicillin-binding proteins for penicillin [122].

Non-the-less, there are five INVDOCK missed targets whose structure is available. These are TNF-α [61] and acetylcholinesterase [62] for aspirin, leucine aminopeptidase [66] and cathepsin B [68] for gentamicin, and COMT for neomycin [74].

The rate of "correct predictions" from this work may be estimated from the ratio between the number of predicted targets and that of experimentally known targets. Given that the number of INVDOCK predicted and experimentally determined targets is 38 and 68, respectively, the rate of "correct prediction" is 56%. It is noted, however, the relevant 3D structure of 22 of the experimental targets is unavailable. Therefore, these targets are outside the scope of INVDOCK and should be excluded in a more appropriate estimate of rate of correct prediction. This way the rate becomes 83%. If one further excludes the three other targets that involve drug-target covalent bonding, the rate is changed to 89%.

Likewise, the rate of "missed targets" can be estimated from the ratio between the number of INVDOCK missed targets and that of experimentally determined targets. This rate is 5%. The rate of "false-positive" may also be estimated as the ratio between the number of INVDOCK targets without experimental validation and that of the experimental targets, which is 63% when those experimental targets without relevant structure are excluded. Because the number of experimental targets is incomplete, the computed rate of correct prediction and that of the "false-positive" here does not add up to 1. This suggests that, without complete knowledge of all toxicity targets, these computed rates might not fully describe the quality of INVDOCK computations.

3.10. Factors that might affect the quality of a ligand–protein inverse docking search

Several reasons might contribute to the discrepancy between INVDOCK results and experimental finding. It is not expected that exhaustive experiments have been done to determine all protein targets of a drug. Thus, discrepancy might arise for those identified proteins that lack relevant experimental information. Limited number of protein entries available in our cavity database is also expected to result in "missed hit". Moreover, some of the PDB structures may be of little relevance to binding study for a particular molecule. These include entries containing an incomplete section or a chain, protein mutants that are structurally different from the corresponding proteins investigated in experiments, ligand-bound proteins whose conformation is relevant only to a specific set of ligands, macromolecular complexes unrelated to a particular biological process studied experimentally. Docking of a molecule to such an irrelevant structure may thus generate a "false hit". Anticipated rapid progress in structural genomics [29] is expected to provide a more diverse set of relevant structures. Knowledge from study of protein functions [123] also facilitates the selection of relevant structures in determination of protein targets related to a particular cellular or physiological condition. Insufficient modeling of drug-induced conformational change may also affect the quality of docking. Even though a ligand might be docked to a receptor site, the strength of binding may be insufficient to cause a toxicity problem, due to docking error or the intrinsic weakness of binding. Efforts have been made towards a better modeling of conformation change and other aspects of docking [82], which is expected to improve the accuracy of docking. Moreover, in some cases, drug-binding may not produce expected toxic effect because multiple metabolic pathways may compensate for the interruption of one particular pathway. Progress in our understanding and modeling of these pathways will help better evaluating the potential effect of drug-binding [124].

A lack of consideration of protein profiles such as gene expression pattern and protein levels may also result in a discrepancy between INVDOCK search and observations. Many experimental studies of ligand–protein interactions are based on the investigation of cell lines or other assays. Observation of molecular events related to a particular ligand–protein interaction requires that the protein under study be at a sufficient level in the system being investigated. If such a level is not reached at a particular setting, then the corresponding experiment is not useful in probing the binding of a molecule to that protein. Proteins not expressed or at too low levels in a particular disease process are unlikely a good therapeutic target. Advance in proteomics is providing rapidly growing information about the profiles of proteins inside cells [10]. Incorporation of this information into an inverse docking procedure can enable the prediction of more relevant protein targets.

Discrepancy can be caused by a neglect of pharmacokinetic and metabolic profile of a molecule. The action of a molecule requires it to achieve an adequate concentration in the fluid bathing the target tissue. The concentration of a molecule is determined by its pharmacokenetic and metabolic profile. Therefore, information about this profile is important in prediction of protein targets that a small molecule can reach at sufficient concentration. Development in pharmacokinetics and drug metabolism [14] is providing more and more information in this regard.

The quality of an inverse docking procedure also depends on the algorithm and force fields used in docking and scoring. Ideally, in an inverse docking procedure, optimum binding mode should be searched in a cavity. Although more CPU-time demanding, such a search may become practical if one can improve the over-all inverse docking search-speed. For instance, the search-speed of INVDOCK can be significantly improved by distributing database search via parallel

computing. Cavity models may also be further refined to reduce the search space. CPU time saved from these and other improvements may be used for searching optimum binding modes. There has been progress in refining docking/scoring algorithm and force fields particularly in the areas of protein flexibility [82] as well as ligand flexibility [15–27]. Knowledge and new methods/force fields gained from these studies can also be incorporated into an inverse docking procedure.

The scoring function can also influence the performance of INVDOCK search. A stricter condition on the energy threshold can easily result in "missed hits". Likewise, a too relaxed energy threshold may lead to the over production of "false hits". The empirical formula for the energy threshold used in this study seems to give reasonable results for the two agents studied. However, additional study is needed to more extensively evaluate and further refine the energy threshold.

4. Conclusion

Ligand–protein inverse docking has been employed as a method for computer-aided prediction of potential toxicity problems and for identification of protein targets that may result in side effects of small drug molecules. The results on eight therapeutic agents indicate the potential capability of this approach. Performance and applicability of this approach need to be further enhanced by refinement of docking and scoring algorithms and by incorporation of new information from advances in structural genomics, proteomics, protein function, pharmacokinetics and drug metabolism.

References

[1] R.W. Brecher, Risk assessment, Toxicol. Pathol. 25 (1997) 23–31.
[2] J.A. DiMasi, N.R. Bryant, L. Lasagna, New drug development in the United States from 1963 to 1990, Clin. Pharmacol. Ther. 50 (1991) 471–486.
[3] J. Drews, Strategic choices facing the pharmaceutical industry: a case for innovation, Drug. Discov. Today 2 (1997) 72–78.
[4] J.A. Zivin, Understanding clinical trials, Sci. Am. 282 (2000) 69–75.
[5] D. Gerhold, T. Rushmore, C.T. Caskey, DNA chips: promising toys have become powerful tools, Trends Biochem. Sci. 24 (1999) 168–173.
[6] E.F. Nuwaysir, M. Bittner, J. Trent, J.C. Barrrett, Microarrays and toxicology: the advent of toxicogenomics, Mol. Carcinog. 24 (1999) 153–159.
[7] C.P. Rodi, R.T. Bunch, S.W. Curtiss, L.D. Kier, M.A. Cabonce, Revolution through genomics in investigative and discovery toxicology, Toxicol. Pathol. 27 (1999) 107–110.
[8] M.D. Todd, R.G. Ulrich, Emerging technologies for accelerated toxicity evaluation of potential drug candidates, Curr. Opin. Drug Discov. Develop. 2 (1999) 58–68.
[9] J.F. Waring, R.G. Ulrich, The impact of genomics-based technologies of drug safety evaluation, Annu. Rev. Pharmacol. Toxicol. 40 (2000) 335–352.
[10] M.D. Barratt, Integrating computer prediction systems with in vitro methods towards a better understanding of toxicology, Toxicol. Lett. 102/103 (1998) 617–621.
[11] N. Greene, P.N. Judson, J.J. Langowski, C.A. Marchant, Knowledge-based expert systems for toxicity and metabolism prediction: DEREK, StAR and METEOR, SAR QSAR Environ. Res. 10 (1999) 299–314.
[12] S.P. Niculescu, K.L. Kaiser, T.W. Schultz, Modeling the toxicity of chemicals to Tetrahymena pyriformis using molecular fragment descriptors and probabilistic neural networks, Arch. Environ. Contam. Toxicol. 39 (2000) 289–298.
[13] A. Persidis, Proteomics. An ambitious drug development platform attempts to link gene sequence to expressed phenotype under various physiological states, Nature Biotech. 16 (1998) 393–394.
[14] J.H. Lin, A.Y. Lu, Role of pharmacokinetics and metabolism in drug discovery and development, Pharmacol. Rev. 49 (1997) 404–449.
[15] J. Wang, P.A. Kollman, I.D. Kuntz, Flexible ligand docking: a multistep strategy approach, Proteins 36 (1999) 1–19.
[16] G. Jones, P. Willet, R.C. Glen, A.R. Leach, R. Taylor, Development and validation of a genetic algorithm for flexible docking, J. Mol. Biol. 267 (1997) 727–748.
[17] D.K. Gehlhaar, G.M. Verkhivker, P.A. Rejto, C.J. Sherman, D.B. Fogel, L.J. Fogel, S.T. Freer, Molecular recognition of the inhibitor AG-1343 by HIV-1 protease: conformationally flexible docking by evolutionary programming, Chem. Biol. 2 (1995) 317–324.
[18] G.M. Morris, D.S. Goodsell, R. Huey, A.J. Olson, Distributed automated docking of flexible ligands to proteins: parallel applications of AutoDock 2.4., J. Comput. Aided Mol. Des. 10 (1996) 293–304.
[19] H.J. Bohm, The computer program LUDI: a new method for the de novo design of enzyme inhibitors, J. Comput. Aided Mol. Des. 6 (1992) 61–78.
[20] M. Rarey, B. Kramer, T. Lengauer, G. Klebe, A fast flexible docking method using an incremental construction algorithm, J. Mol. Biol. 261 (1996) 470–489.
[21] T. Ewing, I.D. Kuntz, Critical evaluation of search algorithms for automated molecular docking and database screening, J. Comput. Chem. 18 (1997) 1175–1189.
[22] W. Welch, J. Ruppert, A.N. Jain, Hammerhead: fast, fully automated docking of flexible ligands to protein binding sites, Chem. Biol. 3 (1996) 449–462.
[23] M.B. Eisen, D.C. Wiley, M. Karplus, R.E. Hubbard, Hook — a program for finding novel molecular architectures that satisfy the chemical and steric requirements of a molecular binding site, Proteins 19 (1994) 199–221.
[24] C. McMartin, R.S. Bohacek, QXP: powerful, rapid computer algorithms for structure-based drug design, J. Comput. Aided Mol. Des. 11 (1997) 333–344.
[25] N.S. Blom, J. Sygusch, High resolution fast quantitative docking using Fourier domain correlation techniques, Proteins 27 (1997) 493–506.
[26] C. A, C.W. Murray, D.E. Clark, D.R. Westhead, M.D. Eldridge, Flexible docking using Tabu search and an empirical estimate of binding affinity, Proteins 33 (1998) 367–382.
[27] D.M. Lorber, B.K. Shoichet, Flexible ligand docking using conformational ensembles, Protein Sci. 7 (1998) 938–950.
[28] B. Rost, C. Sander, Bridging the protein sequence-structure gap by structure predictions, Annu. Rev. Biophys. Biomol. Struct. 25 (1996) 113–136.
[29] A. Sali, 100,000 protein structures for the biologist, Nature Struct. Biol. 5 (1998) 1029–1032.
[30] Y.Z. Chen, D.G. Zhi, Ligand–protein inverse docking and its potential use in computer search of protein targets of a small molecule, Proteins 43 (2001) 217–226.
[31] I.D. Kuntz, J.M. Blaney, S.J. Oatley, R. Langridge, T.E. Ferrin, A geometric approach to macromolecule–ligand recognitions, J. Mol. Biol. 161 (1982) 269–288.
[32] J. Baynes, M.H. Dominiczak, Medical Biochemistry. Mosby, St. Louis, 1999.
[33] K.B. Wallace, A.A. Starkov, Mitochondrial targets of drug toxicity, Annu. Rev. Pharmacol. Toxicol. 40 (2000) 353–388.

[34] A.G. Rauws, E.J. de Waal, J.W. van der Laan, Sense and non-sense in toxicity assessment of medicinal products, Adv. Drug Res. 30 (1997) 15–72.
[35] T.J. Monks, S.S. Lau, The pharmacology and toxicology of polyphenolic-glutathione conjugates, Annu. Rev. Pharmacol. Toxicol. 38 (1998) 229–255.
[36] J.A. Michell, E.M. Gillam, L.A. Stanley, E. Sim, Immunotoxic side-effects of drug therapy, Drug Saf. 5 (1990) 168–178.
[37] A.-N.T. Kong, S. Mandlekar, R. Yu, W. Lei, A. Fasanmande, Pharmacodynamics and toxicodynamics of drug action: signaling in cell survival and cell death, Pharmaceut. Res. 16 (1999) 790–798.
[38] M.V. Martinov, A.G. Plotnikov, V.M. Vitvitsky, F.I. Ataullakhanov, Deficiencies of glycolytic enzymes as a possible cause of hemolytic anemia, Biochim. Biophys. Acta 1474 (2000) 75–87.
[39] E. Beutler, Galactosemia: screening and diagnosis, Clin. Biochem. 24 (1991) 293–300.
[40] J.I. Wolfsdorf, I.A. Holm, D.A. Weinstein, Glycogen storage diseases. Phenotypic, genetic, and biochemical characteristics, and therapy, Endocrinol. Metab. Clin. North Am. 28 (1999) 801–823.
[41] W. Zheng, S. Ren, J.H. Graziano, Manganese inhibits mitochondrial aconitase: a mechanism of manganese neurotoxicity, Brain Res. 799 (1998) 334–342.
[42] E.M. Coughlin, E. Christensen, P.L. Kunz, K. S, V. Walker, N.R. Dennis, R.A. Chalmers, O.N. Elpeleg, D. Whelan, R.J. Pollitt, V. Ramesh, R. Mandell, V.E. Shih, Molecular analysis and prenatal diagnosis of human fumarase deficiency, Mol. Genet. Metab. 63 (1998) 254–262.
[43] K.B. Wallace, A.A. Starkov, Mitochondrial targets of drug toxicity, Annu. Rev. Pharmacol. Toxicol. 40 (2000) 353–388.
[44] K. Li, Y. Li, J.M. Shelton, J.A. Richardson, E. Spencer, Z.J. Chen, X. Wang, R.S. Williams, Cytochrome C deficiency causes embryonic lethality and attenuates stress-induced apoptosis, Cell 101 (2000) 389–399.
[45] M.L. Cunningham, P.S. Ringrose, B.R. Lokesh, Inhibition of the genotoxicity of bleomycin by superoxide dismutase, Mutat. Res. 135 (1984) 199–202.
[46] T.S. Wang, Y.F. Shu, Y.C. Liu, K.Y. Jan, H. Huang, Glutathione peroxidase and catalase modulate the genotoxicity of arsenite, Toxicology 121 (1997) 229–237.
[47] L.A. Esposito, S. Melov, A. Panov, B.A. Cottrell, D.C. Wallace, Mitochondrial disease in mouse results in increased oxidative stress, Proc. Natl. Acad. Sci. U.S.A. 96 (1999) 4820–4825.
[48] M.E. Hinsdale, D.A. Hamm, P.A. Wood, Effects of short-chain acyl-CoA dehydrogenase deficiency on development expression of metabolic enzyme genes in the mouse, Biochem. Mol. Med. 57 (1996) 106–115.
[49] R. Surtees, N. Blau, The neurochemistry of phenylketonuria, Eur. J. Pediatr. 159 (Suppl. 2) (2000) S109–S113.
[50] S. Suzuki, S. Hatashima, Y. Shinzawa, O. Niwa, R. Tamatani, Toxicity of neomycin on enzyme activities of kidney and duodenal mucosa in vivo: organ specificity and species difference between rats and mice, Comput. Biochem. Physiol. C Pharmacol. Toxicol. Endocrinol. 109 (1994) 77–92.
[51] E. Padovan, T-cell response in penicillin G allergy, Clin. Exp. Allergy 28 (Suppl. 4) (1998) 33–36.
[52] F.S. Messiha, Leu-enkephalin, tamoxifen and ethanol interactions: effects on motility and hepatic ethanol metabolizing enzymes, Gen. Pharmacol. 21 (1990) 45–48.
[53] J. Bryla, T. Lietz, R. Jarzyna, M. Michalik, J. Pietkiewicz, Differential in vivo and in vitro effect of gentamicin on glutamate synthesis and glutamate deamination in rabbit kidney-cortex tubules and mitochondria, Pharmacol. Res. 26 (1992) 367–375.
[54] L.D. DeLeve, Effect of decreased glutathione levels in hereditary glutathione synthetase deficiency on dibromoethane-induced genotoxicity in human fibroblasts, Mutat. Res. 389 (1997) 291–297.
[55] E.F. Nuwaysir, D.A. Daggett, V.C. Jordan, H.C. Pitot, Phase II enzyme expression in rat liver in response to the antiestrogen tamoxifen, Cancer Res. 56 (1996) 3704–3710.
[56] M. Tornwall, P.T. Mannisto, Acute toxicity of three new selective COMT inhibitors in mice with special emphasis on interactions with drugs increasing catecholaminergic neurotransmission, Pharmacol. Toxicol. 69 (1991) 64–70.
[57] B. Zhang, K. MacNaul, D. Szalkowski, Z. Li, J. Berger, D.E. Moller, Inhibition of adipocyte differentiation by HIV protease inibitors, J. Clin. Endocrinol. Metab. 84 (1999) 4274–4277.
[58] H.M. Dodds, L.P. Rivory, The mechanism for the inhibition of acetylcholinesterases by irinotecan (CPT-11), Mol. Pharmacol. 56 (1999) 1346–1353.
[59] W.M. Keung, B.L. Vallee, Daidzin: a potent, selective inhibitor of human mitochondrial aldehyde dehydrogenase, Proc. Natl. Acad. Sci. U.S.A. 90 (1993) 1247–1251.
[60] L.L. Dever, P.A. Oruwari, W.E. Figueroa, C.A. O'Donovan, R.H. Eng, Hyperglycemia associated with protease inhibitors in an urban HIV-infected minority patient population, Annu. Pharmacother. 34 (2000) 580–584.
[61] S. Fiorucci, E. Antonelli, L. Santucci, O. Morelli, M. Miglietti, B. Federici, R. Mannucci, P. Del Soldato, A. Morelli, Gastrointestinal safety of nitric oxide-derived aspirin is related to inhibition of ICE-like cysteine proteases in rats, Gastroenterology 116 (1999) 1089–1106.
[62] M. Schwarz, D. Glick, Y. Loewenstein, H. Soreq, Engineering of human cholinesterases explains and predicts diverse consequences of administration of various drugs and poisons, Pharmacol. Ther. 67 (1995) 283–322.
[63] K. Oketani, M. Murakami, H. Fujisaki, T. Wakabayashi, K. Hotta, Effect of geranylgeranylacetone on aspirin-induced changes in gastric glycoproteins, Jpn. J. Pharmacol. 33 (1983) 593–601.
[64] M. El Mouedden, G. Laurent, M.P. Mingeot-Leclercq, P.M. Tulkens, Gentamicin-induced apoptosis in renal cell lines and embryonic rat fibroblasts, Toxicol. Sci. 56 (2000) 229–239.
[65] A.B. Abdel-Naim, M.H. Abdel-Wahab, F.F. Attia, Protective effects of Vitamin E and probucol against gentamicin-induced nephrotoxicity in rats, Pharmacol. Res. 40 (1999) 183–187.
[66] A.A. Fadel, H.A. Larkin, Gentamicin-induced nephrotoxicosis in lambs, Res. Vet. Sci. 61 (1996) 187–192.
[67] J.H. Dominguez, C.C. Hale, M. Qulali, Studies of renal injury. I. Gentamicin toxicity and expression of basolateral transporters, Am. J. Physiol. 270 (1996) F245–253.
[68] C.J. Olbricht, M. Fink, E. Gutjahr, Alterations in lysosomal enzymes of the proximal tubule in gentamicin nephrotoxicity, Kidney Int. 39 (1991) 639–646.
[69] Y. Fukuda, A.S. Malmborg, A. Aperia, Gentamicin inhibition of Na(+), K(+)-ATPase in rat kidney cells, Acta Physiol. Scand. 14 (1991) 27–34.
[70] K.Y. Hostetler, L.B. Hall, Inhibition of kidney lysosomal phospholipases A and C by aminoglycoside antibiotics: possible mechanism of aminoglycoside toxicity, Proc. Natl. Acad. Sci. U.S.A. 79 (1982) 1663–1667.
[71] D.A. Evans, D.O. Jacobs, A. Revhaug, D.W. Wilmore, The effects of tumor necrosis factor and their selective inhibition by ibuprofen, Annu. Surg. 209 (1989) 312–321.
[72] R.V. Srinivas, D. Middlemas, P. Flynn, A. Fridland, Human immunodeficiency virus protease inhibitors serve as substrates for multidrug transporter proteins MDR1 and MRP1 but retain antiviral efficacy in cell lines expressing these transporters, Antimicrob. Agents Chemother. 42 (1998) 3157–3162.
[73] C.B. Washington, G.E. Duran, M.C. Man, B.I. Sikic, T.F. Blaschke, Interaction of anti-HIV protease inhibitors with the multidrug transporter *P*-glycoprotein (*P*-gp) in human cultured cells, J. Acquir. Immune. Defic. Syndr. Hum. Retrovirol. 19 (1998) 203–209.
[74] A. el Ayadi, M. Errami, Interactions between neomycin and cerebral dopaminergic and serotoninergic transmission in rats, Therapie 54 (1999) 595–599.
[75] J.A. Segal, B.D. Harris, Y. Kustova, A. Basile, P. Skolnick, Aminoglycoside neurotoxicity involves NMDA receptor activation, Brain Res. 815 (1999) 270–277.

[76] D.W. Schwertz, J.I. Kreisberg, M.A. Venkatachalam, Effects of aminoglycosides on proximal tubule brush border membrane phosphatidylinositol-specific phospholipase C, J. Pharmacol. Exp. Ther. 231 (1984) 48–55.
[77] F.M. Richards, Areas, volumes, packing and protein structure, Annu. Rev. Biophys. Bioeng. 6 (1977) 151–176.
[78] W.D. Cornell, P. Cieplak, C.I. Bayly, I.R. Gould, K.M. Merz Jr., D.M. Ferguson, D.C. Spellmeyer, T. Fox, J.W. Caldwell, P.A. Kollman, A second generation force field for the simulation of proteins and nucleic acids, J. Am. Chem. Soc. 117 (1995) 5179–5197.
[79] N.C. Baird, Simulation of hydrogen bonding in biological systems: ab initio calculations for NH_3–NH_3 and NH_3–NH_4^+, Int. J. Quantum Chem. Symp. 1 (1974) 49–53.
[80] Y.Z. Chen, E.W. Prohofsky, The role of a minor groove spine of hydration in stabilizing poly(dA)–poly(dT) against fluctuational interbase H-bond disruption in the premelting temperature regime, Nucleic Acids Res. 20 (1992) 415–419.
[81] Y.Z. Chen, E.W. Prohofsky, Premelting base pair opening probability and drug-binding constant of a daunomycin-polyd(GCAT)–polyd(ATGC) complex, Biophys. J. 66 (1994) 820–826.
[82] H.A. Carlson, J.A. McCammon, Accommodating protein flexibility in computational drug design, Mol. Pharmacol. 57 (2000) 213–218.
[83] http://www.mdli.com/cgi/dynamic/welcome.html.
[84] http://www.sci.ouc.bc.ca/chem/molecule/molecule.html.
[85] M.C. Nicklaus, G.W. Milne, D. Zaharevitz, Chem-X and CAMBRIDGE. Comparison of computer generated chemical structures with X-ray crystallographic data, J. Chem. Inf. Comput. Sci. 33 (1993) 639–646.
[86] K.R. Olson, Poisoning and Drug Overdose, Prentice-Hall, Englewood Cliffs, NJ, 1990.
[87] R.T. Gentry, E. Baraona, I. Amir, R. Roine, Z.W. Chayes, R. Sharma, C.S. Lieber, Mechanism of the aspirin-induced rise in blood alcohol levels, Life Sci. 65 (1999) 2505–2512.
[88] J.F. Glasgow, B. Middleton, R. Moore, A. Gray, J. Hill, The mechanism of inhibition of beta-oxidation by aspirin metabolites in skin fibroblasts from Reye's syndrome patients and controls, Biochim. Biophys. Acta 1454 (1999) 115–125.
[89] G.B. Villanueva, N. Allen, Acetylation of antithrombin III by aspirin, Semin. Thromb. Hemost. 12 (1986) 213–215.
[90] I. Puscas, M. Coltau, R. Pasca, Non-steroidal anti-inflammatory drugs activate carbonic anhydrase by a direct mechanism of action, J. Pharmacol. Exp. Ther. 277 (1996) 1464–1466.
[91] J.W. Dekker, E. Nizankowska, M. Schmitz-Schumann, K. Pile, G. Bochenek, A. Dyczek, W.O. Cookson, A. Szczeklik, Aspirin-induced asthma and HLA-DRB1 and HLA-DPB1 genotypes, Clin. Exp. Allergy 27 (1997) 574–577.
[92] A. Smokovitis, N. Kokolis, T. Ploumis, Great variation in the response of tissue plasminogen activator activity, plasminogen activator inhibition and plasmin inhibition to endotoxin, aspirin and endotoxin after administration of aspirin, Thromb. Res. 50 (1988) 495–505.
[93] W.P. Newmann, R.G. Brodows, Aspirin causes tissue insensitivity to insulin in normal man, J. Clin. Endocrinol. Metab. 57 (1983) 1102–1106.
[94] Y. Kawashima, S. Nakagawa, Y. Tachibana, H. Kozuka, Effects of peroxisome proliferators on fatty acid-binding protein in rat liver, Biochim. Biophys. Acta 754 (1983) 21–27.
[95] D.X. Zhu, J.L. Zhao, L. Mo, H.L. Li, Drug allergy: identification and characterization of IgE-reactivities to aspirin and related compounds, J. Investig. Allergol. Clin. Immunol. 7 (1997) 160–168.
[96] S. Suzuki, S. Takamura, J. Yoshida, Y. Shinzawa, O. Niwa, R. Tamatani, Comparison of gentamicin nephrotoxicity between rats and mice, Comput. Biochem. Physiol. C Pharmacol. Toxicol. Endocrinol. 112 (1995) 15–28.
[97] A.A. Abdel-Gayoum, B.H. Ali, K.M. Abdel-Razig, A.A. Bashir, K. Ghywarsha, Effect of gentamicin-induced nephrotoxicity on some carbohydrate metabolic pathways in the rat renal cortex, Arch. Toxicol. 68 (1994) 643–647.
[98] P. Sandhya, S. Mohandass, P. Varalakshmi, Role of DL-alpha-lipoic acid in gentamicin-induced nephrotoxicity, Mol. Cell Biochem. 145 (1995) 11–17.
[99] S. Ghosh, S. Dasgupta, Gentamicin-induced inhibition of steriodogenic enzymes in rat testis, Indian J. Physiol. Pharmacol. 43 (1999) 247–250.
[100] J.A. Vale, T.J. Meredith, Acute poisoning due to non-steroidal anti-inflammatory drugs: clinical features and management, Med. Toxicol. 1 (1986) 12–12.
[101] K.J. Bamforth, K. Dalgliesh, M.W. Coughtrie, Inhibition of human liver steroid sulfotransferase activities by drugs: a novel mechanism of drug toxicity? Eur. J. Pharmacol. 228 (1992) 15–21.
[102] A. Mallat, C. Gallois, J. Tao, A. Habib, J. Maclouf, P. Mavier, A.M. Preaux, S. Lotersztajn, Platelet-derived growth factor-BB and thrombin generate positive and negative signals for human hepatic stellate cell proliferation. Role of prostaglandin/cyclic AMP pathway and cross-talk with endothelin receptors, J. Biol. Chem. 273 (1998) 27300–27305.
[103] A. Mukhopadhyay, A.P. Sarnaik, D.R. Deshmukh, Interactions of ibuprofen with influenza infection and hyperammonemia in an animal model of Reye's syndrome, Pediatr. Res. 31 (1992) 258–260.
[104] J.M. Lenhard, J.E. Weiel, M.A. Paulik, E.S. Furfine, Stimulation of Vitamin A(1) acid signaling by the HIV protease inhibitor indinavir, Biochem. Pharmacol. 59 (2000) 1063–1068.
[105] M. Daudon, Drug-induced urinary calculi in 1999, Prog. Urol. 9 (1999) 1023–1033.
[106] K. Fent, E. Mayer, G. Zbinden, Nephrotoxicity screening in rats: a validation study, Arch. Toxicol. 61 (1988) 349–358.
[107] U. Pirvola, L. Xing-Qun, J. Virkkala, M. Saarma, C. Murakata, A.M. Camoratto, K.M. Walton, J. Ylikoski, Rescue of hearing, auditory hair cells, and neurons by CEP-1347/KT7515, an inhibitor of C-JUN N-terminal kinase activation, J. Neurosci. 20 (2000) 43–50.
[108] E.N. Mikkelsen, M. Brannvall, Virtanen, L.A. Kirsebom, Inhibition of RNase P RNA cleavage by aminoglycosides, Proc. Natl. Acad. Sci. U.S.A. 96 (1999) 6155–6160.
[109] D.J. Stacey, W.G. McLean, Cytoskeletal protein mRNA expression in the chick utricle after treatment in vitro with aminoglycoside antibiotics: effects of insulin, iron chelators and cyclic nucleotides, Brain Res. 871 (2000) 319–332.
[110] T. Lietz, J. Bryla, The effect of various aminoglycoside antibiotics on glycogen phosphorylase activity in liver and kidney medulla of rabbit, Acta Biochim. Pol. 37 (1990) 187–190.
[111] G. Polidoro, G. Del Boccio, C. Di Ilio, R. Piccolomini, G. Ravagnan, G. Federici, In vitro interaction of penicillin Gs and cephalosporins with human placenta GSH *S*-transferase, Res. Commun. Chem. Pathol. Pharmacol. 46 (1984) 411–423.
[112] N. Ohi, B. Aoki, T. Kuroki, M. Matsumoto, K. Kojima, T. Nehashi, Semisynthetic beta-lactam antibiotics. III. Effect on antibacterial activity and COMT-susceptibility of chlorine-introduction into the catechol nucleus of 6-[(*R*)-2-[3-(3,4-dihydroxybenzoyl)-3-(3-hydroxypropyl)-1-ureiophenylacetamido]penicillanic acid, J. Antibiot. Tokyo 40 (1987) 22–28.
[113] J.P. Vallete, G. Nicot, J.P. Charmes, L. Merle, D. Benevent, C. Leroux-Robert, Low molecular weight proteins as urinary markers of aminoglycoside nephrotoxicity in man, Proc. Eur. Dial. Transplant Assoc. 16 (1979) 596–602.
[114] R.E. Favoni, A.D. Cupis, Steroidal and non-steroidal oestrogen antagonists in breast cancer: basic and clinical appraisal, Trands Pharmacol. Sci. 19 (1998) 406–415.
[115] S.S. Hendler, The Doctor's Vitamin and Mineral Encyclopedia, Fireside Press, 1990.
[116] J.M. Rivers, Safety of high-level Vitamin C ingestion, Int. J. Vitam. Nutr. Res. Suppl. 30 (1989) 95–102.
[117] H. Gerster, High-dose Vitamin C: a risk for persons with high iron stores? Int. J. Vitam. Nutr. Res. 69 (1999) 67–82.
[118] A.D. Abell, M.J. Ratcliffe, J. Gerrard, Ascorbic acid-based inhibitors of alpha-amylases, Bioorg. Med. Chem. Lett. 8 (1998) 1703–1706.

[119] A. Martinez, S. Olafsdottir, J. Haavik, T. Flatmark, Inactivation of purified phenylalanine hydroxylase by dithiolthreitol, Biochem. Biophys. Res. Commun. 182 (1992) 92–98.
[120] F. D'Agnillo, A.I. Alayash, Site-specific modifications and toxicity of blood substitutes. The case of diaspirin cross-linked hemoglobin, Adv. Drug Deliv. Rev. 40 (2000) 199–212.
[121] L.T. Wade, J.G. Kenna, J. Caldwell, Immunochemical identification of mouse hepatic protein adducts derived from the non-steroidal anti-inflammatory drugs diclofenac, sulindac, and ibuprofen, Chem. Res. Toxicol. 10 (1997) 546–555.
[122] S. Yan, G.A. Bohach, D.L. Stevens, Persistent acylation of high-molecular weight penicillin-binding proteins by penicillin induces the postantibiotic effect in *Streptococcus pyogenes*, J. Infect. Dis. 170 (1994) 609–614.
[123] P. Bork, T. Dandekar, Y. Diaz-Lazcoz, F. Eisenhaber, M. Huynen, Y.J. Yuan, Predicting function: from genes to genome and back, J. Mol. Biol. 283 (1998) 707–725.
[124] I. Goryanin, T.C. Hodgman, E. Selkov, Mathematical simulation and analysis of cellular metabolism and regulation, Bioinformatics 15 (1999) 749–758.