Biotechnology and Molecular Biology Reviews Vol. 2(5), pp. 123-132, December, 2007
ISSN 1538-2273 © 2007 Academic Journals

Standard Review

Computational aspects of thematics: application to protein tyrosine phosphatase role in diabetes mellitus

Allam Appa Rao

Center for Biotechnology, Andhra University College of Engineering, Visakhapatnam-530003, India. E-mail: principal@aucevizag.ac.in. Phone: +91- 8941- 216037, +91- 9394073150. Fax: +91-891-2747969.

Accepted 23 November, 2007

Computation plays an important role in functional genomics and proteomics. Theoretical Microscopic Titration Curves (Thematics) are being employed to predict active binding sites of enzymes. The principal reasons are that the pace of discovery of new proteins is increasing, outpacing the ability to characterize them in conventional biochemical and structural techniques; in addition, advances in computational, structural and force data are used in an interative manner to improve accuracy of active site prediction. From methods using amino acid and nucleotide sequences evidence is available that residues in the enzyme core are selected for stability while those at the surface, which are sites of protein interaction, trade off stability for ligand interactivity. Thematics is a computational method that predicts chemical and electrostatic properties of residues in enzymes and utilizes information contained in those predictions to identify active sites. Discussion of the chemical basis for the predictive powers of Thematics is featured in this paper, and the role of protein tyrosine phosphatase in type 2 diabetes discussed briefly.

Key words: Computational aspects, thematics, protein tyrosine phosphatase, diabetes mellitus.

TABLE OF CONTENT

INTRODUCTION

Target identification might seem to be the sole domain of bioinformatics methods. Computational science or computer science has been playing a major role in these methods. Bioinformatics methods have to find the genes that are up or down regulated in certain diseases and then point out which proteins are expressed by these genes. These proteins might either be enzymes catalyzing certain metabolic reactions or be proteins involved in signaling pathways. Manual analysis of such a gigantic data is impossible, computer analysis became a necessity. Access to powerful methods for prediction of binding constants for ligands to target receptors is highly desirable in the drug discovery process. Computational methods for the estimation of binding constants (or, equivalently, free energies of binding) can be particularly fruitful when the structure of the receptor is known. Predictions of the affinity of a given inhibitor for different enzyme variants and mutants may be equally valuable.

Enzymes are protein molecules, which perform a variety of cellular and regulatory functions by interacting with

ligands at specific locations called 'active sites.' The active site depends on a variety of factors from the arrangement of protein chain in space and time (secondary, tertiary and quaternary structures). The active site is dynamic with interactions at a variety of physical and chemical factors in the surroundings. With more data flooding can be annotated, computational methods are employed in this area (Terribilini, 2006). Predicting the active site leads to better understanding of a proteins function and therefore can be useful to design drugs rationally that can modulate interaction of proteins by modifying their structure as well as to establish interactions among networks of proteins (Liang, 2006; Halperin, 2003).

The motivation for 'reverse engineering' of protein structure and function is the assumption that sequences of naturally occurring proteins are defined by selective pressure evolutionarily; the pressure is exerted by a balance of function and stability (Koehl, 2002). Studies comparing stability and solubility of proteins from their sequences have shown that core residues in proteins appear to follow the need for stability (Koehl, 2002).

Efforts were made to use structure of known proteins and predict the amino acid sequences so that the information can be utilized to predict structure of n Sequence of amino acids (primary structure) to can be utilized to predict structure of new proteins whose amino acid sequences alone are known (Raja, 2000). Many protein sequence design methods have been employed such as homology modeling in which attempts are made to match unknown to known, and then thread the sequence to a known backbone based on energy expression. Energy expression is in turn determined by search algorithms such as stochastic and deterministic.

'Lock and key' Concept and computational prediction

Emil Fischer, a chemist, summarized the action of enzymes as follows: 'In order to be able to act chemically on one another, an enzyme and its substrate must fit together like a lock and key.' The original 'lock and key' concept has been revised and refined to describe the site where substrate and enzyme interact. The region of the enzyme containing the binding and catalytic sites is called the active site (Palmar, 2004). It usually occurs near or on the surface of the protein as a cleft for the substrate to fit. The active site is located in three dimensions, with amino acids or cofactors comprising binding and active sites. These arrangements are further affected by amino acids in the microenvironment of the active site.

Identification of binding and catalytic sites is a complex process, involving a variety of processes including the use of substrate analogues, modification of amino acid side chains of enzymes, generation of fusion protein and by site-directed mutagenesis (Palmar, 2004; Kennelly, 2006) In essence, to understand enzyme action at the active site requires not only the identity of amino acids, their spatial arrangement. Information about enzyme structure has been traditionally obtained by x-ray crystallography and NMR spectrometry. Considering the complexity of the physical and chemical properties in which the interactions occur, computational prediction requires, besides, the order of amino acids, the various physical and chemical interactions of atoms lining the active site gorge. Once the raw data is available, computational, mathematical and statistical techniques are employed to predict the active site.

Various computational methods for estimating ligand-binding affinity have been devised, each representing a different choice of computational demand versus accuracy (Ajay, 1995). Free energy perturbation (FEP) theory (Kolllman, 1993; Jorgensen, 1989; Beveridge, 1989) combined with conformational sampling by molecular dynamics (MD) or Monte Carlo (MC) simulation can be used to calculate free energy changes upon small modifications of the ligand or receptor (Singh, 1988a, b). However, FEP calculations become quite complicated and computationally expensive for structurally dissimilar inhibitors and for calculation of absolute free energies of binding. In the latter case, precautions must also be taken to ensure a correct standard state when molecules are annihilated or created (Cammon, 1997). Scoring functions applied to single conformations of the docked complex is a more empirical approach to affinity predicttion. They are generally based on identifying individual points of intermolecular interaction such as hydrogen bonds, ionic interactions and hydrophobic interactions, as well as entropy estimates, in a given conformation of the receptor-ligand complex and as signing a binding energy score to each contributing factor (Bohm, 1994; Wallqvist, 1995; Smyithe, 1996; Jain, 1996). These methods include estimates of the effect of solvation, either implicitly by their parameterization to fit a set of experimental values or explicitly by measuring the change in the solvent-accessible surface area upon binding. There are also examples of utilizing scaled molecular mechanics energies for minimized structures to obtain binding energy estimates (Holloway, 1995).

Yuko Tsuchiya et al (Yuko, 2004) analyzed 63 protein-DNA complexes by focusing their attention on the shape of the molecular surface of the protein and DNA, along with the electrostatic potential on the surface, and constructed a new statistical evaluation function to make predictions of DNA interaction sites on protein molecular surfaces. Recently Kengo et al. (2007) have developed a method to predict ligand-binding sites in a new protein structure by searching for similar binding sites in the Protein Data Bank (PDB).

The similarities are measured according to the shapes of the molecular surfaces and their electrostatic potentials. They provided new web server, to interface the search method (URL: http://eF-site.hgc.jp/eF-seek). It has been observed that the density of conserved residue positions is higher at the interface regions of interacting

protein surfaces, except for antibody–antigen complexes, where a very low number of conserved positions is observed at the interface regions. Yuhua et al. (2005) identified putative interacting regions on the surface of interacting partners. Using the free energy along with conservation information and other descriptors used in the literature for ranking docking solutions, such as shape complementarity and pair potentials, they developed a global ranking procedure that significantly improves the docking results by giving top ranks to near-native complex structures

Molecular docking

Assigning function to uncharacterized enzymes discovered through genome projects has provided a great challenge to the fields of informatics, enzymology and structural biology. Docking potential ligands into flexible models of protein structures and docking potential high-energy intermediates, rather than substrates, into known structures are two new computational methods that have provided a much-needed boost to the field (Karen, 2007). The use of structure based docking correctly predicts the function of BC0371 enzyme of enolase family (Ling, 2007) and S-adenosylhomocysteine deaminase of an enzyme from the amidohydrolase family (Jhoannes, 2007). In this work the authors approached the problem with the rational that because proper interaction, and thus corrected computational scoring are the reflection of fit of enzyme to ligand, the use of more complementary ligand will produce superior results. Since the 1990s, hundreds of genomes have been sequenced, providing the great promise of new drug targets, a comprehensive understanding of cell metabolism and a deep well of knowledge for protein engineering (Marvin, 2003). Reliable molecular docking and estimation of binding affinities remain an important goal in structure based drug design applications (Blaney, 1993; Jones, 1995; Lynbard, 1995; Rosenfeld, 1995; Gschwend, 1996; Lengauer, 1996). Given a novel inhibitor against an enzyme of known structure, it is important to predict how the ligand is going to bind. Successful prediction allows one to prioritize synthetic resource into producing compounds, which are more likely to be useful and biologically active. Docking studies are also important when designing new ligands in order to confirm the binding mode and to form the basis for an energetic assessment of their binding affinity. Docking methodology can be used to screen compound databases (Kuntz, 1982; Miller, 1994; Welch, 1996) provided the docking is fast enough and there is a method to provide a reasonably reliable ranking order for the biological activity of the ligands. Kmuncek et al. (????) explained the applicability of evaluate the applicability of automated molecular docking techniques and quantum mechanical calculations to the construction of a set of structures of enzyme-substrate complexes for use in Comparative binding energy analysis to obtain 3D structure-activity relationships (Jan, 2003).

Thematics (Theoretical microscopic titration curves)

Thematics is based on finite difference Poisson-Boltzmann (FDPB) methods for calculating the electrical potential function for a complex array of charges, coupled with a Monte-Carlo procedure for determining the average charge as a function of pH for the ionizable residues in that protein's calculated potential. This method identifies reactive sites, including residues involved in catalysis (Ondrechen, 2001). This technique utilizes Deep View-Swiss Pdb Viewer method to obtain the electrical potential function for the protein, followed by calculation of the predicted titration curves for all of the ionizable residues in the protein structure. The novel feature of Thematics is that it extracts information from the shapes of the theoretical titration curves for the ionizable residues in the protein structure, as derived from a FDPB calculation. In a protein the residues generally considered to be ionizable are all of the Arg, Asp, Glu, His, Lys, and Tyr residues, all Cys residues that are not involved in disulfide bridges, plus the N- and C- termini. A typical equation for the pH as a function of the concentrations of an acid HA and its conjugate base as reported (Jaeju, 2005).

In an enzyme, the protein environment influences the chemical and electrostatic properties of active site residues in order to facilitate catalysis and recognition. Any active-site residue that acts as a catalytic Brønsted acid (/base) must then act as a Brønsted base (/acid) in order to restore itself to its initial state for the next turnover cycle. A perturbed titration curve increases the range of conditions over which a residue may act as both an acid and a base, as is required by the definition of a catalyst. Similarly, such unusual titration behavior expands the pH range over which a residue may exist in both charged and neutral forms, thus assisting reversible recognition. Finding these perturbed titration properties not only identifies the active site but may also give clues about the type of chemistry that is catalyzed by that active site.

In proteins, active sites usually comprise hydrophobic pockets that involve side-chain atoms. Methods which use interaction energies between the receptor and different probes to locate energetically favorable sites are complex. They need the assignment of proton locations and partial charges to the receptor atoms. While van der Waals energies can indicate sterically available regions, the long-range nature of electrostatic potentials make the interpretation of energy levels difficult. The purpose of the Site Finder application in MOE is to calculate possible active sites in a receptor from the 3D atomic coordinates of the receptor. The underlying concept is that the pH dependent electrical potential caused by ionizable residue, perturbs the substrate and the catalytically active residues, so that efficient proton

transfer is enabled between them over the desired pH range (Fariselli, 2002; Yan, 2006)

Prediction of interface residues and sites

Identification of protein-protein interaction sites (Rao, 2006) and detection of specific amino acid residues that contribute to the specificity and strength of protein interactions is an important problem with applications ranging from rational drug design to analysis of metabolic and signal transduction networks. Computational methods, mostly sequence based, were developed to predict interface sites at differing resolution: at the entire domain, a sequence or at the level of each amino acid (Halperin, 2003). In general, interface residues are predicted by their known characteristics, evolutionary conservation of residues and by residue energy distribution on protein surfaces (Liang, 2006). Known characteristics include the fact that protein interfaces occur at the planar and accessible paths, their surface is less flexible. Shehadi et al. (Ihsan, 2005) suggests that the Thematics can be applicable to a much larger set of proteins for which an experimentally determined structure is unavailable. With a few exceptions, the predicted active sites in the comparative model structures are similar to that of the corresponding template structure.

Correlated mutations and its detection

It has long been suggested that, in the course of evolution, residue substitutions, which tend to destabilize a particular structure must be compensated by other substitutions that confer greater stability on that structure (Tuffery, 1999). Altschuh et al. (Alstchuh, 1987) analyzed the amino acid substitutions in the coat protein structure of tobacco mosaic virus and seven related viruses and showed that some pairs of positions with identical patterns of amino acid substitutions are close together using correlated mutations, residues close to interaction site are expected to mutate simultaneously during evolution, in contrast to convolution approach where simultaneously mutations are looked at two interacting partners rather than a single protein (Halperin, 2003) Sequence-based methods depend on the genomic contex, that is, the primary structure, on gene order conservation and on domain homologues (Halperin, 2003). It should be emphasized that such predictive methods need an iterative experimental characterization and production, even though it assures a ‘comprehensive measure of the compatibility of a sequence with a structure’ (Raha, 2000). In addition, a variety of statistical predictive methods are employed using neural networks and machine learning techniques (Raha, 2000). Often a combination of methods are used, involving neural networks, docking and superimposition (Li, 2005). Such predictive methods sometimes use well known phage display libraries (Site Light method) (Halperin, 2003).

The remarkable conservation of protein structure, compared to that of sequences, suggests that, in the course of evolution, residue substitutions which tend to destabilize a particular structure must be compensated by other substitutions that confer greater stability on that structure. Given the compactness of proteins, spatially close residues are expected to undergo the compensatory process. Surprisingly, approaches designed to detect such correlated changes have led, until now, only to limited success in detecting pairs of residues adjacent in the three-dimensional structures (Friesner, 2004). Thematics achieves a high success rate of interaction site predicttion; about 86% correct or partially correct using CatRes/CSA annotations only and about 93% with an expanded reference set (Ying, 2007). Success rates for catalytic residue prediction are similar to those of other structure-based methods, but with substantially better precision and lower false positive rates. The method requires only the structure of the query protein as input. Thematics predictions may be obtained via the web from structures in PDB format at: http://pfweb.chem.neu.edu/thematics/submit.html

Method of Raj Chakrabarti

Given the fact that most of the amino acid sequences in a protein are optimized for structural stability and are often buried within the structure, attempts were made to perform a more comprehensive physico-chemical and electrical interaction map of amino acids at the surface of the protein which are likely to be sites of interactions with other proteins and chemicals (Li, 2005).

In their original study, Chakrabarti et al chose a set of 10 amino acid residues that form essential contacts to the ligand for sequence optimization. The essential contacts were identified by residues necessary for hydrogen-bonds, salt bridges, van der Waals forces or hydrophobic contacts. The latter was obtained using data and by multiple sequence alignments (Chakrabarti, 2005). The three principal steps in the procedure included determining the lowest-energy protein structure for each residue, following sampling of side-chain conformation and calculation of ligand-binding affinity for each. In this way it was observed that most binding-site amino acids were ‘optimized for simple scoring functions based on ligand-binding affinity, under the constraint that residues involved in catalysis are restricted to catalytically favorable conformations (Chakrabarti, 2005).

Protein topology and stability

The sequences of naturally occurring proteins are defined by evolutionary selective pressure, which is controlled by a fine balance of function, stability, and kinetics. Although most random mutations of sequences are unlikely to enhance stability or function, they can be accepted by natural selection as long as they are neutral (or near neutral). As a consequence, the size of the sequence space

space compatible with a given protein fold is very large (although small compared with the full space a protein sequence can explore, whose size is 20^N, where N is the number of residues of the protein). The number of compatible amino acids at a given position in a protein is structure-dependent: some local structures such as tight turns have energetic constraints that can be satisfied only by small amino acids such as glycine, alanine, or proline. In an exploration of sequence space that was associated with a given protein attribute, all- atom models were used along with a physical energy function (Chakrabarti, 2005). It was found that the volume of sequence space was defined by protein length, protein topology and stability of protein fold. Folding free energy function selects naturally available protein sequences in the core but not on the surface (Koehl, 2002). This is utilized to search sequence space for sequences that satisfy stability constraints for known protein structures, which are then compared with naturally occurring counterparts. Information of native sequence appears to be 'encoded' in the backbone structure of protein (Jarmillo, 2002). Core amino acid sequences were on an average 50% identical to their native counterparts. Surface region identity scores were much lower than core amino acid sequences (Heinz, 1992). Newer methods such as Hidden Markov models are being explored to improve performance of identifying appropriate sequences. It has been suggested that sequences the surface of enzymes may have been selected at least in part, for mediating interactions at the expense of protein stability. Amino acids at the surface were therefore optimized not for protein stability, but for functional reasons (Heinz, 1992).

Many proteins maintain their structure while undergoing extensive mutations. For example, alanine substitution of 10 consecutive residues in bacteriophage T4 lysozyme leads to only minor structural differences (Cordes, 1999). On the other hand, a single double mutation can generate a dramatic structural change, as observed in the Arc repressor for which the interchange of the sequence position of residues 11 and 12 leads to a new structure in which each β-strand is replaced by an α-helix (Bogarad, 1999). These seemingly conflicting results have lead to a complicated picture of protein sequence evolution: it is not clear whether a protein fold can evolve into a new fold by accumulation of simple point mutations (Drexler, 1981). As a first step toward a better understanding of evolution, studies have focused on characterizing the protein sequence space compatible with a given protein structure, the so-called inverse folding problem (Pabo, 1983; Arnold,1998). A large range of methods, including *in vitro* experiments mimicking evolution (Arnold, 1998 a, b; Desjarlais, 1995) and fully automated computer protein design (Dahiyat, 1997; Hellinga, 1998; Muegge, 2001), has been proposed for searching sequences that would stabilize a given protein structure with improved stability or with a new activity.

In summary there has been synergistic use of data and methods obtained from different sources to give meaningful biological solutions to the structure and func-tion of proteins, that are available in geometric progression, faster than can be characterized by hitherto tradi-tional methods. Computational methods offer simulation and analytical opportunities to characterize, identify and modify proteins with appropriate characteristics to catalyze the specific biochemical reaction that is required.

The accurate prediction of binding affinity is one of the most important tasks in de novo ligand design to direct the synthesis of potential modeled ligands bound to receptors. The continuously growing number of protein structures and protein structural models, resulting from the knowledge of genomes, represents enormous promise for development of drugs and other bioactive compounds. Full utilization of the potential requires an armory of computational methods for prediction of ligand–protein-binding affinities. Structure-based prediction methods with various levels of accuracy and speed are available today: fast docking and scoring approaches (Leach, 2006) for reduction of vast chemical libraries to several hundreds of probable binders, the second-pass methods selecting a dozen or so compounds with a high probability of success, and the most sophisticated methods like Free Energy Perturbation (Akash, 2007) and Thermodynamic Integration (King, 1998) which can e.g., discern the affinity of closely related analogs.

Protein tyrosine phosphatase and its relationship to diabetes

The global prevalence of diabetes mellitus has increased continuously and it has been predicted that the number of adult diabetics will double within 30 years. Almost 250 million people, nearly 6% of adults in the world, have diabetes (Mayor, 2006). The problem with diabetes is that, even if the exact same mutation caused it in everyone, it would look different from person to person and family to family, depending on environmental influences, the genetic background it's laid upon, and modifier genes (Martin, 1992 a, b). Its expression would be variable. Furthermore, studies have shown that diabetes is not simple; it's genetically complex, involving multiple genes, and multiple gene-environment interactions

Type 1 diabetes mellitus (T1D) results from an autoimmune destruction of beta-cells in genetically predisposed individuals (Castano, 1990). In prediction of T1D the presence of circulating autoantibodies to insulin (IAA), glutamic acid decarboxylase (GADA) and the protein tyrosine phosphatase like IA-2 (IA-2A) are important and multiple autoantibodies confer a higher risk of developing T1D (Bingley, 1994), especially when combined with the risk genotype (Kimpimaki, 2002).

Type 2 diabetes, formerly known as non-insulin-dependent diabetes (NIDDM), accounts for most cases of diabetes mellitus worldwide. It is estimated that in 2005 there were approximately 160 million individuals with the

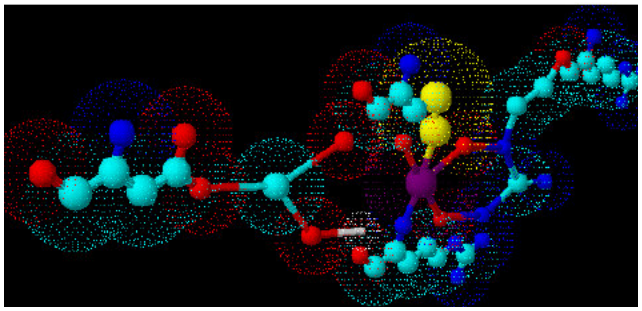

Figure.1. 3D structure of PTP 1B enzyme phosphate formation.

disease and that this number is likely to double by 2025 (Swan, 1982). Swan GW first applied the engineering optimal control theory to equations describing insulin and glucose interactions (Jannik, 2003). T2DM result from complex interaction between genetic and environmental factors. The high incidence of T2DM in certain populations and among first-degree relatives of T2DM patients, as well as the high concordance in identical twins, provides strong evidence that genetic factors underlie susceptibility to T2DM (Elebein, 2002).

Protein tyrosine phosphatase 1B (PTP1B) (Figure 1) has been shown to be involved in the negative regulation of both insulin and leptin action at the *in vitro, ex vivo* and *in vivo* levels. A growing body of human genetic data also support the hypothesis that PTP1B has an important role in insulin signaling and possibly in obesity in humans. The protein tyrosine phosphatases (PTP) constitute a family of closely related key regulatory enzymes that dephosphorylate phosphotyrosine residue in their protein substrates.

Malfunctions of PTP activity are linked to various diseases, ranging from cancer to neurological disorders and diabetes. The family of protein tyrosine phosphatases, which is encoded by 100 genes in humans, plays a critical in the regulation of signal transduction. Recently a variety of links between aberrant PTP function and human diseases has been defined (Liu, 2003) Protein tyrosine phosphatase 1B (PTP1B) has been implicated as one of the key regulators of insulin and leptin signal transduction pathways. Inhibiting PTP1B action using antisense oligonucletides and small molecule inhibitors represents novel therapeutic approach for the treatment of insulin resistance, type II diabetes, and obesity. Liu Gang reviewed the recentaction, particularly small molecules PTP1B inhibitors, and inhibitors with drug like properties (Zhong, 2002). Zhong (2002) summarized the correlation of PTP summarized the correlation of PTP structure and the correlation of PTP structure and function from mutagenesis experiments, in this the molecular basis for PTP1B substrate reorgani zation is discussed, a powerful strategy is presented for creating specific and high affinity bidentate PTP inhibitors that simultaneously bind both the active site and unique adjacent site (Ebina, 1985).

The major sites of autophosphorylation within the pro-tein tyrosine kinase (PTK) domain of the β subunit are Tyr-1158, Tyr-1162, and Tyr-1163 (using the numbering of (White, 1988) which are contained within the PTK activation segment.

Autophosphorylation of all three sites is required for maximal activation (Patti, 1998). Following activation, additional autophosphorylation reactions occur in the juxtamembrane segment and the C-terminal tail of the β subunit. IRK-catalyzed phosphorylation of proteins, such as IRS-1 and IRS-2, creates high-affinity binding sites for proteins that contain Src homology 2 (SH2) domains, such as phosphatidylinositol 3-kinase (PI3-K), resulting in the assembly of the multiprotein signaling complexes that mediate the effects of insulin on cellular metabolism, growth, and glucose homeostasis (Ahmad, 1997).

In the context of insulin signaling, protein phosphorylation is a reversible process, regulated by the coordinated action of protein kinases and phosphatases. Therefore, dephosphorylation of the insulin receptor by protein tyrosine phosphatases (PTPs) is also a critical component of the control of insulin signaling.

Numerous studies have demonstrated that in humans, and in animal mo-dels, insulin resistance in type 2 diabetes and obesity is accompanied by increases in PTP activity and increases in the level of expression of defined members of the PTP family.

In skeletal muscle and adipose tissue from insulin-resistant obese and diabetic subjects, the increased PTP activity has been linked primarily to changes in the expression of the receptor-like PTP LAR (leukocyte common Antigen related) and the cytoplasmic enzyme PTP1B (Ahmad, 1995).

Studies in cell culture models have indicated a functional link between LAR and dephosphorylation of the insulin receptor (Kennedy, 1999).

However, disruption of the *LAR* gene in mice yields a complex phenotype consistent with a postreceptor defect in insulin signaling but associated with impaired activation of downstream signals, such as PI 3-kinase (Desmarais, 1999). Therefore, attention has focused on PTP1B as a major regulator of insulin signaling.

The observation that peptides with tandem pTyr residues bind with high affinity to PTP1B is in agreement with a recent study by Ramachandran and colleagues who observed that pep-tides containing two adjacent nonhydrolyzable analogs of pTyr (difluorophosphono-methyl-Phe) were potent inhibitors of the challenges associated with developing PTP1B advances in various approaches for attenuating PTP1B (Alastair, 2006).

One of the main challenges for the design of potent inhibitors is to overcome the small size and polar nature of the active-site pocket. In the case of PTP1B, bidentate inhibitors have been designed that target two surface pockets. This concept could also be applicable to other family members. The polarity of the active site and difficulties in identifying potent pTyr mimetic compounds

resulted in phosphatase inhibitors with unfavorable pharmacological properties. To generate inhibitors that can be used *in vivo*, these properties must be improved.

Recently Alastair J. Barr and Stefan Knapp constructed phylogenetic tree for PTP 1D which will be useful for development of new PTP inhibitors (Kennedy, 1999). Kennedy et al reviewed its role in diabetes and obesity (Kenn, 2000; Burke, 1998)

Design of small molecule compounds inhibiting the enzymatic function of PTP1B is of great medicinal interest. It is actively being pursued in many academic (Moller, 2000) and industrial (Kole, 1996) organizations. It is appreciated that PTP1B is a negative regulator of insulin receptor signaling and a suitable drug target for treatment of insulin resistance associated with diabetes and obesity (Bandyopadhya, 1997). Clinical studies have found a correlation between insulin resistance states and levels of PTP1B expression in muscle and adipose tissues (Ahmad, 1997; Elchelby, 1999). Recent PTP1B knockout studies revealed that mice lacking functional PTP1B exhibit increased sensitivity toward insulin and are resistant to obesity (Chernoff, 1990).

PTP1B is a prototypical intracellular protein tyrosine phosphatase found in a wide variety of human tissues (Salmeen, 2000). The exact roles of PTP1B in relation to insulin resistance are not fully understood. It has been demonstrated that the interaction of insulin with its receptor leads to the phosphorylation of certain tyrosine residues (1158, 1162 and 1163) within the receptor protein, thus activating the receptor kinase (Ahmad, 1995). PTP1B, probably together with other phosphatases (LAR, PTPα and SH-PTP2) (Tonks, 2001), dephosphorylate the activated insulin receptor. Recent studies with LAR knocked out mice have not shown altered glucose homeostasis, however (Malamas, 2000). The dephosphorylated insulin receptor loses the tyrosine kinase activity that is required for further down-stream signaling (Sarmiento, 2000). Inhibition of PTP1B prevents, at least to some extent, the activated insulin receptor from being inactivated. In addition to the above function, PTP1B dephosphorylates a number of other receptor tyrosine kinases, including the EGF receptor and the PDGF receptor (Tuffery, 1999). Protein- tyrosine phosphatases (PTPs), together with protein –tyrosine kinases, control the tyrosine phosphorylation state in the cell (Figure 1), which is important for cellular activities such as growth , differentiation ,motility ,cell-cell interacttions ,metabolism, gene transcription, and the immune response(Hunter, 1995; Tonks, 1996). PTPs constitute a growing family of transmembrane (receptor- like) and intracellular enzymes (cytoplasmic) that rival the protein tyrosine kinase in terms of structural diversity and complexity. Extensive biochemical and structural studies of PTPs during the past ten years have led to a detailed understanding of the mechanism by which PTPs catalyze phosphate monoester hydrolysis. PTP1B can effectively accommodate peptides of varying sequence, in a catalytically viable fashion that differs from that observed in the PTP1B.DADEpYL structure .considering that the region surrounding the active site is covered with many charged cervices and protrusions, compensatory interact-tions may be available. A further understanding of the specific functional role that PTPs play in cellular signaling and how they are mechanically able to carry out these roles requires a detailed description of the structural features that control PTP substrate specificity. Inspite of the remarkable progress in the identification and characterrization of new PTPs distinguish and in the understanding of the PTP catalysis, the molecular basis by which PTPs distinguish and recognize the diverse substrates that they encounter in the cell has still not been firmly established. This is partly because the physiological substrates for most PTPs remain unknown and it remains a major obstacle to obtain quantities of specifically and stoichiometrically tyrosine phosphory-lated proteins required for detailed enzymological studies. So the application of Thematics for the study of this type of complex proteins will be useful for *in silico* drug discovery studies. (Figure 2) represents the proposed comparison view of active site prediction / drug target identification using proteomic and thematic approaches.

Conclusion

PTP1B is tractable to structure-based drug design, and the crystal structure is well known. Recently, the crystal structure of the closely related T-cell protein tyrosine phosphatase (TCPTP) has become available, raising the possibility for designing selective inhibitors (Johnson, 2002). It is concluded that PTP1B is playing an important role not only in type 2 diabetes but in other diseases like cardiovascular and alzheimer's. As previously reported by us for mathematical, bioinformatic and structural analysis of complex proteins in diabetes and its complications is useful to understand the pathophysiology of diseases (Bhremeramba, 2007; Rao, 2006; Rao, 2006; Sridhar, 2006; Sridhar, 2005). Thematics are useful for further development of specific inhibitors to proteins (PTP1B), potential anti-diabetic drugs, chemical probes for clarification of the roles of PTP1B in normal cellular processes as well as in pathogenic pathways. It is proved beyond doubt that the computational analysis do necessary for any of the protein like PTP which a complex protein structurally.

ACKNOWLEDGMENT

Author is thankful to Srinubabu Gedela for his help during manuscript preparation.

REFERENCES

Ahmad F, Azevedo JL, Cortright R, Dohm GL, Goldstein BJ (1997). Alterations in Skeletal Muscle Protein-Tyrosine Phosphatase Activity

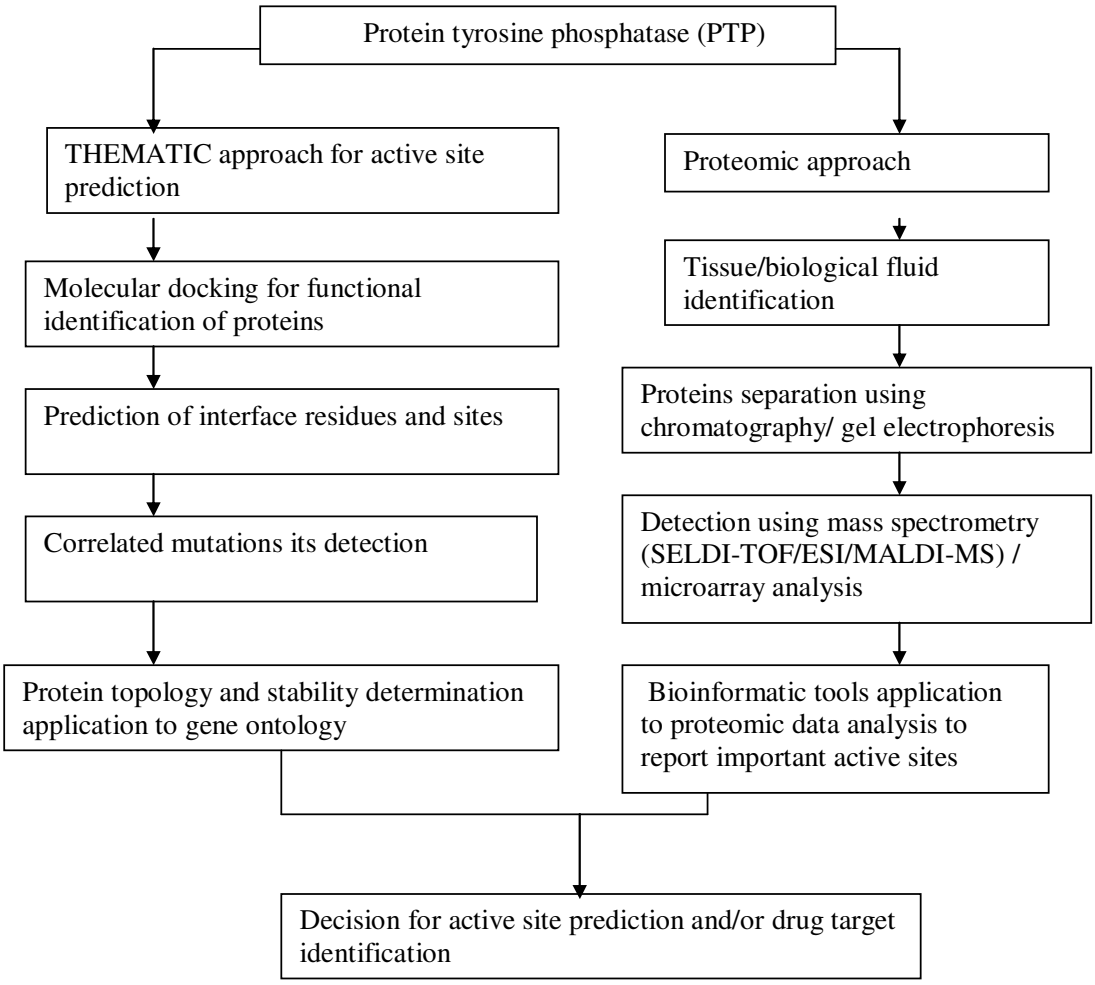

Figure 2. Proposed comparison view of proteomic and thematics approaches for active site prediction; SELDI-TOF: Surface enhanced laser desorption ionization time of flight mass spectrometry; ESI: electro spray ionization mass spectrometry; MALDI: Matrix assisted laser desorption ionization mass spectrometry.

and Expression in Insulin-resistant Human Obesity and Diabetes. Metabolism 46 : 1140-1145

Ahmad F, Azevedo JL, Cortright R, Dohm GL, Goldstein BJ (1997). Alterations in Skeletal Muscle Protein-Tyrosine Phosphatase Activity and Expression in Insulin-resistant Human Obesity and Diabetes J. Clin. Invest. 100: 449–458.Ahmad F, Considine RV, Bauer TL, Ohannesian JP, Marco CC, Goldstein BJ (1997). Improved sensitivity to insulin in obese subjects following weight loss is accompanied by reduced protein-tyrosine phosphatases in adipose tissue .Metabolism. 46:1140-1145.

Ahmad F, Goldstein BJ (1995). Protein-Tyrosine phosphatases:Emerging Targets for Therapeutic invention ofinType2 Diabetes and related states of Insulin Resistance. Biochim. Biophys. Acta. 1248: 57-63.

Ahmad F, Goldstein BJ, (1995). J. Biol. Chem. 272: 448–453.

Ajay, Murcko MA (1995). Computational methods. to. predict. binding. free-energy. in. ligand-. receptor complexes. J. Med. Chem. 38: 4953-4967

Akash KÆ, Stefan B (2007). Improved estimation of ligand–macromolecule binding affinities by linear response approach using a combination of multi-mode MD simulation and QM/MM methods. J. Comput. Aided. Mol. Des. 21: 131–137.

Alastair J, Barr K, Stefan K, T (2006), MAPK-specific tyrosine phosphatases: new targets for drug discovery?. Pharmacol. Sci. 27: 525-530.

Altschuh D, Lesk AM, Bloomer AC, Klug A (1987). Normal modes of symmetric protein assemblies. Application to the tobacco mosaic virus protein disk. J. Mol. Biol. 193: 693-698.

Arnold FH (1998). When blind is better: Protein design by evolution. Nat. Biotechnol. 16: 617-622.

Arnold FH (1998). Design by directed evolution. Acc. Chem. Res. 31: 125-129.

Arnold FH (1998). Enzyme engineering reaches the boiling point Proc. Natl. Acad. Sci. USA. 95: 2035-2039.

Bandyopadhyay D, Kusari A, Kenner KA, Liu F, Chernoff J, Gustafson TA, Kusari J (1997). Protein tyrosine phosphatase role in type 1diabetes J. Biol. Chem 272:1639-1643.

Beveridge DL, DiCapua FM (1989). Free energy via molecular simulations: application to chemical and biochemical system .Annu. Rev. Biophys. Chem. 18: 439-445.

Bhremeramba, Sridhar GR, Rao AA (2007). 'Diabetes Molecular Genetics, Signaling Pathways and Integrated Physiology (J1), sponsored by Biovitrum AB, January 14-19, Keystone Resort, Keystone, Colorado, USA. p. 87.

Bingley PJ, Christie MR, Bonifacio E, Bonfanti R, Shattock M, Fonte MT (1994). Combined analysis of autoantibodies improves prediction of IDDM in islet cell antibody-positive relatives. Diabetes. 43:1304–1310.

Blaney JM, Dixon JS, (1993). Perspect. Drug Discov. Design. 1: 301-306

Bogarad L, Deem M (1999). Random multi-recombinant PCR for the construction of combinatorial protein libraries. Proc. Natl. Acad. Sci. USA. 96: 2591-2597.
Böhm HJ (1994). Protein-protein interactions as targets for antiviral chemotherapy J. Comput.-Aided Mol. Design. 8:243-247.
Burke Jr, TR, Zhang ZY (1998). Protein-tyrosine phosphatases: Structure, mechanism, and inhibitor discovery. Biopolymers. 47: 225-229.
Castano L, Eisenbarth GS (1990). Type-I diabetes: a chronic autoimmune disease of human mouse and rat. Annu. Rev. Immunol. 8: 647–679.
Chakrabarti R, Kilbanov AM, Friesner RA (2005). Computational prediction of native protein ligand-binding and enzyme active site sequences. Proc. Natl. Acad. Sci. USA. 102: 10153-10158.
Chen H, Wertheimer SJ, Lin CH, Amrein KE, Burn P, Quon MJ (1997). protein tyrosine phosphatase inhibition. J. Biol. Chem. 272: 8026-8031.
Chernoff J, Schievella AR, Jost CA, Erikson RL, Neel BG (1990). Cloning of a cDNA for a Major Human Protein-Tyrosine-Phosphatase. Proc. Natl. Acad. Sci. USA. 87: 2735-2739.
Cordes M, Walsh N, McKnight C , Sauer R (1999). Sequence determinants of a conformational switch in a protein structure Sci. 284: 325-329
Dahiyat BI, Mayo SL (1997). De Novo Protein Design: Fully Automated Sequence Selection. Sci. 278: 82-86.
Desjarlais J, Handel T (1995) . strategies in protein design Protein Sci. 4: 2006-2009.
Desmarais S, Friesen RW, Zamboni R, Ramachandran C., Difluro(phosphono) methyl]phenylalanine-containing peptide inhibitors of protein tyrosine phosphatases., Biochem J (1999). 337: 219–223.
Drexler KE (1981). Concepts and schemes for the re-engineering of physical protein modules: generating nanodevices via targeted replacements with constrained amino acids. Proc. Natl. Acad. Sci. USA 78: 5275-5279.
Ebina Y, Ellis L, Jarnagin K, Edery M, Graf L, Clauser E, Ou JH , Masiarz F, Kan YW, Goldfine ID (1985). The human insulin receptor cDNA: The structural basis for hormone-activated transmembrane signalling. Cell. 40: 747–758.
Elbein SC (2002). Perspective: The Search for Genes for Type 2 Diabetes in the Post-Genome Era. Endocrinology. 143: 2012–2018.
Elchelby M, Payette P, Michaliszyn E, Cromlish W, Collins S, Lee LA, Normandin D, Cheng A, Himms HJ, Chan CC, Ramachandran C, Gresser MJ, Tremblay ML, Kennedy BP (1999).enzyme inhibition prediction. Sci. 283:1544.
Fariselli P, Pazos F, Valencia A, Casadio R (2002). Prediction of protein-protein interaction sites in hetero-complexes withneural networks Eur. J. Biochem. 269:1356-1361
Friesner RA, Banks JL, Murphy RB, Halgren TA, Klicic JJ, Mainz DT, Repasky MP, Knoll EH, Shelley M, Perry JK, Shaw DE, Francis P,Shenkin PS (2004). A new approach for rapid, accurate docking and scoring,1. Method and assessment of Docking. J. Med. Chem. 47:1739-49.
Gschwend DA, Good AC, Kuntz ID (1996) Molecular interaction sites of interleukin-2 RNA and methods of modulating the same . J. Mol. Recognition. 9: 175-179.
Lengauer T, Rarey M (1996). Transcription Factors and Neoplasia Vistas in Novel Drug Design Curr. Opin. Struct. Biol. 6: 402-407.
Halperin I, Wolfson H, Nussinov R (2003). Epitope Mapping: The First Step in Developing Epitope-Based Vaccines. Drug. Dev. Protein Sci.12:1344-59.
Head RD, Smythe ML, Oprea TL, Waller CL, Greene S, Marshall GR (1996). Validate - A New Method for the Receptor-Based Prediction of Binding Affinities of Novel Ligands. J.Am.Chem. Soc.118: 3959-3962.
Heinz, D, Baase W, Matthews B (1992). Folding and function of a T4 lysozyme containing 10 consecutive alanines illustrate the redundancy of information in an amino acid sequence. Proc. Natl. Acad. Sci. USA . 89: 3751-3755.
Hellinga HW (1998). Computational protein engineering. Nat. Struct. Biol. 5:525-529.
Holloway MK, Wai JM, Halgren TA, Fitzgerald PM, Vacca JP, Dorsey BD, Levin RB, Thompson Wj, Chen LJ, deSolms SJ, Gaffin N, Ghosh AK, Giuliani EA, Graham SL, Guare JP, Hungate RW, Lyle TA, Sanders WM, Tucker TJ, Wiggins M, Wiscount CM, Woltersdorf OW, Young SD, Darke PL, Zugay JA (1995) Molecular mechanics calculations on protein-ligand complexes. Med. Chem. 38: 305-309.
Hunter T (1995). Protein kinases and phosphatases: the yin and yang of protein phosphorylation and signaling. Cell. 80: 225-236.
Ihsan AS, Alexej A, Alper U, Ying W, Leonel F, Murga M 02115, Usa, Valentin I, Mary JO (2005) J. Bioinformatics. Comput. Biol. 3: 127-143.
J. Clin. Invest. Proteinin tyrosine phosphatase in obesity. 100: 449-453.
Jaeju KO, Leonel F, Murga, Ying W, Mary JO (2005), Prediction of active sites for protein structures from computed chemical properties Bioinformatics. 21: i258-i265.
Jain AN (1996). Scoring noncovalent protein-ligand interactions:A continuous differentiable function tuned to compute binding affinities.J. Comput.-Aided Mol. Design.10:427-432.
Jan K, Michal B, Santos L, Federico G, Rebecca C. Wade JD (2003). J. Comput-Aided Mol. Design. 17: 299–311.
Jannik N, Andersen Nicholas K, tonks (2003). protein tyrosine phosphatase based therapeutics: lessons from PTP1B, Protein phosphatases, Springer Berlin publisher.
Jarmillo A, Wernisch L, Hery S, Woodak SJ (2002). High-Resolution Protein Design with Backbone Freedom. Proc. Natl. Acad. Sci. USA. 99:13554-13559.
Johannes CH, Ricardo MA, Alexander A, Fedorov E, Steven C, Almo B K, Shoichet, Frank M (2007). Nature published online 1 July 2007 (doi: 10.1038/nature 05981).
Johnson TO, Ermolieff J, Michael RJ (2002). Probing acid replacements of thiophene PTP1B inhibitors . Nat. Rev.Drug. Disc.1: 696-699.
Jones G, Willett P (1995). Dose-dependent Activation of Antiapoptotic and Proapoptotic Pathways by Ethanol Treatment in Human Vascular Endothelial Cells. Curr. Opin. Biotech. 6: 652-658.
Jorgensen WL (1989). Free Energy calculations: A Breakthrough for Modeling Organic Chemistry in Solution. Chem. Res. 22: 184-189.
Karen NA (2007). PTP in diabetes. Nat. Chem. Biol. 3: 452-453.
Kenn BP, Kennedy BP, Ramachandran C (2000). Biochemical Pharmacol. 60: 877-883.
Kennedy BP (1999). Role of protein tyrosine phosphatase-1B in diabetes and obesity. Biomed Pharmacother. 53: 466-470.
Kennelly PJ, Rodwell VW (2006). In: Murray RK, Granner DK, Rodwell VW Harper's Illustrated Biochemistry. McGraw Hill, Boston. 27 ed, pp. 49-60.
Kimpimaki T, Kulmala P, Savola K, Kupila A, Korhonen S, Simell T (2002) . Natural History of ß-Cell Autoimmunity in Young Children with Increased Genetic Susceptibility to Type 1 Diabetes Recruited from the General Population. J. Clin. Endocrinol. Metab. 87: 4572–4579.
King H, Aubert RE, Herman WH (1998). Global burden of diabetes. 1995–2025: prevalence numerical estimates and projections. Diabetes Care. 21: 1414–1431.
Koehl P, Levit M (2002). Roles of mutation and recombination in the evolution of protein thermodynamics. Proc. Natl. Acad. Sci. USA. 99: 120-125
Koehl P, Levitt M (2002). Structural Determinants of the Rate of Protein Evolution in Yeast. Proc Natl Acad. Sci. USA. 99:1280-5.
Kole HK, Garant MJ, Kole S, Bernier M (1996). Protein tyrosine phosphatase role in type 2 diabetes. J. Biol. Chem. 271:14302-14307.
Kollman P (1993). Computational and Experimental tools for Studying Protein Structure and Stability. Chem. Rev. 93: 2395-2417.
Kuhlman B, Baker D (2000). Structural Bioinformatics: Methods, Concepts and Application to Blood Coagulation Proteins.Proc. Natl. Acad. Sci. USA. 97:10383-8.
Kuntz ID, Blaney JM, Oatley SJ, Langridge R, Ferrin TE (1982). Molecular speleology: the exploration of crevices in proteins for prediction of binding sites, design of drugs and analysis of surface recognition. J. Mol. Biol. 161: 269-276.
Leach AR, Shoichet BK, Peishoff CE (2006). Prediction of protein-ligand interactions. Docking and scoring: successes and gaps. J. Med. Chem. 49:5851-5856.
Li B, Gallin WJ (2005). A7DB: a relational database for mutational phy-

siological and pharmacological data related to the α7 nicotinic acetylcholine receptor. BMC Structur. Biol. 5: 16-21.
Liang S, Zhand C, Liu S, Zhou Y (2006). Protein binding site prediction using an empirical scoring function. Nucleic. Acids. Res.34:3698-3707. Ling S, Chakrapani K, Alexander A, Fedorov EV, Margaret E, Shoshana B, Heidi J, Patricia C, Steven CA, Matthew PJ, John AG (2007). Roles of galectins in chronic inflammatory microenvironments. Nat. Chem. Biol. 3:486-491.
Liu G (2003). Protein Tyrosine Phosphatase 1B Inhibition: Opportunities and Challenges. Curr. Med. Chem. 10: 1407-1421.
Lybrand TP (1995). Ligand-protein docking and rational drug design. urr. Opin. Struct. Biol. 5: 224-229.
Malamas MS, Sredy J, Gunawan I, Mihan B, Sawicki DR, Seestaller L, Sullivan D, Flam BR (2000). Œ-Ketocarboxylic acid-based inhibitors of protein tyrosine phosphatases J. Med. Chem. 43: 995-1002.
Martin BC, Warram JH, Krolewski AS, Bergman RN, Soeldner JS, Kahn CR (1992) . Role of glucose and insulin resistance in development of type 2 diabetes mellitus: results of a 25-year follow-up study. Lancet. 340: 925–929.
Martin BC, Warram JH, Rosner B, Rich SS, Soeldner JS ,Krolewski AS (1992). Familial clustering of insulin sensitivity. Diabetes. 41: 850–854.
Marvin EF, Gary MJ, David G, Thomassen C, EO, Aristides P (2004) Harnessing the power of genomics for energy and the environment Sci. 300: 290 – 293.
Mayor S (2006). News Severe adverse reactions prompt call for trial design changes. BMJ. 9: 191-1194.
Miller MD, Kearsley SK, Underwood DJ Sheridan RP (1994). Structure activity relationship approaches and applications J. Comput.-Aided Mol. Design. 8: 153-159.
Møller NPH, Iversen LF, Andersen HS, McCormack JG (2000). PTP1B inhibited drugs.,Curr. Opin. Drug Dis. Dev. 3: 527-533.
Muegge I, Rarey M (2001). Proteins in computer sci. Rev. Comput. Chem. 17:1-8.
Ondrechen MJ, Clifton JG, Ringe D (2001). Protein structure to function: insights from computation Proc Natl Acad Sci USA. 98: 12473-12478.
Pabo C (1983). Molecular technology: Designing proteins and peptide. Nature (London). 301: 200-205.
Palmer T (2004) Enzymes. Affiliated East West Press, New Delhi. pp. 175-909.
Palmer T (2004). Enzymes. Affiliated East West Press, New Delhi. pp. 67-75.
Patti ME, Kahn CR (1998). The insulin receptor--a critical link in glucose homeostasis and insulin action. J. Basic Clin. Physiol. Pharmacol. 9: 89–109.
Raha K, Wollacott AM, Italia MJ, Desjarlais JR (2000).). Prediction of aminoacid sequence from structure. Protein Sci. 9:1106-19.
Rao AA, Bhaskar D (2006). Integration specificity of phage c31 integrase in Human genome Int. J. Bioinformatics. 1: 1-5
Rao AA, Bhremeramba, Sridhar GR (2006). Proceedings of 18th IEEE International Conference on Tools with Artificial Intelligence (ICTAI'06). pp. 810-821.
Rao AA, Sridhar GR (2006). Butylcholinesterase connection between type 2 diabetes and Alzheimer's disease. Lipids in Health and Disease. 5: 28-32.
Rao AA., Proceedings of 18th IEEE International Conference on Tools with Artificial Intelligence (ICTAI'06). pp. 12-13.
Ren JM, Li PM, Zhang WR, Sweet LJ, Cline G, Shulman GI, Livingston JN, Goldstein BJ, . Transgenic mice deficient in the LAR protein-tyrosine phosphatase exhibit profound defects in glucose homeostasis. Diabetes 1998). 47: 493–497.
Rosenfeld R, Vajda S, Delisi C (1995). Ann. Rev. Biophys. Biomol. Struct. 24: 677-683.
Salmeen A, Andersen JN, Myers MP, Tonks NK, Barford D (2000). Molecular basis for recognition and dephosphorylation of the activation segment of the insulin Cell. 6: 1401-1406.
Sarmiento M, Puius YA, Vetter SW, Keng YF, Wu L, Zhao Y, Lawrence DS, Almo SC, Zhang ZY (2000). Structural Basis of Plasticity in Protein Tyrosine Phosphatase 1B Substrate Recognition Biochem. 39: 8171-8176.
Singh UC (1988). Probing the Salt Bridge in the Dihydrofolate Reductase-Methotrexate Complex by Using the Coordinate-Coupled Free-Energy Perturbation Method. Proc. Natl. Acad. Sci. USA. 85: 4280-4585.
Singh UC, Benkovic S, McCammon JA, Gilson K, Given JA, Bush BL (1997). Computation of affinity and selectivity: Binding of 2,4-diaminopteridine and 2,4-diaminoquinazoline inhibitors to dihydrofolate reductases Biophys. J. 72: 1047-1053.
Sridhar GR, Nirmala G, Rao AA, Madhavi AS, Sreelatha S, Sudha Rani J, Vijayalakshmi P (2005). Serum butyrylcholinesterase in type 2 diabetes mellitus: a biochemical and bioinformatics approach. Lip. Health and Dis. 4: 18-23.
Sridhar GR, PV Lakshmi, Allam Appa Rao (2006). Phylogenetic tee construction. J. Assoc. Physicians.India. 54:122-128.
Swan GW (1982). An optimal control model of diabetes mellitus, Bull. Mathematical. boil. 44 : 793-808.
Terribilini M, Lee JH, Yan C, Jernigan RL, Honavar V, Dobbs D (2006). Prediction of RNA binding sites in proteins from amino acid sequence. RNA.12:1450-1462.
Tonks NK, Neel BG (1996). Protein Tyrosine Phosphatases in the Vessel Wall. 87: 365-368.
Tonks NK, Neel BG (2001). The T-cell protein tyrosine phosphatase is phosphorylated on Ser-304 by cyclin-dependent protein kinases in mitosis. Curr. Opin. Cell Biol 13: 182-187.
Tuffery P, Durand M, Darlu P (1999). How possible is the detection of correlated mutations?, Theor. Chem. Acc. 101: 9-14.
Wallqvist A, Jernigan RL, Covell DG (1995). A preference-based free-energy parameterization of enzyme-inhibitor binding. Applications to HIV-1-protease inhibitor design Protein Sci. 4:1881-1903
Welch W, Ruppert J , Jain AN (1996). Evidence for the Evolution of Bdelloid Rotifers Without Sexual Reproduction or Genetic Exchange. Chem. Biol. 3: 449-455.
White MF, Shoelson SE, Keutmann H, Kahn CR, (1988). A cascade of tyrosine autophosphorylation in the beta-subunit activates the phosphotransferase of the insulin receptor. J. Biol. Chem. 263: 2969–2980.
Yan C, Terribilini M, Wu F, Jernigan RL, Dobbs D, Honavar V (2006). Alternative Splicing: New Insights from Global Analyses BMC. Bioinformatics. 7: 262-268.
Ying W, Jaeju K, Leonel F, Murga, Mary JO (2007). Active site prediction fo rcomparative model structures with Thematics. BMC Bioinformatics. 2007. 8: 119-125.
Yuhua D, Boojala VB, Reddy, Yiannis N, Kaznessis (2005).Physicochemical and residue conservation calculations to improve the ranking of protein–protein docking solutions Protein Sci. 14:316-328.
Yuko T, Kengo K , Haruki N (2004). Protein structure function bioinformatics. 55: 885-894.Kengo K, Yoichi M, Haruki N (2007). Identification of protein biochemical functions by similarity search using the molecular surface database eF-site. Nucleic Acids Res. 35: W398–W402.
Zhong YZ (2002). Annual. Rev. Pharmcol. Toxicol. 42: 209-234