2015

Insights into the Development of Chemotherapeutics Targeting PFKFB Enzymes

Robert Blake Crochet
Louisiana State University and Agricultural and Mechanical College

INSIGHTS INTO THE DEVELOPMENT OF CHEMOTHERAPEUTICS TARGETING PFKFB ENZYMES

A Dissertation

Submitted to the Graduate Faculty of the
Louisiana State University and
Agricultural and Mechanical College
in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

in

The Department of Biological Sciences

by
Robert Blake Crochet
B.S., Louisiana State University, 2009
December 2015

ACKNOWLEDGMENTS

I would like to express my sincere appreciation and thanks to the many people who have guided, supported, and encouraged me throughout my graduate studies. I am deeply indebted to my advisor, Dr. Yong-Hwan Lee, for making this work possible. No words of thanks are enough for his mentorship, guidance, support and patience. I also thank the members of my committee, Dr. Grover Waldrop, Dr. Naohiro Kato, and Dr. Michal Brylinski for their invaluable support in improving my research project. Additional thanks are owed to my fellow lab members: Dr. Jeong-Do Kim, Dr. Michael Cavalier, and Dr. Minsuh Seo for the contributions they have made to my education.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

The PFKFB enzymes control the primary checkpoint in the glycolytic pathway and are implicated in a multitude of diseases: from cancer, to schizophrenia, to diabetes, and heart disease. The inducible isoform, PFKFB3, is known to be associated with the upregulation of glycolysis in many cancers. The first study within this work investigates the potential for using tier-based approaches of virtual screening to target small molecule kinases, with PFKFB3 serving as a case study. For this investigation, bioactive compounds for PFKFB3 were identified from a compound library of 1364 compounds via high-throughput screening, with bioactive compounds being further characterized as either competitive or non-competitive for F6P. Using the F6P-competitive compounds, several structure based docking programs were assessed individually and in conjunction with a pharmacophore screening. The results showed that the tiered virtual screening approach, using pharmacophore screening in addition to structure-based docking, improved enrichments rates in 80% of cases, reduced CPU costs up to 7-fold, and lessened variability among different structure-based docking methods.

The second study investigates the structural and kinetic characteristics of citrate inhibition on the heart PFKFB isoenzyme, PFKFB2. High levels of citrate, an intermediate of the TCA cycle, signify an abundance of biosynthetic precursors and that additional glucose need not be degraded for this purpose. Previous studies have noted that citrate acts as an important negative feed-back mechanism to limit glycolytic activity by inhibiting PFKFB enzymes, yet the structural and mechanistic details of citrate's inhibition had not been determined. To study the molecular basis for citrate inhibition, the three-dimensional structures of the human and bovine PFKFB2 orthologues were solved, each in complex with citrate. For both cases, citrate primarily occupied the binding site of Fructose-6-phosphate (F6P), competitively blocking F6P from

binding. Additionally, a carboxy arm of citrate extended into the γ-phosphate binding site of ATP, sterically and electrostatically blocking the catalytic binding mode for ATP. In the human orthologue, which utilized AMPPNP as an ATP analogue, conformational changes were observed in the 2-kinase domain as well as the binding mode for AMPPNP. This study gives new insights as to how the citrate-mediate negative feedback loop influences glycolytic flux through PFKFB enzymes.

CHAPTER 1: REVIEW OF LITERATURE

1.1 Glycolysis

The most pressing need of all cells in the body is for an immediate source of energy. Although ATP can be produced in different ways, virtually all cells possess the capacity to convert glucose to lactate with the concomitant synthesis of ATP. This conversion, otherwise known as glycolysis, is the main energy source in many prokaryotes, eukaryotic cells lacking mitochondria, and eukaryotic cells under low-oxygen conditions (Lodish 2000). In eukaryotes, the pathway consist of ten cytosolic enzymes that catalyze the steps involved in the oxidation glucose to pyruvate. Under aerobic conditions, the pyruvate molecules enter the mitochondria where they are further oxidized to carbon dioxide through the citric acid cycle and the respiratory chain, ultimately producing up to 36 ATP molecules. Conversely, in anaerobic environments, pyruvate remains in the cytosol where it is converted to lactate in order to regenerate oxidized nicotinamide adenine dinucleotide (NAD+), a substrate necessary for additional rounds of glycolysis. In both circumstances, glycolysis also serves to provide vitally important precursors for macromolecular synthesis (Fell 1997; Flatt 1995; Fothergill-Gilmore 1987; Mcgilvery 1979).

1.2 The Glycolytic Pathway Is Tightly Controlled

Because all cells require a constant supply of energy to function properly, flux through the glycolytic pathway must be continuously adjusted in response to fluctuating conditions, both inside and outside the cell (Metallo and Vander Heiden 2013). The rate by which glucose is converted to pyruvate is primarily regulated to meet two major cellular needs: (1) maintain adequate levels of ATP and (2) provision molecular building blocks for macromolecular synthesis, such as the formation of fatty acids and DNA (Kim et al. 2006; Wegner et al. 2015).

To ensure these needs are met, glycolysis is subject to stringent regulation at several key steps (Bertram 2000; DeBerardinis and Thompson 2012).

In mammals, Phosphofructokinase (PFK-1) serves as the most important control point in the glycolytic pathway, marking glucose's first point of commitment to the glycolytic pathway. In this reaction, fructose-6-phosphate (F6P) is phosphorylated at the C1 carbon to form fructose-1,6-bisphosphate (F-1,6-P_2). Importantly, this reaction serves as the rate-limiting step for the entire glycolytic pathway and is thus critical in determining glycolytic flux; with increased PFK-1 activity resulting in the activation of glycolysis (Banaszak et al. 2011; Dunaway et al. 1988; Dunaway and Kasten 1988; Evans et al. 1981). Since PFK-1 functions as the central gatekeeper to the glycolytic pathway, its activity must be tightly controlled. Complex allosteric regulation, among other regulatory mechanisms, link PFK-1 activity to the energy status of the cell (Furuya and Uyeda 1980b; Reinhart and Lardy 1980; Yuan et al. 1990). Two indicators of cellular energy, ATP and AMP, both influence PFK-I activity. High levels of ATP act to inhibit the enzyme by lowering its affinity for F6P; conversely, AMP acts to reverse this inhibition. This ATP/AMP ratio provides a mechanism through which the enzymatic activity of PFK-I can be linked to the energetic status of the cell, preventing unnecessary glucose degradation when cellular energy levels are adequate (Al Hasawi et al. 2014; Berg 2012; Webb et al. 2015).

Several metabolites from metabolic pathways found downstream of glycolysis also serve to regulate PFK-1 activity, and thus glycolytic flux. For example, PFK-1 is inhibited by the phosphoenolpyruvate (PEP) generated in the later stages of glycolysis. In 1980, fructose 2,6-bisphosphate (F-2,6-P_2) was identified as a potent activator of phosphofructokinase. Fructose 2,6-bisphosphate was found to elevate PFK-1 activity by diminishing the inhibitory effects of ATP and concomitantly increasing the enzymes affinity for fructose-6-phosphate. By allowing

PFK-1 to bypass the inhibitory effect of ATP, F-2,6-P_2 acts to uncouple the link between the cells energy status and the rate of glycolytic flux (Furuya and Uyeda 1980a; Wegner et al. 2015).

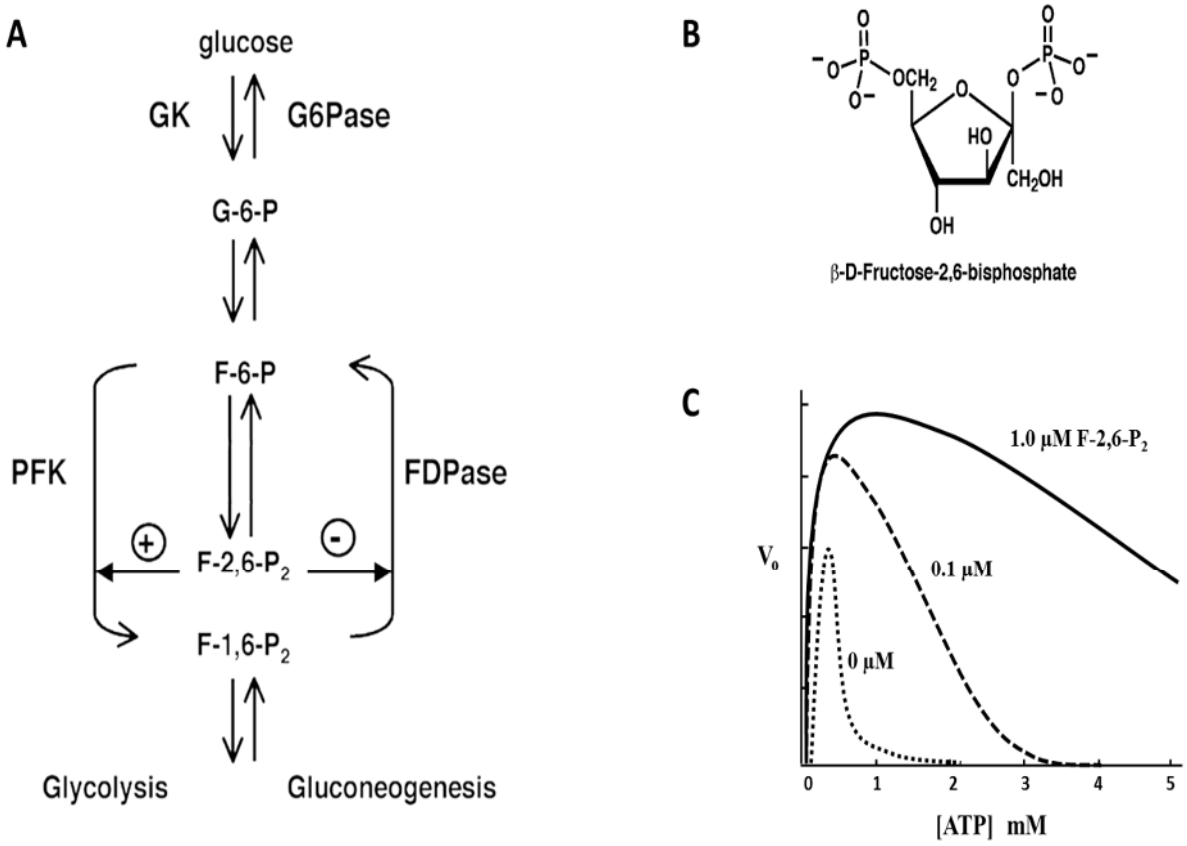


Figure 1.1. Fructose-2,6-Bisphosphate, the most potent allosteric activator of glycolysis. (A) Reciprocal regulation of the glycolytic and gluconeogenic pathways by F-2,6-P_2. (B) The chemical structure of fructose-2,6-bisphosphate. (C) Initial velocity curve showing the effects of varying concentrations of F-2,6-P_2 on ATP inhibition.

1.3 The Bifunctional Enzyme, PFKFB

As previously noted, the concentration of F-2,6-P_2 significantly influences the overall rate of glycolysis via allosteric activation of the rate-limiting enzyme, PFK-1. The intracellular concentration of F-2,6-P_2 is controlled by a family of bifunctional 6-phosphofructo-2-kinase/fructose-2,6-bisphosphatases (PFKFB) that are encoded by four, tissue-specific, genes (O. Minchenko et al. 2003). Although different in size, kinetics, and regulatory properties, each of the four isoforms has kinase and phosphatase activities, allowing for both the synthesis and degradation of F-2,6-P_2. Thus cellular levels of F-2,6-P_2 are determined by a balance between its production and consumption, which are handled two clearly divided functional domains (Leipe et al. 2002; Leipe et al. 2003; Rigden 2008).

Each PFKFB monomer consists of a single polypeptide chain that can be subdivided into four distinct regions: the 6-phosphofructo-2-kinase (6-PF-2-K) domain, the fructose-2,6-bisphosphatase (F-2,6-P_2ase) domain, and two regulatory domains. The two catalytic domains are conserved among the different tissue isoforms with the sequence identity for both domains being >85%. These two domains are connected by a partially flexible linker region that allows for small rotational deviations among isoforms, possibly influencing binding kinetics and/or catalytic rates. Flanking the functional N-terminal and C-terminal domains are two regulatory domains. Unlike the catalytic domains, little sequence similarity is shared among the regulatory domains of the four isoforms. Studies have demonstrated these domains to have very distinct posttranslational modification profiles (Cavalier et al. 2012; Hasemann et al. 1996; Kim et al. 2006; Y. H. Lee et al. 2003).

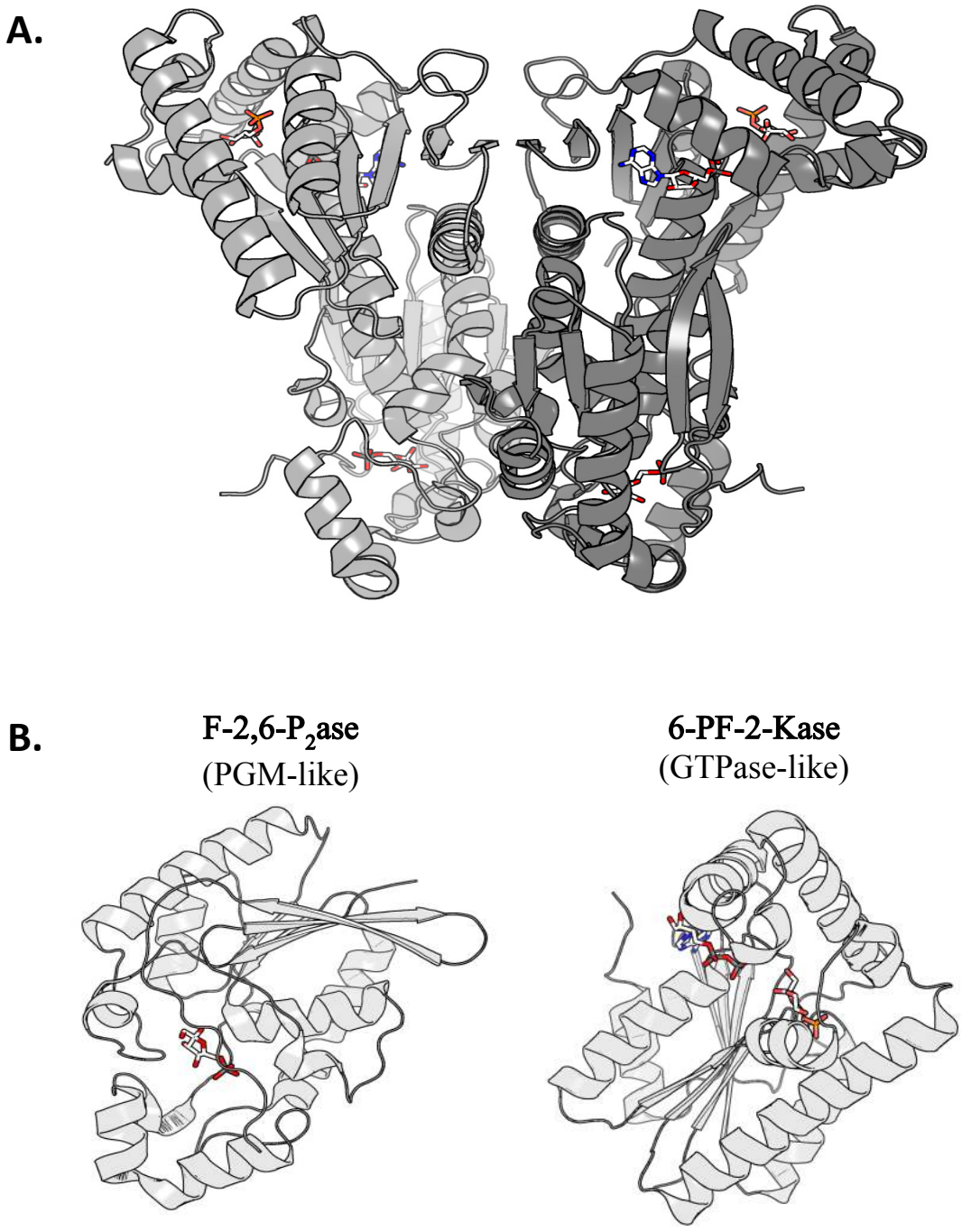


Figure 1.2. Structure of human PFKFB3. (A) Structure of the human PFKFB3 functional homodimer in a head-to-head conformation. Likewise, PFKFB1 and PFKFB4 are also functional homodimers with the same head-to-head arrangement. (B) Ribbon diagrams are used to represent the two catalytic domains of the bifunctional enzyme.

The N-terminal half of the polypeptide forms the kinase domain, which is responsible for the ATP-dependent phosphorylation of F6P to form F-2,6-P_2. The domain consist of a six stranded, predominantly-parallel β-sheet neighbored by seven α-helices. Forming the active site, are the Walker A (GXXXGKT/S) and B (ZZZZD, with Z being any hydrophobic residue) motifs, which facilitate ATP binding as well as the octahedral Mg^{2+} ion coordination between the protein and the β- γ- phosphates of ATP (Hasemann et al. 1996; Ramakrishnan et al. 2002; Walker et al. 1982). Together, these motifs resemble the P-loop NTPase domain of adenylate kinase (Leipe et al. 2003).

For catalysis, the two substrates, ATP and F6P, are positioned and oriented such that a strong, direct, substrate-substrate interaction is achieved between the O_2 oxygen of F6P and the γ-phosphate of ATP (Kim et al. 2007). Upon the substrate-assisted deprotonation of the O_2 oxygen, F6P can perform a nucleophilic attack on the γ-phosphate of ATP. This is followed by an "in-line" transfer of phosphate accompanied with the inversion of the phosphate stereochemistry (Kim et al. 2007; Kitajima et al. 1984; Klahn et al. 2006; Lahiri et al. 2003; Wittinghofer 2006).

Conversely, the C-terminal half of the polypeptide forms the bisphosphatase domain, which catalytically-opposes the reaction performed by the kinase domain. Structurally, the domain is globular, much like the kinase domain, and contains a six-stranded, predominantly-parallel β-sheet neighbored by several α-helices (Furuya et al. 1982; Van Schaftingen and Hers 1981). Overall, the F-2,6-P_2ase domain is homologous with the phosphoglycerate mutase family of enzymes (Bodansky 1972; Fothergill-Gilmore and Watson 1989).

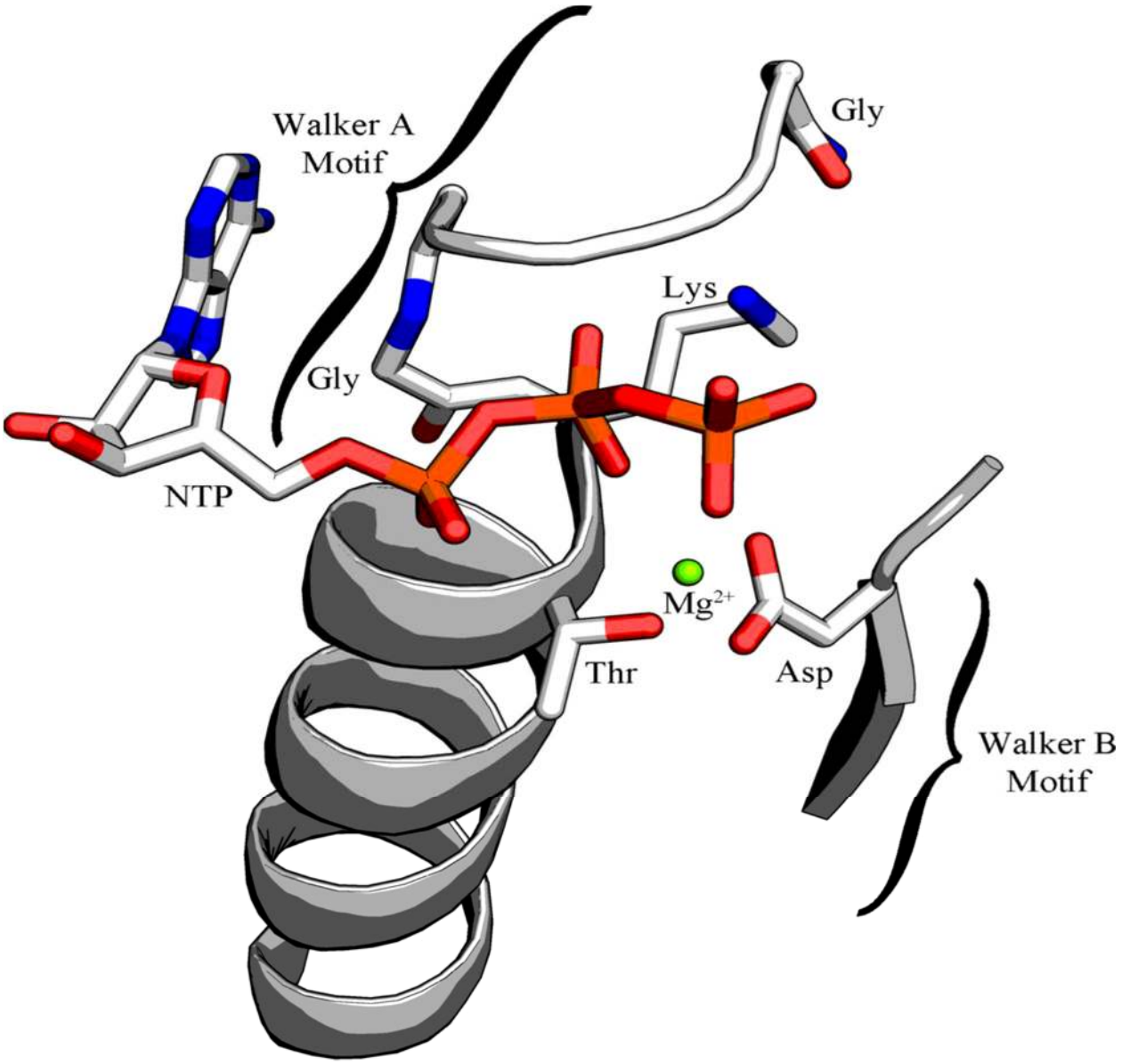


Figure 1.3. P-loop NTPase. Typical binding of a nucleotide triphosphate by the Walker A and B motifs of a P-loop NTPase fold. (Cavalier et al. 2012)

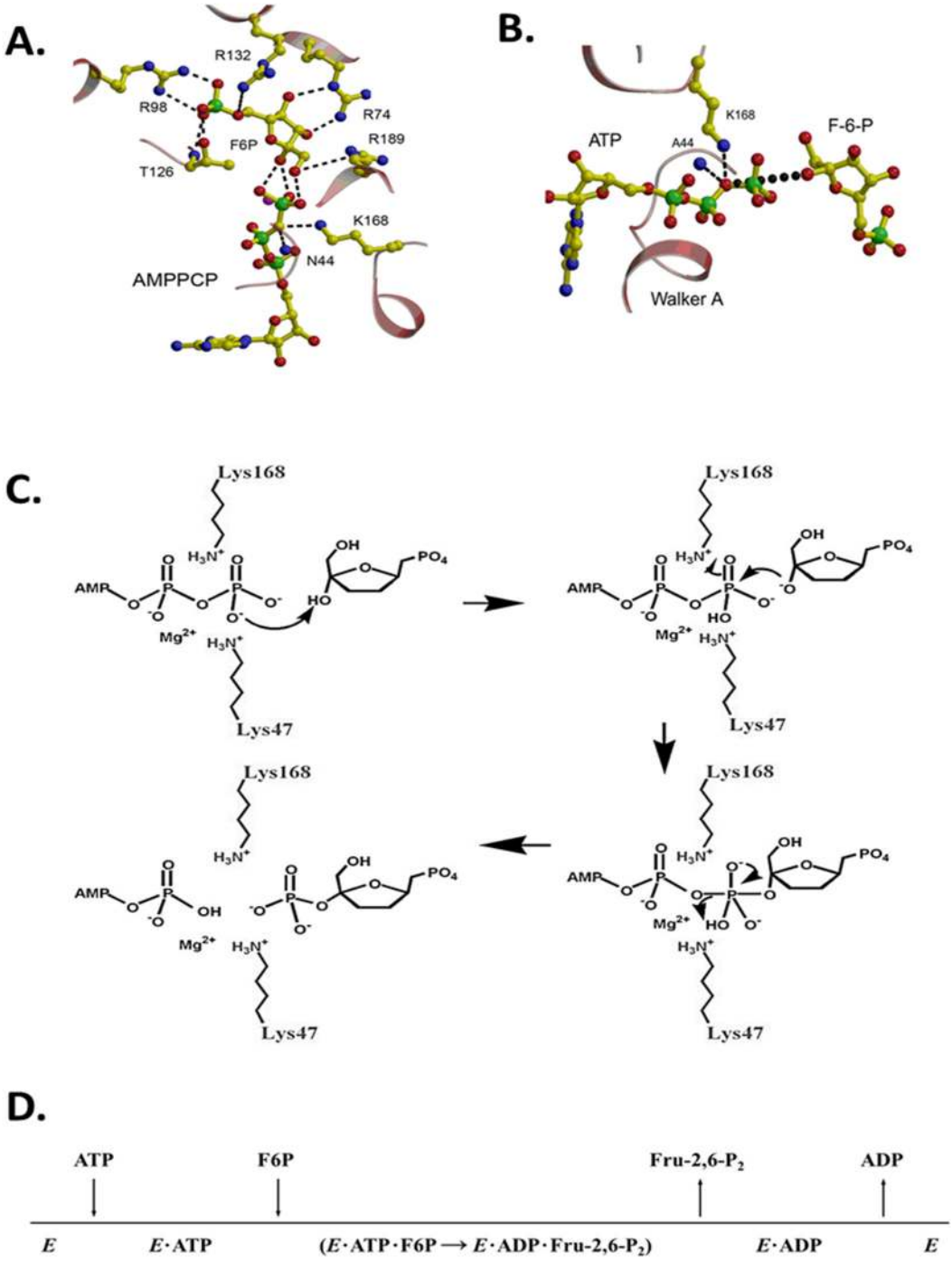


Figure 1.4. Substrate binding and mechanism. (A) Direct substrate/substrate interactions generated from the pseudo-Michaelis complex using AMPPCP and F6P. (B) In-line positioning of the bridge oxygen, γ-Phosphate, and 2-O of F-6-P. (C) Reaction Mechanism of PFKFB3 6-PF-2Kase domain. (D) Ordered Sequential Reaction of the 6-PF-2Kase domain.

In vertebrates, four tissue-specific PFKFB isoforms, encoded by four distinct genes, act to tightly regulate levels of glycolytic flux within particular tissues: liver (PFKFB1), heart (PFKFB2), inducible (PFKFB3), and testis (PFKFB4). Despite having highly similar kinase and bisphosphatase catalytic cores, isoforms differ among their kinase and phosphatase activities, tissue distributions, and its regulatory response to post-translational modifications (O. Minchenko et al. 2003; Pilkis et al. 1995).

1.3.1 Liver Isoform (PFKFB1)

The PFKFB1 gene, which codes for the liver isoform, contains 17 exons encoding three alternatively spliced transcript variants (L, M and F) that differ only within their first exon. The 12 consecutive exons (2-13) shared among all three variants correspond to the kinase and phosphatase catalytic cores, with each being encoded by six exons (Algaier and Uyeda 1988; Lange and Pilkis 1990; Taniyama et al. 1988). The first exon of the L isoform encodes 32 amino acids and gives rise to largest splice variant, which notably contains a phosphorylatable Ser^{32} residue (El-Maghrabi et al. 1982a; El-Maghrabi et al. 1982b). This splice variant is expressed mainly in the liver, but is also found in the skeletal muscle and white adipose tissue. In contrast with the L isoform, the first exon for the M variant only codes for 9 unrelated amino acids that are devoid of a phosphorylation site. The M isoform is predominantly expressed in the skeletal muscle and white adipose tissue. The F mRNAs, on the other hand, are very similar to the M-isoform, the only difference being a truncation of the first exon. Additionally, expression for the F variant is predominant in fibroblasts and fetal tissue (Rider et al. 2004; Ros and Schulze 2013).

Unlike the M and F variants, the L variant is subject to regulatory control via phosphorylation at the N-terminus (Y. H. Lee et al. 2003). Such regulation is observed in response to glucagon, which stimulates a cAMP-dependent protein kinase to phosphorylate an L

variant-specific Ser^{32} (Murray et al. 1984). The result of phosphorylation is an inactivation of 6-PF-2-K activity with a concomitant increase of F-2,6-P_2ase activity, shifting the kinase-to-phosphatase activity ratio from 1.4 to 0.2 (El-Maghrabi et al. 2001). However, because the M and F splice variants lack an N-terminal phosphorylation site, their catalytic actities are not subject PKA mediated regulation. The rationale underlying the differential regulation between splice variants can be understood when comparing liver and muscles isoforms. Because the liver primarily responsible for controlling blood glucose levels, it is imperative that glycolytic flux be regulated so that glucose can be stored when plentiful and released when scares. Muscles, on the other hand, are not responsible for controlling blood glucose levels, reducing the need for such regulation (Darville et al. 1992; Dupriez et al. 1993).

1.3.2 Heart Isoform (PFKFB2)

The PFKFB2 gene, which codes for the heart isoform, contains 15 exons encoding two alternatively spliced transcript variants. The proteins differ at the C-terminal due to an alternative splicing deletion of exon 15, giving rise to a truncated (54 kDa) and full-length (58 kDa) variant. The removal of exon 15, which is responsible for encoding a C-terminal sequence containing several phosphorylation sites, is associated with a loss of regulatory control. For the full length protein, phosphorylation at the C-terminus by PKA is associated with an increase kinase activity and a decrease in bisphosphatase activity, which is in direct opposition to the regulatory effects experienced by PFKFB1 upon phosphorylation (Chikri and Rousseau 1995; Heine-Suner et al. 1998; Ros and Schulze 2013).

As for the overall structure, PFKFB2 shares a highly conserved catalytic core with the liver isoform. The N-terminus of PFKFB2, on the other hand, lacks phosphorylation sites,

making the C-terminus responsible for all regulatory control of the enzyme. (Deprez et al. 1997; Ros and Schulze 2013).

1.3.3 Inducible Isoform (PFKFB3)

PFKFB3, the gene product of pfkfb3, has termed the brain/placental PFK-2, inducible PFK-2 (iPFK-2), and ubiquitous PFK-2 (uPFK-2) due to its ubiquitous expression among tissues. More recently, PFKFB3 has also been referred to as the 'cancer isoform' due to its overexpression and persistent, regulatory activation within numerous tumor cell lines (Kessler and Eschrich 2001; Manzano et al. 1998; Navarro-Sabate et al. 2001). Among the four isoforms, PFKFB3 has been demonstrated to have the highest kinase/phosphatase activity ratio and is thus associated with elevated F-2,6-P_2 levels in tissues where it is expressed. In fact, the bisphosphatase domain of PFKFB3 is almost inactive, making its contribution to F-2,6-P_2 levels negligible when considering the overwhelming high kinase activity of the enzyme (Sakakibara et al. 1999). The 2006 crystal structure of PFKFB3 shed light on the enzyme's unusually high kinase activity as well as its abnormally low bisphosphatase activity (Kim et al. 2006). The structure revealed an N-terminal, β-hairpin motif that makes contact with residues of the bisphosphatase catalytic site, possibly inhibiting F-2,6-P_2ase activity. Moreover, the structure also revealed that the sequence of PFKFB3's kinase domain, as compared with other isoforms, has decreased conformational flexibility in regions near the ATP and F6P substrate loops, likely resulting in enhanced substrate binding (Kim et al. 2006).

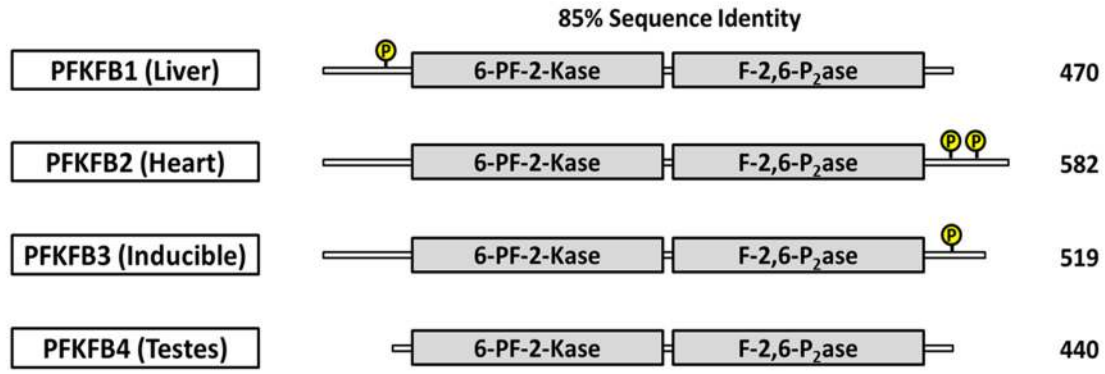
PFKFB3 expression is stimulated by hypoxia through the activation of the HIF-1 transcription factor. In addition, certain oncogenes, such as c-*myc,* c-*ras*, and c-*src,* have been shown to exert complementary effects on PFKFB3 expression as well as cell proliferation (Chesney et al. 1999; Chesney et al. 2005). Regulatory activation, on the other hand, involves

phosphorylation of a C-terminal Ser^{461} by PKA, PKB, PKC, or AMPPK. Such phosphorylation serves to elevate PFKFB3's kinase-to-bisphosphatase activity ratio from 700:1 to 2400:1, resulting in dramatic increases in glycolytic flux (Ros and Schulze 2013; Sakakibara et al. 1997). Studies have found that in many cancers, PFKFB3 is both highly expressed and highly phosphorylated, suggesting that PFKFB3 is strongly linked to the elevated glycolytic activity within cancer cells (Dihazi et al. 2003; Lewis et al. 1997; Shim et al. 1998; Telang et al. 2006)

1.3.4 Testis Isoform (PFKFB4)

The testis isoform of PFKFB, the gene product of pfkfb4, shares a highly conserved catalytic core and C-terminus with the liver isoform. However, unlike other isoforms, PFKFB4 is not subject to regulation by phosphorylation as it is devoid of phosphorylation sites on both N- and C-termini (Sakata et al. 1991). Testis tissues are known to have persistently high glycolytic activity which is in line with the enzyme's kinetic profile. Compared to the liver isoform, the 2-Kase:2-Pase activity ratio is approximately three times higher for the testis isoform. Moreover, because glycolytic activity is persistently high within testis tissues, regulatory control of the enzyme is not needed, explaining the lack of phosphorylation sites (Manzano et al. 1999; Rider et al. 2004; Ros and Schulze 2013).

A.



B.

KINETIC PROPERTIES OF THE PFKFB ISOFORMS

	2-Kase		2-Pase		2-K/2-P Ratio	
	V_{max}	K_m^{F6P}	V_{max}	K_m	Native	Phospho
TESTIS	90	85	22	21	4	NA
LIVER	57	26	45	0.5	1.5	0.44
CANCER	142	11	0.2	> 40	700	2800
HEART	180	50	2.2	40	80	700

Units for V_{max} is U/g and K_m µM.

Figure 1.5. Summary of Phosphorylation Sites and Enzyme Kinetics of PFKFB Isoforms. (A) Schematic showing the empirically known phosphorylation sites of each tissue-specific isoform of PFKFB. (B) Table containing the kinetic properties of each PFKFB isoform.

1.4 Glycolytic Modification in Cancer Cells: The Warburg Effect

While normal cells generate a majority of their energy and biosynthetic precursors through a combination of glycolytic and oxidative metabolism, cancer cells dramatically reprogram their metabolism to strongly favor aerobic glycolysis (R. G. Jones and Thompson

2009; Yeluri et al. 2009). Numerous studies have demonstrated that glycolytic rates in rapidly-growing tumor cells can be up to 200 fold higher than those of their normal tissues counterparts, even in the presence of abundant oxygen. Typically, however, cancer cells in rapid growth phases experience hypoxia ─ an insufficient supply of oxygen ─ and have limited nutrient exchanges with neighboring cells (Bertram 2000; Kim et al. 2006; Pelicano et al. 2006). To survive under such harsh conditions, cancer cells must reprogram their metabolic pathways to be less dependent on oxygen availability and interactions with neighboring cells (Gatenby and Gillies 2004; Moreno-Sanchez et al. 2007). This altered metabolism results from oncogene activation and/or the loss of tumor suppressor genes in multiple signaling pathways.

High levels of glycolytic metabolism confer several competitive advantages to cancer cells. Firstly, although glycolysis yields less ATP than oxidative phosphorylation (OXPHOS), the speed at which energy can be produced is much higher from aerobic glycolysis. Additionally, because cancer cells are less reliant on mitochondrial OXPHOS for energy, fewer reactive oxygen species (ROS) are generated, which are cytotoxic to both normal and cancer cells (Bui and Thompson 2006; Holley et al. 2012). Secondly, increased glycolytic flux provides precursors necessary for the biosynthesis of nucleotides, lipids, and many other metabolites that are essential for survival and proliferation. Thirdly, besides providing cells with ample energy and biosynthetic building blocks, cancer cells suppress the growth of healthy neighboring cells by altering the nearby microenvironments (Pfeiffer et al. 2001). One way they do this is through consuming the majority of diffusible substrates, leaving little for normal cells. A second is through lowering extracellular pH levels, which are established by enhanced lactate production, promoting apoptosis of neighboring normal cells while enhancing invasion and metastasis characteristics of cancer cells (Gatenby and Gillies 2004).

Enhanced glycolysis observed in cancer cells is, at least, partially attributed to an impairment of mitochondrial function, including decreased expression of certain mitochondrial transporters and oxidative enzymes, a truncated TCA cycle, fewer mitochondria per cell, a defective electron-transport chain, higher levels of natural inhibitors for the mitochondrial ATP synthase, and increased sensitivity of mitochondrial DNA to oxidative stress (Bertram et al. 2001 (Bertram 2000; DeBerardinis and Thompson 2012; Xu et al. 2005). Studies have demonstrated that oncogenes such as c-ras, c-src, and c-myc upregulate the majority of glycolytic enzymes, including PFKFB3, hexokinase-2, and lactate dehydrogenase (Levine and Puzio-Kuter 2010). Additionally, the activation of hypoxia-inducible factor-1 (HIF-1) plays a fundamental role in promoting glycolysis in cancer cells. Although HIF-1 is present in normal cells, it is continuously degraded under aerobic conditions. In cancer cells HIF-1α, a regulatory subunit of HIF-1, is overexpressed and, in addition, stabilized by hypoxic tumor microenvironment, resulting in the upregulation of most glycolytic enzymes (Brahimi-Horn and Pouyssegur 2006; Caro 2001; Ke and Costa 2006). This phenomenon of enhanced glycolysis in cancer cells has been termed the Warburg effect (Warburg 1956).

1.5 PFKFB and Cancer Metabolism

In most eukaryotic cells, PFK-1 experiences strong inhibition from cellular concentrations of ATP under physiological conditions; such inhibition often plays a pivotal role in allowing cells to respond to nutritional and hormonal signals at the levels of transcription, translation, and post-translational modification (Ros and Schulze 2013). Cancer cells, on the other hand, frequently bypass this regulatory mechanism by increasing cellular concentrations of F-2,6-P_2 (Yalcin et al. 2009). Numerous studies have demonstrated that the steady-state

concentration of F-2,6-P_2 is much higher among cancer cell lines and is frequently associated with elevated rates of proliferation, cell-cycle progression, and transformation (Clem et al. 2008).

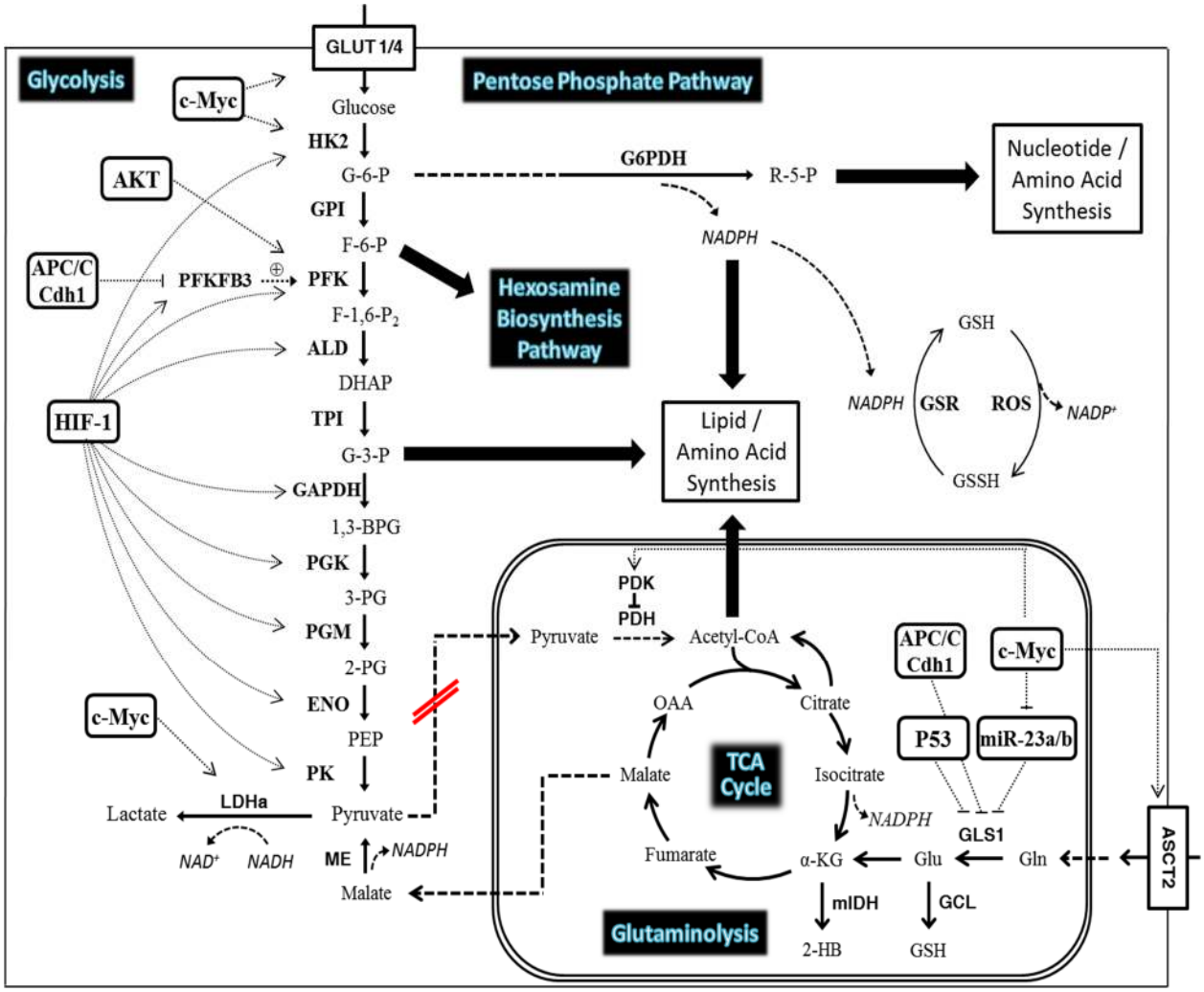


Figure 1.6. Overview of the Metabolic Reprogramming of Tumor Cells. Diagram showing the molecular mechanisms responsible for the Warburg Phenotype.

The inducible PFKFB3 is the predominant isoform in most neoplastic cells including cancer (Chesney et al. 1999; Navarro-Sabate et al. 2001). Unlike the other isoforms, PFKFB3 is ubiquitously expressed in almost all cells via the actions of c-Myc, the downstream transcription

activator of RAS and various growth factors, or the hypoxia-inducible factor 1 upon the onset of hypoxia, a condition common to most fast growing cell populations (Dang and Semenza 1999; A. Minchenko et al. 2002; Mor et al. 2011; Obach et al. 2004). With its exceptionally high 2-Kase activity and low 2-Pase activity, PFKFB3 can maintain F-2,6-P_2 levels over 10-fold higher than other isoforms, providing a clue to the mechanism underlying vigorous glycolysis of cancer, long known as the 'Warburg effect' (Pilkis et al. 1995; Warburg 1956). As such, expression of PFKFB3 is suppressed by p53 and, moreover, in the presence of p53, TiGAR (p53-inducible regulator of glycolysis and apoptosis regulator) down-regulates the cellular F-2,6-P_2 level, to suppress tumorigenic transformation and proliferation (Green and Chipuk 2006; Maddocks et al. 2013; A. Minchenko et al. 2002; O. Minchenko et al. 2003). Moreover, studies have shown that the overexpression of PFKFB3 in several cancer cell lines occurs through a HIF-1 dependent mechanism. Furthermore, it was found that suppression of PFKFB3 via gene silencing resulted in a reduction of cellular F-2,6-P_2 and depressed glycolytic flux, reducing cancer cell viability (Calvo et al. 2006). Since PFKFB3 is the dominant form in most cancer cells, it is an emerging target of cancer chemotherapy.

1.6 PFKFB as a Chemotherapeutic Target

The importance of F-2,6-P_2 to cancer's altered metabolism has made the bifunctional enzyme a popular target for development of anti-neoplastic agents and diagnostic procedures (Ros and Schulze 2013). One such example is a molecular imaging technique known as positron emission tomography (PET), which uses a glucose analog probe, 2-deoxy-2-[18F]fluoro- D-glucose (FDG), to explore the possibility of cancer metastasis on the basis of high glucose consumption associated with the presence of solid tumors (Maddocks et al. 2013; Wechalekar et al. 2005). With regard to anti-neoplastic agents, the dependency of cancer cells on enhanced glycolysis for

energy production, in combination with the suppression of mitochondrial OXPHOS to prevent harmful ROS accumulation, gives rise to several potential targets for therapeutic intervention.

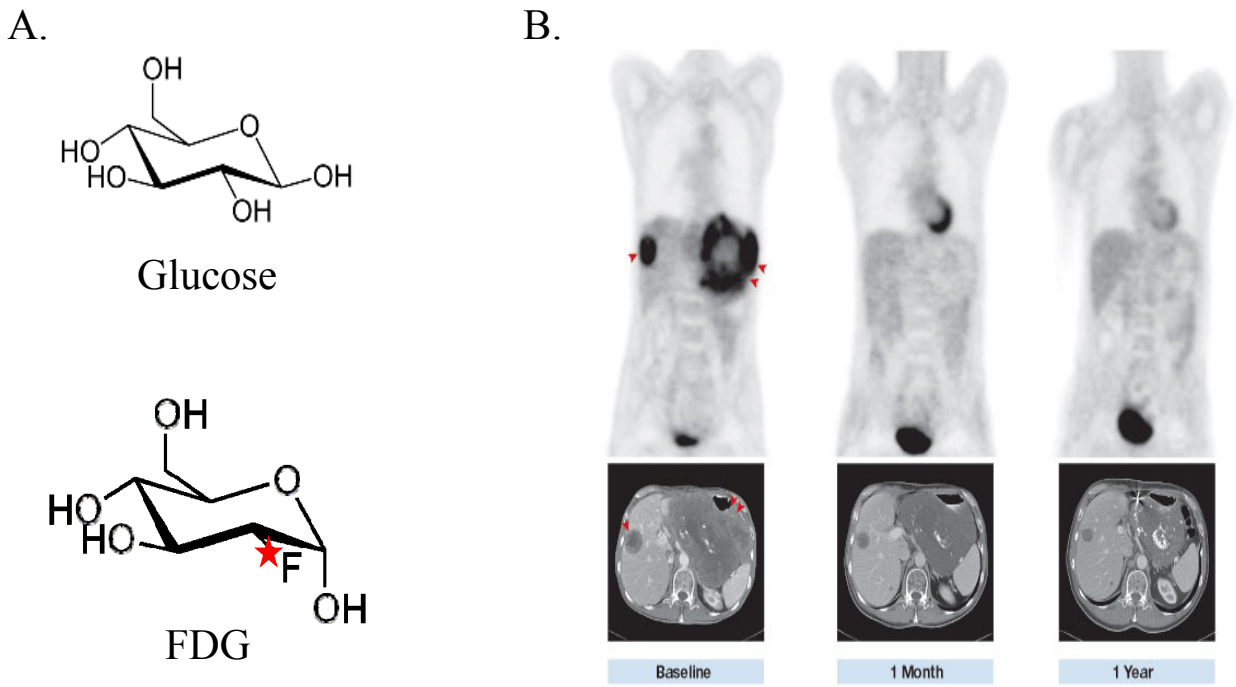


Figure 1.7. Exploiting the Warburg Effect using Positron Emission Tomography (PET). (A) Glucose and the glucose analogue probe FDG (2-deoxy-2-[18F]fluoro- D-glucose). (B) FDG-PET/CT taken at several timepoints during chemotherapy treatment. (Lagaru et al. 2009)

Because PFKFB3 displays the highest 2-Kase:2-Pase activity ratio, it is hypothesized that PFKFB3 may prove to be an attractive target to cancer therapy (Atsumi et al. 2002; Atsumi et al. 2005; Mor et al. 2011). Studies have shown that gene silencing studies involving PFKFB3 results in a reduction of glycolysis and suppression of cancer cell viability and growth, supporting the assumption of PFKFB3 being a promising target for cancer chemotherapeutics. Below, the most promising drug candidates for PFKFB3, and the current stage of each, are briefly described.

3-(3-Pyridinyl)-1-(4-pyridinyl)-2-propen-1-one

3-(3-pyridinyl)-1-(4-pyridinyl)-2-propen-1-one (3PO) was developed based on the structure derived from another isoform, PFKFB4. It blocks fructose-6-phosphate from entering the active site of the kinase domain. *In vivo* studies on neoplastic cells have confirmed that 3PO does suppress glycolysis. Moreover, in studies with melanoma cell lines, it was quantitatively determined that the concentrations of Fru-2, 6-P2, ATP, lactate, NAD+, and NADH are reduced upon 3PO introduction. The isoform selectivity of 3PO is not yet known (Clem et al. 2008).

N4A and N4A Derivatives

Compared to 3PO, N4A (5,6,7,8-tetrahydroxy-2-(4-hydroxyphenyl)chromen-4-one) is a PFKFB3 inhibitor based on the crystal structure of PFKFB3. Its derivatives are suggested to be improved from the subsequent structure-guided optimization. They bind to the F-6-P pocket of PFKFB3 and inhibit its F-2,6-P2 synthesis, leading to low glycolysis and, ultimately, death of cultured cancer cells with no random toxicity. N4A and its derivatives are functionally selective to PFKFB3 and have been shown to induce death of cultured HeLa and T47D cancer cells by efficiently blocking glycolysis (Seo et al. 2011).

PFK-158

PFK158 is a derivative of 3-(3-pyridinyl)-1-[4-pyridinyl]-2-propen-1-one (3PO) that exhibits better potency and pharmacokinetic properties than the parent molecule. PFK-158 was discovered and developed by Advanced Cancer Therapeutics (ACT) and is currently undergoing Phase I clinical trials. Mouse model studies demonstrated ~80% growth inhibition with advanced solid malignancies, with comparable dosages being well tolerated in both rats and dogs. This molecule represents the first-in-man and first-in-class inhibitor for PFKFB3 that

prevents tumor cells from using glucose as a fuel source for survival, growth and metastasis (Chesney et al. 2015; Natarajan et al. 2011).

CHAPTER 2: REVIEW OF LITERATURE

2.1 Drug Discovery and Development

The process of discovering and developing novel therapeutics is complex, expensive, and time-consuming (Dickson and Gagnon 2004b, 2004a; DiMasi et al. 2003). According to a report published in 2014 by the Tufts Center for the Study of Drug Development (CSDD), the cost of developing a prescription drug that gains market approval is approximately $2.6 billion, a 145% increase, correcting for inflation, over the estimates from 2003 (Hughes et al. 2011; Kaitin 2015). However, despite a two-fold increase in production costs, the time associated with bringing a drug to market has remained stagnant, with each new drug taking on average 12-15 years to develop. Furthermore, despite increasing financial investments in pharmaceutical R&D, there has been a steep rise in the attrition rate of drug candidates (C. P. Adams and Brantner 2010; Mahajan and Gupta 2010a, 2010b). This is currently one of the most significant challenges facing pharmaceutical industry as a whole. An overview of the drug development pipeline, from discovery to market approval is shown in figure 2.1.

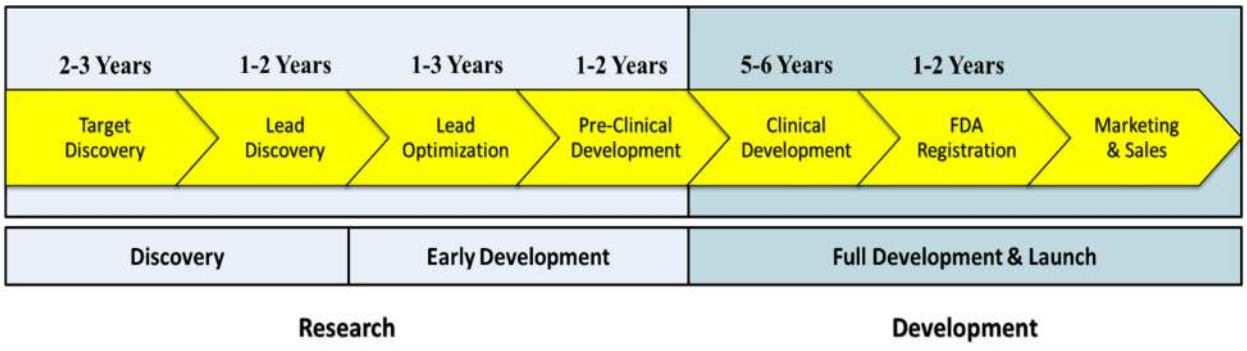


Figure 2.1. Standard Drug Development Pipeline. Model showing the distinct phases of drug discovery and development as well as the associated time and cost.

Initially, there must be a disease or symptom for which there is a need for a novel drug. The first step in the actual discovery process is target identification and validation, which typically occurs via basic research on molecular pathways (Hughes et al. 2011). Once a potential target has been selected, a cell-based assay must be designed so that the biological activity of the compounds can be tested (S. Fox et al. 2006; M. P. Fox and Fox 2006; Moore and Rees 2001). In the lead discovery phase, compounds which modulate the biological activity of the drug target are identified through the screening of large compound collections *in vitro*. Subsequently in lead optimization, the most promising leads are subjected to several rounds of chemical modification to improve potency and pharmacokinetics properties of each compound. Next, the most bioactive compounds generated from lead optimization are tested for efficacy and toxicity in the preclinical development stage using animal models (Hughes et al. 2011). On average, only 1 of every 50 compounds from the pre-clinical phase is tested in human clinical trials, and of these compounds, only 1 of 5 compounds gets registered by the FDA (Hay et al. 2014). Because of the logarithmic nature of compound elimination during drug development, the costs steeply rise after each stage in the process (Lloyd and McElwee 2011). It is therefore of vital importance to drug developers to keep failure rates as low as possible during the later stages of development. Traditionally, this is done by simultaneously testing several protein targets against huge chemical libraries to select the best protein and lead compound set . More recently, drug developers have found that a less expensive solution is to use computational techniques to prioritize drug targets and lead candidates to maximize the likelihood of developing a successful drug.

2.2 Virtual Screening

Over the past two decades, high-throughput screening (HTS) has been the standard method for the identifying new, bioactive compounds. However, because of the complex,

expensive, and time-consuming nature of HTS, computational screening, otherwise known as virtual screening (VS) has become increasingly popular in pharmaceutical research. In essence, virtual screening and high-throughput screening have the same goal (i.e. the identification of novel bioactive compounds (HITs)), yet operate using different philosophies. Whereas HTS aims to experimentally test increasingly large numbers of compounds in the most efficient manner, virtual screening deviates from this 'screen everything' paradigm and attempts to rationalize compound selection strategy to reduce the number of candidates for experimental evaluation and promote the prioritization of physical resources to 'high-potential' candidates (Braga et al. 2014; Cheeseright et al. 2011; Ripphausen et al. 2011). Despite the critical role of HTS in the early stages of drug discovery, the many technical advances, and the tremendous improvements in HTS capacity, VS has still become an established discipline in pharmaceutical research (Bajorath 2002; Kitchen et al. 2004; Macarron 2006; Macarron et al. 2011; Macarron and Luengo 2011).

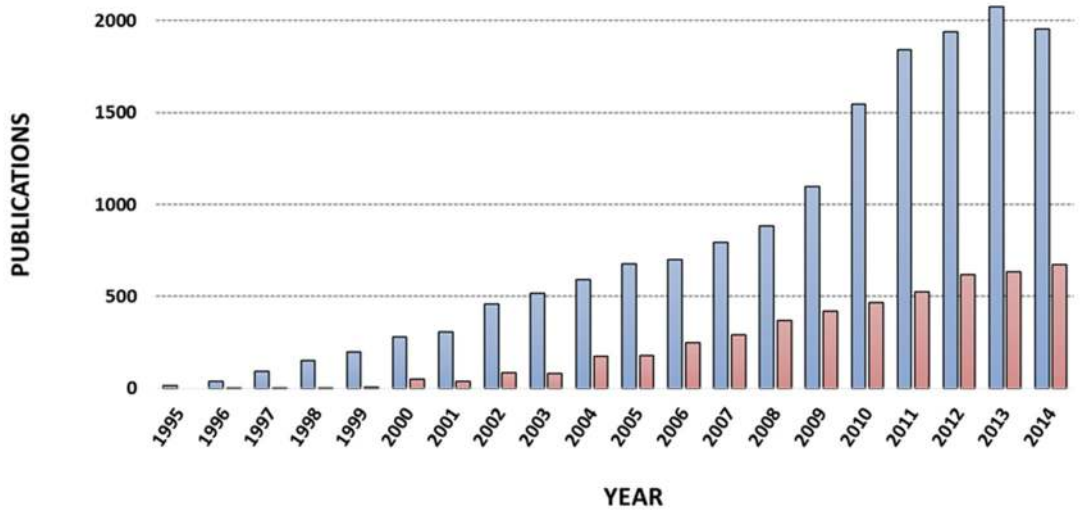


Figure 2.2. Chronological overview of the number of yearly publications for high-throughput (Blue bars) and virtual screening methods (Red bars).

Virtual screening consist of a spectrum of computation methods designed to efficiently search large chemical libraries *in silico* for compounds likely to be active for a biological target of interest. The efficiency of this technique depends on both performance of the screening methods and the content of the database to be screened (Polishchuk et al. 2013). Traditionally, the practice consisted of searching a company's propriety database of compounds comprising the structures of millions of real molecules from a corporate collection (Cheeseright et al. 2011). Over time, as methods have improved, so too has the size of the chemical database that can be feasibly screened. Virtual screening is now being used to screen hundreds of millions of virtual compounds, created from enumerating combinatorial libraries in chemical space (Boehm 2011; X. H. Ma et al. 2012b). HTS, however, can only test compounds that are physically acquirable. Herein lies a tremendous advantage unique to virtual screening. As of 2013, an ensemble of academic, commercial and propriety databases record approximately 60 million structures of existing chemical compounds. Chemical space, on the other hand, which represents all possible (i.e. energetically stable) chemical compounds is estimated to be ~10^{33}, making the number of compounds known to exist less than 1 in 6×10^{24} of those that presumed possible (Polishchuk et al. 2013; Reymond and Awale 2012). Virtual screening methods are largely into two approaches, ones that take the 3D structure of the target protein into account (structure-based drug design, SBDD), and ones that rely on information derived from one or more known bioactive molecules (ligand-based virtual screening, LBVS). Generally speaking, SBDD aims to identify bioactive small molecules from 'docking' them into a binding site of a target protein (Kitchen et al. 2004; Lyne 2002). LBVS, on the other hand, uses physicochemical properties from one or more known bioactive molecules to identify novel structures with the desired bioactivity (Eckert and Bajorath 2007). Such methods are not mutually exclusive, and given

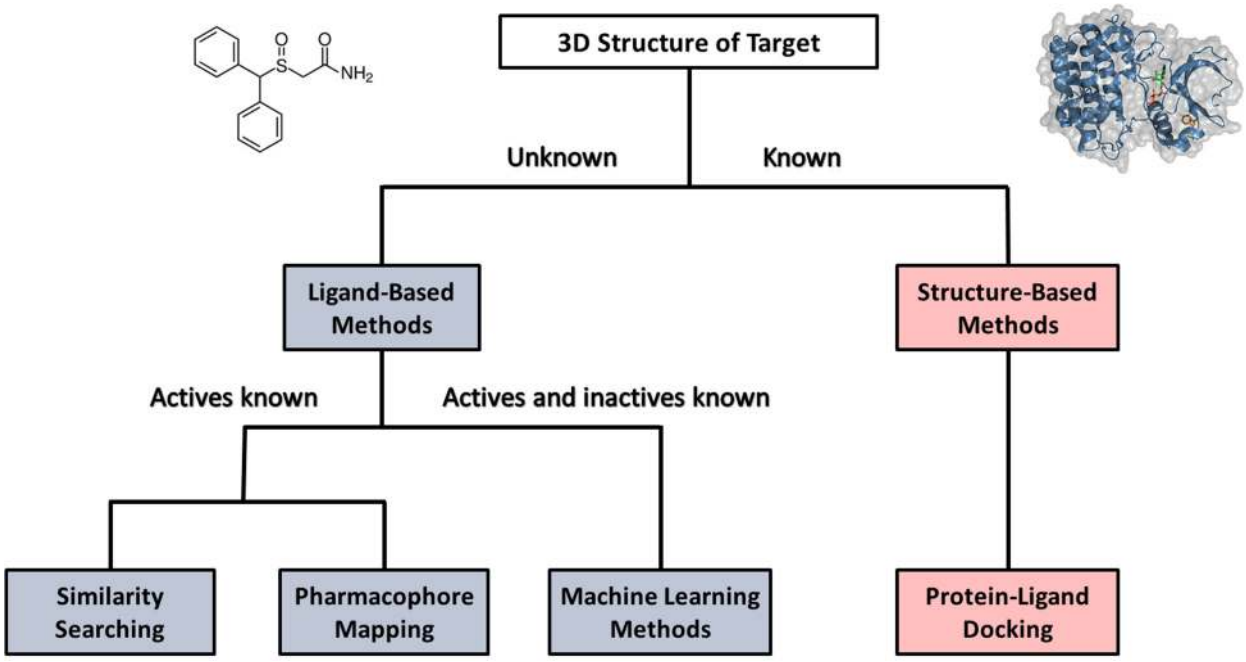


Figure 2.3. Approaches to Virtual Screening. Shown are the two main categories of virtual screening: structure-based and ligand-based.

the prerequisites of each, can be used in concert (D. L. Ma et al. 2012a; Peltason 2008). The two principal types of virtual screening are discussed below.

2.3 Docking-Based Approaches

Molecular docking can broadly be defined as a computational tool ─ belonging to the field of structure-based drug design ─ that attempts to exploit the enzymatic concepts of lock-and-key and/or induced fit binding models to provide useful predictions about potential ligand-protein (or RNA) interactions (Berg 2012; Blaney 2012; Hughes et al. 2011; Yuriev et al. 2011). Since its beginnings in the early 1980's, docking, along with tremendous developments in chemistry, physics, informational technology, biochemistry, and computer technology, has become an essential tool in drug discovery (Chen 2015; Ferreira et al. 2015). Initially, docking studies served primarily as a means to predict binding poses and estimate approximate binding

affinities between protein-ligand complexes from scoring interaction energies (Ferreira et al. 2015). Early successes in binding pose predictions set the groundwork for future development on docking programs and led to the exploration of docking as a tool to identify HITs during the early stages of drug discovery (Brooijmans and Kuntz 2003; Krumrine et al. 2003). Now, docking is considered an essential tool for lead discovery and is associated with a myriad of success stories and FDA approved drugs (Millan 2012).

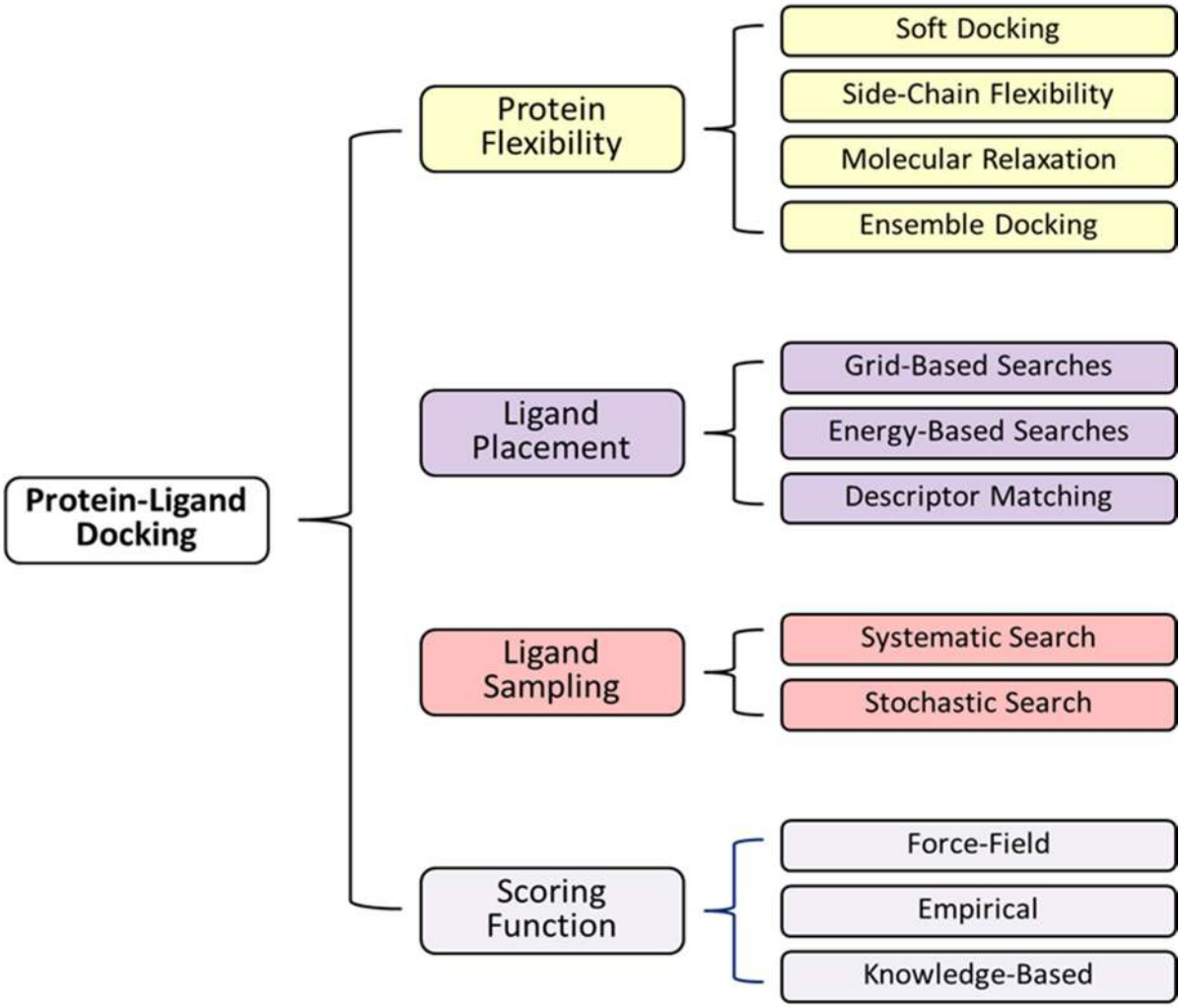


Figure 2.4. Classification of methods for protein-ligand docking.

In molecular docking, the 3-demensional shape of a binding site is used to optimize the ligand's position and orientation to best fit the receptor. In biological systems, the driving force facilitating such interactions arises from complementarities in both shape and electrostatics of the ligand to the binding site surface (Bikker et al. 2009; Halperin et al. 2002). In addition to coulombic interactions such as hydrogen bonds, salt bridges, π-cation, and π-π interactions, hydrophobic forces such as Van der Waals interactions also contribute to binding affinity (Bikker et al. 2009; Landry and Gies 2008). In reality, such interactions are best described by quantum mechanics; however, due to the complexity of biological systems, quantum theory cannot be applied directly (Gao et al. 2002; Goodwin et al. 2002).

Molecular docking can be divided into two separate problems: sampling and scoring. Sampling refers to the generation of putative ligand binding conformations/orientations near a binding-site of a target receptor and can be subdivided into two facets, ligand sampling and protein flexibility (Bikker et al. 2009; Ferreira et al. 2015; Krumrine et al. 2003). Scoring, on the other hand, refers to the evaluation of binding tightness of an individual ligand for a particular receptor using either physical or empirical energy scoring functions. During a docking campaign, sampling and scoring processes are carried out recursively for each ligand until convergence at a minimum energy conformation. The top binding pose, namely the one with the lowest fitness score, is predicted to be the binding mode for the compound (Danishuddin and Khan 2015; Halperin et al. 2002). Predictions of absolute binding free energies, however, remains a significant limitation of docking methods due to simplifications and approximations made to reduce computational demands. Below we will examine each of the major stages. Additionally, the importance, potential, and drawbacks of side chain and backbone flexibility in

structure based docking will also be briefly explored (Cavasotto and Orry 2007; Gane and Dean 2000; Huang et al. 2010; Orry et al. 2006).

2.3.1 Protein Flexibility

Until recently, docking algorithms exhibited little or no flexibility of molecules in the active site, typically treating proteins as completely rigid structures. However, the basic assumption made by these programs, in that proteins are rigid molecules, is a serious limitation because induced fit and conformational flexibility are often critical factors in complex formation (B-Rao et al. 2009; Durrant and McCammon 2010). This is problematic since the process of docking uses crystal-structures and/or homology models, which are only a single representation of one of the many potential protein conformations found in solution. Hence, adjustments to flexible side chains in the binding site are often necessary for proper ligand binding, which is underscored by the results from "cross-docking" experiments (Bissantz et al. 2000; Warren et al. 2006). In cross-docking tests, a flexible ligand is docked into a receptor crystal structure that was crystallized with another ligand bound. From experiments using cross-docking, it was shown that success rates in reproducing the known bound conformation of the ligand is significantly lower in cross-docking experiments than in regular docking experiments (Durrant and McCammon 2010).

Although the existence of ligand-induced flexibility was reported in scientific journals nearly 50 years ago, for a very long time, it was generally considered to be rare or inconsequential. However, recent statistical studies of the Protein Data Bank revealed that nearly 85% of all known proteins contain between one and three flexible residues in the active site. Additionally, another study reported that rigid receptor dockings predictions are incorrect in predicting the binding pose in 40–60% of all studied cases (B-Rao et al. 2009). Many factors

have hindered the development of methods for dealing with ligand-induced protein conformational changes. Among these is a lack of a clear understanding of what kinds of changes may be induced by a ligand in a given protein. Additionally, the numerous degrees of freedom involved in incorporating protein flexibility during docking have posed computational challenges (Barril and Fradera 2006a; Carlson 2002b, 2002a; Sotriffer 2015).

Dealing with protein flexibility is a multi-variable problem that entails everything from side-chain movement, to backbone-loop movements, to major domain rearrangements (Halperin et al. 2002). The earliest attempts to mimic small amounts of protein flexibility applied soft potentials to rigid receptors to allow for some overlap toleration between the ligand and the protein (Lexa and Carlson 2012; May and Zacharias 2005). Such soft scoring functions allowed for some plasticity of the receptor but are limited in their scope to small levels of protein flexibility. Although limited in their potential for simulating true flexibility, soft docking methods are computationally efficient and are still being actively used (Mangoni et al. 1999).

Survey methods are a more recent alternative to modeling protein flexibility. They come in a variety of implementations, but in general, they are similar in principle to the approach of exploring multiple ligand conformations for simulating ligand flexibility (Barril and Fradera 2006b; Corbeil and Moitessier 2009; Corbeil et al. 2009). One type of survey method for modeling protein flexibility deals with the incorporation of rotamer libraries to simulate side-chain flexibility (Koveal et al. 2013). However, since this model can't handle backbone changes, it fails to fully simulate ligand-induced flexibility (McGovern et al. 2002). Another survey method attempts to simulate flexibility by using ensembles of rigid-receptors. Studies on the use of ensembles to simulate protein flexibility have generally found that ensemble-based methods are more accurate in ranking scores, yet are computationally more expensive and also require

multiple high-resolution crystal structures (B-Rao et al. 2009). The last major survey method used to represent protein flexibility incorporates molecular dynamic simulations to simulate both backbone and side chain movement. In this method, induced fit is able to be accounted for as well as explicit solvents and ionic effects. The drawback to MD based survey methods resides in the extremely high computational costs, which are impractical for high-throughput screening jobs even when using computing clusters (Cosconati et al. 2012; Fischer et al. 2014).

Because of the complications associated with protein flexibility, we have only recently attained the requisite computational resources for flexible docking to be considered a feasible approach for high-throughput screening (Huang et al. 2010). However, despite having now acquired the computational means for flexible docking, that is not to say that the problem is solved. The jump from rigid to flexible/semi-flexible receptors is just problem of computational resources, but also of complexity (Fischer et al. 2014). By allowing for receptor flexibility, there are enormous increases in the degrees of freedom in the binding calculations; which is taking a complex system and adding additional complexity. However, despite the difficulties, the potential exist for dramatic improvements in binding predictions from incorporating flexibility. By increasing the variables in the system, we are better able to mimic real-world systems, which should ultimately allow for binding predictions to more closely correlate to reality (Huang et al. 2010; Sotriffer 2015).

2.3.2 Search Algorithms

Docking algorithms generally consist of two separate search engines; one for the placement of a ligand within the receptor's binding site and the other for sampling the ligand's internal degrees of freedom that arise from rotational bonds, stereoisomers, and alternative ring conformations (Bikker et al. 2009; Dias et al. 2008a; Dias and de Azevedo 2008; Yuriev et al.

2011). The first process, refers to the orientation algorithms, which describe the surface of the protein using mathematical models. Generally this class of algorithms can be divided into three categories: grid search, descriptor matching, and energy-based searches. The second process, termed ligand sampling, is responsible for dealing with the flexibility of ligands and can be divided into two categories: systematic searches and stochastic searches (Ferreira et al. 2015; Halperin et al. 2002).

2.3.2.1 Ligand Orientation Algorithms

GRID-BASED Grid-based searches are a class of orientation algorithms that exhaustively search the binding site by moving a rigid ligand through the six translational degrees of freedom associated with three-dimensional space in a systematic, stepwise fashion (Ferreira et al. 2015; Halperin et al. 2002; G. S. Wu et al. 2003) For this method, performance and accuracy are directly correlated with the size of the search grid and the step increments. Traditionally, such methods were developed to search through Cartesian space; however, more recent methods typically favor the use of Fourier space as a means to reduce search times (Brooijmans and Kuntz 2003; Millan 2012; G. S. Wu et al. 2003). However, due to the large number of degrees of freedom associated with the flexibility of ligands, this method is typically used for protein-protein docking where both molecules are treated as rigid bodies (Cai et al. 2002; Danishuddin and Khan 2015; Millan 2012; Yuriev and Ramsland 2013).

DESCRIPTOR MATCHING Descriptor matching searches that aim to be more efficient than grid-based approaches by representing the active site as a spacial array of descriptor features opposed to a search grid of the protein's surface (Ferreira et al. 2015; McGaughey et al. 2007). To accomplish this, descriptor matching algorithms uses receptor-generated property features similar to those in pharmacophore screening, these descriptors allow ligands to be matched and

oriented, based on predefined tolerances. Typically, large numbers of orientations are possible, in which case, each is explored, optimized, and scored to generate the final pose position(s) (Brooijmans and Kuntz 2003; Millan 2012). The two most commonly cited technologies to employ this search engine are FlexX and DOCK, both of which operate by means of fragment-growing. Despite their differences, these programs all share a generalized scheme for descriptor matching; usually sets of at least three to four receptor features are matched to the same number of complementary ligand features, after which an assessment is made to evaluate how well the ligand atoms match those of the receptor, both geometrically as well as chemically (Dias et al. 2008b; Huang et al. 2010; Warren et al. 2006). However, because of the nature of descriptor matching, these methods require that subjective decisions be made on the importance of the various functional groups contained within ligands, because a good choice of the base fragment is essential for the proper placement and orientation of the ligand (Ferreira et al. 2015). Hence, poor choices in the weighting of functional groups can significantly affect the quality of the docking process.

ENERGY-BASED METHODS Energy-based methods, also known as force field-based methods, are a class of orientation algorithms that utilize molecular mechanics force fields, in conjunction with molecular dynamics or energy minimization algorithms, to explore the energy landscape of each ligand's surface (McGaughey et al. 2007; Warren et al. 2006). In this method, an energy minima for a ligand on the surface of a protein must is first identified and then assessed in terms of complementarity to the receptor (Huang et al. 2010). Notably, such energy minimizations are a type of local search and thus are only appropriate for optimizing binding conformations generated by other search engines. Despite the drawback of being a local search method, force-field based approaches have proven useful when combined with other search

algorithms that prevent local minima traps. Such modified search algorithms can be seen in AutoDock, GOLD, and Monte Carlo searches (Bikker et al. 2009; Dias and de Azevedo 2008; Ferreira et al. 2015).

2.3.2.2 Ligand Sampling Algorithms

SYSTEMATIC These algorithms aim to explore all the degrees of freedom in a molecule by imposing slight, incremental variations in the structural parameters to generate all possible ligand binding conformations (Brooijmans and Kuntz 2003; Huang et al. 2010). The algorithm operates by probing the energy landscape of the conformational space and, through numerous iterations of search and evaluation cycles, converges to the minimum energy solution (Ferreira et al. 2015). For docking, there are three types of systematic searches: exhaustive, fragmentation, conformational ensemble.

EXHAUSTIVE METHODS Exhaustive searches are the most straightforward of the systematic algorithms, in that flexible-ligand docking is carried out by systematically rotating all rotatable bonds within a ligand at a given interval. The obvious advantage of this method is its sampling completeness, which can be theoretically perfect given infinitely small increments. Practically, however, the method is limited by its thoroughness, with higher dimensionality from increasing rotatable bonds resulting in a conformational explosion. To address this limitation, some docking programs utilize geometric/chemical constraints to limit conformer generation during the initial screening stages. FRED and Glide are two popular examples of method.

FRAGMENTATION METHODS Fragmentation searches utilize incremental construction algorithms for the placement and subsequent construction of ligands within the active site. In this process, ligands are initially divided into rigid and flexible regions with one or more rigid

fragments being designated as an 'anchor' (Brooijmans and Kuntz 2003; Kitchen et al. 2004). Once the rigid anchors have been defined, they are then docked into the active site. After, a fragment is added in all possible conformations to all placements. The process repeats, with only the best placements taken to the next iteration, until each fragment has been placed and the ligand is fully reconstructed (Daga et al. 2014; Rarey et al. 1996).

CONFORMATIONAL ENSEMBLES In conformation ensemble methods, libraries of pre-generated conformations are used to represent ligand flexibility. Library conformations are typically only calculated once, which serves to reduce computational costs associated with docking at expense of disk space and granular optimizations. In such cases, exhaustive, systematic screening is feasible, making stochastic methods superfluous (Danishuddin and Khan 2015; Yuriev et al. 2011). Similarly, ligands devoid of rotatable bonds also lack internal degrees of freedom, thus requiring the search algorithm to only search the six-translational and rotational degrees of freedom in the target binding site (Bikker et al. 2009; Ferreira et al. 2015).

STOCHASTIC In stochastic methods, ligand binding conformations and orientations are sampled through the introduction of random changes to conformational and rotational/translational spaces of a ligand. For this, the algorithm generates an ensemble of ligand conformers that populate a wide range of the energy landscape. This strategy is important in that it helps avoid the final solution being trapped in a local minima as opposed to the global energy minima. A pre-defined probability function is then used to evaluate each conformer and decide whether it should be accepted or rejected. In molecular docking, there are four major implementations of stochastic algorithms: Monte Carlo methods, evolutionary algorithms, swarm optimization methods, and Tabu search methods.

MONTE CARLO METHODS In Monte Carlo-based methods, an initial configuration is iteratively refined through a series of random steps that can either accepted or rejected based on a Boltzmann probability function (Metropolis criterion). In the context of docking, Monte Carlo-based methods treat ligands holistically, with new configurations being randomly from perturbations in rotational, translational, and torsional degrees of freedom. After each move, the system is minimized and the energy of the new system is calculated. The acceptance or rejection of the new system is controlled by the Metropolis algorithm (Metropolis et al. 1953), which ensures that systems with more favorable configurations (lower energy configurations) are always accepted, while systems higher in energy are only probabilistically accepted. Consequently, since some higher energy systems are accepted, the search can also proceed "uphill" on the energy landscape, allowing for some energy barriers to be crossed to find a deeper energy minima.

GENETIC ALGORITHMS Genetic algorithms are class of stochastic optimization techniques that aim address the high computational cost associated with stochastic methods with methods inspired by the principles of evolution and natural selection. In this method an initial population of ligand conformers is randomly generated, where the orientation, position, and conformation of each conformer is encoded in a chromosome of genes for rotation, translation, and torsion variables. Next, a set of genetic operators (mutation, crossover) is then applied to this population to produce a new generation, whereby individuals with superior fitness have a higher likelihood of 'carrying over' into the next generation. This process continues until a particular termination criterion is reached, such as a minimal root-mean-square deviation (RMSD) of the fittest individual, a constant optimal fitness, or simply a maximum number of generations.

SWARM OPTIMIZATION METHODS Swarm optimization methods attempt to discover an optimal solution in a search space by modeling swarm intelligence. In the method, movements of a ligand mode, represented by particles, through the search space is influenced by its own search history and by the information of the best positions of its neighbors (Huang et al. 2010; Janson et al. 2008).

TABU SEARCHES In Tabu search methods, the probability of acceptance of a configuration is dependent on the previously explored areas in a ligand's conformational space. The random change will be accepted if the RMSD between the current ligand binding conformation and any of the previously recorded solutions is more than the cutoff; otherwise, the random change will be rejected (Huang et al. 2010; Westhead et al. 1997).

2.3.2.3 Scoring Functions

Current docking technologies typically utilize scoring functions in one of two ways. One uses the full scoring function to evaluate a protein-ligand conformation, after which, a subsequent round of modifications to the system is carried out by the search algorithm, and then the same scoring function is re-applied to evaluate the new structure The alternative, instead, applies a two-tier scoring function, in which, a reduced function is used to direct the search and is followed by a more rigorous scoring function to rank the final structures. To estimate the binding potential for a ligand to a protein's active site, current docking technologies typically employ any of three differing classes of scoring functions: force-field methods, empirical methods, and knowledge-based methods (Charifson et al. 1999; Clark 2002; Moitessier et al. 2008; Muegge et al. 2001; Muegge and Rarey 2001; Sprous et al. 2002).

Force Field Force field-based scoring functions are a class of scoring functions that estimate the binding affinity of a protein-ligand complex by summing the contributions of bonded (bond stretching, dihedral variation, and angle bending) and nonbonded terms (van der Waals and electrostatics) in a general master function (Ferreira et al. 2015). The Coulomb and Lennard-Jones potentials, representing the electrostatic and van der Waals terms respectively, commonly used as nonbonded interaction terms in force fields are well suited to sensitively monitor the geometric quality of a binding mode. Such force field functions are derivative of both experimental and *ab initio* quantum mechanics; as they depend on empirical parameterization, care must be taken to ensure that the force field is adequately parameterized for the molecules being investigated (Danishuddin and Khan 2015; Sotriffer 2015).

Many docking programs make use of force field terms to score and optimize binding poses, with Glide (Friesner et al. 2004) and AutoDock (Goodsell et al. 1996; Morris et al. 1996) being two prominent examples. However, pure force field-based scoring functions suffer from a major limitation, in that they ignore most solvent effects and solute entropies, and the calculated scores are just energies or enthalpies, opposed to free energies. Hence, force field terms alone are not sufficient for ranking different ligands or estimating binding affinities. Accordingly, for such a task, desolvation terms, at a minimum, are needed for accurate predictions. Shoichet et al was the first to add the effects of the solvent on protein-ligand interactions using an implicit solvent model (Bikker et al. 2009). Generalized Born/Surface Area (GB/SA) (Ghosh et al. 1998) and Poisson-Boltzmann/Surface Area (PB/SA) (Baker 2005) are archetypal examples of implicit solvent models incorporated into docking programs (Mysinger and Shoichet 2010; Wei et al. 2002). In addition to solvent related limitations, pure force field based scoring functions also suffered from charge bias, causing a phenomenon in which charged molecules tended to outscore

other less charged molecules (Danishuddin and Khan 2015). Newer strategies in force field based scoring functions now typically employ a tiered approach to score and post-process the generated pose predictions while also taking into account implicit solvent contributions such as those described in the Generalized Born and Poisson-Boltzmann methods (Huang et al. 2010; Mysinger and Shoichet 2010). These enhanced scoring functions typically result in better agreement with true binding free energies and the predicted rankings of ligands.

Empirical Regression analysis among structural descriptors and binding data of protein-ligand complexes provides another approach to scoring functions. Because of their reliance on experimental data for both crystal structures and binding energies, the resulting functions are typically termed "empirical scoring functions" (Sotriffer 2015; Wang and Busemeyer 2013). Empirical scoring functions are based on the assessment of localized, chemically intuitive interactions between a ligand and protein in a noncovalent complex. Following the assumption of additivity, the scoring function is fit to reproduce experimental data, such as binding affinities, by summing a set of weighted empirical energy terms such as, electrostatic energy, hydrogen bonding energy, van der Walls energy, entropy energy, desolvation energy, and hydrophobicity energy (Kitchen et al. 2004; Sotriffer 2015). The conceptual design of empirical scoring functions derive from the idea that binding energies can be approximated by a sum of individual, uncorrelated terms. When dealing with many complexes of known structure and affinity (training set), the descriptors are first calculated and then correlated with the experimental binding affinity by assigning weights to regression analysis process or other related statistical methods, such as machine learning techniques (e.g., Support Vector Machines) (Ballester and Mitchell 2010; Sotriffer 2015; Zilian and Sotriffer 2013). The resulting function is then fit to score protein-ligand complexes and obtain binding energy

estimates. The quality of empirical scoring functions critically depends on the size, composition, and quality of the training set, with respect to both structure and affinity data (H. J. Bohm et al. 1999; Kalliokoski et al. 2013; Tame 1999).

Knowledge based Structural databases of protein-ligand complexes contain a wealth of information beyond the value of the individual structure. Knowledge-based scoring functions make use of this structural information. In essence, knowledge-based functions are aim to reproduce experimental binding conformations in contrast to empirical functions which are trained to reproduce binding energies. Generally, this type of scoring function is modeled using relatively simple interaction-pair potentials which are based on their frequency of occurrence in the training set being used (Kitchen et al. 2004; Sotriffer 2015). The method operates on the assumption that close intermolecular interactions between certain types of atoms or functional groups occur more frequently than one would expect by a random distribution are more likely to be energetically favorable and therefore contribute positively to binding affinity(Wang and Busemeyer 2013). Compared to empirical and force field scoring functions, the knowledge-based scoring functions offer a good balance between speed and accuracy. As with empirical scoring functions, a disadvantage of knowledge based scoring functions is that fundamentally rely on information implicitly encoded in limited sets of protein-ligand complex structures, making their predictiveness limited by the training set size and quality. (Ferreira et al. 2015; Huang et al. 2010).

2.4 Ligand-Based Approaches

Ligand-based approaches capitalize on the fact that ligands similar to bioactive ligands are, themselves, more likely to be bioactive than random ligands. Generally, such approaches commonly consider 2D or 3D chemistry, shape, electrostatics, and interaction points to assess

similarity, and are most frequently employed when a particular biological target is unknown or when the crystal structure of a known target is unavailable (Acharya et al. 2011; Bacilieri and Moro 2006; Maggiora and Shanmugasundaram 2011). The major focus lies of these methods lies on three related issues: (i) measure ligand similarity fast and accurately, (ii) devise techniques for molecular representation, alignment, and feature matching, (iii) assess relationships between structure and activity (SAR) using information contained within the ligands (Braga et al. 2014; Maggiora and Shanmugasundaram 2011).

2.4.1 Similarity Based Methods

For any two molecules, the number of features that can be directly compared are quite large (i.e. shape, size, charge, conformation, etc.). In addition, numerous descriptors and algorithms can be used to capture and quantify essential characteristics that can be used for similarity metrics (Amaratunga et al. 2015; Braga et al. 2014; Wawer et al. 2008). For these reasons, many types of similarity measures exist.

2D Fingerprints and Structural Keys Fingerprint screening, is a fast, relatively simple screening method that using binary strings ─ encoding the presence or absence of substructural fragments ─ to measurement intermolecular structural similarity between a target and reference molecule (Cabrera et al. 2015). Using a fragment dictionary, a set of chemical features for each molecule is defined and a bit set to either zero (0) or one (1), depending on whether the substructure exists in the molecule or not. This gives rise to a chemical fingerprint holding molecular descriptors for the target and reference molecules in the form of a binary bit string (Willett 2006). A comparison of the bitstrings of each molecule (target and reference), allows for a calculation of the similarity among the two molecules using the Tanimoto coefficient. Despite 2D fingerprints having proved valuable to drug discovery projects, they suffer from

several weaknesses. For example, changing a single atom in a ring structure may change the fingerprint from being nearly identical to almost completely different. Furthermore, 2D fingerprints lack spacial information. Two compounds that have very different structural connectivity can nonetheless adopt similar 3D orientations and thus confer similar biological effects (Braga et al. 2014).

3D Shape Methods 3D shape-based approaches, or Gaussian-based methods, represent each atomic site within a molecule with a spherically symmetric Gaussian function and measures the volume over all Gaussians. Using the amount of volume overlap, estimations of shape similarity can be assessed and scored (Cabrera et al. 2015; Haigh et al. 2005). However, calculations of volume overlap between two molecules are highly dependent on their relative position in 3D space, thus, for 3D-shape based methods to accurately assess similarity, they must first obtain the best possible superimposition through an optimization or fitting procedure (Ballester et al. 2009; Grant et al. 1996). For scoring, only volume overlap of the optimally aligned molecules is calculated, making the comparison practically independent of the atom types and bonding patterns present in both query and search molecules. As with the fingerprint methods, the Tanimoto coefficient is the most common metric used to asses volume overlap (Peltason 2008; Peltason et al. 2010; Peltason and Bajorath 2011).

2.4.2 Pharmacophore Based Methods

The concept of the pharmacophore, in the modern sense, was first put forth in the late 1960s by Monty Kier. His idea, simply stated, was to create a 3-D spatial arrangement of functional groups essential for biological activity; a three–dimensional pattern that emerges from a set of biologically active molecules (Kier and Roche 1967). For drug design, the features of a pharmacophore represent the location and type of functional group that a candidate molecule

should possess for it to be bioactive against a particular target (Leach et al. 2010). Traditionally, pharmacophore models were comprised of the six main chemical features: positive charged group, negative charged group, hydrogen bond acceptor, hydrogen bond donor, hydrophobic group, and aromatic ring. In modern practice, a pharmacophore model can be established in either a ligand-based manner, by superimposing bioactive molecules and extracting the location and type of shared chemical features essential for bioactivity, or in a structure-based manner, by probing likely interaction points between the ligand(s) and macromolecular target. Both approaches have been applied successfully and extensively in virtual screening (Akram 2015; Braga et al. 2014; Guner and Bowen 2014; Villoutreix et al. 2009).

In the absence of a macromolecular target structure, ligand-based pharmacophore modeling has become an indispensable tool for the early stages of drug discovery. Typically this involves extracting common chemical features from the 3D structures of several known bioactive ligands, collectively generating a map of the essential interactions between the ligand and the macromolecular target In general, pharmacophore generation using multiple ligands involves two main steps: the creation of conformers based on the conformational flexibility and the subsequent alignment of select conformers, and the determination of shared chemical features that represent essential interaction points with the receptor. The handling of ligand conformational flexibility and conducting conformer alignments represent the key techniques in ligand-based pharmacophore modeling, but also the major difficulties associated with the technique (Akram 2015; Guner and Bowen 2014; Stahura and Bajorath 2005; Xue et al. 2005).

Structure-based pharmacophore modeling, on the other hand, directly incorporates the 3D structure of a macromolecular target. The protocol involves an analysis of the chemical features of the active site and their spatial relationships, from which a subsequent pharmacophore model

can be assembled with complementary features (Pirhadi et al. 2013). The structure-based pharmacophore modeling approach can be further classified into two subcategories: receptor-ligand-complex-based and receptor-(without ligand)-based. The receptor-ligand-complex-based approach is convenient in pinpointing the ligand-binding site on the receptor and visually determining the key interaction points between ligands and macromolecule. The obvious limitation of this approach is the need for crystal structure of the receptor–ligand complex. This can be overcome by the receptor-only approach (Akram 2015). However, without a bound ligand, identifying key interaction points within the active site is extremely difficult, often leading to a pharmacophore model containing unnecessary features. Hence, when possible, a receptor-ligand based approach should be implemented (G. Bohm and Jaenicke 1992; Horvath and Jeandenans 2003; Horvath 2011).

2.4.3 QSAR Based Methods

The last and most frequently utilized ligand-based method is the quantitative structure-activity relationship, QSAR. QSAR models are a means used to correlate molecular activities with physiochemical descriptors from a set of related compounds. QSAR modeling techniques can be very useful in predicting bioactivity in experimentally untested compounds (Braga et al. 2014). Furthermore, QSARS have proven helpful in lead identification during the early stages of virtual screening as well as a means to prioritize synthetic efforts (Jorissen and Gilson 2005). Despite their usefulness, QSARS require a significant amount of accurate kinetics data to be employed.

CHAPTER 3: INVESTIGATING COMBINATORIAL APPROACHES IN VIRTUAL SCREENING ON PFKFB3: A CASE STUDY FOR SMALL MOLECULE KINASES

3.1 Abstract

Fruitful efforts toward improving the predictiveness in tier-based approaches to virtual screening (VS) have mainly focused on protein kinases. Despite their significance as drug targets, small molecule kinases have been rarely tested with these approaches. In this section, we investigate the efficacy of a pharmacophore screening-combined structure-based docking approach on the human inducible 6-Phosphofructo-2-kinase/Fructose-2,6-bisphosphatase, an emerging target for cancer chemotherapy.

Six out of a total 1,364 compounds from NCI's Diversity Set II were selected as true actives via throughput screening. Using a database constructed from these compounds, five programs were tested for structure-based docking (SBD) performance, of which, MOE showed the highest enrichments and second highest screening rates. Separately, using the same database, pharmacophore screening was performed, reducing 1,364 compounds to 287 with no loss in true actives, yielding an enrichment of 4.75. When SBD was retested with the pharmacophore filtered database, 4 of the 5 SBD programs showed significant improvements to enrichment rates at only 2.5% of the database, with a 7-fold decrease in an average VS time. Our results altogether suggest that combinatorial approaches of VS technologies are easily applicable to small molecule kinases and, moreover, that such methods can decrease the variability associated with single-method SBD approaches.

3.2 Introduction

Compared to the traditional, costly methods of high throughput screening (HTS), the search for the potential drugs with computational technologies, also called virtual screening (VS), continues to grow in popularity (Sousa et al. 2010; Tuccinardi 2009; Villoutreix et al. 2009). Unlike HTS, which requires constant human intervention, even with the usage of robots, VS needs only a good computational resource and, more attractively, can be performed in a 'fire-and-forget' manner with only minimal human input once properly started. Although each has its own distinct advantages, these methods share a common goal in that they both aim to identify a small number of true biological positives amidst a vast amount of biological negatives (Klebe 2006; Parker and Bajorath 2006).

A computational drug search, or VS, is carried out to find biologically positive compounds from compound databases, which typically contain molecular information on compounds that range in number from thousands to millions. Depending on the screening factors, VS involves diverse strategies: molecular features of the known ligand characteristics, pharmacophore screening (PhS); ligand alignments based on both structural and physical characteristics of ligands, ligand similarity analysis (LSA); and the interactional relationship between ligands and their target receptors, structure-based docking (SBD) (Guido et al. 2008). Depending on the availability and capacity of computational resources, an overall strategy of VS can be varied to employ either one, or combinations, of these technologies to appropriately meet the user's needs.

Assuming that data handling is performed in an organized manner, PhS can be carried out in a very short amount of time using even a personal desk computer. SBD, on the other hand, requires an extensive amount of calculations of Gibbs free energy changes involved in

the various ligand receptor interactions (Guido et al. 2008). Consequently, a high capacity computational resource, such as a processor cluster, is usually necessary for moderate-to-large SBD projects. However, even with such a facility, the amount of required computation time is often still very considerable, a circumstance owing to the fact that the vast majority of SBD time is consumed performing extensive calculations on outright negatives simply because of low hit rates associated with non-enriched databases. To improve the cost efficiency of SBD, efforts have been made to combine SBD with other less-time consuming technologies (Maiorov and Sheridan 2005). As a result, tiered screening strategies, in which ATP was considered as single ligand, have been successful in protein kinases (PK), one of the most popular drug targets up to date (Guido et al. 2008; K. Lee et al. 2010; Nagarajan et al. 2010; Workman and van Montfort 2010).

Despite their significance as drug targets, studies for strategic VS of small molecule kinases (SMK) are extremely rare compared to those of PK (Marsden and Knapp 2008; Workman and van Montfort 2010). To explore how well the concepts of tiered screening translate to SMK, we sought to design a dual-step screening protocol that could accurately and efficiently identify potent inhibitors through a combination of pharmacophore screening (PhS) and structure-based docking (SBD). Both PK and SMK have a second substrate pocket in addition to the widely conserved ATP pocket. Compared to PK, SMKs can have one, or possibly two, ligands that can bind the non-ATP binding site, most of which are known from biochemical/structural characterizations. We speculated that those known ligands may serve as a good resource of pharmacophore and may serve to add more target specificity. Thus, our study was targeted at the second substrate binding site, the pocket for the acceptor of the phosphoryl group from ATP.

For this project, the human inducible 6-phosphofructo-2-kinase/fructose-2,6-bisphosphatase (PFKFB3), a newly emerging drug target for potential cancer chemotherapeutics, was taken as the receptor (Chesney 2006). The kinase domains of the four PFKFB isoforms catalyze the synthesis of fructose-2,6-bisphosphate (F-2,6-P_2), which is the most potent allosteric stimulator of glycolysis, using adenosine triphosphate (ATP) as the phosphoryl donor and fructose-6-phosphate (F-6-P) as the acceptor (Pilkis et al. 1995; Rider et al. 2004). Predominantly expressed in neoplastic cells by the action of HIF-1 among the four isoforms, PFKFB3, with its kinase activity at least 10x that of the second most active isoform, rapidly increases the level of Fructose-2,6-bisphosphate (F-2,6-P_2) (Kim et al. 2006; Yalcin et al. 2009). As a result, glycolysis in neoplastic cells such as cancer is very active, which has been long known as the Warburg effect (Hsu and Sabatini 2008; Kroemer and Pouyssegur 2008). Recent studies have suggested that PFKFB3, whose expression is induced during cancer growth, is one of the most causative molecules of the Warburg effect (Atsumi et al. 2002; Yalcin et al. 2009). To explore the possibility of PFKFB3 as a new cancer therapeutic target, we determined the crystal structure of PFKFB3 to 2.1Å and elucidated the catalytic mechanism for F-2,6-P_2 synthesis at the molecular level (Kim et al. 2006). The resulting molecular structure/function information has been used as the foundation of this study.

In this report, we investigate the efficacy of combinatorial screening of PFKFB3 targeting for the F-6-P site. To evaluate the efficacy of our tiered approach, we developed our own database of active/non-active molecules from the National Cancer Institute's (NCI) Diversity Set II through a biochemical throughput study, since no non-ATP datasets of competitive inhibitors for the second substrate. Using the actives as reference compounds,

we explored applications from the two most common approaches in computational drug discovery: pharmacophore screening and structure based docking. Herein, we present the results from our combinatorial approach, demonstrating the enrichment of a 1,364 compound database, containing six true actives 'T-actives', to a resultant database of 287 compounds, still containing six 'T-actives', while using only one seventh of the computational resources required for standard docking procedures.

3.3 Materials and Methods

Biochemical throughput screening.

The recombinant human PFKFB3 was expressed and purified as described elsewhere. The 1,364 individual compounds of NCI's Diversity Set II were acquired from the NCI. The 2-kinase activity assay for throughput screening was performed using a F-2,6-P_2 assay modified for 96-well plates from the conventional method (van Schaftingen et al. 1982). This assay consisted of two sequential steps: F-2,6-P_2 production by PFKFB3 and allosteric activation of PFK-1 by produced F-2,6-P_2. The first reaction, F-2,6-P_2 synthesis by PFKFB3, was started by adding 130 nM PFKFB3 to mixtures containing 20 mM pH 8.0 TES, 1 mM DTT, 2 mM $MgCl_2$, 50 µM F-6-P, 50 µM ATP, 0.5% Tween, and 10 µM of each inhibitor. This reaction was allowed to run for 10 minutes at 25°C and then stopped by the addition of 0.1 M KOH. Aliquots of 1-4 µL of the first reaction were transferred, after pH neutralization, to the reactions of the second step, which consisted of 50 mM pH 8.0 Tris-HCl, 0.2 mM NADH, 5 mM DTT, 1 mM F-6-P, 2 mM $MgCl_2$, 0.70 units/mL Aldolase, 0.45 units/mL GDH, 0.60 units/mL TIM, and 10 mU PPi:PFK. The reactions of seconds step were started by adding 0.5 mM sodium pyrophosphate (PPi) and were then measured for changes in absorbance at 340 nm over a period of 30 minutes.

For molecules showing strong inhibition in the throughput assay, a study of the steady-state inhibition kinetics was carried out using a method in which the concentrations of F-6-P, ATP, and/or inhibitors were varied according to experimental purposes.

Ligand Library Design.

The virtual ligands of NCI's Diversity Set II were obtained from NCI's Developmental therapeutics program (DTP). The ligands were acquired in SMILEs format and standardized using a template to ensure that each was minimized, pH adjusted, and devoid of salts and other non-ligand contaminants. Additional adjustments, such as explicit hydrogens and force fields were added in a program-dependent fashion depending on manufacturer recommendations. Using the standardized ligand database, an additional conformer database was generated for pharmacophore filtering. For this, MOE was used to generate a conformer database based on default settings with a maximum of 250 conformers per ligand and no post-generation refinement.

Structure-Based Virtual Docking.

Virtual docking was carried out using 5 different programs, DOCK, VINA, FlexX, MOE, and GOLD in an effort to determine the suitability of each for PFKFB3 screening (Ewing et al. 2001; Group 2010; G. Jones et al. 1997; Rarey et al. 1996; Trott and Olson 2010). For this, the protein structure '2AXN' was used as the receptor macromolecule for docking (Kim et al. 2006). The details of program-specific conditions are included in Supplemental Information.

Pharmacophore Screening.

MOE was used to generate pharmacophores from three molecules known to bind to the F-6-P pocket of the PFKFB3: Fructose-6-Phosphate (F-6-P), Ethylenediaminetetraacetic acid

(EDTA), and Phosphoenol Pyruvate (PEP). The conformations for these molecules were obtained from the structural data (*PDB ID:* 2AXN, 2DWO, 2I1V, 2DWP). By superimposing the conformers, property features were extracted and merged and tolerance values were adjusted in accordance with results through retro-fitting. After refinement, eight features were chosen to be included in the final pharmacophore map; however, only 5 features are required to be met at any one time for a compound to pass the filter. Additionally, inclusion and exclusion spheres were added and constraint allowances were adjusted for preference. All pharmacophore searches were carried out within MOE.

3.4 Results

Biochemical Throughput Screening.

To generate a framework of actives and non-actives for VS, a throughput study of 1,364 NCI compounds was carried out. The inhibition extent of 10 μM of each compound, in substrate saturation conditions, was quantified and the results of the top 50 compounds are shown in Fig. 3.1. An arbitrary cutoff was chosen at 75% inhibition to describe compounds that were to be considered 'potential' actives. Based on this cutoff, 10 compounds were identified from the original 1,364.

To select the true positives, the 10 potential actives were subsequently tested for specificity for the F-6-P site, because the VS was targeted for the F-6-P site. Using conventional steady state inhibition kinetics, 6 compounds were selected as the 'true actives (T-actives)' and listed in Fig. 3.2. All T-actives exhibit competitive inhibition against F-6-P and uncompetitive against ATP, as a representative example, NSC278631, is shown in Fig. 3.3. The Ki's for each compound was determined to be at or below 20 μM.

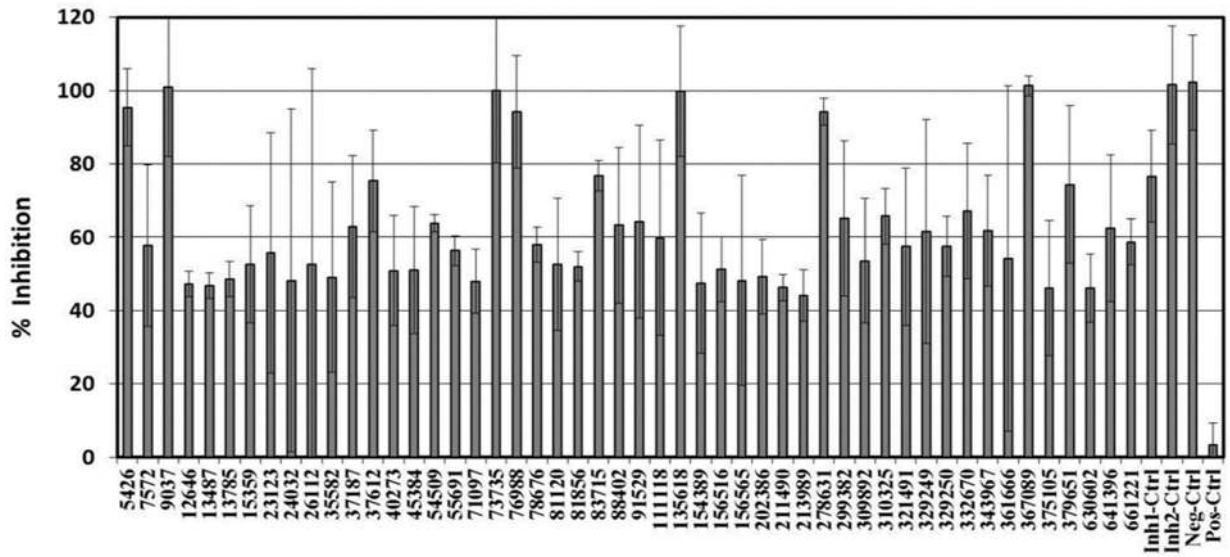


Figure 3.1. Identification of potent PFKFB3 inhibitors via a single-dose (10 μM) primary screening assay. The top 50 PFKFB3 inhibitors of NCI's Diversity set II are shown in relation to four experimental controls. *Inh1* and *inh2* are in-house inhibitors that have been tested and shown to target the PFKFB3 kinase domain. The controls, *Pos* and *Neg*, depict the uninhibited presence and complete absence of PFKFB3, respectively, and thus were used to represent the theoretical maximum and minimum inhibition values by which all screening compounds were compared.

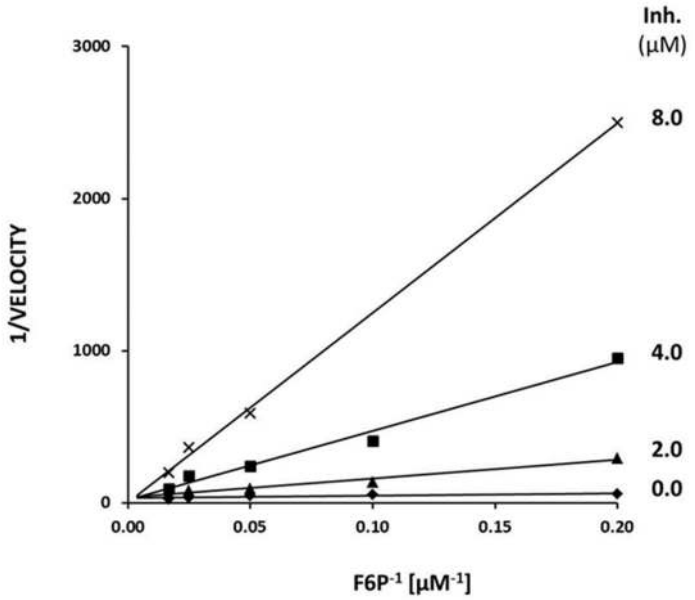


Figure 3.2. Inhibition for the PFKFB3 2-kinase by NSC278631. A double-reciprocal plot shows the competitive inhibition for NSC278631 against F6P. The lines represent varying inhibitor concentrations and were generated by data fitting using the program GraphPad Prism.

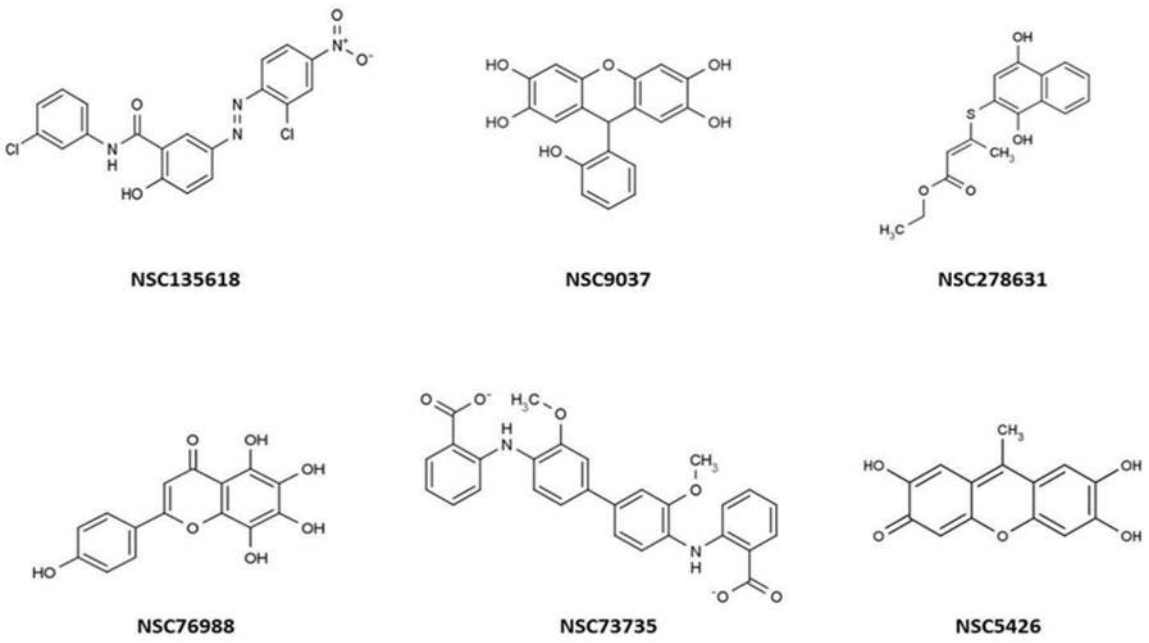


Figure 3.3. The selected actives from the throughput screening of the NCI Diversity Set II.

Pharmacophore Screening.

Using ligands already known to bind to the F-6-P site from crystallographic evidence, namely, F-6-P, F-2,6-P_2, EDTA, and PEP, a pharmacophore model was built and used to screen the NCI diversity set via MOE's pharmacophore screening module (Fig. 3.4) (Kim et al. 2006; Kim et al. 2007). Overall, from this filtering process, the database size was reduced from 1364 to 287 ligands while retaining 6 out of 6 'T-actives'. The results of this procedure demonstrate a significant reduction in non-actives and no reduction in actives. The total screening time was 206 seconds on a 2GHz processor with a conformer database creation time of 9911 seconds.

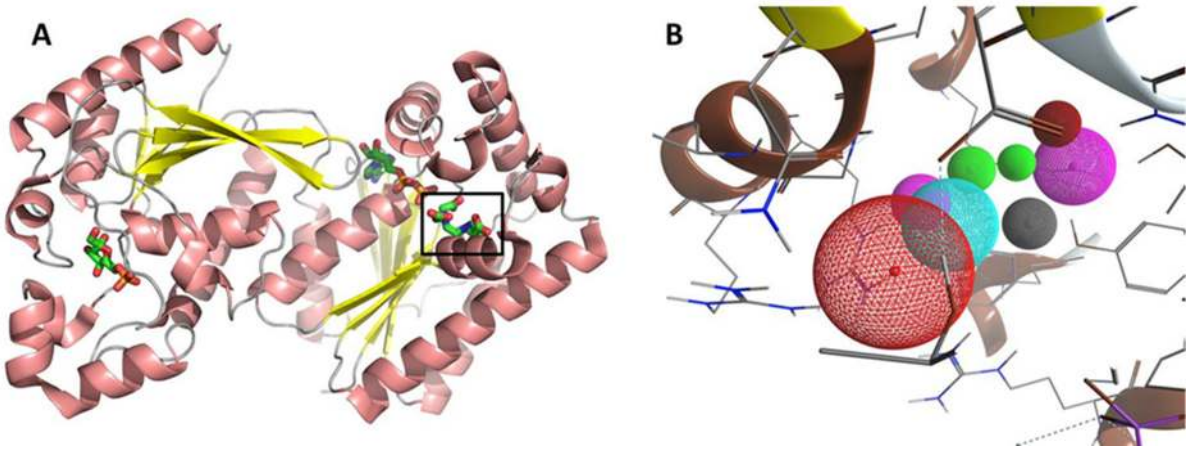


Figure 3.4. Pharmacophore map used in PFKFB3 virtual screening. (a) A ribbon diagram of PFKFB3 in complex with three ligands, ADP, EDTA, and F-6-P (*PDB ID:* 2AXN). The boxed ligand, EDTA, is occupying the F-6-P site of the kinase, which is the target site for our screening protocol. (b) A magnified view of the F-6-P site including a pharmacophore feature map. The feature map consists of 8 spheres of varying sizes and chemical properties, with at least 5 being needed to be met for a ligand to pass the filter. *Pharmacophore features: Red (AccP|AccS), Light-Blue (HydS|AccP), Magenta (ML&(AccS|AccP)), Green (HydP|HydS), Dark Red (Ani&(AccS|AccP)), Gray (ML)*.

Performance comparisons of docking programs.

Because it has been demonstrated in numerous studies that the efficacy of a SBD program directly ties to the target protein, we chose to test the individual performances of several SBD programs. Using PFKFB3, a full database evaluation was conducted to compare the enrichment factors of five popular SBD technologies (Fig. 3.5). The results revealed that each of the tested SBD technologies significantly enriched the NCI diversity set II database. However, as seen in other studies, the enrichment rates varied significantly according to the SBD technology (Kruger and Evers 2010; Li et al. 2010; Plewczynski et al. 2011). For comparison purposes, we investigated the enrichment at two database sizes, 2.5% and 10%. MOE performed best, showing higher enrichments at all database sizes. The other SBD programs were more

varied in their performances with VINA having the second highest enrichment rates at 2.5% and GOLD at 10%.

Combinatorial Screening Efficacies.

To measure the efficacy of the combinatorial screening protocol, the pharmacophore filtering results were subsequently docked using each of the SBD technologies. For this, the PhS enriched database, consisting of the 287 hit molecules with all actives present, was docked and the enrichment rates were evaluated at 2.5% and 10% database sizes (Fig. 3.6). The results demonstrate improved enrichment rates for four of the five SBG technologies at 2.5% database size and five of five at 10% database size compared with docking-only methods. Additionally, it was determined that the application of the combined protocol, greatly reduced the variability of the incorporated SBD technologies, changing the enrichment differences between the highest and lowest scoring technologies from 13 to 6.5 and 4.9 to 3.2 at 2.5% and 10% database sizes, respectively. Reductions in the overall time were also witnessed, showing nearly a 7-fold decrease in the average total time for a complete database screening using the tiered approach.

3.5 Discussion

In this work, the enrichment capability of a commonly employed, tier-based virtual screening approach was evaluated using the small molecule kinase 6-Phosphofructo-2-kinase/Fructose-2,6-bisphosphatase (PFKFB3). For this, biochemical throughput data was generated using the National Cancer Institute's Diversity Set II compound library, serving as a metric upon which enrichment rates were calculated (Fig. 3.1). The aim was to evaluate the

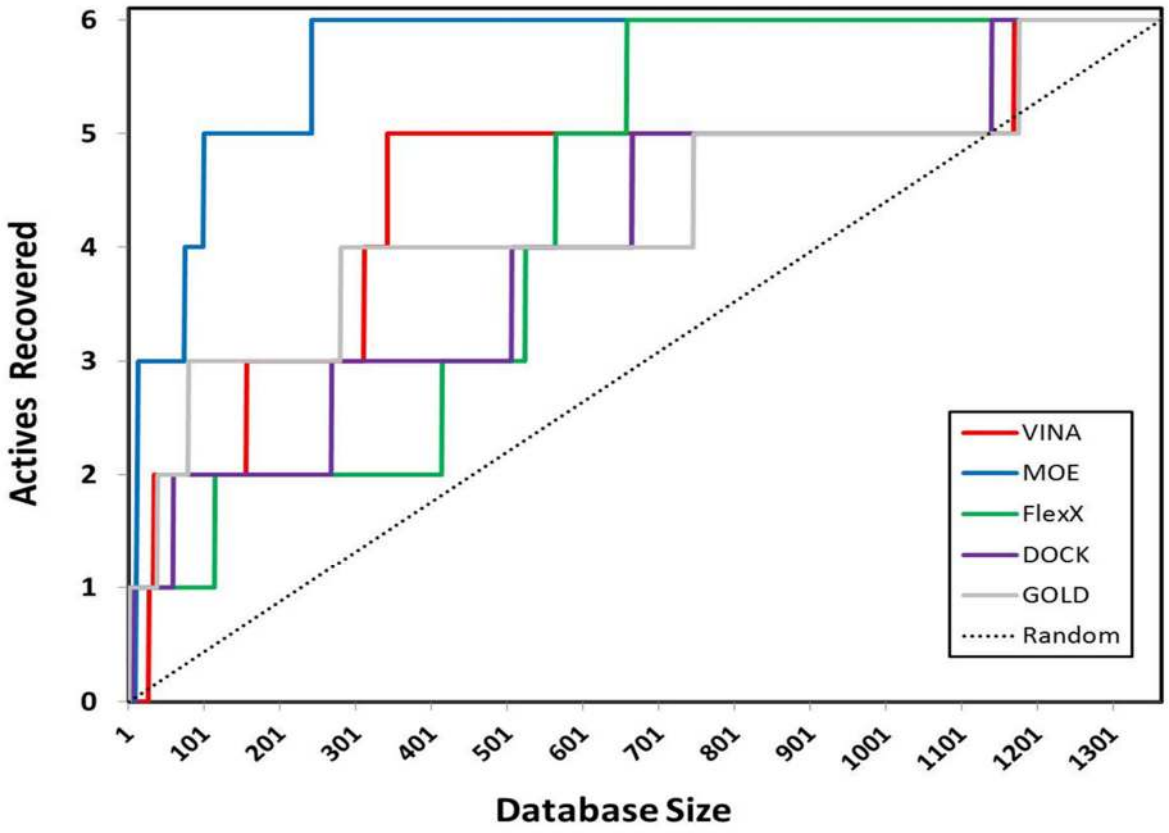


Figure 3.5. Enrichment Comparison of Popular SBD Technologies on PFKFB3. Full Database enrichment rates were calculated for comparison between VINA (red), MOE (blue), FlexX (green), DOCK (purple), GOLD (gray), and random (dotted).

efficacy of the combined pharmacophore-docking protocol in relation to docking-only methods using a representative protein from the small molecule kinase family. Additionally, an investigation into the predictiveness of the major docking technologies was conducted; seeking to identify which offered best enrichments for PFKFB3.

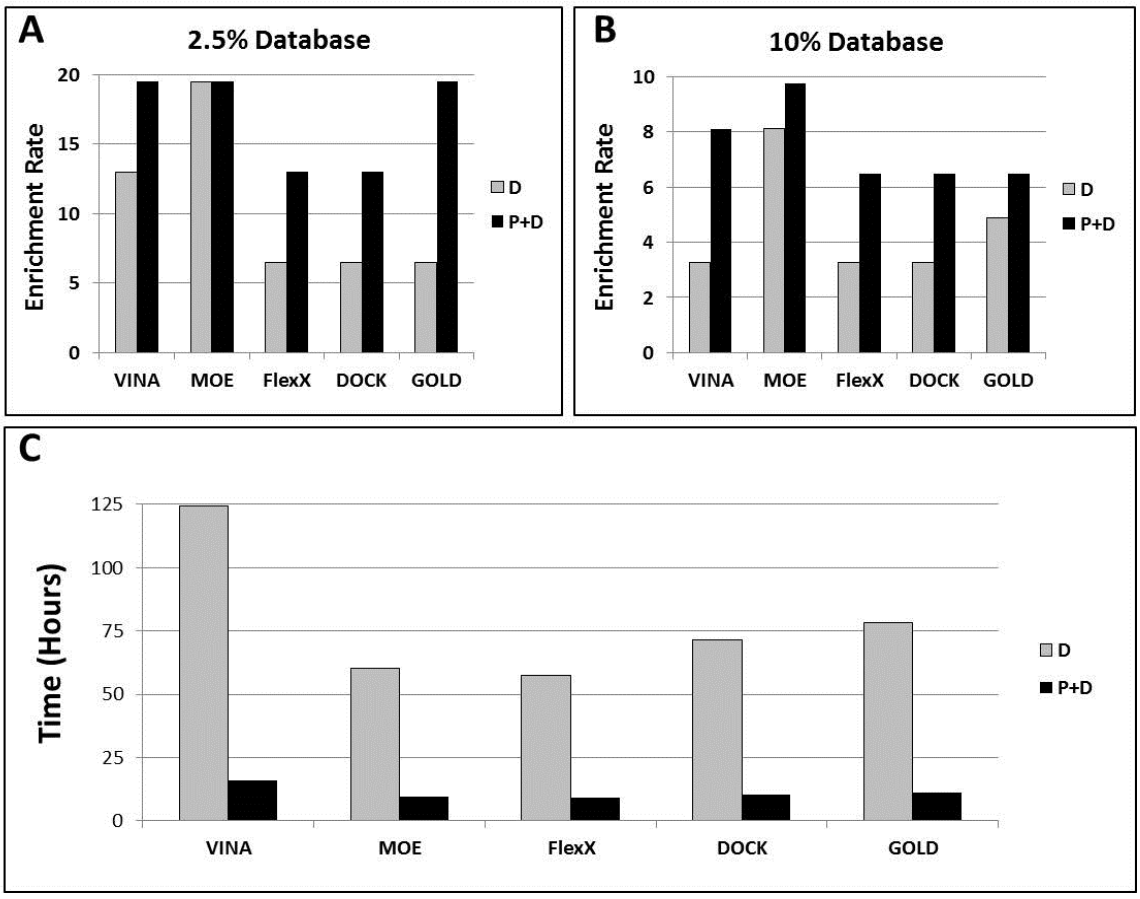


Figure 3.6. Performance comparison between sequential and non-sequential virtual screening protocols. (a and b). Calculated enrichment rates at differing database sizes for Docking (D) and Pharmacophore+Docking (P+D) screening protocols. Shown on the x-axis are the five docking programs, each of which was evaluated individually, and, in conjunction with, a pharmacophore filter as previously described. (c) Measured screening times for Docking (D) and Pharmacophore+Docking (P+D) screening protocols. Shown on the x-axis are the five docking programs, each of which was evaluated individually, and, in conjunction with, a pharmacophore filter as previously described.

The results of our investigation revealed that tiered screening, in the case of PFKFB3, in which, pharmacophore filtering precedes structure based docking, is clearly favorable to any individual screening method, both in efficacy and efficiency (Fig. 3.6). An analysis of our results can be evaluated in terms of three factors: speed, accuracy, and consistency.

The accuracy of in silico predictions is always a critical factor of virtual screening. Hence, our investigation into tiered approaches of virtual screening initially focused on differences in enrichment rates between single SBD and combined PhS-SBD approaches. The results from our testing revealed that, in the case of PFKFB3, the enrichment factors of the combinatorial strategy was always comparable to, or in excess of, the values demonstrated from the single-run structure based simulations. Such results were not entirely unexpected when considering the results of previous studies demonstrating similar findings for single-ligand proteins and protein kinases (Swann et al. 2011; Tan et al. 2008). However, since no similar, large-scale, studies of small molecule kinases are yet published, it was not known if such trends would extend to small molecule kinases. The findings presented here should serve to lessen our knowledge gap involving the enrichment efficacy of tiered approaches to virtual screening for small molecule kinases.

Another finding reported here involves the changes to SBD consistency when incorporated into tiered-based approaches. From our single-method docking results, it was seen that vast differences exist between the enrichment rates of differing technologies. However, upon analysis of the PhS-SBD tiered approach, it was clear that most of the differences in enrichment among the SBD technologies efficacy had significantly lessened. Such findings are promising, in that they help to lessen the need to test multiple SBD technologies for any given project. Because it is well known that different SBD technologies offer differing performances depending on the target macromolecule, such findings suggest that the need to test many SBD technologies can be significantly lessened by using tiered approaches.

Lastly, and most apparent among our findings, was the obvious advantage of the tiered-based approach in reducing the computational time necessary to screen the Diversity Set II database. Despite having higher enrichment rates, the tiered screening was, on average, 7-fold faster than its single-run counterparts. The principle behind such an increase hinges on the role of pre-filtering. By using a loose, lightly-selective filter early on, many of the more egregious ligands can be removed prior to the computationally intensive step of docking, thus significantly reducing the time necessary for any virtual screening project.

The results altogether indicates that a combinatorial approach for VS is far more efficient than a blind run of time- and resource-consuming simulated docking. However, it has to be admitted that there are some shortfalls in combinatorial approaches. First, it requires at least one ligand that is already known to bind to a target receptor and, practically, the more the better. In reality, this problem is almost ignorable for most of the active drug target proteins. It is because only a few proteins whose functions are not known are selected as drug targets. For most cases, when a protein was selected as a drug target, its function is clearly defined and, accordingly, at least one or two ligands are already known.

Second, as apparent from the results of our studies, there is no missing 'T-actives' from the pharmacophore screening. And while this is not necessarily a problem, and was desired as part of our ultimate aim, such a favorable 'T-active' recovery from pharmacophore screening is dependent upon the experience of the modeler. For this project, the creator set out with the intention of designing a loosely selective map rather than one maximizing enrichment at the expensive of recovered 'T-actives'. Such a choice was made upon the basis of the size of the

database to be screened, however, we do not anticipate that alternative selections based on enrichment preferences, would dramatically influence the benefits of combinatorial screening.

In this paper, we demonstrate that the combined use of ligand- and structure-based screening for small molecule kinases can decrease screening times and increase enrichment rates as compared with single step screening approaches.

CHAPTER 4: CRYSTAL STRUCTURE OF 6-PHOSPHOFRUCTO-2-KINASE/FRUCTOSE-2,6-BISPHOSPHATASE (PFFKB2) AND THE INHIBITORY INFLUENCE OF CITRATE ON SUBSTRATE BINDING.

4.1 Abstract

The heart-specific isoform of 6-phosphofructo-2-kinase/fructose-2,6-bisphosphatase (PFKFB2) is an important regulator of glycolytic flux in cardiac cells. Here, we present the crystal structures of two PFKFB2 orthologues, human and bovine, at resolutions of 2.0 and 1.8Å, respectively. Citrate, a TCA cycle intermediate and well-known inhibitor of PFKFB2, was present in the 2-kinase domain of both orthologues, occupying the fructose-6-phosphate binding-site while also extending into the γ-phosphate binding pocket of ATP. This steric and electrostatic occlusion of the γ-phosphate site proved highly consequential to the binding of co-complexed ATP analogues. The bovine structure, which co-crystallized with ADP, closely resembled the overall structure of PFKFB isoforms, with ADP mimicking the catalytic binding mode of ATP. The human structure, on the other hand, co-complexed with AMPPNP, which, unlike ADP, contains a γ-phosphate. The presence of this γ-phosphate made adoption of the catalytic ATP binding mode impossible for AMPPNP, forcing it to bind atypically with concomitant conformational changes to the ATP binding-pocket. Inhibition kinetics were used to validate the structural observations, confirming citrate's inhibition mechanism as competitive for F6P and mixed for ATP. Together, these structural and kinetic data establish a molecular basis for citrate's negative feed-back loop of the glycolytic pathway via PFKFB2.

4.2 Introduction

The bifunctional enzyme, 6-phosphofructo-2kinase/fructose-2,6-bisphosphatase (PFKFB), catalyzes both the synthesis and hydrolysis of fructose2,6-bisphosphate (Fru-2,6-P_2) as a functional homodimer (C. Wu et al. 2006). The dual enzymatic activities are provided by the two separate catalytic domains existing in a single protein subunit, to control the concentration of Fru-2,6-P_2, the most potent allosteric activator of 6-phosphofructo-1-kinase (PFK), the rate-limiting enzyme of glycolysis (El-Maghrabi and Pilkis 1984; Pilkis and Granner 1992; Pilkis et al. 1995). An elevated cellular concentration of Fru-2,6-P_2 increases glycolytic flux, whereas a lowered concentration of Fru-2,6-P2 decreases glycolysis. Reflecting the diversity of tissues and their physiological functions, different tissue-specific isoforms of PFKFB from four distinct genes (*pfkfb*1-4) are expressed, each with different kinetic properties, and a single isoform generally predominates in each tissue: PFKFB1, the liver form; PFKFB2, the heart form; PFKFB3, the inducible form; and PFKFB4, the testis form (Okar et al. 2001).

Because of the uniqueness in its structure/function relationships, the PFKFB enzyme system has long been targets of structure/function studies. As a result, a significant amount of functional data has been produced and the crystal structures of the rat liver Fru-2,6-P2ase domain, the rat testis form (PFKFB4), the human liver form (PFKFB1), and human cancer form (PFKFB3) have been determined (Hasemann et al. 1996; Kim et al. 2006; Y. H. Lee et al. 2003). These data altogether allowed us to understand the molecular mechanism of catalytic reactions and regulation of this enzyme system. Some are now serving as a molecular foundation for development of therapeutics for diabetes and cancer (Garber 2004; A. Minchenko et al. 2002).

However, one of the most fundamental questions yet to be answered about this enzyme system is how each PFKFB isoform performs its differential function in hosting tissues that have specific physiological roles and, accordingly, different optimum conditions for glucose metabolism. Reflecting such differences, the PFKFB isozymes have been shown to have different kinetic properties, as summarized in Figure 4.5. These kinetic differences suggest that glucose metabolism is uniquely related to the physiological roles of the given tissues and not just a simple housekeeping function for energy production (Y. H. Lee et al. 2003; Pilkis et al. 1995).

As an effort to address the tissue-type specifically differentiated structure/function relationships of PFKFB isoforms, we determined the crystal structures of heart isoforms of PFFKB, PFKFB2, from both *H. sapiens* and *B. taurus*. Analysis of their structures and kinetics in comparison with those already known from the previous PFKFB studies suggested a regulatory mechanism as yet unknown and we introduce the results here.

4.3 Materials and Methods

Crystallization of PFKFB2 Homologs

The human and bovine PFKFB2 isoforms with 6xHis tagged at the N-termini were overexpressed in *Escherichia coli* BL21 C41(DE3) and purified using Ni-NTA affinity columns in conjunction with SP Sepharose cation exchange columns. Following purification by SP Sepharose, the proteins were dialyzed with pH 8.0 20 mM Tris•HCl, 10 mM NaP_i, 5 mM β-Mercaptoethanol , and 5% glycerol and then concentrated to 8.0 mg/ml with Millipore centrifugal filter concentrators. Crystallization of the bovine PFKFB2 (bPFKFB2) were prepared via sitting drop vapor diffusion using a 1:1 mixture of protein sample to a mother liquor, pH 7.5 100mM HEPES, pH 7.0 0.5% tacsimate, 13-16% polyethylene glycol 3350, and 3% dioxane.

The crystals were grown to a dimension of 0.2 x 0.2 x 0.05 mm within 2-3 weeks of incubation at 20°C. Crystals of the human PFKFB2 (hPFKFB2) were also prepared with the sitting drop vapor diffusion method using a mother liquor, pH 6.0 100mM MES, 0.5-3.0% polyethylene glycol 8000, 13-16% polyethylene glycol 3350, and 3% dioxane. Data quality crystals of hPFKFB2 were grown after 3-5 weeks of incubation at 12°C.

Data Collection and Processing

The crystals were soaked for cryo-protection and liganding in cryoprotectant solutions, containing the aimed ligands, for 0.5 to 2 hours prior to flash freezing at 77K using liquid N_2. All cryoprotectant solutions were prepared by enriching the reservoir solution of each crystal with 35% ethylene glycol. The diffraction data was collected at both beamline 6C of the Pohang Accelerator Laboratory, Pohang, Korea, using the CCD detector Quantum 210 (ADSC) with a source wavelength of 1.23986 Å or at the Gulf Coast Consortium Protein Crystallography Beamline (PX1) in the Center for Advanced Microstructures and Devices (CAMD), Louisiana State University, Baton Rouge, LA., using a Mar 165 mm CCD detector with a source wavelength of 1.3808 Å. All diffraction data was processed and scaled using HKL2000.

Crystals of hPFKFB2 belong to the primitive orthorhombic space group $P2_12_12_1$, having unit cell dimensions of a = 106.5 Å, b = 113.9 Å, c = 133.2 Å and an asymmetric unit consisting of two monomers. Crystals of bPFKFB2 belong to the C-centered orthorhombic space group, $C222_1$, with unit cell dimensions being a = 82.1 Å, b = 169.5 Å, c = 85.3 Å, α =90, β =90, γ =90 and an asymmetric unit consisting of one monomer. Statistics of the reflection data are summarized in Table 1.

Table 4.1 Statistics of reflection data and structure refinements. Liganding denotes the states in the 2-Kinase/2-Phosphatase. His-P is phosphorylated at His258. $R_{sym} = \Sigma_h(\Sigma_j|I_{h,j} - \langle I_h \rangle|/ \Sigma I_{h,j})$, where h=set of Miller indices, j=set of observations of reflection h, and $\langle I_h \rangle$=the mean intensity. RMSD values are deviations from ideal values. $R_{crys} = \Sigma_h \| F_{o,h} | - |F_{c,h} \| / \Sigma_h |F_{o,h}|$. R_{free} was calculated using 5% of the complete data set randomly excluded from refinement. The numbers in parentheses represent values from the highest resolution shell.

Homologue	Human	Bovine
Liganding	ATP•Citrate/His-P•F6P	ADP•Citrate/His-P•Citrate
Space Group	$P2_12_12_1$	$C222_1$
Unit Cell Dimensions (Å)	106.5 × 113.9 × 133.2	82.1 × 169.5 × 85.3
Resolution Range (Å)	38.55-2.01	40.0 - 1.82
Reflections [$F \geq \sigma(F)$]	106633	52166
Completeness (%)	0.98 (87.3)	97.3 (81.5)
Redundancy	4.2 (3.4)	10.9 (10.9)
$I/\sigma(I)$	16.2 (2.3)	12.5 (1.3)
R_{sym}	0.059 (0.420)	0.050 (0.284)
R_{crys}	0.1490	0.1386
R_{free}	0.1751	0.1618
No. of Amino Acids	845	424
No. of Protein Atoms	6925	3461
No. of Hetero Atoms	194	62
RMSD		
Bond Lengths (Å)	0.016	0.009
Angles (°)	1.40	1.100
Dihedral Angles (°)	16.747	19.090
Mean B factor	39.70	25.82
Protein Atoms (Å^2)	39.00	24.60
Hetero Atoms (Å^2)	42.97	22.84
Water Atoms (Å^2)	45.53	34.59

Structure determination and refinement

The structures of hPFKFB2 and bPFKFB2 were determined by molecular replacement using the Phaser software module implemented in the PHENIX program suite. Initial models for both orthologues were determined using the human liver PFKFB (1K6M) as a search model (Y. H. Lee et al. 2003). The final structures were achieved after iterated model rebuilding and refinement using PHENIX and Coot (P. D. Adams et al. 2010; Emsley and Cowtan 2004). The R_{crys}/R_{free} of the final models are 0.202/0.225 and 0.156/0.211 for hPFKFB2 and bPFKFB2, respectively.

The final models revealed the residues 31-450 out of 505 of hPFKFB2 and those of 28-450 of bPFKFB2. The missing residues are all from both the N- and C- terminal regulatory domains and are considered disordered, based on the results from mass spectroscopy of melt crystals (data not shown). For the human model, 93.2% of the 420 revealed residues lie within the 'most favorable region' of the main chain dihedral angle distribution, whereas none are found in the 'disallowed region'. Similarly, for the 423 residues of the bovine orthologue, 90.9% are within the "most favorable region" and none are in the "disallowed region".

4.4 Results and Discussion

Overall Structures of PFKFB2

As summarized in the previous section and Table 1, crystal structures of the two PFKFB2 orthologs were determined ─ human (hPFKFB2) to 2.1Å and bovine (bPFKFB2) to 1.8Å ─ to elucidate any structural/functional differences in heart-type PFKFB compared to other isoforms. Similar to the other PFKFB isoforms, the crystal structure of the heart isozyme, for both human and bovine orthologues, shows a PFKFB-typical head-to-head homodimer arrangement (Figure

1a) with a dimeric interface constituted primarily through crystal contacts made between two kinase domains.

Each PFKFB2 monomer consists of a single polypeptide chain that can be subdivided into four distinct regions: the 6-phosphofructo-2-kinase (2-Kase) domain, the fructose-2,6-bisphosphatase (2-Pase) domain, and two regulatory domains. The two catalytic domains are conserved among the different tissue isoforms with the sequence identity 76-83%, whereas the N-terminal and C-terminal domains are highly variable. Among 505 amino acid residues of PFKFB2, residues 1-37 and 451-505 constitute the N- and C- terminal regulatory domains, whereas residues 38-248 and 249-450 constitute the 2-Kase and 2-Pase domains, respectively.

The sequence identity shared between the human and bovine orthologs is larger than 95%, supporting the notion that the two PFKFB2 orthologs play a physiological role common to both human and swine. To analyze the structural differences between the two orthologs, Cα traces of hPFKFB2's 2-Kase and 2-Pase domains were superimposed, separately, onto corresponding domains of bPFKFB2 (Figure 1b). It was not appropriate to compare the two domains together, because differences in liganding states of the two orthologs cause differences in the domain-domain interface. Supporting the sequence identity, the two catalytic domains showed a high degree of overall similarity, with RMSD values of 1.46 Å and 0.68 Å, for the 2-Kase and 2-Pase domains, respectively. Although the discrepancy in the 2-Kase domains is higher than expected, no significant structural differences between the two orthologs, which would implicate coinciding differences in the structure/function relationship, are apparent.

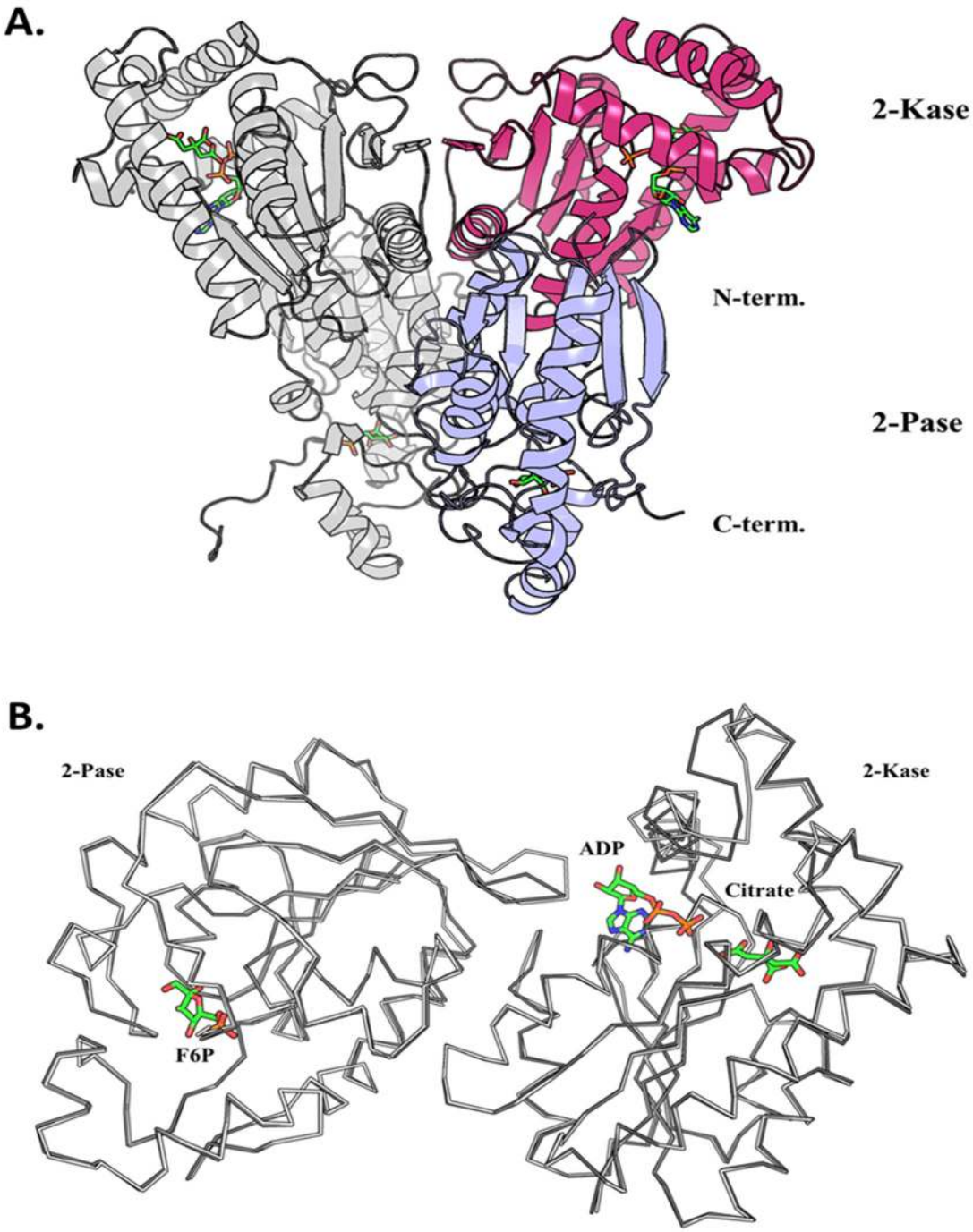


Figure 4.1. Dimeric arrangement of heart form and a comparison of the structures of the human and bovine orthologues. (A) Head-to-head homodimer arrangement of the human PFKFB2. Shown in gray is a complete monomer, while the red and blue colors depict the 2-Kase and 2-Pase domains, respectively. Ligands are colored by element and are shown in stick form. (B) Superimposed C-alpha traces of bovine (white) and human (light gray) kinase and phosphatase domains. The phosphatase domains contain Fru-6-P from the bovine orthologue in the F-2,6-P_2 binding site. The kinase domains with ADP and Citrate from the bovine orthologue in the ATP and F6P binding sites, respectively.

The unexpectedly high RMSD for the 2-Kase domains is likely due to differences in liganding states of the two orthologs despite similar liganding conditions during crystallization. For both the structures, Fru-6-P and citrate were in common, whereas AMPPNP ─ a non-hydrolyzable ATP analog ─ and ADP were exclusively included fill the ATP pockets of hPFKFB2 and bPFKFB2, respectively. Structural dissimilarities were concentrated to localized regions within-or-near the 2-Kase's active site and correlate with citrate's binding to the 2-Kase domain. Despite Fru-6-P being present in crystallization conditions for both orthologues, it didn't complex with the 2-kinase domain of either, regardless of concentration. Instead, citrate, from an earlier purification step, occupies the Fru-6-P binding pocket. Interestingly, the binding of citrate appears to significantly influence ATP but not ADP binding to the 2-Kase domain. The details will be discussed in the following sections.

Conserved catalytic binding of ATP and ADP to the 2-Kase domain

Both hPFKFB2 and bPFKFB2 were co-crystallized in the presence of ADP and AMPPNP, respectively, which were clearly visible within the ATP-binding pocket of the 2-Kase domain based on $|F_o|$ - $|F_c|$ omit maps. The binding of ADP in bPFKFB2 is analogous to the catalytic mode of ADP and ATP, which were well characterized and elucidated in previous studies of the PFKFB protein system (Kim et al. 2007; Y. H. Lee et al. 2003). The adenine moiety resides in a partly solvent-exposed hydrophobic pocket stabilized primarily through CH-π contributions made by Gly51, Tyr54, and Val221. A single hydrogen bond is formed between the adenine N6 and the Oδ of Asn168. The nucleotide diphosphates, on the other hand, occupy a highly anionic phosphate binding loop created by a Walker A motif ─ conserved as -^{49}GLPARGKT^{56}- in all PFKFB isoforms (Walker et al. 1982). Additional hydrogen bonds to the phosphate moieties are contributed from the side chains of Lys172 and Tyr428. This binding is

well conserved among all PFKFB isoforms. However, Tyr428 plays a different role upon binding of ATP in the presence of citrate bound to 2-Kase.

Citrate binding to the 2-Kase domain

There was no Fru-6-P molecule bound to the 2-Kinase domain despite being present in the crystallization mixture, because the 2-Kase, which follows an ordered reaction, requires ATP binding prior to Fru-6-P binding. Instead, citrate was found to be bound to the Fru-6-P pocket of the 2-Kase domains of both hPFKFB2 and bPFKFB2 with strong hydrogen bonds to Thr130 and Tyr197 and salt bridges with Arg78 and Arg102. Except for Arg102, which is conserved though, all of these residues are known to be involved in the catalytic binding of Fru-6-P to the 2-Kase domain. This binding mode of citrate is conserved between the bovine and human orthologs. It is known that the 2-Kase of PFKFB2, the heart isoform, is very sensitive to inhibition from products of TCA cycle such as succinate and citrate. A previous structural study has suggested that succinate behaves like a competitive inhibitor against Fru-6-P (Galluzzi et al. 2013).

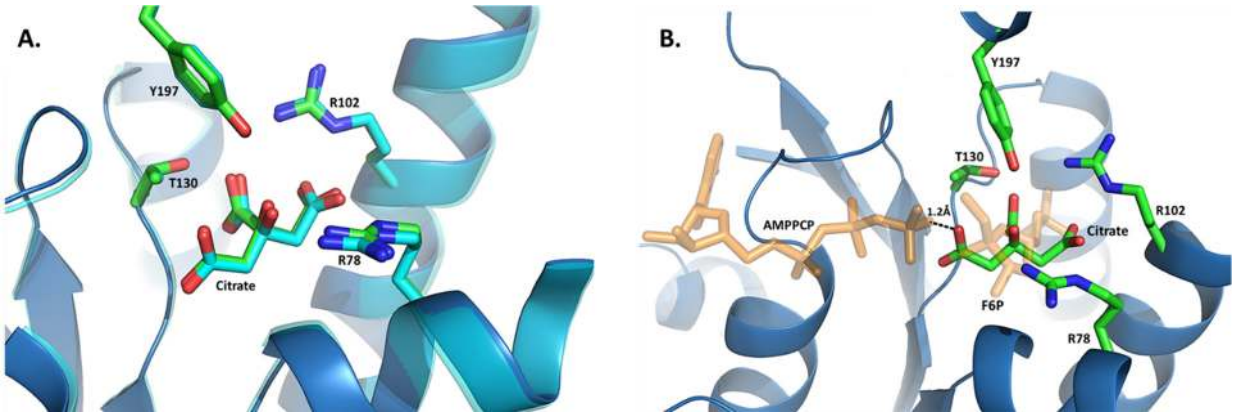


Figure 4.2. Citrate binding in the 2-Kase domain. (A) Superimposed structures of the citrate binding pocket for human (blue) and bovine (light gray) orthologues. Residues forming interactions with citrate are represented as sticks, with a ribbon diagram representing the mainchain. Human PFKFB2 is depicted as a blue ribbon with sticks containing green carbon atoms. In contrast, bovine PFKFB2 is represented as a light gray ribbon with sticks containing white carbon atoms. (B) AMPPCP and F6P from PFKFB3 (2DWP) overlaid onto the citrate binding site of human PFKFB2. The distance between the carboxy arm from citrate and nearest atom from the overlaid AMPPCP is shown. Both AMPPCP and F6P are semi-transparent colored orange.

However, citrate bound to PFKFB2 suggests an inhibition pattern, which is different from that of succinate, which was previously studied (Yuen et al. 1999). While primarily located in the Fru-6-P binding site, citrate, unlike succinate, impedes the γ-phosphate of AMPPNP from occupying its normal location for the Fru-2,6-P_2 synthesis. When citrate from hPFKFB2 was superimposed onto a pseudo-Michaelis complex of PFKFB3 (2DWP), one of the three carboxy group arms of citrate is only 1.2Å away from the γ-phosphate of AMPPCP (Kim et al. 2007). Such close proximity of two negatively charged acidic groups is both sterically and electrostatically unfavorable for ATP binding but ADP binding is not significantly influenced. Thus, it has been suggested that ATP binding to the 2-Kase domain in the known catalytic mode is not favored in the presence of citrate. And, considering perfect conservation of all the residues for citrate, binding of citrate to the four PFKFB isoforms would be very similar.

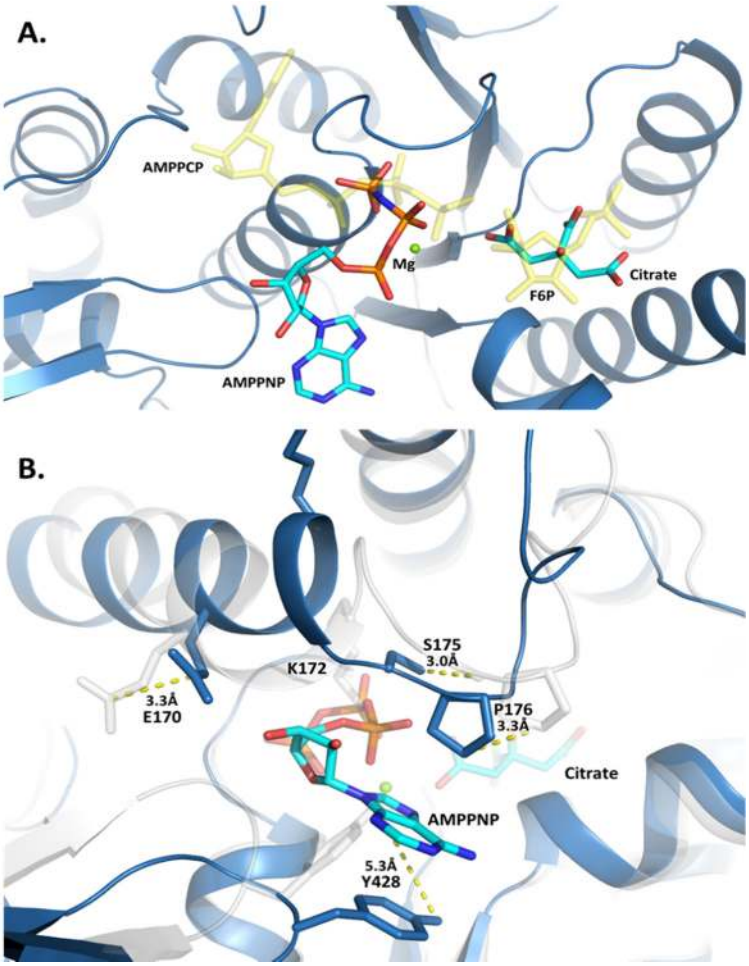


Figure 4.3. Ligand Binding within the 2-Kase Domain. (A) Human PFKFB2 complexed with AMPPNP and Citrate. Superimposed are AMPPCP and F6P from the pseudo-substrate complex with PFKFB3 (2DWP), showing the catalytic ATP and F6P binding modes, respectively. (B) Ribbon diagram view of human PFKFB2 (blue) super-imposed onto the bovine orthologue (light gray). The sidechains and their positional differences are shown for residues involved in AMPPNP binding.

An alternative ATP binding mode to the 2-Kase domain caused by citrate binding

Not surprisingly, the binding mode of AMPPNP in the presence of citrate is markedly different from the conventional mode of catalytic ATP binding, which is required for the Fru-2,6-P_2 synthesis by the 2-Kase domain and conserved among all PFKFB isoforms. When AMPPNP from the hPFKFB2•AMPPNP complex was superimposed onto ADP of bPFKFB2•ADP complex, as can be seen in Figure 3A, the ribose and adenine rings of AMPPNP are rotated 115° and 180°, respectively, relative to those of ADP. Furthermore, the centroid of AMPPNP's adenine ring is 14Å shifted outward from that of ADP. Consequently, few molecular interactions are shared among the bindings of adenosine moieties of AMPPNP and ADP to PFKFB2 and, thus, the nucleoside binding sites for AMPPNP are remarkably different from that of ADP (Figure 3B). From the AMPPNP-complexed structure, it can be seen that the α-carbon of Pro176 shifts ~3.3Å, placing the residue's R-group perpendicular to AMPPNP's planar ring, allowing for CH-π contributions. More significant contributions come by way of π-π stacking interactions with the phenol side chain of Tyr428, which is otherwise involved in interaction with the phosphate moieties. For the π-π stacking between the adenine ring and Tyr428, the β-hairpin containing Tyr428 shifted 4Å and the phenol sidechain rotated away from the position for its conventional interaction with the α-phosphate moiety of ADP (Figure 3C).

As a consequence, the γ-phosphate of AMPNP occupies the site for the β-phosphate of ADP or catalytic ATP, sharing many contact residues with the α- and β- phosphates of ADP (Figure 3B). Of the five residues forming the Walker A motif, Tyr54 is the only notable absence. As for sidechain interactions, the phosphate stabilizing hydrogen bond of Tyr428 is replaced by Ser174. Interactions with Lys172, on the other hand, are completely lost, with the residue flipping outward towards the solvent. This atypical binding of AMPPNP is inappropriate for the

2-Kase catalytic reaction and appeared to be caused by the citrate binding to 2-Kase as described in the following section.

Citrate inhibition at the active site of the 2-Kinase domain

Although previous studies have reported citrate to inhibit PFKFB enzymes, particularly the heart isotype, the inhibition mechanism has remained unknown (Ros and Schulze 2013). From the human structure of PFKFB2 presented here, it can now be seen that citrate competitively blocks F6P binding while simultaneously prohibiting ATP, or any γ-phosphate containing analogues, from binding in a catalytically active position. Inhibition studies were conducted to validate these structural observations. Non-linear regression analyses of the data supported the structural findings, showing citrate to act as a competitive inhibitor for F6P with a K_i of 80.3±6.54μM and non-competitive inhibitor for ATP with a K_{is} of 50.23±8.01μM and K_{ii} of 58.75±9.9μM, as shown in the double-reciprocal plot (Figure 4.4). The mixed inhibition agrees with the observation that citrate influences but does not completely block ATP from binding.

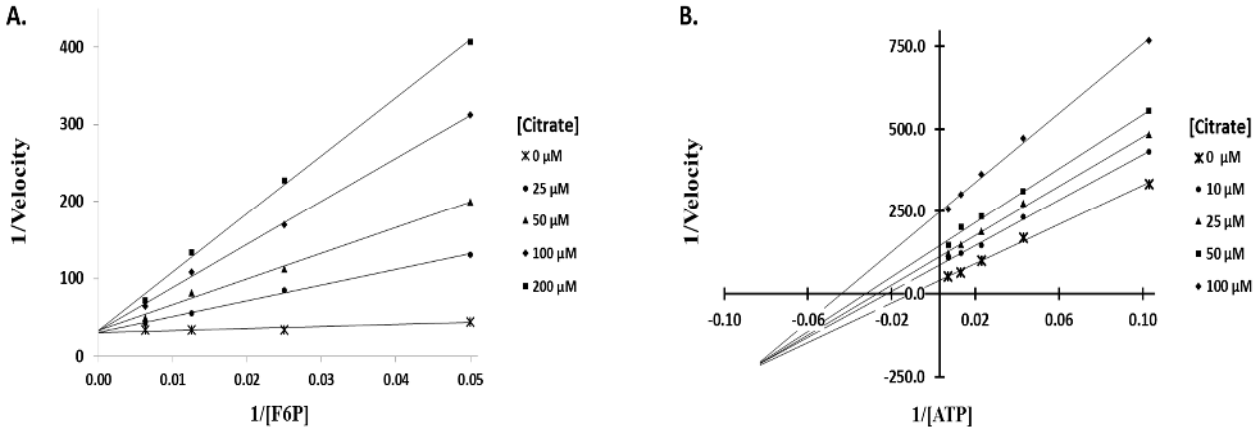


Figure 4.4. 2-Kase inhibition by Citrate. (A) Double reciprocal plot with F6P as the variable substrate. (B) Double reciprocal plot with ATP as the variable substrate.

Table 4.2. Citrate inhibition properties on wild-type and mutant PFKFB2.

Human PFKFB2	Variable Substrate	Fixed Substrate	Citrate Kis (μM)	Citrate Kii (μM)	Inhibition
Wild-Type	ATP	F6P	50.23±8.01	58.75±9.90	NC
	F6P	ATP	80.03±6.54		C
Y428A	ATP	F6P	2.43±1.02	3.05±1.39	NC
	F6P	ATP	67.30±4.23		C
Inhibition: UC, uncompetitive; C, competitive; NC, noncompetitive					

To validate biological significance of the citrate binding and the consequentially altered binding of ATP, a site-directed mutant, Y428A, the π-π stack provider for the altered ATP binding mode was mutated to alanine, was prepared and its kinetic properties analyzed in comparison to the wild-type, as shown in Table 4.2. The Y428A mutant PFKFB2 showed a kinetic pattern similar to wild-type hPFKFB2, yet exhibited a notable 20-fold increase in affinity for citrate from 50.2 μM to 2.4 μM. All other kinetic parameters showed a modest increase in the mutant's affinity for substrates. Interestingly, the inhibitory effect of citrate was associated with a near 10-fold increase in V_{max}. This difference is likely attributable to a reduction in the non-catalytic, inhibitory binding of ATP in response to the lost π-π stacking interactions between phenyl ring and the adenine base of the inhibitory ATP binding.

The apparent differences in kinetic properties between wild-type and the mutant strongly suggest that Tyr428 makes significant contributions to citrate inhibition of 2-Kase by allowing ATP to bind in a non-catalytic, inhibitory fashion. As a consequence, citrate is not only an apparent competitive inhibitor against Fru-6-P but also makes ATP act as an apparent inhibitor.

The present study suggests that the inhibitory ATP binding significantly depends on Tyr428 mediated π-π stack interactions between the adenine base moiety of ATP and the tyrosine side chain. When these π-π stacking interactions were prevented, as with the Y428A mutant, the inhibitory binding of ATP was decreased, making the overall inhibition potency of citrate weaker.

The suggested model is in agreement with the previously accepted model of energy metabolism within cardiac cells. It has been well established that the heart cells depend mainly on oxidation of fatty acids and ketone bodies as energy source and that glycolysis is mostly down regulated in the heart through the so called "glucose-sparing effect". This suggests a possible model of the glucose-sparing effect. An unnecessary increase in the myocardial levels of Fru-2,6-P_2 , the most potent allosteric activator of the glycolysis rate-regulator protein, phosphofructokinase, is avoided by the citrate-sensitive 2-Kase domain in PFKFB2 isoform.

CHAPTER 5: CONCLUSIONS

Regulation of glucose metabolism is a vitally important aspect of cell physiology for both normal and diseased conditions. Many regulatory elements are involved in the process of regulating glycolytic flux; chief among these are the tissue-specific isozymes of PFKFB. Responsible for controlling the rate-limiting step of glycolysis, namely the conversion of F6P to F-1,6-P_2 by PFK-1, PFKFB isozymes catalyze both the synthesis and degradation of F-2,6-P_2 ─ the most potent allosteric activator of the entire glycolytic pathway. Through the regulation of cellular concentrations of F-2,6-P_2, the PFKFB isozymes exert tremendous influence over glycolytic flux, and consequently, cellular energy production. It's not surprising that alterations in expression levels of the PFKFB isozymes are tied to many diseases; such alterations have been shown to be a hallmark of virtually all types of cancer, often being required for both malignant transformation and neoplastic growth.

The inducible isoform, PFKFB3, has been implicated in the enhanced glycolytic activity observed in cancer cells, otherwise known as the Warburg effect. With a catalytic activity more than 10-fold higher than any other isoforms, PFKFB3 dramatically increases the concentration of F-2,6-P_2, allowing for 10- to 100-fold increases in glycolytic activity within growing tumors. This enhanced glycolysis has been reported to be crucial for tumorigenesis, with studies showing that genetic disruptions of PFKFB3 lead to apoptotic cell death of proliferation cancer cells. Accordingly, inhibition of this cancer-specific glycolysis has been suggested as a strategy to kill cancer cells, with PFKFB3 being the primary therapeutic target. Typically, the first step in the drug development process begins with computational screening, a technique used to reduce the number of molecules to be physically tested, allowing for faster and cheaper drug development. However, studies of the efficacy of virtual screening of small molecule kinases, the protein

family to which PFKFB3 belongs, are sparse, with none exploring the advantages and/or disadvantages of using tiered approaches. The first study presented in this work addressed this issue by investigating the efficacy of a pharmacophore screening-combined, structure-based docking approach on the human inducible PFKFB3 isoform. Five different structure-based docking programs were evaluated individually, and in conjunction, with a pharmacophore screening procedure using novel, bioactive compounds for PFKFB3. The study revealed significant differences in enrichment rates and screening times among single-method docking programs, with no discernable correlation between computational intensity and enrichment rates; whereas tiered approaches ─ pharmacophore screening in addition to structure-based docking ─ improved enrichments rates in 80% of cases, reduced CPU costs up to 7-fold, and lessened enrichment variability among the different structure-based docking methods. These findings are the first to demonstrate that tiered virtual screening approaches can be successfully applied to PFKFB3, and by extension small molecule kinases, by showing that such approaches can increase efficacy, reduce time, and normalize variability relative to structure-based docking methods alone.

The structural and kinetic characteristics of citrate, a well-known inhibitor of PFKFB enzymes, were also investigated using the heart isoform, PFKFB2. Citrate serves as an important negative feed-back mechanism to limit glycolytic activity through the inhibition of PFKFB enzymes, especially in cardiac cells, however, until now, the structural and mechanistic details of citrate's inhibition were wholly unknown. In this work, we established a molecular basis for citrate's inhibition through a combination of X-ray crystallography and inhibition kinetics. Citrate-complexed structures of the human and bovine PFKFB2 orthologues were solved to resolutions of 2.0 and 1.8Å, respectively. In both structures, citrate was observed to

fully occupy the binding site of Fructose-6-phosphate (F6P), competitively blocking F6P from binding, while also extending into the γ-phosphate binding site of ATP. This steric and electrostatic occlusion of the γ-phosphate binding site was observed to be highly consequential to the binding mode of the molecule residing in the ATP binding pocket. The bovine structure, which had ADP co-complexed with citrate, closely resembled other isoforms both in backbone and side-chain placement, with ADP closely mimicking the catalytic binding mode of ATP. The human structure, however, had AMPPNP co-complexed with citrate, which, unlike ADP, has a γ-phosphate. The presence of this γ-phosphate forces the adoption of a non-catalytic binding mode of AMPPNP in response to occlusion of the γ-phosphate binding site by citrate, prompting significant conformational changes within the ATP binding pocket of the 2-kinase domain. Inhibition kinetics were used to validate the structural observations and confirmed citrate's inhibition mechanism to be competitive for F6P and mixed for ATP. Prior to this work, no inhibitor had been identified as having a dual-action inhibition mechanism for PFKFB enzymes. Elucidation of citrates dual-action inhibition mechanism now enables citrate to be used as a foundation from which new, more potent inhibitors are able to be developed for PFKFB enzymes and thus aiding in the development of new therapeutic agents targeting diseases associated with PFKFB enzymes.

REFERENCES

Acharya, C., et al. (2011), 'Recent advances in ligand-based drug design: relevance and utility of the conformationally sampled pharmacophore approach', *Curr Comput Aided Drug Des,* 7 (1), 10-22.

Adams, C. P. and Brantner, V. V. (2010), 'Spending on new drug development1', *Health Econ,* 19 (2), 130-41.

Adams, P. D., et al. (2010), 'PHENIX: a comprehensive Python-based system for macromolecular structure solution', *Acta Crystallogr D Biol Crystallogr,* 66 (Pt 2), 213-21.

Akram, H. (2015), *In Silico Drug Discovery and Design* (CRC Press).

Al Hasawi, N., Alkandari, M. F., and Luqmani, Y. A. (2014), 'Phosphofructokinase: a mediator of glycolytic flux in cancer progression', *Crit Rev Oncol Hematol,* 92 (3), 312-21.

Algaier, J. and Uyeda, K. (1988), 'Molecular cloning, sequence analysis, and expression of a human liver cDNA coding for fructose-6-P,2-kinase:fructose-2,6-bisphosphatase', *Biochem Biophys Res Commun,* 153 (1), 328-33.

Amaratunga, D., Cabrera, J., and Lee, Y. S. (2015), 'Resampling-based similarity measures for high-dimensional data', *J Comput Biol,* 22 (1), 54-62.

Atsumi, T., et al. (2002), 'High expression of inducible 6-phosphofructo-2-kinase/fructose-2,6-bisphosphatase (iPFK-2; PFKFB3) in human cancers', *Cancer Res,* 62 (20), 5881-7.

Atsumi, T., et al. (2005), 'Expression of inducible 6-phosphofructo-2-kinase/fructose-2,6-bisphosphatase/PFKFB3 isoforms in adipocytes and their potential role in glycolytic regulation', *Diabetes,* 54 (12), 3349-57.

B-Rao, C., Subramanian, J., and Sharma, S. D. (2009), 'Managing protein flexibility in docking and its applications', *Drug Discovery Today,* 14 (7-8), 394-400.

Bacilieri, M. and Moro, S. (2006), 'Ligand-based drug design methodologies in drug discovery process: an overview', *Curr Drug Discov Technol,* 3 (3), 155-65.

Bajorath, J. (2002), 'Integration of virtual and high-throughput screening', *Nat Rev Drug Discov,* 1 (11), 882-94.

Baker, N. A. (2005), 'Biomolecular applications of Poisson-Boltzmann methods', *Reviews in Computational Chemistry, Vol 21,* 21, 349-79.

Ballester, P. J. and Mitchell, J. B. (2010), 'A machine learning approach to predicting protein-ligand binding affinity with applications to molecular docking', *Bioinformatics,* 26 (9), 1169-75.

Ballester, P. J., Finn, P. W., and Richards, W. G. (2009), 'Ultrafast shape recognition: evaluating a new ligand-based virtual screening technology', *J Mol Graph Model,* 27 (7), 836-45.

Banaszak, K., et al. (2011), 'The crystal structures of eukaryotic phosphofructokinases from baker's yeast and rabbit skeletal muscle', *J Mol Biol,* 407 (2), 284-97.

Barril, X. and Fradera, X. (2006a), 'Incorporating protein flexibility into docking and structure-based drug design', *Expert Opin Drug Discov,* 1 (4), 335-49.

--- (2006b), 'Incorporating protein flexibility into docking and structure-based drug design', *Expert Opinion on Drug Discovery,* 1 (4), 335-49.

Berg, J. M. (2012), *Biochemistry* (7th edn.: W. H. Freeman and Company).

Bertram, J. S. (2000), 'The molecular biology of cancer', *Mol Aspects Med,* 21 (6), 167-223.

Bikker, J. A., et al. (2009), 'Kinase Domain Mutations in Cancer: Implications for Small Molecule Drug Design Strategies', *Journal of Medicinal Chemistry,* 52 (6), 1493-509.

Bissantz, C., Folkers, G., and Rognan, D. (2000), 'Protein-based virtual screening of chemical databases. 1. Evaluation of different docking/scoring combinations', *Journal of Medicinal Chemistry,* 43 (25), 4759-67.

Blaney, J. (2012), 'A very short history of structure-based design: how did we get here and where do we need to go?', *Journal of Computer-Aided Molecular Design,* 26 (1), 13-14.

Bodansky, O. (1972), 'Acid phosphatase', *Adv Clin Chem,* 15, 43-147.

Boehm, M. (2011), *Virtual Screening: Principles, Challenges, and Practical Guidelines* (Wiley-VCH Verlag GmbH & Co.).

Bohm, G. and Jaenicke, R. (1992), 'Correlation functions as a tool for protein modeling and structure analysis', *Protein Sci,* 1 (10), 1269-78.

Bohm, H. J., Banner, D. W., and Weber, L. (1999), 'Combinatorial docking and combinatorial chemistry: design of potent non-peptide thrombin inhibitors', *J Comput Aided Mol Des,* 13 (1), 51-6.

Braga, R. C., et al. (2014), 'Virtual screening strategies in medicinal chemistry: the state of the art and current challenges', *Curr Top Med Chem,* 14 (16), 1899-912.

Brahimi-Horn, C. and Pouyssegur, J. (2006), 'The role of the hypoxia-inducible factor in tumor metabolism growth and invasion', *Bull Cancer,* 93 (8), E73-80.

Brooijmans, N. and Kuntz, I. D. (2003), 'Molecular recognition and docking algorithms', *Annu Rev Biophys Biomol Struct,* 32, 335-73.

Bui, T. and Thompson, C. B. (2006), 'Cancer's sweet tooth', *Cancer Cell,* 9 (6), 419-20.

Cabrera, P. J., et al. (2015), 'Evaluation of Tris-Bipyridine Chromium Complexes for Flow Battery Applications: Impact of Bipyridine Ligand Structure on Solubility and Electrochemistry', *Inorg Chem*.

Cai, W., Shao, X., and Maigret, B. (2002), 'Protein-ligand recognition using spherical harmonic molecular surfaces: towards a fast and efficient filter for large virtual throughput screening', *J Mol Graph Model,* 20 (4), 313-28.

Calvo, M. N., et al. (2006), 'PFKFB3 gene silencing decreases glycolysis, induces cell-cycle delay and inhibits anchorage-independent growth in HeLa cells', *FEBS Lett,* 580 (13), 3308-14.

Carlson, H. A. (2002a), 'Protein flexibility is an important component of structure-based drug discovery', *Curr Pharm Des,* 8 (17), 1571-8.

--- (2002b), 'Protein flexibility and drug design: how to hit a moving target', *Curr Opin Chem Biol,* 6 (4), 447-52.

Caro, J. (2001), 'Hypoxia regulation of gene transcription', *High Alt Med Biol,* 2 (2), 145-54.

Cavalier, M. C., et al. (2012), 'Molecular basis of the fructose-2,6-bisphosphatase reaction of PFKFB3: transition state and the C-terminal function', *Proteins,* 80 (4), 1143-53.

Cavasotto, C. N. and Orry, A. J. (2007), 'Ligand docking and structure-based virtual screening in drug discovery', *Curr Top Med Chem,* 7 (10), 1006-14.

Charifson, P. S., et al. (1999), 'Consensus scoring: A method for obtaining improved hit rates from docking databases of three-dimensional structures into proteins', *Journal of Medicinal Chemistry,* 42 (25), 5100-09.

Cheeseright, T. J., Mackey, M. D., and Scoffin, R. A. (2011), 'High content pharmacophores from molecular fields: a biologically relevant method for comparing and understanding ligands', *Curr Comput Aided Drug Des,* 7 (3), 190-205.

Chen, Y. C. (2015), 'Beware of docking!', *Trends in Pharmacological Sciences,* 36 (2), 78-95.

Chesney, J. (2006), '6-phosphofructo-2-kinase/fructose-2,6-bisphosphatase and tumor cell glycolysis', *Curr Opin Clin Nutr Metab Care,* 9 (5), 535-9.

Chesney, J., et al. (2015), 'Targeting the sugar metabolism of tumors with a first-in-class 6-phosphofructo-2-kinase (PFKFB4) inhibitor', *Oncotarget,* 6 (20), 18001-11.

Chesney, J., et al. (2005), 'Targeted disruption of inducible 6-phosphofructo-2-kinase results in embryonic lethality', *Biochem Biophys Res Commun,* 331 (1), 139-46.

Chesney, J., et al. (1999), 'An inducible gene product for 6-phosphofructo-2-kinase with an AU-rich instability element: role in tumor cell glycolysis and the Warburg effect', *Proc Natl Acad Sci U S A,* 96 (6), 3047-52.

Chikri, M. and Rousseau, G. G. (1995), 'Rat gene coding for heart 6-phosphofructo-2-kinase/fructose-2,6-bisphosphatase: characterization of an unusual promoter region and identification of four mRNAs', *Biochemistry,* 34 (27), 8876-84.

Clark, R. D. (2002), 'Getting past diversity in assessing virtual library designs', *Journal of the Brazilian Chemical Society,* 13 (6), 788-94.

Clem, B., et al. (2008), 'Small-molecule inhibition of 6-phosphofructo-2-kinase activity suppresses glycolytic flux and tumor growth', *Mol Cancer Ther,* 7 (1), 110-20.

Corbeil, C. R. and Moitessier, N. (2009), 'Docking Ligands into Flexible and Solvated Macromolecules. 3. Impact of Input Ligand Conformation, Protein Flexibility, and Water Molecules on the Accuracy of Docking Programs', *Journal of Chemical Information and Modeling,* 49 (4), 997-1009.

Corbeil, C. R., Therrien, E., and Moitessier, N. (2009), 'Modeling Reality for Optimal Docking of Small Molecules to Biological Targets', *Current Computer-Aided Drug Design,* 5 (4), 241-63.

Cosconati, S., et al. (2012), 'Protein flexibility in virtual screening: the BACE-1 case study', *J Chem Inf Model,* 52 (10), 2697-704.

Daga, P. R., Polgar, W. E., and Zaveri, N. T. (2014), 'Structure-based virtual screening of the nociceptin receptor: hybrid docking and shape-based approaches for improved hit identification', *J Chem Inf Model,* 54 (10), 2732-43.

Dang, C. V. and Semenza, G. L. (1999), 'Oncogenic alterations of metabolism', *Trends Biochem Sci,* 24 (2), 68-72.

Danishuddin, M. and Khan, A. U. (2015), 'Structure based virtual screening to discover putative drug candidates: necessary considerations and successful case studies', *Methods,* 71, 135-45.

Darville, M. I., Antoine, I. V., and Rousseau, G. G. (1992), 'Characterization of an enhancer upstream from the muscle-type promoter of a gene encoding 6-phosphofructo-2-kinase/fructose-2,6-bisphosphatase', *Nucleic Acids Res,* 20 (14), 3575-83.

DeBerardinis, R. J. and Thompson, C. B. (2012), 'Cellular metabolism and disease: what do metabolic outliers teach us?', *Cell,* 148 (6), 1132-44.

Deprez, J., et al. (1997), 'Phosphorylation and activation of heart 6-phosphofructo-2-kinase by protein kinase B and other protein kinases of the insulin signaling cascades', *J Biol Chem,* 272 (28), 17269-75.

Dias, R. and de Azevedo, W. F., Jr. (2008), 'Molecular docking algorithms', *Curr Drug Targets,* 9 (12), 1040-7.

Dias, R., et al. (2008a), 'Evaluation of molecular docking using polynomial empirical scoring functions', *Curr Drug Targets,* 9 (12), 1062-70.

Dias, R., et al. (2008b), 'Evaluation of Molecular Docking Using Polynomial Empirical Scoring Functions', *Current Drug Targets,* 9 (12), 1062-70.

Dickson, M. and Gagnon, J. P. (2004a), 'Key factors in the rising cost of new drug discovery and development', *Nat Rev Drug Discov,* 3 (5), 417-29.

--- (2004b), 'The cost of new drug discovery and development', *Discov Med,* 4 (22), 172-9.

Dihazi, H., Kessler, R., and Eschrich, K. (2003), 'Glucose-induced stimulation of the Ras-cAMP pathway in yeast leads to multiple phosphorylations and activation of 6-phosphofructo-2-kinase', *Biochemistry,* 42 (20), 6275-82.

DiMasi, J. A., Hansen, R. W., and Grabowski, H. G. (2003), 'The price of innovation: new estimates of drug development costs', *J Health Econ,* 22 (2), 151-85.

Dunaway, G. A. and Kasten, T. P. (1988), 'Physiological implications of the alteration of 6-phosphofructo-1-kinase isozyme pools during brain development and aging', *Brain Res,* 456 (2), 310-6.

Dunaway, G. A., et al. (1988), 'Analysis of the phosphofructokinase subunits and isoenzymes in human tissues', *Biochem J,* 251 (3), 677-83.

Dupriez, V. J., et al. (1993), 'Characterization of a hepatoma mRNA transcribed from a third promoter of a 6-phosphofructo-2-kinase/fructose-2,6-bisphosphatase-encoding gene and controlled by ets oncogene-related products', *Proc Natl Acad Sci U S A,* 90 (17), 8224-8.

Durrant, J. D. and McCammon, J. A. (2010), 'Computer-aided drug-discovery techniques that account for receptor flexibility', *Curr Opin Pharmacol,* 10 (6), 770-4.

Eckert, H. and Bajorath, J. (2007), 'Molecular similarity analysis in virtual screening: foundations, limitations and novel approaches', *Drug Discov Today,* 12 (5-6), 225-33.

El-Maghrabi, M. R. and Pilkis, S. J. (1984), 'Rat liver 6-phosphofructo 2-kinase/fructose 2,6-bisphosphatase: a review of relationships between the two activities of the enzyme', *J Cell Biochem,* 26 (1), 1-17.

El-Maghrabi, M. R., et al. (1982a), 'Cyclic AMP-dependent phosphorylation of rat liver 6-phosphofructo 2-kinase, fructose 2,6-bisphosphatase', *Biochem Biophys Res Commun,* 106 (3), 794-802.

El-Maghrabi, M. R., et al. (2001), '6-phosphofructo-2-kinase/fructose-2,6-bisphosphatase: suiting structure to need, in a family of tissue-specific enzymes', *Curr Opin Clin Nutr Metab Care,* 4 (5), 411-8.

El-Maghrabi, M. R., et al. (1982b), 'Regulation of rat liver fructose 2,6-bisphosphatase', *J Biol Chem,* 257 (13), 7603-7.

Emsley, P. and Cowtan, K. (2004), 'Coot: model-building tools for molecular graphics', *Acta Crystallogr D Biol Crystallogr,* 60 (Pt 12 Pt 1), 2126-32.

Evans, P. R., Farrants, G. W., and Hudson, P. J. (1981), 'Phosphofructokinase: structure and control', *Philos Trans R Soc Lond B Biol Sci,* 293 (1063), 53-62.

Ewing, T. J., et al. (2001), 'DOCK 4.0: search strategies for automated molecular docking of flexible molecule databases', *Journal of Computer-Aided Molecular Design,* 15 (5), 411-28.

Fell, D. (1997), *Understanding the control of metabolism.* (London: Portland Press).

Ferreira, L. G., Oliva, G., and Andricopulo, A. D. (2015), 'Target-based molecular modeling strategies for schistosomiasis drug discovery', *Future Medicinal Chemistry,* 7 (6), 753-64.

Fischer, M., et al. (2014), 'Incorporation of protein flexibility and conformational energy penalties in docking screens to improve ligand discovery', *Nat Chem,* 6 (7), 575-83.

Flatt, J. P. (1995), 'Carbohydrate balance and food intake regulation', *Am J Clin Nutr,* 62 (1), 155-7.

Fothergill-Gilmore, L. A. (1987), 'Evolution in glycolysis', *Biochem Soc Trans,* 15 (5), 993-5.

Fothergill-Gilmore, L. A. and Watson, H. C. (1989), 'The phosphoglycerate mutases', *Adv Enzymol Relat Areas Mol Biol,* 62, 227-313.

Fox, M. P. and Fox, L. M. (2006), 'Resolving design problems in equivalency trials', *J Pediatr,* 149 (1), 12-6.

Fox, S., et al. (2006), 'High-throughput screening: update on practices and success', *J Biomol Screen,* 11 (7), 864-9.

Friesner, R. A., et al. (2004), 'Glide: a new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy', *J Med Chem,* 47 (7), 1739-49.

Furuya, E. and Uyeda, K. (1980a), 'An activation factor of liver phosphofructokinase', *Proc Natl Acad Sci U S A,* 77 (10), 5861-4.

--- (1980b), 'Regulation of phosphofructokinase by a new mechanism. An activation factor binding to phosphorylated enzyme', *J Biol Chem,* 255 (24), 11656-9.

Furuya, E., Yokoyama, M., and Uyeda, K. (1982), 'Regulation of fructose-6-phosphate 2-kinase by phosphorylation and dephosphorylation: possible mechanism for coordinated control of glycolysis and glycogenolysis', *Proc Natl Acad Sci U S A,* 79 (2), 325-9.

Galluzzi, L., et al. (2013), 'Metabolic targets for cancer therapy', *Nat Rev Drug Discov,* 12 (11), 829-46.

Gane, P. J. and Dean, P. M. (2000), 'Recent advances in structure-based rational drug design', *Curr Opin Struct Biol,* 10 (4), 401-4.

Gao, F., et al. (2002), 'Optimizing higher throughput methods to assess drug-drug interactions for CYP1A2, CYP2C9, CYP2C19, CYP2D6, rCYP2D6, and CYP3A4 in vitro using a single point IC50', *Journal of Biomolecular Screening,* 7 (4), 373-82.

Garber, K. (2004), 'Energy boost: the Warburg effect returns in a new theory of cancer', *J Natl Cancer Inst,* 96 (24), 1805-6.

Gatenby, R. A. and Gillies, R. J. (2004), 'Why do cancers have high aerobic glycolysis?', *Nat Rev Cancer,* 4 (11), 891-9.

Ghosh, A., Rapp, C. S., and Friesner, R. A. (1998), 'Generalized born model based on a surface integral formulation', *Journal of Physical Chemistry B,* 102 (52), 10983-90.

Goodsell, D. S., Morris, G. M., and Olson, A. J. (1996), 'Automated docking of flexible ligands: applications of AutoDock', *J Mol Recognit,* 9 (1), 1-5.

Goodwin, J. T., et al. (2002), 'In silico methodologies in accelerating drug development.', *Abstracts of Papers of the American Chemical Society,* 223, U348-U48.

Grant, D., Long, W. F., and Williamson, F. B. (1996), 'Similarity and dissimilarity in aspects of the binding to heparin of Ca2+ and Zn2+ as revealed by potentiometric titration', *Biochem Soc Trans,* 24 (2), 203S.

Green, D. R. and Chipuk, J. E. (2006), 'p53 and metabolism: Inside the TIGAR', *Cell,* 126 (1), 30-2.

Group, Chemical Computing (2010), 'MOE (Molecular Operating Environment) 2010', (2010 edn.; Montreal, Canada: Chemical Computing Group).

Guido, R. V., Oliva, G., and Andricopulo, A. D. (2008), 'Virtual Screening and Its Integration with Modern Drug Design Technologies', *Current Medicinal Chemistry,* 15 (1), 37-46.

Guner, O. F. and Bowen, J. P. (2014), 'Setting the Record Straight: The Origin of the Pharmacophore Concept', *Journal of Chemical Information and Modeling,* 54 (5), 1269-83.

Haigh, J. A., et al. (2005), 'Small molecule shape-fingerprints', *J Chem Inf Model,* 45 (3), 673-84.

Halperin, I., et al. (2002), 'Principles of docking: An overview of search algorithms and a guide to scoring functions', *Proteins-Structure Function and Genetics,* 47 (4), 409-43.

Hasemann, C. A., et al. (1996), 'The crystal structure of the bifunctional enzyme 6-phosphofructo-2-kinase/fructose-2,6-bisphosphatase reveals distinct domain homologies', *Structure,* 4 (9), 1017-29.

Hay, M., et al. (2014), 'Clinical development success rates for investigational drugs', *Nat Biotechnol,* 32 (1), 40-51.

Heine-Suner, D., et al. (1998), 'Sequence and structure of the human 6-phosphofructo-2-kinase/fructose-2,6-bisphosphatase heart isoform gene (PFKFB2)', *Eur J Biochem,* 254 (1), 103-10.

Holley, A. K., Dhar, S. K., and St Clair, D. K. (2012), 'Curbing cancer's sweet tooth: Is there a role for MnSOD in regulation of the Warburg effect?', *Mitochondrion*.

Horvath, D. (2011), 'Pharmacophore-based virtual screening', *Methods Mol Biol,* 672, 261-98.

Horvath, D. and Jeandenans, C. (2003), 'Neighborhood behavior of in silico structural spaces with respect to in vitro activity spaces--a benchmark for neighborhood behavior assessment of different in silico similarity metrics', *J Chem Inf Comput Sci,* 43 (2), 691-8.

Hsu, P. P. and Sabatini, D. M. (2008), 'Cancer cell metabolism: Warburg and beyond', *Cell,* 134 (5), 703-7.

Huang, S. M., et al. (2010), 'Therapeutic protein-drug interactions and implications for drug development', *Clin Pharmacol Ther,* 87 (4), 497-503.

Hughes, J. P., et al. (2011), 'Principles of early drug discovery', *Br J Pharmacol,* 162 (6), 1239-49.

Janson, S., Merkle, D., and Middendorf, M. (2008), 'Molecular docking with multi-objective particle swarm optimization', *Applied Soft Computing,* 8 (1), 666-75.

Jones, G., et al. (1997), 'Development and validation of a genetic algorithm for flexible docking.', *Journal of molecular biology,* 267 (3), 727-48.

Jones, R. G. and Thompson, C. B. (2009), 'Tumor suppressors and cell metabolism: a recipe for cancer growth', *Genes Dev,* 23 (5), 537-48.

Jorissen, R. N. and Gilson, M. K. (2005), 'Virtual screening of molecular databases using a support vector machine', *J Chem Inf Model,* 45 (3), 549-61.

Kaitin, K (2015), 'Tufts CSDD', (Tufts University).

Kalliokoski, T., Olsson, T. S., and Vulpetti, A. (2013), 'Subpocket analysis method for fragment-based drug discovery', *J Chem Inf Model,* 53 (1), 131-41.

Ke, Q. and Costa, M. (2006), 'Hypoxia-inducible factor-1 (HIF-1)', *Mol Pharmacol,* 70 (5), 1469-80.

Kessler, R. and Eschrich, K. (2001), 'Splice isoforms of ubiquitous 6-phosphofructo-2-kinase/fructose-2,6-bisphosphatase in human brain', *Brain Res Mol Brain Res,* 87 (2), 190-5.

Kier, L. B. and Roche, E. B. (1967), 'Medicinal chemistry of the mesoionic compounds', *J Pharm Sci,* 56 (2), 149-68.

Kim, S. G., et al. (2006), 'Crystal structure of the hypoxia-inducible form of 6-phosphofructo-2-kinase/fructose-2,6-bisphosphatase (PFKFB3): a possible new target for cancer therapy', *J Biol Chem,* 281 (5), 2939-44.

Kim, S. G., et al. (2007), 'A direct substrate-substrate interaction found in the kinase domain of the bifunctional enzyme, 6-phosphofructo-2-kinase/fructose-2,6-bisphosphatase', *J Mol Biol,* 370 (1), 14-26.

Kitajima, S., Sakakibara, R., and Uyeda, K. (1984), 'Kinetic studies of fructose 6-phosphate,2-kinase and fructose 2,6-bisphosphatase', *J Biol Chem,* 259 (11), 6896-903.

Kitchen, D. B., Stahura, F. L., and Bajorath, J. (2004), 'Computational techniques for diversity analysis and compound classification', *Mini Rev Med Chem,* 4 (10), 1029-39.

Klahn, M., Rosta, E., and Warshel, A. (2006), 'On the mechanism of hydrolysis of phosphate monoesters dianions in solutions and proteins', *J Am Chem Soc,* 128 (47), 15310-23.

Klebe, G. (2006), 'Virtual ligand screening: strategies, perspectives and limitations', *Drug Discovery Today,* 11 (13-14), 580-94.

Koveal, D., et al. (2013), 'Ligand Binding Reduces Conformational Flexibility in the Active Site of Tyrosine Phosphatase Related to Biofilm Formation A (TpbA) from Pseudomonas aeruginosa', *Journal of Molecular Biology,* 425 (12), 2219-31.

Kroemer, G and Pouyssegur, J (2008), 'Tumor cell metabolism: cancer's Achilles' heel', *Cancer Cell,* 13, 472-82.

Kruger, D. M. and Evers, A. (2010), 'Comparison of Structure- and Ligand-Based Virtual Screening Protocols Considering Hit List Complementarity and Enrichment Factors', *Chemmedchem,* 5 (1), 148-58.

Krumrine, J., et al. (2003), 'Principles and methods of docking and ligand design', *Methods Biochem Anal,* 44, 443-76.

Lahiri, S. D., et al. (2003), 'The pentacovalent phosphorus intermediate of a phosphoryl transfer reaction', *Science,* 299 (5615), 2067-71.

Landry, Y. and Gies, J. P. (2008), 'Drugs and their molecular targets: an updated overview', *Fundamental & Clinical Pharmacology,* 22 (1), 1-18.

Lange, A. J. and Pilkis, S. J. (1990), 'Sequence of human liver 6-phosphofructo-2-kinase/fructose-2,6-bisphosphatase', *Nucleic Acids Res,* 18 (12), 3652.

Leach, A. R., et al. (2010), 'Three-Dimensional Pharmacophore Methods in Drug Discovery', *Journal of Medicinal Chemistry,* 53 (2), 539-58.

Lee, K., et al. (2010), 'Pharmacophore modeling and virtual screening studies for new VEGFR-2 kinase inhibitors', *European Journal of Medicinal Chemistry,* 45 (11), 5420-27.

Lee, Y. H., et al. (2003), 'Tissue-specific structure/function differentiation of the liver isoform of 6-phosphofructo-2-kinase/fructose-2,6-bisphosphatase', *J Biol Chem,* 278 (1), 523-30.

Leipe, D. D., Koonin, E. V., and Aravind, L. (2003), 'Evolution and classification of P-loop kinases and related proteins', *J Mol Biol,* 333 (4), 781-815.

Leipe, D. D., et al. (2002), 'Classification and evolution of P-loop GTPases and related ATPases', *J Mol Biol,* 317 (1), 41-72.

Levine, A. J. and Puzio-Kuter, A. M. (2010), 'The control of the metabolic switch in cancers by oncogenes and tumor suppressor genes', *Science,* 330 (6009), 1340-4.

Lewis, B. C., et al. (1997), 'Identification of putative c-Myc-responsive genes: characterization of rcl, a novel growth-related gene', *Mol Cell Biol,* 17 (9), 4967-78.

Lexa, K. W. and Carlson, H. A. (2012), 'Protein flexibility in docking and surface mapping', *Quarterly Reviews of Biophysics,* 45 (3), 301-43.

Li, X., et al. (2010), 'Evaluation of the Performance of Four Molecular Docking Programs on a Diverse Set of Protein-Ligand Complexes', *Journal of Computational Chemistry,* 31 (11), 2109-25.

Lloyd, B. K. and McElwee, P. R. (2011), 'Trends over time in characteristics of pharmaceutical drug-related ambulance attendances in Melbourne', *Drug Alcohol Rev,* 30 (3), 271-80.

Lodish, H. (2000), *Molecular Cell Biology* (5th edn.: W. H. Freeman).

Lyne, P. D. (2002), 'Structure-based virtual screening: an overview', *Drug Discov Today,* 7 (20), 1047-55.

Ma, D. L., et al. (2012a), 'Discovery of a natural product-like c-myc G-quadruplex DNA groove-binder by molecular docking', *PLoS One,* 7 (8), e43278.

Ma, X. H., et al. (2012b), 'Virtual screening methods as tools for drug lead discovery from large chemical libraries', *Curr Med Chem,* 19 (32), 5562-71.

Macarron, R. (2006), 'Critical review of the role of HTS in drug discovery', *Drug Discov Today,* 11 (7-8), 277-9.

Macarron, R. and Luengo, J. I. (2011), 'Yin and Yang in medicinal chemistry: what does drug-likeness mean?', *Future Med Chem,* 3 (5), 505-7.

Macarron, R., et al. (2011), 'Impact of high-throughput screening in biomedical research', *Nat Rev Drug Discov,* 10 (3), 188-95.

Maddocks, O. D., et al. (2013), 'Serine starvation induces stress and p53-dependent metabolic remodelling in cancer cells', *Nature,* 493 (7433), 542-6.

Maggiora, G. M. and Shanmugasundaram, V. (2011), 'Molecular similarity measures', *Methods Mol Biol,* 672, 39-100.

Mahajan, R. and Gupta, K. (2010a), 'Adaptive design clinical trials: Methodology, challenges and prospect', *Indian J Pharmacol,* 42 (4), 201-7.

--- (2010b), 'Food and drug administration's critical path initiative and innovations in drug development paradigm: Challenges, progress, and controversies', *J Pharm Bioallied Sci,* 2 (4), 307-13.

Maiorov, V. and Sheridan, R. P. (2005), 'Enhanced virtual screening by combined use of two docking methods: getting the most on a limited budget', *J. Chem. Inf. Model.,* 45 (4), 1017-23.

Mangoni, R., Roccatano, D., and Di Nola, A. (1999), 'Docking of flexible ligands to flexible receptors in solution by molecular dynamics simulation', *Proteins-Structure Function and Bioinformatics,* 35 (2), 153-62.

Manzano, A., et al. (1999), 'Cloning, expression and chromosomal localization of a human testis 6-phosphofructo-2-kinase/fructose-2,6-bisphosphatase gene', *Gene,* 229 (1-2), 83-9.

Manzano, A., et al. (1998), 'Molecular cloning, expression, and chromosomal localization of a ubiquitously expressed human 6-phosphofructo-2-kinase/ fructose-2, 6-bisphosphatase gene (PFKFB3)', *Cytogenet Cell Genet,* 83 (3-4), 214-7.

Marsden, B. D. and Knapp, S. (2008), 'Doing more than just the structure-structural genomics in kinase drug discovery', *Curr Opin Chem Biol,* 12 (1), 40-5.

May, A. and Zacharias, M. (2005), 'Accounting for global protein deformability during protein-protein and protein-ligand docking', *Biochim Biophys Acta,* 1754 (1-2), 225-31.

McGaughey, G. B., et al. (2007), 'Comparison of topological, shape, and docking methods in virtual screening', *J Chem Inf Model,* 47 (4), 1504-19.

Mcgilvery, R. W. (1979), 'Fuel for Breathing', *American Review of Respiratory Disease,* 119 (2), 85-88.

McGovern, S. L., et al. (2002), 'A common mechanism underlying promiscuous inhibitors from virtual and high-throughput screening', *Journal of Medicinal Chemistry,* 45 (8), 1712-22.

Metallo, C. M. and Vander Heiden, M. G. (2013), 'Understanding metabolic regulation and its influence on cell physiology', *Mol Cell,* 49 (3), 388-98.

Metropolis, N., et al. (1953), 'Equation of State Calculations by Fast Computing Machines', *Journal of Chemical Physics,* 21 (6), 1087-92.

Millan, Alex and (2012), *Drug Design Strategies : Quantitative Approaches* (RSC Publishing).

Minchenko, A., et al. (2002), 'Hypoxia-inducible factor-1-mediated expression of the 6-phosphofructo-2-kinase/fructose-2,6-bisphosphatase-3 (PFKFB3) gene. Its possible role in the Warburg effect', *J Biol Chem,* 277 (8), 6183-7.

Minchenko, O., Opentanova, I., and Caro, J. (2003), 'Hypoxic regulation of the 6-phosphofructo-2-kinase/fructose-2,6-bisphosphatase gene family (PFKFB-1-4) expression in vivo', *FEBS Lett,* 554 (3), 264-70.

Moitessier, N., et al. (2008), 'Towards the development of universal, fast and highly accurate docking/scoring methods: a long way to go', *British Journal of Pharmacology,* 153, S7-S26.

Moore, K. and Rees, S. (2001), 'Cell-based versus isolated target screening: how lucky do you feel?', *J Biomol Screen,* 6 (2), 69-74.

Mor, I., Cheung, E. C., and Vousden, K. H. (2011), 'Control of glycolysis through regulation of PFK1: old friends and recent additions', *Cold Spring Harb Symp Quant Biol,* 76, 211-6.

Moreno-Sanchez, R., et al. (2007), 'Energy metabolism in tumor cells', *FEBS J,* 274 (6), 1393-418.

Morris, G. M., et al. (1996), 'Distributed automated docking of flexible ligands to proteins: parallel applications of AutoDock 2.4', *J Comput Aided Mol Des,* 10 (4), 293-304.

Muegge, I. and Rarey, M. (2001), 'Small Molecule Docking and Scoring', *Reviews in Computational Chemistry, Vol 17,* 17, 1-60.

Muegge, I., Heald, S. L., and Brittelli, D. (2001), 'Simple selection criteria for drug-like chemical matter', *J Med Chem,* 44 (12), 1841-6.

Murray, K. J., et al. (1984), 'Amino acid sequence of the phosphorylation site of rat liver 6-phosphofructo-2-kinase/fructose-2,6-bisphosphatase', *J Biol Chem,* 259 (12), 7673-81.

Mysinger, M. M. and Shoichet, B. K. (2010), 'Rapid Context-Dependent Ligand Desolvation in Molecular Docking', *Journal of Chemical Information and Modeling,* 50 (9), 1561-73.

Nagarajan, S., et al. (2010), 'IKK beta inhibitors identification part II: Ligand and structure-based virtual screening', *Bioorganic & Medicinal Chemistry,* 18 (11), 3951-60.

Natarajan, N., et al. (2011), 'Novel immunotherapeutic agents and small molecule antagonists of signalling kinases for the treatment of metastatic melanoma', *Drugs,* 71 (10), 1233-50.

Navarro-Sabate, A., et al. (2001), 'The human ubiquitous 6-phosphofructo-2-kinase/fructose-2,6-bisphosphatase gene (PFKFB3): promoter characterization and genomic structure', *Gene,* 264 (1), 131-8.

Obach, M., et al. (2004), '6-Phosphofructo-2-kinase (pfkfb3) gene promoter contains hypoxia-inducible factor-1 binding sites necessary for transactivation in response to hypoxia', *J Biol Chem,* 279 (51), 53562-70.

Okar, D. A., et al. (2001), 'PFK-2/FBPase-2: maker and breaker of the essential biofactor fructose-2,6-bisphosphate', *Trends Biochem Sci,* 26 (1), 30-5.

Orry, A. J., Abagyan, R. A., and Cavasotto, C. N. (2006), 'Structure-based development of target-specific compound libraries', *Drug Discov Today,* 11 (5-6), 261-6.

Parker, C. N. and Bajorath, J. (2006), 'Towards unified compound screening strategies: A critical evaluation of error sources in experimental and virtual high-throughput screening', *Qsar & Combinatorial Science,* 25 (12), 1153-61.

Pelicano, H., et al. (2006), 'Glycolysis inhibition for anticancer treatment', *Oncogene,* 25 (34), 4633-46.

Peltason, L. (2008), *Chemoinformatics Approaches to Virtual Screening* (Royal Society of Chemistry).

Peltason, L. and Bajorath, J. (2011), 'Computational analysis of activity and selectivity cliffs', *Methods Mol Biol,* 672, 119-32.

Peltason, L., Iyer, P., and Bajorath, J. (2010), 'Rationalizing three-dimensional activity landscapes and the influence of molecular representations on landscape topology and the formation of activity cliffs', *J Chem Inf Model,* 50 (6), 1021-33.

Pfeiffer, T., Schuster, S., and Bonhoeffer, S. (2001), 'Cooperation and competition in the evolution of ATP-producing pathways', *Science,* 292 (5516), 504-7.

Pilkis, S. J. and Granner, D. K. (1992), 'Molecular physiology of the regulation of hepatic gluconeogenesis and glycolysis', *Annu Rev Physiol,* 54, 885-909.

Pilkis, S. J., et al. (1995), '6-Phosphofructo-2-kinase/fructose-2,6-bisphosphatase: a metabolic signaling enzyme', *Annu Rev Biochem,* 64, 799-835.

Pirhadi, S., Shiri, F., and Ghasemi, J. B. (2013), 'Methods and Applications of Structure Based Pharmacophores in Drug Discovery', *Current Topics in Medicinal Chemistry,* 13 (9), 1036-47.

Plewczynski, D., et al. (2011), 'Can We Trust Docking Results? Evaluation of Seven Commonly Used Programs on PDBbind Database', *Journal of Computational Chemistry,* 32 (4), 742-55.

Polishchuk, P. G., Madzhidov, T. I., and Varnek, A. (2013), 'Estimation of the size of drug-like chemical space based on GDB-17 data', *J Comput Aided Mol Des,* 27 (8), 675-9.

Ramakrishnan, C., Dani, V. S., and Ramasarma, T. (2002), 'A conformational analysis of Walker motif A [GXXXXGKT (S)] in nucleotide-binding and other proteins', *Protein Eng,* 15 (10), 783-98.

Rarey, M., Wefing, S., and Lengauer, T. (1996), 'Placement of medium-sized molecular fragments into active sites of proteins', *J Comput Aided Mol Des,* 10 (1), 41-54.

Reinhart, G. D. and Lardy, H. A. (1980), 'Rat liver phosphofructokinase: kinetic activity under near-physiological conditions', *Biochemistry,* 19 (7), 1477-84.

Reymond, J. L. and Awale, M. (2012), 'Exploring chemical space for drug discovery using the chemical universe database', *ACS Chem Neurosci,* 3 (9), 649-57.

Rider, M. H., et al. (2004), '6-phosphofructo-2-kinase/fructose-2,6-bisphosphatase: head-to-head with a bifunctional enzyme that controls glycolysis', *Biochem J,* 381 (Pt 3), 561-79.

Rigden, D. J. (2008), 'The histidine phosphatase superfamily: structure and function', *Biochem J,* 409 (2), 333-48.

Ripphausen, P., Nisius, B., and Bajorath, J. (2011), 'State-of-the-art in ligand-based virtual screening', *Drug Discov Today,* 16 (9-10), 372-6.

Ros, S. and Schulze, A. (2013), 'Balancing glycolytic flux: the role of 6-phosphofructo-2-kinase/fructose 2,6-bisphosphatases in cancer metabolism', *Cancer Metab,* 1 (1), 8.

Sakakibara, R., et al. (1997), 'Characterization of a human placental fructose-6-phosphate, 2-kinase/fructose-2,6-bisphosphatase', *J Biochem,* 122 (1), 122-8.

Sakakibara, R., et al. (1999), 'Tissue distribution of placenta-type 6-phosphofructo- 2-kinase/fructose-2,6-bisphosphatase', *Biochem Biophys Res Commun,* 257 (1), 177-81.

Sakata, J., Abe, Y., and Uyeda, K. (1991), 'Molecular cloning of the DNA and expression and characterization of rat testes fructose-6-phosphate,2-kinase:fructose-2,6-bisphosphatase', *J Biol Chem,* 266 (24), 15764-70.

Seo, M., et al. (2011), 'Structure-based development of small molecule PFKFB3 inhibitors: a framework for potential cancer therapeutic agents targeting the Warburg effect', *PLoS One,* 6 (9), e24179.

Shim, H., et al. (1998), 'A unique glucose-dependent apoptotic pathway induced by c-Myc', *Proc Natl Acad Sci U S A,* 95 (4), 1511-6.

Sotriffer, C. A. (2015), *In Silico Drug Discovery and Design: Theory, Methods, Challenges, and Applications* (Protein-Ligand Docking: From Basic Principles to Advanced Applications: CRCpress).

Sousa, S. F., et al. (2010), 'Virtual screening in drug design and development', *Combinatorial Chemistry & High Throughput Screening,* 13 (5), 442-53.

Sprous, D. G., et al. (2002), 'Assessing the quality of virtual screening results for combinatorial libraries.', *Abstracts of Papers of the American Chemical Society,* 224, U344-U44.

Stahura, F. L. and Bajorath, M. (2005), 'New methodologies for ligand-based virtual screening', *Current Pharmaceutical Design,* 11 (9), 1189-202.

Swann, S. L., et al. (2011), 'A unified, probabilistic framework for structure- and ligand-based virtual screening', *Journal of medicinal chemistry,* 54 (5), 1223-32.

Tame, J. R. (1999), 'Scoring functions: a view from the bench', *J Comput Aided Mol Des,* 13 (2), 99-108.

Tan, L., et al. (2008), 'Integrating structure- and ligand-based virtual screening: comparison of individual, parallel, and fused molecular docking and similarity search calculations on multiple targets', *Chemmedchem,* 3 (10), 1566-71.

Taniyama, M., et al. (1988), 'Isozymes of fructose 6-phosphate,2-kinase:fructose-2,6-bisphosphatase in rat and bovine heart, liver, and skeletal muscle', *Biochem Biophys Res Commun,* 157 (3), 949-54.

Telang, S., et al. (2006), 'Ras transformation requires metabolic control by 6-phosphofructo-2-kinase', *Oncogene,* 25 (55), 7225-34.

Trott, O. and Olson, A. J. (2010), 'AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading', *Journal of Computational Chemistry,* 31 (2), 455-61.

Tuccinardi, T. (2009), 'Docking-Based Virtual Screening: Recent Developments', *Combinatorial Chemistry & High Throughput Screening,* 12 (3), 303-14.

Van Schaftingen, E. and Hers, H. G. (1981), 'Phosphofructokinase 2: the enzyme that forms fructose 2,6-bisphosphate from fructose 6-phosphate and ATP', *Biochem Biophys Res Commun,* 101 (3), 1078-84.

van Schaftingen, E., Davies, D. R., and Hers, H. G. (1982), 'Fructose-2,6-bisphosphatase from rat liver', *Eur J Biochem,* 124 (1), 143-9.

Villoutreix, B. O., Eudes, R., and Miteva, M. A. (2009), 'Structure-Based Virtual Ligand Screening: Recent Success Stories', *Combinatorial Chemistry & High Throughput Screening,* 12 (10), 1000-16.

Walker, J. E., et al. (1982), 'Distantly related sequences in the alpha- and beta-subunits of ATP synthase, myosin, kinases and other ATP-requiring enzymes and a common nucleotide binding fold', *EMBO J,* 1 (8), 945-51.

Wang, Z. and Busemeyer, J. R. (2013), 'A quantum question order model supported by empirical tests of an a priori and precise prediction', *Top Cogn Sci,* 5 (4), 689-710.

Warburg, O. (1956), 'On the origin of cancer cells', *Science,* 123 (3191), 309-14.

Warren, G. L., et al. (2006), 'A critical assessment of docking programs and scoring functions', *Journal of Medicinal Chemistry,* 49 (20), 5912-31.

Wawer, M., et al. (2008), 'Structure-activity relationship anatomy by network-like similarity graphs and local structure-activity relationship indices', *J Med Chem,* 51 (19), 6075-84.

Webb, B. A., et al. (2015), 'Structures of human phosphofructokinase-1 and atomic basis of cancer-associated mutations', *Nature,* 523 (7558), 111-4.

Wechalekar, K., Sharma, B., and Cook, G. (2005), 'PET/CT in oncology--a major advance', *Clin Radiol,* 60 (11), 1143-55.

Wegner, A., et al. (2015), 'How metabolites modulate metabolic flux', *Curr Opin Biotechnol,* 34, 16-22.

Wei, B. Q. Q., et al. (2002), 'A model binding site for testing scoring functions in molecular docking', *Journal of Molecular Biology,* 322 (2), 339-55.

Westhead, D. R., Clark, D. E., and Murray, C. W. (1997), 'A comparison of heuristic search algorithms for molecular docking', *J Comput Aided Mol Des,* 11 (3), 209-28.

Willett, P. (2006), 'Similarity-based virtual screening using 2D fingerprints', *Drug Discov Today,* 11 (23-24), 1046-53.

Wittinghofer, A. (2006), 'Phosphoryl transfer in Ras proteins, conclusive or elusive?', *Trends Biochem Sci,* 31 (1), 20-3.

Workman, P. and van Montfort, R. L. M. (2010), 'PI(3) KINASES Revealing the delta lady', *Nature Chemical Biology,* 6 (2), 82-83.

Wu, C., et al. (2006), 'Roles for fructose-2,6-bisphosphate in the control of fuel metabolism: beyond its allosteric effects on glycolytic and gluconeogenic enzymes', *Adv Enzyme Regul,* 46, 72-88.

Wu, G. S., et al. (2003), 'Detailed analysis of grid-based molecular docking: A case study of CDOCKER - A CHARMm-based MD docking algorithm', *Journal of Computational Chemistry,* 24 (13), 1549-62.

Xu, R. H., et al. (2005), 'Inhibition of glycolysis in cancer cells: a novel strategy to overcome drug resistance associated with mitochondrial respiratory defect and hypoxia', *Cancer Res,* 65 (2), 613-21.

Xue, L., Stahura, F. L., and Bajorath, J. (2005), 'Chemoinformatics: Perspectives and challenges', *Chemometrics and Chemoinformatics,* 894, 41-53.

Yalcin, A., et al. (2009), 'Regulation of glucose metabolism by 6-phosphofructo-2-kinase/fructose-2,6-bisphosphatases in cancer', *Exp Mol Pathol,* 86 (3), 174-9.

Yeluri, S., et al. (2009), 'Cancer's craving for sugar: an opportunity for clinical exploitation', *J Cancer Res Clin Oncol,* 135 (7), 867-77.

Yuan, Z., et al. (1990), 'The role of fructose 2,6-bisphosphate in glycolytic oscillations in extracts and cells of Saccharomyces cerevisiae', *Eur J Biochem,* 192 (3), 791-5.

Yuen, M. H., et al. (1999), 'A switch in the kinase domain of rat testis 6-phosphofructo-2-kinase/fructose-2,6-bisphosphatase', *Biochemistry,* 38 (38), 12333-42.

Yuriev, E. and Ramsland, P. A. (2013), 'Latest developments in molecular docking: 2010-2011 in review', *J Mol Recognit,* 26 (5), 215-39.

Yuriev, E., Agostino, M., and Ramsland, P. A. (2011), 'Challenges and advances in computational docking: 2009 in review', *J Mol Recognit,* 24 (2), 149-64.

Zilian, D. and Sotriffer, C. A. (2013), 'SFCscore(RF): a random forest-based scoring function for improved affinity prediction of protein-ligand complexes', *J Chem Inf Model,* 53 (8), 1923-33.

VITA

Robert Crochet was born in September of 1983, in Baton Rouge, Louisiana. He entered Louisiana State University in 2003, receiving his bachelor's degree in 2008. He then joined the laboratory of Dr. Yong-Hwan Lee at Louisiana State University as a graduate student in the fall of 2010. Robert Crochet will be defending his dissertation in fall of 2015 and he will await the issuing of his Doctor of Philosophy degree in Biochemistry in December 2015.