Georgia State University
ScholarWorks @ Georgia State University

Chemistry Theses | Department of Chemistry

Fall 12-17-2019

Exploring the Dynamics of Glycolytic Regulation Through Two Enzymes: Human Enolase and Liver Pyruvate Kinase

Abdullah Shouaib
ashouaib1@student.gsu.edu

Follow this and additional works at: https://scholarworks.gsu.edu/chemistry_theses

Recommended Citation

Shouaib, Abdullah, "Exploring the Dynamics of Glycolytic Regulation Through Two Enzymes: Human Enolase and Liver Pyruvate Kinase." Thesis, Georgia State University, 2019.
https://scholarworks.gsu.edu/chemistry_theses/134

This Thesis is brought to you for free and open access by the Department of Chemistry at ScholarWorks @ Georgia State University. It has been accepted for inclusion in Chemistry Theses by an authorized administrator of ScholarWorks @ Georgia State University. For more information, please contact scholarworks@gsu.edu.

EXPLORING THE DYNAMICS OF GLYCOLYTIC REGULATION THROUGH TWO ENZYMES: HUMAN ENOLASE AND LIVER PYRUVATE KINASE

by

ABDULLAH SHOUAIB

Under the Direction of Donald Hamelberg, PhD

ABSTRACT

Understanding how conformational dynamics play a role in metabolic regulation is an important objective in various biological disciplines. Using Molecular Dynamics (MD) simulations, we attempt to elucidate the dynamics involved in regulating two glycolytic enzymes: human liver pyruvate kinase (hL-PYK) and human enolase. Despite being metabolically coupled, both enzymes are regulated quite differently: hL-PYK can undergo allosteric modulation, while enolase is competitively inhibited. In the hL-PYK study, we discern a mechanism of allostery induced by the allosteric activator fructose-1,6-bisphosphate, and the inhibitor alanine. In the case of enolase, previous studies have attempted to make isoform-specific inhibitors of the enzyme; to further explore this objective, we compare the dynamics of two conserved isozymes, enolase 1 and 2, through MD simulations. As a proof of concept, we find compounds that discriminate between the two homologues by performing ensemble virtual screening on MD derived free enolase structures.

INDEX WORDS: Molecular Dynamics, Metabolic Regulation, Enolase, Pyruvate Kinase, Allostery, Virtual Screening

EXPLORING THE DYNAMICS OF GLYCOLYTIC REGULATION THROUGH TWO ENZYMES: HUMAN ENOLASE AND LIVER PYRUVATE KINASE

by

ABDULLAH SHOUAIB

A Thesis Submitted in Partial Fulfillment of the Requirements for the Degree of

Master of Science

in the College of Arts and Sciences

Georgia State University

2019

Copyright by
Abdullah Danish Shouaib
2019

EXPLORING THE DYNAMICS OF GLYCOLYTIC REGULATION THROUGH TWO ENZYMES: HUMAN ENOLASE AND LIVER PYRUVATE KINASE

by

ABDULLAH SHOUAIB

Committee Chair: Donald Hamelberg

Committee: Gregory Poon

Shahab Shamsi

Electronic Version Approved:

Office of Graduate Studies

College of Arts and Sciences

Georgia State University

December 2019

DEDICATION

This thesis is in dedication to my family and friends who have supported me along this difficult journey. I have no idea where I'd be without your support and guidance.

ACKNOWLEDGEMENTS

I would like to thank Dr. Hamelberg for granting me the opportunity to work in his lab and get involved in some amazing projects. Though the workload seemed heavy at times, I found myself enjoying even the most difficult aspects of research. The field of Computational Chemistry is one I never imagined I'd be a part of, and I find it absolutely wild to be a part of something as fascinating as this. I'd like to thank my colleagues for offering the support and guidance in times of difficulty, the level of talent and brilliance in this lab is inspiring. I would also like to thank the faculty and staff in the Chemistry Department. Their level of passion and interest will resonate with me for years to come.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

MD	Molecular Dynamics
hL-PYK	Human Liver Pyruvate Kinase
PEP	Phosphoenolpyruvate
FBP	Fructose-1,6-Bisphosphate
2-PG	2-Phsophoglycerate
RMSD	Root Mean Square Deviation
RMSF	Root Mean Square Fluctuation
PCA	Principal Component Analysis
dCNA	Difference Contact Network Analysis

1 INTRODUCTION

1.1 The Role of Conformational Dynamics in Protein Regulation

Cellular metabolism is the fundamental set of chemical reactions that distinguish living organisms from their counterparts. In order to maintain cellular homeostasis, the metabolic enzymes governing metabolism must have regulatory features that shift their catalytic function upon signal.[1] These regulatory mechanisms may vary: some enzymes undergo allosteric modulation, some are regulated directly from the active site, and others are regulated through post-translational modification.[2-4] What these processes all have in common are that they induce regulation by perturbing protein dynamics, and because these enzymes can sample a range of structural conformations, it is essential to understand their dynamics.[5] This study will focus on the dynamics of two metabolically coupled enzymes that have distinct regulatory features: enolase which undergoes inhibition through competitive binding in the active site and the allosteric enzyme human liver pyruvate kinase (hL-PYK).

The interest in understanding the dynamics of metabolic enzymes stems from its relevance to drug discovery and lead-optimization studies. A range of complex diseases that stem from metabolic pathways are still present due to the lack of effective therapeutics. One way to mitigate this issue is to explore structure-function relationships in drug targets (typically enzymes), and to then find compounds that regulate the enzyme through analysis of receptor-ligand interactions.[6-8] Over the years, rational structure-based drug discovery studies have proven beneficial in the clinical realm, with the most notable example being the now approved drug Dorzolamide, an inhibitor to carbonic anhydrase.[9] In fact, a significant number of approved drugs have been found through investigation of x-ray crystal structures of proteins.[10] Yet, as insightful as x-ray crystallographic and NMR data are, these methods are not without their

limitations. Aside from the relatively expensive costs in running these experiments, these techniques ultimately generate static models, and oftentimes neglect the structural dynamics of the protein-ligand complexes as a result.[11] Though cases may exist where molecules bind into a static pocket of an enzyme as detailed by the lock-and-key model, the consensus is that ligands instead bind to an ensemble of receptor conformations, and that upon binding, the ligand induces conformational changes that shift the dynamics to sample a different population.[12] For this reason, incorporating computational tools like molecular dynamics, which inexpensively predicts the intricate motions in macromolecular systems are in demand in drug discovery efforts.[13]

1.2 Purpose of this Study

Our aim in the first part of this study is to elucidate the differences in dynamics between two Enolase homologues: human enolase 1 and 2. Different tissues have varying expressions of enolase isozymes, all of which perform the same function, while specific tumor cells have only expressions of one enolase homologue.[14] As a result, attempts have been made in finding competitive inhibitors that are selective to one homologue.[15-16] The challenge is that these isozymes are highly conserved (>80% sequence similarity).[17] For this reason, we suggest that the structural dynamical differences between these homologues could aid in determining ligand selectivity. In this study, we compare the dynamics of enolase 1 and 2 using MD simulations. In addition, we run MD simulations of enolase 1 and 2 bound to HEX, an inhibitor specific to enolase 2. Finally, we run ensemble virtual screening on both homologues in an attempt to find compounds that could distinguish between the two highly conserved isozymes.

Aside from competitively inhibiting enzymes, there is a growing desirability in developing drugs that exhibit selectivity to allosteric sites of proteins.[18] Allosteric regulation is defined as the transference of dynamical information from a remote region of an enzyme to its catalytic site, thereby altering both the structure and function of the molecular body.[19] Traditionally, allosteric perturbations arise when a ligand or effector molecule binds to a target macromolecule; studies have extended the definition to include perturbations of distal residues.[20] Pyruvate kinase is often regarded as a textbook allosteric protein, for it houses multiple allosteric sites, each responsible for different effector molecules; and though the enzyme's allosteric regulation process is well characterized, the structural details remain elusive.[21] For the second part of this study, we focus on the enzyme human liver pyruvate kinase (hL-PYK). Our aim is to gain insight on the dynamics induced by two allosteric effectors: the activator fructose-1,6-bisphoshate (FBP) and the inhibitor alanine, by comparing MD simulations of hL-PYK in the presence and absence of these allosteric effectors.

1.3 Regulation of Glycolysis

Glycolysis is an ancient well characterized metabolic pathway that involves the breakdown of glucose to form pyruvate and ATP.[22] The significance of glycolysis is that many of its intermediates are essential to different aspects of cellular function, including the pentose phosphate pathway, citric acid cycle, etc.[23] The ten enzymes involved in glycolysis are complicated as well. For one, there can be multiple tissue-specific isozymes representing just one reaction step; these proteins can even be found performing functions outside of catalysis.[24] Because the net production of ATP from glycolysis is extremely low, normal functioning cells will shuttle pyruvate through oxidative phosphorylation.[23] When oxygen levels are low however,

the pyruvate formed from glycolysis is instead reduced to lactate (anaerobic glycolysis).[25] Muscle cells also undergo anaerobic glycolysis, and the lactate generated is normally shuttled out (Cori cycle) where it is converted into glucose by the liver; otherwise, abnormal accumulation of lactate can lead to a variety of issues.[26] Interestingly, tumor cells upregulate anaerobic glycolysis regardless of oxygen levels, leading to the production of lactate.[27-29] This phenomenon, dubbed the Warburg effect, is the reason that a variety of cancer therapeutic studies target glycolysis.

The two glycolytic enzymes in this study are metabolically coupled. Enolase catalyzes the dehydration of 2-phosphoglycerate (2-PG) to phosphoenolpyruvate, which is then converted to pyruvate by pyruvate kinase.[22] Apart from their metabolic proximity, both enzymes share a share a considerable homology.[30] In fact, comparison studies on yeast enolase and pyruvate kinase propose that the two enzymes evolved from a common ancestral protein.[30] The comparisons between both enzymes end here however, since they can be regulated differently. In this study, we individually explore the structural dynamics involved in regulating both enzymes.

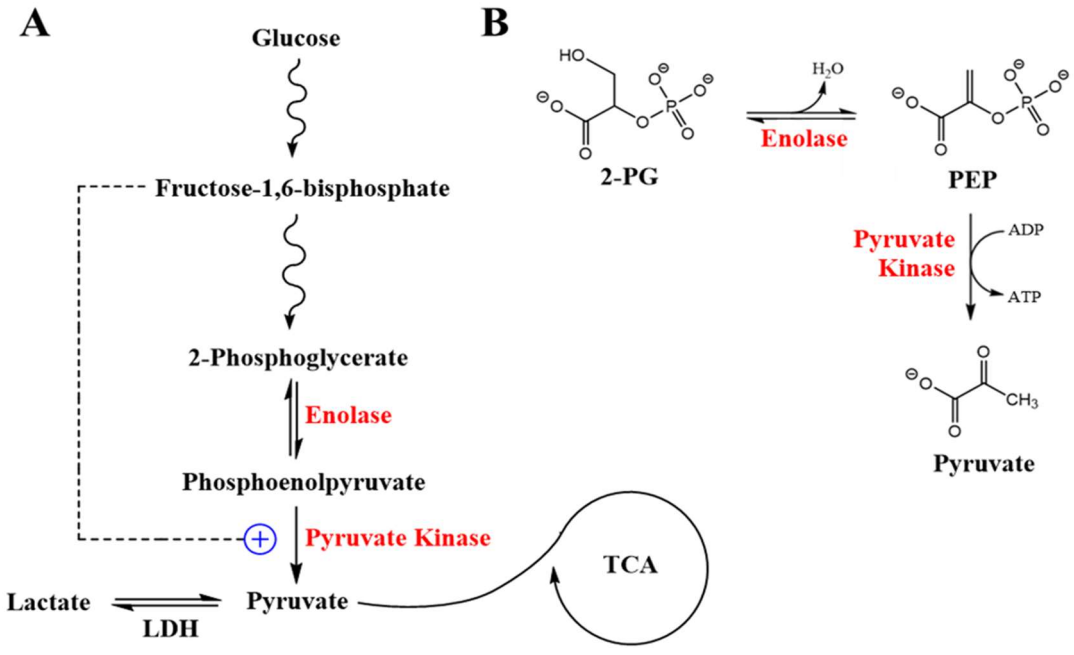


Figure 1.1 Glycolysis with an emphasis on its last two reactions.
(A) The glycolysis pathway. (B) The last two reactions of glycolysis.

2 Computational Chemistry

2.1 Molecular Dynamics

Enzymes and nucleic acids have dynamical characteristics that are oftentimes physically undetectable using conventional experiments.[31] The field of Computational Chemistry presents a solution to this barrier by accounting for the microscopic state ensemble using molecular dynamics, which allows one to study and analyze the structural features of a macromolecular system.[32] Over the years, MD has been used in a variety of investigations including protein folding, simulations of biomolecules, ligand binding studies, etc.[33-35]

Conceptually, two branches of molecular dynamics exist – quantum mechanical (QM) and classical MD. QM molecular dynamics is seen as more precise because it considers the contributions of valence shell electrons in the electrostatic interactions, while attributing classical integration with the nucleus and inner electrons.[36] Though this technique is more explicit, it's a computationally expensive tool as it requires a great deal of time and resources. This is especially the case when dealing with large-scale macromolecules. In contrast, the classical approach creates a coarse-grained model by ignoring the outer valence shell electron contributions, making the technique less computationally expensive.[37]

Classical MD is a statistical mechanics method that centers on Newton's Second Law of Motion, one of the foundational principles of classical mechanics.[38] For every MD simulation, an initial set of coordinates are required. These coordinates can come from x-ray crystallography, NMR solution, and even cryo-EM data. Considering the scope and size of the two enzymes in this study, we will use classical MD simulations to discern the dynamics, using x-ray crystallographic data as our starting coordinates.

2.1.1 Classical Molecular Dynamics

Classical molecular dynamics is based on Newton's Second Law of Motion, where the force (F_i) of a single atom is equal to the mass of the atom (m_i) times the acceleration of that atom (a_i). The acceleration is the derivative of velocity (v), and velocity is the derivative of displacement (r_i).[39]

$$F_i = m_i a_i = m_i \frac{dv}{dt} = m_i \frac{d^2 r_i(t)}{dt^2} \qquad (1)$$

The purpose of MD is to calculate the motions of a molecular system; but to find these motions, the forces acting upon each atom needs to be calculated.[40] Since the force is equal to the negative gradient of a scalar potential energy function (equation 2), one can take the potential energy of all interacting atoms $U(r_1,\ldots,r_N)$ as a function of their positions.[40] Then, the gradient (∇) of that function with respect to the displacement of the atoms can be taken. Finally, the force acting on every atom can be calculated.

$$F_i = -\nabla_{ri} U(r_1, \ldots, r_n) = -\left(\frac{\partial U}{\partial x_i}, \frac{\partial U}{\partial y_i}, \frac{\partial U}{\partial z_i}\right) \qquad (2)$$

The potential of each individual atom is governed by a set of bonded and nonbonded potential terms: the bond valence, length, and dihedral angle potentials, as well as the nonbonded van der Waals and electrostatic potentials.[39] The first potential in equation 3 describes the harmonic vibrational motions between atom pairs that are covalently bonded, where l_i is the bond length and l_{i0} is the equilibrium potential. The next potential, the angular potential, describes the angular vibrational motion that occurs between bonded atoms, where b_i is the harmonic force constant, where θ_i is the bond angle, and θ_{i0} is the equilibrium potential. The third potential in equation 3, the torsion angle potential, describes the dihedral motions, where four consecutively linked atoms (two sets of three atoms, with two in common) are accounted for; the c_i term

represents heights of rotational barriers, and *n* denotes periodicity. The next two potential terms describe nonbonded potentials between atom pairs i and j. The van der Waals potential, which accounts for interatomic forces between unpaired atoms, is modeled by the Lennard-Jones 6-12 potential. In this term, *r* is the distance between the two atoms, σ_{ij} is the distance where the intermolecular potential between the atom pairs is zero, and ε is the well depth. Lastly, the electrostatic potential, modelled by Coulomb's Law, describes the attractive and repulsive forces that occur from charged atoms. In this term, the charges are denoted by q_i and q_j, while the distance between the two atoms is denoted by r_{ij}.

$$U(r_1, \ldots, r_n) = \sum_{bonds} \frac{a_i}{2}(l_i - l_{i0})^2 + \sum_{angles} \frac{b_i}{2}(\theta_i - \theta_{i0})^2 + \sum_{torsions} \frac{c_i}{2}(1 + \cos(nw_i - \gamma_i)) + \sum_{atoms\ i,j} 4\varepsilon_{ij}\left[\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}}\right)^{6}\right] + \sum_{atoms\ i,j} k\frac{q_i q_j}{r_{ij}} \qquad (3)$$

Upon calculations of the potential terms, the force vectors of each atom can be evaluated and used to compute their motions. The Verlet integration algorithm (equation 4) is used to calculate the trajectories of the individual atoms by their position and velocity vectors, with respect to time.[41]

$$r_i(t + \Delta t) \cong 2r_i(t) - r_i(t - \Delta t) + \frac{F_i(t)}{m_i}\Delta t^2 \qquad (4)$$

The Verlet algorithm relies on initial atomic coordinates and a proper choice of a timestep. Conventionally, the initial atomic coordinates are taken from x-ray crystallographic or NMR data. As for the choice of a timestep, a femtosecond timestep is considered ideal; one must be aware that choosing a timestep too small can lead to unnecessarily small atomic motions, while drastically increasing it could give rise to inaccurate motions.[42]

2.1.2 Current State and Outlook of MD

Molecular dynamics studies have proven useful in unearthing a multitude of biologically relevant phenomena. One successful example of MD applied to drug discovery involved a study of Eph tyrosine kinases, where a snapshot representative of the overall conformation of the simulation trajectories were taken and used as a receptor for high-throughput screening.[43] This was proven useful, given that the x-ray crystallographic structures of the enzymes did not sample the conformations necessary for adequate docking. MD studies have also successfully led to the finding of allosteric sites, otherwise undetectable experimentally.[44] Despite the number of successes, there are two challenges in the field of MD that should be addressed: the atomic force fields used in simulations need further improvements, and technical limitations hinder us from procedurally running simulations at more robust timescales.[45]

Regarding the first limitation, the force fields used in molecular dynamics simulations are only approximations of quantum mechanical effects.[39] Considering that classical MD has proven reliability in predicting correct molecular motions for macromolecules, this is not always a hinderance; however, the disregard for QM effects can still be a detriment to studies that involve transition metals and ligand binding events.[46] In classical MD, atoms are assigned fixed partial charges; but in reality, electron clouds move, and so the charge of each atom fluctuates based on the surrounding dynamical environment. Of course, this limitation can be marginally mitigated by implementing QM calculations, but such computations are rigorous and expensive. There is a compromise: computational studies can incorporate hybrid techniques, where QM calculations are incorporated in a select area of a macromolecule, such as the active site of an enzyme or charge calculations of a coordinating transition metal, while the rest of the molecular body is governed by classical mechanics.[48]

Another significant challenge in MD, one that is more pressing, is the issue of sampling. Across the biological world, it is known that a variety of macromolecules can explore multiple conformational ensembles over long periods of time.[49] Unfortunately, MD simulations are currently limited to microsecond timescales (with notable exceptions), and some important dynamical motions can even take form after hours.[50] There are proposed solutions to sampling, however. One such solution is accelerated molecular dynamics (aMD), which involves the reduction of large energy barriers, enabling proteins to sample conformational states that would otherwise be inaccessible from conventional simulations.[51] Though artifacts could arise from this method, it is nonetheless a feasible solution. Other solutions include replica exchange molecular dynamics (REMD) and high temperature simulations, both of which have proven to be effective.[52]

The biggest driving force towards the progress in MD is of course the hardware. As computational power expands, the ability to simulate larger systems over longer periods of time increases. Where MD simulations were once calculated on CPUs, todays software packages like the updated AMBER suite are designed to run MD calculations on GPUs, enhancing the speed of MD computations exponentially in comparison. Also, Nvidia manages to release new GPUs relatively quickly and this leads to significant increases in computational power. For example, the hL-PYK simulations in this study comprises of over 200,000 atoms, and the simulation ran for ~16 nanoseconds per day on a NVIDIA GTX 980. When the same system ran on an upgraded GTX 1080, the simulation speed increased to approximately 25 ns/day. Ultimately, the progress in computational power allows for increase in simulation timescales, and one can soon look to running routine microsecond-long simulations.

2.2 Methods of Analyzing MD Simulations

2.2.1 Root Mean Square Deviation

In Computational Chemistry, the root mean square deviation (RMSD) method measures the average distances between atoms of molecules superimposed onto one another.[53] Specific to MD simulations, the RMSD calculations are generally performed on the backbone atoms of a set of snapshots acquired from a trajectory; these snapshots are superimposed onto the reference structure, which is usually the starting coordinates of the simulation. When it applies to these studies, RMSD calculations are useful in determining whether a simulation is equilibrated, which is when a structure samples a stable average conformation. Equation 4 defines RMSD, where N denotes the number of snapshots/frames of a trajectory, and denotes the distance between atom and its reference coordinates:

$$RMSD = \sqrt{\frac{1}{N}\sum_{i=1}^{N} r_i{}^2} \quad (5)$$

2.2.2 Root Mean Square Fluctuation

The root mean square fluctuations (RMSF) is the measure of displacement of a set of atoms with respect to the reference structure, averaged over the number of atoms involved.[54] Like the RMSD method, RMSF calculations rely on alignment of backbone atoms onto a reference structure. Unlike RMSD however, RMSF calculations can be useful for analyzing an equilibrated trajectory (where the reference structure is an equilibrated snapshot), because it provides insights on the fluctuations of residues of an enzyme with respect to a sample of stable snapshots.

2.2.3 Principal Component Analysis

Principal component analysis (PCA) is a statistical technique that reduces the number of dimensions of a dataset to describe the most important elements of that dataset.[55] While the datasets used in PCA may have linearly correlated variables, the technique extracts the information into a set of linearly uncorrelated variables referred to as principal components (PC). In the context of molecular dynamics simulations, PCA is useful when applied to the cartesian coordinates of a given simulation. When applying the technique to the coordinates of the backbone atoms in a MD trajectory, the most dominant motions of the molecular system can be captured, and generally those motions are considered the most essential.[56] PCA can also be applied to all heavy atoms of a set of essential residues, such as binding site residues.

2.2.4 Difference Contact Statistics & Network Analysis

To capture dynamical information at the atomic level, one can observe inter-residue contact formations and breakages of a molecular system. Sometimes referred to as contact statistics, the technique considers all heavy atoms of each residue, and measures the distances between both residues.[57] If the distance between the two residues is less than 4.5 Å, then the two residues are in contact; if none of the atoms are within the cutoff, then there is no contact. Once the contact probabilities are mapped for each simulation independently, one can measure the residue-residue contact differences between two simulations. The technique in turn can provide a network of residue contact differences between the compared simulations.[58] The method has been used successfully to detail allosteric mechanisms in proteins, perturbations caused by substrate recognition, etc.[59-61] In this study, contact statistics is applied to both enzymatic systems to capture the perturbations caused by ligand recognition.

2.3 Molecular Docking

Molecular docking is an essential computational tool in predicting intermolecular interactions between two molecules.[62] Conventional docking studies involve searching a designated three-dimensional space within a receptor (target molecule), predicting the binding modes of a ligand through stochastic search methods, and using a proper scoring function to calculate the binding affinities (ΔG) of the ligand poses.[63] The results can garner a list of ligand poses ranked in accordance to the most favorable protein-ligand interactions. The input data is arguably the most important aspect. Factors such as receptor conformation(s), search space, and source of small-molecule structures must all be considered prior to performing docking. Though this technique has proven effective in a variety of structure-based studies, results from docking will often trade-off reliability for time efficiency.[64] For this reason, docking is recommended solely for early stage lead-optimization, where the goal is not to find a potential drug, but to instead use docking results as a rationale for further drug design endeavors. The low computational cost comes incredibly handy however, since the technique can be scaled up to screen/dock thousands (even millions) of compounds.

In this study, the software package Autodock Vina[65] is used to virtually screen four-hundred thousand drug-like molecules against the enolase 1 and 2 homologues. For the purpose of this study, a drug-like molecule is defined as any chemical species that fits the Lipinski's Rule of Five.[66] The structures of the molecules were acquired from the ZINC Database[67], while the receptor coordinates were obtained from equilibrated MD simulations of free enolase 1 and 2. Because the interactions between molecules are dynamic in nature, it's not always ideal to dock ligands onto a single static structure (as the lock and key model entails), but to instead dock ligands onto an ensemble of structures, a method referred to as ensemble docking.[68]

3 The Dynamics of Enolase 1 & 2 Regulation

3.1 Introduction

Enolase is a metalloenzyme that catalyzes the conversion of 2-phosphoglycerate to form phosphoenolpyruvate, the penultimate metabolite in glycolysis.[69] Categorized in the lyase class of enzymes, enolase catalyzes both the reverse hydrolysis (gluconeogenesis) and forward dehydration (glycolysis) reactions, making it crucial for metabolic function. In vertebrates, three glycolytic enolase isoforms (α, β, and γ) exist, and though each homologue is encoded by a different gene, they are all highly conserved.[70] Evolutionary studies suggest that these homologues are a result of a gene duplication event traced over 300 million years ago.[71] Though these genes have varying expressions depending on the tissue cells, they all perform the same function, making enolase a functionally redundant enzyme in cells that express multiple isoforms; this genetic redundancy is exploitable.

Certain tumor cells like Glioblastoma and Hepatocellular Carcinoma have deletions of the 1p36 tumor suppressor locus, which contains the ENO1 gene, meaning that these cells function without α-enolase (enolase 1).[14] These tumor cells can still perform glycolytic function through redundant action of γ-enolase (enolase 2). In contrast, many functioning cells express both the ENO1 and ENO2 genes and can use both for glycolysis. Studies show that by selectively inhibiting Enolase 2, the ENO1-deleted tumor cells would lose glycolytic function, while the remaining cells can rely on enolase 1 for metabolism.[15] This phenomenon of targeting cells based on gene deficiency is known as collateral lethality. The challenge that comes in targeting enolase 2, however, is that the homologues share a conserved active site; in fact, the enzymes share an 83% sequence identity. It is therefore proposed that comparing the dynamics of enolase 1 and 2 can aide in discerning the structural motions that explain ligand specificity.

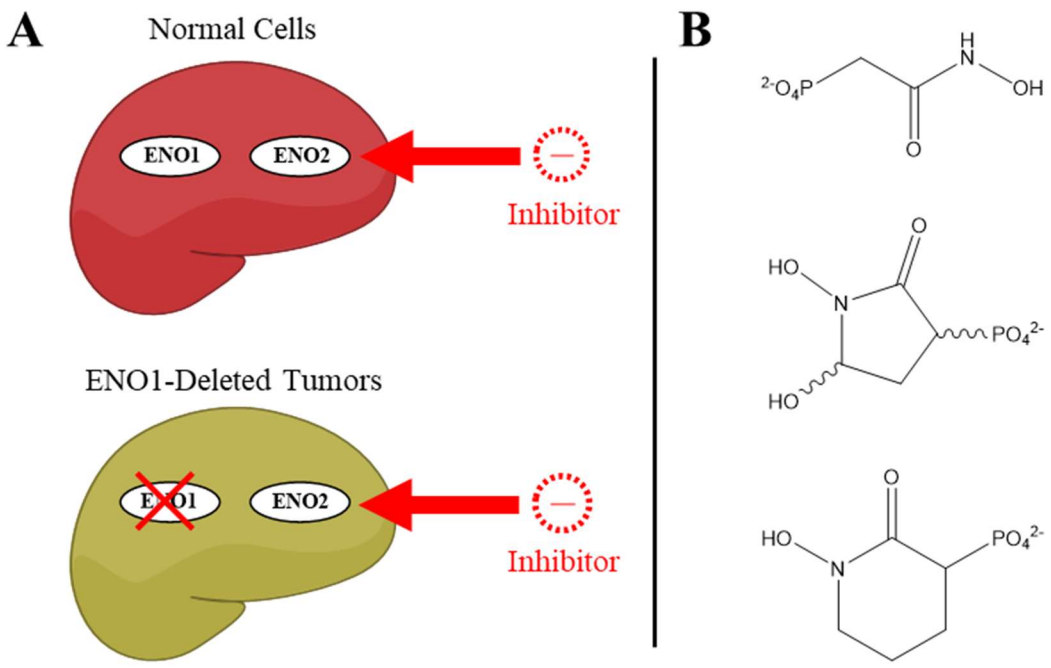


Figure 3.1 Concept of collateral lethality in the context of enolase 1 & 2.
(A) Normal functioning cells can use both isozymes for glycolytic function while specific tumor cells will have ENO1 gene deletion. (B) Compounds that inhibit enolase 2. (Top) Phosphonoacetohydroxamate (PhAH). (Middle) SF2312. (Bottom) HEX.

Recently, a natural phosphono-hydroxamate, SF2312, was found to exhibit selective toxicity towards glioblastoma cells with ENO1 gene deletion.[15] The compound was found to bind more selectively to enolase 2 than its homologue. The molecule was further developed into a prodrug (POMSF), which was then tested *in vivo*.[16] Unfortunately, the inhibitor induced hemolytic anemia in mouse models because of on-target activity against enolase 1 in red-blood cells. In fact, red blood cells have only enolase 1 expression, and so it was determined that the compound was not selective enough towards enolase 2. In response to this problem, a new lead-compound was developed, called POMHEX, which is shown to not produce hemolytic anemia. The HEX-enolase 2 complex was co-crystallized, and the structure has been published in the Protein Data Bank.

The first goal in this study is to discern the differences in dynamics between enolase 1 and 2 through MD simulations in apo-state; in addition, we run MD simulations of the two isozymes complexed with HEX, an inhibitor selective to enolase 2. The second, and perhaps more ambitious goal of this study, is to perform ensemble virtual screening on both homologues. It is important to stress that, like any theoretical technique, molecular docking has its limitations. Where most docking studies attempt to find "hit" molecules with high binding affinities (ΔG) to a target receptor, we attempt to instead compare predicted interactions between a set of ligands and receptors (free enolase 1 and 2). By performing molecular docking of approximately 400,000 small-molecules, one can take the differences in binding affinities (ΔΔG) between both isozymes and compare the resultant protein-ligand interactions.

3.1.1 Structural Analysis of Enolase 1 & 2

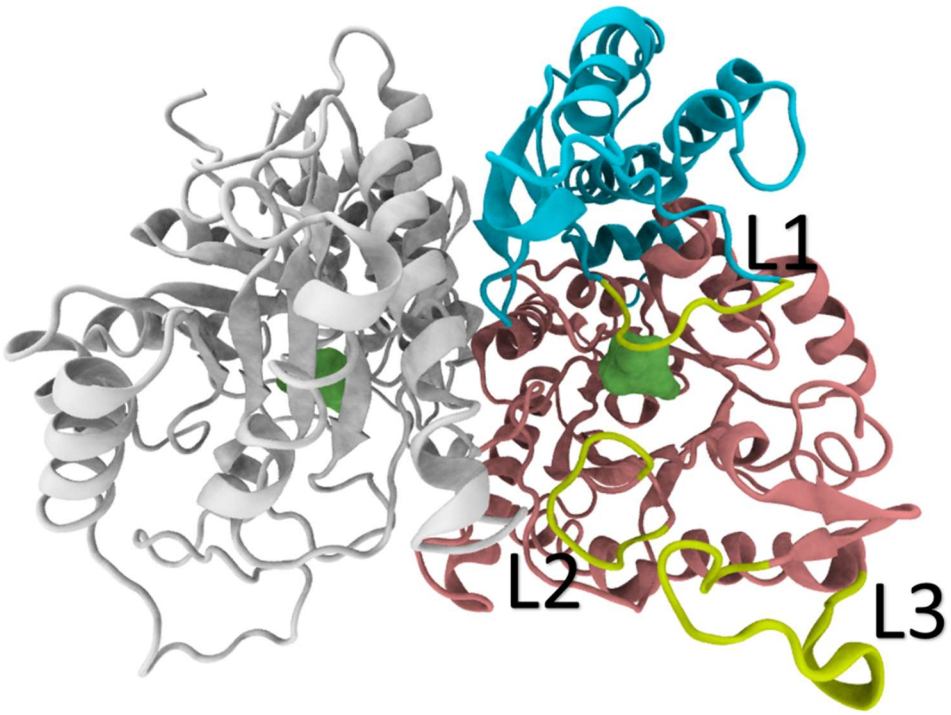


Figure 3.2 Structure of enolase 2 (PDB ID: 4ZCW).
Chain A is color coded: N-terminus domain is cyan, catalytic C-domain is pink, and the loops L1, L2, and L3 are highlighted in yellow. The active site is in green space-fill.

Enolase is a homodimer comprised of two domains: An N-terminus "lid" domain, and a catalytic C-terminus domain comprised of a TIM-barrel fold (eight α-helices and eight β-strands). The enzyme relies on two divalent magnesium cations for function.[72] Upon binding of the first magnesium cation, enolase undergoes a conformation change that "opens" up the active site, allowing for substrate binding.[73] Once a ligand enters the catalytic site and the binding is recognized, a flexible loop L1 (residues 36-43) from the N-terminus domain encapsulates the active site; in addition, two more loops from the TIM-barrel domain, L2 (residues 153-169) and L3 (residues 251-277), also form a "closed conformation" by partially encasing the active site.[74] For the activity of the enzyme to proceed, both the ligand and a set of carboxy residues in the active site must loosely coordinate with a second magnesium cation.[75] Crystallographic studies have shown that competitive inhibitors of enolase can elicit the same L1 loop enclosure that occurs upon substrate binding, but at varying levels.[76]

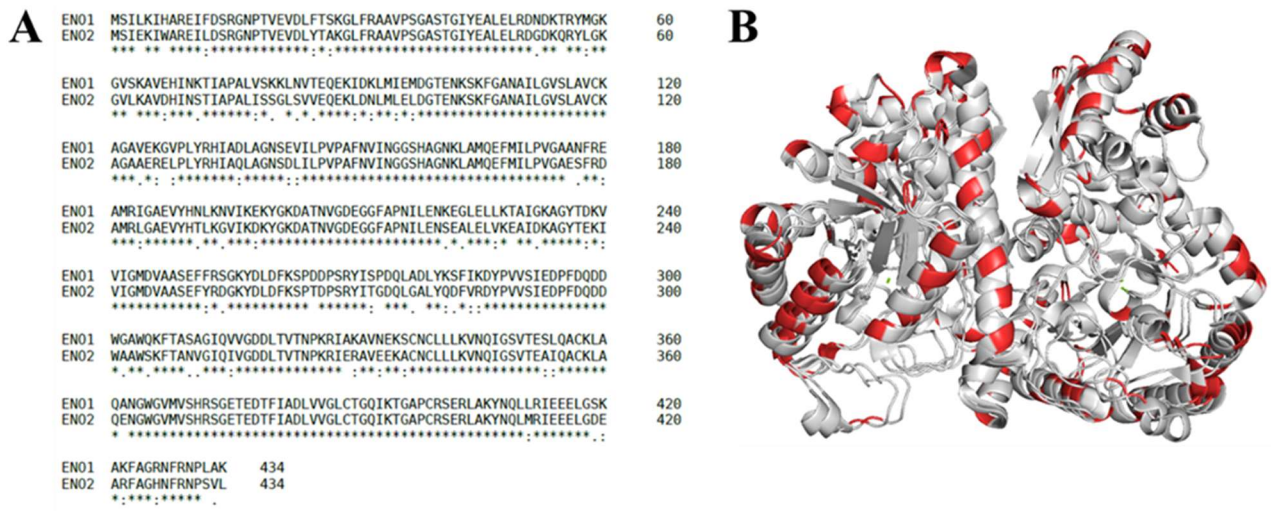


Figure 3.3 Alignment of enolase 1 and 2.
(A) Sequence alignment of enolase 1 and 2 using ClustalW2. (B) Structural superposition of enolase 2 (PDB 4ZCW)[15] and enolase 1 (PDB 2PSN). The residues marked in red are non-conserved, while the residues in cyan are conserved.

Alignment of enolase 1 and 2 reveals that both proteins are highly conserved. Apart from their 83% sequence identity, both isozymes retain the same secondary structural elements. Usually, mechanisms for ligand specificity can be explained by differences in the active site residues; but the catalytic site for both homologues are conserved, which suggests that a set of distal residues could be responsible for the differences between the two isoforms. Based on this observation, it is proposed that by comparing the dynamics of enolase 1 and 2, one can rationalize the atomic-level motions that are responsible for the differences between the two homologues – differences that could be exploited for drug discovery studies.

3.2 Experimental Procedures

3.2.1 Preparation of Enolase 1 & 2 for MD

Four systems were generated for this study: enolase 1 and 2 in apo form, and enolase 1 and 2 bound to the HEX inhibitor. The x-ray crystal structure PDB 5TD9[16] was used for the apo enolase 2 model, while PDB 3B97[74] was used for the apo enolase 1 model. The histidine protonation states were determined by the web-based tool PROPKA (PDB2PQR web server).[77] For consistency, conserved histidine residues between both isoforms were kept with the same protonation states. To generate the bound enolase models, the ligand coordinates were obtained from PDB 5IDZ[16], an x-ray crystal structure of enolase 2 bound to HEX. The PyMOL[78] software was used to structurally superimpose this structure to enolase 1 (PDB 3B97), where the coordinates of HEX were then transferred onto the active sites of the enolase 1 monomers. The resultant PDB structures were ported onto xLeap to verify fidelity of the clipping. The GAFF2 force field was used to model the HEX ligand.[79] MD simulations of enolase 1 and 2 both bound to HEX were performed based on these initial coordinates.

3.2.2 MD Simulation of Enolase 1 & 2

The MD simulations were carried out using the CUDA version of the Amber16 suite of programs[79] along with the modified version of the Cornell et al. ff14SB[80] force field, while the GAFF2[79] force field was used to parameterize the HEX ligand. The simulations were performed on the NVIDIA GeForce GTX 980 GPU. Each system was solvated in a TIP3P[81] periodic octahedron box containing water molecules. Counterions were added to the system to obtain electrostatic neutrality. All simulations were kept at a temperature of 300 K and a constant pressure of 1 bar. Electrostatic interactions were calculated using the particle mesh Ewald (PME) summation method, where long-range nonbonded interactions were accounted for with a 9 Å cutoff.[82] For each system a 2 fs time step was used to solve the equation of motion. After a series of minimizations, the system was equilibrated with a harmonic constraint through five subsequent MD simulations of 1 ns each, where the force constant was reduced from 500 to 5 kcal·Å^2/mole. A final equilibration step ran for 1 ns, where no harmonic constraints were conducted, allowing for the enzyme to move freely. A total of 4.4 μs of simulation data was produced, and the last microsecond from each trajectory was used for analysis. Aside from PyMOL, images were also generated using the VMD[83] software package.

Table 3.1 Summary of MD simulations performed on enolase.

	Simulations	Ligands	Simulation Time
1	Free Enolase 1	2 Mg^{2+}	1.1 μs
2	Free Enolase 2	2 Mg^{2+}	1.1 μs
3	Bound Enolase 1	2 HEX, 2 Mg^{2+}	1.1 μs
4	Bound Enolase 2	2 HEX, 2 Mg^{2+}	1.1 μs

3.2.3 Virtual Screening of Enolase

Virtual screening was performed with the Autodock Vina[65] software. Using the ZINC Database[67], approximately 400,000 compounds following Lipinski's Rule of Five criteria[66] were obtained for virtual screening. The files were already in "pdbqt" format, meaning the Gasteiger charges were already accounted for in each ligand. The method for obtaining proper coordinates of the receptors was a bit more extensive. As mentioned previously, ligand binding events are ultimately dynamic in nature, so an ensemble of receptor structures were obtained for the virtual screening.

To distill the dominant configurations from the MD simulations of apo enolase, PCA was performed on heavy atoms of the active sites of enolase 1 and 2. PC1 and PC2 were plotted, and the probability density distributions were calculated. Four representative snapshots, one at the center of each cluster, was chosen for docking. In total, the four most dominant snapshots from apo enolase 1 and 2 were used as receptors for virtual screening against 400,000 compounds, meaning over 2.8 million docking runs were simulated. AutoGrid version 4.0 was used to generate a periodic box of 13x13x13 Å (1.000 Å spacing) to encapsulate the catalytic site. A binding affinity/score (ΔG) was obtained for each conformation. The virtual screening was performed on a Linux Beowulf cluster of ~270 cores. Two-dimensional docking pose images were generated using PoseView.[84]

3.3 Results & Discussion

3.3.1 The Dynamics of Free Enolase 1 & 2

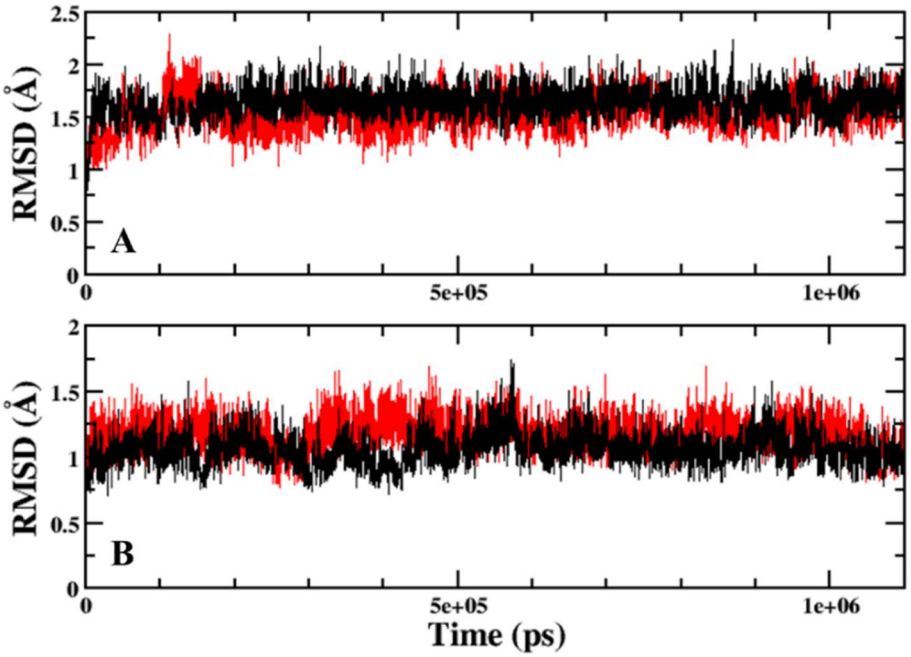


Figure 3.4 RMSD of the backbone atoms of free enolase 1 & 2.
(A) RMSD of free enolase 1 simulation. Chain A is in black, and chain B is in red. (B) RMSD of free enolase 2. Chain A is in black, and chain B is in red.

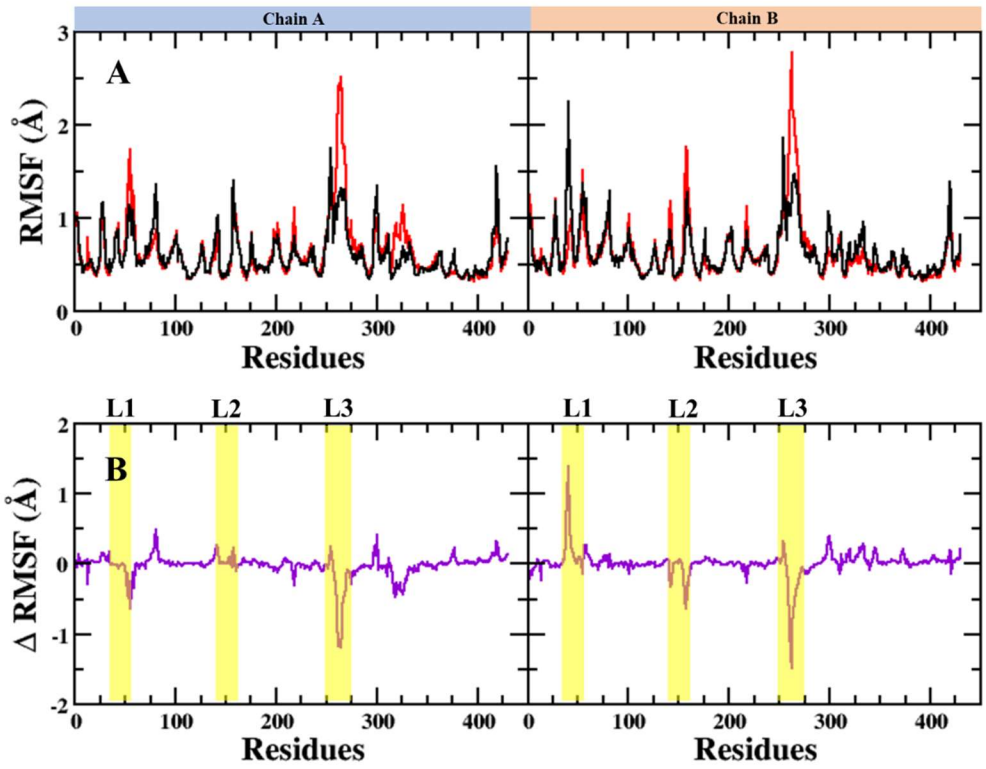


Figure 3.5 RMSF of the backbone atoms of free enolase 1 & 2.
(A) RMSF of free enolase 1 (black) and enolase 2 (red). (B) Difference in RMSF of free enolase 1 and 2 (taken by subtracting the RMSF of enolase 1 to enolase 2).

The RMSD of the MD simulations of enolase 1 and 2 reveals that the structures reach stabilization after the first 100 nanoseconds. Aside from the inferred convergence, it appears that the trajectory of enolase 1 deviates from its starting coordinates more-so than enolase 2. Interestingly, the RMSF of both trajectories show that despite the high sequence and structural homology between both isozymes, their dynamics differ greatly. Enolase 2 appears to fluctuate more than its counterpart. The oscillations of loops L1 and L2, which are highly conserved, differ greatly in enolase 1 and 2. The L3 dynamics fluctuate in both simulations as well. The set of carboxy residues responsible for binding of the second magnesium ion also differ in their fluctuations.

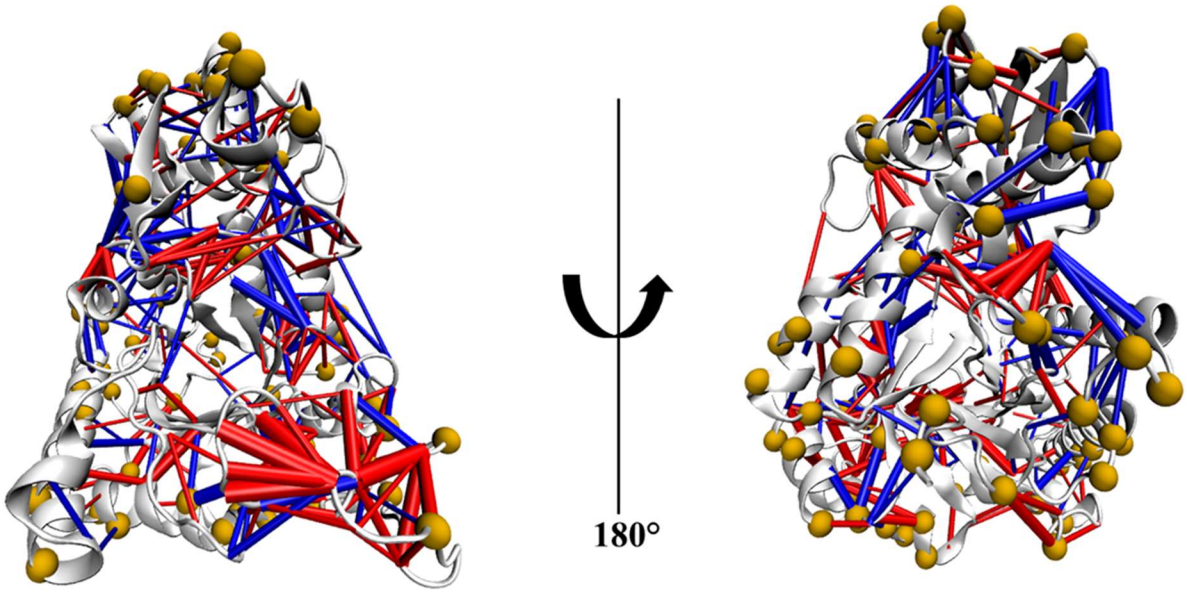


Figure 3.6 Residue−residue dCNA of free enolase 1 and 2.
Difference contact network is **c**alculated by comparing the free/apo enolase 1 to free enolase 2. Blue bars indicate more contact formations from the apo to a substrate-bound network, while red bars indicate more contact breakages. The non-conserved residues are in orange orbs.

Difference residue-residue contact statistics reveal that the two enolase homologues have considerable differences in their dynamics, despite being structurally conserved. A close inspection of the contact network shows that L2 can undergo a great deal of conformational changes. Three residues in this loop are non-conserved: residues 273, 274, and 265. In the enolase 1 trajectory, L2 interacts closely with the enzymatic body. In contrast, L2 is expanded away in the enolase 2 trajectory, resulting in contact breakages. At the active site, the carboxy residues known to coordinate with the second magnesium cation (absent in free enolase) display differences in their contact network as well. This could suggest that prior to ligand binding, enolase 1 and 2 differ in their dynamics, and that these differences in dynamics could influence ligand selectivity.

3.3.2 Dynamics Induced by HEX, an Enolase 2 Specific Inhibitor

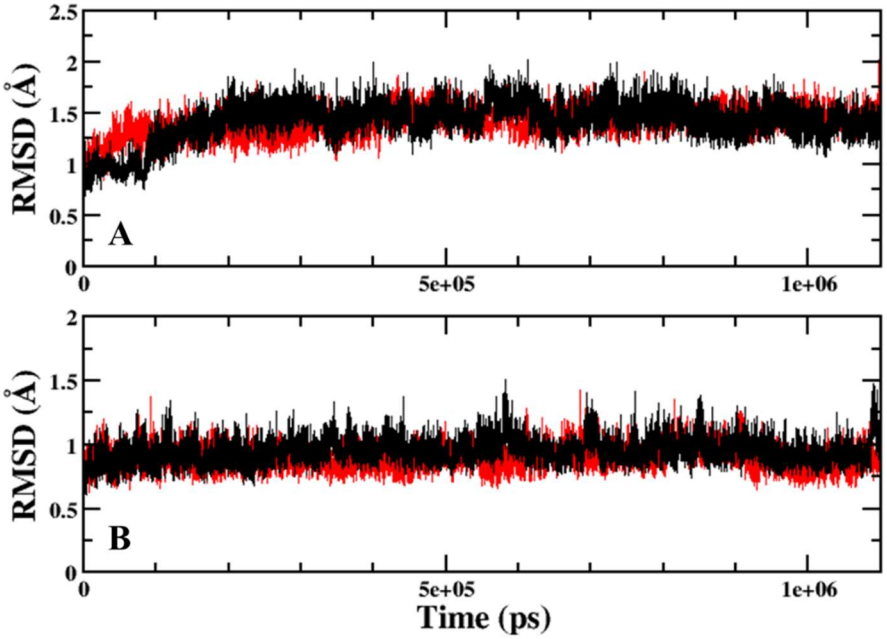


Figure 3.7 RMSD of the backbone atoms of HEX-enolase complex.
(A) RMSD of HEX-bound enolase 1. (B) RMSD of HEX-bound enolase 2. Chain A is in black, and chain B is in red.

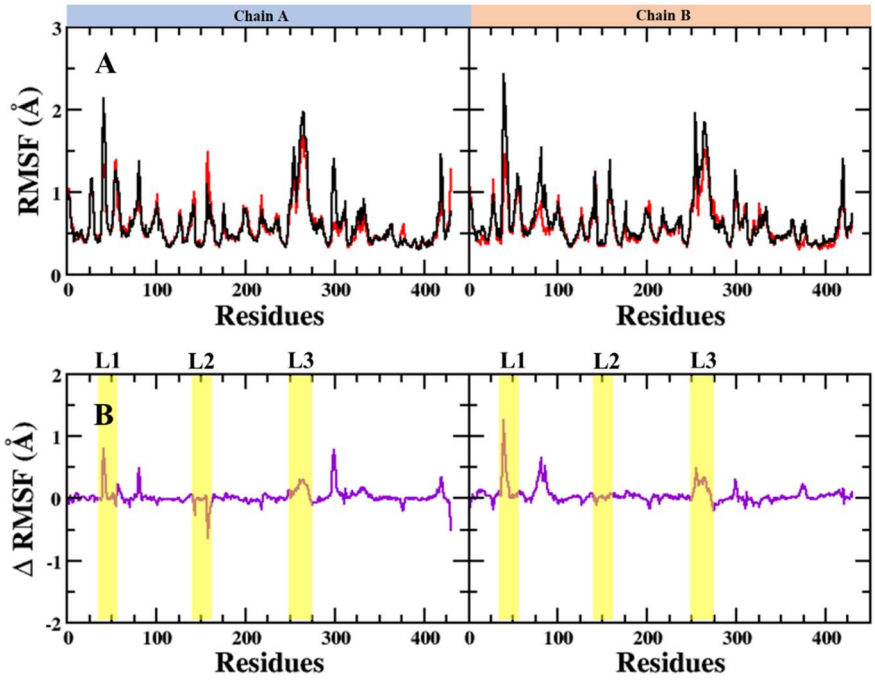


Figure 3.8 RMSF of the backbone atoms of HEX-enolase complex.
(A) RMSF of both HEX-enolase 1 (black) and HEX-enolase 2 (red). (B) Difference in RMSF between enolase 1 and 2.

RMSD reveals that the hex-bound enolase simulations stabilize after 100 nanoseconds. Like the apo variants, enolase 1 deviates from its starting coordinates more-so than enolase 2. Interestingly, the HEX-enolase 1 complex appears to display greater dynamic fluctuations relative to enolase 2. As shown earlier, free enolase 2 fluctuates more so than free enolase 1, but upon binding of HEX, enolase 1 appears to oscillate more than its homologue. This could suggest that the ligand stabilizes enolase 2 to fit a more localized conformation.

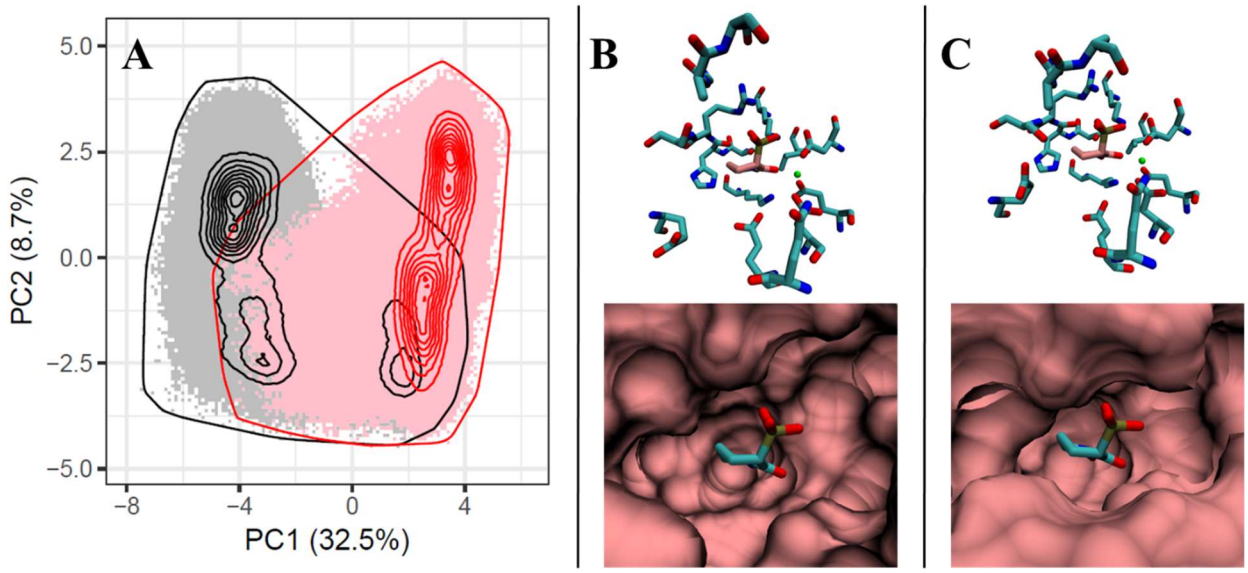


Figure 3.9 PCA of HEX-bound active site.
(A) Plot of PC1 and PC2 with calculated probability density distributions. Enolase 1 is in black, and enolase 2 is in red. (B) Active site of HEX-enolase 1 sampled from the center of the cluster, with the binding pocket at the bottom (surf mode). (C) Active site of HEX-enolase 2 sampled in the center of the cluster, with the binding pocket shown at the bottom (surf mode).

PCA of the catalytic site reveals notable differences between how enolase 1 and 2 recognize HEX binding. The side-chain dynamics of the residues interacting with HEX appear to adopt the same conformation, indicating that HEX induces similar dynamical effects on the interacting residues of both homologues. However, the residues in close proximity to HEX (the

ones not directly in contact) sample different motions. PC1 reveals that the key difference is in the interaction between residues S373 and E210. In HEX-enolase 2, a hydrogen bond interaction is formed in the active site. In the case of HEX-enolase 1, the hydrogen bond isn't formed; also, the L1 loop is slightly sampled away from the ligand. Overall, holo-enolase 1 appears to have a larger catalytic pocket, suggesting that HEX does not exhibit the same effects on the enzyme compared to enolase 2.

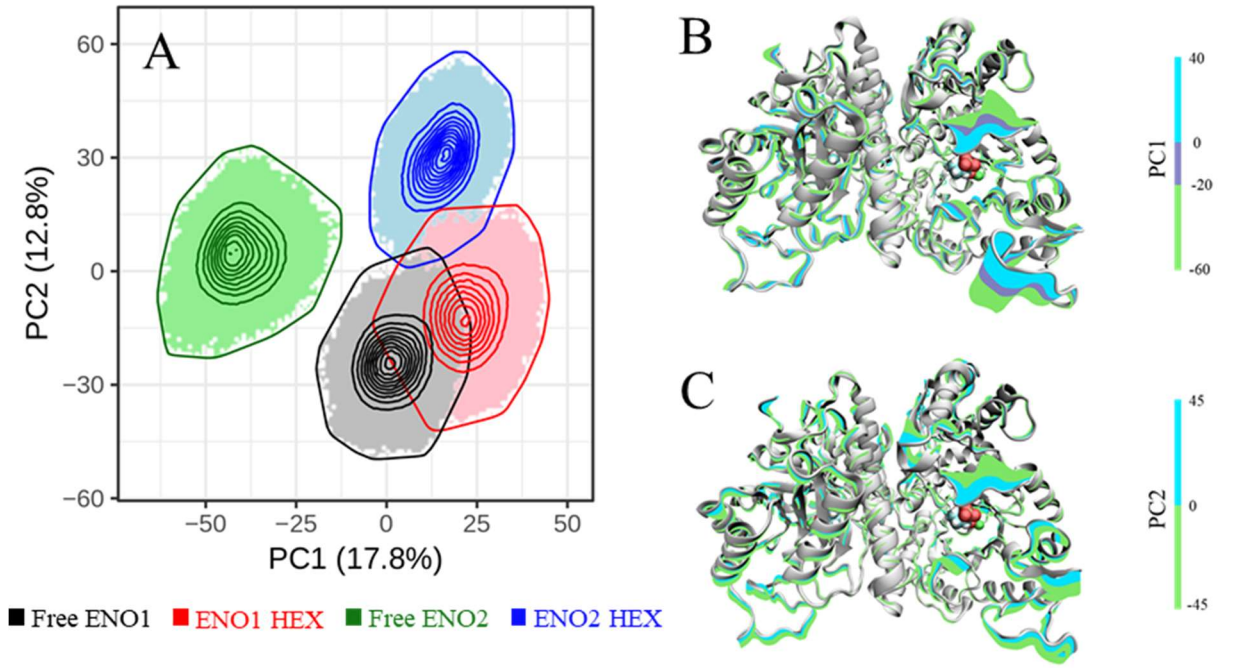


Figure 3.10 PCA of the backbone atoms of the enolase simulations.
(A) Plot of PC1 and PC2 of the four simulations. (B) PC 1 projected onto the backbone atoms of the enolase dimer with respective color scales. (C) PC 2 projected onto the backbone atoms of the enolase dimer with respective color scales.

Principal Component Analysis (PCA) of the backbone atomic coordinates of the enolase simulations reveal that each protein samples distinct conformations. Here, the free and HEX-bound enolase simulations appear to have no overlap, indicating that they sample significant differences in conformation space. This is a stark contrast to free and boud enolase 1, where

there is considerable overlap in PC1 and PC2. This could suggest that HEX elicits the greatest conformational changes to enolase 2, while enolase 1 recognition to HEX is more limited in comparison. PCA also reveals that despite the large degree of homology, the dominant backbone motions differ significantly.

Projection of the principals eignevectors can display the representative backbone motions of the MD simulations. Loops L1 and L3 appear to sample the greatest motions; correlating with previous crystallographic studies, it appears that these loops sample a more “open” conformation in the free enolase simulations, while HEX binding is accompanied by a “closed” conformation in both homologues. The backbone dynamics of L2 are more subtle; in the free enolase 1 and 2 trajectories, L2 fluctuates between an open and closed conformation. In the HEX-bound simulations, this loop slightly projects more outward (away from the catalytic site).

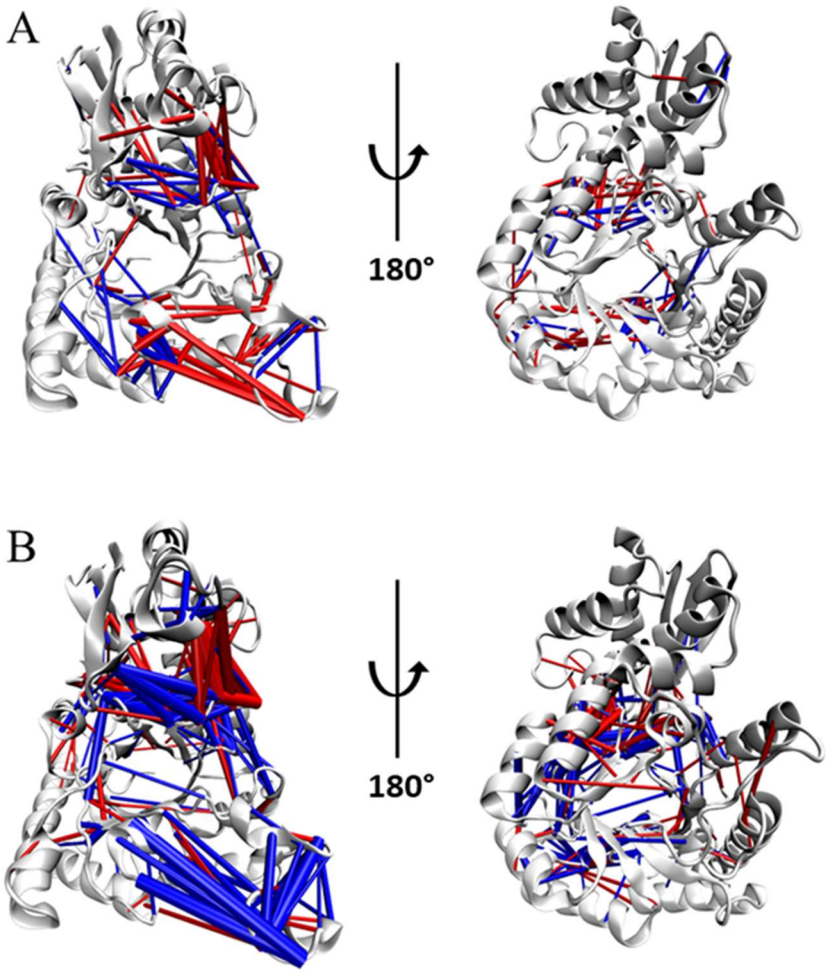


Figure 3.11 Dynamic contact statistics of HEX-enolase 1 and 2.
A) Residue−residue difference contact networks are calculated by comparing the free enolase 1 to the bound enolase 1. (B) Residue-residue difference contact networks comparing free enolase 2 to bound enolase 2. Blue bars indicate more contact formations from the apo to a substrate-bound network, and red bars indicate more contact breakages.

The difference contact statistics of both enolase homologues show that HEX induces greater dynamical changes to enolase 2. The HEX-bound enolase 2 system appears to sample a more "closed" conformation, where L1 and L3 move closer to the active site, forming a large degree of contacts with the rest of the enzyme as a result. In contrast, L2 in bound enolase 1 has L2 form less contacts with the enzyme as opposed to free enolase 1. The L1 loop does appear to enclose the active site upon HEX binding, but the degree to which these changes are broadcasted is much smaller in comparison to the enolase 2 system.

3.3.3 Ensemble Virtual Screening of Enolase 1 & 2

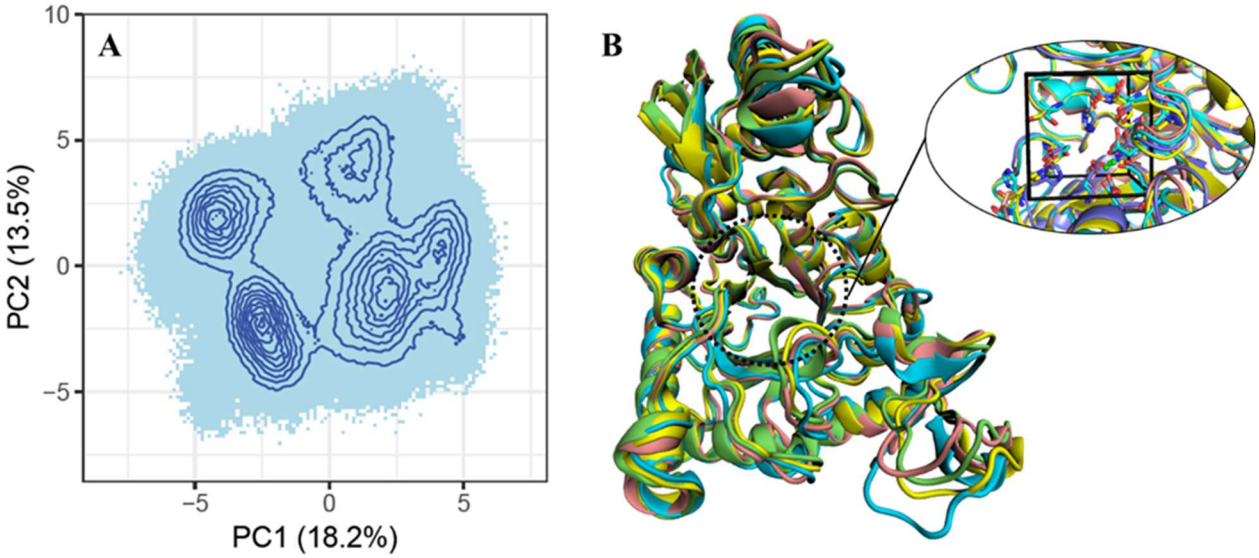


Figure 3.12 Ensemble selection scheme of enolase 2.
(A) PC1 vs PC2 of the active site residues of enolase 2 reveal distinct clusters. (B) The four representative structures acquired from the center of each cluster, superimposed.

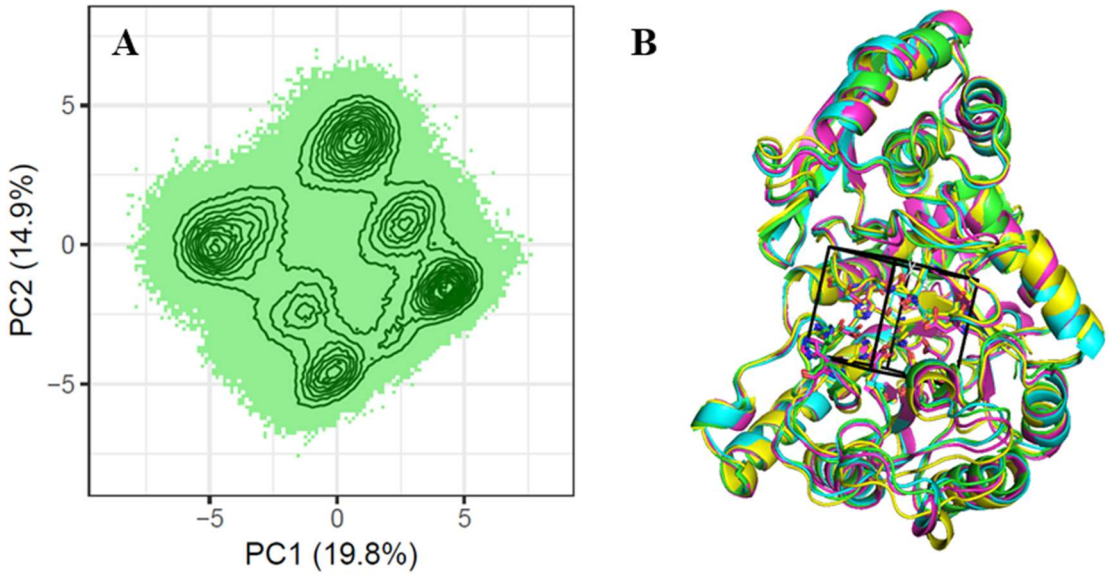


Figure 3.13 Ensemble selection scheme of enolase 1.
(A) PC1 vs PC2 of the active site residues of enolase 1 reveals distinct cluster. (B) The four representative structures acquired from the center of each cluster, superimposed.

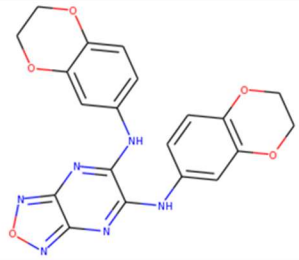
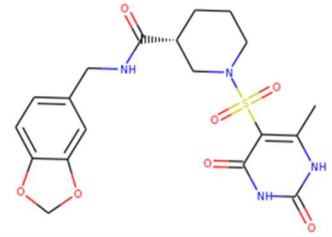
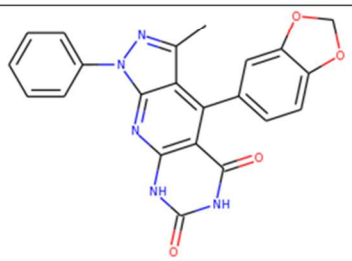
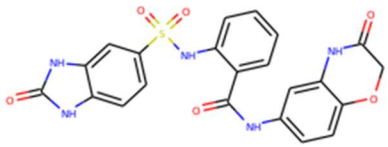
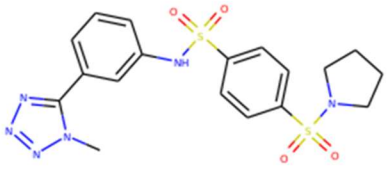
To see if the dynamics of the conserved active site of enolase 1 and 2 could influence ligand specificity, ensemble based virtual screening was performed on the apo trajectories of enolase 1 and 2. PCA was calculated on the active site residues of enolase 1 and 2. PC1 and PC2 were plotted, and one representative snapshot from each cluster was acquired and docked against ~400,000 compounds. A total of eight receptor structures were used (four from enolase 1 and four from enolase 2). Based on the representative snapshots acquired from PCA, loop L2 appears to have a significant degree of structural variability. In some of the receptor snapshots, this loop is orientated closer to the active site; and in some of the other structures, this loop is shifted further away (outside of the periodic box). It is apparent then, that prior to binding, the active site of enolase could potentially sample a range of conformations that would influence ligand binding.

Table 3.2 Control group scores from virtual screening.

Compound	Enolase 1 Scores (kcal/mol)					Enolase 2 Scores (kcal/mol)				
	1	**2**	**3**	**4**	**Average**	**1**	**2**	**3**	**4**	**Average**
2-PG	-5.3	-5.6	-5.9	-5.4	-5.5 ± 0.2	-4.9	-5.8	-5.6	-5.2	-5.4 ± 0.3
PEP	-5	-4.5	-5	-5.3	-5.0 ± 0.2	-4.7	-5.5	-5.3	-4.9	-5.1 ± 0.3
Phah	-5.5	-5.7	-4.9	-5.5	-5.4 ± 0.3	-5.1	-5.5	-5.7	-5.2	-5.4 ± 0.2
HEX	-6.2	-5.3	-5.6	-6	-5.8 ± 0.3	-5.8	-5.6	-5.8	-5.7	-5.7 ± 0.1

A set of known enolase ligands were docked onto the ensemble structures as control trials. As expected, the predicted binding poses all scored favorably to the receptors. In this study, the binding score of substrate 2-PG is used as the cutoff for finding potential inhibitors. In other words, any compound that would score more favorably than 2-PG would be considered as a potential inhibitor for further analysis.

Table 3.3 Top scoring compounds for enolase 1.

Compound	Structure	Enolase 1 Scores (kcal/mole)				Average	Average Deviation
		1	2	3	4		
ZINC000012375172		-9.1	-9.8	-9.2	-7.8	-9.0	± 0.6
ZINC000013118752		-9.1	-9.8	-9.2	-7.8	-9.0	± 0.6
ZINC000009576114		-9.8	-8.5	-8.7	-7.1	-8.5	± 0.7
ZINC000025514335		-8.9	-8.9	-9	-8.5	-8.8	± 0.2
ZINC000014728518		-9.9	-8.3	-7.6	-7.3	-8.3	± 0.8

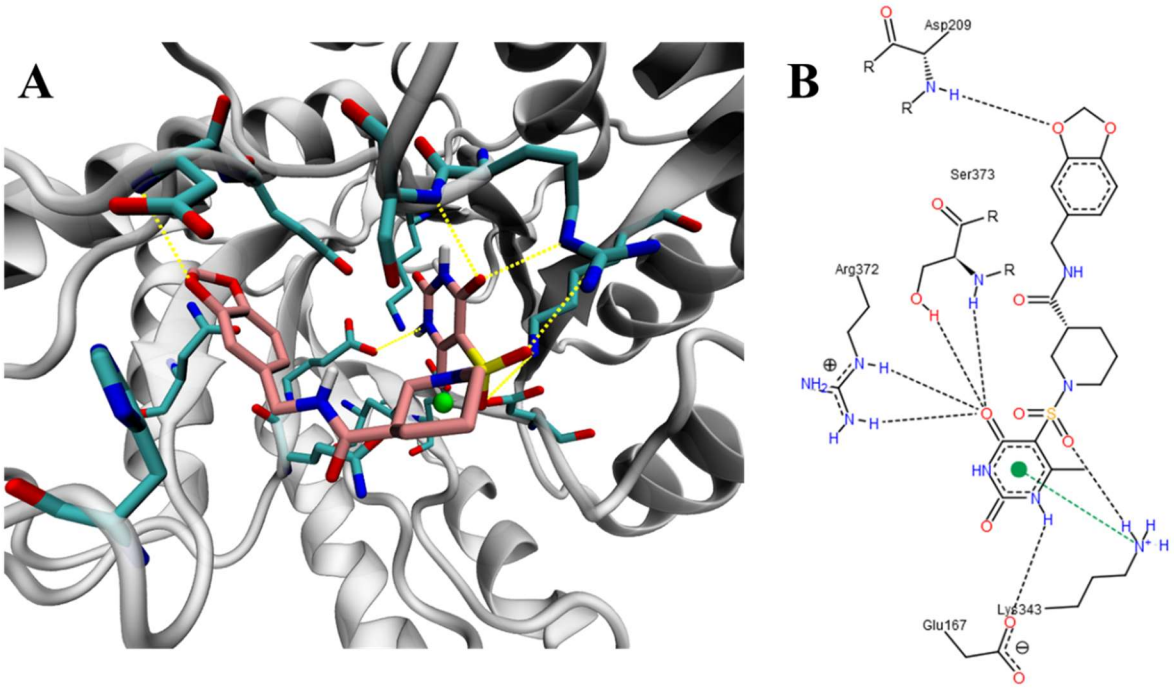
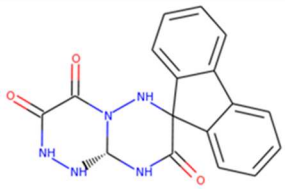
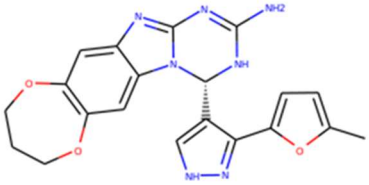
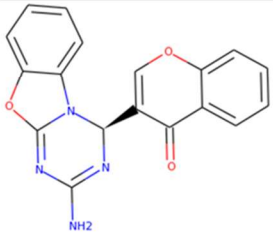
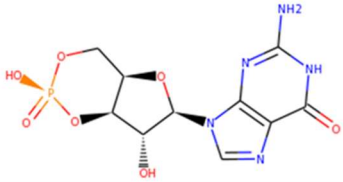


Figure 3.14 Predicted pose of ZINC000013118752 and its interactions to enolase 1.
(A) ZINC000013118752 docked onto the active site of enolase 1. (B) Predicted interactions generated by PoseView.

Three of the top five scoring compounds that dock favorably to enolase 1 comprise of Sulfonamides, a functional group found in a variety of drugs - particularly anti-bacterial and anti-diabetic medications.[85] In the lowest energy poses of these particular compounds, hydrogen bonds form between the sulfate group, and K343 and/or R372. These same compounds also pose favorable binding to Enolase 2; but with less favorable affinity scores (higher ΔG). Another common feature shared by these compounds are that they comprise of amide bonds, benzenes, and/or azaarenes. The enolase 1 structures appear to sample a large space in the binding pocket/cavity, suggesting that the aromatic functional groups can serve as scaffolds that occupy the space.

Table 3.4 Top scoring compounds to enolase 2.

Compound	Structure	Enolase 2 Scores (kcal/mol)				Average	Average Deviation
		1	2	3	4		
ZINC000000351018		-8.6	-8	-8.8	-9.7	-8.8	± 0.5
ZINC000005955022		-8.2	-8.5	-8.8	-8.7	-8.6	± 0.2
ZINC000011035700		-8.1	-8.4	-8.8	-8.7	-8.5	± 0.3
ZINC000012557331		-8	-8.1	-8.3	-8.7	-8.3	± 0.2
ZINC000033494013		-7.4	-7.9	-8.3	-9.3	-8.2	± 0.6

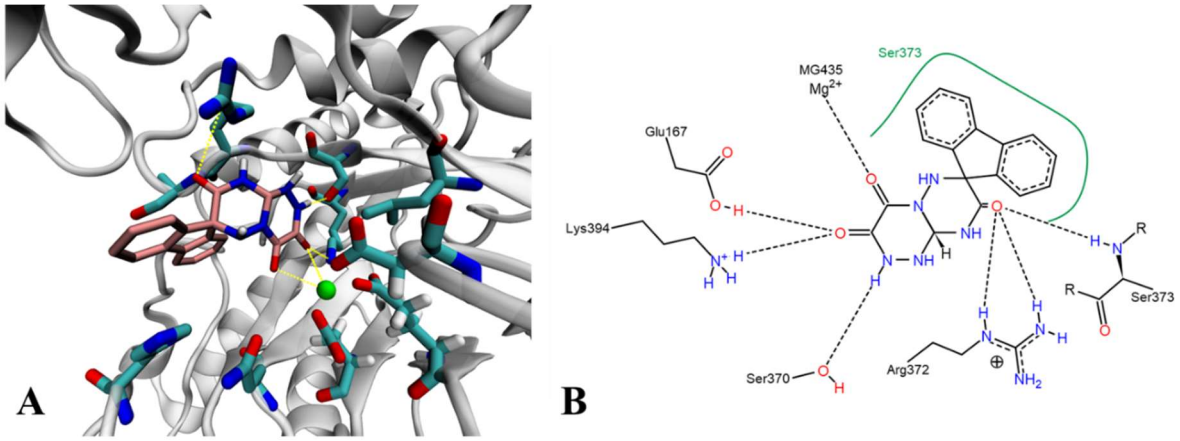
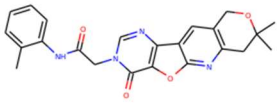
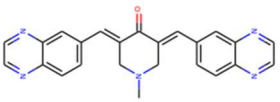
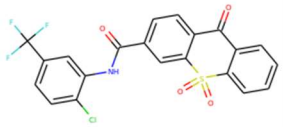


Figure 3.15 Predicted pose of ZINC000005955022 and its interactions with enolase 2. (A) ZINC000005955022 docked onto the active site of enolase 2. (B) Predicted interactions generated by PoseView.

Once again, the top scoring hits for enolase have aromatic functional groups, including benzenes and azaarenes. Interestingly, these compounds appear to have either one or no rotatable bonds; in contrast, the top scoring enolase 1 hits have more rotatable bonds, meaning that those compounds have more degrees of freedom. In addition, the top scoring compound to enolase 2 happen to score just as favorably when docked to enolase 1, but the same can't be said vice versa.

In an attempt to find potential inhibitors more selective to one isozyme, multiple rudimentary comparison methods were employed. One method involved averaging the four binding scores the ligands sampled against the receptors. The differences in binding affinities ($\Delta\Delta G$) between the ligand pose to enolase 1 versus enolase 2 was then measured. Another method involved acquiring the top docking pose score of the four trials, and to then measure the $\Delta\Delta G$ between the two homologues' top hits. It is important to note that Autodock Vina has an inherent error of approximately -2.8 kcal/mol, which was accounted for when comparing the differences between the poses and the receptors.

Table 3.5 Curated list of compounds that bind selectively to enolase 1.

Compound	Structure	Enolase 1 Scores (kcal/mol)					Enolase 2 Scores (kcal/mol)				
		1	2	3	4	Mean	1	2	3	4	Mean
ZINC000013135683		-8.2	-7.7	-7.7	-5.8	-7.4 ± 0.8	8.9	0.6	1.6	6.5	4.4 ± 3.3
ZINC000020115335		-8.1	-7.1	-7.5	-4	-6.7 ± 1.3	10.7	-5.6	2	10.7	4.5 ± 6.3
ZINC000003216746		-7.6	-7.2	-7.9	-7.1	-7.5 ± 0.3	9.4	-0.8	-0.2	8.3	4.2 ± 4.7

Over a dozen compounds were found to (1) pose more favorably to the enolase 1 receptors than the control group ligands, and (2) score poorly when docked to the enolase 2. Out of the dozen or so hits, a list of three compounds were curated in table 3.4. Given that selectively targeting enolase 2 over enolase 1 is a goal in collateral lethality studies, these results came partly as a surprise. There could be multiple explanations as to why not a single compound docked significantly more selectively to enolase 2 than its counterpart. For one, this could suggest that free enolase 1 is more promiscuous. When comparing the MD derived structures, it appears that free enolase 1 has more space in its catalytic pocket.

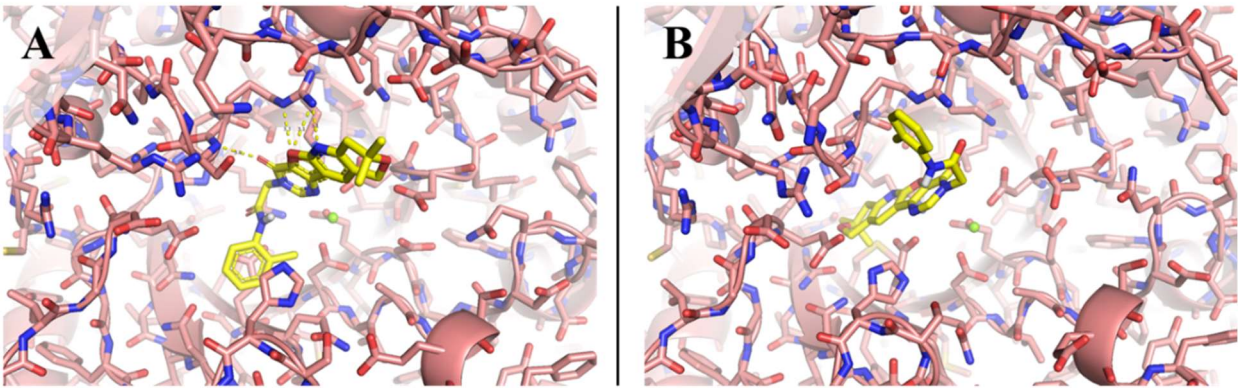


Figure 3.16 Docking pose results of ZINC000013135683.
(A) The compound docked onto the active site of enolase 1 (ensemble 4, binding affinity score is -5.8 kcal/mole). (B) The compound docked onto the active site of enolase 2 (ensemble 1, binding affinity score of 8.9 kcal/mole).

3.4 Conclusions

MD simulations of enolase 1 and 2 reveal that though both enzymes are highly conserved (they share an 83% sequence identity and 94% sequence similarity), their free-state dynamics differ greatly. This could suggest that the conformation of enolase 1 and 2 prior to ligand binding could dictate ligand specificity. As a point of interest, we wanted to explore how an enolase 2 selective inhibitor, called HEX, would influence the dynamics of both homologues. Our findings show that enolase 2 exhibits a great deal of structural changes upon HEX binding, while the enolase 1 response to HEX is less perceived.

MD simulations can provide representative conformations of proteins that can be absent in static crystal structures. For this reason, we set out on an ambitious objective: to screen over 400,000 compounds against MD derived Enolase 1 and 2 structures. In some cases, compounds docked to different representative snapshots more favorably than others. This reaffirms that enolase 1 and 2 are highly dynamic in free state, with L3 particularly influencing the docking pose results. Second, a large number of compounds were capable of binding to enolase 1 more favorably than enolase 2, but not vice versa. This could suggest that enolase 1 is the more promiscuous homologue, where it samples conformations that allow it to access ligands that are otherwise inaccessible to enolase 2. Ultimately, our findings show that docking ligands onto a single static crystal structure alone may not be sufficient, and that if virtual screening were to be used in lead-compound development, then the receptor dynamics should be taken into account.

4 The Dynamics in Allosteric Regulation of hL-PYK

4.1 Introduction

Pyruvate kinase (PYK) is an enzyme that irreversibly catalyzes the final step of glycolysis by transferring phosphate from phosphoenolpyruvate (PEP) to ADP, producing ATP and pyruvate, both of which are essential for cellular function.[86] Typically, the kinase family of proteins are categorized as enzymes that phosphorylate their substrates, but pyruvate kinase is unique because it instead dephosphorylates its own substrate. In mammals, there are four tissue-specific isozymes of pyruvate kinase: M1 (muscle), M2 (fetal), L-Type (liver), and R-Type (erythrocyte) PYK.[87] Due to alternate starting sites, a single gene, PKLR, encodes for both the L and R pyruvate kinases, while a PKM gene encodes for both the M1 and M2 isozymes. The liver isozyme in humans, hL-PYK, has an extensive allosteric regulation scheme that is still not fully understood on a structural level.[88]

During periods of fasting, glucose is produced in the liver through the process of gluconeogenesis; to prevent futile breakdown of this newly generated sugar, glycolysis is downregulated through one of its irreversible steps.[89] The enzyme hL-PYK has several extensive regulatory processes responsible for maintaining this glycolytic/gluconeogenic homeostasis, making it a relevant target in metabolic regulation studies.[90] Through a process of feed-forward stimulation, hL-PYK can be allosterically activated by fructose-1,6-bisphosphate, an earlier glycolytic intermediate. The enzyme can be allosterically inhibited by both alanine and ATP as well. Unfortunately, studies on the structure of hL-PYK are limited, and so the allosteric mechanisms involved in the enzyme are ambiguous; that is of course a point of interest. By running MD simulations of hL-PYK both in the presence and absence of its allosteric effectors, we attempt to discern the dynamics in allosteric regulation of hL-PYK.

4.1.1 Structural Analysis of hL-PYK

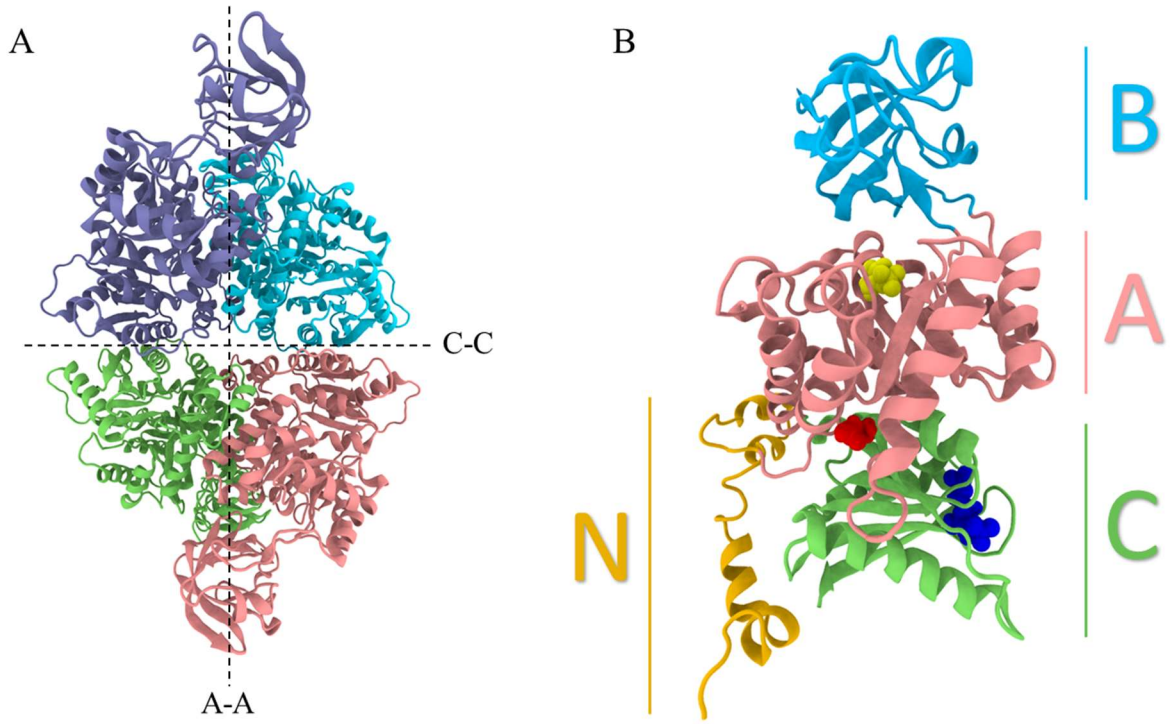


Figure 4.1 Structure of hL-PYK.
(A) The tetrameric structure of hL-PYK with its two interfaces labelled. (B) Monomer of hL-PYK with its labeled four domains. The active site is in yellow space fill, the Alanine allosteric site is in red space fill, and the Fructose-1,6-bisphosphate (FBP) site is in blue space fill.

Structurally, hL-PYK is a homo-tetramer, where each monomer houses four distinguishable domains: A, B, C, and N-terminal domains.[88] The A-domain is a TIM-barrel fold (eight α-helices and eight β-strands) that harbors the active site. Above the catalytic domain, resides the flexible B-domain; previous studies of PYK have shown that this B-domain encloses the active site upon substrate binding, acting as a capsid.[91] Below the A-domain resides the C-domain, comprised of five α-helices and β-strands. The FBP allosteric binding site is housed here, while the alanine allosteric binding site can be found at the interface of domains A and C. The N-terminus domain, which is uniquely present in mammalian PYK, is responsible for the hormonal regulation of hL-PYK.[92]

Each monomer interacts with one another through two interfaces – the A and C interfaces. As aptly named, the A-interface comprise of interactions between the A-domains of each monomer, while the C-domains interact with one another at the C-interface. Previous experimental studies on pyruvate kinase isoforms have shown cooperativity, suggesting that dynamical information can be relayed across monomers through the interfaces.[86] Regarding the active site, the substrate PEP binds in coordination with a divalent magnesium cation, and a monovalent potassium ion. Not only are these ions necessary for enzymatic activity, but an $ADP{\cdot}Mg^{2+}$ complex within the active site is also required.

In previous studies, allosteric regulation of hL-PYK with FBP was explored through a variety of mutations.[93,94] Based on the mutations done near the FBP binding site, Ishwar et.al. hypothesized that in the absence of FBP, the effector loop (residues 527-533) would interact with residue D499 across the C-interface, and that the affinity towards PEP would be lowered in accordance to this conformation.[93] They observed that mutations D499N and W527H led to an increase in PEP affinity without FBP treatment, and further hypothesized that activation of hL-PYK was caused by disrupting the hypothesized interactions between D499 and W527 and/or R528.

In this study, we run MD simulations of hL-PYK both in the presence and absence of FBP. Our findings show that in the absence of FBP, residues D499 indeed interacts with W527 and R528 across the C-interface, supporting the previous hypothesis proposed by Ishwar et al.[93] Our findings also show that FBP binding to hL-PYK disrupts a multitude of interactions across the C-interface. In addition, we run an MD simulation of hL-PYK-alanine complex. Our results show that upon binding of alanine, the dynamics of hL-PYK shift to a more "open" conformation that correlates to motions observed in previous studies on PYK homologues.

4.2 Experimental Procedures

4.2.1 Preparation of hL-PYK for MD Simulations

All crystal structures of hL-PYK have disordered and partly unresolved B-domains, while the rest of the protein is well resolved (and interpretable). This would suggest that the B-domain is capable of adopting multiple conformations even when crystallized. Before MD simulations of hL-PYK could be performed, it was important to first model these highly dynamic missing residues. We used PDB 4IP7, hL-PYK in complex with citrate and FBP, as the starting model.[95] Using PyMOL[78], the coordinates of the missing atoms in the B-domain were acquired through superposition of the PYK monomers from PDB's 4IP7, 4IMA, and 2VGB. The first 25 residues of the N-terminus are unresolved in all crystal structures of hL-PYK due to its large flexibility. The software Modeller[96] was used to approximate the missing N-terminus residues.

In the active site of PDB 4IP7 (the starting model for our MD simulations), citrate is in place of PEP, and Manganese is in place of Magnesium. There is no ADP in the crystallized structure. The coordinates of chain A from PDB 4HYV[97], which contains PEP in the conserved active site, was superimposed to each chain of PDB 4IP7. The coordinates of the PEP were then transferred onto the hL-PYK model. The coordinates of ADP-Mg^{2+} were acquired from superimposing chain A of PDB 3GR4 onto each monomer of hL-PYK. The fidelity of these clippings was confirmed through xLeap. Previous studies have found the alanine (allosteric inhibitor) binding site to be conserved between rabbit muscle pyruvate kinase and hL-PYK. To generate an alanine-bound hL-PYK model, PDB 2G50[98] (from the rabbit muscle isozyme) was superimposed onto each monomer of the hL-PYK system, where the coordinates of alanine were then transferred onto the allosteric site of hL-PYK.

4.2.2 MD Simulations of hL-PYK

The molecular dynamics simulations were carried out using the CUDA version of the Amber16 suite of programs[79] along with the modified version of the Cornell et al. ff14SB[80] force field. Each simulation ran using the NVIDIA GeForce GTX 980 GPU. The x-ray crystal structure PDB 4IP7 (resolution of 1.8 Å) was used as the initial structure of hL-PYK. The GAFF2[79] Force Field was used to parameterize the ligands. The forcefield for ADP was acquired from the Bryce group (AMBER parameter database from Manchester).[99]

Each system was solvated in a TIP3P[35] periodic octahedron box containing water molecules. Counterions were added to the system to obtain electrostatic neutrality. All simulations were kept at a Temperature of 300 K and a constant pressure of 1 bar. Electrostatic interactions were calculated using the particle mesh Ewald (PME) summation method, where long-range nonbonded interactions were accounted for with a 9 Å cutoff.[82] For each system a 2 fs time step was used to solve the equation of motion. After a series of minimizations, the system was equilibrated with a harmonic constraint through five subsequent MD simulations of 1 ns each, where the force constant was reduced from 500 to 5 kcal· Å2/mole. A final equilibration step was performed for 1 ns, where no harmonic constraints were conducted; this allowed the enzyme complex to move freely. All images of the structures were generated through VMD[83].

Table 4.1 Summary of MD simulations performed on hL-PYK.

	MD Simulations	**Ligands**	**Timescale**
1	hL-PYK	4 PEP, 4 ADP, 8 Mg^{2+}, 4 K^{+}	1.1 μs
2	hL-PYK · FBP	4 PEP, 4 ADP, 8 Mg^{2+}, 4 K^{+}, 4 FBP	1.1 μs
3	hL-PYK · Alanine	4 PEP, 4 ADP, 8 Mg^{2+}, 4 K^{+}, 4 Alanine	1.1 μs

4.3 Results & Discussion

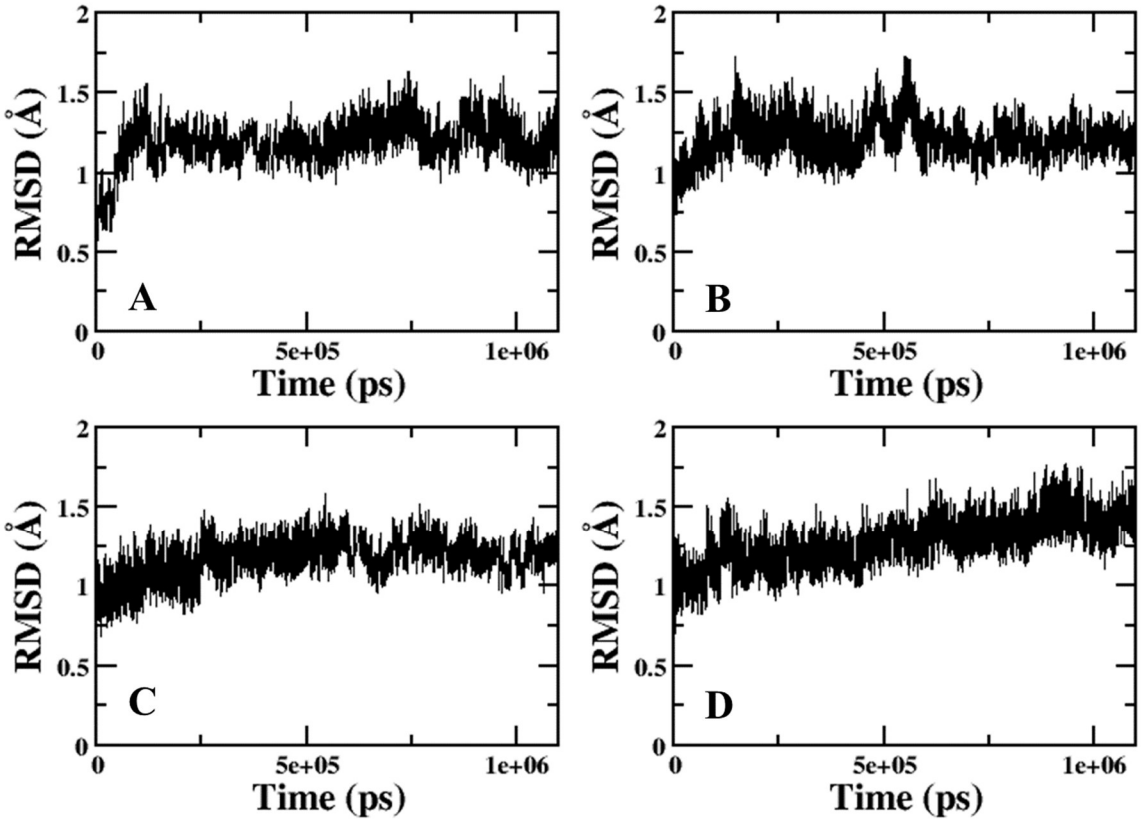


Figure 4.2 RMSD of the hL-PYK monomers.
Each graph represents a separate monomer, with the chain labelled at the bottom left.

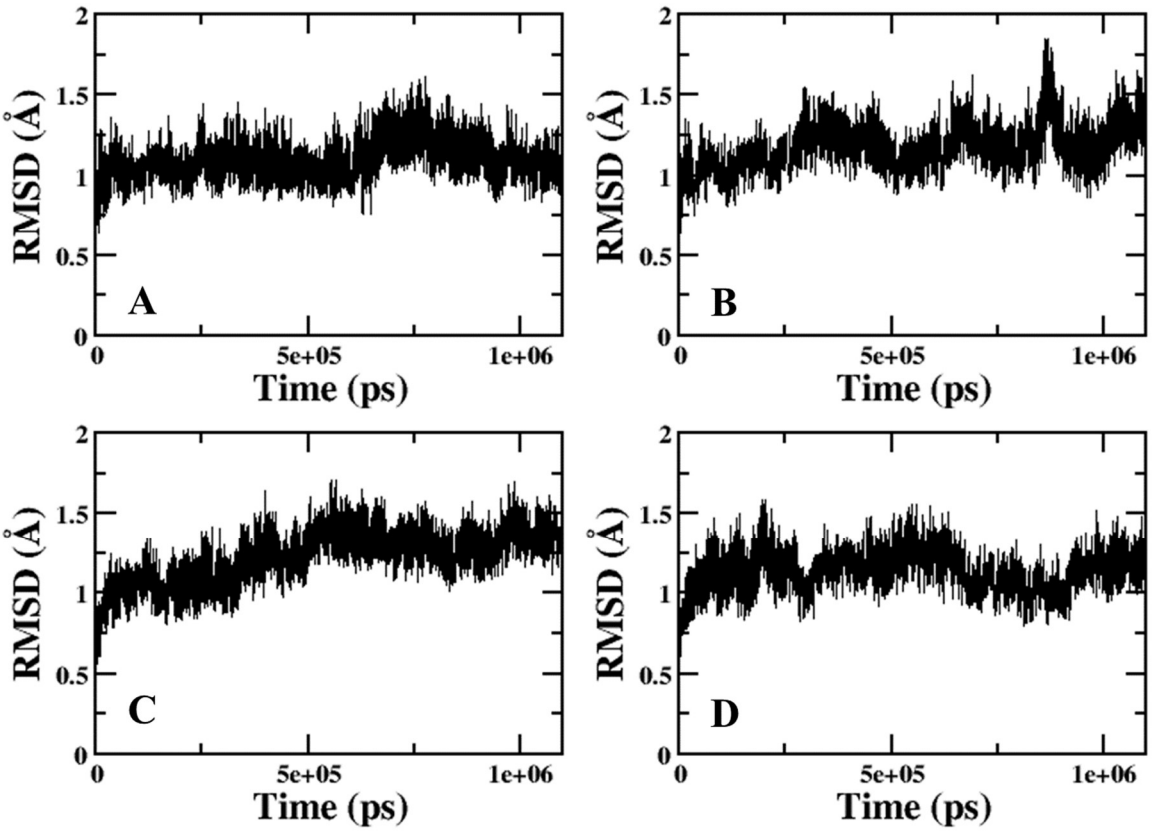


Figure 4.3 RMSD of hL-PYK bound to FBP.
Each graph represents a separate monomer, with the chain labelled at the bottom left.

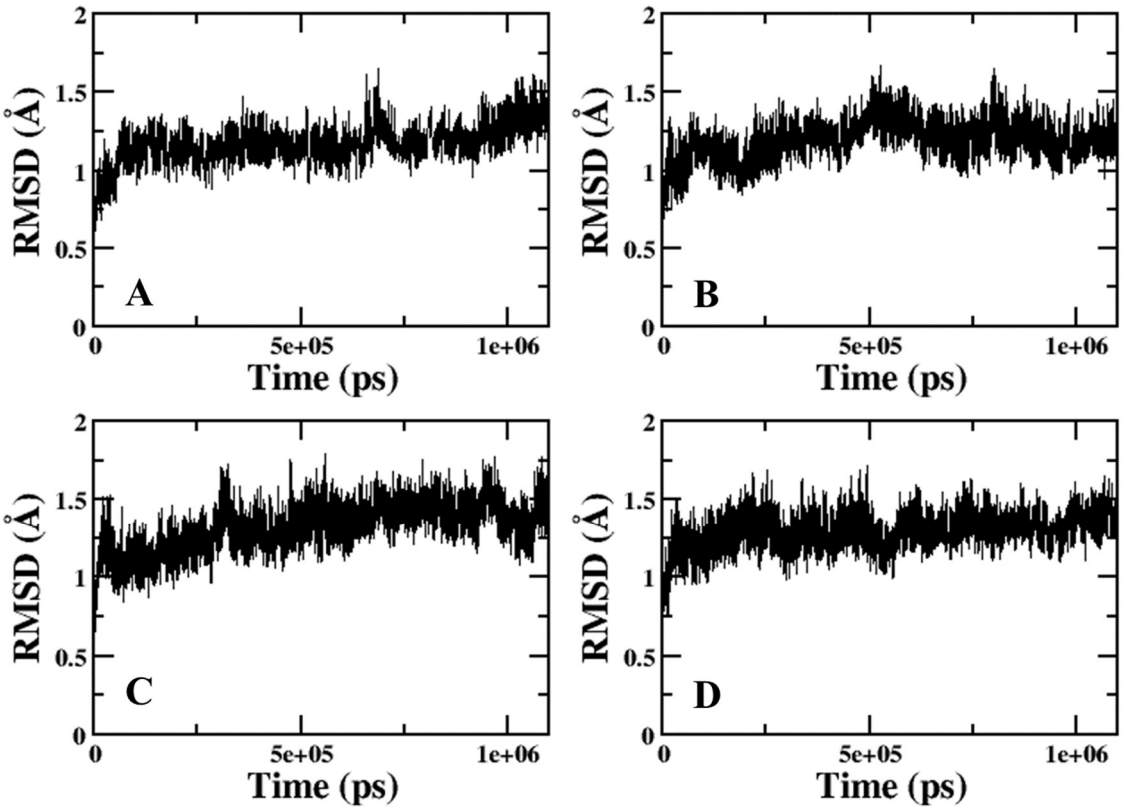


Figure 4.4 RMSD of hL-PYK bound to Alanine.
Each graph represents a separate monomer, with the chain labelled at the bottom left.

As mentioned previously, RMSD is useful in determining convergence of a molecular system. An important consideration here is that the RMSD method relies on rigid-body alignment of an ensemble of structures, and this method of alignment is incredibly sensitive to flexible regions of a molecular body. Therefore, superimposing an entire macromolecule (including the flexible subsets of the molecule) can often lead to unnecessarily large RMSD values. Given that previous crystallographic studies have found both the N and B-domains of PYK to fluctuate greatly, we masked out the first twenty-five residues of the N-terminus domain for all calculations. For the sake of RMSD calculations, the B-domain was also masked out, but all other analysis is done with the inclusion of this flexible domain.

Based on the RMSD, it appears that most chains reach convergence over the microsecond-long simulation, but some notable issues should be considered. Chain D of the standard hL-PYK trajectory (absence of allosteric effector) appears to continuously rise in its RMSD over time, indicating that it may not be fully stabilized. This same characteristic is absent in the other two hL-PYK simulations, where chain D appears to stabilize fairly quickly in the presence of allosteric effectors. Given the difficulty in crystallizing the tetramer, it is plausible to suggest that the starting coordinates of the monomers could sample different conformations.

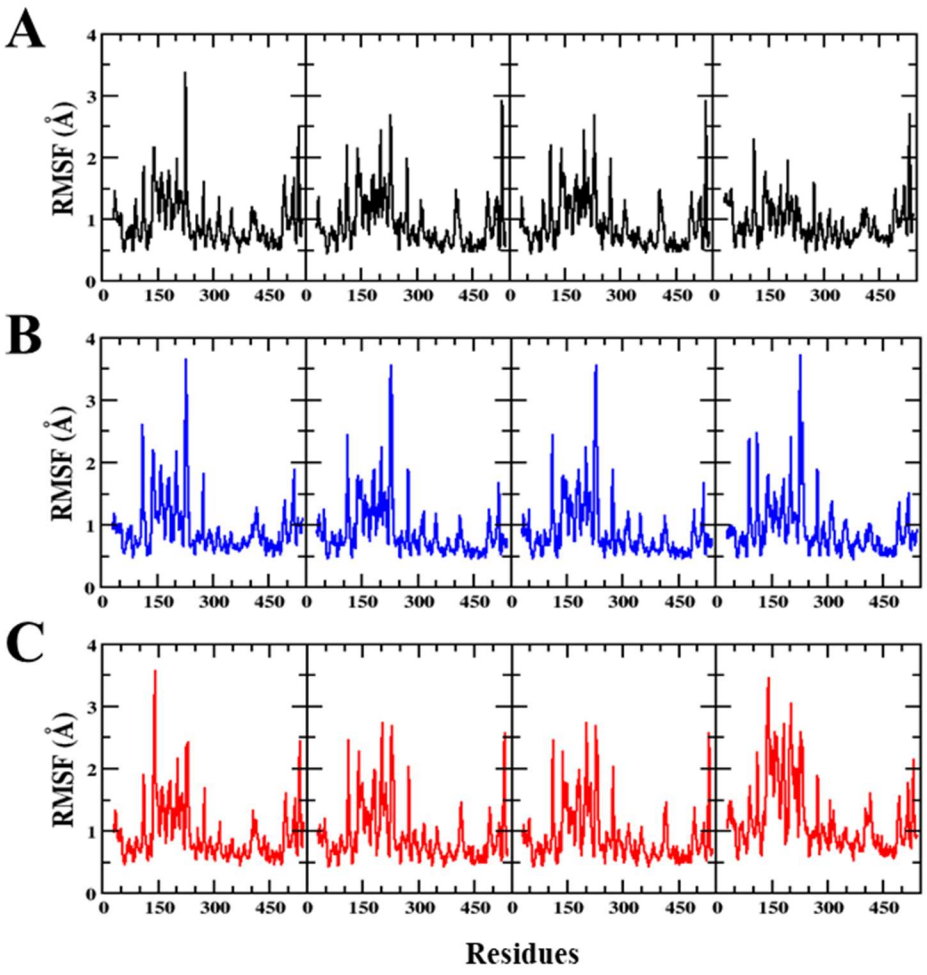


Figure 4.5 RMSF of the hL-PYK simulations.
(A) RMSF of hL-PYK. (B) RMSF of hL-PYK in complex with FBP. (C) RMSF of hL-PYK in complex with alanine. Given that hL-PYK is a homo-tetramer, the panels are ordered from chains A to D.

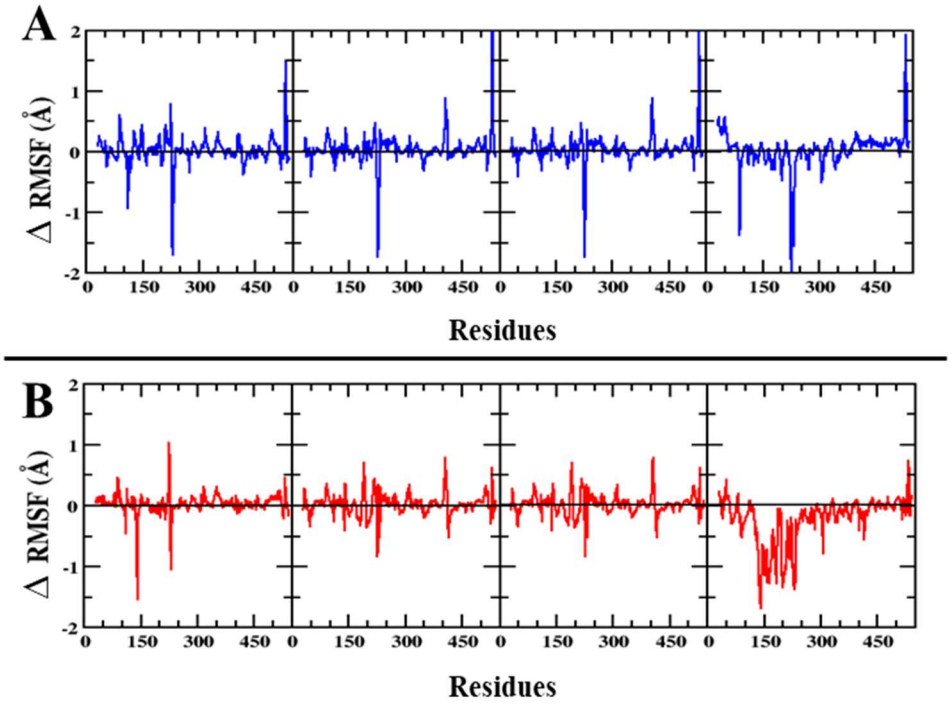


Figure 4.6 Comparison of RMSF of hL-PYK.
(A) Difference in RMSF between hL-PYK and hL-PYK-FBP. (B) Difference in RMSF between hL-PYK and hL-PYK-Alanine. Given that hL-PYK is a homo-tetramer, the panels are ordered from Chains A to D.

Results from the RMSF calculations of the systems show a great deal of dynamical differences occurring in the B-domains (residues 143-230) of each monomer. It appears that upon binding of either FBP or alanine, the B-domains fluctuate greatly. In addition, the FBP effector loops (residues 527-533) undergo large fluctuations in both the FBP bound simulation, and even the alanine-bound hL-PYK system. This could suggest that the dynamics induced by alanine inhibition are linked to the FBP binding site.

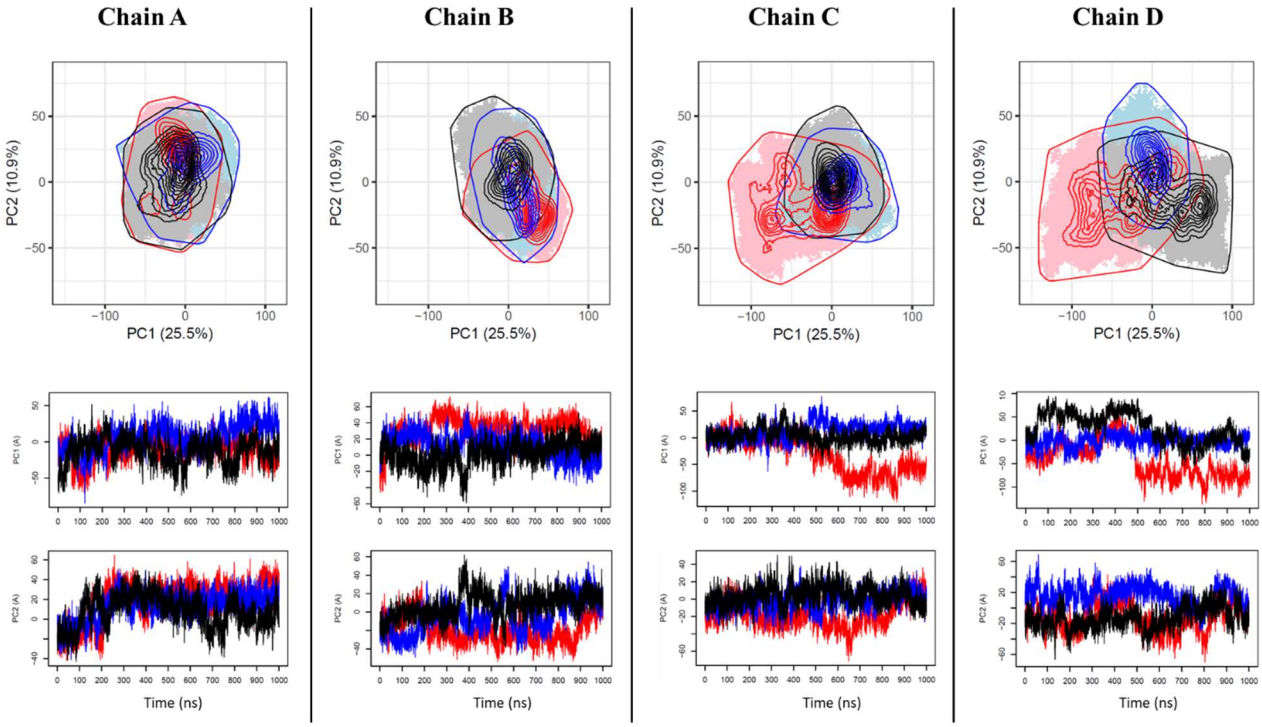


Figure 4.7 PC plots of the PYK monomers.
Top graphs display PC1 vs. PC2 of the monomers. Black is representative of hL-PYK, red represent Alanine bound hL-PYK, and blue represent FBP bound hL-PYK. Bottom plots are PC1 and PC2 with respect to time.

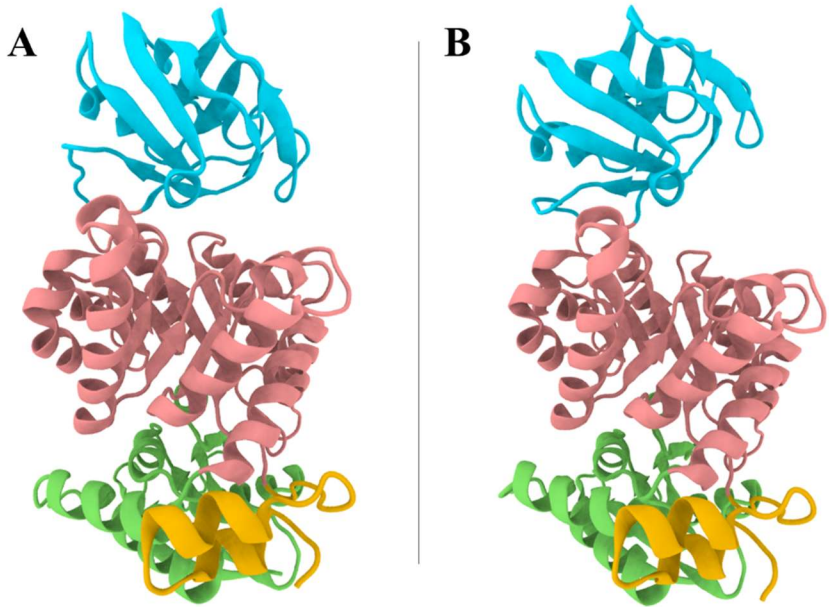


Figure 4.8 Dominant motions revealed by PCA of the monomers.
(A) Conformation of the hL-PYK monomer when PC1 equals 75. (B) Conformation of the hL-PYK monomer when PC1 equals -75.

To capture the most dominant motions at the monomeric level, PCA was performed on the individual chains of the PYK trajectories. As expected, the B-domain exerts most of the motions. Chain A and B of all simulations appear to overlap in PC1 and PC2, indicating that the dominant motions these two chains exert are identical. This could suggest that the dynamics induced by allostery are subtle, and that other techniques could aid in uncovering them. Interestingly, chains C and D of the hL-PYK-alanine simulations exert a different motion after 500 nanoseconds. Here, the B-domain appears to open up, exposing the active site as a result. Previous studies on PYK isozymes have proposed that an inactive PYK would orient itself so that the B-domains would rotate outwards (away from the catalytic site).[100] PC1 of the hL-PYK systems demonstrates that both chains C and D undergo this "opening" movement when bound to alanine.

There are multiple explanations as to why chains A and B of the alanine-PYK complex sample a more closed/rigid conformation in regard to the B-domain. For one, this could be a case where hL-PYK could function as a dimer of dimers, which has been proposed in previous studies. Less optimistically, it is possible that these allosteric transitions require longer timescale simulations. This could suggest that the conformation transition of all monomers is not simultaneous. That being said, the chains C and D conformation shift piqued enough interest for further exploration; in a later section, we compare the dynamics from this particular conformation against the standard hL-PYK dynamics.

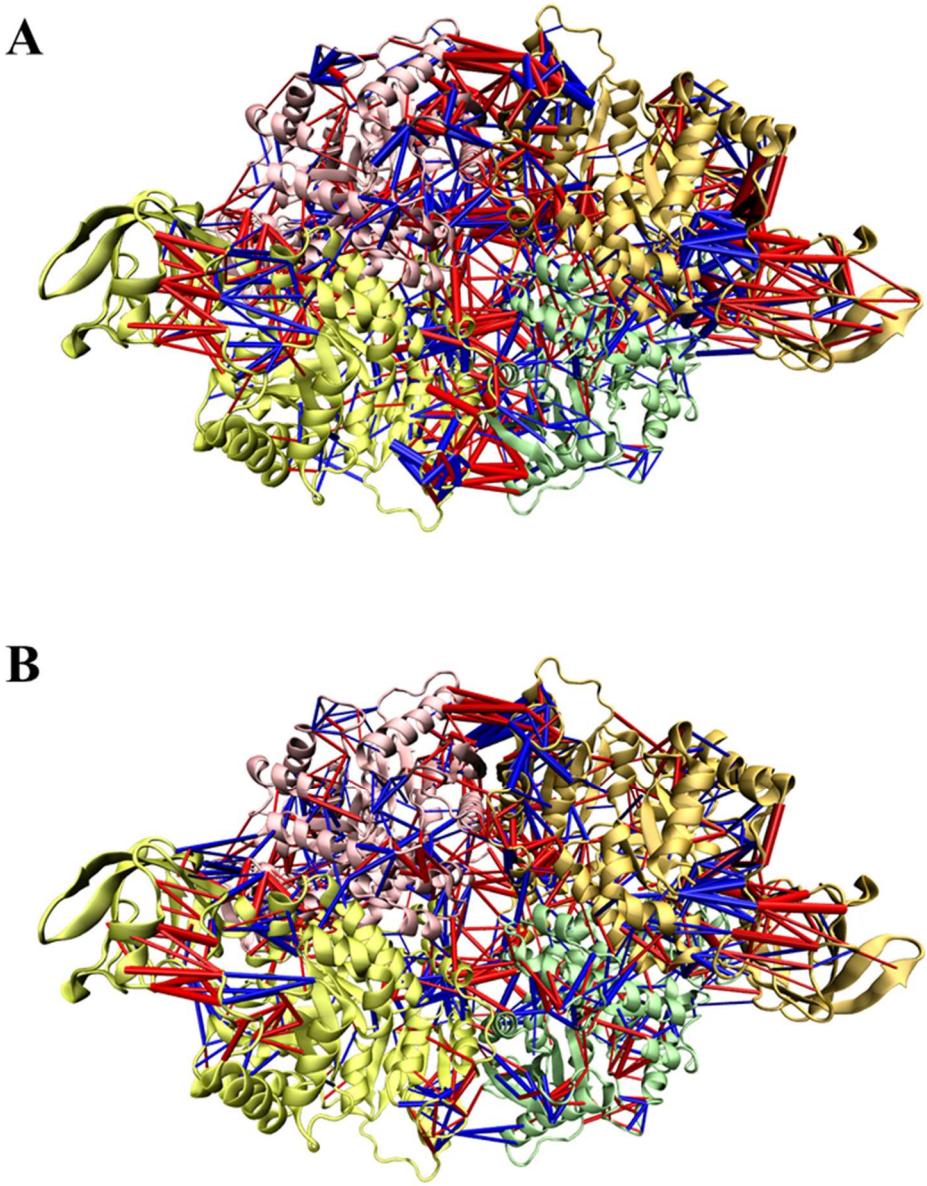


Figure 4.9 Dynamical contacts statistics of the hL-PYK tetramer.
(A) Changes of the residue-residue contacts between hL-PYK and hL-PYK-FBP complex. (B) Changes of the residue-residue contacts between hL-PYK and hL-PYK-alanine. Blue bars indicate more contact formations, and red bars indicate more contact breakages.

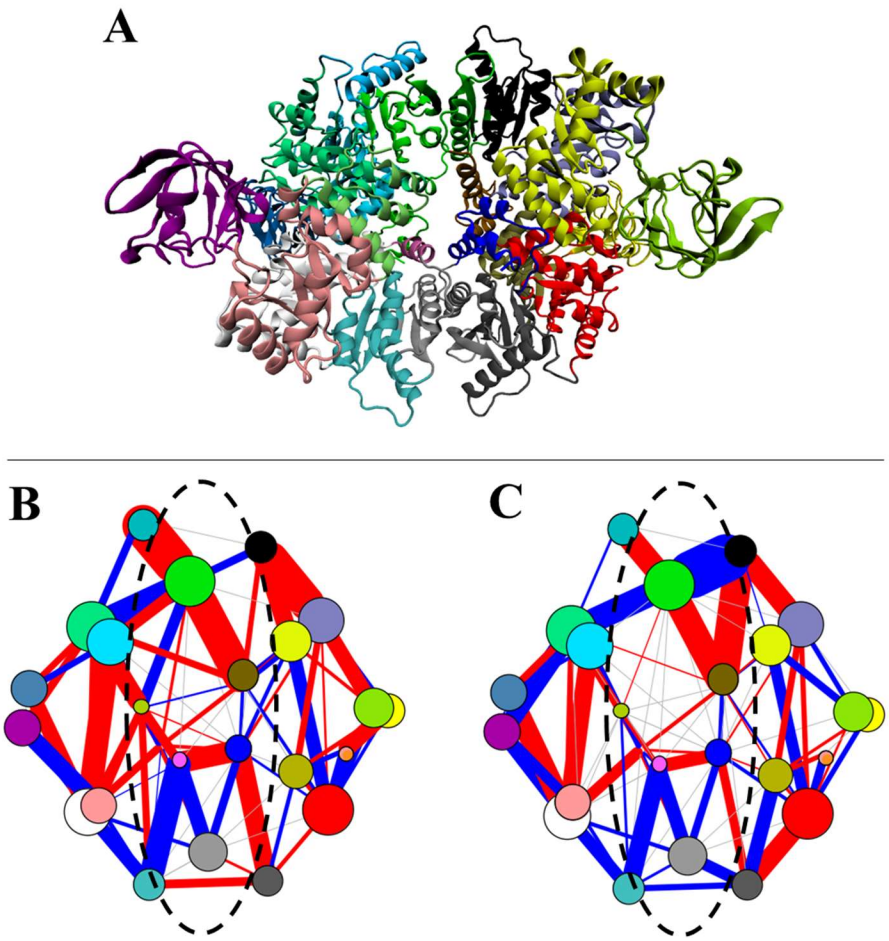


Figure 4.10 Difference contact network analysis of the hL-PYK simulations.
(A) Communities generated from calculating the residue-residue contact statistics of all hL-PYK simulations and generating a consensus network (B) dCNA of hL-PYK vs. hL-PYK-FBP complex. (C) dCNA of hL-PYK vs hL-PYK-alanine complex.

Results from calculating and comparing the contact dynamics of all three simulations reveals that the dynamics between hL-PYK both in the presence and absence of FBP differ greatly. At the C-interface, it is apparent that a large number of contacts are broken, revealing that upon FBP binding, the C-domains from each monomer can shift further away with respect to each other, breaking contacts as a result. When comparing the dynamics of hL-PYK both in the presence and absence of alanine, the C-interface has some lesser contact breakages, but in many cases, more contacts are forming between communities from the C-interface.

4.3.1 The Role of the FBP Binding Site in hL-PYK Regulation

In previous studies on M1 and M2 PYK crystal structures, interactions between D499 and W527, and D499 and R528, were found to occur across the C-interface of the tetramer in the absence of FBP. Residues W527 and R528 reside in the FBP binding site, and crystal structures of M1 and M2 PYK show that upon FBP binding, the backbone atoms of these residues interact with the allosteric effector, leading to a disruption of the previous aforementioned C-interface contacts.

Recent studies on hL-PYK found that by mutating residues D499, W527, and R528, the enzyme would exhibit an increase in substrate affinity without need of the FBP effector, suggesting that the perturbation of these residue-residue contacts could lead to an increase in substrate affinity. Due to the lack of reliable crystal structures of hL-PYK, these contacts are merely a speculation. A very recent crystallographic study on hL-PYK attempted to explore this hypothesis by solving the structures of these PYK mutations, but complications arose from resolving the crystal structures, and the C-interface contacts in the absence of FBP were not observed.

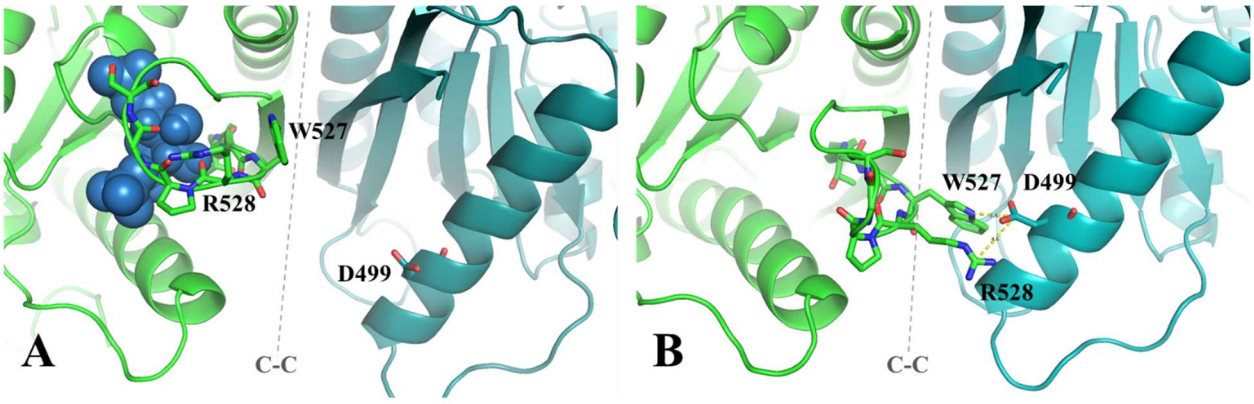


Figure 4.11 Comparison of the FBP effector loop in MD sampled hL-PYK.
(A) The hL-PYK-FBP complex. FBP interacts with the backbone of loop 527-533. (B) hL-PYK in the absence of FBP. The W527 and R528 sidechains from one monomer interact with D499 from another monomer, forming interactions across the C-interface. These interactions have only been hypothesized in hL-PYK.

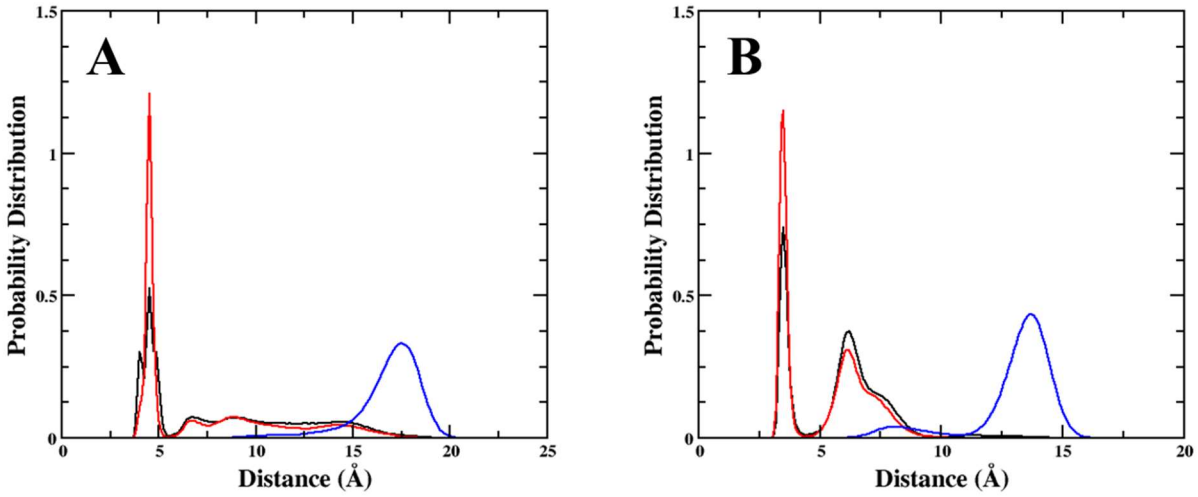


Figure 4.12 Probability distribution of the distances between FBP effector loop and D499 (A) Distance measured from CZ of R528 and CD of D499. (B) Distance measured from NZ of W527 and CD of D499. Black represents hL-PYK, blue represents the hL-PYK-FBP complex, and red represents the hL-PYK-alanine complex.

Results from our MD simulations show that when hL-PYK is simulated without FBP, the allosteric effector loop moves across the C-interface, where residues R528 and W527 from one chain, interacts with D499 from another. When FBP is in complex with hL-PYK, this effector loop is pulled away from the C-interface, and the distance between the side-chains are 10-20 Å. Based on the observed motions in our MD simulations, as well as the previous mutational studies done on residues D499 and W527, we propose that FBP allosterically activates hL-PYK by disrupting C-interface interactions between W527-D499 and R528-D499. Furthermore, difference contact statistics showcase that many residue-residue interactions across the C-interface are disrupted upon FBP binding, insinuating that the shifts of the C-domains away from one another are correlated to PYK activation. We hypothesize that mutations on D499, W527, and R528 would not only lead to an increase in hL-PYK activity, but that the dynamics would also resemble that of FBP bound hL-PYK.

4.3.2 Speculation on the Monomeric Dynamics Induced via Alanine Binding

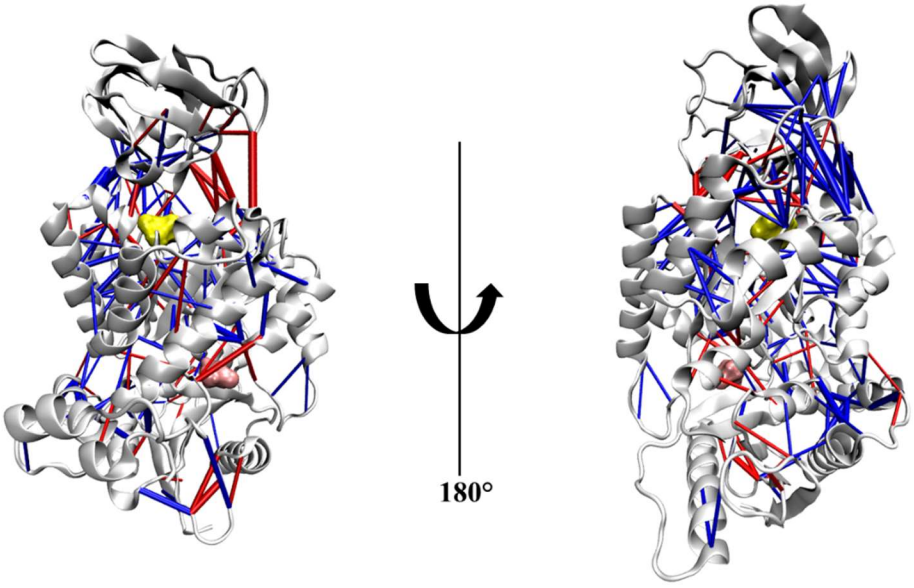


Figure 4.13 Dynamical contact statistics of the last 500 ns of chains C & D.
Changes of the residue-residue contacts between hL-PYK and hL-PYK-alanine at the monomeric level. The last 500 nanoseconds of Chains C and D were used since they exhibited B-domain "opening" motions. The active site is in yellow space-fill, and the alanine binding site is in pink space-fill. Blue bars indicate more contact formations, and red bars indicate more contact breakages.

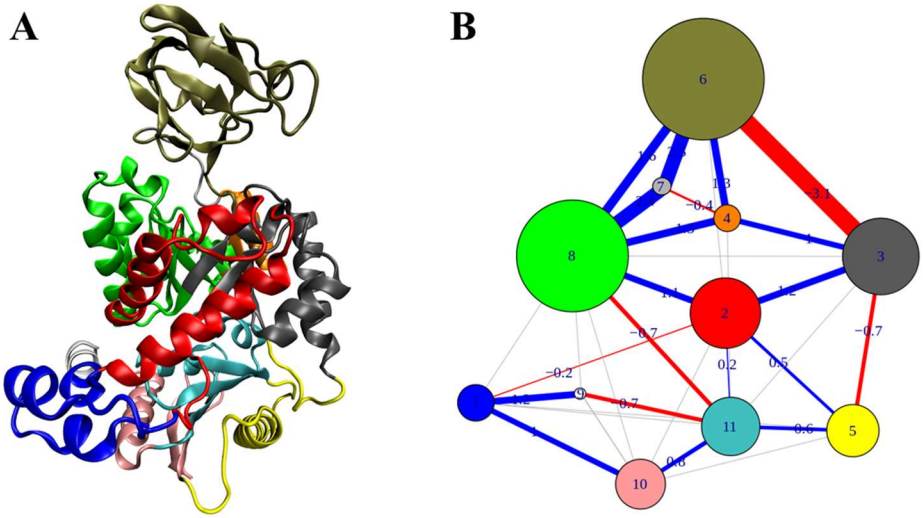


Figure 4.14 Difference contact network analysis of hL-PYK vs alanine-bound hL-PYK.
(A) Communities generated from calculating the residue-residue contact statistics and generating a consensus network. (B) Difference contact network analysis (dCNA) reveals community-level dynamical changes.

Previous studies on PYK isozymes have attributed the enzyme's allosteric inactivation to the "opening" of its B-domain.[100] In an attempt to discern the dynamical changes induced by alanine inhibition of hL-PYK, we ran and analyzed an MD simulation of hL-PYK-alanine complex. PCA revealed that chains C and D began to sample a B-domain opening shift around 500 nanoseconds, while chains A and B managed to remain intact for the entire 1.1 microsecond trajectory. We therefore speculate that these last 500 ns of chains C and D could provide insight on the role alanine plays in the conformational dynamics of hL-PYK, at least at the monomeric level.

Difference contact statistics between hL-PYK and hL-PYK-alanine reveals significant motions of both the A and B domains. In the hL-PYK-alanine complex, the B-domain tilts outward, exposing the catalytic site by forming contacts on one corner (through a tilt) and breaking contacts with the A-domain on the other corner. This entire domain shift reinforces the notion that the B-domain acts as a capsid, where upon binding of the inhibitor alanine, the domain "opens" up and exposes the active site. In addition, residues in the catalytic A-domain (TIM barrel fold) form more contacts with one another in the alanine bound simulation, which could suggest that allosteric inhibition by alanine leads to contraction of the TIM barrel domain.

4.4 Conclusions

Previous structural studies on hL-PYK have had difficulty in coming up with models that would explain the structural conformation shifts induced by allostery. In an attempt to uncover an atomic-level mechanism involved in hL-PYK allosteric regulation, we set out on our exploration using MD simulations of the enzyme both in the presence and absence of the allosteric effectors.

Results from our MD simulations show that in the absence of FBP, a great deal of contacts form across the C-interface, and that the FBP effector loop (527 to 533) from one monomer interacts with D499 from another. Furthermore, we found that FBP binding induces a great deal of contact breakages across the C-interface, with the interactions between W527-D499 and R528-D499 being perturbed greatly. Interestingly, we found that in the MD simulation of alanine bound hL-PYK, these residues interact with one another even more-so, suggesting a possible synergy between the two allosteric effectors. These findings support the hypothesis that residues W527, R528, and D499 are key to allosteric regulation of hL-PYK.

5 CONCLUSIONS

In this study, we explored the dynamics of two enzymes involved in glycolysis in the context of regulation. Despite being metabolically coupled, the two enzymes have distinct regulatory features that make them worth exploring individually.

In the enolase study, we found that the free-state dynamics of the two isozymes, enolase 1 and 2, are very different despite their high conservation. MD simulations of the two enzymes bound to compound HEX, an enolase 2 selective inhibitor, found that the ligand induces greater dynamical changes to enolase 2 than enolase 1. As a point of interest, we embarked on ensemble virtual screening to further distinguish the dynamics of free enolase 1 and 2. Our results led to the discovery of compounds that could be potentially selective to enolase 1, while not a single compound could bind significantly more favorably to enolase 2. This not only reinforces that the dynamics of these two highly conserved homologues are different, but the results could suggest that enolase 1 is more promiscuous.

Given the past difficulties in interpreting the structural motions involved in allosteric regulation of hL-PYK, we embarked on a mission to uncover an atomic-level detailed model of hL-PYK allostery. The results from our MD simulations show that in the absence of allosteric effectors, hL-PYK forms a great deal of contacts across the C-interface, with contacts between W527 and D499 being the most notable ones. Upon binding of FBP, a great deal of contacts across the C-interface, particularly the interaction between W527 and D499, are disrupted. Based on the results of previous mutational studies conducted on the FBP binding site, we propose that FBP activates the enzyme by disrupting contacts across the C-interface, and these contact disruptions are facilitated by the stabilization of the FBP effector loop (residues 527-533). As an extension, we propose that mutating residues D499 or W527 could induce dynamical changes to

hL-PYK that resemble that of FBP. MD simulations of hL-PYK in complex with alanine show that some of the monomers can undergo entire domain motions at the nanosecond timescale. Chains C and D of the trajectory had the entire B-domain shift outwards, exposing the active site as a result. This conformation has been observed in some previous studies on PYK isozymes. Because chain A and B remained in a fixed/overlap conformation throughout the 1.1 microsecond simulation, we conclude that the conformational dynamics induced by alanine are not simultaneous across the monomers. Ultimately, our findings provide some insight on the role dynamics play in glycolysis regulation.

REFERENCES

1. Metallo, C. M.; Vander Heiden, M. G. Understanding Metabolic Regulation and Its Influence on Cell Physiology. *Molecular Cell* **2013**, *49* (3), 388–398.
2. Gunasekaran, K.; Ma, B. Y.; Nussinov, R., Is allostery an intrinsic property of all dynamic proteins? *Proteins-Structure Function and Bioinformatics* **2004**, *57* (3), 433-443.
3. Thoma, J. A.; Koshland, D. E. Competitive Inhibition by Substrate during Enzyme Action. Evidence for the Induced-Fit Theory. *Journal of the American Chemical Society* **1960**, *82* (13), 3329–3333.
4. Engström, L. The Regulation of Liver Pyruvate Kinase by Phosphorylation—Dephosphorylation. *Current Topics in Cellular Regulation* **1978**, 29–51.
5. Daniel, R.; Dunn, R.; Finney, J.; Smith, J. The Role of Dynamics in Enzyme Activity. *Annual Review of Biophysics and Biomolecular Structure* **2003**, *32* (1), 69–92.
6. Teague, S. J. Implications of Protein Flexibility for Drug Discovery. *Nature Reviews Drug Discovery* **2003**, *2* (7), 527–541.
7. Buch, I.; Giorgino, T.; Fabritiis, G. D. Complete Reconstruction of an Enzyme-Inhibitor Binding Process by Molecular Dynamics Simulations. *Proceedings of the National Academy of Sciences* **2011**, *108* (25), 10184–10189.
8. Boehr, D. D.; Nussinov, R.; Wright, P. E. The Role of Dynamic Conformational Ensembles in Biomolecular Recognition. *Nature Chemical Biology* **2009**, *5* (11), 789–796.
9. Supuran, C. T. Structure-Based Drug Discovery of Carbonic Anhydrase Inhibitors. *Journal of Enzyme Inhibition and Medicinal Chemistry* **2012**, *27* (6), 759–772.

10. Greer, J.; Erickson, J. W.; Baldwin, J. J.; Varney, M. D. Application of the Three-Dimensional Structures of Protein Target Molecules in Structure-Based Drug Design. *Journal of Medicinal Chemistry* **1994**, *37* (8), 1035–1054.
11. Davis, A. M.; Teague, S. J.; Kleywegt, G. J. Application and Limitations of X-Ray Crystallographic Data in Structure-Based Ligand and Drug Design. *ChemInform* **2003**, *34* (34).
12. Koshland, D. E. The Key–Lock Theory and the Induced Fit Theory. *Angewandte Chemie International Edition in English* **1995**, *33* (2324), 2375–2378.
13. Durrant, J. D.; Mccammon, J. A. Molecular Dynamics Simulations and Drug Discovery. *BMC Biology* **2011**, *9* (1).
14. Muller, F. L.; Colla, S.; Aquilanti, E.; Manzo, V. E.; Genovese, G.; Lee, J.; Eisenson, D.; Narurkar, R.; Deng, P.; Nezi, L.; Lee, M. A.; Hu, B.; Hu, J.; Sahin, E.; Ong, D.; Fletcher-Sananikone, E.; Ho, D.; Kwong, L.; Brennan, C.; Wang, Y. A.; Chin, L.; Depinho, R. A. Passenger Deletions Generate Therapeutic Vulnerabilities in Cancer. *Nature* **2012**, *488* (7411), 337–342.
15. Leonard, P. G.; Satani, N.; Maxwell, D.; Lin, Y.-H.; Hammoudi, N.; Peng, Z.; Pisaneschi, F.; Link, T. M.; Lee, G. R.; Sun, D.; Prasad, B. A. B.; Francesco, M. E. D.; Czako, B.; Asara, J. M.; Wang, Y. A.; Bornmann, W.; Depinho, R. A.; Muller, F. L. SF2312 Is a Natural Phosphonate Inhibitor of Enolase. *Nature Chemical Biology* **2016**, *12* (12), 1053–1058.
16. Lin, Y.-H.; Satani, N.; Hammoudi, N.; Ackroyd, J. J.; Khadka, S.; Yan, V. C.; Georgiou, D. K.; Sun, Y.; Zielinski, R.; Tran, T.; Pando, S. C.; Wang, X.; Maxwell, D.; Peng, Z.; Pisaneschi, F.; Mandal, P.; Leonard, P. G.; Xu, Q.; Wu, Q.; Jiang, Y.; Czako, B.; Kang,

Z.; Asara, J. M.; Priebe, W.; Bornmann, W.; Marszalek, J. R.; Depinho, R. A.; Muller, F. L. Eradication of ENO1-deleted Glioblastoma through Collateral Lethality. **2018**. **doi: https://doi.org/10.1101/331538**

17. Gerlt, J. A.; Babbitt, P. C.; Jacobson, M. P.; Almo, S. C. Divergent Evolution in Enolase Superfamily: Strategies for Assigning Functions. *Journal of Biological Chemistry* **2011**, *287*(1), 29–34.
18. Nussinov, R.; Tsai, C. J., Allostery in disease and in drug discovery. *Cell* **2013**, *153* (2), 293-305.
19. Tsai, C. J.; Nussinov, R., A Unified View of "How Allostery Works". *Plos Computational Biology* **2014**, *10* (2).
20. Süel, G. M.; Lockless, S. W.; Wall, M. A.; Ranganathan, R. Evolutionarily Conserved Networks of Residues Mediate Allosteric Communication in Proteins. *Nature Structural Biology* **2002**, *10* (1), 59–69.
21. Macpherson, J. A.; Anastasiou, D. Allosteric Regulation of Metabolism in Cancer: Endogenous Mechanisms and Considerations for Drug Design. *Current Opinion in Biotechnology* **2017**, *48*, 102–110.
22. Fothergill-Gilmore, L. A. Evolution in Glycolysis. *Biochemical Society Transactions* **1987**, *15* (5), 993–995.
23. Fernie, A. R.; Carrari, F.; Sweetlove, L. J. Respiratory Metabolism: Glycolysis, the TCA Cycle and Mitochondrial Electron Transport. *Current Opinion in Plant Biology* **2004**, *7* (3), 254–261.
24. Kim, J.-W.; Dang, C. V. Multifaceted Roles of Glycolytic Enzymes. *Trends in Biochemical Sciences* **2005**, *30* (3), 142–150.

25. Green, H. J.; Hughson, R. L.; Orr, G. W.; Ranney, D. A. Interrelationship Between Anaerobic Threshold, Blood Lactate And Muscle Metabolites During Progressive Exercise. *Medicine & Science in Sports & Exercise* **1982**, *14* (2), 160.

26. Gladden, L. B. Lactic Acid: New Roles in a New Millennium. *Proceedings of the National Academy of Sciences* **2001**, *98* (2), 395–397.

27. Warburg, O. The Metabolism Of Tumors In The Body. *The Journal of General Physiology* **1927**, *8* (6), 519–530.

28. Heiden, M. G. V.; Cantley, L. C.; Thompson, C. B. Understanding the Warburg Effect: The Metabolic Requirements of Cell Proliferation. *Science* **2009**, *324* (5930), 1029–1033.

29. Liberti, M. V.; Locasale, J. W. Correction to: 'The Warburg Effect: How Does It Benefit Cancer Cells?' *Trends in Biochemical Sciences* **2016**, *41* (3), 287.

30. Lebioda, L.; Stec, B. Crystal Structure of Enolase Indicates That Enolase and Pyruvate Kinase Evolved from a Common Ancestor. *Nature* **1988**, *333* (6174), 683–686.

31. Creighton, T. E., Protein folding. *Biochem. J.* **1990**, 270 (1), 1-16

32. Cooper, G. M., *The Cell - A Molecular Approach 2nd Edition. Sunderland* (MA): Sinauer Associates: 2000.

33. Sugita, Y.; Okamoto, Y. Replica-Exchange Molecular Dynamics Method for Protein Folding. *Chemical Physics Letters* **1999**, *314* (1-2), 141–151.

34. Huo, S.; Wang, J.; Cieplak, P.; Kollman, P. A.; Kuntz, I. D. Molecular Dynamics and Free Energy Analyses of Cathepsin D−Inhibitor Interactions: Insight into Structure-Based Ligand Design. *Journal of Medicinal Chemistry* **2002**, *45* (7), 1412–1419.

35. Lundborg, M.; Narangifard, A.; Wennberg, C. L.; Lindahl, E.; Daneholt, B.; Norlén, L. Human Skin Barrier Structure and Function Analyzed by Cryo-EM and Molecular Dynamics Simulation. *Journal of Structural Biology* **2018**, *203* (2), 149–161.
36. Mackerell, A. D.; Feig, M.; Brooks, C. L. Extending the Treatment of Backbone Energetics in Protein Force Fields: Limitations of Gas-Phase Quantum Mechanics in Reproducing Protein Conformational Distributions in Molecular Dynamics Simulations. *Journal of Computational Chemistry* **2004**, *25* (11), 1400–1415.
37. Klein, M. L.; Shinoda, W. Large-Scale Molecular Dynamics Simulations of Self-Assembling Systems. *Science* **2008**, *321* (5890), 798–800.
38. Norman, G. E.; Stegailov, V. V. Stochastic Theory of the Classical Molecular Dynamics Method. *Mathematical Models and Computer Simulations* **2013**, *5* (4), 305–333.
39. Meller, J. a., Molecular Dynamics. In *eLS*, John Wiley & Sons, Ltd: 2001.
40. Nosé, S., An extension of the canonical ensemble molecular dynamics method. *Molecular Physics* **1986**, *57* (1), 187-191.
41. Grubmüller, H.; Heller, H.; Windemuth, A.; Schulten, K. Generalized Verlet Algorithm for Efficient Molecular Dynamics Simulations with Long-Range Interactions. *Molecular Simulation* **1991**, *6* (1-3), 121–142.
42. Mazur, A. K. Common Molecular Dynamics Algorithms Revisited: Accuracy and Optimal Time Steps of Störmer–Leapfrog Integrators. *Journal of Computational Physics* **1997**, *136* (2), 354–365.
43. Zhao, H.; Huang, D.; Caflisch, A. Discovery of Tyrosine Kinase Inhibitors by Docking into an Inactive Kinase Conformation Generated by Molecular Dynamics. *ChemMedChem* **2012**, *7* (11), 1983–1990.

44. Lu, S.; Huang, W.; Zhang, J. Recent Computational Advances in the Identification of Allosteric Sites in Proteins. *Drug Discovery Today* **2014**, *19* (10), 1595–1600.

45. Bernardi, R. C.; Melo, M. C.; Schulten, K. Enhanced Sampling Techniques in Molecular Dynamics Simulations of Biological Systems. *Biochimica et Biophysica Acta (BBA) - General Subjects* **2015**, *1850* (5), 872–877.

46. Scheerschmidt, K. Empirical Molecular Dynamics: Possibilities, Requirements, and Limitations. Topics in Applied Physics Theory of Defects in Semiconductors 213–244.

47. Tang, Y.; Suzuki, Y.-I.; Horio, T.; Suzuki, T. Molecular Frame Image Restoration and Partial Wave Analysis of Photoionization Dynamics of NO by Time-Energy Mapping of Photoelectron Angular Distribution. *Physical Review Letters* **2010**, *104* (7).

48. Hong, G.; Cornish, A.; Hegg, E.; Pachter, R. On Understanding Proton Transfer to the Biocatalytic [Fe—Fe]H Sub-Cluster in [Fe—Fe]H2ases: QM/MM MD Simulations. *Biochimica et Biophysica Acta (BBA) - Bioenergetics* **2011**, *1807* (5), 510–517.

49. Shaw, D. E. 166 Millisecond-Long Molecular Dynamics Simulations of Proteins on a Special-Purpose Machine. *Journal of Biomolecular Structure and Dynamics* **2013**, *31* (sup1), 108–108.

50. Popa, I.; Rivas-Pardo, J. A.; Eckels, E. C.; Echelman, D. J.; Badilla, C. L.; Valle-Orero, J.; Fernández, J. M. A HaloTag Anchored Ruler for Week-Long Studies of Protein Dynamics. *Journal of the American Chemical Society* **2016**, *138* (33), 10546–10553.

51. Hamelberg, D.; Mongan, J.; Mccammon, J. A. Accelerated Molecular Dynamics: A Promising and Efficient Simulation Method for Biomolecules. *The Journal of Chemical Physics* **2004**, *120* (24), 11919–11929.

52. Rhee, Y. M.; Pande, V. S. Multiplexed-Replica Exchange Molecular Dynamics Method for Protein Folding Simulation. *Biophysical Journal* **2003**, *84* (2), 775–786.

53. Kuzmanic, A.; Zagrovic, B. Determination of Ensemble-Average Pairwise Root Mean-Square Deviation from Experimental B-Factors. *Biophysical Journal* **2010**, *98* (5), 861–871.

54. Martínez, L. Automatic Identification of Mobile and Rigid Substructures in Molecular Dynamics Simulations and Fractional Structural Fluctuation Analysis. *Plos One* **2015**, *10* (3).

55. Amadei, A.; Linssen, A. B. M.; Berendsen, H. J. C. Essential Dynamics of Proteins. *Proteins: Structure, Function, and Genetics* **1993**, *17* (4), 412–425.

56. Papaleo, E.; Mereghetti, P.; Fantucci, P.; Grandori, R.; Gioia, L. D. Free-Energy Landscape, Principal Component Analysis, and Structural Clustering to Identify Representative Conformations from Molecular Dynamics Simulations: The Myoglobin Case. *Journal of Molecular Graphics and Modelling* **2009**, *27* (8), 889–899.

57. Doshi, U.; Holliday, M. J.; Eisenmesser, E. Z.; Hamelberg, D. Dynamical Network of Residue–Residue Contacts Reveals Coupled Allosteric Effects in Recognition, Catalysis, and Mutation. *Proceedings of the National Academy of Sciences* **2016**, *113* (17), 4735–4740.

58. Yao X-Q.; Momin M.; Hamelberg D. Elucidating Allosteric Communications in Proteins with Difference Contact Network Analysis. *Journal of Chemical Information and Modeling* **2018**, *58*, 1325-1330.

59. Rodriguez-Bussey, I.; Yao, X.-Q.; Shouaib, A. D.; Lopez, J.; Hamelberg, D. Decoding Allosteric Communication Pathways in Cyclophilin A with a Comparative Analysis of

Perturbed Conformational Ensembles. *The Journal of Physical Chemistry B* **2018**, *122* (25), 6528–6535.

60. Momin, M.; Yao, X.-Q.; Thor, W.; Hamelberg, D. Substrate Sequence Determines Catalytic Activities, Domain-Binding Preferences, and Allosteric Mechanisms in Pin1. *The Journal of Physical Chemistry B* **2018**, *122* (25), 6521–6527.
61. Vu, P. J.; Yao, X.-Q.; Momin, M.; Hamelberg, D. Unraveling Allosteric Mechanisms of Enzymatic Catalysis with an Evolutionary Analysis of Residue–Residue Contact Dynamical Changes. *ACS Catalysis* **2018**, *8* (3), 2375–2384.
62. Meng, X.-Y.; Zhang, H.-X.; Mezei, M.; Cui, M. Molecular Docking: A Powerful Approach for Structure-Based Drug Discovery. *Current Computer Aided-Drug Design* **2011**, *7* (2), 146–157.
63. Seeliger, D.; Groot, B. L. D. Ligand Docking and Binding Site Analysis with PyMOL and Autodock/Vina. *Journal of Computer-Aided Molecular Design* **2010**, *24* (5), 417–422.
64. Stelzer, A. C.; Frank, A. T.; Kratz, J. D.; Swanson, M. D.; Gonzalez-Hernandez, M. J.; Lee, J.; Andricioaei, I.; Markovitz, D. M.; Al-Hashimi, H. M. Discovery of Selective Bioactive Small Molecules by Targeting an RNA Dynamic Ensemble. *Nature Chemical Biology* **2011**, *7* (8), 553–559.
65. Trott, O.; Olson, A. J. AutoDock Vina: Improving the Speed and Accuracy of Docking with a New Scoring Function, Efficient Optimization, and Multithreading. *Journal of Computational Chemistry* **2009**.
66. Pollastri, M. P. Overview on the Rule of Five. *Current Protocols in Pharmacology* **2010**.

67. Irwin, J. J.; Shoichet, B. K. ZINC − A Free Database of Commercially Available Compounds for Virtual Screening. *Journal of Chemical Information and Modeling* **2005**, *45* (1), 177–182.

68. Cheng, L. S.; Amaro, R. E.; Xu, D.; Li, W. W.; Arzberger, P. W.; Mccammon, J. A. Ensemble-Based Virtual Screening Reveals Potential Novel Antiviral Compounds for Avian Influenza Neuraminidase. *Journal of Medicinal Chemistry* **2008**, *51* (13), 3878–3894.

69. Wold, F. 18 Enolase. *The Enzymes* **1971**, 499–538.

70. Pancholi, V. Multifunctional α-Enolase: Its Role in Diseases. *Cellular and Molecular Life Sciences* **2001**, *58* (7), 902–920.

71. Tracy, M. R.; Hedges, S. Evolutionary History of the Enolase Gene Family. *Gene* **2000**, *259* (1-2), 129–138.

72. Faller, L. D.; Baroudy, B. M.; Johnson, A. M.; Ewall, R. X. Magnesium Ion Requirements for Yeast Enolase Activity. *Biochemistry* **1977**, *16* (17), 3864–3869.

73. Zhang, E.; Hatada, M.; Brewer, J.; Lebioda, L. Catalytic Metal Ion Binding In Enolase: The Crystal Structure Of Enolase-Mn2 -Phosphonoacetohydroxamate Complex At 2.4 Angstroms Resolution. *Biochemistry* **1994**. *33,* 6295-6300.

74. Kang, H. J.; Jung, S.-K.; Kim, S. J.; Chung, S. J. Structure of Human α-Enolase (hENO1), a Multifunctional Glycolytic Enzyme. *Acta Crystallographica Section D Biological Crystallography* **2008**, *64* (6), 651–657.

75. Schreier, B.; Höcker Birte. Engineering the Enolase Magnesium II Binding Site: Implications for Its Evolution. *Biochemistry* **2010**, *49* (35), 7582–7589.

76. Qin, J.; Chai, G.; Brewer, J. M.; Lovelace, L. L.; Lebioda, L. Fluoride Inhibition of Enolase: Crystal Structure and Thermodynamics. *Biochemistry* **2006**, *45* (3), 793–800.

77. Dolinsky, T. J.; Nielsen, J. E.; Mccammon, J. A.; Baker, N. A. PDB2PQR: An Automated Pipeline for the Setup of Poisson-Boltzmann Electrostatics Calculations. *Nucleic Acids Research* **2004**, *32* (Web Server).

78. DeLano, W. L. *The PyMOL Molecular Graphics System*, 1.1r1; DeLano Scientific, Palo Alto, CA, **2002**.

79. D.A. Case, R.M. Betz, D.S. Cerutti, T.E. Cheatham, III, T.A. Darden, R.E. Duke, T.J. Giese, H. Gohlke, A.W. Goetz, N. Homeyer, S. Izadi, P. Janowski, J. Kaus, A. Kovalenko, T.S. Lee, S. LeGrand, P. Li, C. Lin, T. Luchko, R. Luo, B. Madej, D. Mermelstein, K.M. Merz, G. Monard, H. Nguyen, H.T. Nguyen, I. Omelyan, A. Onufriev, D.R. Roe, A. Roitberg, C. Sagui, C.L. Simmerling, W.M. Botello-Smith, J. Swails, R.C. Walker, J. Wang, R.M. Wolf, X. Wu, L. Xiao and P.A. Kollman. *AMBER 2016*, **2016**, University of California, San Francisco.

80. Maier, J. A.; Martinez, C.; Kasavajhala, K.; Wickstrom, L.; Hauser, K. E.; Simmerling, C. ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from ff99SB. *Journal of Chemical Theory and Computation* **2015**, *11* (8), 3696–3713.

81. Jorgensen, W. L. Revised Tips for Simulations of Liquid Water and Aqueous Solutions. *J. Chem. Phys*. **1982**, *77*, 4156−4163.

82. Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G. A Smooth Particle Mesh Ewald Method. *J. Chem. Phys.* **1995**, *103*, 8577−8593.

83. Humphrey, W.; Dalke, A.; Schulten, K. VMD: Visual Molecular Dynamics. *J. Mol. Graphics* **1996**, *14*, 33−38.

84. Stierand, K.; Rarey, M. PoseView -- Molecular Interaction Patterns at a Glance. *Journal of Cheminformatics* **2010**, *2* (S1).

85. Fujita, T.; Hansch, C. Analysis of the Structure-Activity Relationship of the Sulfonamide Drugs Using Substituent Constants. *Journal of Medicinal Chemistry* **1967**, *10* (6), 991–1000.

86. Llorente, P.; Marco, R.; Sols, A. Regulation of Liver Pyruvate Kinase and the Phosphoenolpyruvate Crossroads. *European Journal of Biochemistry* **1970**, *13* (1), 45-54.

87. Hall, E. R.; Cottam, G. L. Isozymes of Pyruvate Kinase in Vertebrates: Their Physical, Chemical, Kinetic and Immunological Properties. *International Journal of Biochemistry* **1978**, *9* (11), 785–794.

88. Tang, Q.; Fenton, A. W. Whole-protein alanine-scanning mutagenesis of allostery: A large percentage of a protein can contribute to mechanism. *Human Mutation* **2017**, *38* (9), 1132–1143.

89. Scrutton, M. C.; Utter, M.G. The Regulation of Glycolysis and Gluconeogenesis in Animal Tissues. *Annual Review of Biochemistry* **1968**, *37* (1), 249-302.

90. Fenton, A. W., and Hutchinson, M. The pH dependence of the allosteric response of human liver pyruvate kinase to fructose1,6-bisphosphate, ATP, and alanine. *Arch. Biochem. Biophys.* **2009**, *484*, 16−23.

91. Donovan, K. A., Zhu, S., Liuni, P., Peng, F., Kessans, S. A., Wilson, D. J., & Dobson, R. C. (2016). Conformational Dynamics and Allostery in Pyruvate Kinase. *The Journal of biological chemistry* **2016**, *291*(17), 9244–9256.

92. Fenton, A.; Holyoak, T.; Zhang, B.; Deng, J.; Tang, Q.; Prasannan, C. B. Energetic coupling between an oxidizable cysteine and the phosphorylatable N-terminus of human liver pyruvate kinase. *Biophysical Journal* **2013**, *104* (2).

93. Ishwar, A.; Tang, Q.; Fenton, A. W. Distinguishing the Interactions in the Fructose 1,6-Bisphosphate Binding Site of Human Liver Pyruvate Kinase That Contribute to Allostery. *Biochemistry* **2015**, *54* (7), 1516–1524.

94. Tang, Q.; Fenton, A. W. Whole-Protein Alanine-Scanning Mutagenesis of Allostery: A Large Percentage of a Protein Can Contribute to Mechanism. *Human Mutation* **2017**, *38* (9), 1132–1143.

95. Fenton, A.; Holyoak, T.; Zhang, B.; Deng, J.; Tang, Q.; Prasannan, C. B. Energetic Coupling between an Oxidizable Cysteine and the Phosphorylatable N-Terminus of Human Liver Pyruvate Kinase. *Biophysical Journal* **2013**, *104* (2).

96. Webb, B.; Sali, A. Comparative Protein Structure Modeling Using MODELLER. *Current Protocols in Bioinformatics* **2016**.

97. Zhong, W.; Morgan, H. P.; Mcnae, I. W.; Michels, P. A. M.; Fothergill-Gilmore, L. A.; Walkinshaw, M. D. `In Crystallo Substrate Binding Triggers Major Domain Movements and Reveals Magnesium as a Co-Activator Of Trypanosoma Bruceipyruvate Kinase. *Acta Crystallographica Section D Biological Crystallography* **2013**, *69* (9), 1768–1779.

98. Williams, R.; Holyoak, T.; Mcdonald, G.; Gui, C.; Fenton, A. W. Differentiating a Ligands Chemical Requirements for Allosteric Interactions from Those for Protein Binding. Phenylalanine Inhibition of Pyruvate Kinase†,‡. *Biochemistry* **2006**, *45* (17), 5421–5429.

99. Meagher, K. L.; Redman, L. T.; Carlson, H. A. Development of Polyphosphate Parameters for Use with the AMBER Force Field. *Journal of Computational Chemistry* **2003**, *24* (9), 1016–1025.

100. Dombrauckas, J. D.; Santarsiero, B. D.; Mesecar, A. D. Structural Basis for Tumor Pyruvate Kinase M2 Allosteric Regulation and Catalysis. *Biochemistry* **2005**, *44* (27), 9417–9429.