Investigations into the Non-Mevalonate Isoprenoid Biosynthesis Pathway's First Two

Enzymes utilizing Hybrid QM/MM Techniques

by

Justin K. White

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
Department of Chemistry
College of Arts and Sciences
University of South Florida

Major Professor: H. Lee Woodcock, Ph.D.
David J. Merkler, Ph.D.
Wayne C. Guida, Ph.D.
Yu Chen, Ph.D.

Date of Approval:
November 16, 2017

Keywords: QM/MM, computational, 1-deoxy-D-xylulose 5-phosphate synthase,
1-deoxy-D-xylulose 5-phosphate reductoisomerase

ProQuest Number: 10680472

ProQuest 10680472

Dedication

I dedicate this to my constant companion and greatest supporter, Molly Burges.

Acknowledgments

To my family, I would like to express my unending gratitude for the support shown me over the past many years. As I am slogged through this process and begun to confront personal issues, you have been along side me.

My friends of whom, I consider my extended family. You have thought better of me than I have. For that and your undying faith in my potential, I will be eternally grateful. You are always there when I need you, no matter how long it has been since we last spoke.

The love of my life already got the dedication but she deserves so much more. My new wife is always in my corner. Trying to look out for me, even when I don't care about what happens to me. She is my love and rock. She has been there through it all.

I am grateful to my Ph.D. committee of Lee Woodcock, David Merkler, Wayne Guida, and Chair Yu Chen for their guidance and patience over the years. Though I might have pushed the patience portion too much.

My labmates through the years: Christi Whittington, Sai Vankayala, Fiona Kearns, Phillip Hudson, Yura Pevzner, and Jackie Hargis. There might be others I am forgetting but still know that each of you have helped me along the path, if by no other way then being there for me to rant at for a while.

Lastly, I am grateful for the opportunities and guidance provided by Lee and Dave over these many years. I hope I haven't caused you too many causes for stress and know that I do truly appreciate every bit of guidance and patience along the way.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Abstract

Molecular drug design begins with the identification of a problem to solve. This work identifies the growing resistance among human pathogens to current treatments. Once the problem is identified and understood, solutions must be proposed. This one is straight forward, we need new antimicrobial drugs. More specifically, we need to identify novel targets to inhibit. A large portion of antibiotics focus on disruption of macromolecular production while only a few target metabolic systems. Finally, you need to propose solutions based on the information gathered. In order to avoid existing resistance, it is important to avoid the macromolecular route and focus on metabolic enzymes. Preferably, the pathway would have little overlap or similarity with pathways found in the treatment organism. With this in mind, the non-mevalonate (NMA) pathway poses as a very good target for drug design. Many pathogens have been found to be strictly dependent on this pathway while it is absent in humans. Additionally, fosmidomycin has already been shown to inhibit this pathway. Initially thought to just inhibit the 1-deoxy-D-xylulose 5-phosphate (DXP) reductoisomerase (DXR), it has been shown to inhibit several enzymes along the path to a lesser extent. Ideally, this could be repeated or improve upon for future drug design.

With this in mind, the initial stages of the first two enzymes of the NMA pathway were examined utilizing quantum mechanical/molecular mechanical (QM/MM) techniques. The first enzyme was DXP synthase (DXS), which catalyzes a transketolase-like condensation of pyruvate and glyceraldehyde-3-phosphate to produce DXP. DXS and other transketolases are dependent on the thiamine diphosphate (TDP) cofactor, which must be deprotonated of the imidazolium C2 atom producing a highly reactive ylide. A tautomerization occurs prior to this deprotonation to prime the pyrimidinium ring N4 atom to perform the C2 abstraction. The question at hand was the identity of a general base

to perform the N4 abstraction. The results favored a water-mediate mechanism with a higher than usual $\Delta E^{\ddagger}$ of 22.7 kcalcotmol^{-1}. An observation pertaining the tautomerization pertained to the aromaticity of the pyrimidine ring. Upon further investigation, aromaticity was found to play a significant role in the $\Delta E^{\ddagger}$ observed. Aromaticity might contribute 14.2 kcalcotmol^{-1} to the barrier height. This high energy would drive the reaction forward producing the ylide.

Investigation of the DXR enzyme followed this work. Initially, the work was going to focus on the 2 mechanisms proposed for activity, α-ketol rearrangement and retro-aldol/aldol mechanism. Subsequent publications involving secondary kinetic isotope effects (KIEs) add to the pile of evidence supporting the retro-aldol/aldol mechanism. So the project was retooled to investigate the energetic differences between two metal binding modes. The results of this work support a metal coordination across the C3-C4 bond, which eventually extends coordination to include the C2 oxygen. This conformation was help explain the tight binding effecting observation of the putative intermediates (transition states) and aldehyde intermediate. Additionally, as the C2-C3 mode consistently transfers a proton to the phosphate group of DXP or produces an elongated C-O bond, the C2-C3 mode would not be favorable.

Further investigations of these enzymes (e.g. completing the step begin, continuing through the reaction) could provide further illumination into the mechanism of action and possibly reveal new avenues of drug design. Examining the enzymes downstream in the NMA pathway might provide details of interest. Of particular interest is the radical reaction proposed for HDR/IspH. The final step of the pathway produces IDP and DMADP in a 4:1 proportion, which corresponds to the general system requirements for production of the long chain, branched isoprenoids. It would be interesting to compute the mechanism to see if energetics could provide further insights. Additionally, normal mode analysis coupled with vibrational subsystem analysis could identify allosteric sites for feedback sensitivity.

Chapter 1

Introduction

1.1 Isoprenoids

Accounting for nearly 60% of natural product diversity, isoprenoids (or terpenoids), with 55,000 known compounds, comprise the largest family of natural products[56,59]. Many of these compounds serve important biological functions. The lipid-soluble vitamins (A, D, E, And K) and cholesterol are some of the most common examples[161]. Cholesterol is subsequently utilized as a biosynthetic precursor of various steroid hormones, including glucocorticoids, estrogens and androgens. Some synthetic analogs are used in many therapeutic applications. All isoprenoids are derived from two phosphate C_5 isoprene building blocks, isopentenylallyl diphosphate (IDP) and dimethylallyl diphosphate (DMADP). The diversity of isoprenoids comes from number of IDP and DMADP molecules strung together. The string of isoprenes in combination with functional groups such as ketones, aldehydes, alcohols, peroxides, ethers and esters contribute to the considerable diversity amoungst the family of natural products[47].

The family is roughly divided into six major categories (Figure 1.1. These categories are based on the number of isoprene units linked. Monoterpenoids contain two isoprene units and therefore have a ten carbon skeleton. These are the major component of the fragrant oils from leaves, flowers and fruits (e,g, limonene and nerol). Sesquiterpenoids consist of three isoprene units to form 15 carbon cyclic and acyclic compounds. The next category is called diterpenoid and are composed of 20 carbon atoms derived from geranyl

Table 1.1: Examples of Biologically Significant Isoprenoids
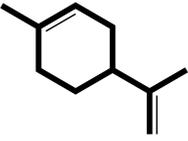
Structure	Description
Limonene	Cyclic terpene utilized for citrus flavor by the cosmetic industry
Abscisic Acid	Plant hormone required for development processes
Retinol	Required for vision and derived from vitamin A
Ubiquinone	Utilized as part of the electron transport chain for cellular respiration
Cholesterol	Precursor to bile acids, steroid hormones and vitamin D; required for cell membrane fluidity
Vitamin E	Antioxidant to reduce the production of reactive oxygen species formation during lipid oxidation
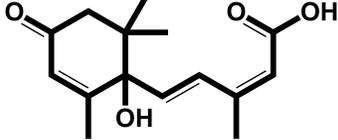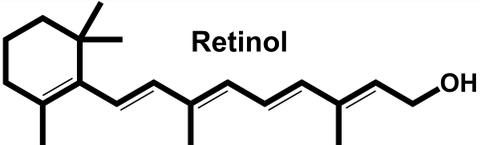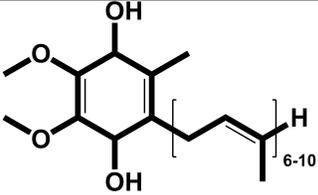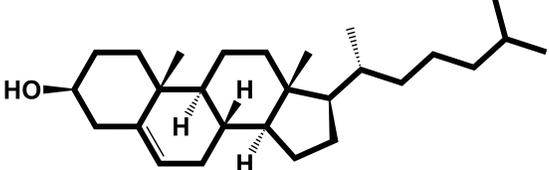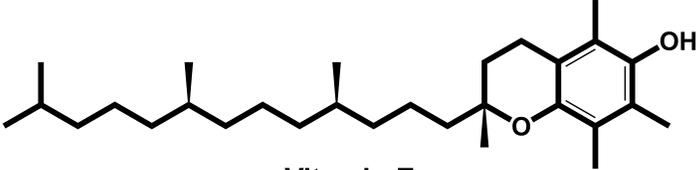
geranoil diphosphate (a 10 carbon unit). Vitamin A, phytohormone and tetrahydro-cannabinol are a few of the more well characterized examples of diterpenes. Ophiobolin A, a fungal metabolite, is an example of the next category, sesterterpenoids. These are derived from 25 carbon framework. Cholesterol are a member of the triterpenoids which are desired from the squalene precursor. Carotenoids are comprised of eight isoprene units to make forty carbon chains with conjugated double bonds. Carotenoids utilize the

absorption properties arising from their conjugated structures to assist in photosynthesis and the prevention of photo-oxidative cellular damage.

1.2 Pathways for Isoprenoid Biosynthesis

1.2.1 Mevalonic Acid Dependent Synthesis of Isoprenoid Building Blocks

The first pathway for isoprenoid biosynthesis was discovered based on interest in cholesterol for the obvious health related interests. During the investigation of cholesterol, researchers discovered deuterium integration originating from labeled acetate via the IDP unit which meant IDP was the direct precursor to cholesterol[16]. Subsequent studies lead to the discovery and characterization of the mevalonate (MVA) pathway in the 1950s named after the key intermediate (3R)-3,5-dihydroxy-3-methylpentanoic acid (mevalonic acid, MVA). For the following decades, the MVA pathway dominated this area of research as it was thought to be the sole route for IDP and DMADP synthesis in living systems. The work in this area lead to a nobel prize in physiology for Lynen and Bloch in 1964, and in Chemistry for Cornforth in 1975[15,40].

As figure 1.1 illustrates[130], the initial step of the MVA pathway is the production of acetoacetyl-CoA via the condensation of two acetyl-CoA molecules catalyzed by acetyl-CoA acetyltransferase. A third acetyl-CoA molecule is attached to the acetoacetyl-CoA to produce 3-hydroxy-3-methylglutaryl-CoA (HMG-CoA) catalyzed via an aldol reaction performed by HMG-CoA synthase. HMG-CoA is reduced by two equivalents of NADPH performed by HMG-CoA reductase producing MVA. The HMG-CoA reduction is the rate-limiting step of this pathway, thus MVA production is the rate-limiting, or key, intermediary step in the pathway. Two consecutive phosphorylations performed by mevalonate kinase and phosphomevalonate kinase produces mevalonate-5-diphosphate (MDP). An ATP-coupled decarboxylation catalyzed by MDP decarboxylase yields IDP[130]. The IDP isomer, DMADP, is produced via two structurally unrelated IDP:DMADP isomerases
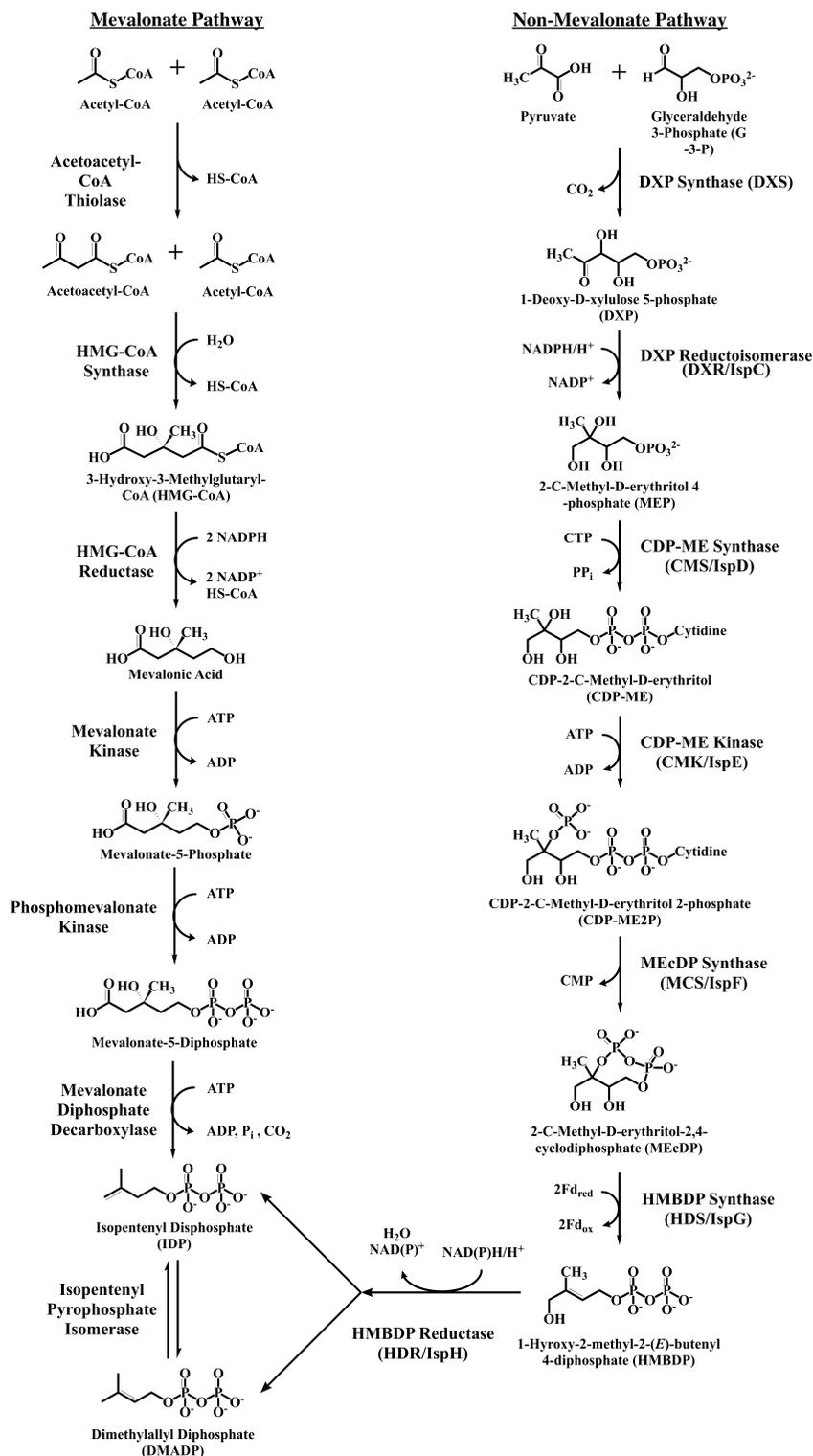
Figure 1.1: An illustration of the complete Mevalonate and Non-Mevalonate pathways culminating in the production of IDP and DMADP.

(Figure 1.1).

1.2.2 Mevalonate-Independent Synthesis of IDP and DMADP

Decades following the MVA pathway discovery, further isotopic labeling studies were performed to trace the fate acetate in producing lycopene, hopanoids, taxol and sterols. The distribution of the labeled material in the subsequent terpenoids was inconsistent with a single source of isoprenic production. Following from these results, several independent research groups discovered a mevalonate-independent (or non-mevalonate, NMA) pathway in eubacteria, green algae, and higher plants. Rohmer et al. identified the conversion of pyruvate to 1-deoxy-D-xylulose 5-phosphate (DXP) as the first step of the NMA pathway by following the incorporation of ^{13}C-labeled pyruvate or glycerol into ubiquinone. Additional independent studies carried out by Arigoni et al. traced the incorporation of [1-^{13}C]- and [2,3,4,5-^{13}C_4]-DXP into the formation of β-carotene, lutein, phytol and sitosterol in cell cultures of *Catharanthus roseus* and demonstrated the involvement of DXP in the NMA pathway. The studies conducted by Arigoni et al. provided further insight into the compartmentalization of isoprenoid synthesis, as well as, the description of the rearrangement proposed by Eisenreich et al. In order to elucidate the NMA pathway (Figure 1.1, further labelling studies were performed and revealed the pathway to be composed of 7 enzymes that catalyze 8 reactions. Background on each of the enzymes found in the NMA pathway can be found in the following paragraphs.

1-Deoxy-D-xylulose 5-Phosphate Synthase (DXS)

The aforementioned conversion of pyruvate to DXP is performed by DXP synthase (DXS) and employs glyceraldehyde-3-phosphate (G3P) (Figure 1.2). DXS is a member a large family of enzymes dependent upon the cofactor thiamine diphosphate (TDP)[176]. Particularly, the structure and reaction catalyzed are highly reminiscent of members of the

subfamily of transketolases (TKs). Structurally, DXS is similar in composition to other members of the subfamily. Each monomer is composed of 3 subunits (I, II and III). In solution, DXS is more commonly found in a homodimer that is functionally significant as each active site communicates with the other. The formation of the active site pocket distinguishes DXS from other members of the TK subfamily. The active site in other TKs exists between domain I of one monomer and domain II of the other monomer II in a twisted conformation that arises via formation of the homodimer. In contrast, the DXS active site resides in a pocket between domains I and II of the same[197]. Despite this distinction, several key residues remained highly conserved with the rest of the subfamily.

Mechanistically speaking, DXS corresponds to the α-hydroxyketone (acyloin) condensation and proceeds via a mechanism highly analogous to other TKs. Prior to substrate binding, the TDP cofactor undergoes a deprotonation of the thiazolium ring forming a carbanion ylide necessary for enzymatic. The usually high pK_a of the thiazolium proton makes this reaction highly unlikely in solution[84]. The active site of DXS and other TKs binds the TDP into a energetically strained 'V'-shaped conformation; which brings the N4 of the pyrimidine ring into proximity of the C2 thiazolium proton[197]. The proximity of nitrogen and strained structure lowers the pK_a significantly[84,125]. Therefore, the carbanion ylide is free to perform a nucleophilic attack on the pyruvate. Subsequently, the electrophilic iminium acts as an electron sink during decarboxylation and results in a carbanion/enamine. The enamine performs a second nucleophilic attack on carbonyl carbon of glyceraldehyde-3-phosphate (G3P). A final deprotonation step releases the new
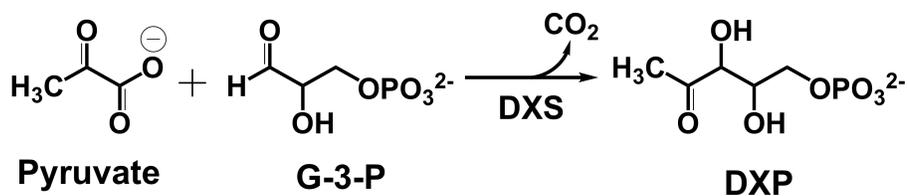
Figure 1.2: The abridged representation of the DXS catalyzed condensation of pyruvate with glyceraldehyde-3-phosphate to produce 1-deoxy-D-xylulose 5-phosphate.

DXP molecule and regenerates the TDP ylide for further catalysis. The substrate binding mechanism of other TKs has been thought to work through either a ping-pong or sequential mechanism[60]. Recent studies of DXS have suggested a random sequential mechanism thus illustrating a further distinction between DXS and other TKs. DXS does show a preferred order involving the formation of the C2α-lactylthiamin diphosphate (LTDP) intermediate as an unusually stable ternary complex of TDP and pyruvate. The hydroxyaldehyde moiety of G3P was found to trigger and accelerate the decarboxylation which produces the enamine utilized in the second nucleophilic reaction[21].

1-Deoxy-D-xylulose 5-Phosphate Reductisomerase (DXR)

The next step in the NMA pathway is actually the first committed step of the pathway. DXP reductoisomerase (DXR) catalyzes the conversion DXP into 2-C-methyl-D-erythritol 4-phosphate (MEP) with preferential dependence on NADPH as reducing agent and Mn^{2+} as a divalent ion[3,94,118,157,179]. The carbon-skeleton rearrangement in this reaction is thought to proceed via the aldehyde intermediate, 2-C-methyl-D-erythrose 4-phosphate (MEsP), which is subsequently reduced by NADPH (Figure 1.3). The idea behind the aldehyde intermediate arose due to similarities between DXR and ketol-acid reductoisomerase (KARI; EC 1.1.1.86); which catalyzes a similar rearrangement-reduction sequence in the conversion of 2-acetolactate to 2,3-dihydroxy-3-methylbutyrate[132]. In both situations, the intermediate has never been directly observed as it is either tightly bound prior to reduction or in such low concentration due to it's transient nature. An experiment by Rohmer and co-workers provided the strongest evidence in support of the MEsP intermediate. The researchers synthesized MEsP and demonstrated its kinetic competency with DXR in the presence of NADPH and Mn^{2+} or Mg^{2+}. They also observed a 7% conversion of MEsP to DXP in the presence of $NADP^+$[37].
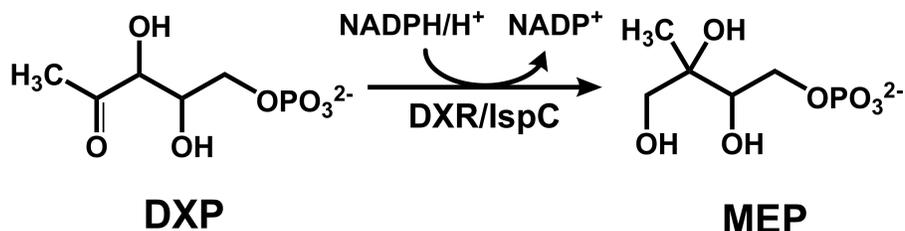
Figure 1.3: The abridged representation of the DXR catalyzed rearrangement coupled reduction of 1-deoxy-D-xylulose 5-phosphate by NADPH to produce 2-C-methyl-D-erythritol 4-phosphate.

Similar to DXS, DXR is most commonly found in a homodimer of the V-shaped monomers producing a saddle-like quaternary structure. The monomers can be further subdivided into three distinct domains. A dinucleotide-binding domain acts as a binding site for the NADPH cofactor is found in the N-terminal region of each monomer. The central domain harbors the catalytic portion of the enzyme and is responsible for the crucial conformational changes required for substrate binding and turnover. The central domain contains a highly flexible loop, which acts as a lid upon substrate binding creating a protected active-site cavity. The final domain is the C-terminal domain consisting of a four-helix bundle and is characterized by its flexibility. The role of this flexible domain is to aid in dimerization[118,151,199,200].

Despite the similarities between DXR and KARI, differences in amino acid sequences and crystal structures suggest different mechanisms for DXR[37,49,95]. Three mechanisms were originally considered for DXR's rearrangement of DXP to MEsP: 1) an α-ketol rearrangement, 2) a retro-aldolization/aldolization and 3) a sequential 1,2-hydride and 1,2-methyl shift[64]. This last proposal was readily eliminated when the ^{13}C-glucose incorporation studies failed to yield the appropriate labeled products. More over, 2-^{13}C and 3,4,5-^{13}C$_3$ labeled DXP experiments strictly yielded [2-^{13}C]- and [1,3,4-^{13}C$_3$]-MEP, respectively, which supports the rejection of the sequential shift mechanism[3,73,160]. The α-ketol (sigmatropic) rearrangement occurs via the migration of the C3-C4 bond to form a C2-C4 bond in order to form MEsP, the aldehyde intermediate. This migration re-

quires a partial positive charge on the C2 atom which can be achieved via protonation or the ketol coordinating with the divalent metal ion. There is evidence to support the latter approach. The metal has been shown to be chelated by the hydroxy groups of DXP. A deprotonation of the C3 hydroxyl group of DXP is required for aldehyde formation Thus, the deprotonation and bond-cleavage/-formation would result in the MEsP[78]. Alternatively, the retro-aldol/aldol reaction begins with the deprotonation of the C4 hydroxyl group followed by cleavage of the C3-C4 bond in a retro-aldolization. The result of this reaction is the formation of a hydroxyacetone enolate and glycoaldehyde phosphate. Subsequently, an aldol condensation will result in the same MEsP intermediate[80,95,109]. Kinetic isotope effects (KIEs) measurements are not compatible with the α-ketol rearrangement mechanism. The hydroxyacetone and glycoaldehyde phosphate intermediates have not been directly observed. Neither have they been successfully incorporated when incubated with DXR and the necessary cofactors. These conflicting results indicate there is further work to be done on the DXR mechanism. Despite the rearrngement mechanism, the subsequent reduction produces MEP in the same way. Deuterium-labeled NADPH and crystal structures have revealed details of the reduction reaction. The *Re* face of MEsP protonated by the *pro-S* hydrogen from the nicotinamide C4 to the C1 of the aldehyde[5,132].

4-Diphosphocytidyl-2C-methyl-D-erythritol Synthase (CMS/IspD)

The discovery of DXS and DXR allowed for the identification and characterization of additional NMA enzymes in quick succession. Utilizing [2-^{14}C]-labeled MEP in *E. coli*, research groups were able to track the formation of new radioactive products[155]. The first product identified was 4-diphosphocytidyl-2C-methyl-D-erythritol (CDP-ME) in nuclear magnetic resonance (NMR) spectroscopic assays (Figure 1.4). A database search of corresponding enzymatic activity pointed researchers towards the *ygbP* gene[101]. Subsequently several experiments showed MEP turnover and IPP production are dependent upon *ygbP*

and several other open reading frames[101,155]. The distribution of the new protein correlated well with the expected presence of the NMA pathway in eubacteria. Following the confirmation of ygbP, now designated IspD (CMS), activity assays revealed the necessity of a divalent cation (Mn^{2+}, Mg^{2+}, or Co^{2+}) with a preference for Mg^{2+}. CMS appears to be substrate specific with low but measurable activity with GTP and ATP. The incorporation of CTP's α-phosphate instead of the β- or γ-phosphates was confirmed utilizing radioactive phosphorous isotopes[101,155].
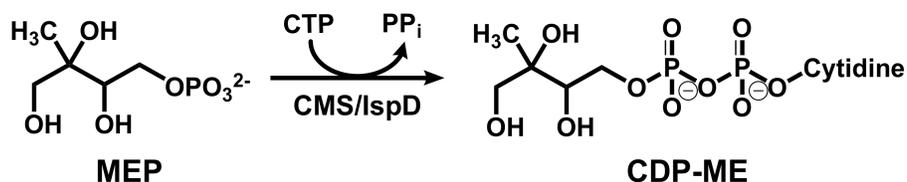
Figure 1.4: The abridged representation of reaction catalyzed by CMS/IspD which attaches a cytidyl group to the phosphate tail of 1-deoxy-D-xylulose 5-phosphate producing 4-diphosphocytidyl-2C-methyl-D-erythritol.

Several crystal structures of CMS have been solved from a variety of organisms[62,88,152]. To continue the trend of the first two enzymes, CMS is found commonly as a homodimer. Each structure revealed strong overall structural conservation with each other and other nucleoside-binding proteins, particularly cytidyltransferases[152]. The domain of each monomer hold a certain characteristic inline with other homologues. One of the domains consists of a so-called β-arm, composed of overlapping parallel and anti-parallel β-strands, which protrudes from the main globular domain at a wide angle. This β-arm acts as a hook-like structure that interlocks closely with another monomer which aids in dimerization[62,88,152]. The tertiary structures with all necessary cofactors provided valuable insights into ligand binding and enzymatic activity. A large network of hydrogen-bonding interactions between ligands and side chains as well as backbone carbonyl and amide groups revealed that both the substrate and products are fixed to the active site. Direct interaction between the protein and cytosine moiety of CTP, in part, explains the selectivity and preference for pyrimidines over purine nucleosides[152]. Basic

residues are proposed to position the triphosphate tail for catalysis. Additionally, these residues might play a role in stabilizing the pentacoordinate transition state during the reaction. Additional phosphate coordinations of MEP and CTP is provided by the Mg^{2+} despite its lack of direct interactions with the enzyme[62,88].

The crystal structures of CMS has lead to the proposal of 2 reaction mechanisms. One proposal involves the formation of a reactive metaphosphate CMP molecule via elimination of a disphosphate group. The metaphosphate CMP is subsequently attacked by the 4-phosphate of MEP to form CDP-ME. The alternative mechanism proposed starts with a direct nucleophilic attack on the α-phosphate of CTP by the 4-phosphate of MEP. The collapse of the pentacoordinate intermediate produces CDP-ME and PPi. Current mutagenesis and structural data favor the second mechanism over the first[152,154].

4-Diphosphocytidyl-2C-methyl-D-erythritol Kinase (CMK/IspE)

The expanding knowledge of the first three enzymes continued to facilitate the discovery of the next downstream catalyst. Genomic analyses showed *E. Coli ychB* and its orthologous sequences showed similar patterns in eubacteria and plants as other NMA genes. Overexpression, purification, and characterization of *ychB* revealed the production of 4-diphosphocytidyl-2C-methyl-D-erythritol 2-phosphate (CDP-MEP) from CDP-ME (Figure 1.5); which corresponds to a phosphorylation of the C2 hydroxy group[117]. Subsequently, the reaction and structure of ychB, later designated CMK or IspE, resembles those catalyzed by the GHMP superfamily of enzymes. In addition to galactokinases and homoserine kinases, two enzymes of the MVA pathway, mevalonate and phosphomevalonate kinases, defines the enzymes of GHMP superfamily[117,185].

Following from the observed sequential conservation, CMK homologues strongly resemble each other as well as other members of the GHMP superfamily. One distinction
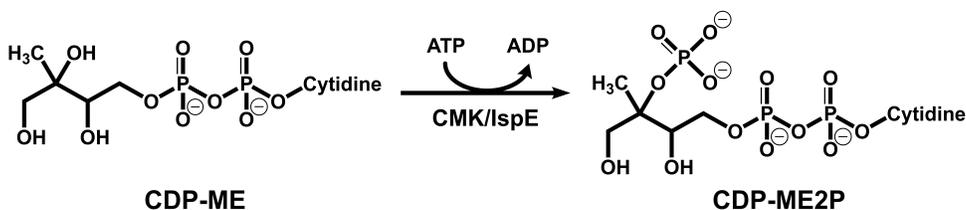
Figure 1.5: The abridged representation of the phosphorylation of 4-diphosphocytidyl-2C-methyl-D-erythritol by CMK/IspE producing 4-diphosphocytidyl-2C-methyl-D-erythritol 2-phosphate.

between CMK and other GHMP enzymes, CMK is commonly found in a monomeric state where as GHMP family members are commonly found in a homodimeric structures[85,126,168,185]. Each CMK monomer consists of predominantly 2 domains. An N-terminal domain responsible for cofactor binding and a C-terminal domain in charge of substate binding. CMK has an overall clamshell-like shape and the catalytic center is formed in an open cavity between the domains after the clamshell closes. This closure brings the substrate and cofactor into proximity in order to promote phosphorylation of CDP-ME[126,168,169,185].

Based on similarities to GHMP kinases, a catalytic mechanism was proposed. Not on direct observations of the actions of the actual enzyme[61,97]. The C2 hydroxyl group of CDP-ME forms hydrogen bonds with the side chains and carboxyl groups of highly conserved lysine and aspartate residues. This network helps to further polarize the hydroxyl group[126,168,185]. Due to this polarization, one of the aspartate residues can act as a base to deprotonate the hydroxyl group. The resulting reactive alkoxide undergoes nucleotide attack of the γ-phosphate of ATP resulting in a similar pentacoordinate intermediate for CMS/IspD. The subsequent collapse of the intermediate results in ADP and CDP-MEP being released with turnover[126,168,185]. A divalent metal ion is required for catalytic activity similar to other GHMP family members. The ion is responsible for positioning and orienting the phosphate moiety in proximity for attack by the acceptor molecule. Additionally, it stabilizes the pentavalent transition state the bond between the β- and

γ-phosphates of ATP[34,61,70,97]. Though, no CMK crystal structure has been observed to contain the metal ion and lack of a highly conserved glutamate residue involved in positioning the metal ion suggests unique role in CMK. Additionally, some have proposed coordinated water molecules might act in place of the metal ion in certain kinases. The exact role of the metal remains to be determined[34,126,168,185].

2C-Methyl-D-erythritol-2,4-cyclodiphosphate Synthase (MCS/IspF)

When CMS was identified as the third enzyme of the NMA pathway, the *ygbP* gene expression was found coupled to another unannotated *ygbB* gene sequence with a few cases even fused inside a single open reading frame[75,155]. Correspondingly, the species distributions of the gene orthologues parallel the presence of NMA based isoprenoid biosynthesis. Attempts to identify the activity of the corresponding protein was determined by challenging the enzyme with CDP-ME and CDP-MEP which produced 2-C-methyl-D-erythritol-3,4-cyclophosphate (MEcP) and 2-C-methyl-D-erythritol-2,4-cyclodiphosphate (MEcDP), respectively[75,181]. MEcDP was found to have been previously detected as a bacterial metabolite. These results support MEcDP as a new intermediate between DXP and IDP or DMADP (Figure 1.6); while MEcP was regarded as an in vitro artifact with no physiological relevance. Subsequently, the name of the enzyme was changed MEcDP synthase or IspF to reflect its function and position in the pathway[75].
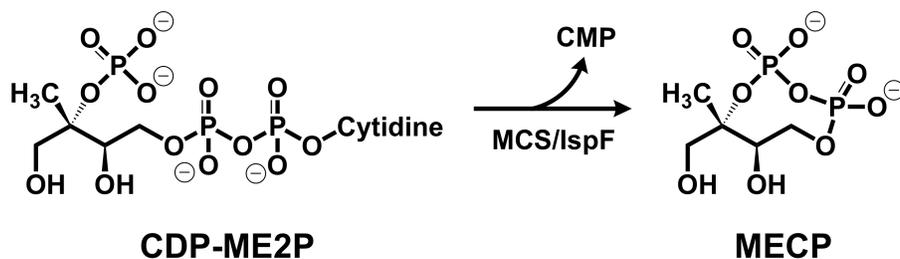
Figure 1.6: The cyclization of 4-diphosphocytidyl-2C-methyl-D-erythritol 2-phosphate and subsequent expulsion of cytidyl release catalyzed by MCS/IspF.

Structural characterizations have been published for the *E. coli*, *Plasmodium falciparum*, *Plasmodium vivax*, and *A. thaliana* MCS enzymes[28,87,139,153,177]. While there are differences in the spacial grouping and composition of the asymmetric unit, the structures all showed the formation of a tightly associated homotrimer. The multimeric assembly buries a large surface area that contributes to the MCS enzyme's heightened stability even in the face of mass spectrometric analysis. This trimeric structural feature is a common feature observed between MCS and any of its wider structural or functional homologues[87,153]. A series of anti-parallel β-sheets form a largely hydrophobic channel at the core of the trimer. This channel is thought to play a role in feedback regulation[87,89,153,177]. The active site is found at the interface of two subunits with both side chains contributing to the catalytic center. Conformational stabilization of the substrate and intermediates is accomplished via interaction with a few key amino acids, and two distinct essential metals, a Zn^{2+} and either a Mg^{2+} or Mn^{2+}. A zinc ion responsible for positioning the cytidyl moiety of the substrate; which itself is tetrahedrally coordinated by an aspartate and two histidine residues as well as the β-phosphate of MEcDP. Both the α- and β-phosphates of the CDP substructure is coordinated and stabilized by either a Mg^{2+} or Mn^{2+}. These phosphate groups additionally play a role in the octahedral coordination of the Mg^{2+} or Mn^{2+} ions with three water molecules and a glutamate residue filling in the rest of the coordination sites[87,153,177].

The intramolecular cyclization of CDP-MEP to MEcDP and concomitant CMP release catalyzed by MCS is thought to proceed via an in-line mechanism. Analogous to the previous two enzymes, the reaction involes the nucleophilic attack on a phosphate moiety thus forming a pentacoordinated transition state. The subsequent collapse of the transition state releases the two products, CMP and MEcDP[87,153,177]. The protective flexible loop closes off the catalytic cavity from the surrounding solvent. Interactions with CDP and MEP substructures via hydrogen-bonding and hydrophobic regions of the cavity accountants for the high degree of selectivity of MCS. The diphosphocytidyl moiety align-

ment is accomplished primarily by the active site metal ions with additional help from hydrogen-bonding and hydrophobic interactions of active site residues. Of particular interest for reactivity, the Zn^{2+} ion increases the electrophilic character of the β-phosphate in addition to aiding in lining up the nucleophilic attack by the 2-phosphate of the MEP moiety. Additionally, the enzyme restricts the flexibility of the substrate bringing the electron donor and acceptor in close proximity[153]. The negative charge of the cyclic transition state is compensated by the positive charge of the 2 metal ions[87,153,177].

1-Hydroxy-2-methyl-2-(E)-butenyl 4-diphosphate Synthase (HDS/IspG)

Unlike the previous upstream catalysts were discovered and characterized in relatively quick succession, the final two steps proved more elusive. The unannotated *gcpE* gene was originally discovered in association with a histidyl tRNA synthetase[59]. Utilizing bioinformatic approaches, an association was observed between the *gcpE* gene and other NMA pathway enzymes were observed to reflect the characteristic distribution patterns. As with DXS, DXR and IspD-IspF, the new gene sequence was detected in various bacterial species, plants and apicomplexa, but not in eukaryotes such as yeast[2,29]. Disruption of this gene was lethal as with the other NMA genes. Isotopic labeling coupled with NMR analysis identified the new intermediate as 1-hydroxy-2-methyl-2-(E)-butenyl 4-phosphate (HMBDP) (Figure 1.7), which can be produced via a reductive deoxygenation of MEcDP[71]. The gene was renamed to IspG to reflect the new position in the NMA pathway. Subsequent, recombinant expression and purification was straightforward, observed activities were low. The presence of three highly conserved cysteines and similarities with sequence motifs of ferrodoxin and aconitase enzymes suggested the presence of a catalytic iron-sulfur, [4Fe4S][71,192]. Further assays performed under anaerobic conditions and the presence of an reducing agent for regenerative purposes resulted in the efficient production of HMBDP from MEcDP. Additional confirmation of the presence of the cluster was UV-vis absorption spectrum, which matched previously observed spectrum of simi-
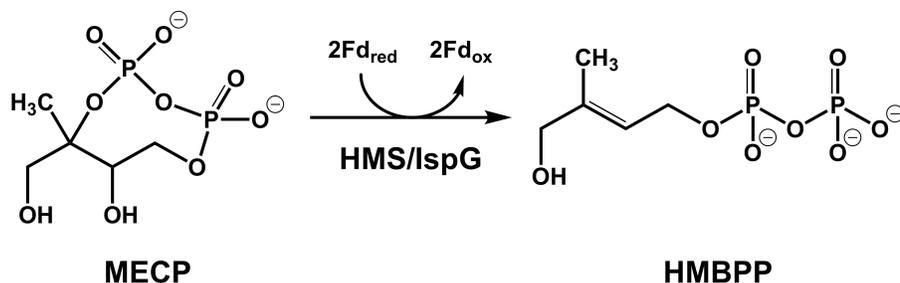
Figure 1.7: The penultimate step of the NMA pathway producing 1-hydroxy-2-methyl-2-(E)-butenyl 4-phosphate from 2-C-methyl-D-erythritol-2,4-cyclodiphosphate.

lar proteins[92,166].

As with DXS, DXR and CMS, HDS is commonly found in a homodimer. Each monomer is composed of two domains[106]. The N-terminal domain is an 8-stranded β-barrel globular subunit similar to the common $(\beta\alpha)_8$-fold of the triose phosphate isomerase (TIM) barrel superfamily. The iron-sulfur cluster is found in the C-terminal domain. Coordination of the cluster is supplied via 3 cysteine residues and a glutamate residue[106,149,150]. MEcDP binds in the active formed between the C-domain of a monomer and the N-domain of the other monomer. All published HDS crystal structures are highly similar; particularly with respect to the N-domain[106]. The C-domains are absent in some structures (probably due to lack of resolution), while the *Plasmidium falciparum* structure has an additional domain. This additional domain is thought to fold into a second TIM barrel to allow for monomeric activity[113,204].

Mechanistic details remained elusive despite the identification of the substrate, product and resolution of several crsytal structures. The reaction was known to involve the [4Fe4S] cluster, elimination of the C3 hydroxyl and a 2 electron reduction. Results of isotopic-exchange experiments, electron paramegnetic resonance (EPR) spectroscopy, and many other experiments permitted the description of the HDS mechanism[23,92,156,166,186,187]. Upon binding MEcDP, a conformational closure causes the displacement of the glutamate residue from the fourth iron while promoting the formation of a covalent bond with the substrate. A deprotonation of the C3 hydroxyl group by a

second glutamate assists in the Fe-O bond formation[106,150,204]. Once bond, the ring of MEcDP opens and closes consistently[198]. The introduction of the first single external electron breaks the ring permenantly either a carbocation or radical (formed via internal electron transfer) and begins the reaction in ernest. The addition of a second external electron produces a C2 carbanion[146,186,187]. Formation and release of HMBDP proceeds via an E_{1cb} elimination results. A localized proton relay change results in the release of H_2O from the cluster and regeneration of the enzyme[146].

1.3   Isoprenoids as Drug Target

Molecular medicine has provided means for mankind to overcome many diseases caused by various microbial life forms. Lately, there has been a growing resistance to current therapies. The discovery of multi-drug resistance forms of many diseases (e.g. tuberculosis and malaria) are poised to return us to the time prior to anti-microbial drugs[96,122,165]. In an age of growing drug resistance, there are very few companies investing in developing novel treatments due to the low returns and high upfront costs. Amongst the now growing list of neglected disease, Tropical diseases (i.e. malaria, leishmaniasis, tuberculosis) represented the most neglected diseases in the world. This due in large part to their concentration in the developing nations. Malaria is one of the most profound problems due to its high morbidity rate and millions of reported cases a year. In the age of modern globalization, growing resistance is a world wide problem[122,183].

Malaria presence a growing international concern. High mortality aside, malaria can cause economic downturns in high areas of infection due to an acutely infected individuals inability to work. Long term effects can be seen by life-long learning impairments caused when children are infected. Malaria is caused by four species of *Plasmodium* but the majority of the mortalities are cause by two of them, *P. falciparum* and *P. vivax*[24,183]. Both show evidence of a growing resistance to long standing therapies such as chloroquine and

fansidar; which has hastened the need to develop novel treatments.

Primarily, drug resistance has been found via mutations in enzymes which reduce the inhibitive effects of the therapy. Additional mutations found in transporter proteins make up a large portion of the remaining resistance[196]. These transporters (i.e. *pfmdr1*) act by removing the drug from the target sites. This is similar to the developing of β-lactamases in infectious bacteria to destroy antibiotics such as penicillin while developing mutations in the target peptidase enzymes. It is a two fold development in resistance.

The NMA pathway has great promise as a target for anti-microbial activity. The most significant benefit is the seeming absence of orthologous enzymes in mammalian cells. Particular absence in humans is a huge benefit. In other words, the entire pathway seems to be absent in humans which rely on the MVA pathway for IDP and DMADP production. The combination of the completion of the human genome project, subsequent expansion in mapping other species genomes, and the identification of the genes and enzymes of the NMA pathway allowed researchers to perform scans. Table 1.2 represents highlights of species and the isoprenoids biosynthesis pathways present. There are a few species that rely on both but have one isolated in an organelle, therefore the products of one of the pathways aren't readily available to use in a crisis[59,63].

As indicated in table 1.2, several protozoal genomes (e.g. *P. falciparum*, and *P. vivax*) have genes corresponding to the NMA pathway. Subsequent studies have revealed these genes to be predominantly located in the apicoplast; which is necessary for survival in the intraerythrocytic and intrahepatic stages of *plasmodium*. The inhibition of the NMA pathway via fosmidomycin can only be save via exogenous introduction of IDP and DMADP suggests this pathway as a new source of anti-malarial drugs[59,201].

Additionally, there is evidence of NMA being a good source of anti-bacterial drugs. A majority of current antibiotics work via the interruption of the biosynthesis of macro-molecular components (e.g. DNA, RNA, cell wall) of the bacterium[6,52]. The remaining

Organism	Non-Mevalonate Pathway							Mevalonate Pathway						
	dxs	ispC	ispD	ispE	ispF	ispG	ispH	hmgs	hmgr	mk	pmk	dpmd	idiI	idiII
Bacteria														
Aquifales (Aquifex aeolicus)	+	+	+	+	+	+	+	−	−	−	−	−	−	−
Chlamydia group (Chlamydophila pneumoniae)	+	+	+	+	+	+	+	−	−	−	−	−	−	−
Cyanobacteria (Synechocystis sp.)	+	+	+	+	+	+	+	−	−	−	−	−	−	+
Deinococcus group (Deinococcus radiodurans)	+	+	+	+	+	+	+	−	−	−	−	−	−	+
Firmicutes														
(Bacillus subtilis)	+	+	+	+	+	+	+	−	−	−	−	−	−	+
(Mycoplasma genitalium)	−	−	−	−	−	−	−	−	−	−	−	−	−	−
(Staphyloccus aureus)	−	−	−	−	−	−	−	+	+	+	+	+	−	+
(Streptomyces coelicolor)	+	+	+	+	+	+	+	−	+	+	+	−	−	−
Proteobacteria														
(Escherichia coli)	+	+	+	+	+	+	+	−	−	−	−	−	+	−
(Rickettsia prowazeckii)	−	−	−	−	−	−	−	−	−	−	−	−	−	+
Spirochaetales														
(Treponema pallidum)	+	+	+	+	+	+	+	−	−	−	−	−	−	−
(Borrelia burgdorferi)	−	−	−	−	−	−	−	+	+	+	+	+	−	+
Thermotogales (Thermotoga maritima)	+	+	+	+	+	+	+	−	−	−	−	−	−	−
Archaea														
Crenarchaeota (Aeropyrum pernix)	−	−	−	−	−	−	−	+	+	+	+	+	−	+
Euryarchaeota (Archaeoglobus fulgidus)	−	−	−	−	−	−	−	+	+	+	+	+	−	+
Eukaryotes														
Animals (Homo sapiens)	−	−	−	−	−	−	−	+	+	+	+	+	+	−
Plants (Arabidopsis thaliana)	+	+	+	+	+	+	+	+	+	+	+	+	+	−
Protozoa (Plasmodium falciparum)	+	+	+	+	+	+	+	−	−	−	−	−	−	−
Yeasts (Saccharomyces cerevisae)	−	−	−	−	−	−	−	+	+	+	+	+	+	−

Table 1.2: Highlights the existence of genes found for both the Mevalonate and Non-Mevalonate Pathways in a variety of life forms.

few target metabolic enzymes. Arigoni et al. conducted a study using bioinformatics to identify 30 *E. coli* necessary for survival which could also be found in other bacterial species. The NMA pathway is amongst these necessary genes[6]. Several pathogenic bacteria, including *E. coli* and *Mycobacterium tuberculosis*, carrying deletions in the NMA genes could only be rescued via exogenous introduction of isoprenoids[55,122]. The presence of the NMA pathway in several pathogenic species but absence in ours indicates this pathway as a key source of novel therapies to combat the growing resistance problem. There are a couple of issues. There are two enzymes belonging to large families of enzymes. DXS and IspE belong the TDP-dependent enzyme and GHMP families, respectively. High sequence similarity between these enzymes and their families poses unintended consequences. This has been observed consistently in work with kinases. Trying to develop a highly specific inhibitor is troubling at best. These considerations should not inhibit our attempts at developing new therapies base on this pathway.

1.4 Computational Methodology

Biochemistry is the study of the chemical reactions involved in biological processes. At the heart of this endeavor are enzymes that facilitate these processes. Hence, it became rapidly apparent a deeper understanding of enzymes was needed. When studying enzymes some key questions are: "What amino acids are participating in the enzymatic action? What are their roles? And what are the possible transition states?"[116,148,182]. In pursuit of answers to these questions, biochemists developed laboratory techniques to probe the relative importance of certain amino acids (AAs) and the motions of these highly dynamic macromolecules. Some of these experimental methods are kinetic isotope effects (KIE); site directed mutagenesis, and Forster resonance energy transfer (FRET). It was hoped that they could give insights into transition state structures and a better view of the molecular level interactions occurring in enzyme active sites. These techniques have contributed significant knowledge of the inner workings of the enzymes, however

they do have limitations. With the advent of macromolecular crystallography and later NMR based methods, the ability to see at the molecular level was greatly enhanced.

Computational biochemistry can provide an even more detailed look into enzyme mechanisms. This includes but is not limited to the investigation of the motion of enzymes, de Novo design of transition state analog inhibitors, and investigating protein-protein interactions[93]. My current focus is in 3 areas: application of computational techniques to elucidate mechanistic detail of two enzymatic processes.

The study of condensed phase chemical and biochemical processes has been major focus for both experimental and computational chemists for several decades now. Though QM approaches for computation are more accurate, the computational cost prohibits the use of these approaches with any biologically relevant systems. This limitation of QM methods was a driving force behind the growing use of more efficient MM methods that are more empirically driven. Significant time and effort has been put into attempting to improve the efficiency of traditional QM methods recently. Though these advances have shown benefits for small molecule chemistry, they are still prohibitively expensive for full scale biochemical applications. A problem that has been mitigated in part through the development of more efficient QM codes and the growing use of hybrid QM/MM methods. Standard methods of QM/MM attempt to couple the cost efficiency of MM methods with the accuracy and precision of QM methods through the division of the system into subsystems. These subsystems are treated at different levels of theory. One region, that is usually made up of the active site or site of most significant interest, is labelled the QM region and treated with the highest level of computational theory. An MM region is also defined and the protein environment that surrounds the QM region. The third region is an interface region that connects the QM and MM regions previously defined. The third region is only necessary if in the course of defining the QM region from the MM region, any bonds found along their borders are split between the regions, becoming the interface

region. A coupled potential is responsible for the inclusion of electrostatic and van der Waals interactions from the QM and MM through the interface region[136,182,188].

Several methods for the implementation of hybrid QM/MM schemes have been reported. Empirical valence bond (EVB) and semi-empirical methods have been employed typically to describe the QM region, due to their relative efficiency in comparison to ab initio QM theory. Though they have been used effectively, several weaknesses that have been well documented. Recently, there has been a bigger push to overcome the deficiencies in these methods through the implementation of more accurate and rigorous computational methods such as ab initio and Density Functional Theory (DFT). Herein, we will be applying QM/MM methodology with the QM region being treated with the more rigorous DFT methodology[162].

A major advantage of using hybrid QM/MM techniques is the ability to compute barriers for biological processes (e.g. NOX production). The relative free energy (Δ G) of each step along the reaction will be calculated in order to ascertain (within relative degrees of certainty) the profile of a proposed mechanism and therefore suggest the most energetically favorable mechanism[136,182,188]. ΔG has been defined as a measure of the driving force behind a reaction. Thus by calculating the driving force, the mostly probable reaction will be uncovered. In addition, key residues involved in the stabilization of the transition state or destabilization of the reactant state will be idenitifed and analyzed for relative electrostatics in determining the relative energetics of different reaction mechanisms.

Chapter 2

Thiamin Diphosphate Activation in 1-deoxy-D-xylulose 5-Phosphate Synthase: Insights into the Mechanism and Underlying Intermolecular Interactions

2.1 ACS Permissions

2.2 Abstract

1-deoxy-D-xylulose 5-phosphate synthase (DXS) is a thiamin diphosphate (TDP) dependent enzyme that marks the beginning of the methylerythritol 4-phosphate isoprenoid biosynthesis pathway. The mechanism of action for DXS is still poorly understood and begins with the formation of a thiazolium ylide. This TDP activation step is thought to proceed through an intramolecular deprotonation by the 4′-aminopyrimidine ring of TDP; however, this step would occur only after an initial deprotonation of its own 4′-amino group. The mechanism of the initial deprotonation has been hypothesized, by analogy to transketolases, to occur via a histidine or an active site water molecule. Results from hybrid quantum mechanical / molecular mechanical (QM/MM) reaction path calculations

reveal an ~10 kcal/mol difference in transition state energies, favoring a water mediated mechanism over direct deprotonation by histidine. This difference was determined to be largely governed by electrostatic changes induced by conformational variations in the active site. Additionally, mutagenesis studies reveal DXS to be an evolutionarily resilient enzyme. Particularly, we hypothesize that residues H82 and H304 may act in a compensatory fashion if the other is lost due to mutation. Further, nucleus-independent chemical shifts (NICS) and aromatic stabilization energy (ASE) calculations suggest that reduction in TDP aromaticity also serves as a factor for regulating ylide formation and controlling reactivity.

2.3 Introduction

Isoprenoids are one of the largest and most diverse families of biomolecules with a number of them essential for life[99,158,161]. An example would be Vitamin A, which plays a role in human growth and development as well as immune system maintenance. Two isoprene molecules are variably employed in the construction of all isoprenoids. Isopentenyl diphosphate (IDP) and dimethylallyl diphosphate (DMADP) are produced via two distinct biosynthetic pathways (Figure 2.1): mevalonate pathway (MVA) and methylerythritol 4-phosphate pathway (MEP pathway)[17,33,191]. The MVA pathway was discovered in the 1950s and was considered the sole pathway until the 1990s when discrepancies in isotopic labeling studies led researchers to hypothesize an alternative, MEP pathway[51,159]. Subsequent, genetic studies have revealed a large variety of life (e.g., algae, bacteria, etc.) to have varying degrees of dependence upon MEP pathway for isoprenoid production; in addition to a number of human pathogens (e.g., *Plasmodium spp.*, and *M. tuberculosis*)[52,66,110,147]. Interestingly, the MEP pathway is absent in all mammalian genomes meaning that the enzymes of this pathway are ideal targets for novel antibiotics and antimalarials[81]. For example, fosmidomycin is known to be an effective agent against several of the *Plasmodium spp.* (malarial pathogens) and targets MEP pathway's second step, 1-deoxy-D-xylulose 5-phosphate reductoisomerase (DXR)[81,100,203].
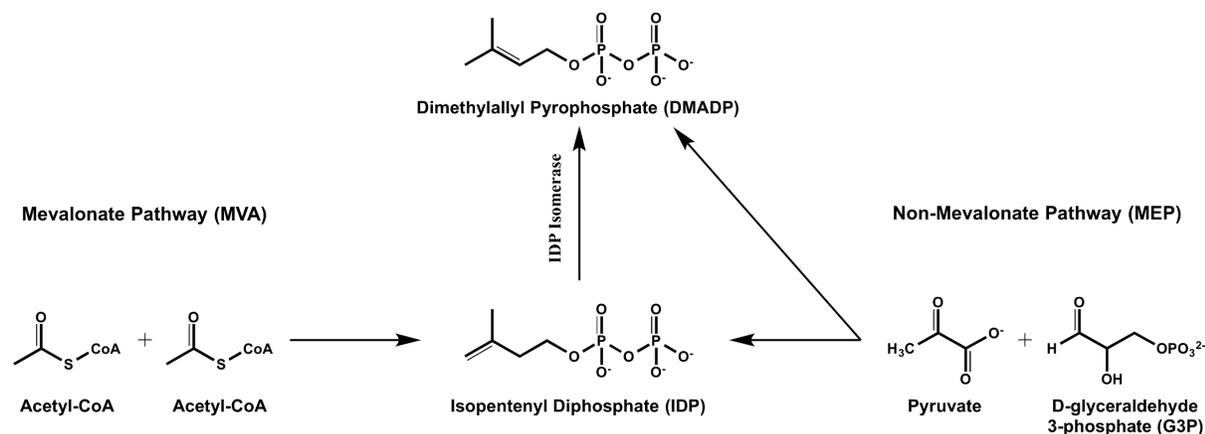
Figure 2.1: Schematic of isoprene production via MVA or MEP pathway. MVA pathway produces DMADP via a secondary enzyme, IDP Isomerase[45,171]. MEP pathway directly synthesizes both isoprene molecules.

MEP pathway is comprised of eight reactions catalyzed by seven enzymes beginning with 1-deoxy-D-xylulose 5-phosphate synthase (DXS)[18,53,69,103,114,115,124,176]. DXS catalyzes the condensation of glyceraldehyde-3-phosphate (G3P) and pyruvate to produce 1-deoxy-d-xylulose 5-phosphate (DXP). Aside from isoprenoid production, DXP is utilized in the production of vitamin B_1 (thiamin) and vitamin B_6 (pyridoxine) biosynthetic pathways[14,53,76] suggesting increased significance to understanding the mechanism of DXS. Further, DXS is believed to be a rate-limiting step due to the observed correlation between isoprenoid product levels and DXS levels.[53] DXS is a member of the thiamin diphosphate (TDP) dependent family of proteins; specifically a member of the transketolase (TK) enzyme subclass and also possesses pyruvate decarboxylase activity.[7,114,176] TKs are a class of TDP dependent enzymes responsible for the transfer of a ketol donor group to an aldehyde or ketone acceptor molecule[41,60,163]. In 2007, Xiang et al. published crystal structures and mutagenesis results of DXS from *E. coli* and *D. radiodurans* and compared them to the E1 subunit of pyruvate dehydrogenase (PDH) and yeast TK (members of the same class of enzymes)[197]. The comparison revealed significant similarities between these four enzymes: 1) each enzyme is composed of three domains (I, II, and III), 2) all possess a TDP cofactor, and 3) all contain a $GDGX_{25-30}N$ motif that

plays a role in producing the twisted 'V' shape of the TDP cofactor in the active site[197]. The strained cofactor conformation has been shown to play a role in lowering the pK_a of a hydrogen on the thiazolium ring's C2 atom (Figure 2.2)[25,26,60,84,125]. The active site of DXS contains a number of strictly/highly conserved residues (e.g., Glu370/372, Asp152/154 in *E. coli/D. radiodurans*, respectively) that are common among TDP dependent enzymes, particularly TKs. These similarities have led researchers to propose a DXS reaction mechanism based, primarily, on mechanistic data of other TK enzymes (Figure 2.2)[41,163,173].
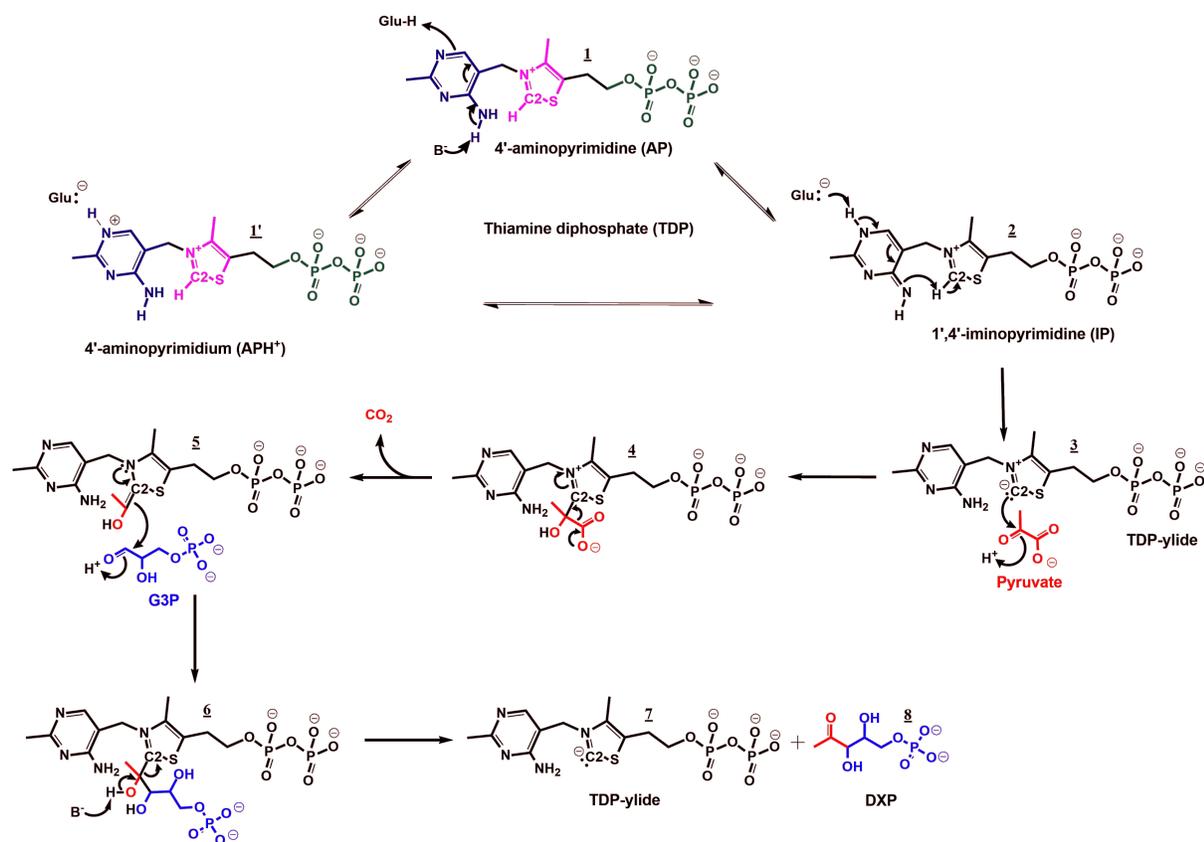
Figure 2.2: Proposed general mechanism for DXP biosynthesis. Pieces of each step are labeled with different colors to indicate where they originate from. Red represents pyruvate and blue are the pieces affiliated with G3P.

Although DXS contains many of the strictly/highly conserved residues of the TDP-dependent superfamily of enzymes (*vide supra*), it displays distinct structural features[197]. Specifically, the domain arrangement of DXS; homodimeric with a deep pocket between

two domains of the same monomer rather than at the dimer interface as is the case with other TDP enzymes. These structural differences logically lead to questions of mechanistic similarity. Until the discovery of DXS, it was believed that all TDP enzymes functioned via a classical "ping-pong" mechanism (i.e., pyruvate binding $\longrightarrow$ CO_2 release $\longrightarrow$ G3P binding)[60]. However, Eubanks and Poulter in 2003 concluded that DXS operates via an ordered mechanism (i.e., irreversible pyruvate binding $\longrightarrow$ G3P binding $\longrightarrow$ CO_2 release) and hypothesized a side reaction for producing CO_2 via binding of a second pyruvate molecule[54]. This hypothesized side reaction was subsquently confirmed by Brammer and Meyers in 2009[19]. However, the following year a steady-state kinetics study examining a herbicide metabolite (ketoclomazone, a derivative of clomazone) provided evidence of a traditional ping-pong mechanism for DXS[123]. Nearly, simultaneously (in 2010), a single-molecule force spectroscopy nano sensor was developed and used to observe an approximate 2-fold binding enhancement of G3P in the presence of pyruvate; suggesting an ordered DXS mechanism[174]. As part of this work, the authors cast doubt on the reliability of previous results based upon assays that measure bulk phenomena rather than single-molecule behavior. To further confound the situation, Meyers and co-workers (in 2011) proposed an unprecedented TDP-based mechanism; G3P and pyruvate were found to bind independently and reversibly. Thus, they concluded DXS functions via a rapid equilibrium, random sequential mechanism[20]. In the following year, Meyers and co-workers revealed a 600 fold acceleration in the decarboxylation of the lactyl-TDP intermediate upon binding of G3P. This result further distinguishes DXS from other TDP-dependent enzymes.

Two recent studies have called into question our understanding of the active sites of the large class of TDP dependent enzymes. For instance, the benzaldehyde lyase (BAL) enzyme is devoid of all but two acid-base residues around the TDP active site: a histidine and highly conserved glutamate[22,121]. Most interesting is the lack of any apparent acid-base residues in glyoxylate carboligase (GCL)[86]. These recent discoveries represent glaring gaps in our understanding of TDP-dependent enzymes and bolster the importance

of investigating distinct related enzymes (i.e., DXS).[21,22,86,121] Here, computation is an ideal partner to experiment.

Of particular interest in this work is the "true first step" of this process: TDP activation, of which significant mechanistic details are still largely uncertain. For example, a proposed mechanism for ylide formation begins with an initial deprotonation of the 4′-aminopyrimidine (AP) state that produces the 1′,4′-iminopyrimidine (IP) state[114,176]. A general base (GB) is required for this deprotonation, however, the identity of this group remains unknown. One hypothesized GB is a highly conserved histidine (His434 in *D. radiodurans* DXS) found proximal to TDP's 4′-amino group[60,82]. The aforementioned mutation studies (i.e., H434A) showed approximately 95% activity retention, which suggests an alternative mechanism. A more recent 2014 study by Querol et al. suggests H431 (*E. coli* equivalent of *D. radiodurans* H434) plays a role in transition state stabilization but not required for catalysis[145]. Additionally, numerous structural differences between DXS and TK enzymes (*vide supra*) support the possibility of an alternative mechanism[197]. One possible alternative mirrors that of human TKs; where a water molecule would replace H434 as the GB with a Gln residue acting to stabilize this via coordination[134,173,190]. This results in two possible TDP activation mechanisms: a water-mediated mechanism (WMM) or direct histidine mechanism (DHM). Even though the WMM utilizes a water molecule as the initial general base, it is possible that H434 plays a role in this mechanism as either a coordination site for the water molecule or as the final location of the proton.

TDP has been shown to exist in four different tautomeric/ionization states (Figure 2.3); however, the exact mechanism for producing the final ylide form remains unclear[8,9,133,141]. Figure 2.3 illustrates two possible mechanisms: (1) a concerted AP to IP conversion followed by ylide formation or (2) a step wise mechanism where the AP is first ionized to a 4′-aminopyrimidinium ion (APH$^+$) and, subsequently, converted to the IP and ylide, respectively. Although direct spectroscopic evidence of the APH$^+$ state remains elusive, its existence has been inferred from alternative experiments (e.g., pH rate profiles, solid-state NMR) and hypothesized to assist in promoting IP formation via

stabilization of the tautomerization reaction[8,83,133]. An elevation in the pK_a of TDP's N1 atom is proposed to account for the APH[+] state's existence; which is justified by its proximity to a strictly conserved glutamate residue. This idea, however, does not consider the possibility of an accompanying elevation in the pK_a of the glutamate residue. Recent studies have determined the pK_a for the N1 atom in DXS to be 7.5[141] and a PROPKA[11,108,138,175] calculation has estimated the E373 residue to have a pK_a of 8.4.[1] These pK_a values suggest an equilibrium between the AP, and APH[+] states; which is consistent with the enzyme stabilizing the IP formation via pK_a modulation. Additionally, Jordan et al. supports the equilibria presented in Figure 2.3; particularly for apo (TDP-bound enzyme lacking substrate) enzymes[83]. Therefore, it is not necessary to select between the step wise or concerted mechanism for the purposes of this study.
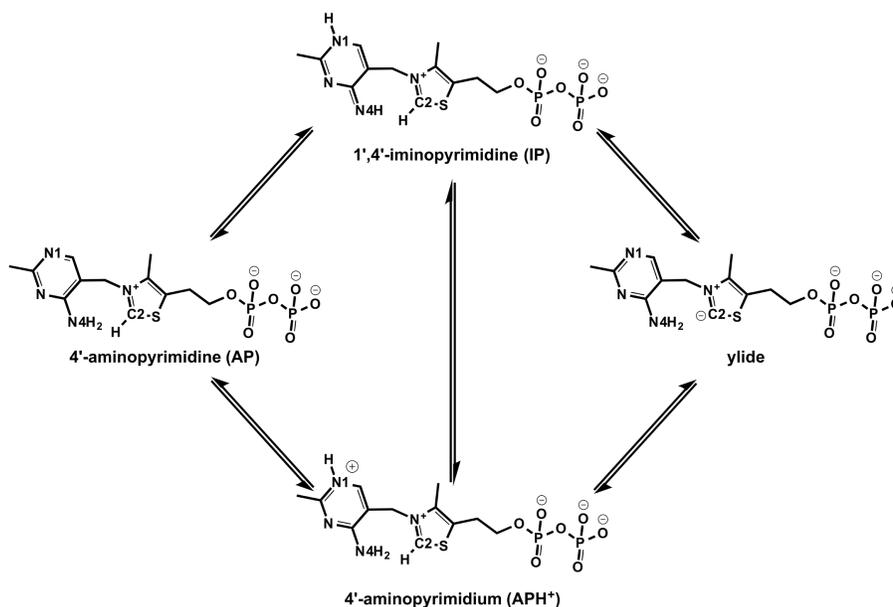
Figure 2.3: Structure and relationship of the 4 possible tautomeric/ionization states proposed for the cofactor of TDP dependent enzymes[8,9,125,133,141]. Key atoms have been given names for reference purposes throughout this article.

---

[1]The pK_a for the N1 atom of DXS was determined via pH rate profile studies. The E373 pK_a was estimated using PROPKA3.1 with TDP but without pyruvate and G3P. The complete output of the PROPKA3.1 calculation can be found in the supporting information.

The rate of activation and turnover of TDP shows a substantial increase when bound to an enzyme rather than in solution[84]. Several factors leading to the increase in activity have been proposed. One factor includes the strained conformation the cofactor adopts upon binding. This conformation places the $4'$-amino group in close proximity to the thiazolium C2 atom; both introducing strain and lowering the pK_a of the C2 hydrogen from 14-19 (depending on solvent) to approximately $9^{65,82}$. The electronics of the pyrimidine ring of TDP would also undoubtedly change during the activation process. These changes could lead to a disruption of the aromaticity and more indirectly influence the energetics of TDP activation. The link between aromaticity and TDP activation has not been investigated previously. Uncovering such a link will provide better understanding of DXS and raise the question if this phenomenon is general for all TDP dependent enzymes (e.g., TK, PDH, etc.).

Herein, a hybrid quantum mechanical/molecular mechanical (QM/MM) study is carried out using the *D. radiodurans* DXS crystal structure[197]. The energy profile of the DHM and WMM are computed to determine the most favorable activation mechanism. Active site electrostatics are also probed to elucidate the stabilizing/destabilizing effects that govern this process. Further, two metrics of aromaticity are employed to quantify this effect and determine its role as a possible driving force in activation of TDP dependent enzymes. In addition to the computational work, a kinetics study, utilizing a coupled enzyme assay, was performed on mutant and wild-type forms of the *D. radiodurans* DXS enzyme. As part of the mutagenesis work, we have re-examined the H434A mutation; which is of particular interest in this study. Our study focuses on the K_M and k_{cat} due to their relationship with substrate affinity and reaction turnover, respectively. The combination of computational and experimental results helps bridge the gap between bulk behavior and atomistic understanding; ultimately leading to new insights into this unique enzyme.

2.4 Methods

2.4.1 Computational Methods

The DXS active homodimer structure was used throughout this study. The crystal structure for the *D. radiodurans* DXS (PDB ID:2O1X)[197] enzyme with TDP bound was processed and parsed via www.charmming.org[127]. A TDP molecule can be broken down into three moieties: a thymine-like pyrimidine ring, a pyrophosphate (residue name utilized in the topology file), and a thiazolium ring(Figure 2.2). Parameters for all three of these have been developed as part of the CHARMM General Force Field (CGenFF)[184]. Final TDP parameters were thus obtained by connecting the respective components and modifying charges (see SI). Parameter validation was done with respect to the TDP crystal conformation based on the RMSD (see SI).

Structural modifications were performed to ensure the active site Glu373 was protonated in agreement with experimental evidence[90,112]. CGenFF and CHARMM22 protein (C22) force fields[119] were used throughout. The system was solvated in a rhombododechedron crystal structure and neutralized with KCl salt to a final concentration of 0.15M. The system was heated from 110K to 310K over 100ps and equilibrated for 200ps at constant pressure (1atm) and temperature (310K). The total system size was then reduced by removing all of the water and salt ions beyond 12Å from the surface of the protein. The reduced structure was then QM/MM minimized, without applying cutoffs, to a gradient tolerance of 0.002 kcal·mol^{-1}·Å^{-1}. All QM/MM calculations employed the Q-Chem4.0\CHARMM[27,170,194] interface at the B3LYP/6-31G* level of theory[13,68,105]. Additionally, the single link atom scheme was used to account for truncation of the QM region and employed group electrostatic exclusions to prevent over polarization of the QM region.

Reaction pathway calculations employed a combination of the replica path method (RPATh) and harmonic distance restraints (RESDi)[36,194,195]. The RPATh method permits the user to divide the system into discrete subsystems (replicas), which are allowed

to change independently of the remaining "environment". A subsystem was defined to be 6.5Å around the QM region, which was comprised of TDP, Glu373, His120, and His434 (Figure 2.4a, 98 QM atoms). The QM region for the WMM included a water molecule (Wat9709) coordinating with the 4′-amino group of the pyrimidine ring of TDP and the Nϵ of His434. Wat9709 was removed from the initial structure prior to QM/MM minimization (*vide supra*) for the DHM. Two replicas of the subsystem were used to model successive steps along the reaction coordinate (δ), which was defined as a linear combination of the bond being broken and the bond being formed (Figure 2.4b, 101 QM atoms). The δ values were defined incrementally for each mechanism starting from the reactant state. The DHM scanned a range of -3.0Å to 3.0Å in increments of 0.3Å with smaller increments of 0.1Å used around the transition state (i.e., -1.0Å to 1.0Å) to provide finer detail. The WMM scanned a range of -1.1Å to 1.1Å in increments of 0.1Å.

The Charge Perturbation Analysis (CPA)[12,42,67,107] technique involves QM/MM single point energy calculations where a single residue's classical charge is scaled to zero to probe its electrostatic contribution. ΔE is computed by taking the difference of the modified (zero-charge residue) and the full QM/MM electronic energy: $\Delta\text{E}_{\text{CPA}} = \text{E}_{\text{elec}}^{ZeroChargeRes}(\text{QM/MM})$ - $\text{E}_{\text{elec}}^{FullMM}(\text{QM/MM})$. CPA calculations were performed for the reactant state (RS) and transition state (TS) of the DHM and WMM as determined by RPATh+RESDi calculations. The reactant state was the starting δ value for each mechanism while the TS corresponded to the point along the path with the highest energy. $\Delta\Delta\text{E}_{\text{CPA}} = \Delta\text{E}_{\text{CPA}}(\text{RS})$ - $\Delta\text{E}_{\text{CPA}}(\text{TS})$ provides insight into stabilizing/destabilizing electrostatic effects with respect to RS and TS. CPA was performed on all 82 residues found within 5Å of the QM region for both mechanisms. To further characterize long range electrostatic changes, the QM/MM dipole moments were calculated around the QM region for the RS and TS of each mechanism. The QM/MM dipoles take into account the external charge contributions of the MM region on the QM region. The calculations were carried out using Q-Chem 4.0[170] and initially visualized in IQmol with final rendering using PyMOL.

Two metrics were employed to quantify aromaticity and gauge the level of significance of it as a possible driving force of ylide formation: nucleus-independent chemical shifts (NICS)[164] and aromatic stabilization energy (ASE)[43,44]. NICS directly measures the aromatic character of a compound[35,39,43,164] while ASE reveals the stabilization/destabilization that arises from the aromaticity of a compound[1]. Due to computational limitations, NICS calculations were performed on reduced versions of the RS and TS subsystems. The reduced subsystems contained 232 or 235 atoms for the DHM or WMM, respectively. The difference of 3 atoms being the absence of Wat9709 from the DHM. Ghost atoms were placed perpendicular to the plane of TDP's pyrimidine ring. Due to the non-symmetric protein active site, the NICS(0) (ring center), NICS(1), and NICS(-1) (atoms 1Å above and below the plane of the ring) will be reported herein[10,35,44,98]. All NICS calculations were performed using Q-Chem 4.0[170] at the B3LYP/6-31G* level of theory[13,68,105]. ASE is typically computed via a reference homodesmotic reaction[1,189]. A homodesmotic reaction must be defined such that equal numbers of each type of atom (sp^3, sp^2, sp) and bond (sp^3-sp^3, sp^3-sp^2, sp^2-sp^2, etc...) exist in both reactants and products[189]. All structures used in ASE calculations were optimized at the B3LYP/6-311+G** level of theory. Energies for each molecule were corrected by subtracting out their respective zero point energy obtained from subsequent frequency calculations.

2.4.2 Experimental Methods

Materials

TDP, pyruvate, G3P, DXP sodium salt, bovine serum albumin, and LB-broth were purchased from Sigma Aldrich. NADPH was purchased from Alexis Biochemical, Ni-NTA resin was purchased from Invitrogen, and β-mercaptoethanol (β-Me) was purchased from Fisher. *E. coli* XL-10 cells, deoxynucleotide mix PCR grade, *pfu*Ultra Hotstart DNA polymerase, QuikChange II site directed mutagenesis kit and acetonitrile (HPLC grade) were purchased from Agilent. The DNA vectors pET28a(+) and pET15b(+) and *E. coli* BL-21 B(DE3) cells were purchased from EMD Biosciences. DNA sequencing services and

primers were purchased from MWG operon. All the other reagents were of the highest quality commercially available.

Cloning of *D. radiodurans* DXS and *E. coli* DXR

A synthetic, codon optimized *D. radiodurans dxs* gene with $5'$-*NdeI* and $3'$-*XhoI* restriction sites in a pMK vector was purchased from Geneart (Germany). The *dxs* gene was excised from the pMK vector and cloned into the *NdeI* and *XhoI* sites of a pET28a(+) vector (*kanamycin* resistance) with an N-terminal His$_6$-tag to yield the pET28a(+)-DXS plasmid. Successful cloning of the *D. radiodurans dxs* gene was confirmed by DNA sequencing at MWG Operon.

A synthetic, codon optimized *E. coli dxr* gene with $5'$-*NdeI* and $3'$-*BamHI* restriction sites in a pMK vector was purchased from Geneart (Germany). The *dxr* gene was excised from pMK vector and cloned into *NdeI* and *BamHI* restriction sites of pET15b(+) vector with a C-terminal His$_6$ tag to yield the pET15b(+)-DXR plasmid. Gene insertion was confirmed by DNA sequencing.

Production of the *D. radiodurans* DXS Mutants

Site-directed mutagenesis was carried out using the QuikChange II site-directed mutagenesis kit. Briefly, the mutagenesis mixture consists of 50-100 ng plasmid pET28a(+)-DXS as a template, 1X PCR reaction buffer, 0.4 mM each of the forward and reverse primer, 0.25 mM dNTP mixture, 5 μL Quik solution, and 2.5 units of *pfu*Ultra hotstart polymerase in a 50 μL reaction. The overlap extension method was used to produce the DXS mutants that were difficult to create via site directed mutagenesis[77]. The sequence of the mutant DNA was confirmed by DNA sequencing.

Assays for DXS Activity

We employed a DXS-DXR coupled assay to determine the wild-type and mutant DXS enzyme activities. In this way, the DXS-dependent production of DXP is ultimately

coupled to the oxidation of NADPH to NADP$^+$ via the DXR enzyme. The solution for the DXS-DXR coupled contained 100 mM HEPES pH 8.0, 100 mM NaCl, 1 mg/mL BSA, 1 mM TDP, 1.5 mM MnCl$_2$, 2 mM β-Me, 0.15 mM NADPH, 0.2 mg/mL DXR, and varying concentrations of pyruvate or G3P[79]. Steady-state kinetic experiments were performed by varying pyruvate or G3P at a fixed saturating concentration of the co-substrate. A DXS-DXR reaction solution was incubated at 37°C for 5 min, the reaction was initiated by addition of 358 nM DXS, and the progress of the reaction monitored spectrophotometrically at 340 nm for the oxidation of NADPH. The DXS and DXR employed in this assay were over-expressed and purified based on the methods presented in the supporting information. Each sample was stored at -80°C until used for the assay. The steady state initial velocity for DXS measured at various concentrations of pyruvate and G3P were fit to equation 1 (see SI for plots) using nonlinear regression analysis in Sigma-Plot 12.0.

$$v \;=\; \frac{V_{max}[S]}{K_M \;+\; [S]} \tag{2.1}$$

2.5 Results and Discussion

A central aim of this investigation is to determine and characterize the mechanism of TDP activation in DXS (Figure 2.2). There are two hypothesized mechanisms acting by a different GB: WMM (Wat9709) and DHM (His434). Though most TK enzymes are thought to rely on a histidine residue as the GB, key structural differences and mutagenesis results suggest DXS might diverge from the majority of TK enzymes[20,54,60,129,141]. The reactant state QM/MM minimized structures (Figure 2.4a, 2.4b) provides some initial insight into this process. Coordination of the water oxygen to the H$_n$ (2.0Å,Figure 2.4b) suggests water could act as the GB. Further, the distance between H$_o$ and N$_\epsilon$ (1.8Å, Figure 2.4b) suggests that this could be the final destination of this proton. Alternatively, in the absence of Wat9709, His434 directly interacts with TDP albeit more distantly (4.2Å, Figure 2.4a)[197].
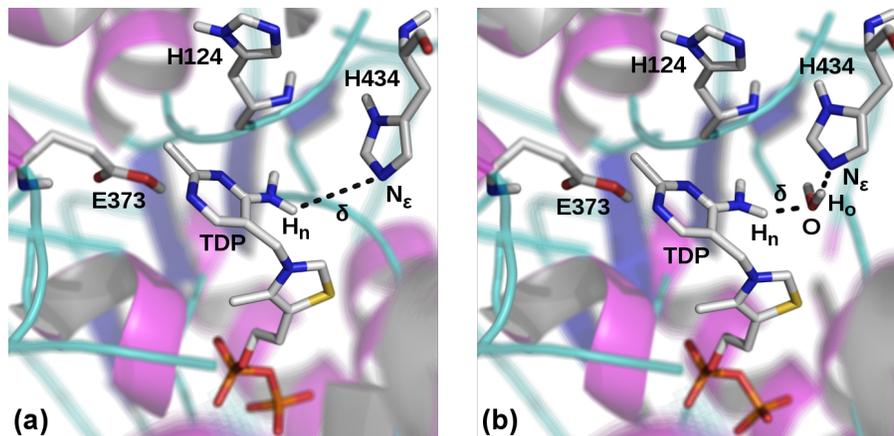
Figure 2.4: Representations of the RS for DHM (a) and WMM (b). The dashed black lines illustrate the proton transfer reaction.

To determine each mechanism's feasibility, the RPATh+RESDi technique was employed and respective minimum energy pathways were computed. A plot of ΔE with respect to δ (Figure 2.5) values illustrates the energetic favorability of the WMM over the DHM. A δ value of 0.3Å corresponds to the TS of both mechanisms. The difference between barriers can partially be explained by a conformational change that occurs during the DHM (Figure 2.6). This involves the movement of His434 into a conformation more favorable for deprotonation of the 4'-amino moiety. His434's movement induces a strain in the protein backbone and perturbs the configuration of the local environment. This change in configuration accounts for a portion of the energetic differences between the WMM and DHM but does not provide a complete explanation. Further, a $\Delta E^{\ddagger}_{\text{WMM}}$ value of 22.7 kcal·mol^{-1} is considerably higher than one might expect for an enzyme catalyzed proton transfer and cannot be explained by a simple conformational change[142].

The reaction pathway calculations applied a restraint to the proton transfer involved in the DHM or WMM. No other restraints were applied to the system. Upon examination of structural changes during the reaction, a second proton was observed to spontaneously transfer from E373 to the N1 atom of TDP's AP ring in both mechanisms (Figure 2.7). Since E373 was included in the QM region, the proton transfer occurred in response to electronic changes encountered during each mechanism. The combination of the re-
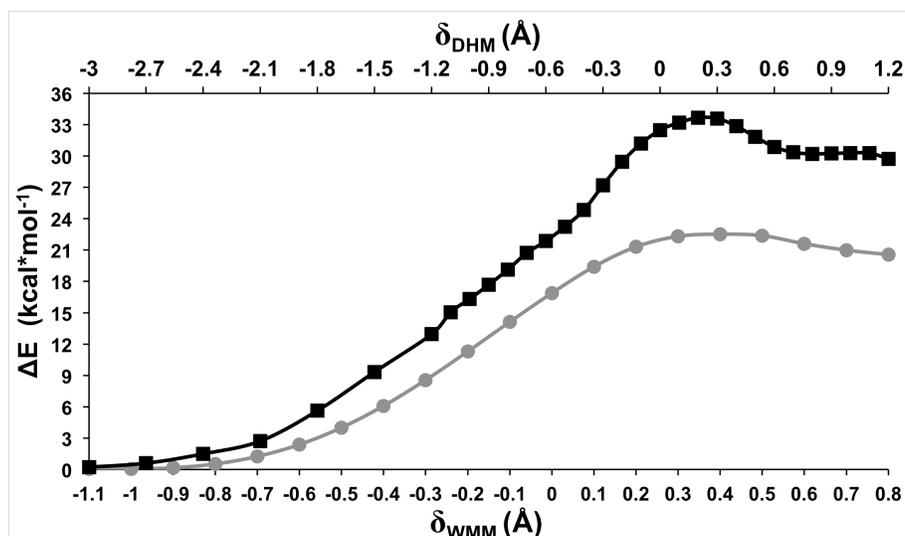
Figure 2.5: Minimum energy profiles computed for the WMM and DHM. The different x-axes are used because of differences in the reaction coordinate ranges for WMM vs DHM; both are associated with the same y-axis. The $\Delta E^{\ddagger}$ are 22.7 kcal·mol^{-1} and 33.7 kcal·mol^{-1} for the WMM (gray circles) and DHM (black squares), respectively.

strained reaction path proton transfer and unrestrained E373 to N1 atom proton transfer represents the tautomerization of the AP to IP state (Figure 2.3). The formation of
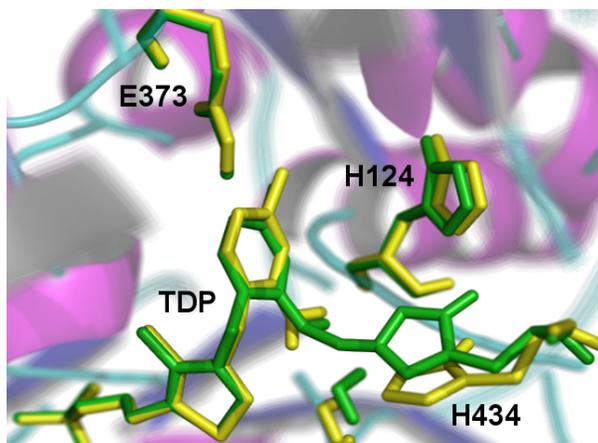
Figure 2.6: Representative conformational changes between the RS (yellow) and TS (green) of the DHM.

the ylide state is dependent upon first forming the IP state. There is some debate in the literature over the exact details of the IP state formation (*vide supra*)[8,9,133,141]. Most studies propose an equilibrium between the AP, APH$^+$, and IP TDP states (Figure 2.3)

particularly for apo enzymes[83]. As highlighted in the introduction, the pK$_a$s of TDP's N1 atom and E373 residue (see SI and Introduction) are approximated to be close to one another using experimental and empirically based computational techniques. The combination of the pK$_a$s and observed responses from QM/MM calculations suggests a concerted mechanism as previously thought. Additional studies are underway to more fully address this unresolved question.
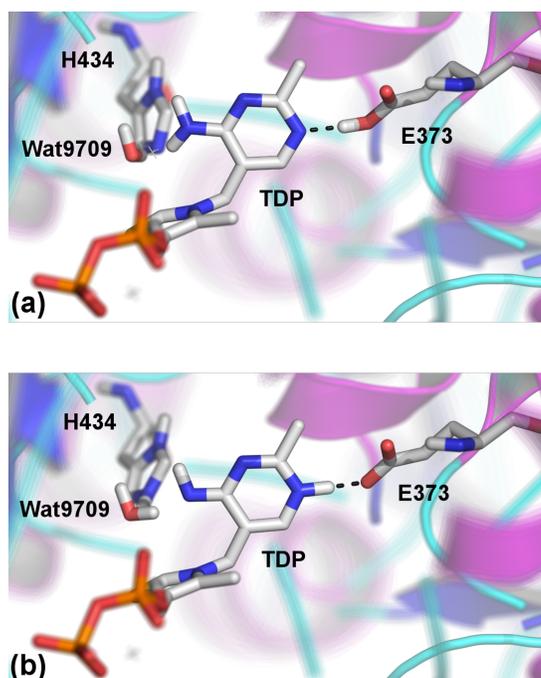
Figure 2.7: Illustration of the proton transfer from E373 to TDP's AP ring during the tautomerization reaction. (a) and (b) represent the reactant and product states, respectively. While this figure only depicts the structures of the WMM, a similar response was observed during the DHM.

The CPA method, which approximates electrostatic contributions of a single active site residue, was used to determine the stabilizing/destabilizing effects of active site residues as a function of both states (i.e., RS vs TS) and mechanisms (i.e., WMM vs DHM). Negative $\Delta\Delta E$ values indicate that a particular residue is more stabilizing towards the TS; whereas positive $\Delta\Delta E$ values show stabilization of the RS. From the 82

	$\Delta\Delta E_{\text{DHM}}$	$\Delta\Delta E_{\text{WMM}}$
K101	2.0	7.5
H51	-10.1	-1.2
K289	-12.3	-2.5
D430	-21.3	-10.8

Table 2.1: $\Delta\Delta E$ values for four residues of interest in the WMM and DHM. Negative $\Delta\Delta E$ values indicate preferential stabilization of the TS; while positive $\Delta\Delta E$ show stabilization of the RS preferentially. All values are in kcal·mol^{-1}.

active site residues examined, there were 4 that showed substantive differences (Table 2.1). Residues found stabilizing the TS were D430, K289, and H51, and Wat10307. K101 were found to preferentially stabilize the RS. K101, H51 and D430 were found in a
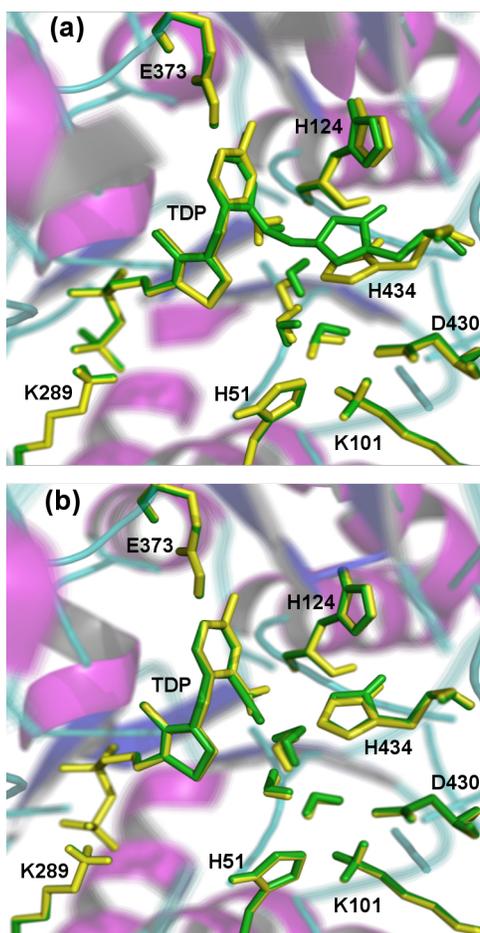
Figure 2.8: Active site conformation of the residues discussed in the CPA results. Images show both the RS (yellow) and TS (green). (a) illustrates the DHM while (b) shows the WMM.

catalytic triad-like configuration in the active site (Figure 2.8). It is unlikely that they play a direct role in this reaction due to their distance from the site of activity (7.2Å). A cluster of water molecules were found to span the distance between the reaction site and triad; which suggests an electrostatic role. K289 coordinates to the negatively charged phosphate tail of TDP (Figure 2.8) and is highly conserved in *D. radiodurans*, as well as other TDP dependent enzymes. In fact, most TDP dependent enzymes are found to require a divalent metal ion and positive residues near the phosphate tail to anchor the cofactor.

The magnitude of $\Delta\Delta E_{DHM}$ values were consistently larger than the magnitudes of $\Delta\Delta E_{WMM}$ values. This behavior is attributed to the structural change that the DHM TS must adopt in order to position His434 for deprotonation of TDP's 4'-amino group. The increased TS stabilization for this mechanism suggests the enzyme is tuned to accommodate alternative activation routes although they may not be the most favorable. For example, active site mutations are a common way that bacteria and other lower life forms (i.e., those that rely on MEP pathway) can adapt to changes in chemical environments. By tuning the DXS active site to stabilize TDP activation via varying general bases, evolutionary fitness is maximized.

To better characterize long range electrostatic effects, QM/MM dipole moments for the RS and TS for each mechanism were computed and visualized (Figure 2.9). The RS dipole moments of both the WMM and DHM were essentially the same. Further, WMM dipoles, both RS and TS, are indistinguishable (Figure 2.9b) whereas the DHM TS dipole moment is significantly perturbed (Figure 2.9a). Again, this effect is attributed to the conformation change His434 undergoes during the DHM and appears to be the underlying source of DXS's ability to stabilize non-water mediated TDP activation.

Herein, we also report experimental kinetics studies of pyruvate and G3P binding and reaction in DXS and several DXS mutants (Table 2.2). For H434A, there exists negligible increase in catalytic rate for pyruvate as well as G3P in comparison to wild-type, respectively. K_M values also slightly increased by 6.1 and 4.6 folds, respectively. G3P's negatively charged phosphate tail is thought to bind in a positively charged region of the active site; which contains the polar H434 residue. Additionally, the negatively charged pyruvate is thought to interact with the same positive region but, not as strongly[197].
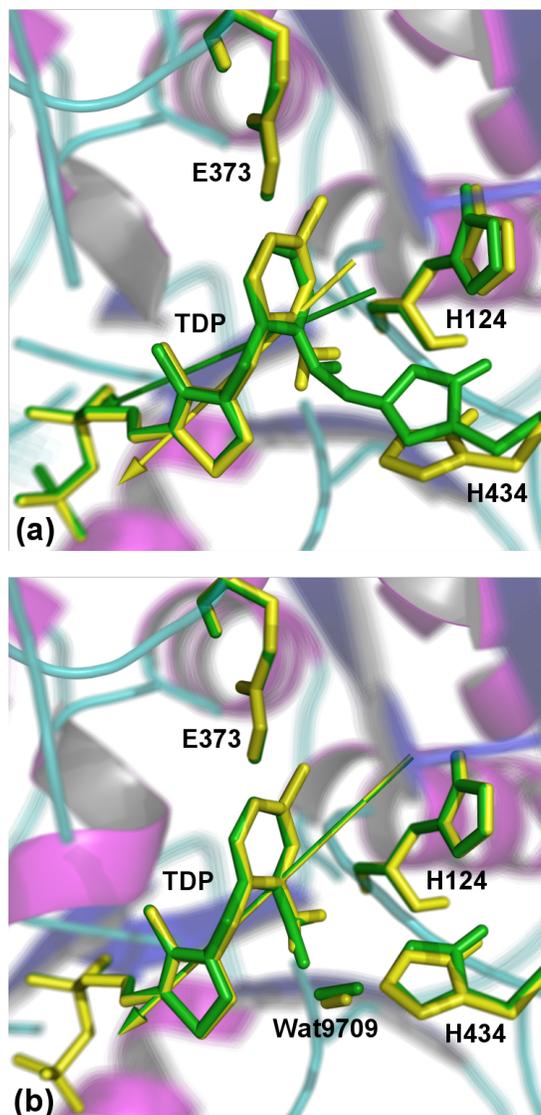
Figure 2.9: Illustrated above are the computed RS (yellow) and TS (green) dipoles of the WMM and DHM. (a) is the DHM and (b) is the WMM.

Therefore, the mutagenesis results suggest that the electrostatic effects that accompany the H434A mutation have a clear destabilizing effect on substate binding while enhancing turnover. This behavior is contrary to what would be expected if H434 is required for initial TDP activation. Thus, the H434A mutant supports the conclusion favoring a WMM for TDP activation.

Another interesting correlation between CPA and mutagenesis results is related to the D430A mutant. As previously discussed, D430 is found in a electrostatic triad of residues that includes K101 and H51 (Figure 2.8). While the k_{cat} for D430A mutant remains relatively unchanged, the K_M for pyruvate and G3P increases 1.9 and 2.4 times, respectively.

42

This behavior indicates a role in substrate binding rather than catalysis, similar to H434. The corresponding residue in yeast TK (D477) has been studied previously[135]. D477 was shown to a have a rather large effect on activity and substrate binding. In comparison, DXS shows only a 50% loss of activity that is caused by decreased substrate affinity. This speaks to the difference between DXS and other TK enzymes and highlights the need to study this unique subclass of enzyme.

Pyruvate				
	K_M (mM)	k_{cat}/K_M (s^{-1}M^{-1})	k_{cat} (s^{-1})	%WT
Wild-type	0.28 ± 0.03	2.6×10^4	7.4 ± 0.3	100
H82A	0.23 ± 0.02	1.7×10^3	0.38 ± 0.01	5.1
H304A	1.7 ± 0.5	5.8×10^2	0.90 ± 0.01	12.1
D430A	0.52 ± 0.5	1.4×10^4	7.2 ± 0.2	97.3
H434A	1.7 ± 0.1	5.9×10^3	9.9 ± 0.2	133.8
G3P				
	K_M (mM)	k_{cat}/K_M (s^{-1}M^{-1})	k_{cat} (s^{-1})	%WT
Wild-type	0.05 ± 0.01	1.5×10^5	7.9 ± 0.4	100
H82A	0.03 ± 0.01	1.3×10^4	0.37 ± 0.02	4.7
H304A	0.08 ± 0.02	1.1×10^4	0.90 ± 0.1	11.4
D430A	0.12 ± 0.01	6.6×10^4	7.7 ± 0.2	97.5
H434A	0.23 ± 0.01	4.2×10^4	9.6 ± 0.3	121.5

Table 2.2: DXS steady-state kinetics data (wild-type and mutants) for both pyruvate and G3P. %WT was determined by comparing the mutant k_{cat} to the wild-type k_{cat}.

Two histidine residues are in close proximity to each other (3.7Å between N_ϵ atoms for H82 and H304) and the center of activity (5.1Å and 5.7Å from the thiazolium C2 atom respectively for H82 and H304) of DXS. Table 2.2 shows that the H82A and H304A mutants produce catalytically defective enzymes resulting in only 2-12% k_{cat} and k_{cat}/K_M values when compared to wild-type. The loss of activity can be explained by their proximity to the thiazolium C2 atom. These residues can assist in stabilizing the α-carbanion\enamine intermediate following pyruvate decarboxylation (Figure 2.2, step 5).While activity in these mutants is significantly retarded, detectable levels of activity are retained. This retention might be explained by the proximity of these two residues to one another. Upon the loss of one histidine, it is possible for the other His residue to recover partial functionality. There is one noticeable difference in the results of these two mutants. The H304A K_M for pyruvate has increased compared to the wild type; while K_M value for H82A

remain similar to the wild-type value. This indicates that while both residues are clearly catalytically important, H304A protrudes into the pyruvate binding site and, therefore, plays a role in binding; which can not be replaced by H82. Thus, accounting for observed differences in mutant K_M values for pyruvate.
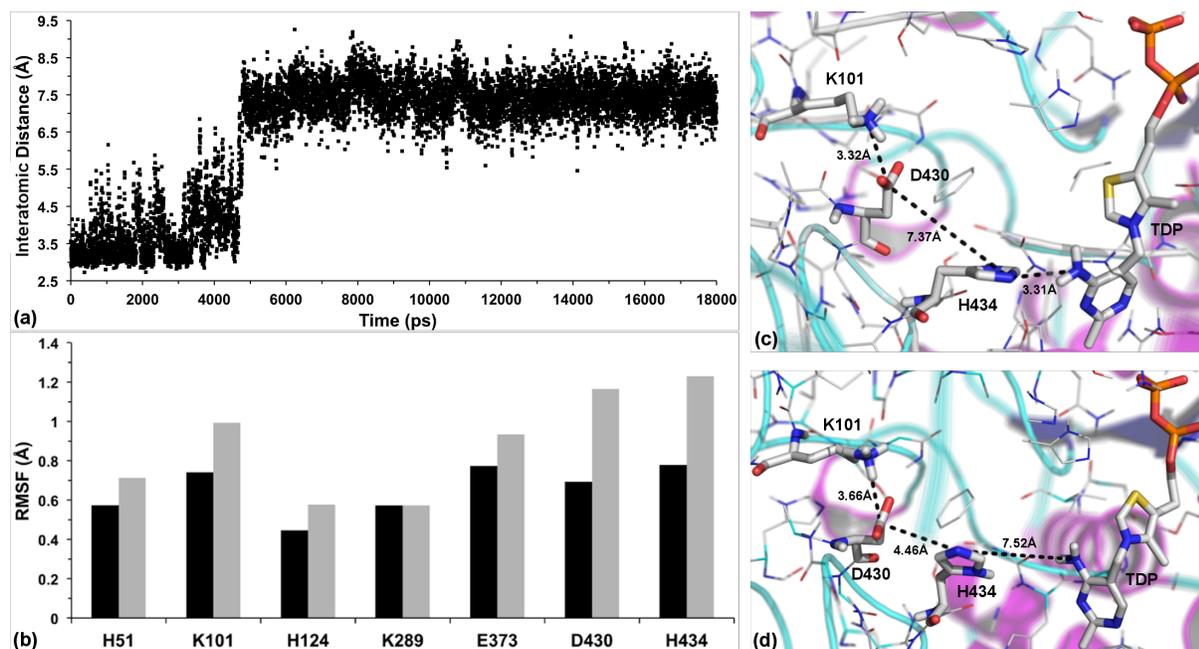
Figure 2.10: Analysis for 18 ns of the unrestrained simulation of the 2O1X DXS structure utilized in this investigation. (a) the distances over time for N_ϵ of H434 to N4′ of TDP's amino group. (b) shows the fluctuations for backbone (black), and side chains (grey) for residues H51, K101, H124, K289, E373, D430, and H434. These residues represent the QM region and key CPA residues previously discussed. (c) snapshot from the 18 ns trajectory with H434 in proximity to the 4′-amino group. (d) is representative of H434 in the second conformation.

With these mutagenesis results, it became apparent that a longer simulation was required to examine active site conformational dynamics. Thus, the 2O1X structure was simulated for an additional 20 ns with the first 2 ns discarded (details found in supplementary information). The trajectories were compared to the QM/MM minimized RS. The distance between the 4′-amino group and the N_ϵ atom of H434 revealed two major conformations (Figure 2.10a). The first conformation lasts for ~5.0 ns and has H434 3.5Å from the AP ring on average. The second conformation has H434 7.1Å from the AP ring on average and remains throughout the simulation. The fact that the 2nd conformation is stable for the majority of the simulation and places the histidine beyond the range of

direct deprotonation of its 4′-amino group provides further support for a water mediated mechanism. Additionally, the backbone and side chain fluctuations were calculated for significant CPA residues (e.g., E373 and H434, Figure 2.10b). The conformational change of H434 to a position proximal to K101 and D430 accounts for the larger side chain fluctuations of K101 and D430 (Figure 2.10). The introduction of H434's imidazole would force K101 and D430 to move in order to accommodate the bulky polar side chain. The combination of the motion of these residues with H434 being the final resting place of the proton abstracted from TDP's amino group suggests a possible regulatory role for H434. H434 could act as a shuttle involved in regenerating the TDP-ylide for further reactions by displacing the abstracted proton onto D430. This perfectly aligns with experimental results showing that the removal of this residue (H434A) slightly increases k_{cat}, allowing any proton transfer from TDP's amino group to D430 to occur more rapidly via a water mediated process (picosecond time scale) rather than the H434 side chain motion that likely occurs on the nanosecond time scale.

	DHM-RS	DHM-TS	ΔNICS
NICS(1)	-7.1	-5.7	1.4
NICS(0)	-5.0	-2.7	2.3
NICS(-1)	-9.3	-6.7	2.6
Average			2.1
	WMM-RS	**WMM-TS**	**ΔNICS**
NICS(1)	-6.8	-4.7	2.1
NICS(0)	-4.6	-2.2	2.5
NICS(-1)	-8.8	-6.3	2.5
Average			2.4

Table 2.3: Calculated NICS values for the WMM and DHM RS and TS. The NICS(0) values are taken from the center of the pyrimidine ring. The NICS(1) and NICS(-1) values are points away and towards a proximal phenylalanine (F398), respectively. A comparison set of benzene (-9.8) and cyclobutadiene (27.6) were computed to show reference aromatic and antiaromatic values, respectively. The average ΔNICS values represent a 29.4% and 35.6% decrease in aromaticity for DHM and WMM, respectively.

TDP reactivity is clearly dependent on the surrounding environment, e.g., rate of reaction increases a billion-fold when bound to an enzyme[84]. Several attempts to determine the underlying energetics have attributed this behavior to the strained 'V' shape

TDP adopts upon binding[25,26,82]. Given the pK$_a$ changes this conformation induces, it is surprising that the barrier to activation (i.e., proton transfer) is significantly higher than expected; $\Delta E^{\ddagger}$=22.7 kcal·mol^{-1} vs. 5-10 kcal·mol^{-1} for typical proton transfers[142]. One possible cause of this is the loss of aromaticity that occurs during ylide formation. Additionally, overestimation of the barrier may be due to the inability to carry out free energy simulations; nevertheless, the energy barrier difference is a more meaningful quantity when seeking to differentiate two possible mechanisms.

To examine the former, i.e., aromaticity effects, both NICS and ASE were computed. NICS calculations estimate the aromaticity of a molecule; negative NICS values indicating aromaticity and positive NICS values antiaromaticity. Table 2.3 reports NICS results for RS and TS of TDP activation via DHM and WMM. An average of the ΔNICS values was used to quantify the relative change in aromatic character. Results indicate the AP ring is aromatic in both the RS and TS with values close to those published for similar pryimidine analogs[1]. However, the TS consistently shows lower aromatic character than the RS; which supports our hypothesis of aromaticity regulating ylide formation.

Calculating the ASE for TDP's AP state should provide additional information about the importance of aromaticity in ylide formation. A homodesmotic reaction (Figure 2.11) provides a reference for determining ASE. Thiamin serves as a model compound for this purpose and represents the key components (e.g., 4′-amino and thiazolium moieties) of TDP. Systems with positive values of ASE are considered to be aromatic, whereas those with negative values are antiaromatic. ASE values are determined as the difference in energies between both halves of the reference reaction (Figure 2.11). Thiamin has an ASE
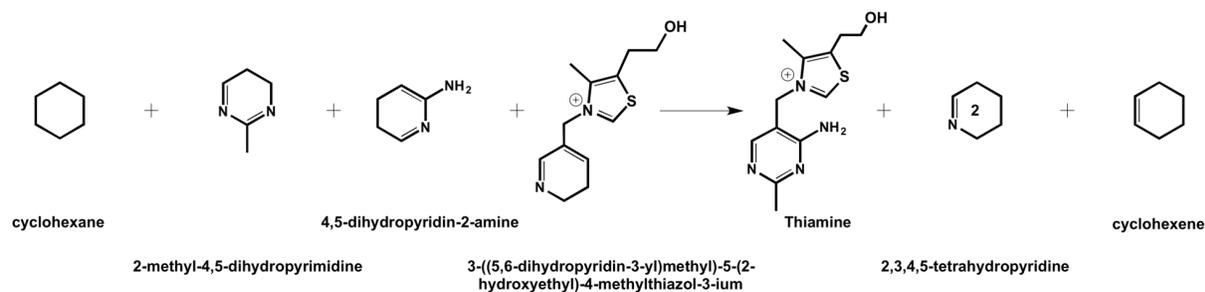
Figure 2.11: Homodesmotic reaction used in evaluating the aromatic stabilization energy for a model TDP.

of 37.6 kcal·mol^{-1}; which is again in close agreement with previously published results of similar pyrimidine derivatives[1].

Combining ASE values with the average decrease in aromaticity (i.e., ΔNICS, Table 2.3), we approximate the stabilization lost at the TS of each mechanism. Aromaticity losses of 13.4 kcal·mol^{-1} and 11.1 kcal·mol^{-1} for WMM and DHM were computed, respectively. Interestingly, we again observe the DHM TS being less destabilized when compared to the WMM TS. This provides further evidence that DXS is well adapted to stabilizing alternative mechanisms of TDP activation. Finally, when the total barrier heights are considered it becomes clear that the loss of aromaticity plays a major role in TDP activation and the initial step of isoprenoid biosynthesis.

2.6 Conclusion

The $\Delta\Delta E^{\ddagger}$ of 10.0 kcal·mol^{-1} difference between the WMM and DHM mechanisms indicates the WMM is the energetically favorable route for ylide formation in DXS. The RPATh+RESDi results seem to suggest the mechanism of proton transfer acts in a concerted fashion proceeding via the tautomeric route between the AP and IP state. Further investigation is ongoing to confirm the relative energetics of a step-wise versus concerted mechanism.

CPA results were indicative of H434 playing a role in long range electrostatic stabilization; which is more clearly illustrated upon examination of the RS and TS active site dipole moments. Mutagenesis studies performed reveal H434 to play a role in substrate binding but not likely a direct role in catalysis. The H434A mutant results reinforce the CPA results. Additionally, a D430A mutant revealed a lower catalytic significance for DXS in comparison to the corresponding yeast TK mutant[135]; again illustrating mechanistic differences. Furthermore, H82A and H304A DXS mutants showed significant decreases in activity (2-12% of wild-type). Given their proximity and retention of measurable activity, it is likely these residues function as back-ups to each other. This comports nicely with computational results that indicate DXS is well suited to functioning via alternative mechanisms (i.e., different general bases), something that would offer a significant evolutionary advantage.

$\Delta E^{\ddagger}_{WMM}$ is significantly higher than what might be expected for a relatively simple proton transfer[142]. CPA results could not account for such behavior. However, upon examination of the 4'-amino moiety, it was evident from structural changes that aromaticity may be changing. The results of NICS and ASE calculations showed that indeed the AP ring was losing aromaticity. If the aromatic contribution is taken into consideration, the new $\Delta E^{\ddagger}_{WMM}$ would be closer to 8.5 kcal·mol^{-1}; which is in the range of similar reactions. This clearly shows that loss of aromaticity plays a key role in controlling activation of TDP in DXS. Further, restoration of this aromaticity upon intramolecular proton transfer from the C2 of the thiazolium ring to the 4'-amino group should ultimately drive the final ylide formation.

2.7 Supporting Information (SI)

The following can be found in Appendix A: PDB files for the transition states of the WMM and DHM; link atom details for QM/MM reaction path calculations; CPA results; topology and parameter files for TDP; methods for the over-expression and purification of DXS and DXR; steady-state initial velocity plots with varying concentrations of G3P and pyruvate; details about 20ns simulations; results of PROPKA3.1 calculations on DXS Chain A.

2.8 Acknowledgments

Chapter 3

Computational Examination of the Magnesium Ion Binding Modes of 1-Deoxy-D-xylulose 5-Phosphate Reductoisomerase

3.1 Introduction

There exists a vast and varied class of natural products derived from two five-carbon iso-prene precursors, isopentenyl diphosphate (IDP) and dimethylallyl diphosphate (DMADP), and serve several essential roles for all living organisms[144]. These are generally known as Isoprenoids. The broad variety of uniques molecules comprising this family are derived via a combination of elongations, rearrangements, cyclizations, and oxidations utilizing IDP and DMADP in various combinations[38]. A few of the important biological roles filled by isoprenoids are prenyl lipids in archaebacteria[46], sterols in eubacteria and eu-karyotes[128], light-harvesting pigments such as carotenoids, electron transport carrier such as ubiquinone and menaquinone, and several growth and development regulators (Figure 3.1)[161]. Additionally, there are several known herbicides or herbivore repellents identified to be isoprenoids[48].

The biosynthesis of the IDP and DMADP building blocks were originally thought to derive from a single enzyme pathway (Figure 3.1[132]. This pathway is known as the mevalonate dependent (MVA) pathway; which was named after the key committed inter-mediate formed from the condensation and reduction of 3 acetyl-CoA molecules producing mevalonic acid (or mevalonate in ionic state). Continuing discrepancies in the results of isotope labeling studies[30,31,57,140,205] led several researchers to postulate the existence of a second yet unidentified pathway. Efforts by researchers such as Rohmer, Arigoni, Lich-
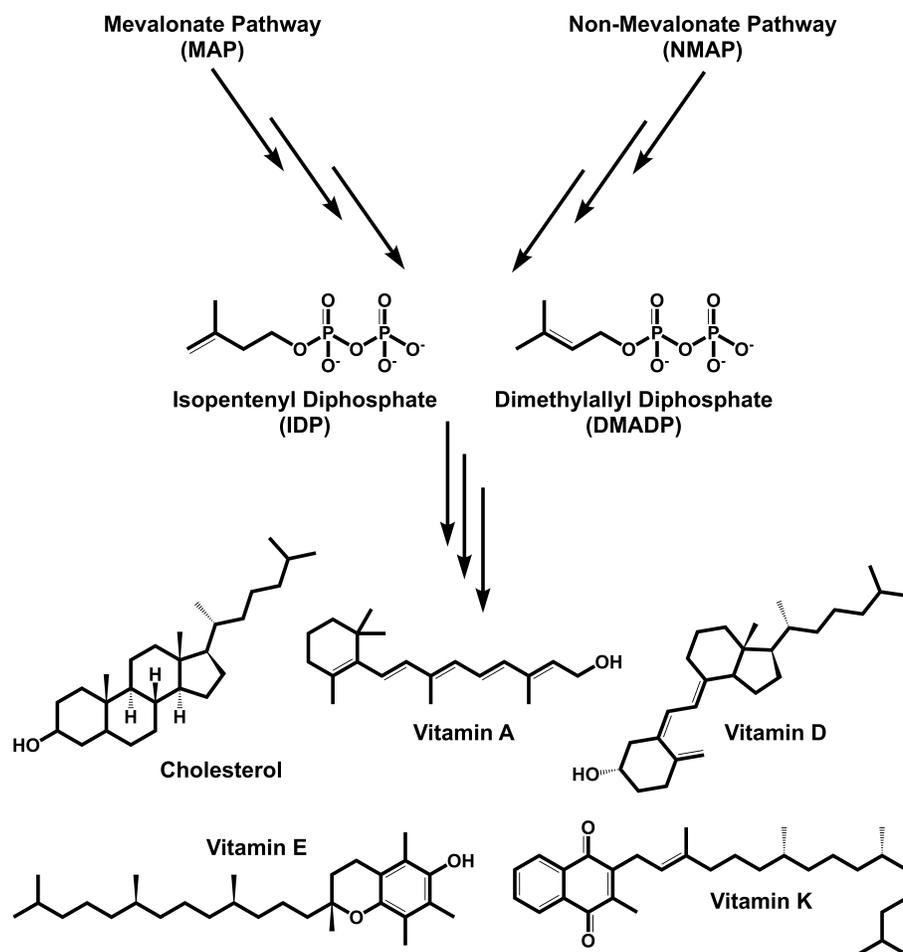
Figure 3.1: Illustration of the two biosynthetic pathways of IDP and DMADP and representative isoprenoids produced from the building blocks

tenthaler, and Seto, *etc.* eventually discovered a new pathway completely distinct from the MVA pathway[32,91,111]. Initially, the pathway had names reflecting the distinction between MVA and the novel, mevalonate-independent or non-mevalonate (NMA) pathway. Seven enzymes catalyzing 8 reactions comprise the NMA pathway[132]. Rohmer and co-workers[159] established the formation of 1-deoxy-D-xylulose 5-phosphate (DXP) via the decarboxyl condensation of pyruvate and glyceraldehyde-3-phosphate (G3P). DXP is additionally utilized as an intermediate for the biosynthesis of vitamins B1 and B6 as well as isoprenoid biosynthesis[114,160]. Therefore, DXP is required and considered the first step of the pathway but not the committing step in the NMA pathway. The succeeding reaction catalyzed by DXP reductoisomerase (DXR) bares the distinction of being

the committed step in the NMA pathway. The product of this reaction, 2-C-methyl-D-erythritol 4-phosphate (MEP), lends its name to the pathway, as the NMA pathway is frequently referred to as the MEP pathway[100].

DXP reductoisomerase catalyzes a carbon-skeleton rearrangement of DXP and subsequently reduced. DXR activity requires a divalent metal cation cofactor and NADPH co-substrate. Out of the divalent metal ions attempted, DXR is activated by only Mn^{2+}, Co^{2+} and Mg^{2+}, in decreasing order respectively[3,94,102,118,179,202]. Additionally, NADH was tested in place of NADPH as the co-substrate with DXRs derived from *E. coli*, *M. tuberculosis* and *S. leopoliensis*. The results were a decrease in activity in some cases as much as a 100-fold decrease in activity. Decreased affinity is responsible for the lost activity primarily due to the loss of the 2′-phosphate of NADPH. Therefore, the phosphate is a binding determinant and not likely directly involved in catalysis since k_{cat} was unaffected[3,179].

Results of isotopic labeling studies demonstrated the required isomerization proceeding via a C3/C2 bond transition. The isomerization results in an aldehyde intermediate 2-C-methyl-D-erythrose 4-phosphate (MEsP), which is subsequently reduced on the *re* face of MEsP by the C4 pro-S hydride of NADPH[4,5]. The proposal of this intermediate was originally based on analogy to ketol-acid reductisomerase (KARI), which catalyzes a similar reaction during the biosynthesis of branched-chain amino acids. As with the KARI reaction, the MEsP aldehyde intermediate has never been directly detected[50,100,179]. Several attempts have been made to isolate the aldehyde intermediate with no success. These results suggest the intermediate might be more transient than originally thought or very tightly bound prior to NADPH reduction[78,179];which has been similarly proposed for KARI[50]. Rohmer and co-workers produced the first compelling evidence supporting the aldehyde intermediate theory by introducing exogenously synthesized MEsP and demonstrating kinetic competency. When incubated with DXR in the presence of NADPH and Mg^{2+} or Mn^{2+}, a factor of 4 and 1.6, respectively, increase in conversion to MEP was observed. While the oxidized coenzyme was present, a 7% conversion of MEsP to DXP

was detectable[37]. Additionally, the K_m for MEsP was found to be greater than DXP by a factor of 4 and 1.6 in the presence of Mg^{2+} or Mn^{2+}, respectively. The argument has been made based on these values against the tight-binding of MEsP. This is flawed though since K_d and K_m are only equal when substrate dissociation is rapid[78,137].

Despite the overall similarities of DXR- and KARI-catalyzed reactions, amino acid differences suggest different mechanism of action[37,49,95]. Three mechanisms were proposed to explain the carbon-skeleton rearrangement: 1) an α-ketol rearrangement, 2) a retro-aldolization/aldolization, and 3) a sequential 1,2-hydride and 1,2-methyl shift[64]. A dismissal of the third mechanism was accomplished based on ^{13}C-glucose and ^{13}-DXP incorporation studies[3,72]. Therefore, further investigations looked to distinguish between the remaining α-ketol rearrangement or retro-aldol/aldol mechanism (Figure 3.2).

The retro-aldol/aldol mechanism should form 2 putative intermediates of glycoaldehyde phosphate and the enolate of hydrxyacetone. If these intermediate could be detected during or following the reaction, it would provide strong evidence in support of the retro-aldol/aldol mechanism. Several attempts were made with no success[58,78,104]. Though the lack of detection is consistent with both mechanisms as the results can be explained as the intermediates fragments are tightly confined to the active site. Additionally, these putative fragments could be so transient, they never truly form. Subsequent, experiments have tended to favor the retro-aldol/aldol mechanism, such as the modification or removal of the C4 hydroxyl group. The α-ketol rearrangement doesn't require the C4 hydroxyl group and therefore any turnover would support. Though turnover was not observed for 1,4-dideoxy-D-xylulose 5-phosphate,the K_i values similar to the K_m indicates a dependence on the C4 moiety for turnover but not binding. The C4 epimer and fluorinated version of DXP produces similar results[143,193]. Due to the relatively good binding of these modified ligands, the retro-aldol/aldol mechanism is favored.

The analogues studies have provided some significant evidence in support of the retro-aldol/aldol mechanism over the α-ketol rearrangement. Kinetic isotope effects (KIEs) provide a means of further probing the mechanism. In order to differentiate between

α-Ketol Rearrangement **Retro-Aldol/Aldol**

DXP

DXP

Methylerythrose phosphate

MEP

NADP$^+$ NADPH
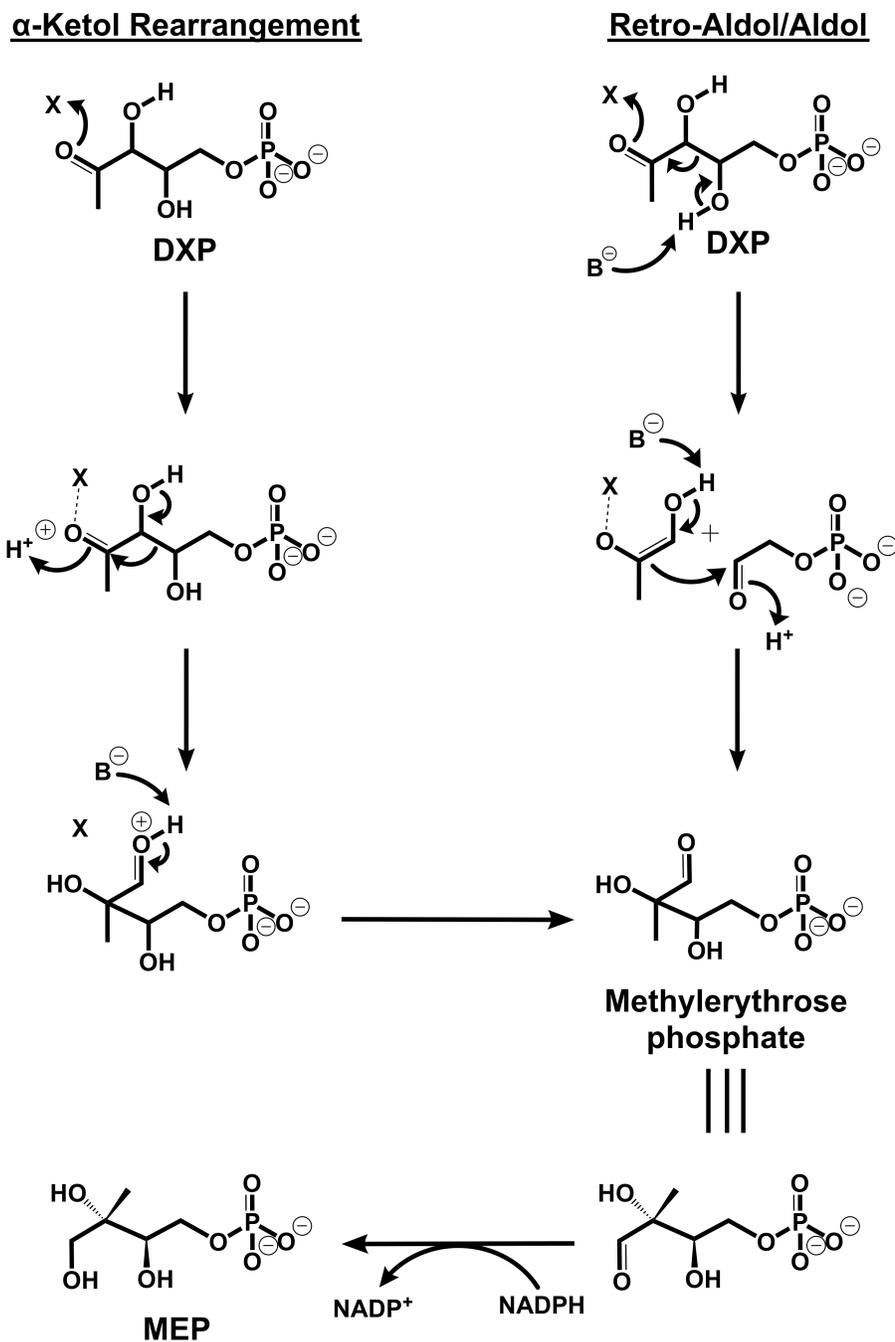
Figure 3.2: The above illustration compares the steps for the α-ketol rearrangement and retro-aldol/aldol mechanisms in a side-by-side view. Each concludes in the aldehyde intermediate 2-C-methyl-D-erythrose 4-phosphate (MEsP), which is subsequently reduced by NADPH to form the 2-C-methyl-D-erythritol 4-phosphate (MEP) product.

the two mechanisms, α-secondary KIEs were measured for [3-^{2}H]- and [4-^{2}H]-DXP. The α-ketol rearrangement predicts a shift from sp^3 to sp^2 at the C3 position while the C4 position remained sp^3, which translates into KIEs>1 and unit KIE values, respectively.

In contrast, the retro-aldol cleavage both C3 and C4 undergo changes from sp^3 to sp^2 with normal KIE values (KIE>1)[131]. The results of 1.04 for [3-^{2}H] and 1.11 for [4-^{2}H]-DXP supports the retro-aldol/aldol mechanism. When compared to muscle aldolase, which has a similar mechanism, the lower KIEs are thought to reflect the partially rate-limiting rearrangement or an early transition state[131]. Finally, a 2D [^{13}C,^{1}H]-HSQC NMR based technique was used to analyze ^{13}13 KIEs. The method measures the reactive competition between light and heavy C substrates in the same mixture with the enzyme. The ratio of ^{13}C/^{12}C represents the KIE. The ratios were measured for 2-, 3-, and 4-^{13}C with results of 1.0031, 1.0303 and 1.0148, respectively[120]. The sigmotropic rearrangement would result in large changes at all locales while retro-aldol predicts larger changes at the C3 and C4 position with little effect on the 2C position. The results of these KIE experiments supports the retro-aldol/aldol mechanism as the most likely mechanism. The only major issue left to challenge the retro-aldol/aldol mechanism is the failure to detect the putative hydroxyacetone and glycoaldehyde intermediates. Currently, the best explanation revolves around the tight binding of these intermediates and/or the molecules exist in such a high energy state, they aren't around long enough to be a true intermediate[132].
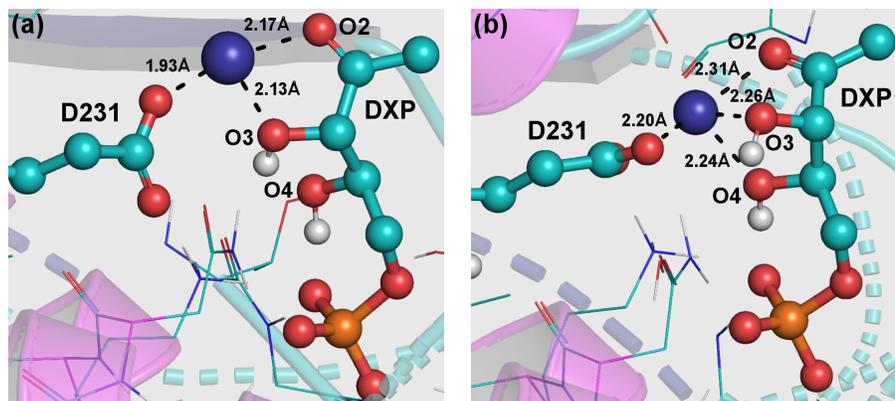
Figure 3.3: Illustration of the C2-C3 and C3-C4 binding modes in the reactant state. These structures were utilized for the purposes the replica path calculcations as the starting point.

The retro-aldol mechanism was originally proposed with the metal ion coordinated between the C2-C3 hydroxyl groups. Results of incubating DXR with Mg^{2+}, NADPH and DXP in the presence of ^{18}O-labeled water to explore the incorporation of the isotope into MEP[109] dispute this proposal. Retro-aldol/aldol mechanism produces carbonyls, if transiently, at each position during the reaction, thus allowing for solvent exchange at the C2, C3, and C4 positions of DXP. Since the only hydroxyl affected was the C2 of DXP, there had to be a protective effect at the C3 and C4 positions. Coordination of the divalent ion would act as protection from solvent exchange, therefore the results suggest a C3-C4 binding mode. This binding mode helps to explain the tight binding of both fragments since the retro-aldol cleavage occurs along the C3-C4 bond[74,118,172,178]. During the bond breaking and subsequent C2-C4 bond forming steps, the Mg^{2+} would remain coordinated to both fragments inhibiting release. It is still possible to interpret the results of these experiments in support of the C2-C3 binding mode. The Lewis acid characteristics of the metals would increase the electrophilicity of the C2 carbonyl thus promoting hydration. Furthermore, the transiency of the intermediates may explain the lack of solvent exchange. Exchange of the C1 or C3 oxygen atoms requires the rate of on-enzyme hydration to rival rates of hydride transfer and aldolization[132].

Examining the energetics of the metal binding modality will be the focus of this work. Mac Sweeney et. al. published a crystal structure of *E. coli* DXR (PDB:1Q0Q) with DXP and NADPH bound in the active site[118]. The experimental results published to this point provide strong support for the retro-aldol/aldol mechanism being the most likely reaction mechanism, so it was decided to focus on this pathway for our calculations. In particular, we focused on the retro-aldol calculation, which is thought to be the true limiting step of this reaction. The putative intermediates are even thought to not be proper intermediates but possibly transition states. QM/MM techniques were utilized to compute the free energy surface of the retro-aldol reaction with the metal ion in the C2-C3 or C3-C4 position.

3.2 Computational Methods

The crystal structure published by Mac Sweeney et. al. (PDB:1Q0Q) was utilized for all calculations in this paper[118]. Although, DXR is generally found to be in a homodimer in solution, there is no evidence currently supporting catalytic interdependence of active sites. Thus, allowing up to focus on a single monomer. The structured was parsed utilizing www.charmming.org[127]. Parameters for DXP were built based on similar structures already found in the CHARMM General Force Field (CGenFF)[184]. The necessary bonds were added based on the most similar structures and the charges were corrected via quantum mechanical calculation. Final validation was performed utilizing the crystal structure as the comparison.

The protein was built and E234 was protonated based on values determined by ProPKA3.1 (see Appendix B)[138,175]. CHARMM22 protein and CGenFF force fields were used throughout these calculations[119]. A Mg^{2+} ion was built separately and added to the composed enzyme. The ion was brought into C2-C3 and C3-C4 orientation via use of the harmonic distance restraint (RESDi) while fixing the rest of the system followed by an unrestrained minimization. The system was solvated in a rhombododechedron crystal structure and neutralized with KCl salt to a final concentration of 0.15M. The system was heated from 110K to 310K over 100ps and equilibrated for 200ps at constant pressure (1atm) and temperature (310K). The total system size was subsequently reduced to cut down on computational costs in the following QM/MM calculations by removing all waters/ions beyond 12Å from the protein surface. The reduced structure was treated to a QM/MM minimization without cut-offs to a tolerance of 0.002 kcal·mol^{-1}·Å^{-1}. All QM/MM calculations employed the Q-Chem4.0\CHARMM[27,170,194] interface at the B3LYP/6-31G* level of theory[13,68,105]. Additionally, the single link atom scheme was used to account for truncation of the QM region and employed group electrostatic exclusions to prevent over polarization of the QM region (23 atoms)[167]. The QM region was defined as the D231, MG, and DXP only. The NADPH molecule was excluded because it is not thought to play a role in the skeletal rearrangement.

Reaction path calculations were performed using the Replica Path (RPATh) in combination with RESDi values to define the steps along the reaction coordinate[36,194,195]. RPATh allows the user to define a subsection of the structure which will be duplicated into replicas. These replicas are free to react normally as the reaction progresses while the larger system is constrained thus cutting down on the computational costs. The replicas utilized in these calculations were included the QM region and a buffer region of 6.5Å around the QM section. For our purposes, two replicas were utilized. In order to provide a buffer from the constrained system, the replicas were defined as all residues within 6.5Å from the QM region. One replica was incrementally progressed along the reaction. This was performed by defining two reaction coordinates (δ_1, δ_2); which were defined with reference to reaction component being controlled.

$$\delta_1 = Bond-Breaking_{C3-C4}; \delta_2 = Bond-Breaking_{O4-H9} - Bond-Forming_{H9-OE2}$$

(3.1)

As the retro-aldolization is composed of two parts, δ_1, corresponding to the breaking of DXP's C3-C4 bond, could be easily be defined while δ_2, corresponding to the deprotonation of the C4 hydroxyl by residue D231, was defined as a linear combination of distances. A two-dimensional energy surface was produced with respect to these reaction coordinates. While the progression along δ_2 was easily defined as beginning at -2.0 (reactant state) and ending at 2.0 (intermediates) as an assumption like previous work, the path of δ_1 was more difficult. Since δ_1 refers to a single bond breaking, the C3-C4 bond of DXP was elongated by 2.0Å. δ_2 was progressed in 0.2Å increments and δ_1 was allowed 0.1Å increments for a 21x21 point 2D surface. After the completion of these calculations, normal mode analysis was utilized to identify the reaction steps corresponding to states of interest ("products")[180]. This was performed by QM calculations utilizing Q-Chem as "freq" jobs. The output frequencies were analyzed for unique asymmetric vibrations corresponding the changes desired.

3.3 Results and Discussion

The first response to the complete 2D-energy surface indicates our initial ranges for the reaction coordinates may have been too broad. Both the C2-C3 and C3-C4 show a range of values produce very strained structural states. A few of these values were repeated to verify with similar results, therefore the rest of the work focused on results prior to δ_2=0.8-2.0 for C2-C3 and δ_2=1.2-2.0 for C3-C4 calculations. The structures present structures representing over extended CO bonds and massively contorted structures.

Before continuing discussions of the energy results, it is important to discuss the identification of the "products"; which correspond to the putative intermediates between the retro-aldol and aldol steps of the DXR reaction. Systems with unique normal modes were found at $\delta_1;\delta_2$=3.40;0.60 for C2-C3 coordinated state and $\delta_1;\delta_2$=3.40-3.60;1.00 for the C3-C4 metal coordination. This results supports the proposal to exclude the results mentioned previously.
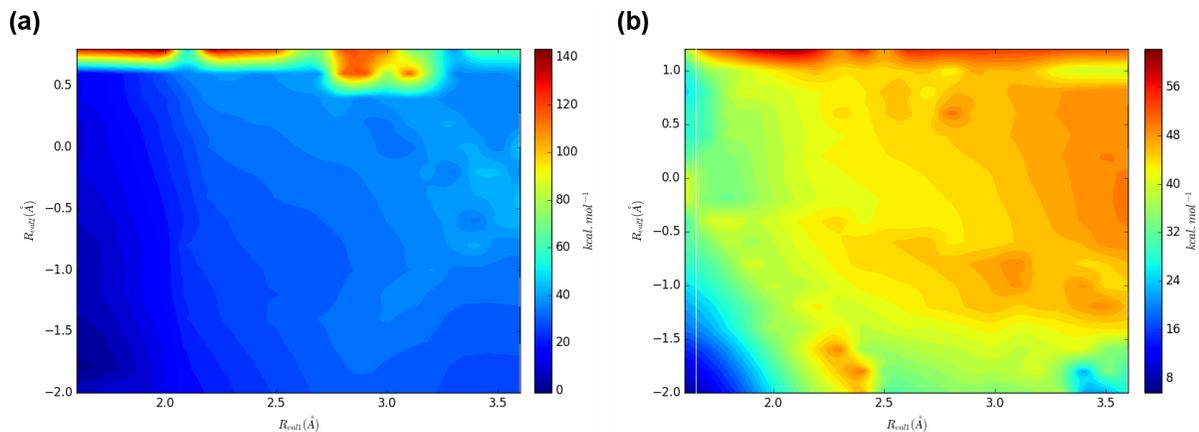
Figure 3.4: Two-dimensional energy surfaces with outlier values removed. With the outliers for the C2-C3 (image (a)) and C3-C4 (image(b)) removed from the surface plot, the details are more easily observed. The valley in the top right of image (b) might indicate a step-wise mechanism.

The results of the reaction path calculations (Figure 3.4) reveal some distinct differences between the binding modes. While the C3-C4 binding mode has distinct peaks and valleys, the C2-C3 surface is ever increasing (Figure 3.4a). This result is not really grounds for dismissal of the binding mode. As the "products" of the retro-aldol reac-
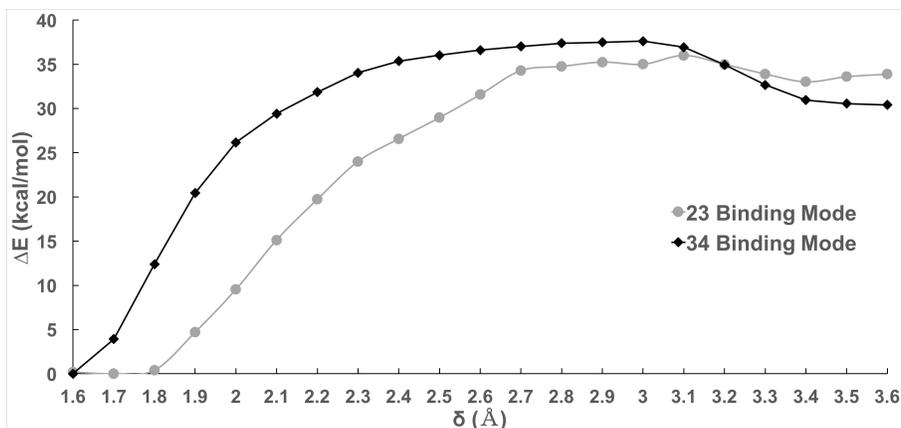
Figure 3.5: One-dimensional representation of the center path across the two-dimensional energy surfaces for the C2-C3 and C3-C4 binding modes.

tion are, in actuality, intermediates or more likely transient transition states during the skeletal rearrangement phase of the DXR reaction. Both binding modes conclude at high energy states (33.0 kcal·mol^{-1} and 30.9 kcal·mol^{-1} for C2-C3 and C3-C4 binding modes, respectively).

For easier comparison, scatter point plots of pathways representing the best path between points were produced (Figure 3.5). The results suggest the binding modes to be rather similar. The barrier energies are 36.0 and 37.6 kcal·mol^{-1} for the C2-C3 and C3-C4 mode, respectively. So overall, the energetics of the C2-C3 compared to C3-C4 isn't sufficient to address the question binding modes. A structural comparison of the "products" in conjunction with the energetics makes for a different outcome. The retro-aldol reaction needs a deprotonation to activate the breaking of the C3-C4 bond producing the hydroxyacetone enolate and glycoaldehyde phosphate "products"[132]. The base is proposed to be the D231 residue found in proximity of the DXP hydroxyl groups, which was controlled by the δ_2 reaction coordinate. Figure 3.6 illustrates the "products" states of each binding mode. The C3-C4 binding mode (Figure 3.6b) are a clear representation of the intended hydroxyacetone enolate, glycoaldehyde phosphate and protonated D231 residue "products". The "product" state for the C2-C3 binding mode (Figure 3.6a) reveals an intermolecular protonation of the phosphate group despite the presence of a

restraint directing proton transfer. There are two possible explanations for this differ-
ence configurations. The C3-C4 binding mode held the C4 hydroxyl group in a favorable
position for deprotonation by the D231 oxygen. Secondly, the metal, acting as a lewis
acid, could have further polarized the O4-H bond, thus promoting deprotonation by the
glutamate.
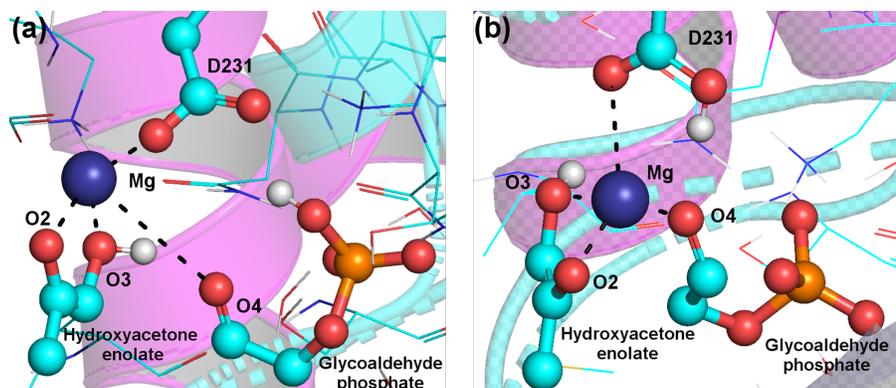
Figure 3.6: The final structures of "products" for the retro-aldol reaction are the hydrox-
yacetone enolate and glycoaldehyde phosphate at the bottom of the two images above.
Image (a) above represents the C2-C3 binding mode that results in a intermolecular pro-
tonation of the phosphate group and preferential coordination with the C2-C3 oxygens.
On the right, image (b) is the C3-C4 binding mode which produced the desired products
and protonation states.

Along with the possible structural highlights for the mechanism, figure 3.6 illustrates
another key difference in the binding modes. Namely, the position of the metal ion after
preparation and RPATh calculations. While the magnesium remains straddling the O2
and O3 atoms of the C2-C3 mode (Figure 3.7a,c), the C3-C4 binding mode actually
transitions into an all oxygen coordination (Figure 3.7b). The conformational change
occurs spontaneously after the heating and equilibration phases of the build phase. Figure
3.7 reveals the differences in coordination between the pre- and post-equilibration steps
for each binding mode. Figure 3.7b shows the Mg^{2+} to rest 2.76 Å from the O2 atom prior
to equilibration, which likely represents a local minimum on the energy potential. After
the injection of energy from heating and equilibration, the distance from the O2 atom
reduces to 2.31 Å. The conformational change observed between these steps occurred
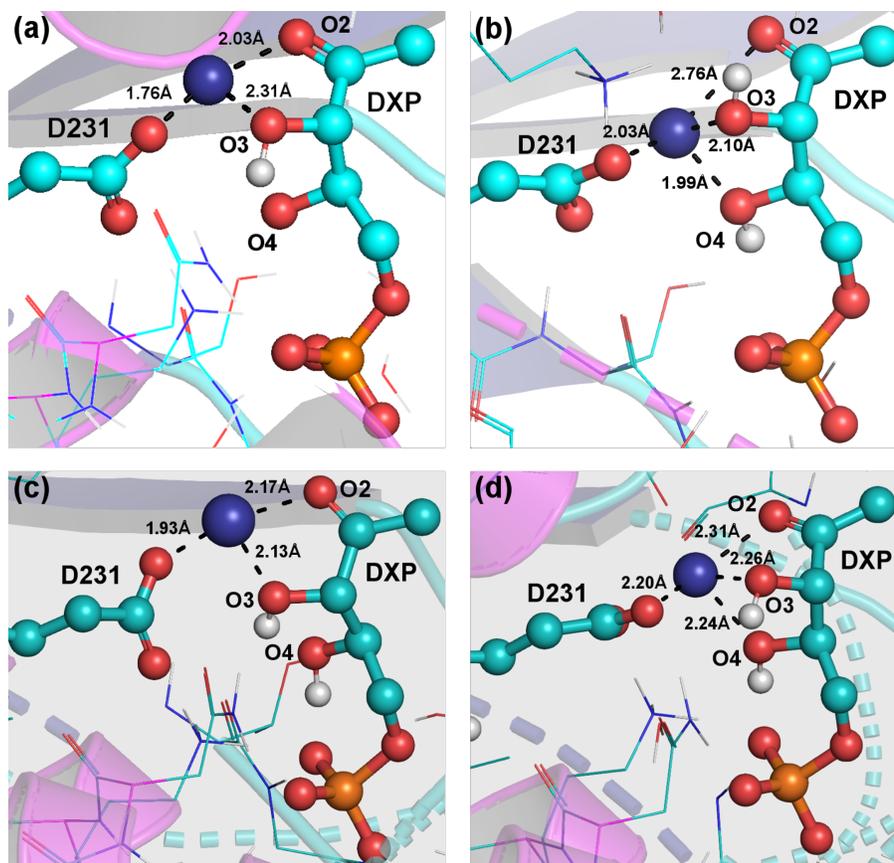
Figure 3.7: Pre-equilibrium and post-equilibrium for for each binding mode. Images (a) and (c) represent the C2-C3 binding mode pre-equilibration and post-equilibration, respectively. Images (b) and (d) represent the C3-C4 binding mode pre-equilibration and post-equilibration, respectively. The side-by-side comparison highlights the changes made.

with no restraints on the system. In contrast, the C2-C3 structures show no interesting in reaching out the O4 atom of DXP. The post-equilibrated structure looks like what one might expect. The originally minimized structure shows some leveling out between the coordination bonds.

The putative intermediates, hydroxyacetone enolate and glycoaldehydephosphate, and reaction intermediate, MEsP, have never been observed directly[78,179], which is the remaining hope for the α-ketol rearrangement mechanism. An explanation is the tight binding of the reactants in the active site[104]. This spontaneous coordination to all DXP oxygens might be further support for the tight binding hypothesis. While the C2-C3 binding mode would only effect the binding of the hydroxyacetone intermediate, the glycoalde-

hyde phosphate might actually be able to leave the active site. Additionally, an all oxygen coordination might aid in the subsequent aldolization by holding all the intermediates together and aiding in stabilizing the C2-C4 bond.

As previously mentioned, the C3-C4 energy surfaces have values with mechanistic implications. Generally, the retro-aldolization requires a deprotonation of the C4 hydroxyl group, and is thought to occur concurrently with the C3-C4 bond breakage (Figure 3.8c). The energetics displayed during these calculations are the first hints of a step-wise retro-aldol reaction beginning with the proton transfer from a highly polarized hydroxyl group and proceeding to the C-C bond breakage. The other valley corresponding to C-C bond breakage occurring first does not contain any structures with unique normal mode frequencies, so this valley is probably an outlier in the data. Figures 3.8a,b are scatter point plots of each proposed step (figure 3.8c, respectively. The $\Delta E^{\ddagger}$ for the deprotonation of 30 kcal·mol^{-1} is considerably higher than one might expect. Therefore, the $\Delta E^{\ddagger}$ indicates the deprotonation as the rate-limiting step. The relatively high energy of the deprotonated state produces a reduced $\Delta E^{\ddagger}$ to the final retro-aldolization "products" produced by C3-C4 bond breakage. These results are far from conclusive but provide a new area of further study.

3.4   Conclusion

The energetics of the two binding modes suggest a preference for the C3-C4 binding mode over the C2-C3 binding mode. It is a slight difference of 2.1 kcal·mol^{-1}; which means the energetics aren't definitive. The combination with the fact the C2-C3 "products" show an intermolecular proton transfer suggests the C2-C3 binding mode to be unfavorable.

In addition to the energetics, configurational differences between the binding modes provide further evidence in support of the C3-C4 binding mode. The spontaneous shift of the C3-C4 mode into a C2-C3-C4 mode enhances the arguments for DXR strongly binding the intermediates; which explains why they haven't been directly observed. Reaction assistance provided by the expanded binding mode could explain the formation
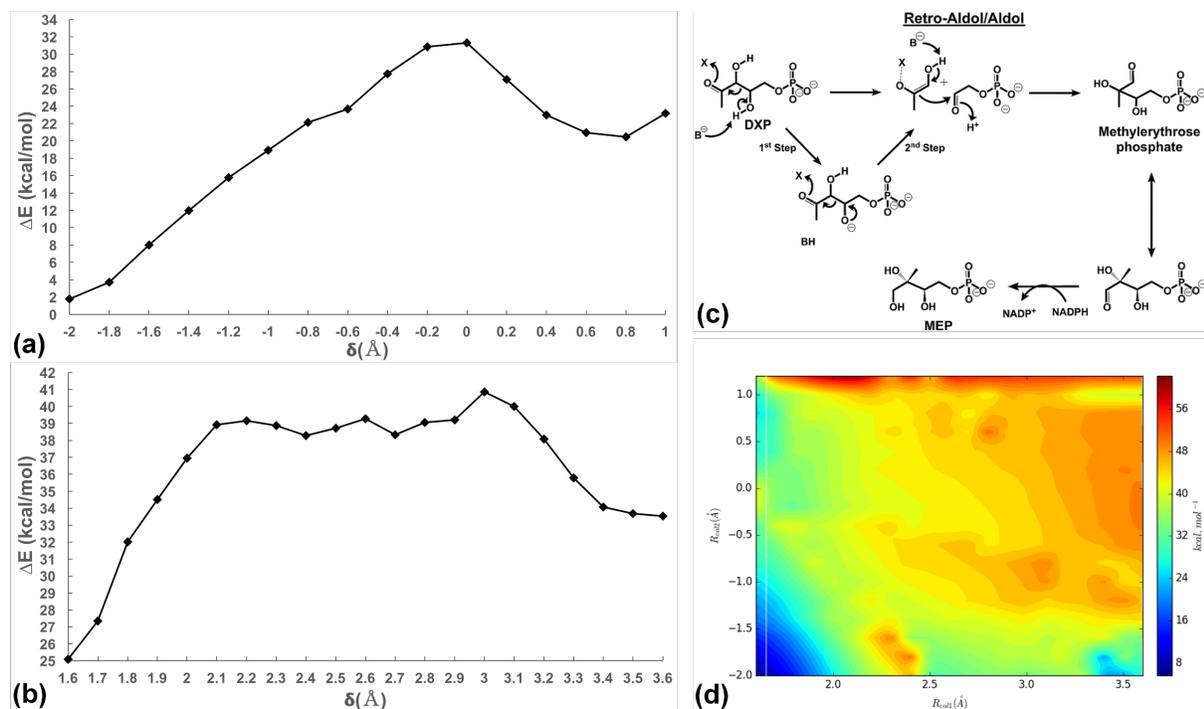
Figure 3.8: Image (a) is the scatter point plot for the deprotonation step while image (b) is the C3-C4 bond breakage. The new step-wise process and old concerted process is illustrated in image (c) with a quick reference to the 2D energy surface for the C3-C4 mode retro-aldol reaction.

of the desired "products" unlike the C2-C3 while holding the intermediate in proximity necessary for the aldolization. This aldolization assistance could provide further evidence these intermediates being truly transition states.

Further work should start with analyzing the changes in active site contributions over the reaction. The mapping of the aldol reaction should be performed starting at the "product" state of the retro-aldol reaction for both binding modes to see if there is a change in preference between the stages of the skeletal rearrangement.

3.5 Supporting Information (SI)

The following can be found in Appendix B: The ProPKA3.0 results of crystal structure for DXR produced by MacKerrell (PDB:1Q0Q).

Chapter 4

Conclusion and Future Work

The work shown in this document represent the initial steps in gaining understanding of the reactions involved in the NMA pathway for Isoprenoid biosynthesis. The work on 1-deoxy-D-xylulose 5-phosphate (DXP) synthase (DXS) focused on the deprotonation of the N4 atom of the thiamine diphosphate (TDP) cofactor, which occurs in preparation of the formation of the ylide via deprotonation of the C2 atom. This step highlights the significance of QM/MM reaction path calculations. Work similar to this can provide insights into reaction steps of an enzyme or even a portion of a step. Kinetics can provide similar insights but are dependent on the step of interest being the rate-limiting step, which provides no assistance with DXS and other TDP dependent enzymes. The ylide activation step is required for activity but happens at the rate of diffusion so experimental practices are currently ineffective. By utilizing computational techniques, it is possible focus on pieces of a reaction and tweeze out pieces of information of significance. Both experiments in this document deal with half of a reaction commonly referred to as a single reaction because they can't be measured experimentally.

4.1   1-Deoxy-D-xylulose 5-Phosphate Synthase Summary and Conclusion

The mechanism of DXS consists of 3 major pieces: ylide activation, pyruvate binding, and pyruvate decarboxylation couples with transferral of acetyl group to glyceraldehyde-3-phosphate (G-3-P). The conclusion is the production of DXP; which is utilized in the production of the isoprenoid precursors of isopentenyl diphosphate (IDP) and dimethy-

lallyl diphosphate (DMADP). The rate-limiting step of the reaction is the pyruvate step. The ylide formation is thought to happen at the rate of diffusion but is required for enzymatic activity since the ylide acts as the reactive center for the enzyme. Prior to activation, a deprotonation happens at the N4 atom of the TDP pyridinium ring producing a tautomeric transition between the 4′-aminopyrimidine (AP) state to the 1′,4′-iminopyrimidine (IP) state. The identity of a general base was unknown which could not be easily determined using experimental techniques. A water-mediate mechanism (WMM) or direct histidine mechanism (DHM) mechanisms were proposed primarily comparison with the other enzymes sharing sequence and structural similarities to other TDP-dependent enzymes and transketolases, in particular. While the active sites of this family of enzymes have a high degree of similarity, the recent discovery of an enzyme deplete of acid/base residues in the active site and a transketolase lacking the requisite histidine provided impulse to investigate DXS further. A reaction coordinate, define as bond-breaking minus bond-forming, was used to incrementally change the system between the reactant and product states

The WMM proposal was found to be preferential by a 11 kcal·mol^{-1} difference in barrier heights. Computational results can be used in tandem with experimental results to help explain or reinforce conclusions made based on the experimental work. A H434A mutant revealed an effect on substrate binding while not effecting turnover. Thus, H434 was proposed not play a direct role in catalysis; which was supported in computationally via charge perturbation analysis. The charge of the H434 residue was artificially turned off and the result compared to the active site when the charge was on. The shift in the dipole resulting from this change illustrated the significance of this residue on long range electrostatics but no direct role in catalysis. Thus, the experimental work was bolstered and explained by the computational results.

Utilizing computational techniques, it is possible to investigate contributing factors otherwise inaccessible through experimental methods. The $\Delta E^{\ddagger}_{WMM}$ was significantly higher than what would normally be expected for a simple proton transfer. A value of

22.7 kcal·mol^{-1} for the WMM path vs 5-10 kcal·mol^{-1} for representative proton transfers suggested some significant contributions. The deprotonation of TDP's N4 atom produces a tautomerization from the AP state to the IP state, which we realized interrupts the aromaticity of the pyrimidinium ring. There is no method for directly computing the change in aromaticity in an enzyme reaction. There is a method for determining the aromaticity of the base TDP molecule and another for determining the percent change in aromaticity in reaction. By combining these methods, it was revealed by taking into account the change in aromaticity the $\Delta E^{\ddagger}_{WMM}$ would be closer to 8.5 kcal·mol^{-1}. The new value being closer to values previously published supported the conclusion of aromaticity playing a part in higher barrier energy. The higher energetic position of the IP state might also act as a driving force in the deprotonation of TDP's C2 atom and production of the ylide required for TDP-dependent activity.

4.2   1-Deoxy-D-xylulose 5-Phosphate Reductoisomerase Summary and Conclusion

When the work began on 1-deoxy-D-xylulose 5-phosphate reductisomerase (DXR), there was much more of a debate in the literature over the last two mechanisms, α-ketol rearrangement or retro-aldol/aldol mechanism. The reduction by NADPH has been pretty well understood since the enzymes initial characterizations. Along the way, papers were published with secondary kinetic isotope effects (KIEs) which supported the retro-aldol/aldol mechanism with the only remaining hope for the α-ketol rearrangement was the refutation of the putative intermediates of the retro-aldol/aldol mechanism, namely hydroxyacetone enolate and glycoaldehyde phosphate. Since the lack of direct observation could be explained with a relatively simple assumption of being tightly bound in the active site requires less assumptions than trying to produce complicated assumptions to explain KIEs, Occam's razor tentatively rules out the α-ketol rearrangement in favor of the retro-aldol/aldol mechanism. So, the work was refocused on a new question pertaining to the binding of the Mg^{+2} ion. Previously, it was thought that the metal bound

across the C2-C3 bond coordinated by their bound oxygens, but the same secondary KIEs shed light on the possibility of a second option. This option being bound across the C3-C4 bond via their oxygens. Besides explaining the observed KIEs, it would help to explain the lack of finding any intermediates since the metal would actually stretch over both intermediates instead of just one like in the C2-C3 mode.

The mechanistic work turned to helping to answer the binding mode question. Initial calculations attempted to utilize a single restraint as was done in the DXS mechanism. A single restraint was found wanting though. The results consistently produced highly strained configurations. In order to have better control, a second reaction coordinate was employed. The coordinates controlled related to the C3-C4 bond breaking and proton transfer from DXP O4 hydroxyl to a carboxyl atom of a glutamate residue 231. Since the first coordinate was a bond breakage, the δ_1 was just stretched from 1.60Å to 3.60Å with 0.1Å increments. Similar to that of DXS, the proton transfer was a combination of the O4-H9 bond breaking and the H9-OE2 bond forming starting at the reactant state and progressing to the positive opposite value (i.e. -2.0Å to 2.0Å) over 0.2Å with the expectation the the final value would be shorter than the final 2.0Å mark.

The 2D-surfaces produced an ever increasing field for the C2-C3 binding mode while the C3-C4 binding mode had contours of interest for the retro-aldol reaction. The reaction consists of a deprotonation and C-C bond breakage producing the putative intermediates. Usually considered to be concerted, the C3-C4 energy surface has a valley at a point corresponding to the proton transfer progressing while the C-C breaking hadn't begun. Proposing the possibility of a step-wise retro-aldol process might be a possibility. The C2-C3 surface lacks any signs suggesting this as a possible conclusion. Both of the final states were found to have comparatively high energies of 30.9 kcal·mol^{-1} and 33.0 kcal·mol^{-1} for the C3-C4 mode and C2-C3 mode, respectively. In addition to the energetics, the structure of the C2-C3 final state has a proton transferred to the phosphate tail of DXP instead of the D231 carboxyl group despite the restraint directing the other way. It is

possible the metal ion plays a role in preferential conformation stability and electron stabilization of the transition state through hydroxyl bond polarization.

Additional structural differences indicate a preference for the C3-C4 binding mode. Upon equilibration, this binding mode spreads across the C2-C3-C4 oxygens thus helping to coordinate the entire molecule. The additional binding would support the claims of tight binding in the active site throughout the reaction. The coordination promote the formation of the C2-C4 bond necessary the skeletal rearrangement. If this coordination is indeed necessary for aldolization, it might explain the lack of turnover when the hydroxyacetone and glycoaldehyde phosphate intermediate were exogenously introduced. The formation of the metal coordination might be very unlikely with the two intermediates compared to the single reactant.

4.3   Future Work

As previously mentioned, the projects above represent parts of a complete step of a reaction. The DXS work focused on the first half of ylide formation while the DXR project focuses on the retro-aldol reaction of the retro-aldol/aldol mechanism. Therefore, the follow up work should look to compute the completion of each step in the reaction. Additionally mapping the following reactions of DXS could provide valuable insight into residue contribution while the reactions are well understood. A project of interest would be the decarboxylation of pyruvate and subsequent transferral to G-3-P. Originally, the decarboxylation was thought to take place prior to G-3-P binding, but recent evidence suggests a pause until G-3-P binding. The energetics of decarboxylation with and without G-3-P and environmental analysis might provide unique insights into DXS.

These two enzymes represent the steps of the NMA pathway. Aspects the downstream enzymes might be ascertained utilizing computational techniques while experimental methods don't have the ability. The IspD enzyme is responsible for the transferral of the CMP group of CTP with a hide degree of specificity. Determining aspects to the specificity might be gleamed via Normal Mode Analysis coupled with Vibrational

Subsystem Analysis. Normal mode computes the frequencies of a system while vibration subsystem analysis determines how the large modes have on a subsystems of interest. Thus, mapping changes brought on via simulations might provide insight in the changes upon binding CTP; which in turn would provide further insights into residues of interest.

A mechanism keenly designed for computational investigation might be that the iron-sulfur cluster dependent IspG and IspH. Of particular interest would be IspH, the enzyme is able to produce by IDP and DMADP in a 4:1 ratio via a radical reaction. The ratio of 4:1 also represents the relative usage in downstream isoprenoids. An investigation into differences in the active site or energetics of the reaction might provide insight into how this eznyme preforms such an operation. There are many enzymes that have undesirable biproducts but no to my knowledge that produce both products of a pathway. The MVA pathway for instance utilizes isomerase to convert between IDP and DMADP.

NMA pathway enzymes are not found in mammalian cells suggest this pathway to be wonderful target to novel anti-biotic research. This is bolster by the fact that fosmidomycin is a known anti-malarial drug and inhibits DXR activity. There are other chemicals going through clinical trials currently with hopes of becoming a cheaper and better treatment for disease. Fosmidomycin has also been shown to inhibit IspD and IspE in addition to DXR though to a reduced extent. So it might even be possible to design a drug that target multiple enzymes, thus producing a stronger anti-microbial compound.

Bibliography

[1] M. Alonso, C. Miranda, N. Martin, and B. Herradon. Chemical applications of neural networks: aromaticity of pyrimidine derivatives. *Physical Chemistry Chemical Physics*, 13(46):20564–20574, 2011. doi: 10.1039/c1cp22001b. URL `<GotoISI>://WOS:000297071400006`.

[2] B. Altincicek, A. K. Kollas, S. Sanderbrand, J. Wiesner, M. Hintz, E. Beck, and H. Jomaa. GcpE is involved in the 2-C-Methyl-D-erythritol 4-Phosphate Pathway of Isoprenoid Biosynthesis in Escherichia coli. *Journal of Bacteriology*, 183(8): 2411–2416, 2001.

[3] A. Argyrou and J. S. Blanchard. Kinetic and Chemical Mechanism of Mycobacterium tuberculosis 1-Deoxy-D-xylulose-5-phosphate Isomeroreductase. *Biochemistry*, 43(14):4375–4384, 2004. doi: 10.1021/bi049974k.

[4] D. Arigoni, S. Sagner, C. Latzel, W. Eisenreich, A. Bacher, and M. H. Zenk. Terpenoid Biosynthesis from 1-Deoxy-D-xylulose in Higher Plants by Intramolecular Skeletal Rearrangement. *Proceedings of the National Academy of Sciences of the United States of America*, 94(20):10600–10605, 9 1997. doi: 10.1073/PNAS.94.20.10600. URL `http://www.ncbi.nlm.nih.gov/pubmed/9380681http://www.ncbi.nlm.nih.gov/pubmed/9380681`.

[5] D. Arigoni, J.-L. Giner, S. Sagner, J. Wungsintaweekul, M. H. Zenk, K. Kis, A. Bacher, and W. Eisenreich. Stereochemical Course of the Reduction Step in the

Formation of 2-C-Methylerythritol from the Terpene Precursor 1-Deoxyxylulose in Higher Plants. *Chemical Communications*, 0(12):1127–1128, 1999. doi: 10.1039/A902216C. URL `http://pubs.rsc.org/en/content/articlepdf/1999/cc/a902216c`.

[6] F. Arigoni, F. Talabot, M. Peitsch, M. D. Edgerton, E. Meldrum, E. Allet, R. Fish, T. Jamotte, M.-L. Curchod, and H. Loferer. A Genome-based Approach for the Identification of Essential Bacterial Genes. *Nature Biotechnology*, 16(9):851–856, 9 1998. doi: 10.1038/nbt0998-851. URL `http://www.nature.com/doifinder/10.1038/nbt0998-851`.

[7] A. M. Bailey, S. Mahapatra, P. J. Brennan, and D. C. Crick. Identification, cloning, purification, and enzymatic characterization of Mycobacterium tuberculosis 1-deoxy-D-xylulose 5-phosphate synthase. *Glycobiology*, 12(12):813–820, 2002. URL `<GotoISI>://WOS:000180359500004`.

[8] A. Balakrishnan, Y. Gao, P. Moorjani, N. S. Nemeria, K. Tittmann, and F. Jordan. Bifunctionality of the Thiamin Diphosphate Cofactor: Assignment of Tautomeric/Ionization States of the 4$'$-Aminopyrimidine Ring when Various Intermediates occupy the Active Sites during the Catalysis of Yeast Pyruvate Decarboxylase. *Journal of the American Chemical Society*, 134(8):3873–3885, 2 2012. ISSN 1520-5126. doi: 10.1021/ja211139c. URL `http://dx.doi.org/10.1021/ja211139c`.

[9] A. Balakrishnan, S. Paramasivam, S. Chakraborty, T. Polenova, and F. Jordan. Solid-state nuclear magnetic resonance studies delineate the role of the protein in activation of both aromatic rings of thiamin. *Journal of the American Chemical Society*, 134(1):665–72, 1 2012. ISSN 1520-5126. doi: 10.1021/ja209856x. URL `http://dx.doi.org/10.1021/ja209856x`.

[10] P. Bao and Z.-H. Yu. New Procedure to Evaluate Aromaticity at the Density Functional Theory, Hartree-Fock, and Post-Self-Consistent Field Levels. *Journal of Computational Chemistry*, 32(2):248–259, 2011. doi: 10.1002/jcc.21614. URL `<GotoISI>://WOS:000285312300007`.

[11] D. C. Bas, D. M. Rogers, and J. H. Jensen. Very fast prediction and rationalization of pKa values for protein-ligand complexes. *Proteins*, 73(3):765–783, 11 2008. ISSN 1097-0134. doi: 10.1002/prot.22102. URL `http://www.ncbi.nlm.nih.gov/pubmed/18498103`.

[12] P. A. Bash, M. J. Field, R. C. Davenport, G. A. Petsko, D. Ringe, and M. Karplus. Computer-Simulation and Analysis of the Reaction Pathway of Triosephosphate Isomerase. *Biochemistry*, 30(24):5826–5832, 1991. doi: 10.1021/bi00238a003. URL `<GotoISI>://WOS:A1991FR44600003`.

[13] A. D. Becke. Density-Functional Thermochemistry .3. The Role of Exact Exchange. *Journal of Chemical Physics*, 98(7):5648–5652, 1993. doi: 10.1063/1.464913. URL `<GotoISI>://WOS:A1993KV99700048`.

[14] T. P. Begley, D. M. Downs, S. E. Ealick, F. W. McLafferty, A. P. G. M. Van Loon, S. Taylor, N. Campobasso, H. J. Chiu, C. Kinsland, J. J. Reddick, and J. Xi. Thiamin Biosynthesis in Prokaryotes. *Archives of Microbiology*, 171(5):293–300, 4 1999. doi: 10.1007/s002030050713.

[15] K. Bloch. Biological Synthesis of Cholesterol. *Science*, 150(3692):19–&, 1965. doi: 10.1126/science.150.3692.19. URL `<GotoISI>://WOS:A19656843600007`.

[16] K. Bloch and D. Rittenberg. ON THE UTILIZATION OF ACETIC ACID FOR CHOLESTEROL FORMATION*. *Journal of Biological Chemistry*, 145: 625–636, 1942. URL `http://www.jbc.org/content/145/2/625.full.pdf?sid=ab0d608a-0e0e-4367-b2d8-e891b9304d04`.

[17] K. Bloch, S. Chaykin, A. H. Phillips, and A. Dewaard. Mevalonic Acid Pyrophosphate and Isopentenylpyrophosphate. *Journal of Biological Chemistry*, 234(10): 2595–2604, 1959. URL `<GotoISI>://WOS:A1959WA41900018`.

[18] F. Bouvier, A. D'Harlingue, C. Suire, R. A. Backhaus, and B. Camara. Dedicated Roles of Plastid Transketolases during the Early Onset of Isoprenoid Biogenesis in Pepper Fruits. *Plant Physiology*, 117(4):1423–1431, 8 1998. doi: 10.1104/pp.117.4. 1423.

[19] L. A. Brammer and C. F. Meyers. Revealing Substrate Promiscuity of 1-Deoxy-D-xylulose 5-Phosphate Synthase. *Organic Letters*, 11(20):4748–4751, 2009. doi: 10.1021/ol901961q. URL `<GotoISI>://WOS:000270461300067`.

[20] L. A. Brammer, J. M. Smith, H. Wade, and C. F. Meyers. 1-Deoxy-D-xylulose 5-Phosphate Synthase Catalyzes a Novel Random Sequential Mechanism. *Journal of Biological Chemistry*, 286(42):36522–36531, 2011. doi: 10.1074/jbc.M111.259747. URL `<GotoISI>://WOS:000296538300037`.

[21] L. A. Brammer Basta, H. Patel, L. Kakalis, F. Jordan, and C. L. Freel Meyers. Defining critical residues for substrate binding to 1-deoxy-D-xylulose 5-phosphate synthase–active site substitutions stabilize the predecarboxylation intermediate C2α-lactylthiamin diphosphate. *The FEBS Journal*, 281(12):2820–37, 6 2014. ISSN 1742-4658. doi: 10.1111/febs.12823. URL `http://www.ncbi.nlm.nih.gov/pubmed/24767541`.

[22] G. S. Brandt, N. Nemeria, S. Chakraborty, M. J. McLeish, A. Yep, G. L. Kenyon, G. A. Petsko, F. Jordan, and D. Ringe. Probing the active center of benzaldehyde lyase with substitutions and the pseudosubstrate analogue benzoylphosphonic acid methyl ester. *Biochemistry*, 47(29):7734–7743, 7 2008. ISSN 1520-4995. doi: 10.1021/bi8004413. URL `http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2729719&tool=pmcentrez&rendertype=abstract`.

[23] W. Brandt, M. A. Dessoy, M. Fulhorst, W. Gao, M. H. Zenk, and L. A. Wessjohann. A Proposed Mechanism for the Reductive Ring Opening of the Cyclodiphosphate MEcPP, a Crucial Transformation in the New DXP/MEP Pathway to Isoprenoids Based on Modeling Studies and Feeding Experiments. *ChemBioChem*, 5(3):311–323, 3 2004. ISSN 14394227. doi: 10.1002/cbic.200300743. URL http://www.ncbi.nlm.nih.gov/pubmed/14997523http://doi.wiley.com/10.1002/cbic.200300743.

[24] J. G. Breman, A. Egan, and G. T. Keusch. *The Intolerable Burden of Malaria: A New Look at the Numbers*. American Society of Tropical Medicine and Hygiene, Northbook, IL, 2001. URL https://www.ncbi.nlm.nih.gov/books/NBK2617/?report=reader.

[25] R. Breslow. The Mechanism of Thiamine Action .2. Rapid Deuterium Exchange in Thiazolium Salts. *Journal of the American Chemical Society*, 79(7):1762–1763, 1957. URL <GotoISI>://WOS:A1957WB80600064.

[26] R. Breslow. On the Mechanism of Thiamine Action .4. Evidence from Studies on Model Systems. *Journal of the American Chemical Society*, 80(14):3719–3726, 1958. URL <GotoISI>://WOS:A1958WB38900063.

[27] B. R. Brooks, C. L. Brooks III, A. D. Mackerell Jr., L. Nilsson, R. J. Petrella, B. Roux, Y. Won, G. Archontis, C. Bartels, S. Boresch, A. Caflisch, L. Caves, Q. Cui, A. R. Dinner, M. Feig, S. Fischer, J. Gao, M. Hodoscek, W. Im, K. Kuczera, T. Lazaridis, J. Ma, V. Ovchinnikov, E. Paci, R. W. Pastor, C. B. Post, J. Z. Pu, M. Schaefer, B. Tidor, R. M. Venable, H. L. Woodcock, X. Wu, W. Yang, D. M. York, and M. Karplus. CHARMM: The Biomolecular Simulation Program. *Journal of Computational Chemistry*, 30(10):1545–1614, 2009. doi: 10.1002/jcc.21287. URL <GotoISI>://WOS:000267269600001.

[28] B. M. Calisto, J. Perez-Gil, M. Bergua, J. Querol-Audi, I. Fita, and S. Imperial. Biosynthesis of isoprenoids in plants: Structure of the 2C-methyl-d-erithrytol 2,4-cyclodiphosphate synthase from ¡i¿Arabidopsis thaliana¡/i¿ . Comparison with the bacterial enzymes. *Protein Science*, 16(9):2082–2088, 9 2007. ISSN 09618368. doi: 10.1110/ps.072972807. URL `http://www.ncbi.nlm.nih.gov/pubmed/17660251http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC2206962http://doi.wiley.com/10.1110/ps.072972807`.

[29] N. Campos, M. Rodríguez-Concepción, M. Seemann, M. Rohmer, and A. Boronat. Identification of gcpE as a Novel Gene of the 2-C-Methyl-D-erythritol 4-Phosphate Pathway for Isoprenoid Biosynthesis in Escherichia coli. *FEBS letters*, 488(3):170–173, 1 2001. URL `http://www.ncbi.nlm.nih.gov/pubmed/11163766`.

[30] D. E. Cane, T. Rossi, and J. P. Pachlatko. The Biosynthesis of Pentalenolactone. *Tetrahedron Letters*, 20(38):3639–3642, 1 1979. doi: 10.1016/S0040-4039(01)95484-X. URL `http://www.sciencedirect.com/science/article/pii/S004040390195484X?via%3Dihub`.

[31] D. E. Cane, T. Rossi, A. M. Tillman, and J. P. Pachlatko. Stereochemical Studies of Isoprenoid Biosynthesis. Biosynthesis of Pentalenolactone from [U-13C6]Glucose and [6-2H2]Glucose. *Journal of the American Chemical Society*, 103(7):1838–1843, 4 1981. doi: 10.1021/ja00397a045. URL `http://pubs.acs.org/doi/abs/10.1021/ja00397a045`.

[32] J. Chappell. Biochemistry and Molecular Biology of the Isoprenoid Biosynthetic Pathway in Plants. *Annual Review of Plant Physiology and Plant Molecular Biology*, 46(1):521–547, 6 1995. doi: 10.1146/annurev.pp.46.060195.002513. URL `http://www.annualreviews.org/doi/10.1146/annurev.pp.46.060195.002513`.

[33] S. Chaykin, J. Law, A. H. Phillips, T. T. Tchen, and K. Bloch. Phosphorylated Intermediates in the Synthesis of Squalene. *Proceedings of the National Academy*

of Sciences of the United States of America, 44(10):998–1004, 1958. doi: 10.1073/ pnas.44.10.998. URL `<GotoISI>://WOS:A1958WJ52800004`.

[34] S. Cheek, H. Zhang, and N. V. Grishin. Sequence and structure classification of kinases. *Journal of molecular biology*, 320(4):855–81, 7 2002. ISSN 0022-2836. URL `http://www.ncbi.nlm.nih.gov/pubmed/12095261`.

[35] Z. F. Chen, C. S. Wannere, C. Corminboeuf, R. Puchta, and P. V. Schleyer. Nucleus-independent Chemical Shifts (NICS) as an Aromaticity Criterion. *Chemical Reviews*, 105(10):3842–3888, 2005. doi: 10.1021/cr030088+. URL `<GotoISI>: //WOS:000232755200014`.

[36] J. W. Chu, B. L. Trout, and B. R. Brooks. A Super-linear Minimization Scheme for the Nudged Elastic Band Method. *Journal of Chemical Physics*, 119(24):12708– 12717, 12 2003. doi: 10.1063/1.1627754.

[37] S. K. Chunduru, G. T. Mrachko, and K. C. Calvo. Mechanism of Ketol Acid Reductoisomerase. Steady-state Analysis and Metal Ion Requirement. *Biochemistry*, 28(2):486–493, 1 1989. ISSN 0006-2960. doi: 10.1021/bi00428a012. URL `http://pubs.acs.org/doi/abs/10.1021/bi00428a012`.

[38] J. D. Connolly and R. A. Hill. *Dictionary of Terpenoids*. CRC Press, 1991. ISBN 9780412257704. URL `https://books.google.com/books?id=FjnKUcD7-B0C`.

[39] C. Corminboeuf, T. Heine, G. Seifert, P. V. Schleyer, and J. Weber. Induced magnetic fields in aromatic n-annulenes - interpretation of NICS tensor components. *Physical Chemistry Chemical Physics*, 6(2):273–276, 2004. doi: 10.1039/b313383b.

[40] J. W. Cornforth, R. H. Cornforth, A. Pelter, M. G. Horning, and G. Popjak. Studies on the Biosynthesis of Cholesterol .7. Rearrangement of Methyl Groups During Enzymic Cyclisation of Squalene. *Tetrahedron*, 5(4):311–339, 1959. doi: 10.1016/ 0040-4020(59)80024-7. URL `<GotoISI>://WOS:A1959WP46200006`.

[41] S. J. Costelloe, J. M. Ward, and P. A. Dalby. Evolutionary analysis of the TPP-dependent enzyme family. *Journal of Molecular Evolution*, 66(1):36–49, 2008. doi: 10.1007/s00239-007-9056-2. URL `<GotoISI>://WOS:000252629000004`.

[42] Q. Cui, M. Elstner, E. Kaxiras, T. Frauenheim, and M. Karplus. A QM/MM implementation of the self-consistent charge density functional tight binding (SCC-DFTB) method. *Journal of Physical Chemistry B*, 105(2):569–585, 2001. doi: 10.1021/jp0029109. URL `<GotoISI>://WOS:000166490900029`.

[43] M. K. Cyranski, T. M. Krygowski, A. R. Katritzky, and P. V. Schleyer. To what extent can aromaticity be defined uniquely? *Journal of Organic Chemistry*, 67(4):1333–1338, 2002. doi: 10.1021/jo016255s. URL `<GotoISI>://WOS:000173901700041`.

[44] M. K. Cyranski, P. V. Schleyer, T. M. Krygowski, H. J. Jiao, and G. Hohlneicher. Facts and Artifacts about Aromatic Stability Estimation. *Tetrahedron*, 59(10): 1657–1665, 2003. doi: 10.1016/s0040-4020(03)00137-6.

[45] T. Dairi, T. Kuzuyama, M. Nishiyama, and I. Fujii. Convergent strategies in biosynthesis. *Natural Product Reports*, 28(6):1054–1086, 2011. doi: 10.1039/c0np00047g. URL `<GotoISI>://WOS:000290993000002`.

[46] M. De Rosa, A. Gambacorta, and A. Gliozzi. Structure, Biosynthesis, and Physicochemical Properties of Archaebacterial Lipids. *Microbiological Reviews*, 50(1): 70–80, 1986. URL `https://www.scopus.com/record/display.uri?eid=2-s2.0-0022578265&origin=inward&txGid=b5e0542cb05a0ce4daffc1d460258205`.

[47] J. S. Dickschat. Isoprenoids in three-dimensional space: the stereochemistry of terpene biosynthesis. *Natural Product Reports*, 28(12):1917–1936, 2011. doi: 10.1039/c1np00063b. URL `<GotoISI>://WOS:000297029900003`.

[48] Duke, Dayan, Romagni, and Rimando. Natural Products as Sources of Herbicides: Current Status and Future Trends. *Weed Research*, 40(1):99–111, 2 2000. ISSN

0043-1737. doi: 10.1046/j.1365-3180.2000.00161.x. URL `http://doi.wiley.com/10.1046/j.1365-3180.2000.00161.x`.

[49] R. Dumas, D. Job, J. Y. Ortholand, G. Emeric, A. Greiner, and R. Douce. Isolation and Kinetic Properties of Acetohydroxy Acid Isomeroreductase from Spinach (Spinacia oleracea) Chloroplasts overexpressed in Escherichia coli. *The Biochemical Journal*, 288 (Pt 3(3):865–74, 12 1992. ISSN 0264-6021. doi: 10.1042/BJ2880865. URL `http://www.ncbi.nlm.nih.gov/pubmed/1472001http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC1131967`.

[50] R. Dumas, V. Biou, F. Halgand, R. Douce, and R. G. Duggleby. Enzymology, Structure, and Dynamics of Acetohydroxy Acid Isomeroreductase. *Accounts of Chemical Research*, 34(5):399–408, 2001. doi: 10.1021/AR000082W. URL `http://pubs.acs.org/doi/abs/10.1021/ar000082w`.

[51] W. Eisenreich, M. Schwarz, A. Cartayrade, D. Arigoni, M. H. Zenk, and A. Bacher. The deoxyxylulose phosphate pathway of terpenoid biosynthesis in plants and microorganisms. *Chemistry & Biology*, 5(9):R221–R233, 1998. doi: 10.1016/s1074-5521(98)90002-3. URL `<GotoISI>://WOS:000076023100002`.

[52] W. Eisenreich, A. Bacher, D. Arigoni, and F. Rohdich. Biosynthesis of Isoprenoids via the Non-Mevalonate Pathway. *Cellular and Molecular Life Sciences*, 61(12):1401–1426, 2004. doi: 10.1007/s00018-004-3381-z.

[53] J. M. Estévez, A. Cantero, A. Reindl, S. Reichler, P. León, J. M. Estevez, A. Cantero, A. Reindl, S. Reichler, and P. Leon. 1-Deoxy-D-xylulose-5-Phosphate Synthase, a Limiting Enzyme for Plastidic Isoprenoid Biosynthesis in Plants. *Journal of Biological Chemistry*, 276(25):22901–22909, 6 2001. ISSN 0021-9258. doi: 10.1074/jbc.M100854200. URL `http://www.ncbi.nlm.nih.gov/pubmed/11264287`.

[54] L. M. Eubanks and C. D. Poulter. Rhodobacter capsulatus 1-deoxy-D-xylulose 5-phosphate synthase: Steady-state kinetics and substrate binding. *Biochem-*

istry, 42(4):1140–1149, 2003. doi: 10.1021/bi0205303. URL `<GotoISI>://WOS:`
`000180695200036`.

[55] S. Falkow and D. Kennedy. Antibiotics, Animals, and People–Again! *Science (New York, N.Y.)*, 291(5503):397, 1 2001. doi: 10.1126/SCIENCE.1058907. URL `http://www.ncbi.nlm.nih.gov/pubmed/11228121`.

[56] R. Firn. *Nature's chemicals the natural products that shaped our world.* Oxford University Press, 2010. ISBN 9780191721700. URL `http://apps.webofknowledge.com/full_record.do?product=WOS&search_` `mode=GeneralSearch&qid=28&SID=2E4kGskQgOEpsoqX4Rp&page=1&doc=1`.

[57] G. Flesch and M. Rohmer. Prokaryotic Hopanoids: The Biosynthesis of the Bacteriohopane Skeleton. Formation of Isoprenic Units from Two Distinct Acetate Pools and a Novel Type of Carbon/Carbon Linkage between a Triterpene and d-Ribose. *European Journal of Biochemistry*, 175(2):405–411, 8 1988. doi: 10.1111/j.1432-1033.1988.tb14210.x. URL `http://doi.wiley.com/10.1111/j.` `1432-1033.1988.tb14210.x`.

[58] D. T. Fox and C. D. Poulter. Mechanistic studies with 2-C-methyl-D-erythritol 4-phosphate synthase from Escherichia coli. *Biochemistry*, 44(23):8360–8368, 2005. doi: 10.1021/bi047312p. URL `http://pubs.acs.org/doi/abs/10.1021/` `bi047312p`.

[59] A. Frank and M. Groll. The Methylerythritol Phosphate Pathway to Isoprenoids. *Chemical Reviews*, 117(8):5675–5703, 4 2017. ISSN 0009-2665. doi: 10.1021/ acs.chemrev.6b00537. URL `http://pubs.acs.org/doi/10.1021/acs.chemrev.` `6b00537`.

[60] R. A. W. Frank, F. J. Leeper, and B. F. Luisi. Structure, mechanism and catalytic duality of thiamine-dependent enzymes. *Cellular and Molecular Life Sciences*, 64

(7-8):892–905, 2007. doi: 10.1007/s00018-007-6423-5. URL `<GotoISI>://WOS:` `000245669400009`.

[61] Z. Fu, M. Wang, D. Potter, H. M. Miziorko, and J.-J. P. Kim. The Structure of a Binary Complex between a Mammalian Mevalonate Kinase and ATP. *Journal of Biological Chemistry*, 277(20):18134–18142, 5 2002. ISSN 0021-9258. doi: 10.1074/jbc.M200912200. URL `http://www.ncbi.nlm.nih.gov/pubmed/11877411http://www.jbc.org/lookup/doi/10.1074/jbc.M200912200`.

[62] M. Gabrielsen, J. Kaiser, F. Rohdich, W. Eisenreich, R. Laupitz, A. Bacher, C. S. Bond, and W. N. Hunter. The crystal structure of a plant 2C-methyl-D-erythritol 4-phosphate cytidylyltransferase exhibits a distinct quaternary structure compared to bacterial homologues and a possible role in feedback regulation for cytidine monophosphate. *The FEBS Journal*, 273(5): 1065–1073, 3 2006. ISSN 1742-464X. doi: 10.1111/j.1742-4658.2006.05133.x. URL `http://www.ncbi.nlm.nih.gov/pubmed/16478479http://doi.wiley.com/10.1111/j.1742-4658.2006.05133.x`.

[63] T. Gräwert, M. Groll, F. Rohdich, A. Bacher, W. Eisenreich, T. Graewert, M. Groll, F. Rohdich, A. Bacher, and W. Eisenreich. Biochemistry of the Non-mevalonate Isoprenoid Pathway. *Cellular and Molecular Life Sciences*, 68(23):3797–3814, 12 2011. ISSN 1420-682X. doi: 10.1007/s00018-011-0753-z. URL `http://link.springer.com/10.1007/s00018-011-0753-z`.

[64] S. Grolle, S. Bringer-Meyer, and H. Sahm. Isolation of the dxr Gene of Zymomonas mobilis and Characterization of the 1-deoxy-D-xylulose 5-phosphate reductoisomerase. *FEMS Microbiology Letters*, 191(1):131–137, 10 2000. doi: 10.1016/s0378-1097(00)00382-7. URL `https://academic.oup.com/femsle/article-lookup/doi/10.1111/j.1574-6968.2000.tb09329.x`.

[65] F. S. Guo, D. Q. Zhang, A. Kahyaoglu, R. S. Farid, and F. Jordan. Is a hydrophobic amino acid required to maintain the reactive V conformation of thiamin at the active center of thiamin diphosphate-requiring enzymes? Experimental and computational studies of isoleucine 415 of yeast pyruvate decarboxylase. *Biochemistry*, 37(38): 13379–13391, 1998. URL `<GotoISI>://WOS:000076088200038`.

[66] S. Handa, D. Ramamoorthy, T. J. Spradling, W. C. Guida, J. H. Adams, K. G. Bendinskas, and D. J. Merkler. Production of Recombinant 1-Deoxy-D-Xylulose-5-Phosphate Synthase from Plasmodium Vivax in Escherichia Coli. *FEBS Open Bio*, 3:124–129, 2013.

[67] J. C. Hargis, S. L. Vankayala, J. K. White, and H. L. Woodcock. Identification and Characterization of Noncovalent Interactions That Drive Binding and Specificity in DD-Peptidases and β-Lactamases. *Journal of Chemical Theory and Computation*, 10(2):855–864, 2 2014. ISSN 1549-9618. doi: 10.1021/ct400968v. URL `http://pubs.acs.org.ezproxy.lib.usf.edu/doi/abs/10.1021/ct400968v`.

[68] Harihara.Pc, J. A. Pople, P. C. Harihara, and J. A. Pople. Influence of Polarization Functions on Molecular-Orbital Hydrogenation Energies. *Theoretica Chimica Acta*, 28(3):213–222, 1973. doi: 10.1007/bf00533485.

[69] M. Harker and P. M. Bramley. Expression of Prokaryotic 1-Deoxy-D-xylulose-5-Phosphatases in Escherichia coli increases Carotenoid and Ubiquinone Biosynthesis. *FEBS Letters*, 448(1):115–119, 4 1999. doi: 10.1016/S0014-5793(99)00360-9.

[70] A. Hartley, S. E. Glynn, V. Barynin, P. J. Baker, S. E. Sedelnikova, C. Verhees, D. de Geus, J. van der Oost, D. J. Timson, R. J. Reece, and D. W. Rice. Substrate Specificity and Mechanism from the Structure of Pyrococcus furiosus Galactokinase. *Journal of Molecular Biology*, 337(2):387–398, 3 2004. ISSN 00222836. doi: 10.1016/j.jmb.2004.01.043. URL `http://www.ncbi.nlm.nih.gov/pubmed/15003454http://linkinghub.elsevier.com/retrieve/pii/S0022283604001147`.

[71] S. Hecht, W. Eisenreich, P. Adam, S. Amslinger, K. Kis, A. Bacher, D. Arigoni, and F. Rohdich. Studies on the Nonmevalonate Pathway to Terpenes: The Role of the GcpE (IspG) Protein. *Proceedings of the National Academy of Sciences of the United States of America*, 98(26):14837–14842, 2001. doi: 10.1073/pnas.201399298.

[72] S. Hecht, F. Kis, W. Eisenreich, S. Amslinger, J. Wungsintaweekul, S. Herz, F. Rohdich, and A. Bacher. Enzyme-assisted Preparation of Isotope-labeled 1-Deoxy-D-xylulose 5-Phosphate. *Journal of Organic Chemistry*, 66(11):3948–3952, 2001. doi: 10.1021/jo0100300.

[73] S. Hecht, J. Wungsintaweekul, F. Rohdich, K. Kis, T. Radykewicz, C. A. Schuhr, W. Eisenreich, G. Richter, and A. Bacher. Biosynthesis of Terpenoids: Efficient Multistep Biotransformation Procedures Affording Isotope-Labeled 2C-Methyl-d-erythritol 4-Phosphate Using Recombinant 2C-Methyl-d-erythritol 4-Phosphate Synthase. *The Journal of Organic Chemistry*, 66(23):7770–7775, 2001. doi: 10.1021/JO015890V. URL `http://pubs.acs.org/doi/abs/10.1021/jo015890v`.

[74] L. M. Henriksson, T. Unge, J. Carlsson, J. Åqvist, S. L. Mowbray, and T. A. Jones. Structures of Mycobacterium tuberculosis 1-Deoxy-D-xylulose-5-phosphate Reductoisomerase Provide New Insights into Catalysis. *The Journal of Biological Chemistry*, 282(27):1990519916, 2007. doi: 10.1074/jbc.M701935200. URL `http://www.jbc.org/content/282/27/19905.full.pdf`.

[75] S. Herz, J. Wungsintaweekul, C. A. Schuhr, S. Hecht, H. Luttgen, S. Sagner, M. Fellermeier, W. Eisenreich, M. H. Zenk, A. Bacher, and F. Rohdich. Biosynthesis of terpenoids: YgbB protein converts 4-diphosphocytidyl-2C-methyl-D-erythritol 2-phosphate to 2C-methyl-D-erythritol 2,4-cyclodiphosphate. *Proceedings of the National Academy of Sciences of the United States of America*, 97(6):2486–2490, 3 2000. ISSN 0027-8424. doi: 10.1073/pnas.040554697. URL `http://www.ncbi.nlm.nih.gov/pubmed/10694574http://www.pubmedcentral.`

nih.gov/articlerender.fcgi?artid=PMC15955http://www.pnas.org/cgi/
doi/10.1073/pnas.040554697.

[76] R. E. Hill, K. Himmeldirk, I. A. Kennedy, R. M. Pauloski, B. G. Sayer, E. Wolf, and I. D. Spenser. The Biogenetic Anatomy of Vitamin B-6: A C-13 NMR Investigation of the Biosynthesis of Pyridoxol in $\textit{Escherichia coli}$. *Journal of Biological Chemistry*, 271(48):30426–30435, 11 1996.

[77] S. N. Ho, H. D. Hunt, R. M. Horton, J. K. Pullen, and L. R. Pease. Site-directed Mutagenesis by Overlap Extension using the Polymerase Chain Reaction. *Gene*, 77(1):51–59, 4 1989. ISSN 03781119. doi: 10.1016/0378-1119(89)90358-2. URL http://www.sciencedirect.com/science/article/pii/0378111989903582.

[78] J.-F. F. Hoeffler, D. Tritsch, C. Grosdemange-Billiard, and M. Rohmer. Isoprenoid Biosynthesis via the Methylerythritol Phosphate Pathway: Mechanistic Investigations of the 1-Deoxy-D-xylulose 5-Phosphate Reductoisomerase. *European Journal of Biochemistry*, 269(18):4446–4457, 9 2002. ISSN 0014-2956. doi: 10.1046/j. 1432-1033.2002.03150.x. URL http://www.ncbi.nlm.nih.gov/pubmed/12230556.

[79] V. Humnabadkar, R. K. Jha, N. Ghatnekar, and S. M. De Sousa. A High-Throughput Screening Assay for Simultaneous Selection of Inhibitors of Mycobacterium tuberculosis 1-Deoxy-D-Xylulose-5-Phosphate Synthase (DXS) or 1-Deoxy-D-Xylulose 5-Phosphate Reductoisomerase (DXR). *Journal of Biomolecular Screening*, 16(3):303–312, 3 2011. ISSN 1552-454X. doi: 10.1177/1087057110394845. URL http://apps.webofknowledge.com/full_record.do?product=UA&search_mode=GeneralSearch&qid=4&SID=3CGKQ16I9hqTlsvBbGH&page=1&doc=1.

[80] A. E. Johnson and M. E. Tanner. Epimerization via CarbonCarbon Bond Cleavage. l-Ribulose-5-phosphate 4-Epimerase as a Masked Class II Aldolase. *Biochemistry*,

37(16):5746–5754, 1998. doi: 10.1021/BI972984J. URL `http://pubs.acs.org/`
`doi/abs/10.1021/bi972984j`.

[81] H. Jomaa, J. Wiesner, S. Sanderbrand, B. Altincicek, C. Weidemeyer, M. Hintz, I. Turbachova, M. Eberl, J. Zeidler, H. K. Lichtenthaler, D. Soldati, and E. Beck. Inhibitors of the Nonmevalonate Pathway of Isoprenoid Biosynthesis as Antimalarial Drugs. *Science*, 285(5433):1573–1576, 1999. doi: 10.1126/science.285.5433.1573.

[82] F. Jordan. Current mechanistic understanding of thiamin diphosphatedependent enzymatic reactions. *Natural Product Reports*, 20(2):184–201, 2003. doi: 10.1039/ b111348h. URL `<GotoISI>://WOS:000182268900002`.

[83] F. Jordan and N. S. Nemeria. Progress in the Experimental Observation of Thiamin Diphosphate-bound Intermediates on Enzymes and Mechanistic Information derived from these Observations. *Bioorganic Chemistry*, 57(0):251–62, 12 2014. ISSN 0045-2068. doi: http://dx.doi.org/10.1016/j.bioorg.2014.08.002. URL `http://www.sciencedirect.com/science/article/pii/S0045206814000674`.

[84] F. Jordan, H. J. Li, and A. Brown. Remarkable stabilization of zwitterionic intermediates may account for a billion-fold rate acceleration by thiamin diphosphate-dependent decarboxylases. *Biochemistry*, 38(20):6369–6373, 1999. URL `<GotoISI>://WOS:000080593700001`.

[85] J. Kalinowska-Tłuścik, L. Miallau, M. Gabrielsen, G. A. Leonard, S. M. McSweeney, and W. N. Hunter. A Triclinic Crystal Form of $\textit{Escherichia coli}$ 4-Diphosphocytidyl-2C-methyl-D-erythritol Kinase and Reassessment of the Quaternary Structure. *Acta Crystallographica Section F - Structural Biology and Crystallization Communications*, 66(3):237–241, 3 2010. ISSN 1744-3091. doi: 10.1107/S1744309109054591. URL `http://www.ncbi.nlm.nih.gov/pubmed/` `20208151http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=` `PMC2833027http://scripts.iucr.org/cgi-bin/paper?S1744309109054591`.

[86] A. Kaplun, E. Binshtein, M. Vyazmensky, A. Steinmetz, Z. Barak, D. M. Chipman, K. Tittmann, and B. Shaanan. Glyoxylate carboligase lacks the canonical active site glutamate of thiamine-dependent enzymes. *Nature: Chemical Biology*, 4(2): 113–8, 2 2008. ISSN 1552-4469. doi: 10.1038/nchembio.62. URL `http://www.ncbi.nlm.nih.gov/pubmed/18176558`.

[87] L. E. Kemp, C. S. Bond, and W. N. Hunter. Structure of 2C-methyl-D-erythritol 2,4-cyclodiphosphate synthase: An essential enzyme for isoprenoid biosynthesis and target for antimicrobial drug development. *Proceedings of the National Academy of Sciences of the United States of America*, 99(10):6591–6596, 5 2002. ISSN 0027-8424. doi: 10.1073/pnas.102679799. URL `http://www.ncbi.nlm.nih.gov/pubmed/11997478http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC124447http://www.pnas.org/cgi/doi/10.1073/pnas.102679799`.

[88] L. E. Kemp, C. S. Bond, and W. N. Hunter. Structure of a tetragonal crystal form of Escherichia coli 2-C-methyl-D-erythritol 4-phosphate cytidylyltransferase. *Acta Crystallographica Section D - Biological Crystallography*, 59(Pt 3):607–610, 3 2003. ISSN 0907-4449. doi: 10.1107/s090744490202365x. URL `http://www.ncbi.nlm.nih.gov/pubmed/12595740`.

[89] L. E. Kemp, M. S. Alphey, C. S. Bond, M. A. J. Ferguson, S. Hecht, A. Bacher, W. Eisenreich, F. Rohdich, and W. N. Hunter. The identification of isoprenoids that bind in the intersubunit cavity of Escherichia coli 2 C -methyl- D -erythritol-2,4-cyclodiphosphate synthase by complementary biophysical methods. *Acta Crystallographica Section D - Biological Crystallography*, 61 (Pt 1):45–52, 1 2005. doi: 10.1107/S0907444904025971. URL `http://www.ncbi.nlm.nih.gov/pubmed/15608374`.

[90] D. Kern, G. Kern, H. Neef, K. Tittmann, M. KillenbergJabs, C. Wikner, G. Schneider, and G. Hubner. How thiamine diphosphate is activated in enzymes. *Science*,

275(5296):67–70, 1997. doi: 10.1126/science.275.5296.67. URL `<GotoISI>://WOS:`
`A1997WA90300049`.

[91] H. Kleinig. The Role of Plastids in Isoprenoid Biosynthesis. *Annual Review of Plant Physiology and Plant Molecular Biology*, 40(1):39–59, 6 1989. ISSN 1040-2519. doi: 10.1146/annurev.pp.40.060189.000351. URL `http://www.annualreviews. org/doi/10.1146/annurev.pp.40.060189.000351`.

[92] A.-K. Kollas, E. C. Duin, M. Eberl, B. Altincicek, M. Hintz, A. Reichenberg, D. Henschker, A. Henne, I. Steinbrecher, D. N. Ostrovsky, R. Hedderich, E. Beck, H. Jomaa, and J. Wiesner. Functional Characterization of GcpE, an Essential Enzyme of the Non-mevalonate Pathway of Isoprenoid Biosynthesis. *FEBS letters*, 532(3):432–6, 12 2002. ISSN 0014-5793. URL `http://www.ncbi.nlm.nih.gov/ pubmed/12482607`.

[93] P. A. Kollman, B. Kuhn, O. Donini, M. Perakyla, R. Stanton, and D. Bakowies. Elucidating the Nature of Enzyme Catalysis utilizing a New Twist on an Old Methodology: Quantum MechanicalFree Energy Calculations on Chemical Reactions in Enzymes and in Aqueous Solution. *Accounts of Chemical Research*, 34(1): 72–79, 2001. doi: 10.1021/AR000032R. URL `http://pubs.acs.org/doi/abs/10. 1021/ar000032r`.

[94] A. T. Koppisch, D. T. Fox, B. S. J. Blagg, and C. D. Poulter. E-coli MEP Synthase: Steady-state Kinetic Analysis and Substrate Binding. *Biochemistry*, 41(1):236–243, 2002. doi: 10.1021/bi0118207.

[95] D. E. Koshland and K. E. Neet. The Catalytic and Regulatory Properties of Enzymes. *Annual Review of Biochemistry*, 37:359–410, 1968. URL `https://www.scopus.com/record/display.uri?eid=2-s2.0-0014235891& origin=inward&txGid=64ceb4f245ac59560d2063ba2da7cc6f`.

[96] A. Koul, E. Arnoult, N. Lounis, J. Guillemont, and K. Andries. The Challenge of New Drug Discovery for Tuberculosis. *Nature*, 469:483–490, 2011. doi: 10.1038/ nature09657. URL `https://www.nature.com/articles/nature09657.pdf`.

[97] S. S. Krishna, T. Zhou, M. Daugherty, A. Osterman, and H. Zhang. Structural Basis for the Catalysis and Substrate Specificity of Homoserine Kinase. *Biochemistry*, 40 (36):10810–10818, 2001. doi: 10.1021/BI010851Z. URL `http://pubs.acs.org/ doi/10.1021/bi010851z`.

[98] T. M. Krygowski, M. K. Cyranski, Z. Czarnocki, G. Hafelinger, and A. R. Katritzky. Aromaticity: A Theoretical Concept of Immense Practical Importance. *Tetrahedron*, 56(13):1783–1796, 2000. doi: 10.1016/s0040-4020(99)00979-5.

[99] T. Kuzuyama and H. Seto. Diversity of the biosynthesis of the isoprene units. *Natural Product Reports*, 20(2):171–183, 2003. doi: 10.1039/b109860h. URL `<GotoISI>://WOS:000182268900001`.

[100] T. Kuzuyama, T. Shimizu, S. Takahashi, and H. Seto. Fosmidomycin, a Specific Inhibitor of 1-Deoxy-D-xylulose 5-Phosphate Reductoisomerase in the Nonmevalonate Pathway for Terpenoid Biosynthesis. *Tetrahedron Letters*, 39(43):7913–7916, 1998.

[101] T. Kuzuyama, M. Takagi, K. Kaneda, T. Dairi, and H. Seto. Formation of 4-(cytidine 5-diphospho)-2-C-methyl-d-erythritol from 2-C-Methyl-d-erythritol 4-Phosphate by 2-C-Methyl-d-erythritol 4-Chosphate Cytidylyltransferase, a New Enzyme in the Nonmevalonate Pathway. *Tetrahedron Letters*, 41(5):703–706, 1 2000. doi: 10.1016/S0040-4039(99)02143-7. URL `http://www.sciencedirect. com/science/article/pii/S0040403999021437?via%3Dihub`.

[102] T. Kuzuyama, S. Takahashi, M. Takagi, and H. Seto. Characterization of 1-Deoxy-D-xylulose 5-Phosphate Reductoisomerase, an Enzyme involved in Isopen-

tenyl Diphosphate Biosynthesis, and Identification of its Catalytic Amino Acid Residues. *Journal of Biological Chemistry*, 275(26):19928–19932, 2000.

[103] B. M. Lange, M. R. Wildung, D. McCaskill, and R. Croteau. A Family of Transketo-lases that Directs Isoprenoid Biosynthesis via a Mevalonate-Independent Pathway. *Proceedings of the National Academy of Sciences of the United States of America*, 95 (5):2100–2104, 1998. URL `http://www.pnas.org/content/95/5/2100.full.pdf`.

[104] S. Lauw, V. Illarionova, A. Bacher, F. Rohdich, and W. Eisenreich. Biosynthesis of Isoprenoids: Studies on the Mechanism of 2C-Methyl-D-erythritol-4-phosphate Synthase. *The FEBS Journal*, 275(16):4060–4073, 2008. doi: 10.1111/j.1742-4658. 2008.06547.x.

[105] C. T. Lee, W. T. Yang, and R. G. Parr. Development of the Colle-Salvetti Correlation-Energy Formula into a Functional of the Electron-Density. *Physical Review B*, 37(2):785–789, 1988. doi: 10.1103/PhysRevB.37.785. URL `<GotoISI>: //WOS:A1988L976200011`.

[106] M. Lee, T. Gräwert, F. Quitterer, F. Rohdich, J. Eppinger, W. Eisen-reich, A. Bacher, and M. Groll. Biosynthesis of Isoprenoids: Crystal Structure of the [4Fe4S] Cluster Protein IspG. *Journal of Molecular Bi-ology*, 404(4):600–610, 12 2010. ISSN 00222836. doi: 10.1016/j.jmb. 2010.09.050. URL `http://www.ncbi.nlm.nih.gov/pubmed/20932974http:// linkinghub.elsevier.com/retrieve/pii/S0022283610010491`.

[107] Y. S. Lee, S. E. Worthington, M. Krauss, and B. R. Brooks. Reaction mechanism of chorismate mutase studied by the combined potentials of quantum mechanics and molecular mechanics. *Journal of Physical Chemistry B*, 106(46):12059–12065, 2002. doi: 10.1021/jp0268718. URL `<GotoISI>://WOS:000179336200023`.

[108] H. Li, A. D. Robertson, and J. H. Jensen. Very Fast Empirical Prediction and Interpretation of Protein pKa Values. *Proteins*, 61:704–721, 2005.

[109] H. Li, J. Tian, W. Sun, W. Qin, and W.-Y. Gao. Mechanistic Insights into 1-Deoxy-D-xylulose 5-Phosphate Reductoisomerase, a Key Enzyme of the MEP Terpenoid Biosynthetic Pathway. *The FEBS Journal*, 280(22):5896–5905, 11 2013. doi: 10.1111/febs.12516. URL `http://doi.wiley.com/10.1111/febs.12516`.

[110] H. K. Lichtenthaler. The 1-deoxy-D-xylulose-5-phosphate pathway of isoprenoid biosynthesis in plants. *Annual Review of Plant Physiology and Plant Molecular Biology*, 50:47–65, 6 1999. ISSN 1040-2519. doi: 10.1146/annurev.arplant.50.1.47. URL `http://www.annualreviews.org/doi/abs/10.1146/annurev.arplant.50.1.47`.

[111] H. K. Lichtenthaler, M. Rohmer, and J. Schwender. Two Independent Biochemical Pathways for Isopentenyl Diphosphate and Isoprenoid Biosynthesis in Higher Plants. *Physiologia Plantarum*, 101(3):643–652, 11 1997. doi: 10.1111/j.1399-3054.1997.tb01049.x. URL `http://doi.wiley.com/10.1111/j.1399-3054.1997.tb01049.x`.

[112] Y. Lindqvist, G. Schneider, U. Ermler, and M. Sundstrom. 3-Dimensional Structure of Transketolase, a Thiamine Diphosphate Dependent Enzyme, at 2.5 $\angstrom$ Resolution. *The EMBO Journal*, 11(7):2373–2379, 1992.

[113] Y.-L. Liu, F. Guerra, K. Wang, W. Wang, J. Li, C. Huang, W. Zhu, K. Houlihan, Z. Li, Y. Zhang, S. K. Nair, and E. Oldfield. Structure, Function and Inhibition of the Two- and Three-domain 4Fe-4S IspG Proteins. *Proceedings of the National Academy of Sciences of the United States of America*, 109(22):8558–8563, 5 2012. doi: 10.1073/pnas.1121107109. URL `http://www.ncbi.nlm.nih.gov/pubmed/22586085http://www.ncbi.nlm.nih.gov/pubmed/22586085http://www.ncbi.nlm.nih.gov/pubmed/22586085`.

[114] L. M. Lois, N. Campos, S. R. Putra, K. Danielsen, M. Rohmer, and A. Boronat. Cloning and characterization of a gene from Escherichia

coli encoding a transketolase-like enzyme that catalyzes the synthesis of D-1-deoxyxylulose 5-phosphate, a common precursor for isoprenoid, thiamin, and pyridoxol biosynthesis. *Proceedings of the National Academy of Sciences of the United States of America*, 95(5):2105–2110, 3 1998. ISSN 0027-8424. URL `http://www.ncbi.nlm.nih.gov/pubmed/9482846http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC19265`.

[115] L. M. Lois, M. Rodriguez-Concepcion, F. Gallego, N. Campos, and A. Boronat. Carotenoid biosynthesis during tomato fruit development: regulatory role of 1-deoxy-D-xylulose 5-phosphate synthase. *The Plant Journal*, 22(6):503–513, 2000.

[116] R. Lonsdale, J. N. Harvey, and A. J. Mulholland. A Practical Guide to Modelling Enzyme-Catalysed Reactions. *Chemical Society Reviews*, 41(8):3025–3038, 2012. doi: 10.1039/c2cs15297e.

[117] H. Lüttgen, F. Rohdich, S. Herz, J. Wungsintaweekul, S. Hecht, C. A. Schuhr, M. Fellermeier, S. Sagner, M. H. Zenk, A. Bacher, and W. Eisenreich. Biosynthesis of terpenoids: YchB protein of Escherichia coli phosphorylates the 2-hydroxy group of 4-diphosphocytidyl-2C-methyl-D-erythritol. *Proceedings of the National Academy of Sciences of the United States of America*, 97(3):1062–7, 2 2000. ISSN 0027-8424. URL `http://www.ncbi.nlm.nih.gov/pubmed/10655484http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC15522`.

[118] A. Mac Sweeney, R. Lange, R. P. M. Fernandes, H. Schulz, G. E. Dale, A. Douangamath, P. J. Proteau, and C. Oefner. The Crystal Structure of E. coli 1-Deoxy-D-xylulose-5-phosphate Reductoisomerase in a Ternary Complex with the Antimalarial Compound Fosmidomycin and NADPH reveals a Tight-binding Closed Enzyme Conformation. *Journal of Molecular Biology*, 345(1):115–127, 2005. doi: 10.1016/j.jmb.2004.10.030.

[119] A. D. MacKerell, D. Bashford, M. Bellott, R. L. Dunbrack, J. D. Evanseck, M. J. Field, S. Fischer, J. Gao, H. Guo, S. Ha, D. Joseph-McCarthy, L. Kuchnir, K. Kuczera, F. T. K. Lau, C. Mattos, S. Michnick, T. Ngo, D. T. Nguyen, B. Prodhom, W. E. Reiher, B. Roux, M. Schlenkrich, J. C. Smith, R. Stote, J. Straub, M. Watanabe, J. Wiorkiewicz-Kuczera, D. Yin, and M. Karplus. All-atom Empirical Potential for Molecular Modeling and Dynamics studies of Proteins. *Journal of Physical Chemistry B*, 102(18):3586–3616, 1998.

[120] K. A. Manning, B. Sathyamoorthy, A. Eletsky, T. Szyperski, and A. S. Murkin. Highly Precise Measurement of Kinetic Isotope Effects Using HDetected 2D [H]-HSQC NMR Spectroscopy. *Journal of the American Chemical Society*, 134(51): 2058920592, 2012. doi: 10.1021/ja310353c. URL `http://pubs.acs.org/doi/pdfplus/10.1021/ja310353c`.

[121] A. Maraite, T. Schmidt, M. B. Ansörge-Schumacher, A. M. Brzozowski, and G. Grogan. Structure of the ThDP-dependent Enzyme Benzaldehyde Lyase Refined to 1.65 $\angstrom$ Resolution. *Acta Crystallographica Section F - Structural Biology and Crystallization Communications*, 63(Pt 7):546–548, 7 2007. ISSN 1744-3091. doi: 10.1107/S1744309107028576. URL `http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2335142&tool=pmcentrez&rendertype=abstract`.

[122] T. Masini, B. S. Kroezen, and A. K. H. Hirsch. Druggability of the Enzymes of the Non-mevalonate-pathway. *Drug Discovery Today*, 18 (23):1256–1262, 2013. doi: 10.1016/j.drudis.2013.07.003. URL `https://ac.els-cdn.com/S1359644613002377/1-s2.0-S1359644613002377-main.pdf?_tid=01c4d722-ca3c-11e7-8ef9-00000aab0f6b&acdnat=1510774654_41ad79b189e84568d5f6b401a18d1c53`.

[123] Y. Matsue, H. Mizuno, T. Tomita, T. Asami, M. Nishiyama, and T. Kuzuyama. The herbicide ketoclomazone inhibits 1-deoxy-D-xylulose 5-phosphate synthase in the 2-C-methyl-D-erythritol 4-phosphate pathway and shows antibacterial activity

against Haemophilus influenzae. *Journal of Antibiotics*, 63(10):583–588, 2010. doi: 10.1038/ja.2010.100. URL `<GotoISI>://WOS:000283978200001`.

[124] P. D. Matthews and E. T. Wurtzel. Metabolic engineering of carotenoid accumulation in Escherichia coli by modulation of the isoprenoid precursor pool with expression of deoxyxylulose phosphate synthase. *Applied Microbiology and Biotechnology*, 53(4):396–400, 4 2000. ISSN 0175-7598. doi: 10.1007/s002530051632. URL `http://link.springer.com/10.1007/s002530051632`.

[125] D. Meyer, P. Neumann, R. Ficner, and K. Tittmann. Observation of a Stable Carbene at the Active Site of a Thiamin Enzyme. *Nature: Chemical Biology*, 9(8): 488–490, 2013.

[126] L. Miallau, M. S. Alphey, L. E. Kemp, G. A. Leonard, S. M. McSweeney, S. Hecht, A. Bacher, W. Eisenreich, F. Rohdich, and W. N. Hunter. Biosynthesis of isoprenoids: Crystal structure of 4-diphosphocytidyl-2C-methyl-D-erythritol kinase. *Proceedings of the National Academy of Sciences of the United States of America*, 100(16):9173–9178, 8 2003. ISSN 0027-8424. doi: 10.1073/pnas.1533425100. URL `http://www.ncbi.nlm.nih.gov/pubmed/12878729http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC170891http://www.pnas.org/cgi/doi/10.1073/pnas.1533425100`.

[127] B. T. Miller, R. P. Singh, J. B. Klauda, M. Hodoscek, B. R. Brooks, and H. L. Woodcock III. CHARMMing: A new, flexible web portal for CHARMM. *Journal of Chemical Information and Modeling*, 48(9):1920–1929, 2008. doi: 10.1021/ci800133b. URL `<GotoISI>://WOS:000259398500018`.

[128] J. T. Mills, S. T. Furlong, and E. A. Dawidowicz. Plasma Membrane Biogenesis in Eukaryotic Cells: Translocation of Newly Synthesized Lipid. *Proceedings of the National Academy of Sciences of the United States of America*, 81:1385–1388, 1984. URL `http://www.pnas.org/content/81/5/1385.full.pdf`.

[129] L. Mitschke, C. Parthier, K. Schroder-Tittmann, J. Coy, S. Ludtke, and K. Tittmann. The Crystal Structure of Human Transketolase and New Insights into Its Mode of Action. *Journal of Biological Chemistry*, 285(41):31559–31570, 2010. doi: 10.1074/jbc.M110.149955.

[130] H. M. Miziorko. Enzymes of the mevalonate pathway of isoprenoid biosynthesis. *Archives of Biochemistry and Biophysics*, 505(2):131–143, 2011. doi: 10.1016/j. abb.2010.09.028. URL `<GotoISI>://WOS:000288286000001`.

[131] J. W. Munos, X. Pu, S. O. Mansoorabadi, H. J. Kim, and H.-w. Liu. A Secondary Kinetic Isotope Effect Study of the 1-Deoxy-D-xylulose-5-phosphate Reductoisomerase-Catalyzed Reaction: Evidence for a Retroaldol-Aldol Rearrangement. *Journal of the American Chemical Society*, 131(6):2048–2049, 2 2009. ISSN 0002-7863. doi: 10.1021/ja807987h. URL `http://pubs.acs.org/doi/abs/10.1021/ja807987h`.

[132] A. S. Murkin, K. A. Manning, and S. A. Kholodar. Mechanism and Inhibition of 1-Deoxy-D-xylulose-5-phosphate Reductoisomerase. *Bioorganic Chemistry*, 57:171–185, 2014. doi: 10.1016/j.bioorg.2014.06.001. URL `https://ac.els-cdn.com/S0045206814000467/1-s2.0-S0045206814000467-main.pdf?_tid=e0ad4cb4-c98b-11e7-92f0-00000aab0f6b&acdnat=1510699007_20041b054596561f9224562d11744b0c`.

[133] N. S. Nemeria, S. Chakraborty, A. Balakrishnan, and F. Jordan. Reaction Mechanisms of Thiamin Diphosphate Enzymes: Defining States of Ionization and Tautomerization of the Cofactor at Individual Steps. *The FEBS Journal*, 276(9): 2432–2446, 5 2009. ISSN 1742-4658. doi: 10.1111/j.1742-4658.2009.06964.x. URL `http://apps.webofknowledge.com/full_record.do?product=UA&search_mode=Refine&qid=4&SID=4FDs3H94dMWK4ZH3IHJ&page=1&doc=2`.

[134] M. Nikkola, Y. Lindqvist, and G. Schneider. Refined Structure of Transketolase from $\textit{Saccharomyces cerevisiae}$ at 2.0 $\angstrom$ Resolution. *Journal of Molecular Biology*, 238(3):387–404, 5 1994. doi: 10.1006/jmbi.1994.1299.

[135] U. Nilsson, L. Meshalkina, Y. Lindqvist, and G. Schneider. Examination of substrate binding in thiamin diphosphate-dependent transketolase by protein crystallography and site-directed mutagenesis. *Journal of Biological Chemistry*, 272(3): 1864–1869, 1997. URL `<GotoISI>://WOS:A1997WD05800071`.

[136] J. Norberg and L. Nilsson. Advances in Biomolecular Simulations: Methodology and Recent Applications. *Quarterly Reviews of Biophysics*, 36(3):257–306, 8 2003. doi: 10.1017/S0033583503003895. URL `http://www.journals.cambridge.org/abstract_S0033583503003895`.

[137] D. B. Northrop. On the Meaning of Km and V/K in Enzyme Kinetics. *Journal of Chemical Education*, 75(9):1153, 9 1998. ISSN 0021-9584. doi: 10.1021/ed075p1153. URL `http://pubs.acs.org/doi/abs/10.1021/ed075p1153`.

[138] M. H. M. Olsson, C. R. Søndergaard, M. Rostkowski, and J. H. Jensen. PROPKA3: Consistent treatment of internal and surface residues in empirical pKa predictions. *Journal of Chemical Theory and Computation*, 7(2):525–537, 2 2011. ISSN 1549-9618. doi: 10.1021/ct100578z. URL `http://dx.doi.org/10.1021/ct100578z`.

[139] P. E. ORourke, J. Kalinowska-Tłuścik, P. K. Fyfe, A. Dawson, and W. N. Hunter. Crystal structures of IspF from Plasmodium falciparum and Burkholderia cenocepacia: comparisons inform antimicrobial drug target assessment. *BMC Structural Biology*, 14(1):1–12, 1 2014. doi: 10.1186/1472-6807-14-1. URL `http://www.ncbi.nlm.nih.gov/pubmed/24410837`.

[140] S. Pandian, S. Saengchjan, and T. S. Raman. An Alternative Pathway for the Biosynthesis of Isoprenoid Compounds in Bacteria. *The Biochemical journal*, 196

(3):675–681, 6 1981. doi: 10.1042/BJ1960675. URL `http://www.ncbi.nlm.nih.gov/pubmed/6274317http://www.ncbi.nlm.nih.gov/pubmed/6274317`.

[141] H. Patel, N. S. Nemeria, L. A. Brammer, C. L. Freel Meyers, and F. Jordan. Observation of Thiamin-Bound Intermediates and Microscopic Rate Constants for their Interconversion on 1-Deoxy-D-xylulose 5-Phosphate Synthase: 600-fold Rate Acceleration of Pyruvate Decarboxylation by D-Glyceraldehyde-3-phosphate. *Journal of the American Chemical Society*, 134(44):18374–18379, 11 2012. ISSN 1520-5126. doi: 10.1021/ja307315u. URL `http://dx.doi.org/10.1021/ja307315u`.

[142] C. L. Perrin and J. B. Nielson. "Strong" Hydrogen Bonds in Chemistry and Biology. *Annual Review of Physical Chemistry*, 48:511–544, 1997. doi: 10.1146/annurev.physchem.48.1.511. URL `http://www.ncbi.nlm.nih.gov/pubmed/9348662`.

[143] C. Phaosiri and P. J. Proteau. Substrate Analogs for the Investigation of Deoxyxylulose 5-phosphate Reductoisomerase Inhibition: Synthesis and Evaluation. *Bioorganic & Medicinal Chemistry Letters*, 14 (21):5309–5312, 2004. doi: 10.1016/j.bmcl.2004.08.023. URL `https://ac.els-cdn.com/S0960894X04010315/1-s2.0-S0960894X04010315-main.pdf?_tid=5c9ffdd8-c7ae-11e7-b929-00000aab0f6b&acdnat=1510493915_b4fe7c108c331c97399bc31ea8b5fad7`.

[144] J. W. Porter and S. L. Spurgeon. *Biosynthesis of Isoprenoid Compounds*. Number v. 1 in Biosynthesis of Isoprenoid Compounds. Wiley, New York, 1981. URL `https://books.google.com/books?id=X2cXAQAAIAAJ`.

[145] J. Querol, A. Boronat, J. J. Centelles, S. Imperial, J. Querol-Audí, A. Boronat, J. J. Centelles, and S. Imperial. Catalytically Important Residues in E. coli 1-Deoxy-D-Xylulose 5-Phosphate Synthase. *Journal of Biosciences and Medicines*, 02(04): 30–35, 6 2014. ISSN 2327-5081. doi: 10.4236/jbm.2014.24006. URL `http://www.scirp.org/journal/PaperInformation.aspx?PaperID=46744&#abstract`.

[146] F. Quitterer, A. Frank, K. Wang, G. Rao, B. O'Dowd, J. Li, F. Guerra, S. Abdel-Azeim, A. Bacher, J. Eppinger, E. Oldfield, and M. Groll. Atomic-Resolution Structures of Discrete Stages on the Reaction Coordinate of the [Fe4S4] Enzyme IspG (GcpE). *Journal of Molecular Biology*, 427 (12):2220–2228, 6 2015. ISSN 00222836. doi: 10.1016/j.jmb.2015.04.002. URL `http://www.ncbi.nlm.nih.gov/pubmed/25868383http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC4433817http://linkinghub.elsevier.com/retrieve/pii/S0022283615002259`.

[147] D. Ramamoorthy, S. Handa, D. J. Merkler, and W. C. Guida. Plasmodium Vivax 1-Deoxy-D-Xylulose-5-Phosphate Synthase: Homology Modeling, Domain Swapping, and Virtual Screening. *Journal of Data Mining Genomics Proteomics*, 1:602–2153, 2014.

[148] K. E. Ranaghan and A. J. Mulholland. Investigations of Enzyme-catalysed Reactions with Combined Quantum Mechanics/Molecular Mechanics (QM/MM) Methods. *International Reviews in Physical Chemistry*, 29(1):65–133, 1 2010. ISSN 0144-235X. doi: 10.1080/01442350903495417. URL `http://www.tandfonline.com/doi/abs/10.1080/01442350903495417`.

[149] I. Rekittke, T. Nonaka, J. Wiesner, U. Demmer, E. Warkentin, H. Jomaa, and U. Ermler. Structure of the E -1-Hydroxy-2-methyl-but-2-enyl-4-diphosphate Synthase (GcpE) from Thermus thermophilus. *FEBS Letters*, 585(3):447–451, 2 2011. ISSN 00145793. doi: 10.1016/j.febslet.2010.12.012. URL `http://www.ncbi.nlm.nih.gov/pubmed/21167158http://doi.wiley.com/10.1016/j.febslet.2010.12.012`.

[150] I. Rekittke, H. Jomaa, and U. Ermler. Structure of the GcpE (IspG)-MEcPP Complex from Thermus thermophilus. *FEBS Letters*, 586(19): 3452–3457, 9 2012. ISSN 00145793. doi: 10.1016/j.febslet.2012.07.070.

URL `http://www.ncbi.nlm.nih.gov/pubmed/22967895http://doi.wiley.com/10.1016/j.febslet.2012.07.070`.

[151] K. Reuter, S. Sanderbrand, H. Jomaa, J. Wiesner, I. Steinbrecher, E. Beck, M. Hintz, G. Klebe, and M. T. Stubbs. Crystal Structure of 1-Deoxy-D-xylulose-5-phosphate Reductoisomerase, a Crucial Enzyme in the Non-mevalonate Pathway of Isoprenoid Biosynthesis. *Journal of Biological Chemistry*, 277(7):5378–5384, 2 2002. ISSN 0021-9258. doi: 10.1074/jbc. M109500200. URL `http://www.ncbi.nlm.nih.gov/pubmed/11741911http://www.jbc.org/lookup/doi/10.1074/jbc.M109500200`.

[152] S. B. Richard, M. E. Bowman, W. Kwiatkowski, I. Kang, C. Chow, A. M. Lillo, D. E. Cane, and J. P. Noel. Structure of 4-diphosphocytidyl-2-C-methylerythritol synthetase involved in mevalonate-independent isoprenoid biosynthesis. *Nature: Structural Biology*, 8(7):641–648, 2001. doi: 10.1038/89691.

[153] S. B. Richard, J.-L. L. Ferrer, M. E. Bowman, A. M. Lillo, C. N. Tetzlaff, D. E. Cane, and J. P. Noel. Structure and Mechanism of 2-C-Methyl-D-erythritol 2,4-Cyclodiphosphate Synthase - An Enzyme in the Mevalonate-independent of Isoprenoid Biosynthetic Pathway. *Journal of Biological Chemistry*, 277(10):8667–8672, 3 2002. ISSN 0021-9258. doi: 10.1074/jbc. C100739200. URL `http://www.ncbi.nlm.nih.gov/pubmed/11786530http://www.jbc.org/lookup/doi/10.1074/jbc.C100739200`.

[154] S. B. Richard, A. M. Lillo, C. N. Tetzlaff, M. E. Bowman, J. P. Noel, and D. E. Cane. Kinetic Analysis of Escherichia coli 2-C-Methyl-d-erythritol-4-phosphate Cytidyltransferase, Wild Type and Mutants, Reveals Roles of Active Site Amino Acids. *Biochemistry*, 43(38):12189–12197, 2004. doi: 10.1021/BI0487241. URL `http://pubs.acs.org/doi/abs/10.1021/bi0487241`.

[155] F. Rohdich, J. Wungsintaweekul, M. Fellermeier, S. Sagner, S. Herz, K. Kis, W. Eisenreich, A. Bacher, and M. H. Zenk. Cytidine 5'-Triphosphate-dependent Biosynthesis of Isoprenoids: YgbP Protein of Escherichia coli catalyzes the Formation of 4-Diphosphocytidyl-2-C-methylerythritol. *Proceedings of the National Academy of Sciences of the United States of America*, 96(21):11758–63, 10 1999. ISSN 0027-8424. doi: 10.1073/pnas.96.21.11758. URL `http://www.ncbi.nlm.nih.gov/pubmed/10518523http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC18359`.

[156] F. Rohdich, F. Zepeck, P. Adam, S. Hecht, J. Kaiser, R. Laupitz, T. Grawert, S. Amslinger, W. Eisenreich, A. Bacher, and D. Arigoni. The Deoxyxylulose Phosphate Pathway of Isoprenoid Biosynthesis: Studies on the Mechanisms of the Reactions catalyzed by IspG and IspH Protein. *Proceedings of the National Academy of Sciences of the United States of America*, 100(4):1586–1591, 2003. doi: 10.1073/pnas.0337742100.

[157] F. Rohdich, S. Lauw, J. Kaiser, R. Feicht, P. Köhler, A. Bacher, and W. Eisenreich. Isoprenoid Biosynthesis in Plants ? 2C-Methyl-d-erythritol-4-phosphate Synthase (IspC Protein) of Arabidopsis thaliana. *The FEBS Journal*, 273(19):4446–4458, 10 2006. ISSN 1742-464X. doi: 10.1111/j.1742-4658.2006.05446.x. URL `http://doi.wiley.com/10.1111/j.1742-4658.2006.05446.x`.

[158] M. Rohmer. The Mevalonate-Independent Methylerythritol 4-Phosphate (MEP) Pathway for Isoprenoid Biosynthesis, including Carotenoids. *Pure and Applied Chemistry*, 71(12):2279–2284, 1999. URL `https://www.iupac.org/publications/pac/pdf/1999/pdf/7112x2279.pdf`.

[159] M. Rohmer, M. Knani, P. Simonin, B. Sutter, and H. Sahm. Isoprenoid Biosynthesis in Bacteria: A Novel Pathway for the Early Steps Leading to Isopentenyl Diphosphate. *The Biochemical Journal*, 295(2):517–524, 10 1993. ISSN 0264-6021. doi: 10.1042/BJ2950517. URL

http://www.ncbi.nlm.nih.gov/pubmed/8240251http://www.pubmedcentral.
nih.gov/articlerender.fcgi?artid=PMC1134910%3CGoto.

[160] M. Rohmer, M. Seemann, S. Horbach, S. Bringer-Meyer, and H. Sahm. Glycer-
aldehyde 3-Phosphate and Pyruvate as Precursors of Isoprenic Units in an Al-
ternative Non-mevalonate Pathway for Terpenoid Biosynthesis. *Journal of the
American Chemical Society*, 118(11):2564–2566, 1 1996. ISSN 0002-7863. doi:
10.1021/ja9538344. URL http://pubs.acs.org/doi/abs/10.1021/ja9538344.

[161] J. C. Sacchettini and C. D. Poulter. Biochemistry - Creating Isoprenoid Diversity.
Science, 277(5333):1788–1789, 1997. doi: 10.1126/science.277.5333.1788.

[162] D. R. Salahub, A. de la Lande, A. Goursot, R. Zhang, and Y. Zhang. Re-
cent Progress in Density Functional Methodology for Biomolecular Modeling. In
M. V. Putz and D. M. P. Mingos, editors, *Applications of Density Functional
Theory to Biological and Bioinorganic Chemistry*, pages 1–64. Springer Berlin
Heidelberg, Berlin, Heidelberg, 2013. ISBN 978-3-642-32750-6. doi: 10.1007/
978-3-642-32750-6{_}1. URL https://doi.org/10.1007/978-3-642-32750-6_
1.

[163] G. Schenk, F. J. Leeper, R. England, P. F. Nixon, and R. G. Duggleby. Inves-
tigation of the mechanistic functions of residues HIS113 and HIS114 in pyruvate
decarboxylase from Zymomonas mobilis: A proposed model in the binding of the
substrate pyruvate. *The FASEB Journal*, 11(9):A1135–A1135, 1997.

[164] P. V. Schleyer and H. J. Jiao. What is aromaticity? *Pure and Applied Chemistry*,
68(2):209–218, 1996.

[165] M. Schlitzer and R. Ortmann. Feeding the Antimalarial Pipeline. *ChemMedChem*, 5
(11):1837–1840, 11 2010. doi: 10.1002/cmdc.201000341. URL http://doi.wiley.
com/10.1002/cmdc.201000341.

[166] M. Seemann, B. T. S. Bui, M. Wolff, D. Tritsch, N. Campos, A. Boronat, A. Marquet, and M. Rohmer. Isoprenoid Biosynthesis through the Methylery-thritol Phosphate Pathway: The (E)-4-Hydroxy-3-methylbut-2-enyl Diphosphate Synthase (GcpE) is a [4Fe4S] Protein. *Angewandte Chemie International Edition*, 41(22):4337–4339, 11 2002. doi: Doi10.1002/1521-3773(20021115) 41:22⟨4337::Aid-Anie4337⟩3.0.Co;2-K. URL `http://www.ncbi.nlm.nih.gov/pubmed/12434382`.

[167] H. M. Senn and W. Thiel. QM/MM Methods for Biomolecular Systems. *Ange-wandte Chemie International Edition*, 48(7):1198–1229, 2 2009. doi: 10.1002/anie. 200802019. URL `http://doi.wiley.com/10.1002/anie.200802019`.

[168] T. Sgraja, M. S. Alphey, S. Ghilagaber, R. Marquez, M. N. Robertson, J. L. Hemmings, S. Lauw, F. Rohdich, A. Bacher, W. Eisenreich, V. Illarionova, and W. N. Hunter. Characterization of Aquifex aeolicus 4-diphosphocytidyl-2C-methyl-D-erythritol kinase - ligand recognition in a template for antimicrobial drug discovery. *The FEBS Journal*, 275(11):2779–2794, 6 2008. ISSN 1742464X. doi: 10.1111/j.1742-4658.2008.06418.x. URL `http://www.ncbi.nlm.nih.gov/pubmed/18422643http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC2655357http://doi.wiley.com/10.1111/j.1742-4658.2008.06418.x`.

[169] S. Shan, X. Chen, T. Liu, H. Zhao, Z. Rao, and Z. Lou. Crystal structure of 4-diphosphocytidyl-2-C-methyl-D-erythritol kinase (IspE) from Mycobacterium tu-berculosis. *The FASEB Journal*, 25(5):1577–1584, 5 2011. ISSN 0892-6638. doi: 10. 1096/fj.10-175786. URL `http://www.ncbi.nlm.nih.gov/pubmed/21282208http://www.fasebj.org/cgi/doi/10.1096/fj.10-175786`.

[170] Y. Shao, L. F. Molnar, Y. Jung, J. Kussmann, C. Ochsenfeld, S. T. Brown, A. T. B. Gilbert, L. V. Slipchenko, S. V. Levchenko, D. P. O'Neill, R. A. DiStasio Jr., R. C. Lochan, T. Wang, G. J. O. Beran, N. A. Besley, J. M. Herbert, C. Y. Lin, T. Van Voorhis, S. H. Chien, A. Sodt, R. P. Steele, V. A. Rassolov, P. E. Maslen,

P. P. Korambath, R. D. Adamson, B. Austin, J. Baker, E. F. C. Byrd, H. Dachsel, R. J. Doerksen, A. Dreuw, B. D. Dunietz, A. D. Dutoi, T. R. Furlani, S. R. Gwaltney, A. Heyden, S. Hirata, C.-P. Hsu, G. Kedziora, R. Z. Khalliulin, P. Klunzinger, A. M. Lee, M. S. Lee, W. Liang, I. Lotan, N. Nair, B. Peters, E. I. Proynov, P. A. Pieniazek, Y. M. Rhee, J. Ritchie, E. Rosta, C. D. Sherrill, A. C. Simmonett, J. E. Subotnik, H. L. Woodcock III, W. Zhang, A. T. Bell, A. K. Chakraborty, D. M. Chipman, F. J. Keil, A. Warshel, W. J. Hehre, H. F. Schaefer III, J. Kong, A. I. Krylov, P. M. W. Gill, and M. Head-Gordon. Advances in Methods and Algorithms in a Modern Quantum Chemistry Program Package. *Physical Chemistry Chemical Physics*, 8(27):3172–3191, 2006. doi: 10.1039/b517914a.

[171] C. G. M. G. Sinead Heuston Maire Begley, C. Hill, S. Heuston, M. Begley, C. G. M. Gahan, and C. Hill. Isoprenoid Biosynthesis in Bacterial Pathogens. *Microbiology*, 158(Pt 6):1389–1401, 6 2012. ISSN 1465-2080. doi: 10.1099/mic.0.051599-0. URL `http://mic.sgmjournals.org.ezproxy.lib.usf.edu/content/158/Pt_6/1389.abstract`.

[172] N. Singh, G. Cheve, M. Avery, and C. McCurdy. Targeting the Methyl Erythritol Phosphate (MEP) Pathway for Novel Antimalarial, Antibacterial and Herbicidal Drug Discovery: Inhibition of 1-Deoxy-D-Xylulose-5-Phosphate Reductoisomerase (DXR) Enzyme. *Current Pharmaceutical Design*, 13(11):1161–1177, 4 2007. doi: 10.2174/138161207780618939. URL `http://www.eurekaselect.com/openurl/content.php?genre=article&issn=1381-6128&volume=13&issue=11&spage=1161`.

[173] C. K. Singleton, J. J. L. Wang, L. Shan, and P. R. Martin. Conserved Residues are Functionally Distinct within Transketolases of Different Species. *Biochemistry*, 35 (49):15865–15869, 12 1996. doi: 10.1021/bi9616920.

[174] X. Sisquella, K. de Pourcq, J. Alguacil, J. Robles, F. Sanz, D. Anselmetti, S. Imperial, and X. Fernàndez-Busquets. A single-molecule force spectroscopy nanosen-

sor for the identification of new antibiotics and antimalarials. *The FASEB Journal*, 24(11):4203–17, 11 2010. ISSN 1530-6860. doi: 10.1096/fj.10-155507. URL `http://www.ncbi.nlm.nih.gov/pubmed/20634351`.

[175] C. R. Søndergaard, M. H. M. Olsson, M. M. Rostkowski, J. H. Jensen, C. R. Sondergaard, M. H. M. Olsson, M. M. Rostkowski, and J. H. Jensen. Improved Treatment of Ligands and Coupling Effects in Empirical Calculation and Rationalization of pK(a) Values. *Journal of Chemical Theory and Computation*, 7(7):2284–2295, 5 2011. ISSN 1549-9618. doi: 10.1021/ct200133y. URL `http://dx.doi.org/10.1021/ct200133y`.

[176] G. A. Sprenger, U. Schorken, T. Wiegert, S. Grolle, A. A. DeGraaf, S. V. Taylor, T. P. Begley, S. BringerMeyer, H. Sahm, U. Schrken, T. Wiegert, S. Grolle, A. A. de Graaf, S. V. Taylor, T. P. Begley, S. Bringer-Meyer, H. Sahm, U. Schörken, T. Wiegert, S. Grolle, A. A. de Graaf, S. V. Taylor, T. P. Begley, S. Bringer-Meyer, and H. Sahm. Identification of a Thiamin-Dependent Synthase in Escherichia coli Required for the Formation of the 1-Deoxy-D-xylulose 5-Phosphate Precursor to Isoprenoids, Thiamin, and Pyridoxol. *Proceedings of the National Academy of Sciences of the United States of America*, 94(24):12857–12862, 11 1997. ISSN 0027-8424. URL `http://www.ncbi.nlm.nih.gov/pubmed/9371765http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC24228http://www.pnas.org/content/94/24/12857.full.pdf`.

[177] S. Steinbacher, J. Kaiser, J. Wungsintaweekul, S. Hecht, W. Eisenreich, S. Gerhardt, A. Bacher, and F. Rohdich. Structure of 2C-methyl-d-erythritol-2,4-cyclodiphosphate synthase involved in mevalonate-independent biosynthesis of isoprenoids. *Journal of Molecular Biology*, 316(1):79–88, 2 2002. ISSN 00222836. doi: 10.1006/jmbi.2001.5341. URL `http://www.ncbi.nlm.nih.gov/pubmed/11829504http://linkinghub.elsevier.com/retrieve/pii/S0022283601953410`.

[178] S. Steinbacher, J. Kaiser, W. Eisenreich, R. Huber, A. Bacher, and F. Rohdich. Structural Basis of Fosmidomycin Action Revealed by the Complex with 2-C-Methyl-D-erythritol 4-Phosphate Synthase (IspC) - Implications for the Catalytic Mechanism and Anti-Malaria Drug Development. *Journal of Biological Chemistry*, 278(20):18401–18407, 5 2003. doi: 10.1074/jbc.M300993200. URL http://www.jbc.org/content/278/20/18401.full.pdfhttp://www.ncbi.nlm.nih.gov/pubmed/12621040.

[179] S. Takahashi, T. Kuzuyama, H. Watanabe, and H. Seto. A 1-Deoxy-D-xylulose 5-Phosphate Reductoisomerase Catalyzing the Formation of 2-C-Methyl-D-elythritol 4-Phosphate in an Alternative Nonmevalonate Pathway for Terpenoid Biosynthesis. *Proceedings of the National Academy of Sciences of the United States of America*, 95(17):9879–9884, 8 1998. ISSN 0027-8424. URL http://www.ncbi.nlm.nih.gov/pubmed/9707569http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC21430.

[180] D. G. Truhlar. Transition State Theory for Enzyme Kinetics. *Archives of Biochemistry and Biophysics*, 582:10–17, 2015. doi: 10.1016/j.abb.2015.05.004. URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4555010/pdf/nihms705440.pdf.

[181] D. L. Turner, H. Santos, P. Fareleira, I. Pacheco, J. LeGall, and A. V. Xavier. Structure determination of a novel cyclic phosphocompound isolated from Desulfovibrio desulfuricans. *The Biochemical Journal*, 285 (Pt 2):387–90, 7 1992. URL http://www.ncbi.nlm.nih.gov/pubmed/1637331.

[182] M. W. Van Der Kamp and A. J. Mulholland. Combined Quantum Mechanics/Molecular Mechanics (QM/MM) Methods in Computational Enzymology. *Biochemistry*, 52(16):2708–2728, 2013. ISSN 00062960. doi: 10.1021/bi400215w.

[183] J.-Y. Van Der Meer and A. K. H. Hirsch. The Isoprenoid-precursor Dependence of Plasmodium spp. *Natural Product Reports*, 29(7):721–728, 2012. doi: 10.

1039/c2np20013a. URL `http://pubs.rsc.org/en/content/articlepdf/2012/np/c2np20013a`.

[184] K. Vanommeslaeghe, E. Hatcher, C. Acharya, S. Kundu, S. Zhong, J. Shim, E. Darian, O. Guvench, P. Lopes, I. Vorobyov, and A. D. MacKerell Jr. CHARMM General Force Field: A Force Field for Drug-Like Molecules Compatible with the CHARMM All-Atom Additive Biological Force Fields. *Journal of Computational Chemistry*, 31(4):671–690, 2010. doi: 10.1002/jcc.21367. URL `<GotoISI>://WOS:000274922000002`.

[185] T. Wada, T. Kuzuyama, S. Satoh, S. Kuramitsu, S. Yokoyama, S. Unzai, J. R. Tame, and S.-Y. Park. Crystal Structure of 4-(Cytidine 5-diphospho)-2- ¡i¿C¡/i¿ -methyl-d-erythritol kinase, an Enzyme in the Non-mevalonate Pathway of Isoprenoid Synthesis. *Journal of Biological Chemistry*, 278(32):30022–30027, 8 2003. ISSN 0021-9258. doi: 10.1074/jbc.M304339200. URL `http://www.ncbi.nlm.nih.gov/pubmed/12771135http://www.jbc.org/lookup/doi/10.1074/jbc.M304339200`.

[186] W. Wang and E. Oldfield. Bioorganometallic Chemistry with IspG and IspH: Structure, Function, and Inhibition of the [Fe ¡sub¿4¡/sub¿ S ¡sub¿4¡/sub¿] Proteins Involved in Isoprenoid Biosynthesis. *Angewandte Chemie International Edition*, 53(17):4294–4310, 4 2014. ISSN 14337851. doi: 10.1002/anie.201306712. URL `http://doi.wiley.com/10.1002/anie.201306712`.

[187] W. Wang, J. Li, K. Wang, C. Huang, Y. Zhang, and E. Oldfield. Organometallic Mechanism of Action and Inhibition of the 4Fe-4S Isoprenoid Biosynthesis Protein GcpE (IspG). *Proceedings of the National Academy of Sciences of the United States of America*, 107(25):11189–11193, 6 2010. doi: 10.1073/pnas.1000264107. URL `http://www.ncbi.nlm.nih.gov/pubmed/20534554`.

[188] A. Warshel. Multiscale Modeling of Biological Functions: From Enzymes to Molecular Machines (Nobel Lecture). *Angewandte Chemie International Edition*, 53 (38):10020–10031, 9 2014. ISSN 14337851. doi: 10.1002/anie.201403689. URL `http://doi.wiley.com/10.1002/anie.201403689`.

[189] S. E. Wheeler, K. N. Houk, P. V. R. Schleyer, and W. D. Allen. A Hierarchy of Homodesmotic Reactions for Thermochemistry. *Journal of the American Chemical Society*, 131(7):2547–2560, 2009. doi: 10.1021/ja805843n. URL `<GotoISI>://WOS:000263576100041`.

[190] C. Wikner, L. Meshalkina, U. Nilsson, M. Nikkola, Y. Lindqvist, M. Sundstrom, and G. Schneider. Analysis of an Invariant Cofactor-Protein Interaction in Thiamin Diphosphate-Dependent Enzymes by Site-Directed Mutagenesis: Glutamic-Acid-418 in Transketolase is Essential for Catalysis. *Journal of Biological Chemistry*, 269(51):32144–32150, 1994.

[191] E. I. Wilding, J. R. Brown, A. P. Bryant, A. F. Chalker, D. J. Holmes, K. A. Ingraham, S. Iordanescu, C. Y. So, M. Rosenberg, and M. N. Gwynn. Identification, Evolution, and Essentiality of the Mevalonate Pathway for Isopentenyl Diphosphate Biosynthesis in Gram-positive Cocci. *Journal of Bacteriology*, 182(15):4319–4327, 8 2000. doi: 10.1128/JB.182.15.4319-4327.2000. URL `http://www.ncbi.nlm.nih.gov/pubmed/10894743`.

[192] M. Wolff, M. Seemann, C. Grosdemange-Billiard, D. Tritsch, N. Campos, M. Rodrguez-Concepción, A. Boronat, and M. Rohmer. Isoprenoid Biosynthesis via the Methylerythritol Phosphate Pathway. (E)-4-Hydroxy-3-methylbut-2-enyl Diphosphate: Chemical Synthesis and Formation from Methylerythritol Cyclodiphosphate by a Cell-Free System from Escherichia coli. *Tetrahedron Letters*, 43 (14):2555–2559, 4 2002. doi: 10.1016/S0040-4039(02)00293-9. URL `http://www.sciencedirect.com/science/article/pii/S0040403902002939?via%3Dihub`.

[193] A. Wong, J. W. Munos, V. Devasthali, K. A. Johnson, and H.-w. Liu. Study of 1-Deoxy-d-xylulose-5-phosphate Reductoisomerase: Synthesis and Evaluation of Fluorinated Substrate Analogues. *Organic Letters*, 6(20):3625–3628, 2004. doi: 10.1021/OL048459B. URL `http://pubs.acs.org/doi/abs/10.1021/ol048459b`.

[194] H. L. Woodcock, M. Hodoscek, and B. R. Brooks. Exploring SCC-DFTB paths for mapping QM/MM reaction mechanisms. *Journal of Physical Chemistry A*, 111(26):5720–5728, 2007. doi: 10.1021/jp0714217. URL `<GotoISI>://WOS: 000247573600017`.

[195] H. L. Woodcock, M. Hodoscek, A. T. B. Gilbert, P. M. W. Gill, H. F. Schaefer, B. R. Brooks, H. L. Woodcock III, M. Hodoscek, A. T. B. Gilbert, P. M. W. Gill, H. F. Schaefer III, and B. R. Brooks. Interfacing Q-chem and CHARMM to perform QM/MM reaction path calculations. *Journal of Computational Chemistry*, 28(9):1485–1502, 7 2007. ISSN 0192-8651. doi: 10.1002/jcc.20587. URL `http://www.ncbi.nlm.nih.gov/pubmed/17334987`.

[196] C. J. Woodrow and S. Krishna. Antimalarial Drugs: Recent Advances in Molecular Determinants of Resistance and their Clinical Significance. *Cellular and Molecular Life Sciences*, 63(14):1586–1596, 2006. doi: 10.1007/s00018-006-6071-1. URL `https://link.springer.com/content/pdf/10.1007%2Fs00018-006-6071-1.pdf`.

[197] S. Xiang, G. Usunow, G. Lange, M. Busch, and L. Tong. Crystal structure of 1-deoxy-d-xylulose 5-phosphate synthase, a crucial enzyme for isoprenoids biosynthesis. *Journal of Biological Chemistry*, 282(4):2676–2682, 1 2007. ISSN 0021-9258. doi: 10.1074/jbc.M610235200. URL `http://www.ncbi.nlm.nih.gov/pubmed/17135236http://www.jbc.org/lookup/doi/10.1074/jbc.M610235200`.

[198] Y. Xiao, D. Rooker, Q. You, C. L. Freel Meyers, and P. Liu. IspG-Catalyzed Positional Isotopic Exchange in Methylerythritol Cyclodiphosphate of the De-

oxyxylulose Phosphate Pathway: Mechanistic Implications. *ChemBioChem*, 12(4):527–530, 3 2011. ISSN 14394227. doi: 10.1002/cbic.201000716. URL `http://www.ncbi.nlm.nih.gov/pubmed/22238143http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC3257810http://doi.wiley.com/10.1002/cbic.201000716`.

[199] S. Yajima, T. Nonaka, T. Kuzuyama, H. Seto, and K. Ohsawa. Crystal structure of 1-Deoxy-D-xylulose 5-Phosphate Reductoisomerase Complexed with Cofactors: Implications of a Flexible Loop Movement upon Substrate Binding. *Journal of Biochemistry*, 131(3):313–317, 3 2002. ISSN 0021-924X. URL `http://www.ncbi.nlm.nih.gov/pubmed/11872159`.

[200] S. Yajima, K. Hara, D. Iino, Y. Sasaki, T. Kuzuyama, K. Ohsawa, and H. Seto. Structure of 1-Deoxy-D-xylulose 5-Phosphate Reductoisomerase in a Quaternary Complex with a Magnesium Ion, NADPH and the Antimalarial Drug Fosmidomycin. *Acta Crystallographica Section F - Structural Biology and Crystallization Communications*, 63(6):466–470, 6 2007. doi: 10.1107/S1744309107024475. URL `http://www.ncbi.nlm.nih.gov/pubmed/17554164`.

[201] E. Yeh and J. L. DeRisi. Chemical Rescue of Malaria Parasites lacking an Apicoplast defines Organelle Function in Blood-Stage Plasmodium falciparum. *PLoS Biology*, 9(8):e1001138, 8 2011. doi: 10.1371/journal.pbio.1001138. URL `http://dx.plos.org/10.1371/journal.pbio.1001138`.

[202] X. H. Yin and P. J. Proteau. Characterization of Native and Histidine-tagged Deoxyxylulose 5-Phosphate Reductoisomerase from the Cyanobacterium $\textit{Synechocystis sp}$ PCC6803. *Biochimica et Biophysica Acta - Proteins and Proteomics*, 1652(1):75–81, 2003. doi: 10.1016/j.bbapap.2003.08.005.

[203] J. Zeidler, J. Schwender, C. Müller, J. Wiesnerb, C. W. Eidemeyerb, E. Beckb, H. Jomaab, H. K. Lichtenthaler, J. Zeidler, J. Schwender, C. Muller, J. Wies-

ner, C. Weidemeyer, E. Beck, H. Jomaa, and H. K. Lichtenthaler. Inhibition of the Non-mevalonate 1-Deoxy-D-xylulose-5-phosphate Pathway of Plant Isoprenoid Biosynthesis by Fosmidomycin. *Zeitschrift Fur Naturforsch. C-a J. Biosci.*, 53 (11-12):980–986, 1998. URL `https://www.degruyter.com/downloadpdf/j/znc.1998.53.issue-11-12/znc-1998-11-1208/znc-1998-11-1208.pdf`.

[204] F. Zepeck, T. Gräwert, J. Kaiser, N. Schramek, W. Eisenreich, A. Bacher, and F. Rohdich. Biosynthesis of Isoprenoids. Purification and Properties of IspG Protein from Escherichia coli. *The Journal of Organic Chemistry*, 70(23):9168–9174, 2005. doi: 10.1021/JO0510787. URL `http://pubs.acs.org/doi/10.1021/jo0510787`.

[205] D. Zhou and R. H. White. Early Steps of Isoprenoid Biosynthesis in Escherichia coli. *The Biochemical Journal*, 273(3):627–634, 2 1991. doi: 10.1042/BJ2730627. URL `http://www.ncbi.nlm.nih.gov/pubmed/1996960`.

Appendix A: Supporting Information for "Thiamine Diphosphate Activation in 1-deoxy-D-xylulose 5-Phosphate Synthase: Insights into the Mechanism and Underlying Intermolecular Interactions"

A.1   Methods

A.1.1   Topology and Parameters for Thiamine Diphosphate (TDP)

TDP has never been parameterized for use in MM calculations. In order to build 1-deoxy-d-xylulose 5-phosphate synthase (DXS), it was necessary to develop a topology and parameter file that would reproduce the TDP crystal structure. TDP has a pyrimidine ring, thiazole ring, and an inorganic phosphate tail; each of which have been developed for use in CHARMM calculations. The topology and parameters for each moiety were used and augmented to account for TDP's final structure. The additional bonds and parameters were determined based on structures with homologous chemical properties. QM calculations at the B3LYP/6-31G* level of theory were used to determine acquire initial charges for undefined atoms (e.g., C1 atom, *vide infra*). These charges were combined with those already established for each moiety. MacKerrell's charge rules (mackerell.umaryland.edu/ff_dev.html) for substituents was followed to account for the linkers between groups. Finally, the atomic charges were balanced via manual manipulation and chemical intuition to equal the final -2.0 charge for TDP. A minimization of TDP was performed using the topology and parameter files to a tolerance of 0.002 kcal·mol^{-1}·Å^{-1}. The minimized structure's RMSD was compared to that of the crystal

structure. The two structures were found to deviation by 0.1Å.

```
*Topology File for Thiamine Diphosphate using CGenFF Atom Types
*
36 1


RESI TDP              -2.00 !
GROUP        ! 0.00
ATOM C1   CG331       0.43 !              H2
ATOM H1   HGA3        0.16 !              |
ATOM H2   HGA3        0.16 !          H1--C1--H3
ATOM H3   HGA3        0.16 !              |
ATOM C2   CG2R64     -0.50 !             C2
ATOM N1   NG2R62     -0.49 !            /  \\
ATOM N2   NG2R62     -0.54 !           /     \\
ATOM C3   CG2R62     -0.04 !          N1      N2
ATOM H4   HGR62       0.11 !          ||       |
ATOM C4   CG2R64      0.45 !          C3      C4
ATOM N3   NG2S3      -0.80 !         / \    // \
ATOM H31  HGP4        0.38 !       H4  \ //    N3--H31
ATOM H32  HGP4        0.38 !            C5      |
ATOM C5   CG2R62      0.14 !            |       H32
GROUP        !  1.00
ATOM C6   CG324      -0.24 !          H5-C6-H6
ATOM H5   HGA2        0.11 !              |
ATOM H6   HGA2        0.11 !        H8     N4(+)
ATOM N4   NG2R52      0.59 !         |    /    \\
ATOM C7   CG2R53     -0.41 ! H7--C10--C8     C7--H71
```

110

```
ATOM H71  HGR52      0.14 !       |   ||       |
ATOM S1   SG2R50     0.73 !       |   ||    /
ATOM C8   CG2R51     0.40 !      H9   C9--S1
ATOM C9   CG2R51    -0.25 !             |
ATOM C10  CG331     -0.54 !      H10--C11--H11
ATOM H7   HGA3       0.12 !             |
ATOM H8   HGA3       0.12 !      H12--C12--H13
ATOM H9   HGA3       0.12 !             |
GROUP       ! -3.00
ATOM C11  CG321     -0.30 !            O1
ATOM H10  HGA2       0.10 !             |
ATOM H11  HGA2       0.10 !      O4==P1==O3 (-)
ATOM C12  CG321     -0.16 !             |
ATOM H12  HGA2       0.10 !            O2
ATOM H13  HGA2       0.10 !             |
ATOM O1   OG303     -0.62 !      O7==P2==O6 (-)
ATOM P1   PG1        1.50 !             |
ATOM O2   OG304     -0.74 !           O5 (-)
ATOM O3   OG2P1     -0.80 !
ATOM O4   OG2P1     -0.80 !
ATOM P2   PG2        1.10 !
ATOM O5   OG2P1     -0.86 !
ATOM O6   OG2P1     -0.86 !
ATOM O7   OG2P1     -0.86 !
ATOM O4   OG2P1     -0.80 !
ATOM P2   PG2        1.10 !
ATOM O5   OG2P1     -0.86 !
ATOM O6   OG2P1     -0.86 !
```

ATOM O7   OG2P1   -0.86 !

BOND C1   H1   C1   H2   C1   H3

BOND C1   C2   N1   C2   N1   C3

BOND C2   N2   N2   C4   C4   N3   N3   H31   N3   H32

BOND C4   C5   C5   C3   C3   H4

BOND C5   C6   C6   H5   C6   H6   C6   N4

BOND N4   C7   C7   H71   N4   C8   C8   C10

BOND C10 H7   C10 H8   C10 H9

BOND C8   C9   C9   S1   C9   C11   S1   C7

BOND C11 H10   C11 H11   C11 C12

BOND C12 H12   C12 H13   C12 O1

BOND O1   P1   P1   O2   P1   O3   P1   O4

BOND O2   P2   P2   O5   P2   O6   P2   O7

IC C2   N1   C3   C5   0.0000 0.0   0.0   0.0 0.0000

IC C3   N1   C2   N2   0.0000 0.0   0.0   0.0 0.0000

IC N1   C2   N2   C4   0.0000 0.0   0.0   0.0 0.0000

IC N2   C4   N3   H31   0.0000 0.0   0.0   0.0 0.0000

*Parameter File for Thiamine Diphosphate

*

BONDS

CG324 NG2R52   400.00   1.4580

CG2R51 CG334   229.63   1.5000

CG2R62 CG324   222.50   1.4900

CG2R62 NG2R62   302.00   1.3430

```
CG2R64 CG331      230.00   1.4780


ANGLES

CG321  CG2R51 SG2R50      45.80   124.00

CG331  CG2R51 NG2R52      45.80   124.00

CG331  CG2R64 NG2R62      45.80   121.00

CG2R62 CG324  HGA2        55.00   110.10

NG2R52 CG324  HGA2        33.43   110.10

CG2R62 CG324  HGP5        49.30   107.50

CG2R51 CG321  CG321       58.35   114.00

CG2R51 NG2R52 CG324       70.00   126.90

CG2R53 NG2R52 CG324       70.00   126.90

CG2R62 CG2R62 CG324       45.80   119.00

CG2R64 CG2R62 CG324       45.80   119.00

NG2R62 CG2R62 HGR62       44.00   115.00

NG2R52 CG2R53 HGR52       32.00   126.00

CG2R62 NG2R62 CG2R64      40.00   110.50

CG2R64 CG2R62 NG2R62      20.00   124.00

CG2R62 CG2R62 NG2R62      85.00   122.90

NG2R52 CG2R53 SG2R50     110.00   117.20

NG2R52 CG324  HGP5        33.43   110.10   22.53   2.179

CG2R64 CG331  HGA3        33.43   110.10   22.53   2.179

CG2R62 CG324  NG2R52      45.00   102.30   35.00   2.101


DIHEDRAL

HGA3   CG331  CG2R64 NG2R62     2.8000  2   180.00

CG321  CG321  OG303  PG2        0.0000  3     0.00

CG331  CG2R64 NG2R62 CG2R62     1.0000  2   180.00


                      113
```

```
CG331  CG2R64 NG2R62 CG2R64      1.0000  2   180.00

CG324  CG2R62 CG2R62 HGR62       2.4000  2   180.00

CG324  NG2R52 CG2R53 HGR52       2.0000  2   180.00

CG331  CG2R51 CG2R51 CG321       1.2000  2   180.00

CG324  NG2R52 CG2R51 CG331       3.0000  2   180.00

NG2S3  CG2R64 CG2R62 CG324       0.0000  2   180.00

CG324  NG2R52 CG2R51 CG2R51      5.4000  2   180.00

CG324  NG2R52 CG2R53 SG2R50      6.0000  2   180.00

CG2R51 CG321  CG321  HGA2        0.1500  3     0.00

CG2R51 CG321  CG321  OG303       0.4000  3   180.00

CG2R62 CG324  NG2R52 CG2R53      0.2300  2   180.00

CG2R62 CG324  NG2R52 CG2R51      0.2300  2   180.00

CG2R51 NG2R52 CG324  HGA2        0.0000  3     0.00

CG2R53 NG2R52 CG324  HGA2        0.0000  3     0.00

CG2R62 CG2R62 CG324  HGA2        0.0000  3     0.00

CG2R64 CG2R62 CG324  HGA2        0.0000  2     0.00

NG2R52 CG2R51 CG331  HGA3        0.1900  3     0.00

SG2R50 CG2R51 CG321  HGA2        0.1900  3     0.00

CG2R51 CG2R51 CG321  CG321       0.2000  1     0.00

CG2R51 CG2R51 CG321  CG321       0.2700  2     0.00

CG2R51 CG2R51 CG321  CG321       0.0000  3     0.00

SG2R50 CG2R51 CG321  CG321       0.1900  3     0.00

CG2R62 CG2R62 CG324  NG2R52      0.1500  2   180.00

CG2R64 CG2R62 CG324  NG2R52      0.1500  2   180.00

CG2R51 NG2R52 CG2R53 HGR52       2.0000  2   180.00

CG2R64 NG2R62 CG2R62 HGR62       4.5000  2   180.00

CG2R53 NG2R52 CG2R51 CG331       3.0000  2   180.00

CG2R53 SG2R50 CG2R51 CG321       8.5000  2   180.00
```

114

```
NG2R62 CG2R62 CG2R62 CG324      0.0000  2   180.00

NG2R62 CG2R64 CG2R62 CG324      3.1000  2   180.00

SG2R50 CG2R51 CG2R51 CG331      3.0000  2   180.00

CG2R64 NG2R62 CG2R62 CG2R62     1.2000  2   180.00

CG2R64 NG2R62 CG2R64 CG2R62     2.0000  2   180.00

NG2R62 CG2R62 CG2R62 CG2R64     3.0000  2   180.00

NG2R52 CG2R53 SG2R50 CG2R51     8.5000  2   180.00

NG2R62 CG2R64 NG2R62 CG2R62     3.1000  2   180.00

NG2R52 CG2R51 CG2R51 SG2R50     7.0000  2   180.00

SG2R50 CG2R53 NG2R52 CG2R51     2.0000  2   180.00


END
```

A.1.2   Extended 20 ns Molecular Dynamics Simulation

The crystal structure for the *D. radiodurans* DXS (PDB ID:201X)[197] enzyme with TDP
bound was processed and parsed via www.charmming.org[127]. The topology and param-
eters for TDP shown above were used in generating the structure used in this extended
simulation. Structural modifications were performed to ensure the active site Glu373 was
protonated in agreement with experimental evidence[90,112]. CGenFF and CHARMM22
protein (C22) force fields[119] were used throughout. The system was solvated in a cubic
crystal structure and neutralized with KCl salt to a final concentration of 0.15M. The
system was heated from 110K to 310K over 100ps. The system was simulated for 20 ns
at constant pressure (1atm) and temperature (310K). For the purposes of analysis, the
first 2 ns were discarded to allow for an equilibration period.

A.1.3 Over-expression and Purification of Wildtype DXS and the DXS Mutants

Plasmids containing the wild type D. radiodurans dxs gene or the mutant dxs gene were transformed into E. coli BL-21 B(DE3) cells and used for protein expression. An overnight culture of E. coli in LB broth containing 50 μg/ml kanamycin was diluted 100-fold, cultured at 37°C until the absorbance at 600 nm reached $\sim$ 0.6, and then cooled to 20°C. Expression was induced by the addition of 0.5 mM isopropyl β-D-1-thiogalactopyranoside (IPTG). The cells were harvested by centrifugation (6,000 $\times$ g for 10 min) after being shaken for 6 hrs at 20°C and the resulting cell pellets stored at -80°C before purification. Cells were thawed and all the purification steps were performed at 4°C. Cells were resuspended in binding buffer (20 mM Tris, 500 mM NaCl, 5 mM imidazole, 10 mM β-Me, pH = 7.5) supplemented with 1 mM phenylmethanesulfonylfluoride (PMSF), 4 μg/mL leupeptin, and 2 μg/mL pepstatin, sonicated using a Heat systems W-380 ultrasonic processor, and centrifuged (16,000 $\times$ g for 20 min) to remove cell debris. The supernatant from the cell lysate was applied to a 1.5 cm $\times$ 5 cm column packed with Ni-NTA resin that had been equilibrated with binding buffer. Non-bound proteins eluted from the column by first washing with 5 column volumes of binding buffer followed by 20 column volumes of wash buffer (20 mM Tris pH 7.5, 500 mM NaCl, 60 mM imidazole, and 10 mM β-Me). The bound DXS (wildtype or mutant) was eluted from the Ni-NTA resin using elution buffer (20 mM Tris pH 7.5, 500 mM NaCl, 250 mM imidazole, and 10 mM β-Me). A flow rate of 1.5 mL/min was maintained through the Ni-NTA column for all the loading and washing steps. DXS-containing fractions containing were combined, exhaustively dialyzed at 4°C against 20 mM Tris pH 7.5, 100 mM NaCl, and 10 mM β-Me, and concentrated by ultrafiltration. The final yield of DXS (wildtype or mutant) was 7-8 mg/L of E. coli culture medium. Enzyme was flash frozen in liquid nitrogen, stored at -80°C. The purity of the DXS (wildtype or the mutant) was evaluated by SDS-PAGE.
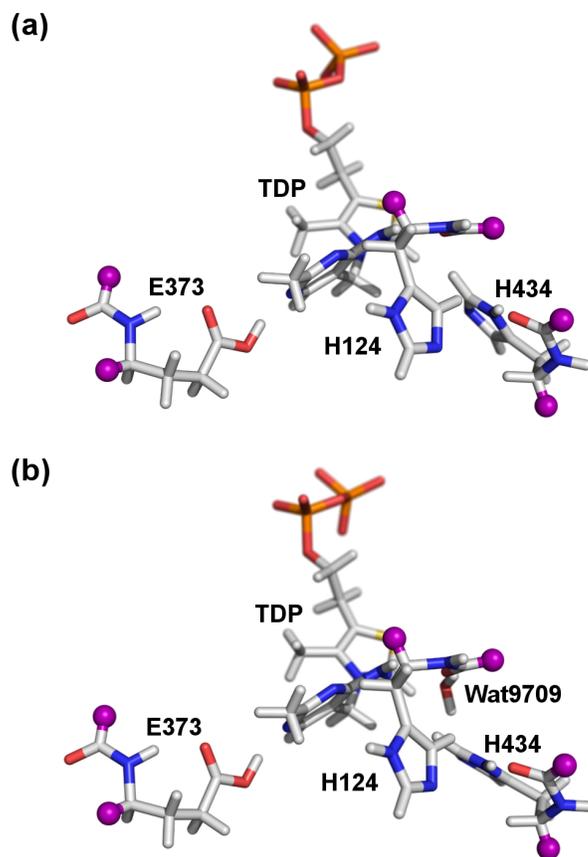
Figure **A.1**: An illustration of the QM regions of both the direct histidine mechanism (DHM, image (a)) and the water-mediated mechanism (WMM, image (b)). DHM's QM region contained 98 atoms; which were made up from H124, E373, H434 and TDP. WMM's QM region was comprised of all the same residues as the DHM with the addition of Wat9709 (reactive water) making the total 101 atoms. Additionally, linker atoms were used the C_O–C_α bond (purple atoms).
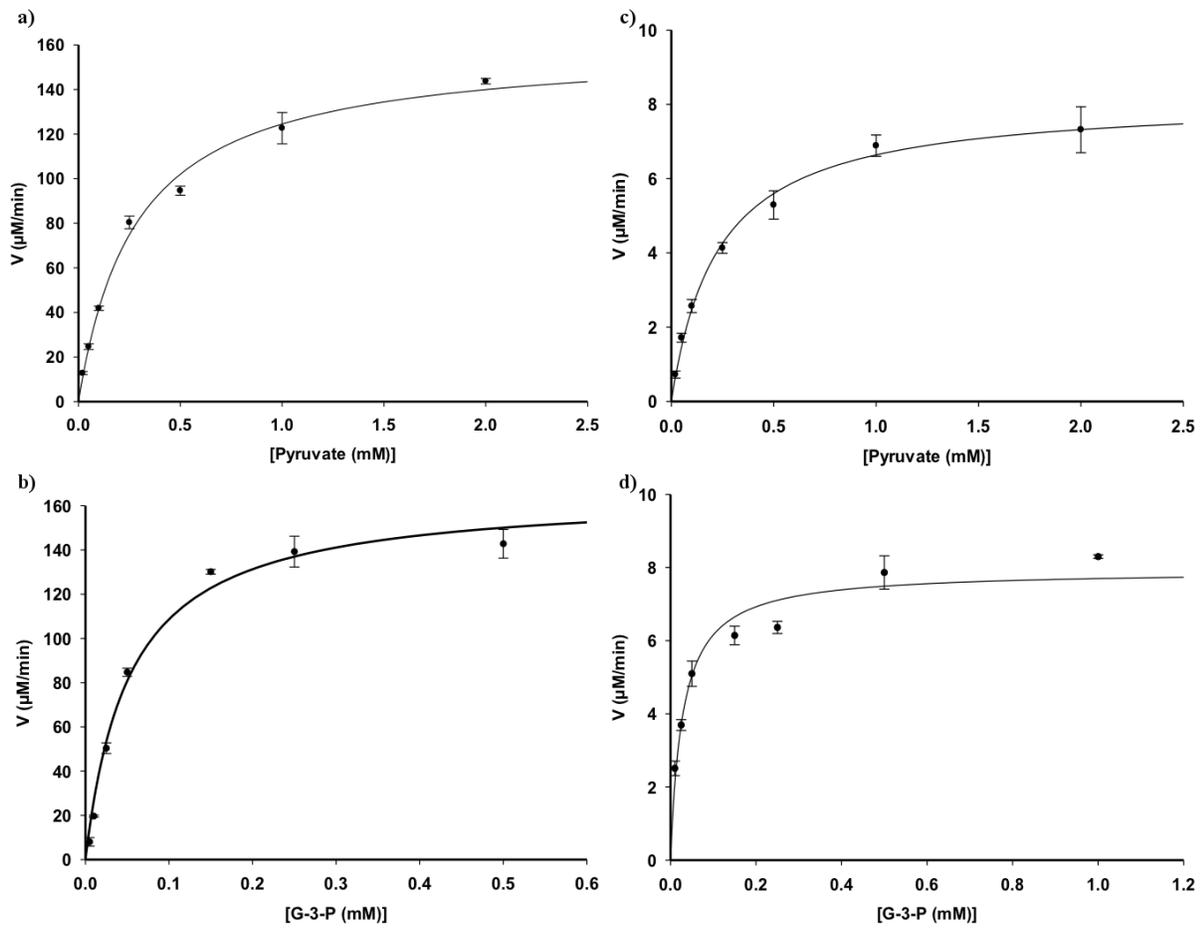
Figure **A.2**: Plots of the initial velocities versus varying concentrations of pyruvate or G3P. Plots a and b are for wild-type DXS while plots c and d represent the H82A mutant.
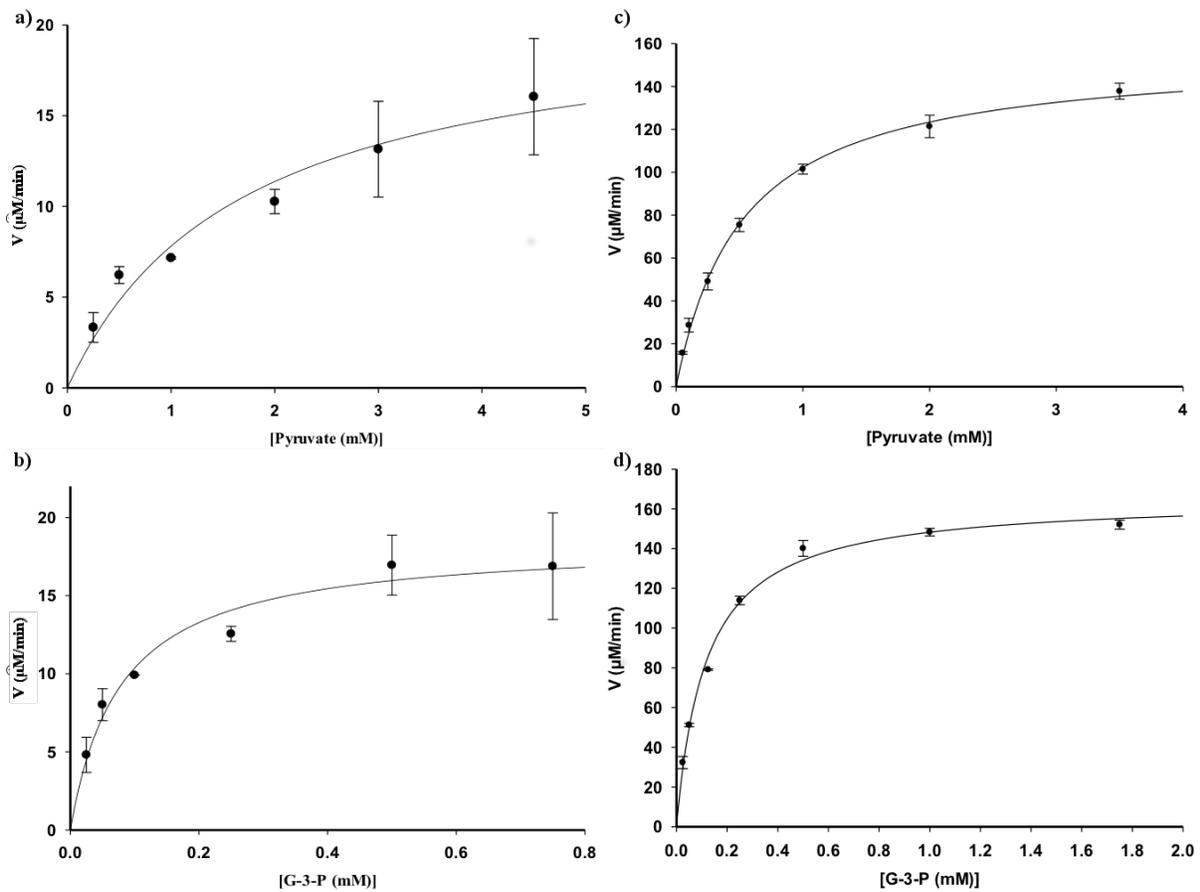
Figure **A.3**: Graphs of initial velocities versus varying concentrations of pyruvate or G3P. Plots a and b are for the H304A mutant. Plots c and d represent the D430A mutant.
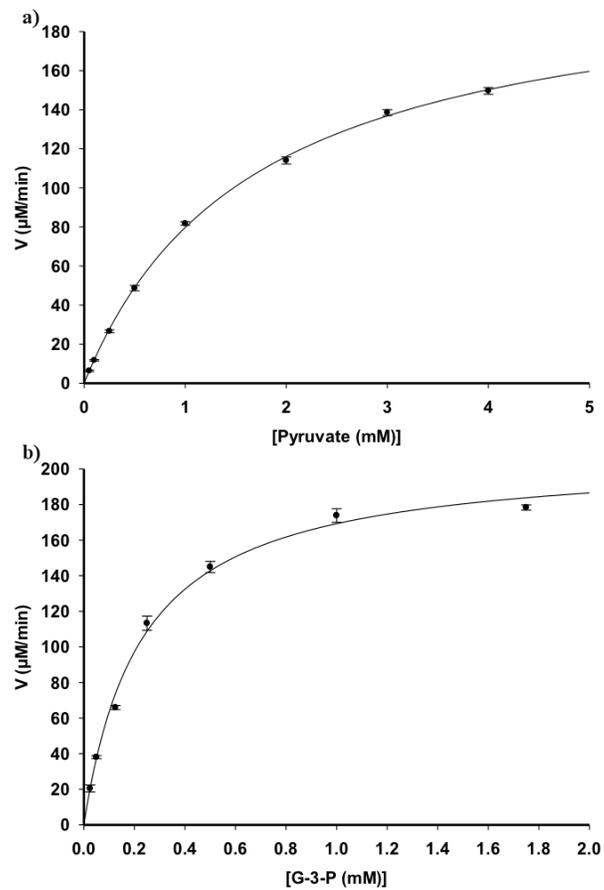
Figure **A.4**: Plots of the initial velocities versus varying concentrations of pyruvate or G3P. All graphs are for the H434A mutant discussed in primary manuscript.

A.2 ProPKA3.1 Results

Below will be found the ProPKA3.1 results for monomer A of DXS:
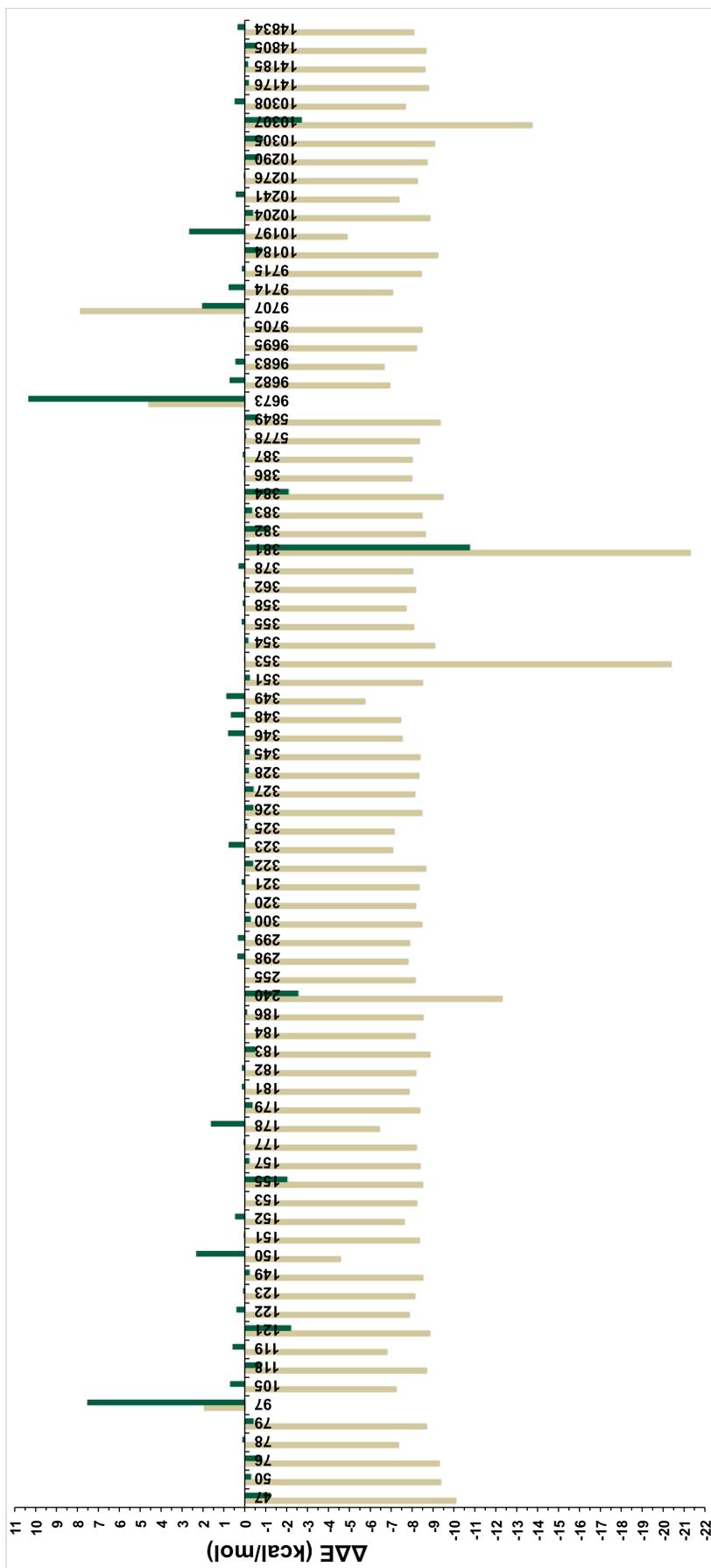
Figure **A.5**: This bar graph represents all of the CPA data acquired for all of the residues within 5.0Å of the QM region. DHM is in gold while WMM is in green. What is immediately apparent is the significantly larger change in magnitude for all DHM results in comparison to that of the WMM. This is thought to come from the large change in the active site dipole.

```
propka3.1                                                            2014-09-29
-------------------------------------------------------------------------------
--                                                                           --
--                   PROPKA: A PROTEIN PKA PREDICTOR                         --
--                                                                           --
--                  VERSION 1.0,  04/25/2004, IOWA CITY                      --
--                            BY HUI LI                                      --
--                                                                           --
--               VERSION 2.0,  11/05/2007, IOWA CITY/COPENHAGEN              --
--                  BY DELPHINE C. BAS AND DAVID M. ROGERS                   --
--                                                                           --
--                  VERSION 3.0,  01/06/2011, COPENHAGEN                     --
--              BY MATS H.M. OLSSON AND CHRESTEN R. SONDERGARD               --
--                                                                           --
--                  VERSION 3.1,  07/01/2011, COPENHAGEN                     --
--              BY CHRESTEN R. SONDERGARD AND MATS H.M. OLSSON               --
-------------------------------------------------------------------------------
```

```
-------------------------------------------------------------------------------
 References:

   Very Fast Empirical Prediction and Rationalization of Protein pKa Values
   Hui Li, Andrew D. Robertson and Jan H. Jensen
   PROTEINS: Structure, Function, and Bioinformatics 61:704-721 (2005)

   Very Fast Prediction and Rationalization of pKa Values for Protein-Ligand Complexes
   Delphine C. Bas, David M. Rogers and Jan H. Jensen
   PROTEINS: Structure, Function, and Bioinformatics 73:765-783 (2008)

   PROPKA3: Consistent Treatment of Internal and Surface Residues in Empirical pKa predictions
   Mats H.M. Olsson, Chresten R. Sondergard, Michal Rostkowski, and Jan H. Jensen
   Journal of Chemical Theory and Computation, 7(2):525-537 (2011)

   Improved Treatment of Ligands and Coupling Effects in Empirical Calculation
    and Rationalization of pKa Values
   Chresten R. Sondergaard, Mats H.M. Olsson, Michal Rostkowski, and Jan H. Jensen
   Journal of Chemical Theory and Computation, (2011)

-------------------------------------------------------------------------------
```

```
 ---------  -----  ------  ---------------------  --------------  --------------  --------------
                           DESOLVATION  EFFECTS      SIDECHAIN       BACKBONE       COULOMBIC
   RESIDUE    pKa   BURIED  REGULAR      RE        HYDROGEN BOND   HYDROGEN BOND   INTERACTION
```

```
---------   -----   ------   ---------   ---------   --------------   --------------   --------------

ASP   9 A    3.88     0 %    0.12  160    0.00    0    0.00 XXX   0 X    0.00 XXX   0 X   -0.05 ARG  43 A
ASP   9 A                                              0.00 XXX   0 X    0.00 XXX   0 X    0.01 ASP  14 A


ASP  14 A    2.41     0 %    0.49  266    0.00    0   -0.72 SER   8 A    0.00 XXX   0 X   -0.00 N+    5 A
ASP  14 A                                             -0.54 THR  10 A    0.00 XXX   0 X   -0.02 ARG  43 A
ASP  14 A                                             -0.32 ARG  94 A    0.00 XXX   0 X   -0.02 HIS  17 A
ASP  14 A                                              0.00 XXX   0 X    0.00 XXX   0 X   -0.27 ARG  94 A


ASP  21 A    2.72     0 %    0.44  281    0.00    0   -0.22 ARG  24 A   -0.78 HIS  17 A   -0.04 LYS  20 A
ASP  21 A                                              0.00 XXX   0 X    0.00 XXX   0 X   -0.01 LYS  23 A
ASP  21 A                                              0.00 XXX   0 X    0.00 XXX   0 X   -0.28 ARG  24 A
ASP  21 A                                              0.00 XXX   0 X    0.00 XXX   0 X   -0.20 HIS  17 A


ASP  60 A    2.08    83 %    3.05  513    1.08    0   -0.61 THR  34 A   -0.07 ASP  60 A   -0.10 LYS 291 A
ASP  60 A                                             -1.58 ARG  38 A    0.00 XXX   0 X   -1.13 ARG  38 A
ASP  60 A                                             -1.60 HIS 262 A    0.00 XXX   0 X   -0.77 HIS 262 A


ASP  70 A    4.70*   20 %    0.75  338    0.17    0   -0.36 ARG  73 A   -0.03 ASP  70 A   -0.09 LYS  20 A
ASP  70 A                                              0.67 ASP  74 A    0.00 XXX   0 X   -0.28 LYS  23 A
ASP  70 A                                              0.00 XXX   0 X    0.00 XXX   0 X   -0.02 HIS 117 A
ASP  70 A                                              0.00 XXX   0 X    0.00 XXX   0 X   -0.08 HIS 147 A
ASP  70 A                                              0.00 XXX   0 X    0.00 XXX   0 X   -0.27 ARG  73 A
ASP  70 A                                              0.00 XXX   0 X    0.00 XXX   0 X    0.43 ASP  74 A


ASP  74 A    1.39*   50 %    1.78  422    0.40    0   -0.67 ASP  70 A   -0.78 ASP  70 A   -0.07 LYS  23 A
ASP  74 A                                             -1.60 HIS 147 A   -0.82 HIS 147 A   -0.05 ARG  73 A
ASP  74 A                                              0.00 XXX   0 X    0.00 XXX   0 X   -0.03 ARG 139 A
ASP  74 A                                              0.00 XXX   0 X    0.00 XXX   0 X   -0.56 HIS 147 A


ASP  79 A    4.78   100 %    4.09  606    0.40    0   -0.73 SER 128 A   -0.81 SER 126 A   -0.22 ARG 401 A
ASP  79 A                                             -0.85 SER 156 A   -0.01 THR 127 A   -0.24 HIS  82 A
ASP  79 A                                              0.00 XXX   0 X   -0.64 SER 128 A    0.00 XXX   0 X


ASP  95 A    3.85     4 %    0.26  292    0.00    0    0.00 XXX   0 X    0.00 XXX   0 X   -0.17 ARG  94 A
ASP  95 A                                              0.00 XXX   0 X    0.00 XXX   0 X    0.03 ASP  14 A
ASP  95 A                                              0.00 XXX   0 X    0.00 XXX   0 X   -0.06 HIS  17 A


ASP  99 A    2.75     0 %    0.31  243    0.00    0   -0.85 LYS 102 A    0.00 XXX   0 X   -0.13 ARG  47 A
ASP  99 A                                              0.00 XXX   0 X    0.00 XXX   0 X   -0.38 LYS 102 A


ASP 118 A    0.14    82 %    3.19  510    0.81    0   -1.60 HIS  87 A   -0.75 LYS 111 A   -0.01 LYS 111 A
ASP 118 A                                             -1.47 ARG  93 A    0.00 XXX   0 X   -0.08 HIS  66 A
ASP 118 A                                              0.00 XXX   0 X    0.00 XXX   0 X   -0.15 HIS 117 A
ASP 118 A                                              0.00 XXX   0 X    0.00 XXX   0 X   -1.47 HIS  87 A
```

124

```
ASP 118 A                                         0.00 XXX   0 X    0.00 XXX   0 X   -0.99 LYS  88 A
ASP 118 A                                         0.00 XXX   0 X    0.00 XXX   0 X   -1.13 ARG  93 A

ASP 140 A   2.00    5 %   0.64  295   0.06   0   -0.84 ARG 139 A    0.00 XXX   0 X   -0.25 LYS 175 A
ASP 140 A                                        -0.74 ARG 174 A    0.00 XXX   0 X   -0.41 ARG 139 A
ASP 140 A                                         0.00 XXX   0 X    0.00 XXX   0 X   -0.26 ARG 174 A

ASP 145 A   3.69    0 %   0.16  216   0.00   0    0.00 XXX   0 X   -0.10 ASP 145 A   -0.04 ARG 139 A
ASP 145 A                                         0.00 XXX   0 X    0.00 XXX   0 X   -0.11 LYS 175 A
ASP 145 A                                         0.00 XXX   0 X    0.00 XXX   0 X    0.04 ASP  74 A
ASP 145 A                                         0.00 XXX   0 X    0.00 XXX   0 X   -0.08 HIS 147 A
ASP 145 A                                         0.00 XXX   0 X    0.00 XXX   0 X    0.03 ASP 276 A

ASP 154 A   4.87*  95 %   3.64  546   1.01   0   -0.18 ASN 183 A   -0.38 ASP 154 A   -4.07 MG   MG A
ASP 154 A                                        -0.04 TDP O22 A   -0.08 GLY 155 A   -0.19 LYS 289 A
ASP 154 A                                         0.28 ASP 182 A   -0.23 ASN 183 A   -0.11 HIS 284 A
ASP 154 A                                         0.00 XXX   0 X    0.00 XXX   0 X    0.86 ASP 182 A
ASP 154 A                                         0.00 XXX   0 X    0.00 XXX   0 X    0.55 TDP O21 A

ASP 171 A   3.96    0 %   0.17  171   0.00   0    0.00 XXX   0 X    0.00 XXX   0 X   -0.01 ARG 174 A

ASP 182 A   1.80   62 %   1.88  456   0.13   0   -0.28 ASP 154 A   -0.55 GLU 184 A   -1.03 MG   MG A
ASP 182 A                                        -1.60 HIS 284 A    0.00 XXX   0 X   -0.55 HIS 284 A

ASP 260 A   3.25   18 %   0.74  332   0.02   0    0.00 XXX   0 X   -0.82 HIS 262 A   -0.01 ARG  38 A
ASP 260 A                                         0.00 XXX   0 X   -0.29 ASN 263 A   -0.06 LYS 291 A
ASP 260 A                                         0.00 XXX   0 X    0.00 XXX   0 X    0.14 ASP  60 A
ASP 260 A                                         0.00 XXX   0 X    0.00 XXX   0 X   -0.26 HIS 262 A

ASP 276 A   3.36    3 %   0.32  290   0.01   0   -0.51 HIS 147 A    0.00 XXX   0 X   -0.02 LYS 175 A
ASP 276 A                                         0.00 XXX   0 X    0.00 XXX   0 X    0.04 ASP  74 A
ASP 276 A                                         0.00 XXX   0 X    0.00 XXX   0 X   -0.28 HIS 147 A

ASP 278 A   3.75    0 %   0.17  212   0.00   0    0.00 XXX   0 X   -0.02 ASP 278 A   -0.15 LYS 175 A
ASP 278 A                                         0.00 XXX   0 X    0.00 XXX   0 X   -0.06 ARG 254 A
ASP 278 A                                         0.00 XXX   0 X    0.00 XXX   0 X    0.01 ASP 276 A

ASP 299 A   3.71   21 %   0.81  341   0.02   0    0.00 XXX   0 X   -0.08 ILE 301 A   -0.04 LYS 289 A
ASP 299 A                                         0.00 XXX   0 X   -0.81 TYR 302 A    0.00 XXX   0 X

ASP 310 A   3.71    0 %   0.43  201   0.00   0    0.00 XXX   0 X   -0.10 ASP 310 A   -0.00 LYS 308 A
ASP 310 A                                         0.00 XXX   0 X   -0.42 THR 313 A    0.00 XXX   0 X

ASP 339 A   3.64   42 %   1.76  398   0.09   0   -0.84 THR 342 A   -0.74 ARG 341 A   -0.21 ARG 341 A
ASP 339 A                                         0.00 XXX   0 X   -0.06 THR 342 A   -0.04 ARG 365 A
ASP 339 A                                         0.00 XXX   0 X    0.00 XXX   0 X   -0.12 ARG 389 A
```

125

```
ASP 368 A    5.57    80 %    2.37  505   0.53     0     0.00 XXX    0 X    -0.40 ALA 348 A    -0.19 ARG 350 A
ASP 368 A                                              0.00 XXX    0 X    -0.69 GLY 370 A     0.05 GLU 189 A
ASP 368 A                                              0.00 XXX    0 X     0.00 XXX    0 X     0.11 GLU 351 A

ASP 404 A    4.06     0 %    0.33  254   0.00     0     0.00 XXX    0 X     0.00 XXX    0 X    -0.22 HIS 408 A
ASP 404 A                                              0.00 XXX    0 X     0.00 XXX    0 X     0.15 ASP 409 A

ASP 409 A    2.62     3 %    0.59  291   0.03     0    -0.85 THR 378 A     0.00 XXX    0 X    -0.32 HIS 408 A
ASP 409 A                                             -0.63 HIS 408 A     0.00 XXX    0 X     0.00 XXX    0 X

ASP 422 A    4.51   100 %    3.67  576   1.07     0    -0.84 SER 325 A    -0.74 ARG 423 A    -0.05 ARG 480 A
ASP 422 A                                             -0.62 ARG 423 A     0.00 XXX    0 X    -1.77 ARG 423 A

ASP 430 A    0.37   100 %    2.39  582   0.00     0    -0.85 LYS 101 A     0.00 XXX    0 X    -0.16 ARG 423 A
ASP 430 A                                              0.00 XXX    0 X     0.00 XXX    0 X    -0.32 ARG 480 A
ASP 430 A                                              0.00 XXX    0 X     0.00 XXX    0 X    -0.03 HIS 304 A
ASP 430 A                                              0.00 XXX    0 X     0.00 XXX    0 X    -0.02 HIS 597 A
ASP 430 A                                              0.00 XXX    0 X     0.00 XXX    0 X    -0.84 HIS  51 A
ASP 430 A                                              0.00 XXX    0 X     0.00 XXX    0 X    -2.03 LYS 101 A
ASP 430 A                                              0.00 XXX    0 X     0.00 XXX    0 X    -1.57 HIS 434 A

ASP 439 A    5.32   100 %    4.06  640   0.42     0    -0.60 SER 396 A    -0.03 ASP 439 A     0.13 ASP 561 A
ASP 439 A                                             -0.42 ARG 477 A     0.00 XXX    0 X    -2.03 ARG 477 A

ASP 456 A    4.07    55 %    1.47  436   0.00     0     0.00 XXX    0 X    -0.22 ALA 458 A    -0.27 LYS 455 A
ASP 456 A                                              0.00 XXX    0 X    -0.67 GLU 459 A    -0.03 ARG 536 A

ASP 471 A    3.58     0 %    0.21  257   0.00     0     0.00 XXX    0 X     0.00 XXX    0 X    -0.08 ARG 389 A
ASP 471 A                                              0.00 XXX    0 X     0.00 XXX    0 X    -0.10 ARG 450 A
ASP 471 A                                              0.00 XXX    0 X     0.00 XXX    0 X    -0.18 HIS 414 A
ASP 471 A                                              0.00 XXX    0 X     0.00 XXX    0 X    -0.06 HIS 470 A

ASP 493 A    2.83     0 %    0.28  240   0.00     0    -0.85 LYS 465 A     0.00 XXX    0 X    -0.02 HIS 470 A
ASP 493 A                                              0.00 XXX    0 X     0.00 XXX    0 X    -0.38 LYS 465 A

ASP 506 A    3.37     0 %    0.31  230   0.00     0     0.00 XXX    0 X    -0.26 ASP 506 A    -0.20 ARG 501 A
ASP 506 A                                              0.00 XXX    0 X     0.00 XXX    0 X    -0.28 ARG 552 A

ASP 507 A    3.96     6 %    0.54  299   0.04     0    -0.09 ARG 554 A     0.00 XXX    0 X    -0.38 ARG 554 A
ASP 507 A                                              0.00 XXX    0 X     0.00 XXX    0 X     0.04 ASP 506 A

ASP 518 A    5.43    62 %    1.23  455   0.76     0     0.00 XXX    0 X    -0.00 ASP 518 A    -0.38 LYS 515 A
ASP 518 A                                              0.00 XXX    0 X     0.00 XXX    0 X     0.02 ASP 456 A

ASP 526 A    3.97     0 %    0.17  169   0.00     0     0.00 XXX    0 X     0.00 XXX    0 X    -0.03 LYS 503 A
```

126

```
ASP 526 A                                    0.00 XXX   0 X    0.00 XXX   0 X   -0.04 LYS 522 A
ASP 526 A                                    0.00 XXX   0 X    0.00 XXX   0 X   -0.11 ARG 615 A
ASP 526 A                                    0.00 XXX   0 X    0.00 XXX   0 X    0.16 GLU 525 A

ASP 542 A   4.02   34 %   1.37  377  0.28  0  0.00 XXX   0 X   -0.81 TRP 499 A   -0.05 ARG 536 A
ASP 542 A                                    0.00 XXX   0 X   -0.01 GLU 544 A    0.06 GLU 544 A
ASP 542 A                                    0.00 XXX   0 X   -0.62 MET 545 A    0.00 XXX   0 X

ASP 561 A   2.66  100 %   3.52  609  0.68  0 -1.60 HIS 604 A   -0.61 ILE 426 A   -0.43 LYS 515 A
ASP 561 A                                    0.00 XXX   0 X    0.00 XXX   0 X   -0.88 ARG 477 A
ASP 561 A                                    0.00 XXX   0 X    0.00 XXX   0 X   -1.82 HIS 604 A

ASP 592 A   4.02    0 %   0.22  202  0.00  0  0.00 XXX   0 X    0.00 XXX   0 X    0.00 XXX   0 X

ASP 610 A   4.24   17 %   0.89  329  0.04  0  0.00 XXX   0 X   -0.03 ASP 610 A    0.01 GLU 601 A
ASP 610 A                                    0.00 XXX   0 X   -0.35 ALA 613 A   -0.12 HIS 604 A

ASP 624 A   3.93    0 %   0.13  161  0.00  0  0.00 XXX   0 X    0.00 XXX   0 X    0.00 XXX   0 X

GLU  28 A   4.41    0 %   0.26  185  0.00  0  0.00 XXX   0 X   -0.26 GLU  28 A   -0.10 ARG  27 A

GLU  35 A   4.59    0 %   0.16  277  0.00  0  0.00 XXX   0 X    0.00 XXX   0 X   -0.06 ARG  38 A
GLU  35 A                                    0.00 XXX   0 X    0.00 XXX   0 X   -0.00 LYS 291 A

GLU  36 A   4.95   28 %   1.84  359  0.22  0  0.00 XXX   0 X   -0.83 LEU  12 A   -0.03 ARG  43 A
GLU  36 A                                    0.00 XXX   0 X   -0.84 LEU  13 A   -0.11 ARG  94 A
GLU  36 A                                    0.00 XXX   0 X    0.00 XXX   0 X    0.02 ASP   9 A
GLU  36 A                                    0.00 XXX   0 X    0.00 XXX   0 X    0.14 ASP  14 A
GLU  36 A                                    0.00 XXX   0 X    0.00 XXX   0 X    0.04 GLU  40 A

GLU  40 A   4.20   57 %   1.62  440  0.17  0 -0.54 ARG  43 A    0.00 XXX   0 X    0.07 ASP  14 A
GLU  40 A                                   -0.69 ARG  94 A    0.00 XXX   0 X    0.08 ASP  95 A
GLU  40 A                                    0.00 XXX   0 X    0.00 XXX   0 X   -0.40 ARG  43 A
GLU  40 A                                    0.00 XXX   0 X    0.00 XXX   0 X   -0.60 ARG  94 A

GLU 103 A   3.66    0 %   0.37  248  0.00  0 -0.85 SER 602 A    0.00 XXX   0 X   -0.33 ARG 606 A
GLU 103 A                                    0.00 XXX   0 X    0.00 XXX   0 X   -0.03 HIS 597 A

GLU 114 A   4.03   64 %   2.45  461  0.71  0 -0.85 LYS  88 A   -0.80 SER 107 A   -0.09 LYS 111 A
GLU 114 A                                   -0.49 ARG  93 A    0.00 XXX   0 X   -0.10 HIS  87 A
GLU 114 A                                    0.00 XXX   0 X    0.00 XXX   0 X   -0.18 HIS 597 A
GLU 114 A                                    0.00 XXX   0 X    0.00 XXX   0 X   -1.01 LYS  88 A
GLU 114 A                                    0.00 XXX   0 X    0.00 XXX   0 X   -0.66 ARG  93 A
GLU 114 A                                    0.00 XXX   0 X    0.00 XXX   0 X    0.57 ASP 118 A

GLU 116 A   4.36    0 %   0.31  255  0.00  0  0.00 XXX   0 X    0.00 XXX   0 X   -0.07 LYS  20 A
```

127

```
  GLU 116 A                                       0.00 XXX    0 X     0.00 XXX    0 X    -0.13 ARG   73 A
  GLU 116 A                                       0.00 XXX    0 X     0.00 XXX    0 X    -0.01 ARG   75 A
  GLU 116 A                                       0.00 XXX    0 X     0.00 XXX    0 X    -0.25 HIS  117 A

  GLU 184 A    5.15    16 %   0.32  325   0.04   0 0.00 XXX    0 X     0.00 XXX    0 X    -0.08 MG    MG A
  GLU 184 A                                       0.00 XXX    0 X     0.00 XXX    0 X     0.15 ASP  154 A
  GLU 184 A                                       0.00 XXX    0 X     0.00 XXX    0 X     0.23 ASP  182 A

  GLU 189 A    3.74    29 %   0.67  362   0.00   0 -0.33 ARG 350 A    -0.61 GLU 189 A    -0.03 LYS  289 A
  GLU 189 A                                       0.00 XXX    0 X     0.00 XXX    0 X    -0.05 ARG  360 A
  GLU 189 A                                       0.00 XXX    0 X     0.00 XXX    0 X     0.03 ASP  299 A
  GLU 189 A                                       0.00 XXX    0 X     0.00 XXX    0 X    -0.43 ARG  350 A

  GLU 266 A    5.09     7 %   0.43  300   0.05   0 0.00 XXX    0 X     0.00 XXX    0 X     0.11 ASP  260 A

  GLU 272 A    4.24     8 %   0.62  304   0.05   0 -0.54 ARG  27 A     0.00 XXX    0 X    -0.40 ARG   27 A

  GLU 297 A    4.13    17 %   0.70  330   0.13   0 0.00 XXX    0 X    -0.64 GLY 292 A    -0.10 ARG   38 A
  GLU 297 A                                       0.00 XXX    0 X     0.00 XXX    0 X    -0.36 LYS  291 A
  GLU 297 A                                       0.00 XXX    0 X     0.00 XXX    0 X    -0.10 HIS  262 A

  GLU 315 A    4.73     0 %   0.12  142   0.00   0 0.00 XXX    0 X     0.00 XXX    0 X     0.11 ASP  310 A

  GLU 330 A    4.67    20 %   0.90  337   0.12   0 -0.14 GLN 485 A     0.00 XXX    0 X    -0.01 LYS  337 A
  GLU 330 A                                      -0.62 TYR 322 A       0.00 XXX    0 X    -0.23 ARG  461 A
  GLU 330 A                                       0.00 XXX    0 X     0.00 XXX    0 X     0.11 GLU  334 A
  GLU 330 A                                       0.00 XXX    0 X     0.00 XXX    0 X     0.05 GLU  357 A

  GLU 334 A    2.73    12 %   0.92  316   0.08   0 -0.66 TRP 491 A    -0.58 GLY 489 A    -0.13 LYS  337 A
  GLU 334 A                                      -0.91 ARG 461 A       0.00 XXX    0 X    -0.49 ARG  461 A

  GLU 351 A    4.75    60 %   2.02  448   0.00   0 -0.85 SER 188 A    -0.07 GLU 351 A    -0.06 LYS  289 A
  GLU 351 A                                      -0.33 ARG 350 A       0.00 XXX    0 X    -0.09 ARG  360 A
  GLU 351 A                                       0.00 XXX    0 X     0.00 XXX    0 X     0.24 GLU  189 A
  GLU 351 A                                       0.00 XXX    0 X     0.00 XXX    0 X     0.04 ASP  299 A
  GLU 351 A                                       0.00 XXX    0 X     0.00 XXX    0 X    -0.66 ARG  350 A

  GLU 357 A    4.56     0 %   0.26  275   0.00   0 0.00 XXX    0 X     0.00 XXX    0 X    -0.11 LYS  337 A
  GLU 357 A                                       0.00 XXX    0 X     0.00 XXX    0 X    -0.04 ARG  360 A
  GLU 357 A                                       0.00 XXX    0 X     0.00 XXX    0 X    -0.04 HIS  362 A

  GLU 373 A    8.47   100 %   3.73  588   0.84   0 0.00 XXX    0 X    -0.53 GLU 373 A     0.57 GLU  374 A
  GLU 373 A                                       0.00 XXX    0 X     0.00 XXX    0 X    -0.64 ARG  401 A

  GLU 374 A    5.56    60 %   1.43  448   0.37   0 0.00 XXX    0 X    -0.36 GLU 374 A    -0.37 ARG  401 A
```

128

```
GLU 413 A   4.63    0 %    0.14  179   0.00    0    0.00 XXX   0 X    0.00 XXX   0 X   -0.01 HIS 414 A

GLU 459 A   5.00   81 %    3.11  507   0.29    0   -0.48 ASN 534 A   -0.84 LYS 455 A    0.45 ASP 456 A
GLU 459 A                                          -0.45 LYS 455 A    0.00 XXX   0 X    0.10 GLU 500 A
GLU 459 A                                          -0.26 ARG 536 A    0.00 XXX   0 X   -0.74 LYS 455 A
GLU 459 A                                           0.00 XXX   0 X    0.00 XXX   0 X   -0.67 ARG 536 A

GLU 498 A   4.82    0 %    0.33  278   0.00    0    0.00 XXX   0 X    0.00 XXX   0 X   -0.24 LYS 495 A
GLU 498 A                                           0.00 XXX   0 X    0.00 XXX   0 X   -0.12 ARG 536 A
GLU 498 A                                           0.00 XXX   0 X    0.00 XXX   0 X    0.19 GLU 500 A
GLU 498 A                                           0.00 XXX   0 X    0.00 XXX   0 X    0.17 ASP 542 A

GLU 500 A   4.53   16 %    0.46  325   0.12    0    0.00 XXX   0 X    0.00 XXX   0 X   -0.18 LYS 455 A
GLU 500 A                                           0.00 XXX   0 X    0.00 XXX   0 X   -0.00 LYS 495 A
GLU 500 A                                           0.00 XXX   0 X    0.00 XXX   0 X   -0.43 ARG 536 A
GLU 500 A                                           0.00 XXX   0 X    0.00 XXX   0 X    0.06 ASP 542 A

GLU 525 A   3.59    0 %    0.30  205   0.00    0   -0.73 LYS 522 A    0.00 XXX   0 X   -0.10 ARG 615 A
GLU 525 A                                           0.00 XXX   0 X    0.00 XXX   0 X   -0.38 LYS 522 A

GLU 543 A   4.98   14 %    0.46  321   0.04    0    0.00 XXX   0 X    0.00 XXX   0 X   -0.12 ARG 547 A
GLU 543 A                                           0.00 XXX   0 X    0.00 XXX   0 X    0.04 ASP 542 A
GLU 543 A                                           0.00 XXX   0 X    0.00 XXX   0 X    0.08 GLU 544 A

GLU 544 A   3.95    0 %    0.21  189   0.00    0   -0.48 ARG 547 A    0.00 XXX   0 X   -0.27 ARG 547 A

GLU 548 A   4.64    0 %    0.17  240   0.00    0    0.00 XXX   0 X    0.00 XXX   0 X   -0.07 ARG 501 A
GLU 548 A                                           0.00 XXX   0 X    0.00 XXX   0 X   -0.05 ARG 552 A
GLU 548 A                                           0.00 XXX   0 X    0.00 XXX   0 X    0.03 ASP 506 A
GLU 548 A                                           0.00 XXX   0 X    0.00 XXX   0 X    0.04 ASP 542 A
GLU 548 A                                           0.00 XXX   0 X    0.00 XXX   0 X    0.03 GLU 544 A

GLU 560 A   6.95   96 %    3.34  550   1.19    0   -0.10 ASN 562 A   -0.76 PHE 568 A   -0.11 ARG 444 A
GLU 560 A                                           0.00 XXX   0 X   -0.84 GLY 569 A   -0.31 ARG 477 A
GLU 560 A                                           0.00 XXX   0 X    0.00 XXX   0 X    0.02 ASP 439 A
GLU 560 A                                           0.00 XXX   0 X    0.00 XXX   0 X    0.02 ASP 561 A

GLU 574 A   3.73    8 %    0.48  305   0.06    0   -0.82 ARG 444 A    0.00 XXX   0 X   -0.01 LYS 539 A
GLU 574 A                                           0.00 XXX   0 X    0.00 XXX   0 X   -0.48 ARG 444 A

GLU 593 A   4.70    0 %    0.14  184   0.00    0    0.00 XXX   0 X    0.00 XXX   0 X   -0.00 ARG 606 A
GLU 593 A                                           0.00 XXX   0 X    0.00 XXX   0 X    0.04 ASP 592 A
GLU 593 A                                           0.00 XXX   0 X    0.00 XXX   0 X    0.02 GLU 596 A

GLU 596 A   4.39   20 %    0.42  337   0.02    0    0.00 XXX   0 X    0.00 XXX   0 X   -0.37 LYS 111 A
GLU 596 A                                           0.00 XXX   0 X    0.00 XXX   0 X   -0.01 ARG 606 A
```

129

propka.ki.ku.dk/pka/2o1x.pka

```
  GLU 596 A                                      0.00 XXX   0 X    0.00 XXX   0 X    0.01 GLU 114 A
  GLU 596 A                                      0.00 XXX   0 X    0.00 XXX   0 X   -0.18 HIS 597 A

  GLU 601 A   4.18    19 %    0.50  334   0.00  0 0.00 XXX   0 X   -0.45 GLU 601 A   -0.05 ARG 480 A
  GLU 601 A                                      0.00 XXX   0 X    0.00 XXX   0 X   -0.16 LYS 515 A
  GLU 601 A                                      0.00 XXX   0 X    0.00 XXX   0 X   -0.15 HIS 604 A

  GLU 620 A   4.26     0 %    0.31  249   0.00  0 -0.22 ARG 586 A  0.00 XXX   0 X   -0.34 ARG 586 A

  HIS  17 A   6.53     0 %   -0.23  253   0.00  0 0.00 XXX   0 X    0.00 XXX   0 X   -0.00 ARG  24 A
  HIS  17 A                                      0.00 XXX   0 X    0.00 XXX   0 X   -0.02 ARG  94 A
  HIS  17 A                                      0.00 XXX   0 X    0.00 XXX   0 X    0.02 ASP  14 A
  HIS  17 A                                      0.00 XXX   0 X    0.00 XXX   0 X    0.20 ASP  21 A
  HIS  17 A                                      0.00 XXX   0 X    0.00 XXX   0 X    0.06 ASP  95 A

  HIS  51 A   1.06   100 %   -2.59  639   0.00  0 0.00 XXX   0 X    0.27 HIS 304 A   -1.49 LYS 101 A
  HIS  51 A                                      0.00 XXX   0 X    0.00 XXX   0 X   -0.10 LYS 289 A
  HIS  51 A                                      0.00 XXX   0 X    0.00 XXX   0 X   -1.11 HIS  82 A
  HIS  51 A                                      0.00 XXX   0 X    0.00 XXX   0 X   -1.26 HIS 304 A
  HIS  51 A                                      0.00 XXX   0 X    0.00 XXX   0 X   -0.27 HIS 434 A
  HIS  51 A                                      0.00 XXX   0 X    0.00 XXX   0 X    0.28 TDP O21 A
  HIS  51 A                                      0.00 XXX   0 X    0.00 XXX   0 X    0.84 ASP 430 A

  HIS  66 A   1.56   100 %   -2.88  603   0.00  0 0.00 XXX   0 X    0.00 XXX   0 X   -0.26 LYS  23 A
  HIS  66 A                                      0.00 XXX   0 X    0.00 XXX   0 X   -0.12 ARG  93 A
  HIS  66 A                                      0.00 XXX   0 X    0.00 XXX   0 X   -1.34 HIS  87 A
  HIS  66 A                                      0.00 XXX   0 X    0.00 XXX   0 X   -0.40 HIS 117 A
  HIS  66 A                                      0.00 XXX   0 X    0.00 XXX   0 X    0.08 ASP 118 A

  HIS  82 A   5.29   100 %   -3.20  664   0.00  0 0.56 TDP O11 A   0.00 XXX   0 X   -0.72 MG   MG A
  HIS  82 A                                      0.83 TDP O21 A    0.00 XXX   0 X   -0.06 LYS 101 A
  HIS  82 A                                      0.00 XXX   0 X    0.00 XXX   0 X   -0.41 LYS 289 A
  HIS  82 A                                      0.00 XXX   0 X    0.00 XXX   0 X    0.24 ASP  79 A
  HIS  82 A                                      0.00 XXX   0 X    0.00 XXX   0 X    1.56 TDP O21 A

  HIS  87 A   4.47   100 %   -3.14  615   0.00  0 1.60 ASP 118 A   0.02 HIS 117 A   -0.46 LYS  88 A
  HIS  87 A                                      0.00 XXX   0 X    0.00 XXX   0 X   -1.23 ARG  93 A
  HIS  87 A                                      0.00 XXX   0 X    0.00 XXX   0 X    0.10 GLU 114 A
  HIS  87 A                                      0.00 XXX   0 X    0.00 XXX   0 X   -0.38 HIS 117 A
  HIS  87 A                                      0.00 XXX   0 X    0.00 XXX   0 X    1.47 ASP 118 A

  HIS 117 A   5.29    49 %   -1.54  419   0.00  0 0.00 XXX   0 X    0.31 SER  71 A   -0.05 LYS  20 A
  HIS 117 A                                      0.00 XXX   0 X    0.00 XXX   0 X   -0.04 LYS  23 A
  HIS 117 A                                      0.00 XXX   0 X    0.00 XXX   0 X   -0.09 ARG  73 A
  HIS 117 A                                      0.00 XXX   0 X    0.00 XXX   0 X   -0.23 ARG  93 A
  HIS 117 A                                      0.00 XXX   0 X    0.00 XXX   0 X    0.02 ASP  70 A
```

130

```
HIS 117 A                                      0.00 XXX   0 X    0.00 XXX   0 X     0.25 GLU 116 A
HIS 117 A                                      0.00 XXX   0 X    0.00 XXX   0 X     0.15 ASP 118 A


HIS 124 A    2.62    92 %   -2.42   540   0.00    0   0.00 XXX   0 X    0.04 THR 433 A    -0.63 ARG 401 A
HIS 124 A                                      0.00 XXX   0 X    0.00 XXX   0 X    -0.86 HIS 434 A


HIS 147 A    8.30    44 %   -1.14   404   0.00    0   1.60 ASP  74 A    0.00 XXX   0 X    -0.10 ARG 139 A
HIS 147 A                                      0.51 ASP 276 A    0.00 XXX   0 X    -0.06 LYS 175 A
HIS 147 A                                      0.00 XXX   0 X    0.00 XXX   0 X     0.08 ASP  70 A
HIS 147 A                                      0.00 XXX   0 X    0.00 XXX   0 X     0.08 ASP 145 A
HIS 147 A                                      0.00 XXX   0 X    0.00 XXX   0 X     0.56 ASP  74 A
HIS 147 A                                      0.00 XXX   0 X    0.00 XXX   0 X     0.28 ASP 276 A


HIS 262 A    6.11    72 %   -2.27   484   0.00    0   1.60 ASP  60 A    0.01 GLY 261 A    -0.61 ARG  38 A
HIS 262 A                                      0.00 XXX   0 X    0.00 XXX   0 X    -0.26 LYS 291 A
HIS 262 A                                      0.00 XXX   0 X    0.00 XXX   0 X     0.26 ASP 260 A
HIS 262 A                                      0.00 XXX   0 X    0.00 XXX   0 X     0.10 GLU 297 A
HIS 262 A                                      0.00 XXX   0 X    0.00 XXX   0 X     0.77 ASP  60 A


HIS 284 A    6.81    59 %   -1.94   446   0.00    0   1.60 ASP 182 A    0.00 XXX   0 X     0.11 ASP 154 A
HIS 284 A                                      0.00 XXX   0 X    0.00 XXX   0 X     0.55 ASP 182 A


HIS 304 A    2.33   100 %   -2.82   619   0.00    0   0.17 TDP O21 A    0.00 XXX   0 X    -0.23 MG   MG A
HIS 304 A                                      0.00 XXX   0 X    0.00 XXX   0 X    -0.06 LYS 101 A
HIS 304 A                                      0.00 XXX   0 X    0.00 XXX   0 X    -1.64 LYS 289 A
HIS 304 A                                      0.00 XXX   0 X    0.00 XXX   0 X    -1.07 HIS  82 A
HIS 304 A                                      0.00 XXX   0 X    0.00 XXX   0 X     0.03 ASP 430 A
HIS 304 A                                      0.00 XXX   0 X    0.00 XXX   0 X     1.46 TDP O21 A


HIS 362 A    5.25    25 %   -1.11   352   0.00    0   0.00 XXX   0 X    0.00 XXX   0 X    -0.13 LYS 337 A
HIS 362 A                                      0.00 XXX   0 X    0.00 XXX   0 X    -0.05 ARG 365 A
HIS 362 A                                      0.00 XXX   0 X    0.00 XXX   0 X     0.04 GLU 357 A


HIS 364 A    6.26     0 %   -0.15   160   0.00    0   0.00 XXX   0 X    0.00 XXX   0 X    -0.09 ARG 365 A


HIS 408 A    7.42     0 %   -0.25   206   0.00    0   0.63 ASP 409 A    0.00 XXX   0 X     0.22 ASP 404 A
HIS 408 A                                      0.00 XXX   0 X    0.00 XXX   0 X     0.32 ASP 409 A


HIS 414 A    5.62    19 %   -0.67   335   0.00    0   0.00 XXX   0 X    0.00 XXX   0 X    -0.39 ARG 450 A
HIS 414 A                                      0.00 XXX   0 X    0.00 XXX   0 X    -0.01 LYS 539 A
HIS 414 A                                      0.00 XXX   0 X    0.00 XXX   0 X     0.01 GLU 413 A
HIS 414 A                                      0.00 XXX   0 X    0.00 XXX   0 X     0.18 ASP 471 A


HIS 434 A    4.72   100 %   -2.87   589   0.00    0   0.00 XXX   0 X    0.00 XXX   0 X    -0.49 LYS 101 A
HIS 434 A                                      0.00 XXX   0 X    0.00 XXX   0 X     1.57 ASP 430 A
```

```
 HIS 470 A    5.65    13 %    -0.83   318   0.00    0    0.00 XXX    0 X    0.00 XXX    0 X   -0.04 ARG 389 A
 HIS 470 A                                                0.00 XXX    0 X    0.00 XXX    0 X   -0.05 ARG 450 A
 HIS 470 A                                                0.00 XXX    0 X    0.00 XXX    0 X   -0.00 LYS 465 A
 HIS 470 A                                                0.00 XXX    0 X    0.00 XXX    0 X    0.06 ASP 471 A
 HIS 470 A                                                0.00 XXX    0 X    0.00 XXX    0 X    0.02 ASP 493 A

 HIS 582 A    6.30     0 %    -0.20   168   0.00    0    0.00 XXX    0 X    0.00 XXX    0 X    0.00 XXX    0 X

 HIS 597 A    5.69    48 %    -1.37   416   0.00    0    0.00 XXX    0 X    0.49 LYS 102 A   -0.13 LYS  88 A
 HIS 597 A                                                0.00 XXX    0 X    0.00 HIS 597 A   -0.12 LYS 101 A
 HIS 597 A                                                0.00 XXX    0 X    0.00 XXX    0 X   -0.10 LYS 111 A
 HIS 597 A                                                0.00 XXX    0 X    0.00 XXX    0 X    0.03 GLU 103 A
 HIS 597 A                                                0.00 XXX    0 X    0.00 XXX    0 X    0.18 GLU 114 A
 HIS 597 A                                                0.00 XXX    0 X    0.00 XXX    0 X    0.02 ASP 430 A
 HIS 597 A                                                0.00 XXX    0 X    0.00 XXX    0 X    0.18 GLU 596 A

 HIS 604 A    7.47    77 %    -2.46   496   0.00    0    1.60 ASP 561 A    0.00 XXX    0 X   -0.00 ARG 477 A
 HIS 604 A                                                0.00 XXX    0 X    0.00 XXX    0 X   -0.24 LYS 515 A
 HIS 604 A                                                0.00 XXX    0 X    0.00 XXX    0 X   -0.01 ARG 606 A
 HIS 604 A                                                0.00 XXX    0 X    0.00 XXX    0 X    0.15 GLU 601 A
 HIS 604 A                                                0.00 XXX    0 X    0.00 XXX    0 X    0.12 ASP 610 A
 HIS 604 A                                                0.00 XXX    0 X    0.00 XXX    0 X    1.82 ASP 561 A

 CYS  45 A   12.61   100 %     3.82   570   0.00    0    0.00 XXX    0 X    0.00 XXX    0 X   -0.15 ARG  47 A
 CYS  45 A                                                0.00 XXX    0 X    0.00 XXX    0 X   -0.06 LYS 101 A

 CYS 420 A   12.64   100 %     3.56   731   0.00    0    0.00 XXX    0 X    0.00 XXX    0 X    0.09 ASP 422 A

 TYR  67 A   10.55    16 %     0.83   325   0.00    0   -0.18 ARG  27 A    0.00 XXX    0 X   -0.31 ARG  27 A
 TYR  67 A                                                0.00 XXX    0 X    0.00 XXX    0 X    0.08 GLU  28 A
 TYR  67 A                                                0.00 XXX    0 X    0.00 XXX    0 X    0.13 GLU 272 A

 TYR  85 A   13.91   100 %     3.39   596   0.00    0    0.00 XXX    0 X   -0.81 HIS  51 A    1.31 CYS  45 A
 TYR  85 A                                                0.00 XXX    0 X    0.00 XXX    0 X   -0.18 ARG  47 A
 TYR  85 A                                                0.00 XXX    0 X    0.00 XXX    0 X    0.21 ASP 430 A

 TYR 255 A   11.49    28 %     1.40   359   0.00    0    0.00 XXX    0 X    0.00 XXX    0 X    0.09 ASP 182 A

 TYR 295 A   11.77    25 %     1.65   352   0.00    0    0.00 XXX    0 X    0.00 XXX    0 X    0.12 ASP 310 A

 TYR 302 A   10.96    23 %     1.44   347   0.00    0    0.00 XXX    0 X   -0.77 ALA 307 A    0.05 ASP 299 A
 TYR 302 A                                                0.00 XXX    0 X    0.00 XXX    0 X    0.24 TYR 316 A

 TYR 316 A   10.47     0 %     0.30   264   0.00    0    0.00 XXX    0 X    0.00 XXX    0 X    0.15 ASP 299 A
 TYR 316 A                                                0.00 XXX    0 X    0.00 XXX    0 X    0.01 GLU 315 A
```

```
TYR 322 A   12.45    30 %    1.50   365   0.00    0    0.62 GLU 330 A    0.00 XXX   0 X    0.03 GLU 334 A
TYR 322 A                                                0.00 XXX   0 X    0.00 XXX   0 X   -0.18 ARG 461 A
TYR 322 A                                                0.00 XXX   0 X    0.00 XXX   0 X    0.48 GLU 330 A

TYR 366 A   13.79    73 %    3.09   485   0.00    0    0.00 XXX   0 X    0.00 XXX   0 X    0.06 GLU 189 A
TYR 366 A                                                0.00 XXX   0 X    0.00 XXX   0 X   -0.44 ARG 350 A
TYR 366 A                                                0.00 XXX   0 X    0.00 XXX   0 X    0.25 GLU 351 A
TYR 366 A                                                0.00 XXX   0 X    0.00 XXX   0 X    0.84 ASP 368 A

TYR 395 A   11.94   100 %    2.90   585   0.00    0    0.00 XXX   0 X    0.00 XXX   0 X   -2.03 ARG 423 A
TYR 395 A                                                0.00 XXX   0 X    0.00 XXX   0 X    0.20 ASP 430 A
TYR 395 A                                                0.00 XXX   0 X    0.00 XXX   0 X   -0.20 ARG 480 A
TYR 395 A                                                0.00 XXX   0 X    0.00 XXX   0 X    1.08 ASP 422 A

TYR 403 A   10.35     0 %    0.31   276   0.00    0    0.00 XXX   0 X    0.00 XXX   0 X    0.05 ASP 404 A

TYR 466 A   10.65    21 %    0.76   339   0.00    0    0.00 XXX   0 X    0.00 XXX   0 X   -0.11 ARG 450 A
TYR 466 A                                                0.00 XXX   0 X    0.00 XXX   0 X    0.01 ASP 471 A
TYR 466 A                                                0.00 XXX   0 X    0.00 XXX   0 X    0.01 GLU 498 A
TYR 466 A                                                0.00 XXX   0 X    0.00 XXX   0 X   -0.02 LYS 495 A

TYR 478 A   16.85   100 %    3.92   663   0.00    0    0.00 XXX   0 X    0.00 XXX   0 X    1.47 CYS 420 A
TYR 478 A                                                0.00 XXX   0 X    0.00 XXX   0 X   -0.17 ARG 423 A
TYR 478 A                                                0.00 XXX   0 X    0.00 XXX   0 X    0.08 TYR 395 A
TYR 478 A                                                0.00 XXX   0 X    0.00 XXX   0 X    1.55 ASP 422 A

TYR 519 A   12.47    48 %    1.57   416   0.00    0    0.00 XXX   0 X    0.00 XXX   0 X    0.12 ASP 518 A
TYR 519 A                                                0.00 XXX   0 X    0.00 XXX   0 X    0.30 ASP 561 A
TYR 519 A                                                0.00 XXX   0 X    0.00 XXX   0 X    0.26 GLU 601 A
TYR 519 A                                                0.00 XXX   0 X    0.00 XXX   0 X    0.22 ASP 610 A

LYS  20 A   10.07     0 %   -0.22   208   0.00    0    0.00 XXX   0 X    0.00 XXX   0 X    0.04 ASP  21 A
LYS  20 A                                                0.00 XXX   0 X    0.00 XXX   0 X   -0.03 ARG  24 A
LYS  20 A                                                0.00 XXX   0 X    0.00 XXX   0 X    0.09 ASP  70 A
LYS  20 A                                                0.00 XXX   0 X    0.00 XXX   0 X   -0.38 ARG  73 A
LYS  20 A                                                0.00 XXX   0 X    0.00 XXX   0 X    0.07 GLU 116 A

LYS  23 A    9.06    49 %   -1.65   419   0.00    0    0.00 XXX   0 X    0.00 XXX   0 X    0.01 ASP  21 A
LYS  23 A                                                0.00 XXX   0 X    0.00 XXX   0 X   -0.01 ARG  24 A
LYS  23 A                                                0.00 XXX   0 X    0.00 XXX   0 X    0.28 ASP  70 A
LYS  23 A                                                0.00 XXX   0 X    0.00 XXX   0 X   -0.07 ARG  73 A
LYS  23 A                                                0.00 XXX   0 X    0.00 XXX   0 X    0.07 ASP  74 A
LYS  23 A                                                0.00 XXX   0 X    0.00 XXX   0 X   -0.07 LYS  20 A

LYS  88 A    8.79    88 %   -3.57   529   0.00    0    0.85 GLU 114 A    0.00 XXX   0 X   -0.94 ARG  93 A
LYS  88 A                                                0.00 XXX   0 X    0.00 XXX   0 X   -0.02 LYS 101 A
```

133

```
LYS  88 A                                   0.00 XXX   0 X    0.00 XXX   0 X   -0.04 LYS 111 A
LYS  88 A                                   0.00 XXX   0 X    0.00 XXX   0 X    1.01 GLU 114 A
LYS  88 A                                   0.00 XXX   0 X    0.00 XXX   0 X    0.99 ASP 118 A

LYS 101 A  10.47  100 %  -2.85  572  0.00   0   0.85 ASP 430 A    0.00 XXX   0 X    0.06 CYS  45 A
LYS 101 A                                   0.00 XXX   0 X    0.00 XXX   0 X   -0.06 ARG  47 A
LYS 101 A                                   0.00 XXX   0 X    0.00 XXX   0 X   -0.07 ARG 480 A
LYS 101 A                                   0.00 XXX   0 X    0.00 XXX   0 X    2.03 ASP 430 A

LYS 102 A  10.91    0 %  -0.44  258  0.00   0   0.85 ASP  99 A    0.00 XXX   0 X   -0.38 ARG  47 A
LYS 102 A                                   0.00 XXX   0 X    0.00 XXX   0 X    0.38 ASP  99 A

LYS 111 A  10.37   25 %  -0.60  352  0.00   0   0.00 XXX   0 X    0.00 XXX   0 X    0.09 GLU 114 A
LYS 111 A                                   0.00 XXX   0 X    0.00 XXX   0 X    0.01 ASP 118 A
LYS 111 A                                   0.00 XXX   0 X    0.00 XXX   0 X    0.37 GLU 596 A

LYS 144 A  10.30    0 %  -0.20  177  0.00   0   0.00 XXX   0 X    0.00 XXX   0 X   -0.00 ARG  75 A

LYS 175 A  10.40    2 %  -0.31  288  0.00   0   0.00 XXX   0 X    0.00 XXX   0 X   -0.30 ARG 139 A
LYS 175 A                                   0.00 XXX   0 X    0.00 XXX   0 X    0.25 ASP 140 A
LYS 175 A                                   0.00 XXX   0 X    0.00 XXX   0 X    0.11 ASP 145 A
LYS 175 A                                   0.00 XXX   0 X    0.00 XXX   0 X   -0.03 ARG 174 A
LYS 175 A                                   0.00 XXX   0 X    0.00 XXX   0 X    0.02 ASP 276 A
LYS 175 A                                   0.00 XXX   0 X    0.00 XXX   0 X    0.15 ASP 278 A

LYS 196 A  10.39    0 %  -0.11  166  0.00   0   0.00 XXX   0 X    0.00 XXX   0 X    0.00 XXX   0 X

LYS 289 A   8.34   98 %  -3.61  556  0.00   0   0.20 TDP O21 A    0.00 XXX   0 X   -1.34 MG   MG A
LYS 289 A                                   0.00 XXX   0 X    0.00 XXX   0 X    0.19 ASP 154 A
LYS 289 A                                   0.00 XXX   0 X    0.00 XXX   0 X    0.03 GLU 189 A
LYS 289 A                                   0.00 XXX   0 X    0.00 XXX   0 X    0.04 ASP 299 A
LYS 289 A                                   0.00 XXX   0 X    0.00 XXX   0 X    0.06 GLU 351 A
LYS 289 A                                   0.00 XXX   0 X    0.00 XXX   0 X    0.25 TDP O13 A
LYS 289 A                                   0.00 XXX   0 X    0.00 XXX   0 X    2.03 TDP O21 A

LYS 291 A  10.53    7 %  -0.30  302  0.00   0   0.00 XXX   0 X    0.00 XXX   0 X    0.00 GLU  35 A
LYS 291 A                                   0.00 XXX   0 X    0.00 XXX   0 X   -0.19 ARG  38 A
LYS 291 A                                   0.00 XXX   0 X    0.00 XXX   0 X    0.10 ASP  60 A
LYS 291 A                                   0.00 XXX   0 X    0.00 XXX   0 X    0.06 ASP 260 A
LYS 291 A                                   0.00 XXX   0 X    0.00 XXX   0 X    0.36 GLU 297 A

LYS 308 A  10.37    0 %  -0.13  195  0.00   0   0.00 XXX   0 X    0.00 XXX   0 X    0.00 ASP 310 A

LYS 337 A  10.47    0 %  -0.28  280  0.00   0   0.00 XXX   0 X    0.00 XXX   0 X    0.01 GLU 330 A
LYS 337 A                                   0.00 XXX   0 X    0.00 XXX   0 X    0.13 GLU 334 A
LYS 337 A                                   0.00 XXX   0 X    0.00 XXX   0 X    0.11 GLU 357 A
```

134

```
   LYS 337 A                                            0.00 XXX   0 X    0.00 XXX   0 X   -0.01 ARG 461 A

   LYS 455 A   10.52    41 %   -1.28   396    0.00    0   0.45 GLU 459 A    0.00 XXX   0 X    0.27 ASP 456 A
   LYS 455 A                                            0.00 XXX   0 X    0.00 XXX   0 X    0.18 GLU 500 A
   LYS 455 A                                            0.00 XXX   0 X    0.00 XXX   0 X   -0.34 ARG 536 A
   LYS 455 A                                            0.00 XXX   0 X    0.00 XXX   0 X    0.74 GLU 459 A

   LYS 465 A   11.18     4 %   -0.55   292    0.00    0   0.85 ASP 493 A    0.00 XXX   0 X    0.38 ASP 493 A

   LYS 495 A   10.68     0 %   -0.09   148    0.00    0   0.00 XXX   0 X    0.00 XXX   0 X    0.24 GLU 498 A
   LYS 495 A                                            0.00 XXX   0 X    0.00 XXX   0 X    0.00 GLU 500 A
   LYS 495 A                                            0.00 XXX   0 X    0.00 XXX   0 X    0.02 TYR 466 A

   LYS 503 A    9.86     9 %   -0.64   306    0.00    0   0.00 XXX   0 X    0.00 XXX   0 X   -0.03 ARG 501 A
   LYS 503 A                                            0.00 XXX   0 X    0.00 XXX   0 X    0.03 ASP 526 A

   LYS 515 A   10.07    68 %   -1.36   471    0.00    0   0.00 XXX   0 X    0.00 XXX   0 X   -0.05 ARG 480 A
   LYS 515 A                                            0.00 XXX   0 X    0.00 XXX   0 X    0.38 ASP 518 A
   LYS 515 A                                            0.00 XXX   0 X    0.00 XXX   0 X    0.43 ASP 561 A
   LYS 515 A                                            0.00 XXX   0 X    0.00 XXX   0 X    0.16 GLU 601 A

   LYS 522 A   11.34     0 %   -0.24   201    0.00    0   0.73 GLU 525 A    0.00 XXX   0 X    0.04 ASP 526 A
   LYS 522 A                                            0.00 XXX   0 X    0.00 XXX   0 X   -0.06 ARG 615 A
   LYS 522 A                                            0.00 XXX   0 X    0.00 XXX   0 X    0.38 GLU 525 A

   LYS 539 A    9.22    41 %   -1.12   396    0.00    0   0.00 XXX   0 X    0.00 XXX   0 X   -0.10 ARG 444 A
   LYS 539 A                                            0.00 XXX   0 X    0.00 XXX   0 X   -0.07 ARG 450 A
   LYS 539 A                                            0.00 XXX   0 X    0.00 XXX   0 X    0.01 GLU 574 A

   ARG  24 A   12.77     0 %   -0.22   193    0.00    0   0.22 ASP  21 A    0.00 XXX   0 X    0.28 ASP  21 A

   ARG  27 A   13.63     0 %   -0.39   282    0.00    0   0.18 TYR  67 A    0.00 XXX   0 X    0.10 GLU  28 A
   ARG  27 A                                            0.54 GLU 272 A    0.00 XXX   0 X    0.31 TYR  67 A
   ARG  27 A                                            0.00 XXX   0 X    0.00 XXX   0 X    0.40 GLU 272 A

   ARG  38 A   12.94    79 %   -2.43   503    0.00    0   1.58 ASP  60 A    0.00 XXX   0 X    0.06 GLU  35 A
   ARG  38 A                                            0.00 XXX   0 X    0.00 XXX   0 X    0.01 ASP 260 A
   ARG  38 A                                            0.00 XXX   0 X    0.00 XXX   0 X    0.10 GLU 297 A
   ARG  38 A                                            0.00 XXX   0 X    0.00 XXX   0 X    1.13 ASP  60 A

   ARG  43 A   12.70    25 %   -0.71   352    0.00    0   0.54 GLU  40 A    0.00 XXX   0 X    0.05 ASP   9 A
   ARG  43 A                                            0.00 XXX   0 X    0.00 XXX   0 X    0.02 ASP  14 A
   ARG  43 A                                            0.00 XXX   0 X    0.00 XXX   0 X    0.03 GLU  36 A
   ARG  43 A                                            0.00 XXX   0 X    0.00 XXX   0 X   -0.13 ARG  94 A
   ARG  43 A                                            0.00 XXX   0 X    0.00 XXX   0 X    0.40 GLU  40 A
```

135

```
ARG   47 A  11.71    35 %   -1.24  378   0.00    0   0.00 XXX   0 X    0.00 XXX   0 X    0.15 CYS   45 A
ARG   47 A                                            0.00 XXX   0 X    0.00 XXX   0 X    0.18 TYR   85 A
ARG   47 A                                            0.00 XXX   0 X    0.00 XXX   0 X    0.13 ASP   99 A

ARG   73 A  12.97     0 %   -0.34  246   0.00    0   0.36 ASP  70 A    0.00 XXX   0 X    0.05 ASP   74 A
ARG   73 A                                            0.00 XXX   0 X    0.00 XXX   0 X    0.13 GLU  116 A
ARG   73 A                                            0.00 XXX   0 X    0.00 XXX   0 X    0.27 ASP   70 A

ARG   75 A  11.65    41 %   -0.85  395   0.00    0   0.00 XXX   0 X    0.00 XXX   0 X    0.01 GLU  116 A

ARG   93 A  13.33    80 %   -2.92  506   0.00    0   0.49 GLU 114 A    0.00 XXX   0 X    0.66 GLU  114 A
ARG   93 A                                            1.47 ASP 118 A    0.00 XXX   0 X    1.13 ASP  118 A

ARG   94 A  13.65    31 %   -1.01  368   0.00    0   0.32 ASP  14 A    0.00 XXX   0 X    0.11 GLU   36 A
ARG   94 A                                            0.69 GLU  40 A    0.00 XXX   0 X    0.17 ASP   95 A
ARG   94 A                                            0.00 XXX   0 X    0.00 XXX   0 X    0.27 ASP   14 A
ARG   94 A                                            0.00 XXX   0 X    0.00 XXX   0 X    0.60 GLU   40 A

ARG  139 A  12.42    36 %   -1.35  381   0.00    0   0.84 ASP 140 A    0.00 XXX   0 X    0.03 ASP   74 A
ARG  139 A                                            0.00 XXX   0 X    0.00 XXX   0 X    0.04 ASP  145 A
ARG  139 A                                            0.00 XXX   0 X    0.00 XXX   0 X   -0.05 ARG  174 A
ARG  139 A                                            0.00 XXX   0 X    0.00 XXX   0 X    0.41 ASP  140 A

ARG  174 A  13.16     0 %   -0.35  247   0.00    0   0.74 ASP 140 A    0.00 XXX   0 X    0.01 ASP  171 A
ARG  174 A                                            0.00 XXX   0 X    0.00 XXX   0 X    0.26 ASP  140 A

ARG  254 A  12.00    10 %   -0.56  309   0.00    0   0.00 XXX   0 X    0.00 XXX   0 X    0.06 ASP  278 A

ARG  273 A  12.32     0 %   -0.18  160   0.00    0   0.00 XXX   0 X    0.00 XXX   0 X    0.00 XXX    0 X

ARG  341 A  12.38     0 %   -0.33  278   0.00    0   0.00 XXX   0 X    0.00 XXX   0 X    0.21 ASP  339 A

ARG  350 A  13.63    43 %   -1.26  403   0.00    0   0.33 GLU 189 A    0.00 XXX   0 X    0.44 TYR  366 A
ARG  350 A                                            0.33 GLU 351 A    0.00 XXX   0 X    0.19 ASP  368 A
ARG  350 A                                            0.00 XXX   0 X    0.00 XXX   0 X    0.43 GLU  189 A
ARG  350 A                                            0.00 XXX   0 X    0.00 XXX   0 X    0.66 GLU  351 A

ARG  360 A  12.32     0 %   -0.22  225   0.00    0   0.00 XXX   0 X    0.00 XXX   0 X    0.05 GLU  189 A
ARG  360 A                                            0.00 XXX   0 X    0.00 XXX   0 X    0.09 GLU  351 A
ARG  360 A                                            0.00 XXX   0 X    0.00 XXX   0 X    0.04 GLU  357 A
ARG  360 A                                            0.00 XXX   0 X    0.00 XXX   0 X   -0.14 ARG  350 A

ARG  365 A  12.24     0 %   -0.29  271   0.00    0   0.00 XXX   0 X    0.00 XXX   0 X    0.04 ASP  339 A

ARG  389 A  11.54    23 %   -0.76  347   0.00    0   0.00 XXX   0 X    0.00 XXX   0 X    0.12 ASP  339 A
ARG  389 A                                            0.00 XXX   0 X    0.00 XXX   0 X    0.08 ASP  471 A
```

```
ARG 389 A                                       0.00 XXX   0 X    0.00 XXX   0 X   -0.40 ARG 341 A

ARG 401 A   11.96    72 %   -1.77  482   0.00    0   0.00 XXX   0 X    0.00 XXX   0 X    0.22 ASP  79 A
ARG 401 A                                       0.00 XXX   0 X    0.00 XXX   0 X    0.37 GLU 374 A
ARG 401 A                                       0.00 XXX   0 X    0.00 XXX   0 X    0.64 GLU 373 A

ARG 423 A   14.85    93 %   -2.40  541   0.00    0   0.62 ASP 422 A    0.00 XXX   0 X    2.03 TYR 395 A
ARG 423 A                                       0.00 XXX   0 X    0.00 XXX   0 X    0.16 ASP 430 A
ARG 423 A                                       0.00 XXX   0 X    0.00 XXX   0 X    0.17 TYR 478 A
ARG 423 A                                       0.00 XXX   0 X    0.00 XXX   0 X    1.77 ASP 422 A

ARG 444 A   12.71    41 %   -1.21  397   0.00    0   0.82 GLU 574 A    0.00 XXX   0 X    0.11 GLU 560 A
ARG 444 A                                       0.00 XXX   0 X    0.00 XXX   0 X    0.48 GLU 574 A

ARG 450 A   10.85    51 %   -1.86  425   0.00    0   0.00 XXX   0 X    0.00 XXX   0 X    0.11 TYR 466 A
ARG 450 A                                       0.00 XXX   0 X    0.00 XXX   0 X    0.10 ASP 471 A

ARG 461 A   12.72    42 %   -1.59  399   0.00    0   0.91 GLU 334 A    0.00 XXX   0 X    0.18 TYR 322 A
ARG 461 A                                       0.00 XXX   0 X    0.00 XXX   0 X    0.23 GLU 330 A
ARG 461 A                                       0.00 XXX   0 X    0.00 XXX   0 X    0.49 GLU 334 A

ARG 477 A   12.67   100 %   -3.47  679   0.00    0   0.42 ASP 439 A    0.00 XXX   0 X    0.31 GLU 560 A
ARG 477 A                                       0.00 XXX   0 X    0.00 XXX   0 X    2.03 ASP 439 A
ARG 477 A                                       0.00 XXX   0 X    0.00 XXX   0 X    0.88 ASP 561 A

ARG 480 A   11.77    71 %   -0.83  481   0.00    0   0.00 XXX   0 X    0.00 XXX   0 X    0.20 TYR 395 A
ARG 480 A                                       0.00 XXX   0 X    0.00 XXX   0 X    0.05 ASP 422 A
ARG 480 A                                       0.00 XXX   0 X    0.00 XXX   0 X    0.32 ASP 430 A
ARG 480 A                                       0.00 XXX   0 X    0.00 XXX   0 X    0.05 GLU 601 A
ARG 480 A                                       0.00 XXX   0 X    0.00 XXX   0 X   -0.52 ARG 423 A

ARG 501 A   11.92    28 %   -0.84  361   0.00    0   0.00 XXX   0 X    0.00 XXX   0 X    0.20 ASP 506 A
ARG 501 A                                       0.00 XXX   0 X    0.00 XXX   0 X    0.07 GLU 548 A
ARG 501 A                                       0.00 XXX   0 X    0.00 XXX   0 X   -0.02 ARG 552 A

ARG 536 A   12.28    53 %   -1.78  430   0.00    0   0.26 GLU 459 A    0.00 XXX   0 X    0.03 ASP 456 A
ARG 536 A                                       0.00 XXX   0 X    0.00 XXX   0 X    0.12 GLU 498 A
ARG 536 A                                       0.00 XXX   0 X    0.00 XXX   0 X    0.43 GLU 500 A
ARG 536 A                                       0.00 XXX   0 X    0.00 XXX   0 X    0.05 ASP 542 A
ARG 536 A                                       0.00 XXX   0 X    0.00 XXX   0 X    0.67 GLU 459 A

ARG 547 A   13.11     0 %   -0.27  204   0.00    0   0.48 GLU 544 A    0.00 XXX   0 X    0.12 GLU 543 A
ARG 547 A                                       0.00 XXX   0 X    0.00 XXX   0 X    0.27 GLU 544 A

ARG 552 A   12.72     0 %   -0.10  150   0.00    0   0.00 XXX   0 X    0.00 XXX   0 X    0.28 ASP 506 A
ARG 552 A                                       0.00 XXX   0 X    0.00 XXX   0 X    0.05 GLU 548 A
```

137

```
ARG 554 A   12.63     0 %    -0.34   264   0.00    0    0.09 ASP 507 A    0.00 XXX   0 X    0.38 ASP 507 A

ARG 586 A   12.70     0 %    -0.36   281   0.00    0    0.22 GLU 620 A    0.00 XXX   0 X    0.34 GLU 620 A

ARG 606 A   12.40     3 %    -0.44   291   0.00    0    0.00 XXX   0 X    0.00 XXX   0 X    0.33 GLU 103 A
ARG 606 A                                                0.00 XXX   0 X    0.00 XXX   0 X    0.00 GLU 593 A
ARG 606 A                                                0.00 XXX   0 X    0.00 XXX   0 X    0.01 GLU 596 A

ARG 615 A   12.25     0 %    -0.45   273   0.00    0    0.00 XXX   0 X    0.00 XXX   0 X    0.10 GLU 525 A
ARG 615 A                                                0.00 XXX   0 X    0.00 XXX   0 X    0.11 ASP 526 A

N+    5 A    7.94     0 %    -0.07    70   0.00    0    0.00 XXX   0 X    0.00 XXX   0 X    0.00 ASP  14 A

TDP N1' A   -0.58   100 %    -3.93   581   0.00    0    0.00 XXX   0 X    0.00 XXX   0 X   -0.48 ARG 401 A
TDP N1' A                                                0.00 XXX   0 X    0.00 XXX   0 X   -0.78 HIS 124 A
TDP N1' A                                                0.00 XXX   0 X    0.00 XXX   0 X   -0.38 HIS 434 A

TDP O13 A    8.30*  100 %     3.41   638   0.00    0   -0.52 SER 156 A    0.00 XXX   0 X   -3.48 MG   MG A
TDP O13 A                                                0.00 XXX   0 X    0.00 XXX   0 X   -0.25 LYS 289 A
TDP O13 A                                                0.00 XXX   0 X    0.00 XXX   0 X    0.03 ASP 182 A
TDP O13 A                                                0.00 XXX   0 X    0.00 XXX   0 X    0.85 ASP  79 A
TDP O13 A                                                0.00 XXX   0 X    0.00 XXX   0 X    0.90 ASP 154 A
TDP O13 A                                                0.00 XXX   0 X    0.00 XXX   0 X    1.36 TDP O21 A

TDP O21 A   -0.86*  100 %     3.86   613   0.00    0   -0.85 SER  54 A    0.00 XXX   0 X   -3.35 MG   MG A
TDP O21 A                                               -0.83 HIS  82 A    0.00 XXX   0 X   -0.28 HIS  51 A
TDP O21 A                                               -0.20 LYS 289 A    0.00 XXX   0 X   -1.56 HIS  82 A
TDP O21 A                                               -0.17 HIS 304 A    0.00 XXX   0 X   -2.03 LYS 289 A
TDP O21 A                                                0.00 XXX   0 X    0.00 XXX   0 X   -1.46 HIS 304 A
```

Coupled residues (marked *) were detected. Please rerun PropKa with the --display-coupled-residues
or -d option for detailed information.

SUMMARY OF THIS PREDICTION
      Group   pKa   model-pKa   ligand atom-type
    ASP   9 A   3.88   3.80
    ASP 14 A   2.41   3.80
    ASP 21 A   2.72   3.80
    ASP 60 A   2.08   3.80
    ASP 70 A   4.70   3.80
    ASP 74 A   1.39   3.80
    ASP 79 A   4.78   3.80
    ASP 95 A   3.85   3.80
    ASP 99 A   2.75   3.80
    ASP 118 A   0.14   3.80

138

```
ASP 140 A      2.00        3.80
ASP 145 A      3.69        3.80
ASP 154 A      4.87        3.80
ASP 171 A      3.96        3.80
ASP 182 A      1.80        3.80
ASP 260 A      3.25        3.80
ASP 276 A      3.36        3.80
ASP 278 A      3.75        3.80
ASP 299 A      3.71        3.80
ASP 310 A      3.71        3.80
ASP 339 A      3.64        3.80
ASP 368 A      5.57        3.80
ASP 404 A      4.06        3.80
ASP 409 A      2.62        3.80
ASP 422 A      4.51        3.80
ASP 430 A      0.37        3.80
ASP 439 A      5.32        3.80
ASP 456 A      4.07        3.80
ASP 471 A      3.58        3.80
ASP 493 A      2.83        3.80
ASP 506 A      3.37        3.80
ASP 507 A      3.96        3.80
ASP 518 A      5.43        3.80
ASP 526 A      3.97        3.80
ASP 542 A      4.02        3.80
ASP 561 A      2.66        3.80
ASP 592 A      4.02        3.80
ASP 610 A      4.24        3.80
ASP 624 A      3.93        3.80
GLU  28 A      4.41        4.50
GLU  35 A      4.59        4.50
GLU  36 A      4.95        4.50
GLU  40 A      4.20        4.50
GLU 103 A      3.66        4.50
GLU 114 A      4.03        4.50
GLU 116 A      4.36        4.50
GLU 184 A      5.15        4.50
GLU 189 A      3.74        4.50
GLU 266 A      5.09        4.50
GLU 272 A      4.24        4.50
GLU 297 A      4.13        4.50
GLU 315 A      4.73        4.50
GLU 330 A      4.67        4.50
GLU 334 A      2.73        4.50
GLU 351 A      4.75        4.50
GLU 357 A      4.56        4.50
```

139

```
GLU  373  A      8.47         4.50
GLU  374  A      5.56         4.50
GLU  413  A      4.63         4.50
GLU  459  A      5.00         4.50
GLU  498  A      4.82         4.50
GLU  500  A      4.53         4.50
GLU  525  A      3.59         4.50
GLU  543  A      4.98         4.50
GLU  544  A      3.95         4.50
GLU  548  A      4.64         4.50
GLU  560  A      6.95         4.50
GLU  574  A      3.73         4.50
GLU  593  A      4.70         4.50
GLU  596  A      4.39         4.50
GLU  601  A      4.18         4.50
GLU  620  A      4.26         4.50
HIS   17  A      6.53         6.50
HIS   51  A      1.06         6.50
HIS   66  A      1.56         6.50
HIS   82  A      5.29         6.50
HIS   87  A      4.47         6.50
HIS  117  A      5.29         6.50
HIS  124  A      2.62         6.50
HIS  147  A      8.30         6.50
HIS  262  A      6.11         6.50
HIS  284  A      6.81         6.50
HIS  304  A      2.33         6.50
HIS  362  A      5.25         6.50
HIS  364  A      6.26         6.50
HIS  408  A      7.42         6.50
HIS  414  A      5.62         6.50
HIS  434  A      4.72         6.50
HIS  470  A      5.65         6.50
HIS  582  A      6.30         6.50
HIS  597  A      5.69         6.50
HIS  604  A      7.47         6.50
CYS   45  A     12.61         9.00
CYS  420  A     12.64         9.00
TYR   67  A     10.55        10.00
TYR   85  A     13.91        10.00
TYR  255  A     11.49        10.00
TYR  295  A     11.77        10.00
TYR  302  A     10.96        10.00
TYR  316  A     10.47        10.00
TYR  322  A     12.45        10.00
TYR  366  A     13.79        10.00
```

```
     TYR 395 A     11.94          10.00
     TYR 403 A     10.35          10.00
     TYR 466 A     10.65          10.00
     TYR 478 A     16.85          10.00
     TYR 519 A     12.47          10.00
     LYS  20 A     10.07          10.50
     LYS  23 A      9.06          10.50
     LYS  88 A      8.79          10.50
     LYS 101 A     10.47          10.50
     LYS 102 A     10.91          10.50
     LYS 111 A     10.37          10.50
     LYS 144 A     10.30          10.50
     LYS 175 A     10.40          10.50
     LYS 196 A     10.39          10.50
     LYS 289 A      8.34          10.50
     LYS 291 A     10.53          10.50
     LYS 308 A     10.37          10.50
     LYS 337 A     10.47          10.50
     LYS 455 A     10.52          10.50
     LYS 465 A     11.18          10.50
     LYS 495 A     10.68          10.50
     LYS 503 A      9.86          10.50
     LYS 515 A     10.07          10.50
     LYS 522 A     11.34          10.50
     LYS 539 A      9.22          10.50
     ARG  24 A     12.77          12.50
     ARG  27 A     13.63          12.50
     ARG  38 A     12.94          12.50
     ARG  43 A     12.70          12.50
     ARG  47 A     11.71          12.50
     ARG  73 A     12.97          12.50
     ARG  75 A     11.65          12.50
     ARG  93 A     13.33          12.50
     ARG  94 A     13.65          12.50
     ARG 139 A     12.42          12.50
     ARG 174 A     13.16          12.50
     ARG 254 A     12.00          12.50
     ARG 273 A     12.32          12.50
     ARG 341 A     12.38          12.50
     ARG 350 A     13.63          12.50
     ARG 360 A     12.32          12.50
     ARG 365 A     12.24          12.50
     ARG 389 A     11.54          12.50
     ARG 401 A     11.96          12.50
     ARG 423 A     14.85          12.50
     ARG 444 A     12.71          12.50
```

```
     ARG 450 A     10.85       12.50
     ARG 461 A     12.72       12.50
     ARG 477 A     12.67       12.50
     ARG 480 A     11.77       12.50
     ARG 501 A     11.92       12.50
     ARG 536 A     12.28       12.50
     ARG 547 A     13.11       12.50
     ARG 552 A     12.72       12.50
     ARG 554 A     12.63       12.50
     ARG 586 A     12.70       12.50
     ARG 606 A     12.40       12.50
     ARG 615 A     12.25       12.50
     N+    5 A      7.94        8.00
     TDP N1' A     -0.58        5.00               NAR
     TDP O13 A      8.30        6.00               OP
     TDP O21 A     -0.86        6.00               OP
--------------------------------------------------------------------------------------------
--------------------------------------------------------------------------------------------
Free energy of   folding (kcal/mol) as a function of pH (using neutral reference)
     0.00    121.58
     1.00    114.56
     2.00    102.32
     3.00     82.92
     4.00     63.15
     5.00     54.15
     6.00     47.75
     7.00     46.01
     8.00     50.01
     9.00     52.49
    10.00     55.52
    11.00     64.72
    12.00     74.13
    13.00     82.54
    14.00     90.18


The pH of optimum stability is  6.7 for which the free energy is  45.5 kcal/mol at 298K
Could not determine pH values where the free energy is within 80 % of minimum
Could not determine the pH-range where the free energy is negative


Protein charge of folded and unfolded state as a function of pH
    pH  unfolded  folded
     0.00    75.99   72.32
     1.00    75.93   69.11
     2.00    75.28   63.76
     3.00    69.62   53.32
     4.00    43.87   33.19
```

142

```
    5.00        12.26      7.93
    6.00        -3.38     -7.79
    7.00       -16.81    -14.82
    8.00       -22.21    -19.43
    9.00       -25.65    -24.47
   10.00       -36.21    -31.88
   11.00       -53.00    -44.94
   12.00       -65.18    -59.50
   13.00       -83.00    -76.21
   14.00       -89.98    -86.04
 The pI is  6.00 (folded) and  6.00 (unfolded)
```

143

Below will be found the ProPKA3.1 results for monomer B of DXS:

```
propka3.1                                                          2014-09-29
-----------------------------------------------------------------------------
--                                                                         --
--                  PROPKA: A PROTEIN PKA PREDICTOR                        --
--                                                                         --
--                  VERSION 1.0,  04/25/2004, IOWA CITY                    --
--                             BY HUI LI                                   --
--                                                                         --
--                  VERSION 2.0,  11/05/2007, IOWA CITY/COPENHAGEN         --
--                    BY DELPHINE C. BAS AND DAVID M. ROGERS               --
--                                                                         --
--                  VERSION 3.0,  01/06/2011, COPENHAGEN                   --
--              BY MATS H.M. OLSSON AND CHRESTEN R. SONDERGARD             --
--                                                                         --
--                  VERSION 3.1,  07/01/2011, COPENHAGEN                   --
--              BY CHRESTEN R. SONDERGARD AND MATS H.M. OLSSON             --
-----------------------------------------------------------------------------


-----------------------------------------------------------------------------
 References:

   Very Fast Empirical Prediction and Rationalization of Protein pKa Values
   Hui Li, Andrew D. Robertson and Jan H. Jensen
   PROTEINS: Structure, Function, and Bioinformatics 61:704-721 (2005)

   Very Fast Prediction and Rationalization of pKa Values for Protein-Ligand Complexes
   Delphine C. Bas, David M. Rogers and Jan H. Jensen
   PROTEINS: Structure, Function, and Bioinformatics 73:765-783 (2008)

   PROPKA3: Consistent Treatment of Internal and Surface Residues in Empirical pKa predictions
   Mats H.M. Olsson, Chresten R. Sondergard, Michal Rostkowski, and Jan H. Jensen
   Journal of Chemical Theory and Computation, 7(2):525-537 (2011)

   Improved Treatment of Ligands and Coupling Effects in Empirical Calculation
    and Rationalization of pKa Values
   Chresten R. Sondergaard, Mats H.M. Olsson, Michal Rostkowski, and Jan H. Jensen
   Journal of Chemical Theory and Computation, (2011)


-----------------------------------------------------------------------------


 ---------  -----  ------   --------------------   --------------   --------------   --------------
                            DESOLVATION  EFFECTS      SIDECHAIN        BACKBONE        COULOMBIC
   RESIDUE    pKa   BURIED   REGULAR      RE        HYDROGEN BOND    HYDROGEN BOND    INTERACTION
```

```
--------    -----   ------   ---------   ---------    --------------      --------------      --------------

ASP   9 B   3.68     0 %     0.15  186    0.00    0    0.00 XXX   0 X      0.00 XXX   0 X     -0.11 N+    8 B
ASP   9 B                                               0.00 XXX   0 X      0.00 XXX   0 X     -0.19 ARG  43 B
ASP   9 B                                               0.00 XXX   0 X      0.00 XXX   0 X     -0.01 ARG  94 B
ASP   9 B                                               0.00 XXX   0 X      0.00 XXX   0 X      0.02 ASP  14 B
ASP   9 B                                               0.00 XXX   0 X      0.00 XXX   0 X      0.02 GLU  40 B

ASP  14 B   1.43     0 %     0.47  244    0.00    0   -0.80 SER   8 B      0.00 XXX   0 X     -0.14 N+    8 B
ASP  14 B                                              -0.77 THR  10 B      0.00 XXX   0 X     -0.01 ARG  43 B
ASP  14 B                                              -0.81 ARG  94 B      0.00 XXX   0 X     -0.02 HIS  17 B
ASP  14 B                                               0.00 XXX   0 X      0.00 XXX   0 X     -0.28 ARG  94 B

ASP  21 B   2.19     0 %     0.48  275    0.00    0   -0.68 ARG  24 B     -0.77 HIS  17 B     -0.08 LYS  20 B
ASP  21 B                                               0.00 XXX   0 X     -0.01 GLY  18 B     -0.23 HIS  17 B
ASP  21 B                                               0.00 XXX   0 X      0.00 XXX   0 X     -0.33 ARG  24 B

ASP  60 B   2.17    85 %     3.16  519    1.16    0   -0.50 THR  34 B     -0.15 ASP  60 B     -0.13 LYS 291 B
ASP  60 B                                              -1.62 ARG  38 B      0.00 XXX   0 X     -1.17 ARG  38 B
ASP  60 B                                              -1.60 HIS 262 B      0.00 XXX   0 X     -0.79 HIS 262 B

ASP  70 B   3.34     8 %     0.59  304    0.07    0   -0.81 ARG  73 B      0.00 XXX   0 X     -0.05 LYS  20 B
ASP  70 B                                               0.00 XXX   0 X      0.00 XXX   0 X     -0.21 LYS  23 B
ASP  70 B                                               0.00 XXX   0 X      0.00 XXX   0 X      0.33 ASP  74 B
ASP  70 B                                               0.00 XXX   0 X      0.00 XXX   0 X     -0.01 HIS 117 B
ASP  70 B                                               0.00 XXX   0 X      0.00 XXX   0 X     -0.04 HIS 147 B
ASP  70 B                                               0.00 XXX   0 X      0.00 XXX   0 X     -0.32 ARG  73 B

ASP  74 B   2.18    53 %     1.80  429    0.42    0   -1.60 HIS 147 B     -0.77 ASP  70 B     -0.13 LYS  23 B
ASP  74 B                                               0.00 XXX   0 X     -0.73 HIS 147 B     -0.05 ARG  73 B
ASP  74 B                                               0.00 XXX   0 X      0.00 XXX   0 X     -0.00 ARG 139 B
ASP  74 B                                               0.00 XXX   0 X      0.00 XXX   0 X     -0.55 HIS 147 B

ASP  79 B   4.72   100 %     3.96  616    0.37    0   -0.69 SER 128 B     -0.82 SER 126 B     -0.25 ARG 401 B
ASP  79 B                                              -0.73 SER 156 B     -0.03 THR 127 B     -0.21 HIS  82 B
ASP  79 B                                               0.00 XXX   0 X     -0.67 SER 128 B      0.00 XXX   0 X

ASP  95 B   3.93     0 %     0.21  274    0.00    0    0.00 XXX   0 X      0.00 XXX   0 X     -0.12 ARG  94 B
ASP  95 B                                               0.00 XXX   0 X      0.00 XXX   0 X      0.01 ASP  14 B
ASP  95 B                                               0.00 XXX   0 X      0.00 XXX   0 X     -0.05 HIS  17 B
ASP  95 B                                               0.00 XXX   0 X      0.00 XXX   0 X      0.07 GLU  40 B

ASP  99 B   2.94     0 %     0.30  236    0.00    0   -0.69 LYS 102 B      0.00 XXX   0 X     -0.10 ARG  47 B
ASP  99 B                                               0.00 XXX   0 X      0.00 XXX   0 X     -0.37 LYS 102 B

ASP 118 B   0.13    81 %     3.24  507    0.81    0   -1.60 HIS  87 B     -0.83 LYS 111 B     -0.01 LYS 111 B
```

```
ASP 118 B                                    -1.59 ARG  93 B    0.00 XXX   0 X   -0.07 HIS  66 B
ASP 118 B                                     0.00 XXX   0 X    0.00 XXX   0 X   -0.15 HIS 117 B
ASP 118 B                                     0.00 XXX   0 X    0.00 XXX   0 X   -1.43 HIS  87 B
ASP 118 B                                     0.00 XXX   0 X    0.00 XXX   0 X   -0.96 LYS  88 B
ASP 118 B                                     0.00 XXX   0 X    0.00 XXX   0 X   -1.09 ARG  93 B

ASP 140 B   2.50    3 %   0.64  290   0.04  0 -0.55 ARG 139 B    0.00 XXX   0 X   -0.17 LYS 175 B
ASP 140 B                                    -0.54 ARG 174 B    0.00 XXX   0 X   -0.44 ARG 139 B
ASP 140 B                                     0.00 XXX   0 X    0.00 XXX   0 X   -0.27 ARG 174 B

ASP 145 B   3.72    0 %   0.21  247   0.00  0  0.00 XXX   0 X    0.00 XXX   0 X   -0.09 ARG 139 B
ASP 145 B                                     0.00 XXX   0 X    0.00 XXX   0 X   -0.16 LYS 175 B
ASP 145 B                                     0.00 XXX   0 X    0.00 XXX   0 X    0.06 ASP  74 B
ASP 145 B                                     0.00 XXX   0 X    0.00 XXX   0 X    0.03 ASP 140 B
ASP 145 B                                     0.00 XXX   0 X    0.00 XXX   0 X   -0.13 HIS 147 B

ASP 154 B   5.46*  100 %   4.05  577   1.25  0  0.00 XXX   0 X   -0.28 ASP 154 B   -4.07 MG   MG B
ASP 154 B                                     0.00 XXX   0 X   -0.09 GLY 155 B   -0.40 LYS 289 B
ASP 154 B                                     0.00 XXX   0 X   -0.64 ASN 183 B    0.13 GLU 184 B
ASP 154 B                                     0.00 XXX   0 X   -0.01 GLU 184 B   -0.08 HIS 284 B
ASP 154 B                                     0.00 XXX   0 X   -0.06 MET 185 B    0.94 ASP 182 B
ASP 154 B                                     0.00 XXX   0 X    0.00 XXX   0 X    0.91 TDP O21 B

ASP 171 B   3.95    0 %   0.15  163   0.00  0  0.00 XXX   0 X    0.00 XXX   0 X    0.00 XXX   0 X

ASP 182 B   2.62   69 %   2.61  475   0.12  0 -1.60 HIS 284 B   -0.67 GLU 184 B   -1.08 MG   MG B
ASP 182 B                                     0.00 XXX   0 X    0.00 XXX   0 X   -0.56 HIS 284 B

ASP 260 B   3.44   15 %   0.61  324   0.02  0  0.00 XXX   0 X   -0.74 HIS 262 B   -0.11 LYS 291 B
ASP 260 B                                     0.00 XXX   0 X   -0.01 ASN 263 B    0.11 ASP  60 B
ASP 260 B                                     0.00 XXX   0 X    0.00 XXX   0 X   -0.24 HIS 262 B

ASP 276 B   3.91    0 %   0.28  276   0.00  0  0.00 XXX   0 X    0.00 XXX   0 X   -0.01 LYS 175 B
ASP 276 B                                     0.00 XXX   0 X    0.00 XXX   0 X   -0.01 ARG 273 B
ASP 276 B                                     0.00 XXX   0 X    0.00 XXX   0 X    0.02 ASP 145 B
ASP 276 B                                     0.00 XXX   0 X    0.00 XXX   0 X   -0.20 HIS 147 B
ASP 276 B                                     0.00 XXX   0 X    0.00 XXX   0 X    0.02 GLU 272 B
ASP 276 B                                     0.00 XXX   0 X    0.00 XXX   0 X    0.01 ASP 278 B

ASP 278 B   3.67    0 %   0.16  200   0.00  0  0.00 XXX   0 X    0.00 XXX   0 X   -0.22 LYS 175 B
ASP 278 B                                     0.00 XXX   0 X    0.00 XXX   0 X   -0.07 ARG 254 B

ASP 299 B   3.78   21 %   0.83  340   0.03  0  0.00 XXX   0 X   -0.09 ILE 301 B    0.00 XXX   0 X
ASP 299 B                                     0.00 XXX   0 X   -0.79 TYR 302 B    0.00 XXX   0 X

ASP 310 B   3.34    0 %   0.39  201   0.00  0 -0.14 THR 313 B   -0.65 ALA 312 B    0.00 XXX   0 X
```

147

```
ASP 310 B                                          0.00 XXX    0 X   -0.06 THR 313 B    0.00 XXX    0 X

ASP 339 B    3.08    40 %   1.46  393   0.07    0  -0.46 TRP 335 B   -0.58 ARG 341 B   -0.22 ARG 341 B
ASP 339 B                                         -0.69 THR 342 B    0.00 XXX    0 X   -0.17 ARG 365 B
ASP 339 B                                          0.00 XXX    0 X    0.00 XXX    0 X   -0.13 ARG 389 B

ASP 368 B    5.38    82 %   2.39  512   0.45    0   0.00 XXX    0 X   -0.32 ALA 348 B   -0.21 ARG 350 B
ASP 368 B                                          0.00 XXX    0 X   -0.80 GLY 370 B    0.06 GLU 189 B

ASP 404 B    4.07     0 %   0.33  254   0.00    0   0.00 XXX    0 X    0.00 XXX    0 X   -0.21 HIS 408 B
ASP 404 B                                          0.00 XXX    0 X    0.00 XXX    0 X    0.15 ASP 409 B

ASP 409 B    2.72     2 %   0.55  287   0.01    0  -0.65 THR 378 B    0.00 XXX    0 X   -0.32 HIS 408 B
ASP 409 B                                         -0.68 HIS 408 B    0.00 XXX    0 X    0.00 XXX    0 X

ASP 422 B    4.89   100 %   3.67  576   1.05    0  -0.79 SER 325 B   -0.61 ARG 423 B   -0.01 ARG 480 B
ASP 422 B                                         -0.52 ARG 423 B    0.00 XXX    0 X   -1.70 ARG 423 B

ASP 430 B    0.22   100 %   2.30  575   0.00    0  -0.85 LYS 101 B    0.00 XXX    0 X   -0.18 ARG 423 B
ASP 430 B                                          0.00 XXX    0 X    0.00 XXX    0 X   -0.38 ARG 480 B
ASP 430 B                                          0.00 XXX    0 X    0.00 XXX    0 X   -0.03 HIS 304 B
ASP 430 B                                          0.00 XXX    0 X    0.00 XXX    0 X   -0.01 HIS 597 B
ASP 430 B                                          0.00 XXX    0 X    0.00 XXX    0 X   -0.77 HIS  51 B
ASP 430 B                                          0.00 XXX    0 X    0.00 XXX    0 X   -2.03 LYS 101 B
ASP 430 B                                          0.00 XXX    0 X    0.00 XXX    0 X   -1.61 HIS 434 B

ASP 439 B    4.77   100 %   3.84  645   0.42    0  -0.85 SER 396 B    0.00 XXX    0 X    0.11 ASP 561 B
ASP 439 B                                         -0.52 ARG 477 B    0.00 XXX    0 X   -2.03 ARG 477 B

ASP 456 B    4.44    56 %   1.74  438   0.00    0   0.00 XXX    0 X   -0.00 ASP 456 B   -0.35 LYS 455 B
ASP 456 B                                          0.00 XXX    0 X   -0.00 ALA 458 B   -0.04 ARG 536 B
ASP 456 B                                          0.00 XXX    0 X   -0.79 GLU 459 B    0.09 GLU 500 B

ASP 471 B    3.61     0 %   0.20  260   0.00    0   0.00 XXX    0 X    0.00 XXX    0 X   -0.01 ARG 341 B
ASP 471 B                                          0.00 XXX    0 X    0.00 XXX    0 X   -0.18 ARG 389 B
ASP 471 B                                          0.00 XXX    0 X    0.00 XXX    0 X   -0.05 ARG 450 B
ASP 471 B                                          0.00 XXX    0 X    0.00 XXX    0 X   -0.10 HIS 414 B
ASP 471 B                                          0.00 XXX    0 X    0.00 XXX    0 X   -0.06 HIS 470 B

ASP 493 B    3.31     0 %   0.25  231   0.00    0  -0.34 LYS 465 B    0.00 XXX    0 X   -0.03 LYS 495 B
ASP 493 B                                          0.00 XXX    0 X    0.00 XXX    0 X   -0.02 HIS 470 B
ASP 493 B                                          0.00 XXX    0 X    0.00 XXX    0 X   -0.35 LYS 465 B

ASP 506 B    3.61     0 %   0.24  235   0.00    0   0.00 XXX    0 X   -0.11 ASP 506 B   -0.18 ARG 501 B
ASP 506 B                                          0.00 XXX    0 X    0.00 XXX    0 X   -0.18 ARG 552 B
ASP 506 B                                          0.00 XXX    0 X    0.00 XXX    0 X    0.04 ASP 507 B
```

148

```
ASP 507 B    3.21     3 %    0.48  291   0.02    0   -0.71 ARG 554 B    0.00 XXX   0 X   -0.38 ARG 554 B

ASP 518 B    4.66    53 %    0.89  429   0.36    0    0.00 XXX   0 X    0.00 XXX   0 X   -0.34 LYS 515 B
ASP 518 B                                            0.00 XXX   0 X    0.00 XXX   0 X   -0.06 LYS 522 B

ASP 526 B    3.84     0 %    0.15  171   0.00    0    0.00 XXX   0 X    0.00 XXX   0 X   -0.03 LYS 503 B
ASP 526 B                                            0.00 XXX   0 X    0.00 XXX   0 X   -0.07 ARG 615 B

ASP 542 B    4.13    35 %    1.53  380   0.31    0    0.00 XXX   0 X   -0.80 TRP 499 B   -0.06 ARG 536 B
ASP 542 B                                            0.00 XXX   0 X   -0.64 MET 545 B    0.00 XXX   0 X

ASP 561 B    4.23   100 %    3.39  621   0.87    0    0.00 XXX   0 X   -0.68 ILE 426 B   -0.28 LYS 515 B
ASP 561 B                                            0.00 XXX   0 X    0.00 XXX   0 X   -0.95 ARG 477 B
ASP 561 B                                            0.00 XXX   0 X    0.00 XXX   0 X   -1.92 HIS 604 B

ASP 592 B    4.04     0 %    0.24  208   0.00    0    0.00 XXX   0 X    0.00 XXX   0 X    0.00 XXX   0 X

ASP 610 B    4.13    19 %    0.96  335   0.04    0    0.00 XXX   0 X   -0.03 ASP 610 B   -0.10 HIS 604 B
ASP 610 B                                            0.00 XXX   0 X   -0.54 ALA 613 B    0.00 XXX   0 X

ASP 624 B    3.93     0 %    0.13  171   0.00    0    0.00 XXX   0 X    0.00 XXX   0 X   -0.01 ARG 554 B
ASP 624 B                                            0.00 XXX   0 X    0.00 XXX   0 X    0.01 ASP 507 B

GLU  28 B    4.26     0 %    0.28  194   0.00    0    0.00 XXX   0 X   -0.42 GLU  28 B   -0.10 ARG  27 B

GLU  35 B    4.72    20 %    0.37  338   0.00    0    0.00 XXX   0 X    0.00 XXX   0 X   -0.19 ARG  38 B
GLU  35 B                                            0.00 XXX   0 X    0.00 XXX   0 X   -0.01 LYS 291 B
GLU  35 B                                            0.00 XXX   0 X    0.00 XXX   0 X    0.04 ASP  60 B

GLU  36 B    4.72    21 %    1.47  339   0.17    0    0.00 XXX   0 X   -0.63 LEU  12 B   -0.06 N+    8 B
GLU  36 B                                            0.00 XXX   0 X   -0.82 LEU  13 B   -0.02 ARG  43 B
GLU  36 B                                            0.00 XXX   0 X    0.00 XXX   0 X   -0.10 ARG  94 B
GLU  36 B                                            0.00 XXX   0 X    0.00 XXX   0 X    0.04 ASP   9 B
GLU  36 B                                            0.00 XXX   0 X    0.00 XXX   0 X    0.13 ASP  14 B
GLU  36 B                                            0.00 XXX   0 X    0.00 XXX   0 X    0.04 GLU  40 B

GLU  40 B    3.24    46 %    1.47  409   0.11    0   -0.46 ARG  43 B    0.00 XXX   0 X    0.08 ASP  14 B
GLU  40 B                                           -1.59 ARG  94 B    0.00 XXX   0 X   -0.35 ARG  43 B
GLU  40 B                                            0.00 XXX   0 X    0.00 XXX   0 X   -0.52 ARG  94 B

GLU 103 B    4.02     0 %    0.34  244   0.00    0   -0.45 SER 602 B    0.00 XXX   0 X   -0.35 ARG 606 B
GLU 103 B                                            0.00 XXX   0 X    0.00 XXX   0 X   -0.03 HIS 597 B

GLU 114 B    3.93    64 %    2.45  461   0.66    0   -0.65 LYS  88 B   -0.80 SER 107 B   -0.07 LYS 111 B
GLU 114 B                                           -0.74 ARG  93 B    0.00 XXX   0 X   -0.11 HIS  87 B
```

149

```
GLU 114 B                                           0.00 XXX    0 X     0.00 XXX    0 X    -0.16 HIS 597 B
GLU 114 B                                           0.00 XXX    0 X     0.00 XXX    0 X    -1.03 LYS  88 B
GLU 114 B                                           0.00 XXX    0 X     0.00 XXX    0 X    -0.72 ARG  93 B
GLU 114 B                                           0.00 XXX    0 X     0.00 XXX    0 X     0.59 ASP 118 B

GLU 116 B    4.37    0 %   0.20  233   0.00    0     0.00 XXX    0 X     0.00 XXX    0 X    -0.05 LYS  20 B
GLU 116 B                                           0.00 XXX    0 X     0.00 XXX    0 X    -0.11 ARG  73 B
GLU 116 B                                           0.00 XXX    0 X     0.00 XXX    0 X    -0.01 ARG  75 B
GLU 116 B                                           0.00 XXX    0 X     0.00 XXX    0 X    -0.17 HIS 117 B

GLU 184 B    5.09   15 %   0.32  324   0.04    0     0.00 XXX    0 X     0.00 XXX    0 X    -0.06 MG   MG B
GLU 184 B                                           0.00 XXX    0 X     0.00 XXX    0 X    -0.01 ARG 199 B
GLU 184 B                                           0.00 XXX    0 X     0.00 XXX    0 X     0.29 ASP 182 B
GLU 184 B                                           0.00 XXX    0 X     0.00 XXX    0 X     0.02 GLU 189 B
GLU 184 B                                           0.00 XXX    0 X     0.00 XXX    0 X    -0.02 HIS 284 B

GLU 189 B    5.03   25 %   0.42  350   0.31    0     0.00 XXX    0 X     0.00 XXX    0 X    -0.20 ARG 350 B

GLU 266 B    5.63   18 %   0.78  333   0.14    0     0.00 XXX    0 X     0.00 XXX    0 X     0.22 ASP 260 B

GLU 272 B    3.65    7 %   0.62  300   0.05    0    -0.30 ARG  27 B     0.00 XXX    0 X    -0.39 ARG  27 B
GLU 272 B                                          -0.57 ARG 273 B     0.00 XXX    0 X    -0.24 ARG 273 B

GLU 297 B    3.97   10 %   0.47  308   0.06    0    -0.15 LYS 291 B    -0.42 GLY 292 B    -0.05 ARG  38 B
GLU 297 B                                           0.00 XXX    0 X     0.00 XXX    0 X    -0.35 LYS 291 B
GLU 297 B                                           0.00 XXX    0 X     0.00 XXX    0 X    -0.09 HIS 262 B

GLU 315 B    4.70    0 %   0.14  142   0.00    0     0.00 XXX    0 X     0.00 XXX    0 X     0.06 ASP 310 B

GLU 330 B    4.39   20 %   0.82  337   0.11    0    -0.38 GLN 485 B     0.00 XXX    0 X    -0.00 LYS 337 B
GLU 330 B                                          -0.61 TYR 322 B     0.00 XXX    0 X    -0.16 ARG 461 B
GLU 330 B                                           0.00 XXX    0 X     0.00 XXX    0 X     0.12 GLU 334 B

GLU 334 B    3.20   17 %   0.92  328   0.13    0    -0.37 TRP 491 B    -0.76 GLY 489 B    -0.13 LYS 337 B
GLU 334 B                                          -0.60 ARG 461 B     0.00 XXX    0 X    -0.50 ARG 461 B

GLU 351 B    5.63   55 %   1.78  434   0.00    0    -0.09 ARG 350 B     0.00 XXX    0 X    -0.10 ARG 360 B
GLU 351 B                                           0.00 XXX    0 X     0.00 XXX    0 X     0.07 GLU 189 B
GLU 351 B                                           0.00 XXX    0 X     0.00 XXX    0 X     0.05 ASP 299 B
GLU 351 B                                           0.00 XXX    0 X     0.00 XXX    0 X     0.07 ASP 368 B
GLU 351 B                                           0.00 XXX    0 X     0.00 XXX    0 X    -0.63 ARG 350 B

GLU 357 B    4.58    0 %   0.25  261   0.00    0     0.00 XXX    0 X     0.00 XXX    0 X    -0.13 LYS 337 B
GLU 357 B                                           0.00 XXX    0 X     0.00 XXX    0 X    -0.04 ARG 360 B
GLU 357 B                                           0.00 XXX    0 X     0.00 XXX    0 X     0.05 GLU 330 B
GLU 357 B                                           0.00 XXX    0 X     0.00 XXX    0 X    -0.04 HIS 362 B
```

150

```
GLU 373 B    7.98   100 %    3.55   587   0.83   0    0.00 XXX   0 X   -0.74 GLU 373 B    0.51 GLU 374 B
GLU 373 B                                               0.00 XXX   0 X    0.00 XXX   0 X   -0.67 ARG 401 B

GLU 374 B    5.25    57 %    1.27   440   0.34   0    0.00 XXX   0 X   -0.49 GLU 374 B   -0.36 ARG 401 B

GLU 413 B    4.64     0 %    0.14   177   0.00   0    0.00 XXX   0 X    0.00 XXX   0 X   -0.01 HIS 414 B

GLU 459 B    5.19    79 %    3.13   503   0.24   0   -0.82 ASN 534 B   -0.76 LYS 455 B    0.46 ASP 456 B
GLU 459 B                                              -0.61 ARG 536 B    0.00 XXX   0 X    0.35 GLU 500 B
GLU 459 B                                               0.00 XXX   0 X    0.00 XXX   0 X   -0.51 LYS 455 B
GLU 459 B                                               0.00 XXX   0 X    0.00 XXX   0 X   -0.78 ARG 536 B

GLU 498 B    4.73     0 %    0.41   278   0.00   0    0.00 XXX   0 X    0.00 XXX   0 X   -0.23 LYS 495 B
GLU 498 B                                               0.00 XXX   0 X    0.00 XXX   0 X   -0.12 ARG 536 B
GLU 498 B                                               0.00 XXX   0 X    0.00 XXX   0 X    0.17 ASP 542 B

GLU 500 B    3.59    25 %    0.70   350   0.09   0   -0.85 LYS 455 B    0.00 XXX   0 X   -0.35 ARG 536 B
GLU 500 B                                               0.00 XXX   0 X    0.00 XXX   0 X   -0.49 LYS 455 B

GLU 525 B    4.56     0 %    0.21   213   0.00   0    0.00 XXX   0 X    0.00 XXX   0 X   -0.04 LYS 503 B
GLU 525 B                                               0.00 XXX   0 X    0.00 XXX   0 X   -0.18 LYS 522 B
GLU 525 B                                               0.00 XXX   0 X    0.00 XXX   0 X   -0.07 ARG 615 B
GLU 525 B                                               0.00 XXX   0 X    0.00 XXX   0 X    0.12 ASP 526 B
GLU 525 B                                               0.00 XXX   0 X    0.00 XXX   0 X    0.03 GLU 628 B

GLU 543 B    4.84    14 %    0.40   321   0.03   0    0.00 XXX   0 X    0.00 XXX   0 X   -0.13 ARG 547 B
GLU 543 B                                               0.00 XXX   0 X    0.00 XXX   0 X    0.02 ASP 542 B
GLU 543 B                                               0.00 XXX   0 X    0.00 XXX   0 X    0.03 GLU 544 B

GLU 544 B    4.58     0 %    0.12   172   0.00   0    0.00 XXX   0 X    0.00 XXX   0 X   -0.14 ARG 547 B
GLU 544 B                                               0.00 XXX   0 X    0.00 XXX   0 X    0.07 ASP 542 B
GLU 544 B                                               0.00 XXX   0 X    0.00 XXX   0 X    0.03 GLU 548 B

GLU 548 B    4.43     0 %    0.19   219   0.00   0    0.00 XXX   0 X    0.00 XXX   0 X   -0.01 ARG 501 B
GLU 548 B                                               0.00 XXX   0 X    0.00 XXX   0 X   -0.28 ARG 552 B
GLU 548 B                                               0.00 XXX   0 X    0.00 XXX   0 X    0.03 ASP 542 B

GLU 560 B    7.65    93 %    3.47   543   1.59   0    0.00 XXX   0 X   -0.72 PHE 568 B   -0.12 ARG 444 B
GLU 560 B                                               0.00 XXX   0 X   -0.82 GLY 569 B   -0.28 ARG 477 B
GLU 560 B                                               0.00 XXX   0 X    0.00 XXX   0 X    0.02 ASP 561 B

GLU 574 B    3.77    11 %    0.50   311   0.07   0   -0.80 ARG 444 B    0.00 XXX   0 X   -0.02 LYS 539 B
GLU 574 B                                               0.00 XXX   0 X    0.00 XXX   0 X   -0.49 ARG 444 B

GLU 593 B    4.70     0 %    0.14   193   0.00   0    0.00 XXX   0 X    0.00 XXX   0 X    0.04 ASP 592 B
```

151

```
   GLU  593 B                                          0.00 XXX    0 X     0.00 XXX    0 X     0.02 GLU 596 B

   GLU  596 B   4.44    26 %    0.59   354    0.00    0   0.00 XXX    0 X     0.00 XXX    0 X    -0.41 LYS 111 B
   GLU  596 B                                          0.00 XXX    0 X     0.00 XXX    0 X    -0.01 ARG 606 B
   GLU  596 B                                          0.00 XXX    0 X     0.00 XXX    0 X     0.01 GLU 114 B
   GLU  596 B                                          0.00 XXX    0 X     0.00 XXX    0 X    -0.25 HIS 597 B

   GLU  601 B   4.45    18 %    0.48   331    0.00    0   0.00 XXX    0 X    -0.19 GLU 601 B    -0.08 ARG 480 B
   GLU  601 B                                          0.00 XXX    0 X     0.00 XXX    0 X    -0.15 LYS 515 B
   GLU  601 B                                          0.00 XXX    0 X     0.00 XXX    0 X    -0.11 HIS 604 B

   GLU  620 B   4.62     0 %    0.26   240    0.00    0   0.00 XXX    0 X     0.00 XXX    0 X    -0.14 ARG 586 B

   GLU  628 B   4.29     0 %    0.36   195    0.00    0  -0.22 ARG 615 B     0.00 XXX    0 X     0.03 ASP 526 B
   GLU  628 B                                          0.00 XXX    0 X     0.00 XXX    0 X    -0.38 ARG 615 B

   HIS   17 B   6.51     0 %   -0.25   239    0.00    0   0.00 XXX    0 X     0.00 XXX    0 X    -0.02 ARG  24 B
   HIS   17 B                                          0.00 XXX    0 X     0.00 XXX    0 X    -0.02 ARG  94 B
   HIS   17 B                                          0.00 XXX    0 X     0.00 XXX    0 X     0.02 ASP  14 B
   HIS   17 B                                          0.00 XXX    0 X     0.00 XXX    0 X     0.23 ASP  21 B
   HIS   17 B                                          0.00 XXX    0 X     0.00 XXX    0 X     0.05 ASP  95 B

   HIS   51 B   0.43   100 %   -2.59   642    0.00    0   0.00 XXX    0 X     0.00 XXX    0 X    -1.66 LYS 101 B
   HIS   51 B                                          0.00 XXX    0 X     0.00 XXX    0 X    -0.12 LYS 289 B
   HIS   51 B                                          0.00 XXX    0 X     0.00 XXX    0 X    -1.12 HIS  82 B
   HIS   51 B                                          0.00 XXX    0 X     0.00 XXX    0 X    -1.24 HIS 304 B
   HIS   51 B                                          0.00 XXX    0 X     0.00 XXX    0 X    -0.25 HIS 434 B
   HIS   51 B                                          0.00 XXX    0 X     0.00 XXX    0 X     0.15 TDP O21 B
   HIS   51 B                                          0.00 XXX    0 X     0.00 XXX    0 X     0.77 ASP 430 B

   HIS   66 B   1.51   100 %   -2.89   590    0.00    0   0.00 XXX    0 X     0.00 XXX    0 X    -0.34 LYS  23 B
   HIS   66 B                                          0.00 XXX    0 X     0.00 XXX    0 X    -0.09 ARG  93 B
   HIS   66 B                                          0.00 XXX    0 X     0.00 XXX    0 X    -1.32 HIS  87 B
   HIS   66 B                                          0.00 XXX    0 X     0.00 XXX    0 X    -0.41 HIS 117 B
   HIS   66 B                                          0.00 XXX    0 X     0.00 XXX    0 X     0.07 ASP 118 B

   HIS   82 B   4.51   100 %   -3.32   669    0.00    0   0.01 GLN  83 B     0.00 XXX    0 X    -0.74 MG   MG B
   HIS   82 B                                          0.53 TDP O11 B     0.00 XXX    0 X    -0.08 LYS 101 B
   HIS   82 B                                          0.71 TDP O21 B     0.00 XXX    0 X    -0.54 LYS 289 B
   HIS   82 B                                          0.00 XXX    0 X     0.00 XXX    0 X     0.21 ASP  79 B
   HIS   82 B                                          0.00 XXX    0 X     0.00 XXX    0 X     1.23 TDP O21 B

   HIS   87 B   4.64   100 %   -3.12   614    0.00    0   1.60 ASP 118 B     0.07 HIS 117 B    -0.46 LYS  88 B
   HIS   87 B                                          0.00 XXX    0 X     0.00 XXX    0 X    -1.11 ARG  93 B
   HIS   87 B                                          0.00 XXX    0 X     0.00 XXX    0 X     0.11 GLU 114 B
   HIS   87 B                                          0.00 XXX    0 X     0.00 XXX    0 X    -0.37 HIS 117 B
```

152

```
HIS  87 B                                    0.00 XXX   0 X   0.00 XXX   0 X    1.43 ASP 118 B

HIS 117 B   5.19   49 %   -1.46  419   0.00   0   0.00 XXX   0 X   0.18 SER  71 B   -0.06 LYS  20 B
HIS 117 B                                    0.00 XXX   0 X   0.00 XXX   0 X   -0.01 LYS  23 B
HIS 117 B                                    0.00 XXX   0 X   0.00 XXX   0 X   -0.07 ARG  73 B
HIS 117 B                                    0.00 XXX   0 X   0.00 XXX   0 X   -0.22 ARG  93 B
HIS 117 B                                    0.00 XXX   0 X   0.00 XXX   0 X    0.01 ASP  70 B
HIS 117 B                                    0.00 XXX   0 X   0.00 XXX   0 X    0.17 GLU 116 B
HIS 117 B                                    0.00 XXX   0 X   0.00 XXX   0 X    0.15 ASP 118 B

HIS 124 B   2.70   92 %   -2.38  539   0.00   0   0.00 XXX   0 X   0.11 THR 433 B   -0.70 ARG 401 B
HIS 124 B                                    0.00 XXX   0 X   0.00 XXX   0 X   -0.83 HIS 434 B

HIS 147 B   7.66   47 %   -1.24  412   0.00   0   1.60 ASP  74 B   0.00 XXX   0 X   -0.07 ARG 139 B
HIS 147 B                                    0.00 XXX   0 X   0.00 XXX   0 X   -0.04 LYS 175 B
HIS 147 B                                    0.00 XXX   0 X   0.00 XXX   0 X    0.04 ASP  70 B
HIS 147 B                                    0.00 XXX   0 X   0.00 XXX   0 X    0.13 ASP 145 B
HIS 147 B                                    0.00 XXX   0 X   0.00 XXX   0 X    0.20 ASP 276 B
HIS 147 B                                    0.00 XXX   0 X   0.00 XXX   0 X    0.55 ASP  74 B

HIS 262 B   5.90   72 %   -2.33  484   0.00   0   1.60 ASP  60 B   0.00 GLY 261 B   -0.57 ARG  38 B
HIS 262 B                                    0.00 XXX   0 X   0.00 XXX   0 X   -0.42 LYS 291 B
HIS 262 B                                    0.00 XXX   0 X   0.00 XXX   0 X    0.24 ASP 260 B
HIS 262 B                                    0.00 XXX   0 X   0.00 XXX   0 X    0.09 GLU 297 B
HIS 262 B                                    0.00 XXX   0 X   0.00 XXX   0 X    0.79 ASP  60 B

HIS 284 B   6.61   66 %   -2.12  466   0.00   0   1.60 ASP 182 B   0.00 XXX   0 X   -0.03 ARG 199 B
HIS 284 B                                    0.00 XXX   0 X   0.00 XXX   0 X    0.08 ASP 154 B
HIS 284 B                                    0.00 XXX   0 X   0.00 XXX   0 X    0.02 GLU 184 B
HIS 284 B                                    0.00 XXX   0 X   0.00 XXX   0 X    0.56 ASP 182 B

HIS 304 B   1.71  100 %   -2.79  617   0.00   0   0.00 XXX   0 X   0.00 XXX   0 X   -0.33 MG   MG B
HIS 304 B                                    0.00 XXX   0 X   0.00 XXX   0 X   -0.12 LYS 101 B
HIS 304 B                                    0.00 XXX   0 X   0.00 XXX   0 X   -1.62 LYS 289 B
HIS 304 B                                    0.00 XXX   0 X   0.00 XXX   0 X   -1.08 HIS  82 B
HIS 304 B                                    0.00 XXX   0 X   0.00 XXX   0 X    0.03 ASP 430 B
HIS 304 B                                    0.00 XXX   0 X   0.00 XXX   0 X    1.12 TDP O21 B

HIS 362 B   5.16   21 %   -1.05  340   0.00   0   0.00 XXX   0 X   0.00 XXX   0 X   -0.12 LYS 337 B
HIS 362 B                                    0.00 XXX   0 X   0.00 XXX   0 X   -0.22 ARG 365 B
HIS 362 B                                    0.00 XXX   0 X   0.00 XXX   0 X    0.04 GLU 357 B

HIS 364 B   6.34    0 %   -0.15  160   0.00   0   0.00 XXX   0 X   0.00 XXX   0 X   -0.01 ARG 365 B

HIS 408 B   7.47    0 %   -0.23  202   0.00   0   0.68 ASP 409 B   0.00 XXX   0 X    0.21 ASP 404 B
HIS 408 B                                    0.00 XXX   0 X   0.00 XXX   0 X    0.32 ASP 409 B
```

153

```
HIS 414 B    5.56     20 %    -0.67   338   0.00      0    0.00 XXX    0 X    0.00 XXX    0 X   -0.37 ARG 450 B
HIS 414 B                                                   0.00 XXX    0 X    0.00 XXX    0 X   -0.00 LYS 539 B
HIS 414 B                                                   0.00 XXX    0 X    0.00 XXX    0 X    0.01 GLU 413 B
HIS 414 B                                                   0.00 XXX    0 X    0.00 XXX    0 X    0.10 ASP 471 B

HIS 434 B    4.76    100 %    -2.87   585   0.00      0    0.00 XXX    0 X    0.00 XXX    0 X   -0.48 LYS 101 B
HIS 434 B                                                   0.00 XXX    0 X    0.00 XXX    0 X    1.61 ASP 430 B

HIS 470 B    5.59     12 %    -0.87   316   0.00      0    0.00 XXX    0 X    0.00 XXX    0 X   -0.06 ARG 389 B
HIS 470 B                                                   0.00 XXX    0 X    0.00 XXX    0 X   -0.05 ARG 450 B
HIS 470 B                                                   0.00 XXX    0 X    0.00 XXX    0 X   -0.01 LYS 465 B
HIS 470 B                                                   0.00 XXX    0 X    0.00 XXX    0 X    0.06 ASP 471 B
HIS 470 B                                                   0.00 XXX    0 X    0.00 XXX    0 X    0.02 ASP 493 B

HIS 582 B    6.27      0 %    -0.23   170   0.00      0    0.00 XXX    0 X    0.00 XXX    0 X    0.00 XXX    0 X

HIS 597 B    5.56     49 %    -1.34   418   0.00      0    0.00 XXX    0 X    0.03 LYS 101 B   -0.11 LYS  88 B
HIS 597 B                                                   0.00 XXX    0 X    0.20 LYS 102 B   -0.08 LYS 101 B
HIS 597 B                                                   0.00 XXX    0 X    0.00 XXX    0 X   -0.08 LYS 111 B
HIS 597 B                                                   0.00 XXX    0 X    0.00 XXX    0 X    0.03 GLU 103 B
HIS 597 B                                                   0.00 XXX    0 X    0.00 XXX    0 X    0.16 GLU 114 B
HIS 597 B                                                   0.00 XXX    0 X    0.00 XXX    0 X    0.01 ASP 430 B
HIS 597 B                                                   0.00 XXX    0 X    0.00 XXX    0 X    0.25 GLU 596 B

HIS 604 B    5.95     80 %    -2.48   504   0.00      0    0.00 XXX    0 X    0.00 XXX    0 X   -0.01 ARG 477 B
HIS 604 B                                                   0.00 XXX    0 X    0.00 XXX    0 X   -0.18 LYS 515 B
HIS 604 B                                                   0.00 XXX    0 X    0.00 XXX    0 X   -0.01 ARG 606 B
HIS 604 B                                                   0.00 XXX    0 X    0.00 XXX    0 X    0.11 GLU 601 B
HIS 604 B                                                   0.00 XXX    0 X    0.00 XXX    0 X    0.10 ASP 610 B
HIS 604 B                                                   0.00 XXX    0 X    0.00 XXX    0 X    1.92 ASP 561 B

CYS  45 B   12.62    100 %     3.73   575   0.00      0    0.00 XXX    0 X    0.00 XXX    0 X   -0.09 ARG  47 B
CYS  45 B                                                   0.00 XXX    0 X    0.00 XXX    0 X   -0.02 LYS 101 B

CYS 420 B   12.82    100 %     3.73   737   0.00      0    0.00 XXX    0 X    0.00 XXX    0 X    0.09 ASP 422 B

TYR  67 B   10.59     19 %     0.87   335   0.00      0   -0.16 ARG  27 B    0.00 XXX    0 X   -0.27 ARG  27 B
TYR  67 B                                                   0.00 XXX    0 X    0.00 XXX    0 X    0.08 GLU  28 B
TYR  67 B                                                   0.00 XXX    0 X    0.00 XXX    0 X    0.08 GLU 272 B

TYR  85 B   13.82    100 %     3.47   601   0.00      0    0.00 XXX    0 X   -0.84 HIS  51 B    1.13 CYS  45 B
TYR  85 B                                                   0.00 XXX    0 X    0.00 XXX    0 X   -0.12 ARG  47 B
TYR  85 B                                                   0.00 XXX    0 X    0.00 XXX    0 X    0.18 ASP 430 B

TYR 255 B   11.64     34 %     1.62   377   0.00      0    0.00 XXX    0 X    0.00 XXX    0 X    0.09 ASP 182 B
```

154

```
  TYR 255 B                                     0.00 XXX    0 X     0.00 XXX    0 X    -0.07 ARG 199 B

  TYR 295 B  10.62    0 %    0.50  247   0.00    0    0.00 XXX    0 X     0.00 XXX    0 X     0.09 ASP 310 B
  TYR 295 B                                     0.00 XXX    0 X     0.00 XXX    0 X     0.03 GLU 315 B

  TYR 302 B  11.37   23 %    1.43  347   0.00    0    0.00 XXX    0 X    -0.35 ALA 307 B     0.05 ASP 299 B
  TYR 302 B                                     0.00 XXX    0 X     0.00 XXX    0 X     0.24 TYR 316 B

  TYR 316 B  10.51    0 %    0.32  271   0.00    0    0.00 XXX    0 X     0.00 XXX    0 X     0.19 ASP 299 B
  TYR 316 B                                     0.00 XXX    0 X     0.00 XXX    0 X     0.00 GLU 315 B

  TYR 322 B  12.12   26 %    1.38  355   0.00    0   -0.25 GLN 485 B     0.00 XXX    0 X     0.05 GLU 334 B
  TYR 322 B                                     0.61 GLU 330 B     0.00 XXX    0 X    -0.14 ARG 461 B
  TYR 322 B                                     0.00 XXX    0 X     0.00 XXX    0 X     0.47 GLU 330 B

  TYR 366 B  12.68   70 %    2.64  477   0.00    0   -0.48 ARG 350 B     0.00 XXX    0 X     0.03 GLU 189 B
  TYR 366 B                                     0.00 XXX    0 X     0.00 XXX    0 X    -0.51 ARG 350 B
  TYR 366 B                                     0.00 XXX    0 X     0.00 XXX    0 X     0.21 GLU 351 B
  TYR 366 B                                     0.00 XXX    0 X     0.00 XXX    0 X     0.79 ASP 368 B

  TYR 395 B  12.09  100 %    2.96  586   0.00    0    0.00 XXX    0 X     0.00 XXX    0 X    -2.03 ARG 423 B
  TYR 395 B                                     0.00 XXX    0 X     0.00 XXX    0 X     0.20 ASP 430 B
  TYR 395 B                                     0.00 XXX    0 X     0.00 XXX    0 X    -0.16 ARG 480 B
  TYR 395 B                                     0.00 XXX    0 X     0.00 XXX    0 X     1.12 ASP 422 B

  TYR 403 B  10.37    0 %    0.33  277   0.00    0    0.00 XXX    0 X     0.00 XXX    0 X     0.04 ASP 404 B

  TYR 466 B  10.67   20 %    0.82  337   0.00    0    0.00 XXX    0 X     0.00 XXX    0 X    -0.14 ARG 450 B
  TYR 466 B                                     0.00 XXX    0 X     0.00 XXX    0 X     0.01 ASP 471 B
  TYR 466 B                                     0.00 XXX    0 X     0.00 XXX    0 X    -0.02 LYS 495 B

  TYR 478 B  16.80  100 %    3.86  663   0.00    0    0.00 XXX    0 X     0.00 XXX    0 X     1.46 CYS 420 B
  TYR 478 B                                     0.00 XXX    0 X     0.00 XXX    0 X    -0.15 ARG 423 B
  TYR 478 B                                     0.00 XXX    0 X     0.00 XXX    0 X     0.10 TYR 395 B
  TYR 478 B                                     0.00 XXX    0 X     0.00 XXX    0 X     1.54 ASP 422 B

  TYR 519 B  12.32   48 %    1.49  415   0.00    0    0.00 XXX    0 X     0.00 XXX    0 X     0.14 ASP 518 B
  TYR 519 B                                     0.00 XXX    0 X     0.00 XXX    0 X     0.32 ASP 561 B
  TYR 519 B                                     0.00 XXX    0 X     0.00 XXX    0 X     0.21 GLU 601 B
  TYR 519 B                                     0.00 XXX    0 X     0.00 XXX    0 X     0.17 ASP 610 B

  LYS  20 B  10.29    0 %   -0.15  205   0.00    0    0.00 XXX    0 X     0.00 XXX    0 X     0.08 ASP  21 B
  LYS  20 B                                     0.00 XXX    0 X     0.00 XXX    0 X    -0.05 ARG  24 B
  LYS  20 B                                     0.00 XXX    0 X     0.00 XXX    0 X     0.05 ASP  70 B
  LYS  20 B                                     0.00 XXX    0 X     0.00 XXX    0 X    -0.19 ARG  73 B
  LYS  20 B                                     0.00 XXX    0 X     0.00 XXX    0 X     0.05 GLU 116 B
```

```
LYS  23 B    8.57    58 %   -2.24  443   0.00    0   0.00 XXX   0 X   0.00 XXX   0 X    0.21 ASP  70 B
LYS  23 B                                              0.00 XXX   0 X   0.00 XXX   0 X   -0.02 ARG  73 B
LYS  23 B                                              0.00 XXX   0 X   0.00 XXX   0 X    0.13 ASP  74 B

LYS  88 B    8.66    90 %   -3.58  532   0.00    0   0.65 GLU 114 B   0.00 XXX   0 X   -0.88 ARG  93 B
LYS  88 B                                              0.00 XXX   0 X   0.00 XXX   0 X   -0.02 LYS 111 B
LYS  88 B                                              0.00 XXX   0 X   0.00 XXX   0 X    1.03 GLU 114 B
LYS  88 B                                              0.00 XXX   0 X   0.00 XXX   0 X    0.96 ASP 118 B

LYS 101 B   10.50   100 %   -2.70  567   0.00    0   0.85 ASP 430 B   0.00 XXX   0 X    0.02 CYS  45 B
LYS 101 B                                              0.00 XXX   0 X   0.00 XXX   0 X   -0.05 ARG  47 B
LYS 101 B                                              0.00 XXX   0 X   0.00 XXX   0 X   -0.15 ARG 480 B
LYS 101 B                                              0.00 XXX   0 X   0.00 XXX   0 X    2.03 ASP 430 B

LYS 102 B   10.84     0 %   -0.39  248   0.00    0   0.69 ASP  99 B   0.00 XXX   0 X   -0.34 ARG  47 B
LYS 102 B                                              0.00 XXX   0 X   0.00 XXX   0 X    0.37 ASP  99 B

LYS 111 B   10.41    21 %   -0.58  339   0.00    0   0.00 XXX   0 X   0.00 XXX   0 X    0.07 GLU 114 B
LYS 111 B                                              0.00 XXX   0 X   0.00 XXX   0 X    0.01 ASP 118 B
LYS 111 B                                              0.00 XXX   0 X   0.00 XXX   0 X    0.41 GLU 596 B

LYS 144 B   10.28     0 %   -0.21  174   0.00    0   0.00 XXX   0 X   0.00 XXX   0 X   -0.01 ARG  75 B

LYS 175 B   10.63     0 %   -0.22  269   0.00    0   0.00 XXX   0 X   0.00 XXX   0 X   -0.18 ARG 139 B
LYS 175 B                                              0.00 XXX   0 X   0.00 XXX   0 X    0.17 ASP 140 B
LYS 175 B                                              0.00 XXX   0 X   0.00 XXX   0 X    0.16 ASP 145 B
LYS 175 B                                              0.00 XXX   0 X   0.00 XXX   0 X   -0.02 ARG 174 B
LYS 175 B                                              0.00 XXX   0 X   0.00 XXX   0 X    0.01 ASP 276 B
LYS 175 B                                              0.00 XXX   0 X   0.00 XXX   0 X    0.22 ASP 278 B

LYS 196 B    9.82    20 %   -0.68  338   0.00    0   0.00 XXX   0 X   0.00 XXX   0 X    0.00 XXX   0 X

LYS 289 B    8.42   100 %   -3.62  575   0.00    0   0.80 TDP O21 B   0.00 XXX   0 X   -2.08 MG   MG B
LYS 289 B                                              0.00 XXX   0 X   0.00 XXX   0 X    0.40 ASP 154 B
LYS 289 B                                              0.00 XXX   0 X   0.00 XXX   0 X    0.39 TDP O13 B
LYS 289 B                                              0.00 XXX   0 X   0.00 XXX   0 X    2.03 TDP O21 B

LYS 291 B   10.57    18 %   -0.49  332   0.00    0   0.15 GLU 297 B   0.00 XXX   0 X    0.01 GLU  35 B
LYS 291 B                                              0.00 XXX   0 X   0.00 XXX   0 X   -0.18 ARG  38 B
LYS 291 B                                              0.00 XXX   0 X   0.00 XXX   0 X    0.13 ASP  60 B
LYS 291 B                                              0.00 XXX   0 X   0.00 XXX   0 X    0.11 ASP 260 B
LYS 291 B                                              0.00 XXX   0 X   0.00 XXX   0 X    0.35 GLU 297 B

LYS 308 B   10.37     0 %   -0.13  174   0.00    0   0.00 XXX   0 X   0.00 XXX   0 X    0.00 XXX   0 X
```

156

```
LYS  337 B   10.52     0 %    -0.24   271    0.00    0     0.00 XXX    0 X     0.00 XXX    0 X      0.00 GLU  330 B
LYS  337 B                                                     0.00 XXX    0 X     0.00 XXX    0 X      0.13 GLU  334 B
LYS  337 B                                                     0.00 XXX    0 X     0.00 XXX    0 X      0.13 GLU  357 B

LYS  455 B   11.53    31 %    -1.00   368    0.00    0     0.85 GLU  500 B   0.00 XXX    0 X      0.35 ASP  456 B
LYS  455 B                                                     0.00 XXX    0 X     0.00 XXX    0 X     -0.18 ARG  536 B
LYS  455 B                                                     0.00 XXX    0 X     0.00 XXX    0 X      0.51 GLU  459 B
LYS  455 B                                                     0.00 XXX    0 X     0.00 XXX    0 X      0.49 GLU  500 B

LYS  465 B   10.84     2 %    -0.36   288    0.00    0     0.34 ASP  493 B   0.00 XXX    0 X      0.35 ASP  493 B

LYS  495 B   10.68     0 %    -0.10   165    0.00    0     0.00 XXX    0 X     0.00 XXX    0 X      0.03 ASP  493 B
LYS  495 B                                                     0.00 XXX    0 X     0.00 XXX    0 X      0.23 GLU  498 B
LYS  495 B                                                     0.00 XXX    0 X     0.00 XXX    0 X      0.02 TYR  466 B

LYS  503 B   10.11     8 %    -0.45   304    0.00    0     0.00 XXX    0 X     0.00 XXX    0 X     -0.01 ARG  501 B
LYS  503 B                                                     0.00 XXX    0 X     0.00 XXX    0 X      0.04 GLU  525 B
LYS  503 B                                                     0.00 XXX    0 X     0.00 XXX    0 X      0.03 ASP  526 B

LYS  515 B   10.25    58 %    -1.00   444    0.00    0     0.00 XXX    0 X     0.00 XXX    0 X     -0.02 ARG  480 B
LYS  515 B                                                     0.00 XXX    0 X     0.00 XXX    0 X      0.34 ASP  518 B
LYS  515 B                                                     0.00 XXX    0 X     0.00 XXX    0 X      0.28 ASP  561 B
LYS  515 B                                                     0.00 XXX    0 X     0.00 XXX    0 X      0.15 GLU  601 B

LYS  522 B   10.60     0 %    -0.14   220    0.00    0     0.00 XXX    0 X     0.00 XXX    0 X      0.06 ASP  518 B
LYS  522 B                                                     0.00 XXX    0 X     0.00 XXX    0 X      0.18 GLU  525 B

LYS  539 B    9.01    40 %    -1.36   392    0.00    0     0.00 XXX    0 X     0.00 XXX    0 X     -0.11 ARG  444 B
LYS  539 B                                                     0.00 XXX    0 X     0.00 XXX    0 X     -0.04 ARG  450 B
LYS  539 B                                                     0.00 XXX    0 X     0.00 XXX    0 X      0.02 GLU  574 B

ARG   24 B   13.22     0 %    -0.28   201    0.00    0     0.68 ASP   21 B   0.00 XXX    0 X      0.33 ASP   21 B

ARG   27 B   13.34     0 %    -0.38   282    0.00    0     0.16 TYR   67 B   0.00 XXX    0 X      0.10 GLU   28 B
ARG   27 B                                                     0.30 GLU  272 B   0.00 XXX    0 X      0.27 TYR   67 B
ARG   27 B                                                     0.00 XXX    0 X     0.00 XXX    0 X      0.39 GLU  272 B

ARG   38 B   13.02    80 %    -2.51   505    0.00    0     1.62 ASP   60 B   0.00 XXX    0 X      0.19 GLU   35 B
ARG   38 B                                                     0.00 XXX    0 X     0.00 XXX    0 X      0.05 GLU  297 B
ARG   38 B                                                     0.00 XXX    0 X     0.00 XXX    0 X      1.17 ASP   60 B

ARG   43 B   12.72    21 %    -0.70   341    0.00    0     0.46 GLU   40 B   0.00 XXX    0 X      0.19 ASP    9 B
ARG   43 B                                                     0.00 XXX    0 X     0.00 XXX    0 X      0.01 ASP   14 B
ARG   43 B                                                     0.00 XXX    0 X     0.00 XXX    0 X      0.02 GLU   36 B
ARG   43 B                                                     0.00 XXX    0 X     0.00 XXX    0 X     -0.12 ARG   94 B
ARG   43 B                                                     0.00 XXX    0 X     0.00 XXX    0 X      0.35 GLU   40 B
```

157

```
ARG  47 B  11.77   29 %  -1.05  363   0.00    0   0.00 XXX   0 X   0.00 XXX   0 X   0.09 CYS  45 B
ARG  47 B                                          0.00 XXX   0 X   0.00 XXX   0 X   0.12 TYR  85 B
ARG  47 B                                          0.00 XXX   0 X   0.00 XXX   0 X   0.10 ASP  99 B

ARG  73 B  13.49    0 %  -0.30  237   0.00    0   0.81 ASP  70 B   0.00 XXX   0 X   0.05 ASP  74 B
ARG  73 B                                          0.00 XXX   0 X   0.00 XXX   0 X   0.11 GLU 116 B
ARG  73 B                                          0.00 XXX   0 X   0.00 XXX   0 X   0.32 ASP  70 B

ARG  75 B  11.63   41 %  -0.88  395   0.00    0   0.00 XXX   0 X   0.00 XXX   0 X   0.01 GLU 116 B

ARG  93 B  13.83   78 %  -2.81  501   0.00    0   0.74 GLU 114 B   0.00 XXX   0 X   0.72 GLU 114 B
ARG  93 B                                          1.59 ASP 118 B   0.00 XXX   0 X   1.09 ASP 118 B

ARG  94 B  15.05   18 %  -0.88  333   0.00    0   0.81 ASP  14 B   0.00 XXX   0 X   0.01 ASP   9 B
ARG  94 B                                          1.59 GLU  40 B   0.00 XXX   0 X   0.10 GLU  36 B
ARG  94 B                                          0.00 XXX   0 X   0.00 XXX   0 X   0.12 ASP  95 B
ARG  94 B                                          0.00 XXX   0 X   0.00 XXX   0 X   0.28 ASP  14 B
ARG  94 B                                          0.00 XXX   0 X   0.00 XXX   0 X   0.52 GLU  40 B

ARG 139 B  12.36   29 %  -1.16  362   0.00    0   0.55 ASP 140 B   0.00 XXX   0 X   0.00 ASP  74 B
ARG 139 B                                          0.00 XXX   0 X   0.00 XXX   0 X   0.09 ASP 145 B
ARG 139 B                                          0.00 XXX   0 X   0.00 XXX   0 X  -0.07 ARG 174 B
ARG 139 B                                          0.00 XXX   0 X   0.00 XXX   0 X   0.44 ASP 140 B

ARG 174 B  12.96    0 %  -0.35  246   0.00    0   0.54 ASP 140 B   0.00 XXX   0 X   0.27 ASP 140 B

ARG 199 B  12.42    0 %  -0.15  195   0.00    0   0.00 XXX   0 X   0.00 XXX   0 X   0.01 GLU 184 B
ARG 199 B                                          0.00 XXX   0 X   0.00 XXX   0 X   0.07 TYR 255 B

ARG 254 B  12.04    8 %  -0.53  304   0.00    0   0.00 XXX   0 X   0.00 XXX   0 X   0.07 ASP 278 B

ARG 273 B  12.77    0 %  -0.48  212   0.00    0   0.57 GLU 272 B   0.00 XXX   0 X   0.01 ASP 276 B
ARG 273 B                                          0.00 XXX   0 X   0.00 XXX   0 X  -0.09 ARG  27 B
ARG 273 B                                          0.00 XXX   0 X   0.00 XXX   0 X   0.24 GLU 272 B

ARG 341 B  12.46    0 %  -0.28  280   0.00    0   0.00 XXX   0 X   0.00 XXX   0 X   0.22 ASP 339 B
ARG 341 B                                          0.00 XXX   0 X   0.00 XXX   0 X   0.01 ASP 471 B

ARG 350 B  13.50   41 %  -1.12  395   0.00    0   0.48 TYR 366 B   0.00 XXX   0 X   0.20 GLU 189 B
ARG 350 B                                          0.09 GLU 351 B   0.00 XXX   0 X   0.51 TYR 366 B
ARG 350 B                                          0.00 XXX   0 X   0.00 XXX   0 X   0.21 ASP 368 B
ARG 350 B                                          0.00 XXX   0 X   0.00 XXX   0 X   0.63 GLU 351 B

ARG 360 B  12.30    0 %  -0.23  231   0.00    0   0.00 XXX   0 X   0.00 XXX   0 X   0.10 GLU 351 B
ARG 360 B                                          0.00 XXX   0 X   0.00 XXX   0 X   0.04 GLU 357 B
```

158

```
ARG 360 B                                  0.00 XXX    0 X    0.00 XXX    0 X   -0.12 ARG 350 B

ARG 365 B   11.33    32 %   -1.35   370   0.00    0    0.00 XXX    0 X    0.00 XXX    0 X    0.17 ASP 339 B

ARG 389 B   11.57    30 %   -0.89   365   0.00    0    0.00 XXX    0 X    0.00 XXX    0 X    0.13 ASP 339 B
ARG 389 B                                  0.00 XXX    0 X    0.00 XXX    0 X    0.18 ASP 471 B
ARG 389 B                                  0.00 XXX    0 X    0.00 XXX    0 X   -0.34 ARG 341 B

ARG 401 B   11.90    73 %   -1.89   486   0.00    0    0.00 XXX    0 X    0.00 XXX    0 X    0.25 ASP  79 B
ARG 401 B                                  0.00 XXX    0 X    0.00 XXX    0 X    0.36 GLU 374 B
ARG 401 B                                  0.00 XXX    0 X    0.00 XXX    0 X    0.67 GLU 373 B

ARG 423 B   14.76    92 %   -2.33   540   0.00    0    0.52 ASP 422 B    0.00 XXX    0 X    2.03 TYR 395 B
ARG 423 B                                  0.00 XXX    0 X    0.00 XXX    0 X    0.18 ASP 430 B
ARG 423 B                                  0.00 XXX    0 X    0.00 XXX    0 X    0.15 TYR 478 B
ARG 423 B                                  0.00 XXX    0 X    0.00 XXX    0 X    1.70 ASP 422 B

ARG 444 B   12.71    41 %   -1.19   397   0.00    0    0.80 GLU 574 B    0.00 XXX    0 X    0.12 GLU 560 B
ARG 444 B                                  0.00 XXX    0 X    0.00 XXX    0 X    0.49 GLU 574 B

ARG 450 B   10.92    48 %   -1.77   416   0.00    0    0.00 XXX    0 X    0.00 XXX    0 X    0.14 TYR 466 B
ARG 450 B                                  0.00 XXX    0 X    0.00 XXX    0 X    0.05 ASP 471 B

ARG 461 B   12.33    39 %   -1.57   391   0.00    0    0.60 GLU 334 B    0.00 XXX    0 X    0.14 TYR 322 B
ARG 461 B                                  0.00 XXX    0 X    0.00 XXX    0 X    0.16 GLU 330 B
ARG 461 B                                  0.00 XXX    0 X    0.00 XXX    0 X    0.50 GLU 334 B

ARG 477 B   12.71   100 %   -3.57   679   0.00    0    0.52 ASP 439 B    0.00 XXX    0 X    0.28 GLU 560 B
ARG 477 B                                  0.00 XXX    0 X    0.00 XXX    0 X    2.03 ASP 439 B
ARG 477 B                                  0.00 XXX    0 X    0.00 XXX    0 X    0.95 ASP 561 B

ARG 480 B   11.82    71 %   -0.85   481   0.00    0    0.00 XXX    0 X    0.00 XXX    0 X    0.16 TYR 395 B
ARG 480 B                                  0.00 XXX    0 X    0.00 XXX    0 X    0.01 ASP 422 B
ARG 480 B                                  0.00 XXX    0 X    0.00 XXX    0 X    0.38 ASP 430 B
ARG 480 B                                  0.00 XXX    0 X    0.00 XXX    0 X    0.08 GLU 601 B
ARG 480 B                                  0.00 XXX    0 X    0.00 XXX    0 X   -0.46 ARG 423 B

ARG 501 B   11.73    29 %   -0.91   362   0.00    0    0.00 XXX    0 X    0.00 XXX    0 X    0.18 ASP 506 B
ARG 501 B                                  0.00 XXX    0 X    0.00 XXX    0 X    0.01 GLU 548 B
ARG 501 B                                  0.00 XXX    0 X    0.00 XXX    0 X   -0.06 ARG 552 B

ARG 536 B   12.63    58 %   -1.85   443   0.00    0    0.61 GLU 459 B    0.00 XXX    0 X    0.04 ASP 456 B
ARG 536 B                                  0.00 XXX    0 X    0.00 XXX    0 X    0.12 GLU 498 B
ARG 536 B                                  0.00 XXX    0 X    0.00 XXX    0 X    0.35 GLU 500 B
ARG 536 B                                  0.00 XXX    0 X    0.00 XXX    0 X    0.06 ASP 542 B
ARG 536 B                                  0.00 XXX    0 X    0.00 XXX    0 X    0.78 GLU 459 B
```

159

```
ARG 547 B   12.49     0 %   -0.28   208    0.00    0    0.00 XXX    0 X    0.00 XXX    0 X    0.13 GLU 543 B
ARG 547 B                                                 0.00 XXX    0 X    0.00 XXX    0 X    0.14 GLU 544 B


ARG 552 B   12.73     0 %   -0.23   222    0.00    0    0.00 XXX    0 X    0.00 XXX    0 X    0.18 ASP 506 B
ARG 552 B                                                 0.00 XXX    0 X    0.00 XXX    0 X    0.28 GLU 548 B


ARG 554 B   13.25     0 %   -0.35   266    0.00    0    0.71 ASP 507 B    0.00 XXX    0 X    0.01 ASP 624 B
ARG 554 B                                                 0.00 XXX    0 X    0.00 XXX    0 X    0.38 ASP 507 B


ARG 586 B   12.35     0 %   -0.28   261    0.00    0    0.00 XXX    0 X    0.00 XXX    0 X    0.14 GLU 620 B


ARG 606 B   12.38     3 %   -0.47   291    0.00    0    0.00 XXX    0 X    0.00 XXX    0 X    0.35 GLU 103 B
ARG 606 B                                                 0.00 XXX    0 X    0.00 XXX    0 X    0.01 GLU 596 B


ARG 615 B   12.80     0 %   -0.46   279    0.00    0    0.22 GLU 628 B    0.00 XXX    0 X    0.07 GLU 525 B
ARG 615 B                                                 0.00 XXX    0 X    0.00 XXX    0 X    0.07 ASP 526 B
ARG 615 B                                                 0.00 XXX    0 X    0.00 XXX    0 X    0.38 GLU 628 B


N+    8 B    8.05     0 %   -0.26   136    0.00    0    0.00 XXX    0 X    0.00 XXX    0 X    0.11 ASP   9 B
N+    8 B                                                 0.00 XXX    0 X    0.00 XXX    0 X    0.14 ASP  14 B
N+    8 B                                                 0.00 XXX    0 X    0.00 XXX    0 X    0.06 GLU  36 B


TDP N1' B   -0.57   100 %   -3.89   578    0.00    0    0.00 XXX    0 X    0.00 XXX    0 X   -0.55 ARG 401 B
TDP N1' B                                                 0.00 XXX    0 X    0.00 XXX    0 X   -0.77 HIS 124 B
TDP N1' B                                                 0.00 XXX    0 X    0.00 XXX    0 X   -0.35 HIS 434 B


TDP O13 B    8.33*  100 %    3.53   644    0.00    0   -0.78 SER 156 B    0.00 XXX    0 X   -3.48 MG   MG B
TDP O13 B                                                 0.00 XXX    0 X    0.00 XXX    0 X   -0.39 LYS 289 B
TDP O13 B                                                 0.00 XXX    0 X    0.00 XXX    0 X    0.07 ASP 182 B
TDP O13 B                                                 0.00 XXX    0 X    0.00 XXX    0 X    0.79 ASP  79 B
TDP O13 B                                                 0.00 XXX    0 X    0.00 XXX    0 X    1.20 ASP 154 B
TDP O13 B                                                 0.00 XXX    0 X    0.00 XXX    0 X    1.38 TDP O21 B


TDP O21 B   -1.66*  100 %    3.87   616    0.00    0   -0.85 SER  54 B    0.00 XXX    0 X   -4.07 MG   MG B
TDP O21 B                                                -0.58 ASN 183 B    0.00 XXX    0 X   -0.15 HIS  51 B
TDP O21 B                                                -0.71 HIS  82 B    0.00 XXX    0 X   -1.23 HIS  82 B
TDP O21 B                                                -0.80 LYS 289 B    0.00 XXX    0 X   -2.03 LYS 289 B
TDP O21 B                                                 0.00 XXX    0 X    0.00 XXX    0 X   -1.12 HIS 304 B
```

Coupled residues (marked *) were detected. Please rerun PropKa with the --display-coupled-residues
or -d option for detailed information.

SUMMARY OF THIS PREDICTION
      Group   pKa   model-pKa   ligand atom-type
    ASP   9 B   3.68   3.80

```
ASP   14 B      1.43          3.80
ASP   21 B      2.19          3.80
ASP   60 B      2.17          3.80
ASP   70 B      3.34          3.80
ASP   74 B      2.18          3.80
ASP   79 B      4.72          3.80
ASP   95 B      3.93          3.80
ASP   99 B      2.94          3.80
ASP  118 B      0.13          3.80
ASP  140 B      2.50          3.80
ASP  145 B      3.72          3.80
ASP  154 B      5.46          3.80
ASP  171 B      3.95          3.80
ASP  182 B      2.62          3.80
ASP  260 B      3.44          3.80
ASP  276 B      3.91          3.80
ASP  278 B      3.67          3.80
ASP  299 B      3.78          3.80
ASP  310 B      3.34          3.80
ASP  339 B      3.08          3.80
ASP  368 B      5.38          3.80
ASP  404 B      4.07          3.80
ASP  409 B      2.72          3.80
ASP  422 B      4.89          3.80
ASP  430 B      0.22          3.80
ASP  439 B      4.77          3.80
ASP  456 B      4.44          3.80
ASP  471 B      3.61          3.80
ASP  493 B      3.31          3.80
ASP  506 B      3.61          3.80
ASP  507 B      3.21          3.80
ASP  518 B      4.66          3.80
ASP  526 B      3.84          3.80
ASP  542 B      4.13          3.80
ASP  561 B      4.23          3.80
ASP  592 B      4.04          3.80
ASP  610 B      4.13          3.80
ASP  624 B      3.93          3.80
GLU   28 B      4.26          4.50
GLU   35 B      4.72          4.50
GLU   36 B      4.72          4.50
GLU   40 B      3.24          4.50
GLU  103 B      4.02          4.50
GLU  114 B      3.93          4.50
GLU  116 B      4.37          4.50
GLU  184 B      5.09          4.50
```

161

```
GLU  189  B        5.03          4.50
GLU  266  B        5.63          4.50
GLU  272  B        3.65          4.50
GLU  297  B        3.97          4.50
GLU  315  B        4.70          4.50
GLU  330  B        4.39          4.50
GLU  334  B        3.20          4.50
GLU  351  B        5.63          4.50
GLU  357  B        4.58          4.50
GLU  373  B        7.98          4.50
GLU  374  B        5.25          4.50
GLU  413  B        4.64          4.50
GLU  459  B        5.19          4.50
GLU  498  B        4.73          4.50
GLU  500  B        3.59          4.50
GLU  525  B        4.56          4.50
GLU  543  B        4.84          4.50
GLU  544  B        4.58          4.50
GLU  548  B        4.43          4.50
GLU  560  B        7.65          4.50
GLU  574  B        3.77          4.50
GLU  593  B        4.70          4.50
GLU  596  B        4.44          4.50
GLU  601  B        4.45          4.50
GLU  620  B        4.62          4.50
GLU  628  B        4.29          4.50
HIS   17  B        6.51          6.50
HIS   51  B        0.43          6.50
HIS   66  B        1.51          6.50
HIS   82  B        4.51          6.50
HIS   87  B        4.64          6.50
HIS  117  B        5.19          6.50
HIS  124  B        2.70          6.50
HIS  147  B        7.66          6.50
HIS  262  B        5.90          6.50
HIS  284  B        6.61          6.50
HIS  304  B        1.71          6.50
HIS  362  B        5.16          6.50
HIS  364  B        6.34          6.50
HIS  408  B        7.47          6.50
HIS  414  B        5.56          6.50
HIS  434  B        4.76          6.50
HIS  470  B        5.59          6.50
HIS  582  B        6.27          6.50
HIS  597  B        5.56          6.50
HIS  604  B        5.95          6.50
```

162

```
CYS   45 B     12.62         9.00
CYS  420 B     12.82         9.00
TYR   67 B     10.59        10.00
TYR   85 B     13.82        10.00
TYR  255 B     11.64        10.00
TYR  295 B     10.62        10.00
TYR  302 B     11.37        10.00
TYR  316 B     10.51        10.00
TYR  322 B     12.12        10.00
TYR  366 B     12.68        10.00
TYR  395 B     12.09        10.00
TYR  403 B     10.37        10.00
TYR  466 B     10.67        10.00
TYR  478 B     16.80        10.00
TYR  519 B     12.32        10.00
LYS   20 B     10.29        10.50
LYS   23 B      8.57        10.50
LYS   88 B      8.66        10.50
LYS  101 B     10.50        10.50
LYS  102 B     10.84        10.50
LYS  111 B     10.41        10.50
LYS  144 B     10.28        10.50
LYS  175 B     10.63        10.50
LYS  196 B      9.82        10.50
LYS  289 B      8.42        10.50
LYS  291 B     10.57        10.50
LYS  308 B     10.37        10.50
LYS  337 B     10.52        10.50
LYS  455 B     11.53        10.50
LYS  465 B     10.84        10.50
LYS  495 B     10.68        10.50
LYS  503 B     10.11        10.50
LYS  515 B     10.25        10.50
LYS  522 B     10.60        10.50
LYS  539 B      9.01        10.50
ARG   24 B     13.22        12.50
ARG   27 B     13.34        12.50
ARG   38 B     13.02        12.50
ARG   43 B     12.72        12.50
ARG   47 B     11.77        12.50
ARG   73 B     13.49        12.50
ARG   75 B     11.63        12.50
ARG   93 B     13.83        12.50
ARG   94 B     15.05        12.50
ARG  139 B     12.36        12.50
ARG  174 B     12.96        12.50
```

163

```
      ARG 199 B    12.42      12.50
      ARG 254 B    12.04      12.50
      ARG 273 B    12.77      12.50
      ARG 341 B    12.46      12.50
      ARG 350 B    13.50      12.50
      ARG 360 B    12.30      12.50
      ARG 365 B    11.33      12.50
      ARG 389 B    11.57      12.50
      ARG 401 B    11.90      12.50
      ARG 423 B    14.76      12.50
      ARG 444 B    12.71      12.50
      ARG 450 B    10.92      12.50
      ARG 461 B    12.33      12.50
      ARG 477 B    12.71      12.50
      ARG 480 B    11.82      12.50
      ARG 501 B    11.73      12.50
      ARG 536 B    12.63      12.50
      ARG 547 B    12.49      12.50
      ARG 552 B    12.73      12.50
      ARG 554 B    13.25      12.50
      ARG 586 B    12.35      12.50
      ARG 606 B    12.38      12.50
      ARG 615 B    12.80      12.50
      N+    8 B     8.05       8.00
      TDP N1' B    -0.57       5.00                NAR
      TDP O13 B     8.33       6.00                OP
      TDP O21 B    -1.66       6.00                OP
 -------------------------------------------------------------------------------------------------------
 -------------------------------------------------------------------------------------------------------
 Free energy of   folding (kcal/mol) as a function of pH (using neutral reference)
    0.00    124.42
    1.00    117.04
    2.00    104.84
    3.00     86.18
    4.00     66.66
    5.00     57.47
    6.00     50.36
    7.00     47.60
    8.00     50.78
    9.00     52.28
   10.00     54.93
   11.00     63.98
   12.00     72.65
   13.00     80.85
   14.00     88.96
```

164

```
The pH of optimum stability is  6.8 for which the free energy is  47.4 kcal/mol at 298K
Could not determine pH values where the free energy is within 80 % of minimum
Could not determine the pH-range where the free energy is negative

Protein charge of folded and unfolded state as a function of pH
     pH   unfolded   folded
   0.00      76.99    73.05
   1.00      76.93    69.92
   2.00      76.28    65.06
   3.00      70.59    54.84
   4.00      44.63    33.90
   5.00      12.51     7.85
   6.00      -3.35    -8.48
   7.00     -16.81   -15.46
   8.00     -22.21   -20.12
   9.00     -25.65   -25.07
  10.00     -36.21   -31.88
  11.00     -53.03   -45.35
  12.00     -65.42   -60.25
  13.00     -83.76   -76.79
  14.00     -90.95   -86.50
The pI is  5.40 (folded) and  5.80 (unfolded)
```

Appendix B: Supporting Information for "Computational Examination of the Magnesium Ion Binding Modes of 1-Deoxy-D-xylulose 5-Phosphate Reductoisomerase"

B.1 Results from the ProPKA3.0 Calculations of a DXR with Substrates Bound

Below will be found the ProPKA3.0 results for of crystal structure, PDB:1Q0Q, for DXR with bound NADPH and DXP:

```
propka3.0, revision 182                                                         2015-01-13
-------------------------------------------------------------------------------------------
--                                                                                       --
--                          PROPKA: A PROTEIN PKA PREDICTOR                               --
--                                                                                       --
--                        VERSION 1.0,  04/25/2004, IOWA CITY                            --
--                                   BY HUI LI                                           --
--                                                                                       --
--                        VERSION 2.0,  11/05/2007, IOWA CITY/COPENHAGEN                 --
--                        BY DELPHINE C. BAS AND DAVID M. ROGERS                         --
--                                                                                       --
--                        VERSION 3.0,  xx/xx/2010, COPENHAGEN                           --
--                        BY MATS H.M. OLSSON AND CHRESTEN R. SONDERGARD                 --
--                                                                                       --
-------------------------------------------------------------------------------------------


-------------------------------------------------------------------------------------------
 References:

   Very Fast Empirical Prediction and Rationalization of Protein pKa Values
   Hui Li, Andrew D. Robertson and Jan H. Jensen
   PROTEINS: Structure, Function, and Bioinformatics 61:704-721 (2005)

   Very Fast Prediction and Rationalization of pKa Values for Protein-Ligand Complexes
   Delphine C. Bas, David M. Rogers and Jan H. Jensen
   PROTEINS: Structure, Function, and Bioinformatics 73:765-783 (2008)

   PROPKA3: Consistent Treatment of Internal and Surface Residues in Empirical pKa predictions
   Mats H.M. Olsson, Chresten R. Sondergard, Michal Rostkowski, and Jan H. Jensen
   Journal of Chemical Theory and Computation, to be submitted (2010)
-------------------------------------------------------------------------------------------


-------------------------------------------------------------------------------------------
 WARNING !

   Propka3.0 is not identical to propka2.0 and does not work with ligands
-------------------------------------------------------------------------------------------
```

RESIDUE	pKa	BURIED	DESOLVATION EFFECTS REGULAR RE		SIDECHAIN HYDROGEN BOND	BACKBONE HYDROGEN BOND	COULOMBIC INTERACTION
ASP 19 A	3.18	29 %	1.00 362	0.13 0	-0.48 ARG 22 A	0.00 XXX 0 X	-0.38 HIS 23 A

167

```
ASP   19 A                                           -0.11 ARG 277 A       0.00 XXX   0 X      -0.33 ARG  22 A
ASP   19 A                                            0.00 XXX   0 X       0.00 XXX   0 X      -0.47 ARG 277 A

ASP   57 A   4.68   27 %   0.80  356   0.01    0      0.00 XXX   0 X       0.00 XXX   0 X       0.08 ASP  58 A

ASP   58 A   2.97    0 %   0.38  241   0.00    0     -0.39 SER  61 A      -0.20 ALA  60 A      -0.07 LYS  37 A
ASP   58 A                                            0.00 XXX   0 X      -0.54 SER  61 A       0.00 XXX   0 X

ASP   87 A   3.70    1 %   0.38  284   0.01    0     -0.50 GLN  83 A       0.00 XXX   0 X       0.00 XXX   0 X

ASP   93 A   3.29    0 %   0.26  215   0.00    0     -0.36 ARG  52 A       0.00 XXX   0 X      -0.10 ARG  29 A
ASP   93 A                                            0.00 XXX   0 X       0.00 XXX   0 X      -0.05 LYS 118 A
ASP   93 A                                            0.00 XXX   0 X       0.00 XXX   0 X      -0.25 ARG  52 A

ASP   95 A   3.65   17 %   0.65  330   0.24    0      0.00 XXX   0 X      -0.81 GLN   3 A      -0.06 LYS   2 A
ASP   95 A                                            0.00 XXX   0 X       0.00 XXX   0 X      -0.02 ARG  29 A
ASP   95 A                                            0.00 XXX   0 X       0.00 XXX   0 X      -0.14 LYS 118 A
ASP   95 A                                            0.00 XXX   0 X       0.00 XXX   0 X      -0.08 N+    1 A
ASP   95 A                                            0.00 XXX   0 X       0.00 XXX   0 X       0.06 ASP  93 A

ASP  137 A   3.26   10 %   0.65  309   0.07    0     -0.85 TYR 170 A       0.00 XXX   0 X      -0.28 ARG 133 A
ASP  137 A                                            0.00 XXX   0 X       0.00 XXX   0 X      -0.13 LYS 140 A

ASP  150 A   1.23* 100 %   4.83  738   0.26    0     -0.85 LYS 125 A      -0.03 SER 151 A      -2.03 LYS 125 A
ASP  150 A                                           -0.56 GLU 152 A       0.00 XXX   0 X      -1.69 HIS 153 A
ASP  150 A                                           -0.57 GLU 231 A       0.00 XXX   0 X      -0.59 LYS 228 A
ASP  150 A                                           -1.34 GLU 234 A       0.00 XXX   0 X       0.00 XXX   0 X

ASP  172 A   3.56    0 %   0.36  272   0.00    0      0.00 XXX   0 X      -0.03 GLU 174 A       0.00 XXX   0 X
ASP  172 A                                            0.00 XXX   0 X      -0.57 GLN 175 A       0.00 XXX   0 X

ASP  197 A   4.06    0 %   0.26  157   0.00    0      0.00 XXX   0 X       0.00 XXX   0 X       0.00 XXX   0 X

ASP  204 A   3.06    0 %   0.22  214   0.00    0     -0.65 ARG 216 A       0.00 XXX   0 X      -0.31 ARG 216 A

ASP  221 A   4.58  100 %   3.64  679   0.94    0     -0.84 ASN 339 A       0.00 XXX   0 X      -0.29 LYS 125 A
ASP  221 A                                           -0.85 LYS 217 A       0.00 XXX   0 X       0.22 GLU 340 A
ASP  221 A                                            0.00 XXX   0 X       0.00 XXX   0 X      -2.03 LYS 217 A

ASP  264 A   4.07  100 %   3.58  601   0.78    0     -0.84 SER 266 A      -0.08 ASP 264 A      -1.86 ARG 289 B
ASP  264 A                                           -0.50 ARG 289 B      -0.82 SER 266 A       0.00 XXX   0 X

ASP  275 A   4.13   62 %   1.88  455   0.00    0     -0.39 ARG 277 A      -0.68 ARG 277 A      -0.04 LYS 295 A
ASP  275 A                                            0.00 XXX   0 X       0.00 XXX   0 X       0.11 ASP  19 A
ASP  275 A                                            0.00 XXX   0 X       0.00 XXX   0 X      -0.01 HIS  23 A
ASP  275 A                                            0.00 XXX   0 X       0.00 XXX   0 X      -0.55 ARG 277 A
```

168

```
     ASP 298 A    3.40      0 %    0.36   261    0.00      0     0.00 XXX    0 X    -0.46 CYS 300 A    -0.03 LYS 295 A
     ASP 298 A                                                   0.00 XXX    0 X    -0.15 LYS 301 A    -0.08 LYS 301 A
     ASP 298 A                                                   0.00 XXX    0 X     0.00 XXX    0 X    -0.09 ARG 261 B
     ASP 298 A                                                   0.00 XXX    0 X     0.00 XXX    0 X     0.05 GLU 247 B

     ASP 311 A    3.36      0 %    0.37   264    0.00      0     0.00 XXX    0 X    -0.21 ASP 313 A    -0.21 ARG 314 A
     ASP 311 A                                                   0.00 XXX    0 X    -0.39 ARG 314 A     0.00 XXX    0 X

     ASP 313 A    3.96      0 %    0.16   173    0.00      0     0.00 XXX    0 X     0.00 XXX    0 X    -0.19 LYS 319 A
     ASP 313 A                                                   0.00 XXX    0 X     0.00 XXX    0 X     0.19 ASP 311 A

     ASP 355 A    3.98     28 %    1.05   361    0.13      0     0.00 XXX    0 X    -0.81 ARG 352 A    -0.19 ARG 352 A

     ASP 368 A    3.91      0 %    0.14   192    0.00      0     0.00 XXX    0 X     0.00 XXX    0 X    -0.03 ARG 370 A

     ASP 376 A    4.05      0 %    0.17   194    0.00      0     0.00 XXX    0 X     0.00 XXX    0 X     0.08 ASP 377 A

     ASP 377 A    3.27      0 %    0.36   214    0.00      0    -0.12 CYS 374 A    -0.76 GLN 373 A    -0.00 ARG 370 A

     ASP 382 A    4.83     67 %    2.05   469    0.38      0    -1.12 ARG 386 A     0.00 XXX    0 X    -0.05 ARG 390 A
     ASP 382 A                                                   0.00 XXX    0 X     0.00 XXX    0 X     0.59 GLU 340 A
     ASP 382 A                                                   0.00 XXX    0 X     0.00 XXX    0 X    -0.83 ARG 386 A

     GLU  26 A    4.44      0 %    0.08   119    0.00      0     0.00 XXX    0 X     0.00 XXX    0 X    -0.03 N+     1 A
     GLU  26 A                                                   0.00 XXX    0 X     0.00 XXX    0 X    -0.11 HIS   27 A

     GLU  44 A    4.37      1 %    0.36   283    0.00      0    -0.11 ARG  41 A     0.00 XXX    0 X    -0.10 ARG   22 A
     GLU  44 A                                                   0.00 XXX    0 X     0.00 XXX    0 X     0.12 GLU   48 A
     GLU  44 A                                                   0.00 XXX    0 X     0.00 XXX    0 X    -0.41 ARG   41 A

     GLU  48 A    3.41      9 %    0.70   306    0.01      0    -1.40 ARG  22 A     0.00 XXX    0 X    -0.04 ARG   41 A
     GLU  48 A                                                   0.00 XXX    0 X     0.00 XXX    0 X     0.06 ASP   19 A
     GLU  48 A                                                   0.00 XXX    0 X     0.00 XXX    0 X    -0.43 ARG   22 A

     GLU  59 A    4.50      0 %    0.17   166    0.00      0     0.00 XXX    0 X     0.00 XXX    0 X    -0.21 LYS   63 A
     GLU  59 A                                                   0.00 XXX    0 X     0.00 XXX    0 X     0.04 ASP   58 A

     GLU  77 A    4.52      0 %    0.17   226    0.00      0     0.00 XXX    0 X     0.00 XXX    0 X    -0.11 LYS   66 A
     GLU  77 A                                                   0.00 XXX    0 X     0.00 XXX    0 X    -0.04 ARG   75 A

     GLU  92 A    4.61      0 %    0.09   164    0.00      0     0.00 XXX    0 X     0.00 XXX    0 X    -0.14 LYS  118 A
     GLU  92 A                                                   0.00 XXX    0 X     0.00 XXX    0 X     0.14 ASP   93 A
     GLU  92 A                                                   0.00 XXX    0 X     0.00 XXX    0 X     0.01 ASP   95 A

     GLU 126 A    6.06    100 %    3.88   645    0.15      0    -0.68 LYS 217 A    -0.35 GLY 103 A    -0.00 ARG 386 A
```

169

```
  GLU 126 A                                       0.00 XXX    0 X   -0.65 GLU 126 A    0.17 ASP 150 A
  GLU 126 A                                       0.00 XXX    0 X    0.00 XXX    0 X   -0.62 LYS 125 A
  GLU 126 A                                       0.00 XXX    0 X    0.00 XXX    0 X   -1.97 LYS 217 A
  GLU 126 A                                       0.00 XXX    0 X    0.00 XXX    0 X    0.74 ASP 221 A
  GLU 126 A                                       0.00 XXX    0 X    0.00 XXX    0 X    0.90 GLU 340 A

  GLU 152 A  10.96*  100 %   5.00  712   0.81   0  -0.77 LYS 228 A   -0.74 GLU 152 A   -1.03 LYS 125 A
  GLU 152 A                                       0.56 ASP 150 A    0.00 XXX    0 X   -2.03 LYS 228 A
  GLU 152 A                                       1.12 GLU 231 A    0.00 XXX    0 X    1.50 ASP 150 A
  GLU 152 A                                       0.00 XXX    0 X    0.00 XXX    0 X    2.03 GLU 231 A

  GLU 174 A   4.82    0 %   0.20  279   0.00   0   0.00 XXX    0 X    0.00 XXX    0 X    0.12 ASP 172 A

  GLU 192 A   4.49    0 %   0.14  235   0.00   0   0.00 XXX    0 X    0.00 XXX    0 X   -0.13 ARG 191 A
  GLU 192 A                                       0.00 XXX    0 X    0.00 XXX    0 X   -0.02 ARG 314 A

  GLU 231 A   4.51*  100 %   5.25  721   0.66   0  -0.50 ASN 227 A    0.00 XXX    0 X   -2.03 LYS 125 A
  GLU 231 A                                      -0.85 LYS 125 A    0.00 XXX    0 X    2.03 ASP 150 A
  GLU 231 A                                       0.57 ASP 150 A    0.00 XXX    0 X   -1.11 HIS 153 A
  GLU 231 A                                      -1.12 GLU 152 A    0.00 XXX    0 X   -2.03 LYS 228 A
  GLU 231 A                                      -0.85 LYS 228 A    0.00 XXX    0 X    0.00 XXX    0 X

  GLU 234 A  12.81*  100 %   4.27  713   0.66   0   1.34 ASP 150 A    0.00 XXX    0 X   -1.17 LYS 125 A
  GLU 234 A                                       0.00 XXX    0 X    0.00 XXX    0 X   -0.05 LYS 228 A
  GLU 234 A                                       0.00 XXX    0 X    0.00 XXX    0 X    0.18 GLU 126 A
  GLU 234 A                                       0.00 XXX    0 X    0.00 XXX    0 X    0.33 GLU 152 A
  GLU 234 A                                       0.00 XXX    0 X    0.00 XXX    0 X    2.03 ASP 150 A
  GLU 234 A                                       0.00 XXX    0 X    0.00 XXX    0 X    0.72 GLU 231 A

  GLU 247 A   3.44   61 %   1.67  451   0.05   0  -0.78 SER 180 A    0.00 XXX    0 X    0.05 ASP 298 B
  GLU 247 A                                      -1.21 ARG 261 A    0.00 XXX    0 X   -0.84 ARG 261 A

  GLU 273 A   5.10   30 %   0.86  365   0.06   0   0.00 XXX    0 X    0.00 XXX    0 X   -0.32 LYS 295 A
  GLU 273 A                                       0.00 XXX    0 X    0.00 XXX    0 X   -0.12 LYS 301 A
  GLU 273 A                                       0.00 XXX    0 X    0.00 XXX    0 X   -0.00 ARG 261 B
  GLU 273 A                                       0.00 XXX    0 X    0.00 XXX    0 X    0.09 ASP 275 A
  GLU 273 A                                       0.00 XXX    0 X    0.00 XXX    0 X    0.04 ASP 298 A

  GLU 323 A   4.69    6 %   0.36  297   0.03   0   0.00 XXX    0 X    0.00 XXX    0 X   -0.20 LYS 319 A

  GLU 326 A   3.96    2 %   0.35  288   0.01   0  -0.53 ARG 236 A    0.00 XXX    0 X   -0.36 ARG 236 A

  GLU 340 A   3.88   100 %   2.82  569   0.67   0  -0.44 ASN 336 A   -0.77 ALA 104 A   -1.09 LYS 217 A
  GLU 340 A                                      -0.60 ARG 386 A    0.00 XXX    0 X   -1.21 ARG 386 A

  GLU 365 A   4.74    0 %   0.28  234   0.00   0   0.00 XXX    0 X    0.00 XXX    0 X   -0.04 LYS 366 A
```

170

```
GLU 371 A    3.97    19 %    0.78   334    0.06     0     0.00 XXX    0 X    -0.61 GLN 329 A    -0.04 ARG 370 A
GLU 371 A                                                  0.00 XXX    0 X    -0.73 ALA 330 A     0.00 XXX    0 X

GLU 387 A    4.63     0 %    0.20   212    0.00     0     0.00 XXX    0 X     0.00 XXX    0 X    -0.07 ARG 390 A

GLU 392 A    4.20     2 %    0.60   287    0.01     0    -0.04 SER 362 A     0.00 XXX    0 X    -0.32 LYS 366 A
GLU 392 A                                               -0.02 ARG 395 A     0.00 XXX    0 X    -0.14 LYS 391 A
GLU 392 A                                                  0.00 XXX    0 X     0.00 XXX    0 X    -0.38 ARG 395 A

C-  398 A    3.31     0 %    0.11    95    0.00     0     0.00 XXX    0 X     0.00 XXX    0 X     0.00 XXX    0 X

HIS  23 A    6.08     8 %   -0.51   305    0.00     0     0.00 XXX    0 X     0.00 XXX    0 X    -0.08 ARG  22 A
HIS  23 A                                                  0.00 XXX    0 X     0.00 XXX    0 X    -0.22 ARG 277 A
HIS  23 A                                                  0.00 XXX    0 X     0.00 XXX    0 X     0.38 ASP  19 A
HIS  23 A                                                  0.00 XXX    0 X     0.00 XXX    0 X     0.01 ASP 275 A

HIS  27 A    6.02     7 %   -0.39   301    0.00     0     0.00 XXX    0 X     0.00 XXX    0 X    -0.08 LYS   2 A
HIS  27 A                                                  0.00 XXX    0 X     0.00 XXX    0 X    -0.11 N+    1 A
HIS  27 A                                                  0.00 XXX    0 X     0.00 XXX    0 X     0.11 GLU  26 A

HIS 153 A    5.02   100 %   -3.95   718    0.00     0     0.00 XXX    0 X     0.77 HIS 153 A    -0.74 LYS 125 A
HIS 153 A                                                  0.00 XXX    0 X     0.00 XXX    0 X    -0.35 LYS 228 A
HIS 153 A                                                  0.00 XXX    0 X     0.00 XXX    0 X     1.69 ASP 150 A
HIS 153 A                                                  0.00 XXX    0 X     0.00 XXX    0 X     1.11 GLU 231 A

HIS 166 A    5.89    18 %   -0.61   331    0.00     0     0.00 XXX    0 X     0.00 XXX    0 X     0.00 XXX    0 X

HIS 209 A    2.93    86 %   -3.09   522    0.00     0     0.00 XXX    0 X     0.00 XXX    0 X    -0.20 ARG 191 A
HIS 209 A                                                  0.00 XXX    0 X     0.00 XXX    0 X    -0.28 LYS 228 A

HIS 251 A    2.88   100 %   -3.33   569    0.00     0     0.00 XXX    0 X     0.00 XXX    0 X    -0.02 ARG 191 A
HIS 251 A                                                  0.00 XXX    0 X     0.00 XXX    0 X    -0.26 HIS 257 A

HIS 257 A    3.99   100 %   -3.13   667    0.00     0     0.29 GLN 270 A     0.23 HIS 257 A    -0.00 LYS 228 A
HIS 257 A                                                  0.00 XXX    0 X     0.11 HIS 257 A     0.00 XXX    0 X

HIS 282 A    2.01   100 %   -3.10   641    0.00     0     0.00 XXX    0 X     0.00 XXX    0 X    -1.39 ARG 289 A

CYS  15 A   10.34    61 %    1.82   453    0.00     0    -0.02 ARG  41 A     0.00 XXX    0 X    -0.09 ARG  22 A
CYS  15 A                                                  0.00 XXX    0 X     0.00 XXX    0 X    -0.55 ARG  41 A
CYS  15 A                                                  0.00 XXX    0 X     0.00 XXX    0 X     0.02 ASP  19 A
CYS  15 A                                                  0.00 XXX    0 X     0.00 XXX    0 X     0.15 GLU  44 A
CYS  15 A                                                  0.00 XXX    0 X     0.00 XXX    0 X     0.01 GLU  48 A

CYS  46 A   11.92    80 %    2.97   504    0.00     0    -0.05 THR  76 A     0.00 XXX    0 X     0.00 XXX    0 X
```

```
CYS  86 A  10.56    42 %    1.59  399    0.00     0    0.00 XXX   0 X   -0.03 CYS  86 A   -0.15 ARG 115 A
CYS  86 A                                                0.00 XXX   0 X    0.00 XXX   0 X    0.14 ASP  87 A

CYS 131 A  11.75    80 %    2.75  505    0.00     0    0.00 XXX   0 X   -0.02 PHE 135 A    0.01 ASP 137 A
CYS 131 A                                                0.00 XXX   0 X    0.00 XXX   0 X    0.02 CYS 374 A

CYS 207 A  10.78    55 %    2.19  436    0.00     0    0.00 XXX   0 X    0.00 XXX   0 X   -0.53 ARG 216 A
CYS 207 A                                                0.00 XXX   0 X    0.00 XXX   0 X    0.13 ASP 204 A

CYS 300 A  10.02    17 %    0.74  329    0.00     0    0.00 XXX   0 X    0.00 XXX   0 X   -0.03 ARG 261 B
CYS 300 A                                                0.00 XXX   0 X    0.00 XXX   0 X    0.26 ASP 298 A
CYS 300 A                                                0.00 XXX   0 X    0.00 XXX   0 X    0.04 GLU 247 B

CYS 317 A  11.85   100 %    3.20  588    0.00     0   -0.40 THR 224 A    0.00 XXX   0 X   -0.00 LYS 217 A
CYS 317 A                                               -0.32 ASN 360 A    0.00 XXX   0 X    0.38 ASP 221 A

CYS 374 A   9.07     0 %    0.46  224    0.00     0    0.12 ASP 377 A   -0.34 ASP 376 A    0.20 ASP 376 A
CYS 374 A                                                0.00 XXX   0 X   -0.75 ASP 377 A    0.38 ASP 377 A

TYR  53 A  10.17     0 %    0.28  268    0.00     0   -0.05 ARG  52 A    0.00 XXX   0 X   -0.21 ARG  52 A
TYR  53 A                                                0.00 XXX   0 X    0.00 XXX   0 X    0.14 GLU  77 A
TYR  53 A                                                0.00 XXX   0 X    0.00 XXX   0 X    0.01 ASP  93 A

TYR 170 A  11.71    22 %    1.23  342    0.00     0   -0.34 ARG 133 A    0.00 XXX   0 X   -0.47 ARG 133 A
TYR 170 A                                                0.85 ASP 137 A    0.00 XXX   0 X    0.44 ASP 137 A

TYR 232 A  10.66    50 %    1.63  422    0.00     0   -0.66 ARG 236 A    0.00 XXX   0 X   -0.56 ARG 236 A
TYR 232 A                                                0.00 XXX   0 X    0.00 XXX   0 X    0.25 GLU 326 A

TYR 262 A  12.82   100 %    3.71  631    0.00     0    0.00 XXX   0 X    0.00 XXX   0 X    0.13 ASP 264 A
TYR 262 A                                                0.00 XXX   0 X    0.00 XXX   0 X   -1.02 ARG 289 B

TYR 312 A  11.37    48 %    1.26  416    0.00     0    0.00 XXX   0 X    0.00 XXX   0 X    0.10 TYR 232 A

TYR 315 A  11.98    70 %    2.28  476    0.00     0    0.00 XXX   0 X    0.00 XXX   0 X   -0.15 ARG 191 A
TYR 315 A                                                0.00 XXX   0 X    0.00 XXX   0 X    0.01 GLU 192 A
TYR 315 A                                                0.00 XXX   0 X    0.00 XXX   0 X    0.10 ASP 311 A
TYR 315 A                                                0.00 XXX   0 X    0.00 XXX   0 X   -0.26 ARG 314 A

LYS   2 A   9.13    54 %   -1.43  432    0.00     0    0.00 XXX   0 X    0.00 XXX   0 X    0.06 ASP  95 A

LYS  37 A  10.33     9 %   -0.25  306    0.00     0    0.00 XXX   0 X    0.00 XXX   0 X    0.07 ASP  58 A

LYS  63 A  10.62     0 %   -0.09  136    0.00     0    0.00 XXX   0 X    0.00 XXX   0 X    0.21 GLU  59 A
```

172

```
LYS  66 A  10.36    0 %   -0.25  250   0.00     0   0.00 XXX   0 X   0.00 XXX   0 X    0.11 GLU  77 A

LYS 118 A   9.81   22 %   -1.02  344   0.00     0   0.00 XXX   0 X   0.00 XXX   0 X    0.14 GLU  92 A
LYS 118 A                                           0.00 XXX   0 X   0.00 XXX   0 X    0.05 ASP  93 A
LYS 118 A                                           0.00 XXX   0 X   0.00 XXX   0 X    0.14 ASP  95 A

LYS 125 A  13.89  100 %   -5.49  702   0.00     0   0.85 ASP 150 A   0.00 XXX   0 X    1.03 GLU 152 A
LYS 125 A                                           0.85 GLU 231 A   0.00 XXX   0 X    0.29 ASP 221 A
LYS 125 A                                           0.00 XXX   0 X   0.00 XXX   0 X    1.17 GLU 234 A
LYS 125 A                                           0.00 XXX   0 X   0.00 XXX   0 X    0.62 GLU 126 A
LYS 125 A                                           0.00 XXX   0 X   0.00 XXX   0 X    2.03 ASP 150 A
LYS 125 A                                           0.00 XXX   0 X   0.00 XXX   0 X    2.03 GLU 231 A

LYS 140 A  10.29    0 %   -0.27  212   0.00     0   0.00 XXX   0 X   0.00 XXX   0 X   -0.07 ARG 133 A
LYS 140 A                                           0.00 XXX   0 X   0.00 XXX   0 X    0.13 ASP 137 A

LYS 143 A  10.10    0 %   -0.30  184   0.00     0   0.00 XXX   0 X   0.00 XXX   0 X   -0.10 LYS 140 A

LYS 217 A  13.30  100 %   -3.47  645   0.00     0   0.68 GLU 126 A   0.00 XXX   0 X    0.00 CYS 317 A
LYS 217 A                                           0.85 ASP 221 A   0.00 XXX   0 X   -0.12 ARG 386 A
LYS 217 A                                           0.00 XXX   0 X   0.00 XXX   0 X   -0.24 LYS 125 A
LYS 217 A                                           0.00 XXX   0 X   0.00 XXX   0 X    1.97 GLU 126 A
LYS 217 A                                           0.00 XXX   0 X   0.00 XXX   0 X    2.03 ASP 221 A
LYS 217 A                                           0.00 XXX   0 X   0.00 XXX   0 X    1.09 GLU 340 A

LYS 228 A  10.72  100 %   -5.43  704   0.00     0   0.77 GLU 152 A   0.00 XXX   0 X    2.03 GLU 152 A
LYS 228 A                                           0.85 GLU 231 A   0.00 XXX   0 X    0.05 GLU 234 A
LYS 228 A                                           0.00 XXX   0 X   0.00 XXX   0 X   -0.68 LYS 125 A
LYS 228 A                                           0.00 XXX   0 X   0.00 XXX   0 X    0.59 ASP 150 A
LYS 228 A                                           0.00 XXX   0 X   0.00 XXX   0 X    2.03 GLU 231 A

LYS 295 A  10.71    0 %   -0.17  232   0.00     0   0.00 XXX   0 X   0.00 XXX   0 X    0.32 GLU 273 A
LYS 295 A                                           0.00 XXX   0 X   0.00 XXX   0 X    0.04 ASP 275 A
LYS 295 A                                           0.00 XXX   0 X   0.00 XXX   0 X   -0.01 ARG 277 A
LYS 295 A                                           0.00 XXX   0 X   0.00 XXX   0 X    0.03 ASP 298 A

LYS 301 A  10.43    0 %   -0.12  199   0.00     0   0.00 XXX   0 X   0.00 XXX   0 X    0.12 GLU 273 A
LYS 301 A                                           0.00 XXX   0 X   0.00 XXX   0 X    0.08 ASP 298 A
LYS 301 A                                           0.00 XXX   0 X   0.00 XXX   0 X   -0.15 LYS 295 A

LYS 319 A  10.72    0 %   -0.17  209   0.00     0   0.00 XXX   0 X   0.00 XXX   0 X    0.19 ASP 313 A
LYS 319 A                                           0.00 XXX   0 X   0.00 XXX   0 X    0.20 GLU 323 A

LYS 366 A  10.40    0 %   -0.23  216   0.00     0   0.00 XXX   0 X   0.00 XXX   0 X    0.04 GLU 365 A
LYS 366 A                                           0.00 XXX   0 X   0.00 XXX   0 X    0.32 GLU 392 A
LYS 366 A                                           0.00 XXX   0 X   0.00 XXX   0 X   -0.24 ARG 395 A
```

```
LYS 391 A  10.17    0 %   -0.13  174   0.00    0   0.00 XXX   0 X   0.00 XXX   0 X   0.14 GLU 392 A
LYS 391 A                                           0.00 XXX   0 X   0.00 XXX   0 X  -0.15 ARG 395 A
LYS 391 A                                           0.00 XXX   0 X   0.00 XXX   0 X  -0.19 LYS 366 A

ARG  22 A  14.57   20 %   -0.76  337   0.00    0   0.48 ASP  19 A   0.00 XXX   0 X   0.09 CYS  15 A
ARG  22 A                                           1.40 GLU  48 A   0.00 XXX   0 X   0.10 GLU  44 A
ARG  22 A                                           0.00 XXX   0 X   0.00 XXX   0 X   0.33 ASP  19 A
ARG  22 A                                           0.00 XXX   0 X   0.00 XXX   0 X   0.43 GLU  48 A

ARG  29 A  11.94   12 %   -0.53  314   0.00    0   0.00 XXX   0 X   0.00 XXX   0 X   0.10 ASP  93 A
ARG  29 A                                           0.00 XXX   0 X   0.00 XXX   0 X   0.02 ASP  95 A
ARG  29 A                                           0.00 XXX   0 X   0.00 XXX   0 X  -0.16 ARG  52 A

ARG  41 A  13.08   14 %   -0.45  321   0.00    0   0.02 CYS  15 A   0.00 XXX   0 X   0.55 CYS  15 A
ARG  41 A                                           0.11 GLU  44 A   0.00 XXX   0 X   0.04 GLU  48 A
ARG  41 A                                           0.00 XXX   0 X   0.00 XXX   0 X  -0.09 ARG  22 A
ARG  41 A                                           0.00 XXX   0 X   0.00 XXX   0 X   0.41 GLU  44 A

ARG  52 A  12.96    2 %   -0.41  286   0.00    0   0.05 TYR  53 A   0.00 XXX   0 X   0.21 TYR  53 A
ARG  52 A                                           0.36 ASP  93 A   0.00 XXX   0 X   0.25 ASP  93 A

ARG  75 A  12.46    0 %   -0.08  166   0.00    0   0.00 XXX   0 X   0.00 XXX   0 X   0.04 GLU  77 A

ARG 115 A  12.16   18 %   -0.49  331   0.00    0   0.00 XXX   0 X   0.00 XXX   0 X   0.15 CYS  86 A

ARG 133 A  12.71   31 %   -0.89  367   0.00    0   0.34 TYR 170 A   0.00 XXX   0 X   0.28 ASP 137 A
ARG 133 A                                           0.00 XXX   0 X   0.00 XXX   0 X   0.47 TYR 170 A

ARG 191 A  11.47   47 %   -1.30  414   0.00    0   0.00 XXX   0 X   0.00 XXX   0 X   0.13 GLU 192 A
ARG 191 A                                           0.00 XXX   0 X   0.00 XXX   0 X   0.15 TYR 315 A

ARG 196 A  12.41    0 %   -0.09  117   0.00    0   0.00 XXX   0 X   0.00 XXX   0 X   0.00 XXX   0 X

ARG 208 A  12.42    0 %   -0.08  172   0.00    0   0.00 XXX   0 X   0.00 XXX   0 X   0.00 XXX   0 X

ARG 216 A  13.46   13 %   -0.53  318   0.00    0   0.65 ASP 204 A   0.00 XXX   0 X   0.53 CYS 207 A
ARG 216 A                                           0.00 XXX   0 X   0.00 XXX   0 X   0.31 ASP 204 A

ARG 236 A  13.16   43 %   -1.46  402   0.00    0   0.66 TYR 232 A   0.00 XXX   0 X   0.56 TYR 232 A
ARG 236 A                                           0.53 GLU 326 A   0.00 XXX   0 X   0.36 GLU 326 A

ARG 261 A  12.56   73 %   -2.14  485   0.00    0   1.21 GLU 247 A   0.00 XXX   0 X   0.03 GLU 273 B
ARG 261 A                                           0.00 XXX   0 X   0.00 XXX   0 X   0.10 ASP 298 B
ARG 261 A                                           0.00 XXX   0 X   0.00 XXX   0 X   0.02 CYS 300 B
ARG 261 A                                           0.00 XXX   0 X   0.00 XXX   0 X   0.84 GLU 247 A
```

174

```
ARG 277 A  13.08    28 %   -0.90  359   0.00    0    0.11 ASP  19 A   0.00 XXX   0 X   -0.04 ARG  22 A
ARG 277 A                                                0.39 ASP 275 A   0.00 XXX   0 X    0.47 ASP  19 A
ARG 277 A                                                0.00 XXX   0 X   0.00 XXX   0 X    0.55 ASP 275 A

ARG 289 A  12.78   100 %   -3.29  617   0.00    0    0.60 ASP 264 B   0.00 XXX   0 X    1.01 TYR 262 B
ARG 289 A                                                0.00 XXX   0 X   0.00 XXX   0 X    1.96 ASP 264 B

ARG 314 A  12.66     0 %   -0.33  280   0.00    0    0.00 XXX   0 X   0.00 XXX   0 X    0.02 GLU 192 A
ARG 314 A                                                0.00 XXX   0 X   0.00 XXX   0 X    0.21 ASP 311 A
ARG 314 A                                                0.00 XXX   0 X   0.00 XXX   0 X    0.26 TYR 315 A

ARG 352 A  12.55     0 %   -0.14  225   0.00    0    0.00 XXX   0 X   0.00 XXX   0 X    0.19 ASP 355 A

ARG 370 A  12.49     0 %   -0.08  119   0.00    0    0.00 XXX   0 X   0.00 XXX   0 X    0.03 ASP 368 A
ARG 370 A                                                0.00 XXX   0 X   0.00 XXX   0 X    0.04 GLU 371 A
ARG 370 A                                                0.00 XXX   0 X   0.00 XXX   0 X    0.00 ASP 377 A

ARG 386 A  14.66    65 %   -1.60  463   0.00    0    0.60 GLU 340 A   0.00 XXX   0 X    0.00 GLU 126 A
ARG 386 A                                                1.12 ASP 382 A   0.00 XXX   0 X    1.21 GLU 340 A
ARG 386 A                                                0.00 XXX   0 X   0.00 XXX   0 X    0.83 ASP 382 A

ARG 390 A  12.10     2 %   -0.36  287   0.00    0    0.00 XXX   0 X   0.00 XXX   0 X    0.05 ASP 382 A
ARG 390 A                                                0.00 XXX   0 X   0.00 XXX   0 X    0.07 GLU 387 A
ARG 390 A                                                0.00 XXX   0 X   0.00 XXX   0 X   -0.16 ARG 386 A

ARG 395 A  12.64     0 %   -0.27  208   0.00    0    0.02 GLU 392 A   0.00 XXX   0 X    0.38 GLU 392 A

N+    1 A   7.77     0 %   -0.25  225   0.00    0    0.00 XXX   0 X   0.00 XXX   0 X   -0.09 LYS   2 A
N+    1 A                                                0.00 XXX   0 X   0.00 XXX   0 X    0.03 GLU  26 A
N+    1 A                                                0.00 XXX   0 X   0.00 XXX   0 X    0.08 ASP  95 A

ASP  19 B   2.91    26 %    0.94  355   0.11    0   -0.49 ARG  22 B   0.00 XXX   0 X   -0.40 HIS  23 B
ASP  19 B                                               -0.33 ARG 277 B   0.00 XXX   0 X   -0.32 ARG  22 B
ASP  19 B                                                0.00 XXX   0 X   0.00 XXX   0 X   -0.40 ARG 277 B

ASP  57 B   4.81    31 %    0.92  367   0.02    0    0.00 XXX   0 X   0.00 XXX   0 X   -0.02 LYS  37 B
ASP  57 B                                                0.00 XXX   0 X   0.00 XXX   0 X    0.08 ASP  58 B

ASP  58 B   2.80     0 %    0.41  234   0.00    0   -0.79 SER  61 B  -0.38 SER  61 B   -0.24 LYS  37 B

ASP  87 B   3.54     4 %    0.43  292   0.03    0   -0.71 GLN  83 B   0.00 XXX   0 X   -0.00 ARG 115 B

ASP  93 B   2.15     0 %    0.31  245   0.00    0   -0.80 GLN   3 B   0.00 XXX   0 X   -0.05 ARG  29 B
ASP  93 B                                               -0.75 ARG  52 B   0.00 XXX   0 X   -0.05 LYS 118 B
ASP  93 B                                                0.00 XXX   0 X   0.00 XXX   0 X   -0.32 ARG  52 B
```

175

```
ASP   95 B    3.42    17 %    0.62   330    0.21    0    0.00 XXX    0 X   -0.83 GLN    3 B   -0.23 LYS    2 B
ASP   95 B                                                 0.00 XXX    0 X    0.00 XXX    0 X   -0.01 ARG   29 B
ASP   95 B                                                 0.00 XXX    0 X    0.00 XXX    0 X   -0.13 LYS  118 B
ASP   95 B                                                 0.00 XXX    0 X    0.00 XXX    0 X   -0.09 N+     1 B
ASP   95 B                                                 0.00 XXX    0 X    0.00 XXX    0 X    0.07 ASP   93 B

ASP  137 B    3.22     6 %    0.54   297    0.04    0   -0.81 TYR  170 B    0.00 XXX    0 X   -0.26 ARG  133 B
ASP  137 B                                                 0.00 XXX    0 X    0.00 XXX    0 X   -0.09 LYS  140 B

ASP  150 B    6.49*  100 %    4.88   737    0.24    0   -0.78 LYS  125 B   -0.02 SER  151 B    0.15 GLU  126 B
ASP  150 B                                              -0.52 GLU  152 B    0.00 XXX    0 X   -2.03 LYS  125 B
ASP  150 B                                               0.54 GLU  231 B    0.00 XXX    0 X   -0.60 LYS  228 B
ASP  150 B                                              -1.21 GLU  234 B    0.00 XXX    0 X    2.03 GLU  231 B

ASP  172 B    3.44     0 %    0.41   270    0.00    0    0.00 XXX    0 X   -0.04 GLU  174 B    0.00 XXX    0 X
ASP  172 B                                                 0.00 XXX    0 X   -0.73 GLN  175 B    0.00 XXX    0 X

ASP  197 B    4.13     0 %    0.33   165    0.00    0    0.00 XXX    0 X    0.00 XXX    0 X    0.00 XXX    0 X

ASP  204 B    3.70     0 %    0.21   215    0.00    0    0.00 XXX    0 X    0.00 XXX    0 X   -0.21 ARG  208 B
ASP  204 B                                                 0.00 XXX    0 X    0.00 XXX    0 X   -0.10 ARG  216 B

ASP  221 B    4.35   100 %    3.68   679    0.72    0   -0.84 ASN  339 B    0.00 XXX    0 X   -0.33 LYS  125 B
ASP  221 B                                              -0.85 LYS  217 B    0.00 XXX    0 X    0.21 GLU  340 B
ASP  221 B                                               0.00 XXX    0 X    0.00 XXX    0 X   -2.03 LYS  217 B

ASP  264 B    3.88   100 %    3.64   600    0.75    0   -0.85 SER  266 B   -0.09 ASP  264 B   -1.96 ARG  289 A
ASP  264 B                                              -0.60 ARG  289 A   -0.81 SER  266 B    0.00 XXX    0 X

ASP  275 B    3.98    61 %    1.96   451    0.00    0   -0.58 ARG  277 B   -0.69 ARG  277 B   -0.01 LYS  295 B
ASP  275 B                                                 0.00 XXX    0 X    0.00 XXX    0 X    0.10 ASP   19 B
ASP  275 B                                                 0.00 XXX    0 X    0.00 XXX    0 X   -0.02 HIS   23 B
ASP  275 B                                                 0.00 XXX    0 X    0.00 XXX    0 X   -0.59 ARG  277 B

ASP  298 B    3.26     0 %    0.40   265    0.00    0    0.00 XXX    0 X   -0.58 CYS  300 B   -0.10 ARG  261 A
ASP  298 B                                                 0.00 XXX    0 X    0.00 XXX    0 X   -0.10 LYS  295 B
ASP  298 B                                                 0.00 XXX    0 X    0.00 XXX    0 X   -0.17 LYS  301 B

ASP  311 B    3.62     0 %    0.43   282    0.00    0   -0.07 ARG  314 B   -0.32 ARG  314 B   -0.22 ARG  314 B

ASP  313 B    4.03     0 %    0.15   172    0.00    0    0.00 XXX    0 X   -0.02 ASP  313 B   -0.08 LYS  319 B
ASP  313 B                                                 0.00 XXX    0 X    0.00 XXX    0 X    0.17 ASP  311 B

ASP  355 B    3.97    28 %    1.07   361    0.13    0    0.00 XXX    0 X   -0.83 ARG  352 B   -0.20 ARG  352 B
```

176

```
ASP 368 B    3.92      0 %    0.13  197   0.00     0    0.00 XXX    0 X    0.00 XXX    0 X   -0.01 ARG 370 B

ASP 376 B    4.07      0 %    0.18  196   0.00     0    0.00 XXX    0 X    0.00 XXX    0 X    0.09 ASP 377 B

ASP 377 B    3.27      0 %    0.35  208   0.00     0   -0.16 CYS 374 B   -0.71 GLN 373 B   -0.01 ARG 370 B

ASP 382 B    4.82     67 %    2.06  469   0.40     0   -1.14 ARG 386 B    0.00 XXX    0 X   -0.05 ARG 390 B
ASP 382 B                                                0.00 XXX    0 X    0.00 XXX    0 X    0.61 GLU 340 B
ASP 382 B                                                0.00 XXX    0 X    0.00 XXX    0 X   -0.86 ARG 386 B

GLU  26 B    4.35      0 %    0.18  188   0.00     0    0.00 XXX    0 X    0.00 XXX    0 X   -0.03 N+    1 B
GLU  26 B                                                0.00 XXX    0 X    0.00 XXX    0 X   -0.31 HIS  27 B

GLU  44 B    4.46      0 %    0.34  278   0.00     0    0.00 XXX    0 X    0.00 XXX    0 X   -0.09 ARG  22 B
GLU  44 B                                                0.00 XXX    0 X    0.00 XXX    0 X   -0.40 ARG  41 B
GLU  44 B                                                0.00 XXX    0 X    0.00 XXX    0 X    0.11 GLU  48 B

GLU  48 B    3.46      9 %    0.73  306   0.01     0   -1.37 ARG  22 B    0.00 XXX    0 X   -0.03 ARG  41 B
GLU  48 B                                                0.00 XXX    0 X    0.00 XXX    0 X    0.06 ASP  19 B
GLU  48 B                                                0.00 XXX    0 X    0.00 XXX    0 X   -0.43 ARG  22 B

GLU  59 B    4.66      0 %    0.10  155   0.00     0    0.00 XXX    0 X    0.00 XXX    0 X    0.05 ASP  58 B

GLU  77 B    4.72      0 %    0.26  265   0.00     0    0.00 XXX    0 X    0.00 XXX    0 X   -0.04 LYS  66 B

GLU  92 B    4.52      0 %    0.08  165   0.00     0    0.00 XXX    0 X    0.00 XXX    0 X   -0.08 LYS 118 B
GLU  92 B                                                0.00 XXX    0 X    0.00 XXX    0 X    0.02 ASP  93 B

GLU 126 B    5.71    100 %    3.86  643   0.15     0   -0.78 LYS 217 B   -0.35 GLY 103 B   -0.03 ARG 386 B
GLU 126 B                                                0.00 XXX    0 X   -0.65 GLU 126 B   -0.61 LYS 125 B
GLU 126 B                                                0.00 XXX    0 X    0.00 XXX    0 X   -2.03 LYS 217 B
GLU 126 B                                                0.00 XXX    0 X    0.00 XXX    0 X    0.72 ASP 221 B
GLU 126 B                                                0.00 XXX    0 X    0.00 XXX    0 X    0.93 GLU 340 B

GLU 152 B   10.62*   100 %    4.88  707   0.80     0   -0.74 LYS 228 B   -0.74 GLU 152 B   -0.99 LYS 125 B
GLU 152 B                                                0.52 ASP 150 B    0.00 XXX    0 X   -2.03 LYS 228 B
GLU 152 B                                                0.87 GLU 231 B    0.00 XXX    0 X    1.52 ASP 150 B
GLU 152 B                                                0.00 XXX    0 X    0.00 XXX    0 X    2.03 GLU 231 B

GLU 174 B    4.79      0 %    0.16  251   0.00     0    0.00 XXX    0 X    0.00 XXX    0 X    0.12 ASP 172 B

GLU 192 B    4.32      0 %    0.19  245   0.00     0    0.00 XXX    0 X    0.00 XXX    0 X   -0.34 ARG 191 B
GLU 192 B                                                0.00 XXX    0 X    0.00 XXX    0 X   -0.03 ARG 314 B

GLU 231 B    1.51*   100 %    5.13  724   0.66     0   -0.47 ASN 227 B    0.00 XXX    0 X   -0.03 HIS 209 B
GLU 231 B                                               -0.85 LYS 125 B    0.00 XXX    0 X   -2.03 LYS 125 B
```

177

```
GLU 231 B                                    -0.54 ASP 150 B     0.00 XXX    0 X   -1.09 HIS 153 B
GLU 231 B                                    -0.87 GLU 152 B     0.00 XXX    0 X   -2.03 LYS 228 B
GLU 231 B                                    -0.85 LYS 228 B     0.00 XXX    0 X    0.00 XXX    0 X

GLU 234 B  12.78* 100 %   4.37  712   0.66  0   1.21 ASP 150 B     0.00 XXX    0 X   -1.14 LYS 125 B
GLU 234 B                                     0.00 XXX    0 X     0.00 XXX    0 X   -0.06 LYS 228 B
GLU 234 B                                     0.00 XXX    0 X     0.00 XXX    0 X    0.18 GLU 126 B
GLU 234 B                                     0.00 XXX    0 X     0.00 XXX    0 X    0.34 GLU 152 B
GLU 234 B                                     0.00 XXX    0 X     0.00 XXX    0 X    1.97 ASP 150 B
GLU 234 B                                     0.00 XXX    0 X     0.00 XXX    0 X    0.78 GLU 231 B

GLU 247 B   3.30   58 %   1.63  445   0.04  0  -0.80 SER 180 B     0.00 XXX    0 X   -0.80 ARG 261 B
GLU 247 B                                    -1.27 ARG 261 B     0.00 XXX    0 X    0.00 XXX    0 X

GLU 273 B   4.91   28 %   0.85  361   0.10  0  -0.29 LYS 295 B     0.00 XXX    0 X   -0.03 ARG 261 A
GLU 273 B                                     0.00 XXX    0 X     0.00 XXX    0 X    0.09 ASP 275 B
GLU 273 B                                     0.00 XXX    0 X     0.00 XXX    0 X    0.05 ASP 298 B
GLU 273 B                                     0.00 XXX    0 X     0.00 XXX    0 X   -0.36 LYS 295 B

GLU 323 B   4.58    0 %   0.34  282   0.00  0   0.00 XXX    0 X     0.00 XXX    0 X   -0.27 LYS 319 B

GLU 326 B   4.71    0 %   0.16  216   0.00  0   0.00 XXX    0 X     0.00 XXX    0 X   -0.05 ARG 236 B
GLU 326 B                                     0.00 XXX    0 X     0.00 XXX    0 X    0.10 GLU 323 B

GLU 340 B   3.93  100 %   2.89  577   0.65  0  -0.48 ASN 336 B    -0.75 ALA 104 B   -1.03 LYS 217 B
GLU 340 B                                    -0.52 ARG 386 B     0.00 XXX    0 X   -1.33 ARG 386 B

GLU 365 B   4.68    0 %   0.26  223   0.00  0   0.00 XXX    0 X     0.00 XXX    0 X   -0.10 LYS 366 B
GLU 365 B                                     0.00 XXX    0 X     0.00 XXX    0 X    0.02 GLU 392 B

GLU 371 B   3.85   18 %   0.81  333   0.08  0   0.00 XXX    0 X    -0.68 GLN 329 B   -0.14 ARG 370 B
GLU 371 B                                     0.00 XXX    0 X    -0.71 ALA 330 B    0.00 XXX    0 X

GLU 387 B   3.95    0 %   0.22  204   0.00  0  -0.20 ARG 390 B     0.00 XXX    0 X   -0.26 ARG 390 B
GLU 387 B                                     0.00 XXX    0 X     0.00 XXX    0 X   -0.31 LYS 391 B

GLU 392 B   2.80    1 %   0.53  283   0.00  0  -0.85 SER 362 B     0.00 XXX    0 X   -0.00 LYS 391 B
GLU 392 B                                    -0.69 LYS 366 B     0.00 XXX    0 X   -0.31 ARG 395 B
GLU 392 B                                     0.00 XXX    0 X     0.00 XXX    0 X   -0.38 LYS 366 B

C-  398 B   3.30    0 %   0.10  110   0.00  0   0.00 XXX    0 X     0.00 XXX    0 X    0.00 XXX    0 X

HIS  23 B   6.72   10 %  -0.55  309   0.00  0   0.00 XXX    0 X     0.65 HIS  23 B   -0.08 ARG  22 B
HIS  23 B                                     0.00 XXX    0 X     0.00 XXX    0 X   -0.22 ARG 277 B
HIS  23 B                                     0.00 XXX    0 X     0.00 XXX    0 X    0.40 ASP  19 B
HIS  23 B                                     0.00 XXX    0 X     0.00 XXX    0 X    0.02 ASP 275 B
```

178

```
HIS  27 B    6.17    10 %   -0.53   308    0.00     0     0.00 XXX    0 X     0.00 XXX    0 X    -0.00 LYS   2 B
HIS  27 B                                                  0.00 XXX    0 X     0.00 XXX    0 X    -0.10 N+    1 B
HIS  27 B                                                  0.00 XXX    0 X     0.00 XXX    0 X     0.31 GLU  26 B

HIS 153 B    2.68   100 %   -3.92   718    0.00     0     0.00 XXX    0 X     0.00 XXX    0 X    -0.66 LYS 125 B
HIS 153 B                                                  0.00 XXX    0 X     0.00 XXX    0 X    -0.34 LYS 228 B
HIS 153 B                                                  0.00 XXX    0 X     0.00 XXX    0 X     1.09 GLU 231 B

HIS 166 B    5.91    17 %   -0.59   330    0.00     0     0.00 XXX    0 X     0.00 XXX    0 X     0.00 XXX   0 X

HIS 209 B    3.25    85 %   -3.02   519    0.00     0     0.00 XXX    0 X     0.00 XXX    0 X    -0.26 LYS 228 B
HIS 209 B                                                  0.00 XXX    0 X     0.00 XXX    0 X     0.03 GLU 231 B

HIS 251 B    3.40   100 %   -3.33   564    0.00     0     0.02 GLN 253 B    0.48 HIS 251 B    -0.00 ARG 191 B
HIS 251 B                                                  0.00 XXX    0 X     0.00 XXX    0 X    -0.26 HIS 257 B

HIS 257 B    3.97   100 %   -3.14   666    0.00     0     0.28 GLN 270 B    0.23 HIS 257 B    -0.01 LYS 228 B
HIS 257 B                                                  0.00 XXX    0 X     0.12 HIS 257 B     0.00 XXX   0 X

HIS 282 B    2.10   100 %   -3.09   640    0.00     0     0.00 XXX    0 X     0.07 HIS 282 B    -1.39 ARG 289 B

CYS  15 B    9.90    59 %    1.76   447    0.00     0    -0.35 ARG  41 B   -0.04 CYS  15 B    -0.09 ARG  22 B
CYS  15 B                                                  0.00 XXX    0 X     0.00 XXX    0 X     0.01 ASP  19 B
CYS  15 B                                                  0.00 XXX    0 X     0.00 XXX    0 X     0.15 GLU  44 B
CYS  15 B                                                  0.00 XXX    0 X     0.00 XXX    0 X    -0.53 ARG  41 B

CYS  46 B   11.93    77 %    2.93   498    0.00     0     0.00 XXX    0 X     0.00 XXX    0 X     0.00 XXX   0 X

CYS  86 B   10.64    44 %    1.63   404    0.00     0     0.00 XXX    0 X     0.00 XXX    0 X    -0.14 ARG 115 B
CYS  86 B                                                  0.00 XXX    0 X     0.00 XXX    0 X     0.01 ASP  57 B
CYS  86 B                                                  0.00 XXX    0 X     0.00 XXX    0 X     0.14 ASP  87 B

CYS 131 B   11.77    81 %    2.76   509    0.00     0     0.00 XXX    0 X    -0.01 PHE 135 B     0.01 CYS 374 B

CYS 207 B   10.83    54 %    2.01   432    0.00     0     0.00 XXX    0 X     0.00 XXX    0 X    -0.03 ARG 208 B
CYS 207 B                                                  0.00 XXX    0 X     0.00 XXX    0 X    -0.27 ARG 216 B
CYS 207 B                                                  0.00 XXX    0 X     0.00 XXX    0 X     0.12 ASP 204 B

CYS 300 B    9.97    15 %    0.73   324    0.00     0     0.00 XXX    0 X     0.00 XXX    0 X    -0.02 ARG 261 A
CYS 300 B                                                  0.00 XXX    0 X     0.00 XXX    0 X    -0.02 LYS 301 B
CYS 300 B                                                  0.00 XXX    0 X     0.00 XXX    0 X     0.04 GLU 247 A
CYS 300 B                                                  0.00 XXX    0 X     0.00 XXX    0 X     0.25 ASP 298 B

CYS 317 B   12.04   100 %    3.22   592    0.00     0    -0.34 THR 224 B   -0.02 CYS 317 B     0.39 ASP 221 B
CYS 317 B                                                 -0.22 ASN 360 B    0.00 XXX    0 X     0.00 XXX   0 X
```

179

```
CYS 374 B    9.14    0 %    0.47   227   0.00    0    0.16 ASP 377 B   -0.43 ASP 376 B    0.23 ASP 376 B
CYS 374 B                                                0.00 XXX   0 X   -0.68 ASP 377 B    0.38 ASP 377 B

TYR  53 B   10.37    0 %    0.28   275   0.00    0    0.00 XXX   0 X    0.00 XXX   0 X   -0.18 ARG  52 B
TYR  53 B                                                0.00 XXX   0 X    0.00 XXX   0 X    0.24 GLU  77 B
TYR  53 B                                                0.00 XXX   0 X    0.00 XXX   0 X    0.02 ASP  87 B
TYR  53 B                                                0.00 XXX   0 X    0.00 XXX   0 X    0.02 ASP  93 B

TYR 170 B   11.85   25 %    1.32   352   0.00    0   -0.37 ARG 133 B    0.00 XXX   0 X   -0.34 ARG 133 B
TYR 170 B                                                0.81 ASP 137 B    0.00 XXX   0 X    0.44 ASP 137 B

TYR 232 B   10.47   51 %    1.60   423   0.00    0   -0.62 ARG 236 B    0.00 XXX   0 X   -0.59 ARG 236 B
TYR 232 B                                                0.00 XXX   0 X    0.00 XXX   0 X    0.08 GLU 326 B

TYR 262 B   12.79  100 %    3.66   629   0.00    0    0.00 XXX   0 X    0.00 XXX   0 X   -1.01 ARG 289 A
TYR 262 B                                                0.00 XXX   0 X    0.00 XXX   0 X    0.14 ASP 264 B

TYR 312 B   11.10   43 %    1.02   402   0.00    0    0.00 XXX   0 X    0.00 XXX   0 X    0.08 TYR 232 B

TYR 315 B   12.23   72 %    2.54   483   0.00    0    0.00 XXX   0 X    0.00 XXX   0 X   -0.15 ARG 191 B
TYR 315 B                                                0.00 XXX   0 X    0.00 XXX   0 X    0.02 GLU 192 B
TYR 315 B                                                0.00 XXX   0 X    0.00 XXX   0 X    0.14 ASP 311 B
TYR 315 B                                                0.00 XXX   0 X    0.00 XXX   0 X   -0.32 ARG 314 B

LYS   2 B   10.38   16 %   -0.34   325   0.00    0    0.00 XXX   0 X    0.00 XXX   0 X    0.23 ASP  95 B

LYS  37 B   10.63    0 %   -0.14   243   0.00    0    0.00 XXX   0 X    0.00 XXX   0 X    0.02 ASP  57 B
LYS  37 B                                                0.00 XXX   0 X    0.00 XXX   0 X    0.24 ASP  58 B

LYS  63 B   10.43    0 %   -0.07   124   0.00    0    0.00 XXX   0 X    0.00 XXX   0 X    0.00 XXX   0 X

LYS  66 B   10.29    0 %   -0.25   244   0.00    0    0.00 XXX   0 X    0.00 XXX   0 X    0.04 GLU  77 B

LYS 118 B    9.57   27 %   -1.19   358   0.00    0    0.00 XXX   0 X    0.00 XXX   0 X    0.08 GLU  92 B
LYS 118 B                                                0.00 XXX   0 X    0.00 XXX   0 X    0.05 ASP  93 B
LYS 118 B                                                0.00 XXX   0 X    0.00 XXX   0 X    0.13 ASP  95 B

LYS 125 B   13.75  100 %   -5.52   703   0.00    0    0.78 ASP 150 B    0.00 XXX   0 X    0.99 GLU 152 B
LYS 125 B                                                0.85 GLU 231 B    0.00 XXX   0 X    0.33 ASP 221 B
LYS 125 B                                                0.00 XXX   0 X    0.00 XXX   0 X    1.14 GLU 234 B
LYS 125 B                                                0.00 XXX   0 X    0.00 XXX   0 X    0.61 GLU 126 B
LYS 125 B                                                0.00 XXX   0 X    0.00 XXX   0 X    2.03 ASP 150 B
LYS 125 B                                                0.00 XXX   0 X    0.00 XXX   0 X    2.03 GLU 231 B

LYS 140 B   10.38    0 %   -0.19   208   0.00    0    0.00 XXX   0 X    0.00 XXX   0 X   -0.02 ARG 133 B
```

180

```
    LYS 140 B                                        0.00 XXX    0 X    0.00 XXX    0 X    0.09 ASP 137 B

    LYS 143 B  10.43    0 %  -0.07  121   0.00    0   0.00 XXX    0 X    0.00 XXX    0 X    0.00 XXX    0 X

    LYS 217 B  13.29  100 %  -3.53  643   0.00    0   0.78 GLU 126 B    0.00 XXX    0 X   -0.12 ARG 386 B
    LYS 217 B                                        0.85 ASP 221 B    0.00 XXX    0 X   -0.30 LYS 125 B
    LYS 217 B                                        0.00 XXX    0 X    0.00 XXX    0 X    2.03 GLU 126 B
    LYS 217 B                                        0.00 XXX    0 X    0.00 XXX    0 X    2.03 ASP 221 B
    LYS 217 B                                        0.00 XXX    0 X    0.00 XXX    0 X    1.03 GLU 340 B

    LYS 228 B  10.85  100 %  -5.30  707   0.00    0   0.74 GLU 152 B    0.00 XXX    0 X    2.03 GLU 152 B
    LYS 228 B                                        0.85 GLU 231 B    0.00 XXX    0 X    0.06 GLU 234 B
    LYS 228 B                                        0.00 XXX    0 X    0.00 XXX    0 X   -0.66 LYS 125 B
    LYS 228 B                                        0.00 XXX    0 X    0.00 XXX    0 X    0.60 ASP 150 B
    LYS 228 B                                        0.00 XXX    0 X    0.00 XXX    0 X    2.03 GLU 231 B

    LYS 295 B  11.01    0 %  -0.24  252   0.00    0   0.29 GLU 273 B    0.00 XXX    0 X   -0.01 ARG 261 A
    LYS 295 B                                        0.00 XXX    0 X    0.00 XXX    0 X    0.01 ASP 275 B
    LYS 295 B                                        0.00 XXX    0 X    0.00 XXX    0 X    0.10 ASP 298 B
    LYS 295 B                                        0.00 XXX    0 X    0.00 XXX    0 X    0.36 GLU 273 B

    LYS 301 B  10.58    0 %  -0.07  103   0.00    0   0.00 XXX    0 X    0.00 XXX    0 X    0.17 ASP 298 B
    LYS 301 B                                        0.00 XXX    0 X    0.00 XXX    0 X    0.02 CYS 300 B
    LYS 301 B                                        0.00 XXX    0 X    0.00 XXX    0 X   -0.04 LYS 295 B

    LYS 319 B  10.69    0 %  -0.15  203   0.00    0   0.00 XXX    0 X    0.00 XXX    0 X    0.08 ASP 313 B
    LYS 319 B                                        0.00 XXX    0 X    0.00 XXX    0 X    0.27 GLU 323 B

    LYS 366 B  11.02    0 %  -0.48  254   0.00    0   0.69 GLU 392 B    0.00 XXX    0 X    0.10 GLU 365 B
    LYS 366 B                                        0.00 XXX    0 X    0.00 XXX    0 X   -0.17 ARG 395 B
    LYS 366 B                                        0.00 XXX    0 X    0.00 XXX    0 X    0.38 GLU 392 B

    LYS 391 B  10.43    0 %  -0.28  193   0.00    0   0.00 XXX    0 X    0.00 XXX    0 X    0.31 GLU 387 B
    LYS 391 B                                        0.00 XXX    0 X    0.00 XXX    0 X   -0.04 ARG 390 B
    LYS 391 B                                        0.00 XXX    0 X    0.00 XXX    0 X    0.00 GLU 392 B
    LYS 391 B                                        0.00 XXX    0 X    0.00 XXX    0 X   -0.06 ARG 395 B

    ARG  22 B  14.53   20 %  -0.78  337   0.00    0   0.49 ASP  19 B    0.00 XXX    0 X    0.09 CYS  15 B
    ARG  22 B                                        1.37 GLU  48 B    0.00 XXX    0 X    0.09 GLU  44 B
    ARG  22 B                                        0.00 XXX    0 X    0.00 XXX    0 X    0.32 ASP  19 B
    ARG  22 B                                        0.00 XXX    0 X    0.00 XXX    0 X    0.43 GLU  48 B

    ARG  29 B  12.28    0 %  -0.26  262   0.00    0   0.00 XXX    0 X    0.00 XXX    0 X    0.05 ASP  93 B
    ARG  29 B                                        0.00 XXX    0 X    0.00 XXX    0 X    0.01 ASP  95 B
    ARG  29 B                                        0.00 XXX    0 X    0.00 XXX    0 X   -0.03 ARG  52 B
```

```
ARG   41 B  13.35    11 %   -0.40  313   0.00    0    0.35 CYS  15 B   0.00 XXX   0 X    0.40 GLU  44 B
ARG   41 B                                            0.00 XXX   0 X   0.00 XXX   0 X    0.03 GLU  48 B
ARG   41 B                                            0.00 XXX   0 X   0.00 XXX   0 X   -0.08 ARG  22 B
ARG   41 B                                            0.00 XXX   0 X   0.00 XXX   0 X    0.53 CYS  15 B

ARG   52 B  13.25     5 %   -0.50  296   0.00    0    0.75 ASP  93 B   0.00 XXX   0 X    0.18 TYR  53 B
ARG   52 B                                            0.00 XXX   0 X   0.00 XXX   0 X    0.32 ASP  93 B

ARG   75 B  12.44     0 %   -0.06  139   0.00    0    0.00 XXX   0 X   0.00 XXX   0 X    0.00 XXX   0 X

ARG  115 B  12.23    15 %   -0.41  323   0.00    0    0.00 XXX   0 X   0.00 XXX   0 X    0.14 CYS  86 B
ARG  115 B                                            0.00 XXX   0 X   0.00 XXX   0 X    0.00 ASP  87 B

ARG  133 B  12.74    26 %   -0.73  354   0.00    0    0.37 TYR 170 B   0.00 XXX   0 X    0.26 ASP 137 B
ARG  133 B                                            0.00 XXX   0 X   0.00 XXX   0 X    0.34 TYR 170 B

ARG  191 B  12.22    24 %   -0.74  349   0.00    0    0.00 XXX   0 X   0.00 XXX   0 X    0.34 GLU 192 B
ARG  191 B                                            0.00 XXX   0 X   0.00 XXX   0 X    0.15 TYR 315 B
ARG  191 B                                            0.00 XXX   0 X   0.00 XXX   0 X   -0.03 ARG 314 B

ARG  196 B  12.44     0 %   -0.06   74   0.00    0    0.00 XXX   0 X   0.00 XXX   0 X    0.00 XXX   0 X

ARG  208 B  12.51     0 %   -0.23  199   0.00    0    0.00 XXX   0 X   0.00 XXX   0 X    0.21 ASP 204 B
ARG  208 B                                            0.00 XXX   0 X   0.00 XXX   0 X    0.03 CYS 207 B

ARG  216 B  12.47    18 %   -0.40  331   0.00    0    0.00 XXX   0 X   0.00 XXX   0 X    0.10 ASP 204 B
ARG  216 B                                            0.00 XXX   0 X   0.00 XXX   0 X    0.27 CYS 207 B

ARG  236 B  12.36    43 %   -1.40  403   0.00    0    0.62 TYR 232 B   0.00 XXX   0 X    0.59 TYR 232 B
ARG  236 B                                            0.00 XXX   0 X   0.00 XXX   0 X    0.05 GLU 326 B

ARG  261 B  12.70    70 %   -1.99  476   0.00    0    1.27 GLU 247 B   0.00 XXX   0 X    0.00 GLU 273 A
ARG  261 B                                            0.00 XXX   0 X   0.00 XXX   0 X    0.09 ASP 298 A
ARG  261 B                                            0.00 XXX   0 X   0.00 XXX   0 X    0.03 CYS 300 A
ARG  261 B                                            0.00 XXX   0 X   0.00 XXX   0 X    0.80 GLU 247 B

ARG  277 B  13.49    28 %   -0.89  359   0.00    0    0.33 ASP  19 B   0.00 XXX   0 X   -0.02 ARG  22 B
ARG  277 B                                            0.58 ASP 275 B   0.00 XXX   0 X    0.40 ASP  19 B
ARG  277 B                                            0.00 XXX   0 X   0.00 XXX   0 X    0.59 ASP 275 B

ARG  289 B  12.63   100 %   -3.24  618   0.00    0    0.50 ASP 264 A   0.00 XXX   0 X    1.02 TYR 262 A
ARG  289 B                                            0.00 XXX   0 X   0.00 XXX   0 X    1.86 ASP 264 A

ARG  314 B  12.52     8 %   -0.62  303   0.00    0    0.07 ASP 311 B   0.00 XXX   0 X    0.03 GLU 192 B
ARG  314 B                                            0.00 XXX   0 X   0.00 XXX   0 X    0.22 ASP 311 B
ARG  314 B                                            0.00 XXX   0 X   0.00 XXX   0 X    0.32 TYR 315 B
```

182

```
ARG 352 B   12.56     0 %    -0.14  212    0.00     0    0.00 XXX   0 X    0.00 XXX   0 X    0.20 ASP 355 B

ARG 370 B   12.47     0 %    -0.19  169    0.00     0    0.00 XXX   0 X    0.00 XXX   0 X    0.01 ASP 368 B
ARG 370 B                                                  0.00 XXX   0 X    0.00 XXX   0 X    0.14 GLU 371 B
ARG 370 B                                                  0.00 XXX   0 X    0.00 XXX   0 X    0.01 ASP 377 B

ARG 386 B   14.67    69 %    -1.71  475    0.00     0    0.52 GLU 340 B   0.00 XXX   0 X    0.03 GLU 126 B
ARG 386 B                                                  1.14 ASP 382 B   0.00 XXX   0 X    1.33 GLU 340 B
ARG 386 B                                                  0.00 XXX   0 X    0.00 XXX   0 X    0.86 ASP 382 B

ARG 390 B   12.46     3 %    -0.39  290    0.00     0    0.20 GLU 387 B   0.00 XXX   0 X    0.05 ASP 382 B
ARG 390 B                                                  0.00 XXX   0 X    0.00 XXX   0 X    0.26 GLU 387 B
ARG 390 B                                                  0.00 XXX   0 X    0.00 XXX   0 X   -0.16 ARG 386 B

ARG 395 B   12.58     0 %    -0.23  197    0.00     0    0.00 XXX   0 X    0.00 XXX   0 X    0.31 GLU 392 B

N+    1 B    7.66     0 %    -0.25  221    0.00     0    0.00 XXX   0 X    0.00 XXX   0 X   -0.14 LYS   2 B
N+    1 B                                                  0.00 XXX   0 X    0.00 XXX   0 X    0.03 GLU  26 B
N+    1 B                                                  0.00 XXX   0 X    0.00 XXX   0 X   -0.06 ARG  29 B
N+    1 B                                                  0.00 XXX   0 X    0.00 XXX   0 X    0.09 ASP  95 B
```

```
--------------------------------------------------------------------------------------------
  Residues that are found to be 'coupled', i.e. titrates together, has been marked by '*' in the above
  section. Please rerun PropKa with the --display-coupled-residues option for detailed information.
--------------------------------------------------------------------------------------------
SUMMARY OF THIS PREDICTION
     RESIDUE    pKa   pKmodel   ligand atom-type
  ASP  19 A    3.18     3.80
  ASP  57 A    4.68     3.80
  ASP  58 A    2.97     3.80
  ASP  87 A    3.70     3.80
  ASP  93 A    3.29     3.80
  ASP  95 A    3.65     3.80
  ASP 137 A    3.26     3.80
  ASP 150 A    1.23     3.80
  ASP 172 A    3.56     3.80
  ASP 197 A    4.06     3.80
  ASP 204 A    3.06     3.80
  ASP 221 A    4.58     3.80
  ASP 264 A    4.07     3.80
  ASP 275 A    4.13     3.80
  ASP 298 A    3.40     3.80
  ASP 311 A    3.36     3.80
  ASP 313 A    3.96     3.80
  ASP 355 A    3.98     3.80
```

```
ASP 368 A      3.91       3.80
ASP 376 A      4.05       3.80
ASP 377 A      3.27       3.80
ASP 382 A      4.83       3.80
GLU  26 A      4.44       4.50
GLU  44 A      4.37       4.50
GLU  48 A      3.41       4.50
GLU  59 A      4.50       4.50
GLU  77 A      4.52       4.50
GLU  92 A      4.61       4.50
GLU 126 A      6.06       4.50
GLU 152 A     10.96       4.50
GLU 174 A      4.82       4.50
GLU 192 A      4.49       4.50
GLU 231 A      4.51       4.50
GLU 234 A     12.81       4.50
GLU 247 A      3.44       4.50
GLU 273 A      5.10       4.50
GLU 323 A      4.69       4.50
GLU 326 A      3.96       4.50
GLU 340 A      3.88       4.50
GLU 365 A      4.74       4.50
GLU 371 A      3.97       4.50
GLU 387 A      4.63       4.50
GLU 392 A      4.20       4.50
C-  398 A      3.31       3.20
HIS  23 A      6.08       6.50
HIS  27 A      6.02       6.50
HIS 153 A      5.02       6.50
HIS 166 A      5.89       6.50
HIS 209 A      2.93       6.50
HIS 251 A      2.88       6.50
HIS 257 A      3.99       6.50
HIS 282 A      2.01       6.50
CYS  15 A     10.34       9.00
CYS  46 A     11.92       9.00
CYS  86 A     10.56       9.00
CYS 131 A     11.75       9.00
CYS 207 A     10.78       9.00
CYS 300 A     10.02       9.00
CYS 317 A     11.85       9.00
CYS 374 A      9.07       9.00
TYR  53 A     10.17      10.00
TYR 170 A     11.71      10.00
TYR 232 A     10.66      10.00
TYR 262 A     12.82      10.00
```

184

```
TYR 312 A    11.37      10.00
TYR 315 A    11.98      10.00
LYS   2 A     9.13      10.50
LYS  37 A    10.33      10.50
LYS  63 A    10.62      10.50
LYS  66 A    10.36      10.50
LYS 118 A     9.81      10.50
LYS 125 A    13.89      10.50
LYS 140 A    10.29      10.50
LYS 143 A    10.10      10.50
LYS 217 A    13.30      10.50
LYS 228 A    10.72      10.50
LYS 295 A    10.71      10.50
LYS 301 A    10.43      10.50
LYS 319 A    10.72      10.50
LYS 366 A    10.40      10.50
LYS 391 A    10.17      10.50
ARG  22 A    14.57      12.50
ARG  29 A    11.94      12.50
ARG  41 A    13.08      12.50
ARG  52 A    12.96      12.50
ARG  75 A    12.46      12.50
ARG 115 A    12.16      12.50
ARG 133 A    12.71      12.50
ARG 191 A    11.47      12.50
ARG 196 A    12.41      12.50
ARG 208 A    12.42      12.50
ARG 216 A    13.46      12.50
ARG 236 A    13.16      12.50
ARG 261 A    12.56      12.50
ARG 277 A    13.08      12.50
ARG 289 A    12.78      12.50
ARG 314 A    12.66      12.50
ARG 352 A    12.55      12.50
ARG 370 A    12.49      12.50
ARG 386 A    14.66      12.50
ARG 390 A    12.10      12.50
ARG 395 A    12.64      12.50
N+    1 A     7.77       8.00
ASP  19 B     2.91       3.80
ASP  57 B     4.81       3.80
ASP  58 B     2.80       3.80
ASP  87 B     3.54       3.80
ASP  93 B     2.15       3.80
ASP  95 B     3.42       3.80
ASP 137 B     3.22       3.80
```

185

```
ASP 150 B      6.49         3.80
ASP 172 B      3.44         3.80
ASP 197 B      4.13         3.80
ASP 204 B      3.70         3.80
ASP 221 B      4.35         3.80
ASP 264 B      3.88         3.80
ASP 275 B      3.98         3.80
ASP 298 B      3.26         3.80
ASP 311 B      3.62         3.80
ASP 313 B      4.03         3.80
ASP 355 B      3.97         3.80
ASP 368 B      3.92         3.80
ASP 376 B      4.07         3.80
ASP 377 B      3.27         3.80
ASP 382 B      4.82         3.80
GLU  26 B      4.35         4.50
GLU  44 B      4.46         4.50
GLU  48 B      3.46         4.50
GLU  59 B      4.66         4.50
GLU  77 B      4.72         4.50
GLU  92 B      4.52         4.50
GLU 126 B      5.71         4.50
GLU 152 B     10.62         4.50
GLU 174 B      4.79         4.50
GLU 192 B      4.32         4.50
GLU 231 B      1.51         4.50
GLU 234 B     12.78         4.50
GLU 247 B      3.30         4.50
GLU 273 B      4.91         4.50
GLU 323 B      4.58         4.50
GLU 326 B      4.71         4.50
GLU 340 B      3.93         4.50
GLU 365 B      4.68         4.50
GLU 371 B      3.85         4.50
GLU 387 B      3.95         4.50
GLU 392 B      2.80         4.50
C-  398 B      3.30         3.20
HIS  23 B      6.72         6.50
HIS  27 B      6.17         6.50
HIS 153 B      2.68         6.50
HIS 166 B      5.91         6.50
HIS 209 B      3.25         6.50
HIS 251 B      3.40         6.50
HIS 257 B      3.97         6.50
HIS 282 B      2.10         6.50
CYS  15 B      9.90         9.00
```

186

```
CYS  46 B    11.93      9.00
CYS  86 B    10.64      9.00
CYS 131 B    11.77      9.00
CYS 207 B    10.83      9.00
CYS 300 B     9.97      9.00
CYS 317 B    12.04      9.00
CYS 374 B     9.14      9.00
TYR  53 B    10.37     10.00
TYR 170 B    11.85     10.00
TYR 232 B    10.47     10.00
TYR 262 B    12.79     10.00
TYR 312 B    11.10     10.00
TYR 315 B    12.23     10.00
LYS   2 B    10.38     10.50
LYS  37 B    10.63     10.50
LYS  63 B    10.43     10.50
LYS  66 B    10.29     10.50
LYS 118 B     9.57     10.50
LYS 125 B    13.75     10.50
LYS 140 B    10.38     10.50
LYS 143 B    10.43     10.50
LYS 217 B    13.29     10.50
LYS 228 B    10.85     10.50
LYS 295 B    11.01     10.50
LYS 301 B    10.58     10.50
LYS 319 B    10.69     10.50
LYS 366 B    11.02     10.50
LYS 391 B    10.43     10.50
ARG  22 B    14.53     12.50
ARG  29 B    12.28     12.50
ARG  41 B    13.35     12.50
ARG  52 B    13.25     12.50
ARG  75 B    12.44     12.50
ARG 115 B    12.23     12.50
ARG 133 B    12.74     12.50
ARG 191 B    12.22     12.50
ARG 196 B    12.44     12.50
ARG 208 B    12.51     12.50
ARG 216 B    12.47     12.50
ARG 236 B    12.36     12.50
ARG 261 B    12.70     12.50
ARG 277 B    13.49     12.50
ARG 289 B    12.63     12.50
ARG 314 B    12.52     12.50
ARG 352 B    12.56     12.50
ARG 370 B    12.47     12.50
```

187

```
   ARG 386 B    14.67      12.50
   ARG 390 B    12.46      12.50
   ARG 395 B    12.58      12.50
   N+    1 B     7.66       8.00
--------------------------------------------------------------------------------------------
--------------------------------------------------------------------------------------------
Free energy of   folding (kcal/mol) as a function of pH (using neutral reference)
   0.00    115.93
   1.00    115.34
   2.00    112.14
   3.00    101.87
   4.00     84.92
   5.00     75.56
   6.00     70.96
   7.00     69.49
   8.00     73.69
   9.00     83.58
  10.00    103.51
  11.00    129.66
  12.00    151.92
  13.00    168.65
  14.00    180.06


The pH of optimum stability is  6.7 for which the free energy is  69.2 kcal/mol at 298K
Could not determine pH values where the free energy is within 80 % of maximum
Could not determine where the free energy is positive


Protein charge of folded and unfolded state as a function of pH
   pH  unfolded  folded
  0.00     89.99   89.87
  1.00     89.90   88.92
  2.00     89.06   84.77
  3.00     81.91   70.68
  4.00     51.15   40.23
  5.00     14.24   10.39
  6.00     -0.32   -3.36
  7.00    -10.36   -8.99
  8.00    -16.16  -11.56
  9.00    -25.79  -14.95
 10.00    -43.86  -26.00
 11.00    -66.83  -47.76
 12.00    -83.04  -69.48
 13.00   -105.80  -94.70
 14.00   -114.70 -109.22
The pI is  5.65 (folded) and  5.97 (unfolded)
```

188

Appendix C: Copyright Permissions

Figure **C.1**: Reprinted with permission from White, J.K.; Handa, S; Vankayala, S.L.; Woodcock, H.L., Thiamin Diphosphate Activation in 1-Deoxy-d-xylulose 5-Phosphate Synthase: Insights into the Mechanism and Underlying Intermolecular Interactions, *J. Phys. Chem. B*, **2016**, *120* (37), pp 99229934, **DOI:** 10.1021/acs.jpcb.6b07248. Copyright 2016 American Chemical Society.