8-2015

Conformational Dynamics of Cytochrome P450cam Upon Ligand Binding

Ana Virginia Bernal Gomez
University of Tennessee - Knoxville, abernalg@vols.utk.edu

To the Graduate Council:

I am submitting herewith a thesis written by Ana Virginia Bernal Gomez entitled "Conformational Dynamics of Cytochrome P450cam Upon Ligand Binding." I have examined the final electronic copy of this thesis for form and content and recommend that it be accepted in partial fulfillment of the requirements for the degree of Master of Science, with a major in Biochemistry and Cellular and Molecular Biology.

Nitin Jain, Major Professor

We have read this thesis and recommend its acceptance:

Engin Serpersu, Jerome Baudry

Accepted for the Council:
Carolyn R. Hodges

Vice Provost and Dean of the Graduate School

(Original signatures are on file with official student records.)

Conformational Dynamics of Cytochrome P450cam Upon Ligand Binding

A Thesis Presented for the
Master of Science
Degree
The University of Tennessee, Knoxville

Ana Virginia Bernal Gomez
August 2015

ACKNOWLEDGEMENTS

For his guidance, I acknowledge my advisor, Dr. Nitin Jain. For his initial instruction, I acknowledge Nick Lopes. For the use of their neutron facilities, I acknowledge Oak Ridge National Labs Spallation Neutron Source and the National Institutes of Standards and Technology.

ABSTRACT

Cytochrome P450s are a superfamily of metalloenzymes that are responsible for the monooxygenation of their hydrophobic substrates. P450's retain the same general structural scaffold, however are able to bind promiscuously to substrates with distinct physico-chemical properties. It is believed that they possess considerable flexibility in the substrate binding regions as well as their active site to accomplish the diverse binding and catalytic chemistry with high regio- and stereo-specificity. In order to investigate the conformational dynamics inherent in these enzymes, especially in context of binding different ligands, we carried out amide proton exchange studies via NMR spectroscopy on a model P450 system of cytochrome P450cam (CYP101). These studies were performed in both camphor-bound and substrate-free forms of CYP101 and provide information on the motional properties of residues on slow timescales. Comparison of the exchange rates obtained from these studies for the two forms show that overall the substrate-free form is more dynamic than the camphor-bound form on the millisecond-second timescale due to its faster exchange rates, with regions being part of the substrate access site and active site in particular showing the largest differences in exchange rates. This study is the first of its kind looking at the residue-level changes in conformational dynamics upon ligand binding in a P450 and identifies specific regions where slow dynamic changes occur. Results from these studies help provide mechanistic insights into the process of differential ligand recognition by P450s.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I
DYNAMIC NATURE of ENZYMES

1.1 Enzyme Promiscuity Strategies

Enzymes are complex proteins that undertake specific catalytic tasks, allowing substrates to be catalyzed into product. Enzymes can be highly specific, catalyzing one or two substrates, or they can catalyze many substrates. The question of how this is possible has always been prevalent, and there are a few proposed hypotheses. One details how the overall structure and function of members of protein super-families catalyze completely different reactions. A prominent example of this is the α/β hydrolase family (1). While maintaining a catalytic triad, the different enzymes of this family can catalyze the breakage of carbon-carbon bonds, the hydrolysis of carbon-halogen bonds and ester bonds. In their work, the researchers note that there is less than 10% sequence homology between the members of the super-families, but their overall structure remains the essentially the same. This indicates that an enzyme that can perform a different function (or multiple functions) can more quickly diverge from it's ancestors and become more favored evolutionarily (1). There are several examples of such catalytic promiscuity, where one isoform of an enzyme can perform different chemical reactions on different substrates, such as carbonic anhydrase III, which has been known to perform hydration of CO_2 and hydrolysis of carbon esters and phosphoesters (1). However, there are also enzymes that can perform one chemical reaction on various different types of substrates. How these varying catalytic functions are accomplished is an intriguing question.

A hypothesis that is rapidly emerging and being widely accepted as to how these enzymes can be promiscuous is that these enzymes are believed to have high conformational plasticity. This is similar to the plasticity exhibited by intrinsically disordered proteins (IDPs) (2). The hypothesis on the promiscuity of IDPs is that their flexibility can change when bound to different binding partners, but is not necessarily a reflection of the ligand binding (3). A likely mechanism by which this promiscuity is

accomplished is that of conformational selection, or equilibrium dynamics. The conformational selection model is widely believed to operate in promiscuous substrate binding by enzymes, which typically are highly mobile entities. In this model, the enzyme does not have one structure, which fits exactly around the substrate, as per the lock and key model; neither does it have only one conformation, where it may slightly change it's shape to accommodate a substrate, as described by the induced fit model. Conformational selection occurs when the enzyme itself samples multiple conformations at once in solution, and the substrate favorably binds to one of those. The equilibrium of the conformations shifts, and the most favorable conformations are the ones that catalyze with the greatest efficiency.

Unlike IDPs, enzymes do not necessarily have to exhibit flexibility over the entirety of the protein. The equilibrium dynamics can be localized only to the binding or active sites. Parigi and colleagues observe that mutations of the binding site of the IDP, α-synuclein, do not affect the relaxation rates of the binding site, or any other parts of the protein, which is distinct from the equilibrium dynamics mechanism (3). This indicates that any localized perturbations are diffused over the entire protein and may not result in selection of certain conformations with distinct dynamic properties. It should be noted however that their relaxation measurements were done in the pico-nanosecond timescale, so it is possible that this protein is differentially dynamic on other timescales. Another difference between IDPs and enzymes that exhibit equilibrium dynamics, is that there is no stable, favored conformation for catalysis (3). This implies that these enzymes have natural high plasticity and this is how they accommodate multiple substrates, due to the fact that the catalytic efficiency is not greatly affected. The flexibility, or dynamics of the protein, allows it to bind multiple ligands. Aharoni et al observe that the substrate binding region of α-lytic protease has various loops, which increase in flexibility upon mutation. The mutation increases the promiscuous binding behavior of the enzyme by 10^5, while only decreasing the catalytic efficiency by half (4). An important component of the conformational selection model is that the enzyme must have structural flexibility (4, 5).

1.2 Conformational Plasticity of Cytochrome P450's

A prominent class of enzymes that exhibit great conformational plasticity are the ones belonging to the family of cytochrome P450s. Cytochrome P450s are promiscuous heme-containing monoxygenases, present in all kingdoms of life (6) that aid in biosynthesis and the metabolism of many xenobiotics, including drugs in humans. A typical reaction carried out by these enzymes is the monooxygenation of their substrates, in order to make them more soluble (Eq 1).

$$\mathrm{RH} + \mathrm{O_2} + 2\,\mathrm{e^-} + 2\,\mathrm{H^+} \rightarrow \mathrm{ROH} + \mathrm{H_2O} \qquad \text{Eq.1}$$

This reaction requires 2 electrons and 2 protons, which are supplied by NAD(P)H. The electrons are carried by an electron transfer protein, which varies depending on which P450 it is. Another interesting thing about the electrons is that they are transferred one at a time. The catalytic cycle (Figure 1) is as follows: The heme iron is in its +3 oxidation state when the substrate binds. The first electron is transferred following this, converting the iron to +2 oxidation state. Molecular oxygen then binds to the iron and the first electron is transferred from Fe^{+2} to oxygen, creating a radical. At this point, the second electron is transferred to the system, creating an iron-oxene. Next, two protons come and sequester an oxygen atom, forming H_2O which then leaves. This allows the oxygen and iron to form a complex. Rearrangement occurs, which makes the substrate susceptible to hydroxylation, forming the product, which then allows the hydroxylated product to exit the active site.

Monooxygenation reactions such as those carried out by P450s are notoriously difficult to carry out without an enzyme due to the high amount of energy required for activating the C-H bond for oxygen insertion and would require heating the reactants to very high temperatures to facilitate the reaction. Its ability to make such a reaction possible under ambient conditions has earned this enzyme the name, "Nature's blowtorch " (7).

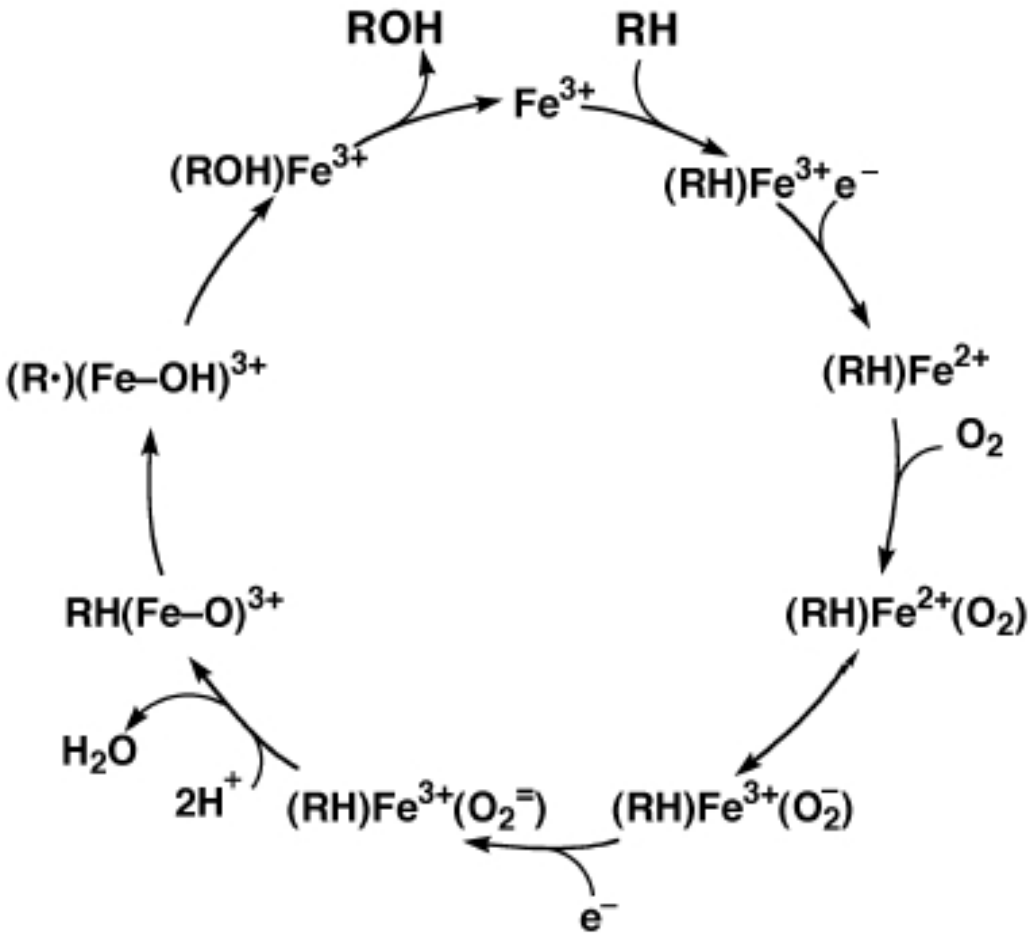

Figure 1: Catalytic Cycle of Cytochrome P450. The substrate RH enters and proceeds to be hydroxylated to the product ROH by a series of steps, including the transfer of two electrons by an electron carrier protein, and addition of oxygen molecule. (8).

A large percentage of pharmaceutical drugs (>90%) are metabolized by several isoforms of the P450 enzyme in humans using the above catalytic cycle. Although these enzymes can metabolize all of these drugs with very distinct physico-chemical properties, the efficiency of catalysis is not always optimal, leading to byproducts and drug toxicity. For example, the well-known non-steroid anti inflammatory drug (NSAID) acetaminophen, known by trade name Tylenol, is a substrate of CYP3A4. One of the intermediates that it produces, known as NAPQI, is a toxin known for interactions with other proteins, hence the high toxicity of this drug. P450's catalyze a hydroxylation and a rearrangement of acetaminophen, which is one more step than the usual hydroxylation (9). The reason for this may be that acetaminophen is not an ideal substrate for P450, which is why it can have toxic intermediates, and therefore toxic side effects. If the precise recognition and binding of various drugs was known, pharmaceutical companies could avoid this to a large extent. Members of the P450 superfamily, whether bacterial or human (Figure 2), retain remarkably similar structural architecture or fold, while adapting to varying needs of substrate recognition and the regio- and stereo- chemistry of the reactions catalyzed using a dynamic rearrangement of their secondary structural features depending on the substrate, cofactor or oxidation state. Thus, one of the main focuses recently, apart from elucidation of P450 structural factors, is the investigation of how this conformational flexibility of cytochrome P450 enzymes allows for modulating of protein-substrate interactions to accommodate substrates of differing physico-chemical properties as well as product specificity. For this reason, studying the dynamics of P450s will help in the fundamental understanding of its substrate promiscuity.

The extent of flexibility of these enzymes became apparent with the discovery by Scott and coworkers that a soluble version of cytochrome 2B4 from rabbit exhibits a 15 Å separation of the components of the substrate binding regions as part of a large open active site, making it capable of large displacements to accommodate substrates of varying size (10). Normally, this open conformation to allow substrate access to the

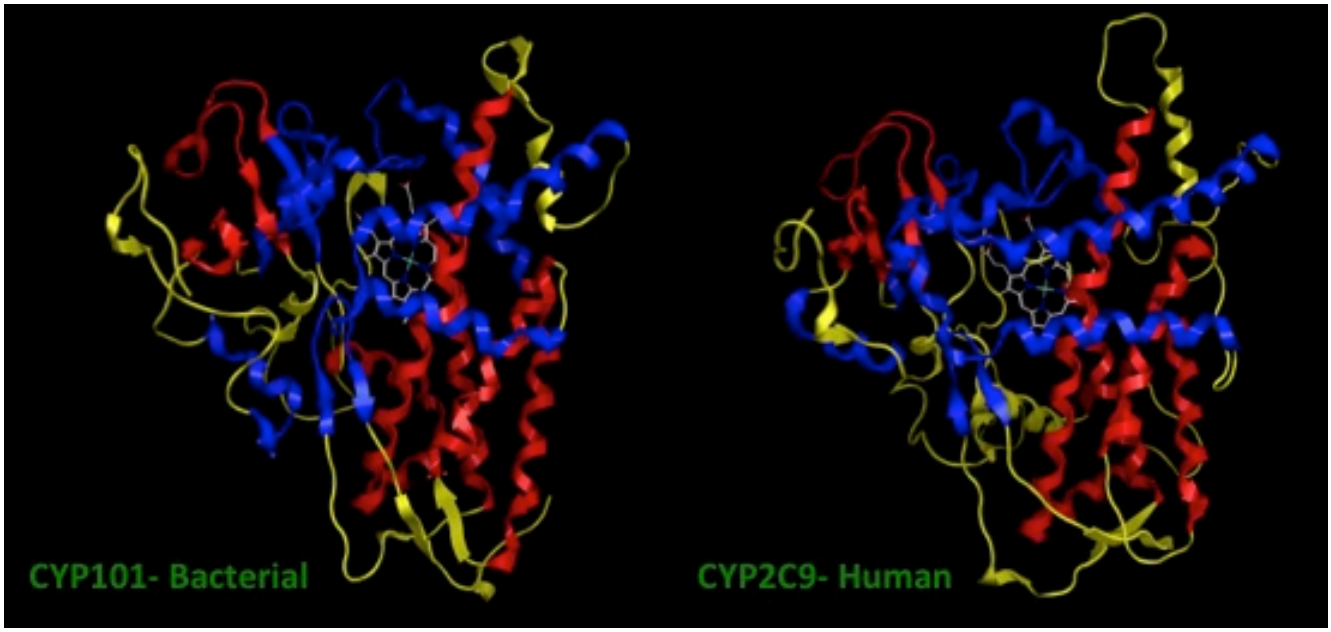

Figure 2: **Structural Fold Similarities of cytochrome P450s.** The structures of a bacterial P450, CYP101 (left panel) and a human P450 isoform, CYP2C9 (right panel) are shown. Although some differences exist, the general structural fold is maintained, despite a sequence similarity of only 10-15% (11).

active site is marked by a cleft of 10 Å or less in any P450 structure solved previously (10, 11). The substrate access channels, present in membrane bound P450's, also revealed openings of only 2-3 Å, which does not provide enough space to allow the substrate to pass through the channel (11-13). Yet, somehow they do. This allowed for the notion to develop that these enzymes have considerable flexibility, which could be exploited to bind diverse substrates, providing impetus for further dynamic studies of P450s by both computational and experimental methods.

Crystal structures of several P450s, both bacterial and from other higher organisms, reveal considerable variability in regions around the active sites of these enzymes. While the active site comprising the heme center is buried inside the protein, the substrates access the active site via an access channel comprised of secondary structure elements collectively known as the substrate recognition sites or SRS (Figure 3). The SRS typically include several helices such as the B', F, G and I helices as well as some beta-sheets such as the β 1, 3 and 5 sheets (6, 11, 14). Other P450's generally exhibit the same SRS, although in the human P450s there are additional identified flexible regions (11, 15). Also, the length of the flexible regions varies between the prokaryotic and eukaryotic P450s, with mammalian P450s having longer F-G and B-C loops, leading to generally increased flexibility in these regions for the mammalian ones and with better adaptability to accommodate substrates of increasing size due to larger binding pockets. Essentially, there have been many studies that have determined that multiple loops and their adjoining helices are a part of the SRS machinery that allows P450s to be so dynamic (15). However, no matter the components, the placement of the SRS is generally in the same location. In spite of this knowledge, it is not entirely clear which specific residues within the various P450s undergo the greatest rate of dynamic change to effect substrate binding, and how that change is affecting the overall dynamics of those SRS.

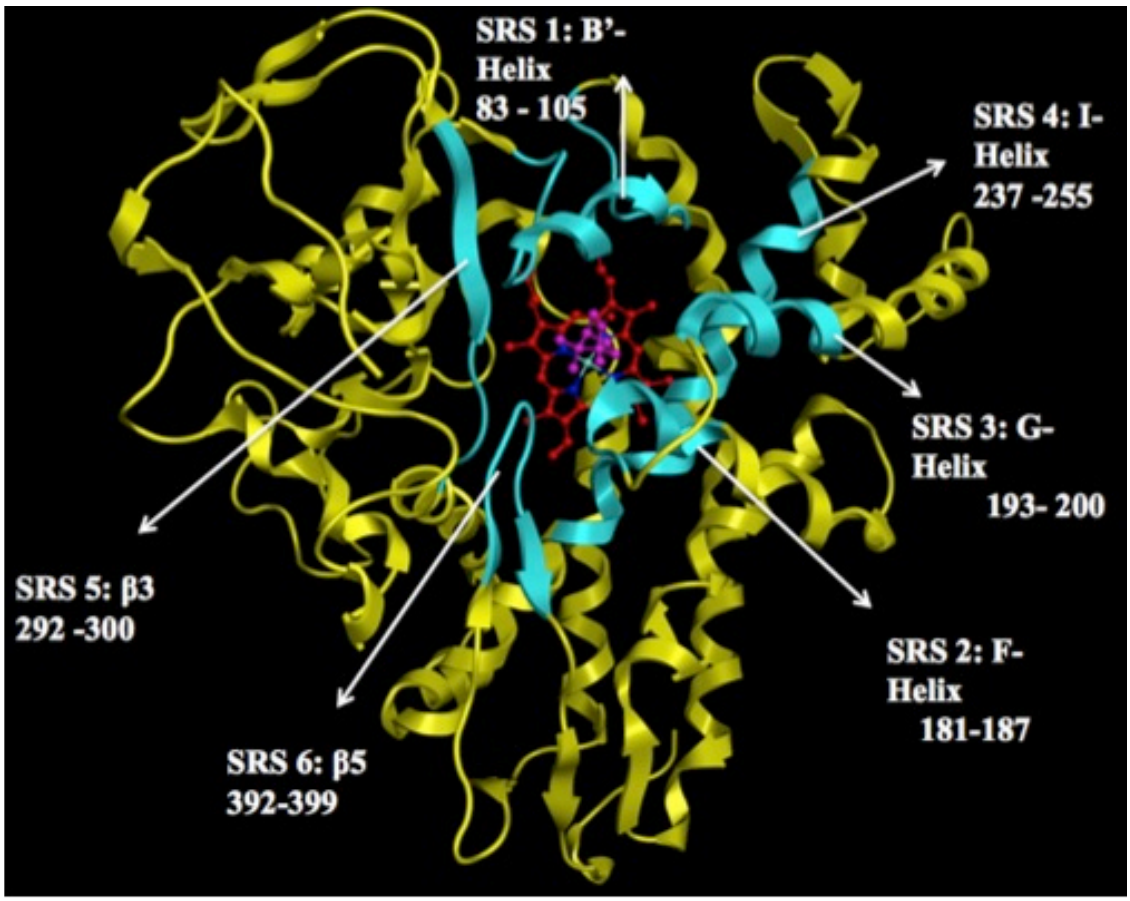

Figure 3: Substrate Recognitions Sites of CYP101. The various substrate recognition sites (SRS's) are structurally highlighted in cyan, and are highly dynamic regions of the protein, which facilitate binding of a multitude of ligands to CYP101. SRS's are numbered, with each numbered SRS containing the structural element as well as residue numbers within that element.

Computational studies by Wade and coworkers have attempted to identify pathways by which substrates may enter into and products may exit from the active site using molecular dynamics simulations on several P450s (13). They found existence of many access channels with some common ones among all P450s that involved several flexible elements of the SRS such as the F-G, B-C helices and loops. These studies again indicate the dynamic nature of the access/egress paths and provide functional rationalization for the structural variability of these regions in the crystal structures solved for the different P450s so far. The dynamic behavior of these regions also makes it difficult to use computational modeling and dock compounds into the SRS or active site of a P450 during the process of computational drug design.

The flexible nature of these enzymes have to be taken into account during the docking process and without sufficient knowledge of the conformational selection a P450 can undergo, it is difficult to simulate the docking process and know which conformation a compound is most likely to bind to, because it is unknown. The mechanism by which these conformations are selected is unknown, and how they switch from one to the other is also unknown. Therefore, experimental approaches to augment the dynamic information during the docking process can help with more efficient discovery of drug-like molecules and prevent unwanted levels of toxicity due to incomplete catalysis or weak binding when applied in practice. Experimental approaches incorporating NMR, EPR, IR and mass spectrometry are increasingly being used (15-17) to characterize the dynamic properties of P450 enzymes, however systematic studies involving several ligands and how the dynamic properties of these enzymes change in response to different ligands have not been carried out yet, which are essential to understand the structural-dynamic relationships in these enzymes. Discovering how the dynamics of various P450 regions are key in ligand recognition and the specific mechanism by which they accomplish differential ligand binding will be a key piece of knowledge in further understanding how exactly these proteins work.

CHAPTER II
INVESTIGATION OF DYNAMICS OF A MODEL P450 ENZYME, CYTOCHROME 450CAM (CYP101)

2.1 Cytochrome P450cam as a Model System

Given the conformational plasticity of cytochrome P450s and variability in the SRS, assessing the dynamic and mechanistic factors that give rise to conformational selection within the individual P450s as a function of ligand (substrate or inhibitor) binding is an important goal among the researchers in the P450 area. A dynamic comparison between distinct ligand bound states could provide further evidence that conformational selection may be a ubiquitous mechanism within the P450 superfamily. The first cytochrome P450 to be structurally and mechanistically characterized in detail was cytochrome P450cam (CYP101) from the bacterium, Pseudomonas putida (18). CYP101 is a hydroxylase for its natural substrate camphor, which the bacterium uses as carbon and energy source. It utilizes two electrons to catalyze its reaction that are transferred to it by its redox partner, putidaredoxin (Pdx), a [2Fe-2S] ferredoxin from the same organism. Since its structural characterization by X-ray crystallography, it has served as model system for further mechanistic investigation, such as elucidation of catalytic mechanism of P450s, ligand entry and binding to active site as well as redox partner binding and electron transfer. CYP101 is thus a well-studied cytochrome and there is much information available on it. It is soluble and the methods of growth and expression have been well optimized. It is also becoming increasingly evident that it is a fairly promiscuous enzyme with ability to bind to many ligands (*vide infra*) and exhibiting conformational flexibility rivaling other P450 enzymes, due to large structural differences (~ 10 Å) between its camphor-bound (Figure 4) and its wide open substrate-free conformation (Figure 5) (19). Therefore, it forms an ideal system to initiate dynamic studies since the knowledge gained by studying this model system will provide insights into how other P450s are also likely to utilize their flexibility in effecting ligand binding due to their structural similarities.

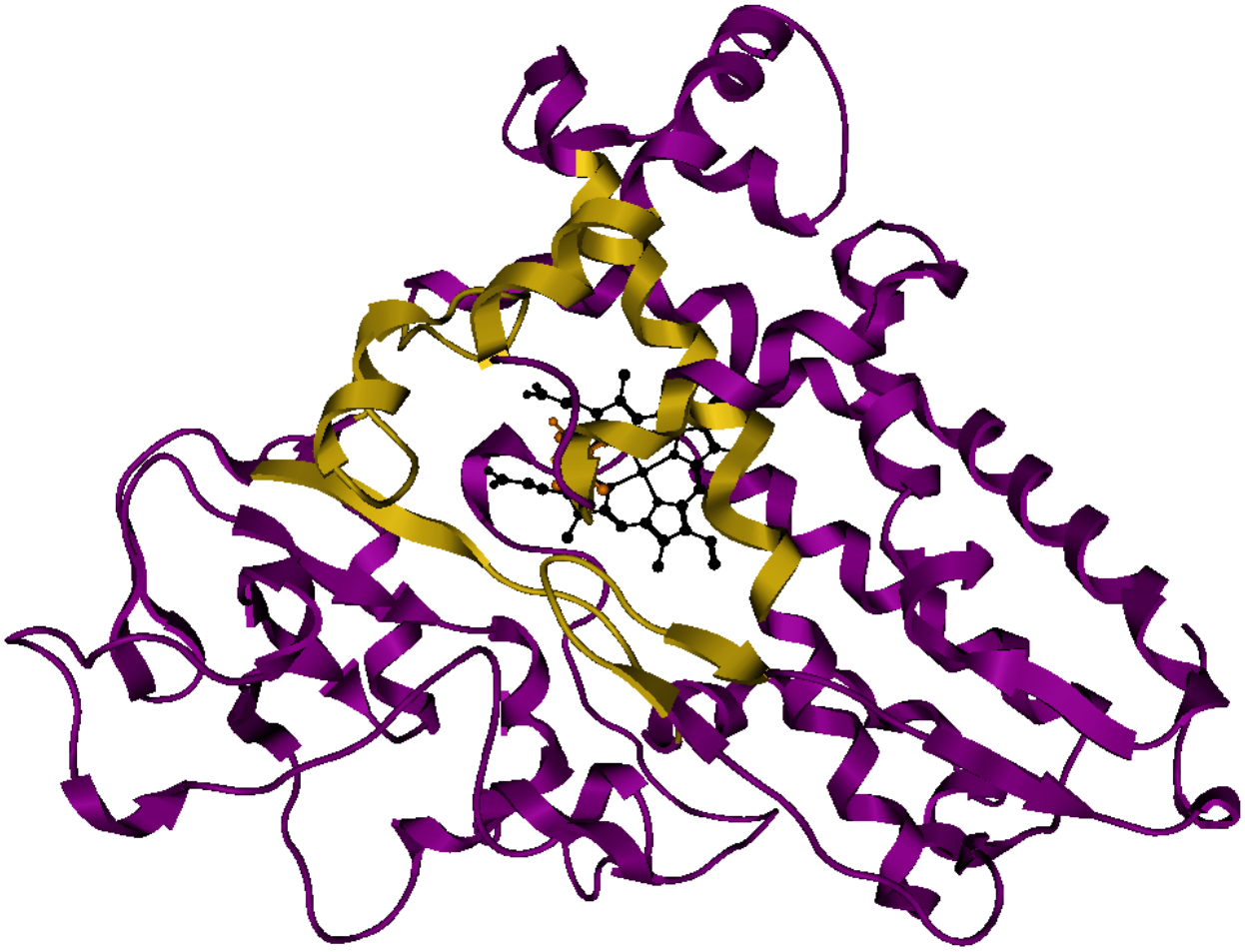

Figure 4: X-ray crystal structure of camphor Bound CYP101. SRS is highlighted in gold. As shown, the various SRS elements are enclosed over the active site, allowing the substrate to remain within the active site. PDB ID: 3L63 (19)

Figure 5: X-ray crystal structure of substrate-free CYP101. The various SRS elements (shown in gold) have moved away from the active site, taking on a more open conformation. PDB ID: 3L62 (19)

CYP101 is roughly triangular in shape with the heme active site embedded deep inside the protein across from the I helix (Figure 4). The substrate camphor accesses the active site via a water-filled access channel lined by the various SRS elements such as the B', F and G helices, β3 and β5 sheets as well as the B-C and F-G loops (Figure 3). Upon entry, camphor is anchored in the active site by hydrogen bonding to side-chain of Tyr 96 and hydrophobic interactions with the surrounding aromatic and aliphatic residues in the active site pocket (Figure 6). The occupancy of camphor in the active site displaces the water coordinating the iron at the 6th coordination site, converting the substrate-free form of CYP101 to the substrate-bound form. This type of conversion upon substrate binding is common to all P450s with their respective substrates. Inhibitors of CYP101 enter and bind via similar mechanism, although they may be directly coordinated to the heme at the 6th coordination site unlike camphor, which sits in close proximity to this site. The 6th site in the camphor-bound form is then coordinated with oxygen, necessary for the hydroxylation reaction, which distinguishes the catalytically competent substrate binding form from the inhibitor binding form (Figure 7).

While it is generally accepted that the SRS facilitate binding of various ligands via a dynamic mechanism, it is not clear how dynamically different the various SRS elements are in the presence of different ligands. It is unknown which specific residues undergo the greatest rate of dynamic change, and how their changes affect the overall dynamics of the various SRS elements. Also, can we discover in which timescale these residues are dynamic and what mechanism controls their motions to allow binding of multiple ligands? Detailed experimental investigations on these questions have not been carried out. If the protein dynamics of CYP101 can be investigated in absence and presence of various ligands, then it would be interesting to know if the conformational selection model is used in CYP101 and generally as a mechanism for all of P450 binding and catalytic events.

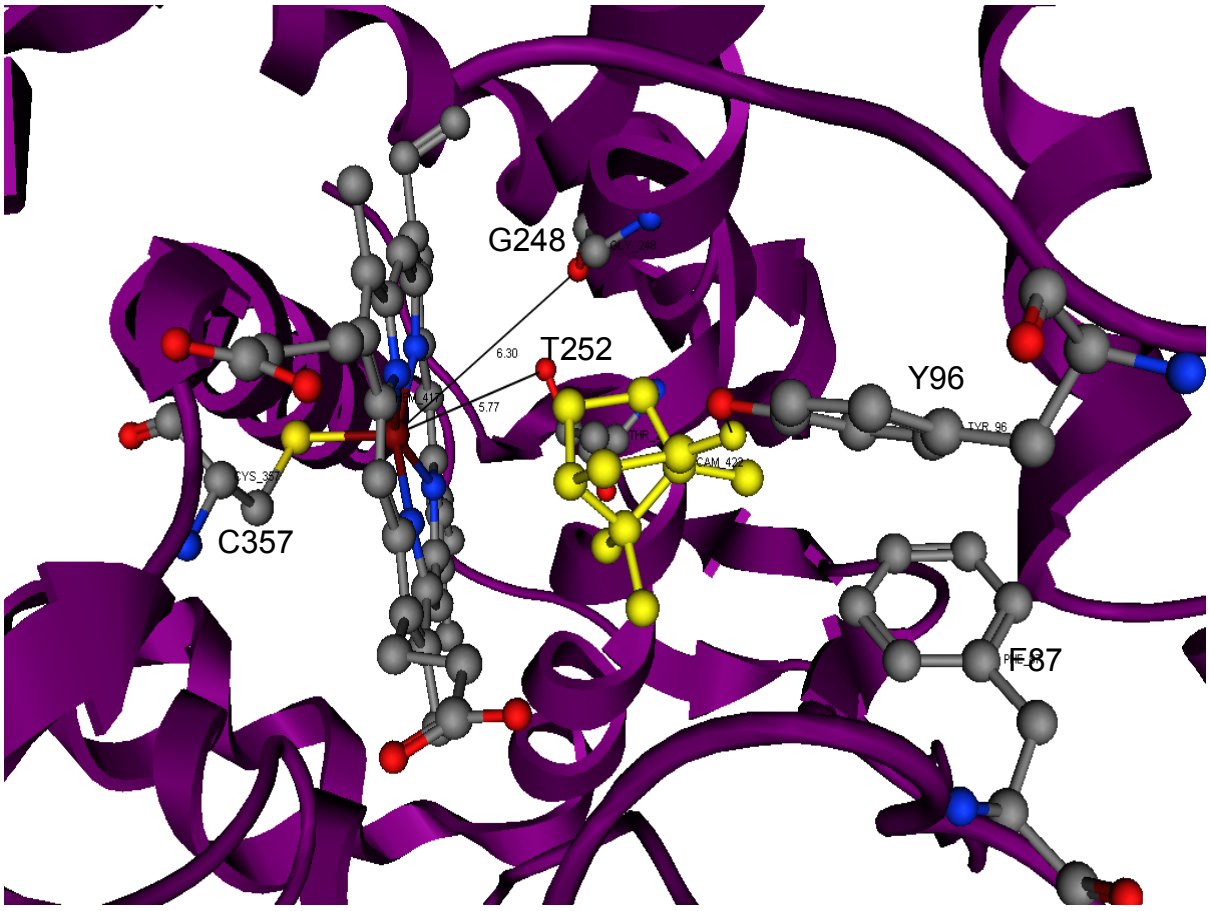

Figure 6: Active Site structure of Camphor-Bound CYP101. The coordination of the 5th heme site is shown with the thiolate ligand formed by cysteine 357 (behind the heme plane) which causes a slight puckering of the heme. The substrate camphor occupies the region near the 6th coordination site. Tyrosine 96 forms a hydrogen bond with camphor and positions it with the 5th carbon available for hydroxylation. PDB ID: 2CPP (18).

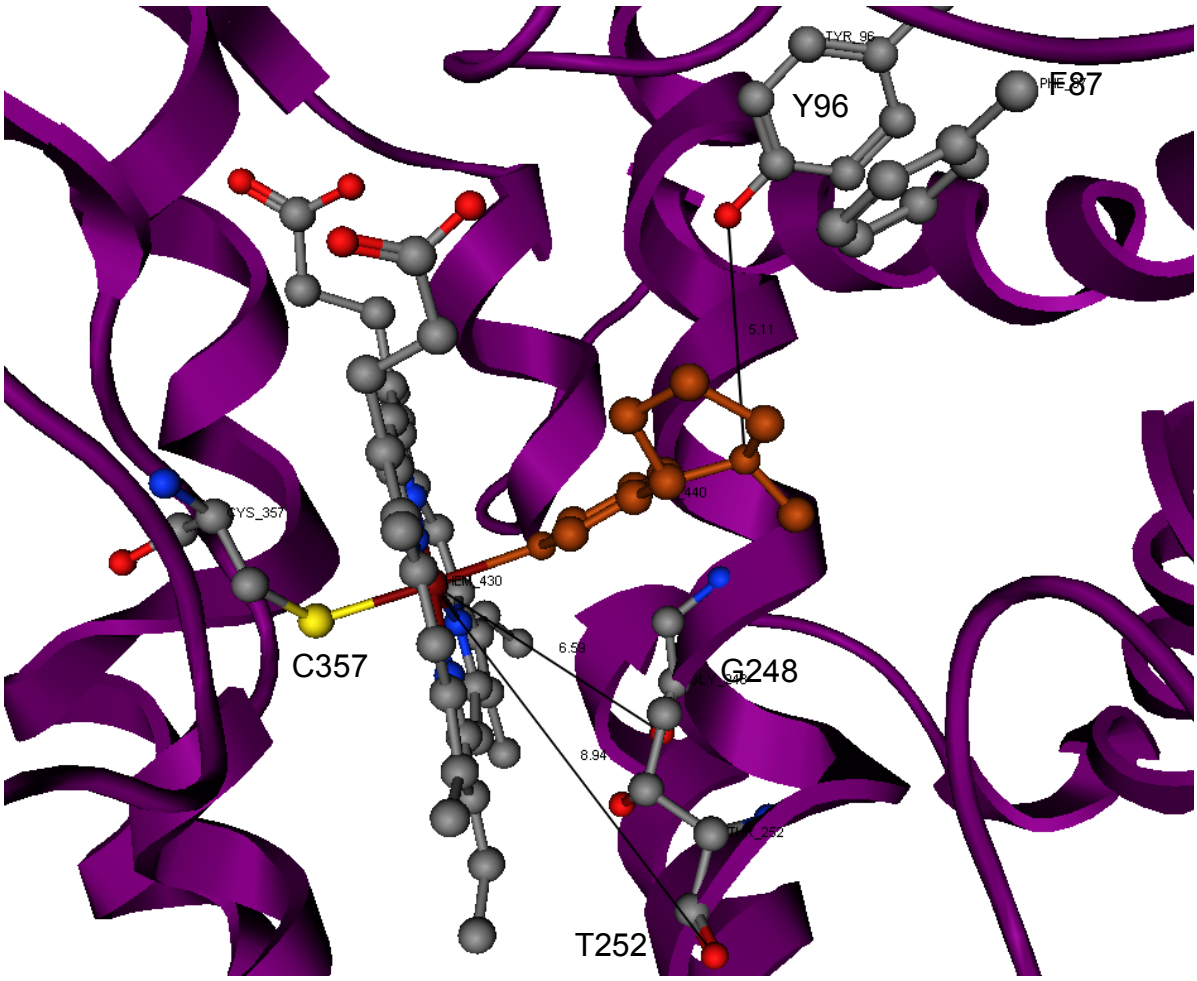

Figure 7: Active site structure of Nicotine-Bound CYP101. The coordination of thiolate ligand, Cys357 to the 5th coordination site is shown. A coordination bond is formed between nicotine and the heme iron at the 6th coordination site, removing the puckering of the heme plane. Tyr96 side-chain also hydrogen bonds to another nitrogen, further holding the nicotine in place. PDB ID: 1P2Y (20)

2.2 Methods for Studying Protein Dynamics

Proteins exist as an ensemble of conformations that are perpetually switching, rather than a static entity. All structures, whether determined by crystallography, NMR or neuron scattering, represent an average of all the conformations a protein can adopt. Proteins use dynamics to facilitate their function. Here, dynamics is defined in the words of Henzler-Wildman and Kern as “any time dependent change in atomic coordinates”(21). There are equilibrium and non-equilibrium dynamic effects, but those at equilibrium are the most relevant to protein function as a result of dynamics.

Dynamics is defined as a time dependent reversible process, and different types of motions occur at different timescales. Very fast motions include bond vibration, which occurs at the femtosecond timescale. Bond rotations and side chain rotamerization occur 10-100 fold slower, at the pico to nanosecond timescale. Loop motions span the nano to microsecond timescales. The millisecond and longer timescales include all other large domain motions, including helical shifts and overall protein breathing (Figure 8). All of these timescales can be studied using different techniques, and there are optimal techniques for each timescale (21). Molecular dynamics simulations are appropriate for the faster timescales, due to the optimization of parameters and providing accurate atomic level description of experimental data. Slower timescale data produced by MD is less accurate when compared to experimental methods. MD simulations have been used with experimental techniques such as neutron scattering to allow visualization of protein dynamic motions on a multitude of timescales, from femtoseconds to microseconds, by selection of trajectories that best fit experimental neutron scattering measurements on these different timescales (22). Although this requires considerable amounts of protein samples to obtain good quality neutron scattering data, in favorable cases neutron scattering methods can be used to contrast amplitude of dynamic motions between various forms of the protein, such as complexed or non-complexed forms and study proteins without the influence of solvent (23).

Forster resonance energy transfer (FRET) is another popular technique used to study protein dynamics. It utilizes two energy transferring fluorophores of differing wavelengths of light. If the two tags come in contact with one another, the emitted photon is of a certain energy, the spectrum of which can be interpreted both in structural and dynamic terms (24). FRET is utilized to determine nano to microsecond dynamics, however does not give high-resolution atomic-level dynamic information (24). Also, since FRET requires the addition of the fluorophore tags on the protein, the protein may not be folded appropriately, depending on the placement and size of the tags. The dynamics may also be affected by the placement of the tags.

Another technique to study protein dynamics is X-ray crystallography, which has the potential to provide information on distinctive dynamic motions at the nanosecond timescale by interpretation of electron density maps as shown in a recent study by Fenwick and colleagues on the enzyme dihydrofolate reductase (25). X-ray crystallography however requires the formation of crystals, which may not necessarily reflect the natural environment of the protein. In order to form the crystals, proteins must pack in a certain way, which may trap the protein in a particular conformation or freeze out a subset of motions observed in solution. Screening techniques for crystal formation also requires a large amount of protein. In spite of these inherent limitations, newer methods in X-ray crystallography such as Laue diffraction methods or femtosecond serial resolution crystallography allow atomic-level resolution of dynamic events by taking snapshots of the proteins in crystals as they perform their function (26). However, these techniques suffer from technical limitations in terms of the kind of systems they can be applied to obtain dynamic information.

NMR spectroscopy is one of the more versatile spectroscopic techniques to study protein dynamics, as it is able to probe picosecond to minute timescales. NMR is also non destructive to the sample and requires relatively low concentrations and volume of protein. There are different NMR methods to probe time-dependent fluctuations in structure on different timescales.

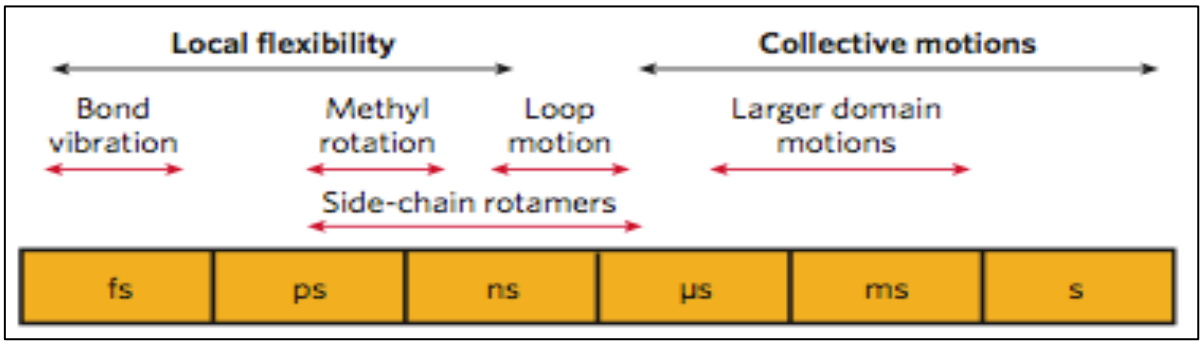

Figure 8: Timescales of different protein motions. Dynamics is a time dependent process, as different motions can be seen on different timescales.

Nuclear spin relaxation rates T_1 and T_2 can be used in a site-specific manner to probe motions in the ps-ns timescale across the protein such as bond vibrations, loop motions, sidechain rotations and backbone motions (27). Lineshape analysis of exchange-broadened resonances and more recently CPMG relaxation dispersion methods have been used to quantify exchange processes in the us-ms timeframe such as secondary structure reorientations, slower loop motions and large-scale domain movements (28, 29). Dynamic processes on the ms-min timescale can be detected by monitoring NMR signal intensities as a function of time that give information about slower motions occurring in dynamic events such as protein folding, conformational disorder and overall protein breathing motions that may affect binding and catalytic rates. A common way of measuring these types of dynamic behavior is to use hydrogen-deuterium exchange (HDX), whereby the exchange of hydrogen atoms between protein and the solvent is monitored over a period of time. The HDX technique can be used in conjunction with either NMR or mass spectrometry, and has been used for many years to observe protein folding and dynamics. One of the first measures of protein dynamics was on the CD2 domain of immunoglobulin, which determined that the domain has an intermediate which forms after global opening, which has highly protected amides, while the native state has non protected amides, which is relevant for the function of this domain (30). Recently, HDX has been used to determine the dynamics of membrane proteins using solid state NMR (30). In the early days of matrix assisted laser desorption/ionization time of flight (MALDI-TOF) mass spectrometry (MS), Mandell and coworkers used the catalytic domain of PKA in order to determine how much the instrument could distinguish between fragments, and deuterated and hydrogenated samples (31). After this time, amide exchange has been a prominent tool in the determination of slow protein dynamics by mass spectrometry.

In HDX NMR, typically the exchange of backbone amide protons with deuterium is monitored after being placed in a completely deuterated solvent such as D_2O. Since deuterium is not detectable in the ^{1}H-^{15}N HSQC experiment, the exchange process can be monitored as a drop in signal intensity in the HSQC spectrum due to loss of protons.

The protons exchange differentially based on their relative exposure to the solvent which depends on the “breathing” motions of the protein, dictated by the equilibrium between the open and closed states of the protein. As such, this type of dynamic motion can give information on the relative accessibility of different regions in the protein to the solvent and is indicative of the nature of conformational dynamics i.e. fast or slow between various forms of the protein. The comparative nature of the HDX data gives a good starting point to characterize the global dynamic properties of proteins and see where more quantitative site-specific dynamic studies on other timescales can be focused in the future.

2.3 Conformational Dynamics of CYP101

The X-ray crystal structures of CYP101 have been solved in both substrate-free and various ligand-bound forms (19). Comparison of the structures reveals that substrate-free CYP101 samples a distinct open conformation marked by retraction of the F and G helices and disordering of the B’ helix (Figure 5), while the camphor-bound form exhibits a closed conformation with the same elements highly ordered due to ligand association (Figure 4). The open conformation is similar to those seen in the structures of CYP101 with tethered substrates (19), suggesting that CYP101 is dynamically capable of undergoing substantial structural changes and visit different conformations sampled by substrate-bound forms, indicative of a conformational selection model operating in ligand binding. Subsequent double electron-electron resonance (DEER) experiments also show existence of this open conformation in solution which converts to a closed conformation upon binding camphor, similar to that observed in the crystal structures (32). Accelerated MD studies on substrate-free CYP101 further reveal existence of a cluster of partially open conformation in addition to the open conformation observed in the crystal structures, suggesting that CYP101 is capable of populating different conformational ensembles (33).

Another experimental study to investigate conformational dynamics of CYP101 was performed by the Ortiz de Montellano group, which looked at how exactly a large substrate such as ketoconazole, a potent antifungal, is able to bind its target P450s (34, 35). Ketoconazole was designed as a target for human CYP3A4, however is able to bind to bacterial P450s such as CYP101. In performing the mutation L244A, they found interesting implications on the binding of inhibitors, where the mutation was able to enhance the conformational plasticity of the active site. Even though the mutation involved replacing a larger amino acid with a smaller amino acid in the active site, the binding of a small inhibitor imidazole showed much lower affinity than the wild-type protein, although the binding of larger inhibitors such as ketoconazole was not affected. Due to the mutation, the I helix shifts inwards preventing imidazole from taking an orientation to bind favorably to the heme, On the other hand, the mutant active site is able to conformationally adjust to accommodate larger substrates such as ketoconazole, demonstrating the high conformational plasticity of the active site. Whether the same conformation is sampled also in the wild-type protein or in the open substrate-free conformation remains to be seen, as the crystal structure of CYP101 with ketoconazole has not been solved yet.

The conformational plasticity of the P450 active site can also be seen from NMR studies on substrate behavior in the active site using paramagnetically-induced spin relaxation of water and substrate in CYP102, which indicated dynamic movement of the substrate in the active site (36). Similar reorientation of substrate has been observed for camphor and a related substrate, norcamphor, in the active site of CYP101 using NMR and fluorescence anisotropy measurements (37). In addition, NMR studies on CYP101 complexation with Pdx showed that binding of Pdx on the proximal side of heme perturbed resonances on the distal side, especially in the substrate binding regions, which was hypothesized to occur as a result of Pdx-enforced selection of a subset of conformations that prevented loss of substrate before catalysis (38). Furthermore, from NMR T_1/T_2 relaxation experiments on oxidized and reduced forms of CYP101, it has been found that the oxidized form of the enzyme is more dynamic than the reduced

form and analysis of amide exchange rates as well as backbone dynamics of the protein from relaxation data indicated that the dynamic differences while delocalized over the entire protein were more pronounced in the SRS regions, providing strong evidence that CYP101 may sample different subsets of conformations in the oxidized and reduced forms (15).

2.4 Preliminary NMR studies on CYP101 in presence of different ligands

The abundance of data from the various crystal structures in ligand-free and ligand-bound forms of CYP101 as well as the numerous spectroscopic studies carried out on CYP101 in different forms strongly suggests that CYP101 makes use of its inherent flexibility in binding ligands of differing physico-chemical properties. However, it is not clear from these studies what specific mechanism is utilized by CYP101 to select for the appropriate conformation(s) from the multitude of conformations available to it. A systematic study investigating the changes in conformational dynamics in the absence and presence of different substrates and inhibitors to delineate specific binding mechanisms has not been performed so far. We therefore carried out preliminary NMR studies using ^{15}N labeled samples of CYP101 in absence and presence of different ligands to obtain insights into the type of structural and dynamic differences apparent with these binding events, since NMR spectra are sensitive to both these parameters. ^{15}N-^{1}H HSQC-TROSY spectra of ^{15}N labeled oxidized CYP101 bound to camphor, nicotine and ketoconazole were collected in presence of saturating amounts of ligands. The resulting spectra were compared with each other and also the ^{15}N-^{1}H HSQC-TROSY spectrum of ligand-free CYP101 (Figure 9). Comparison of the various spectra produced striking observations. Overall, the chemical shifts between the spectra exhibit only small changes, so much so that it is possible to relate the peaks in each form on a one on one basis for most part. The chemical shifts in the spectra are also well dispersed indicating that the protein is similarly and correctly folded in each form.

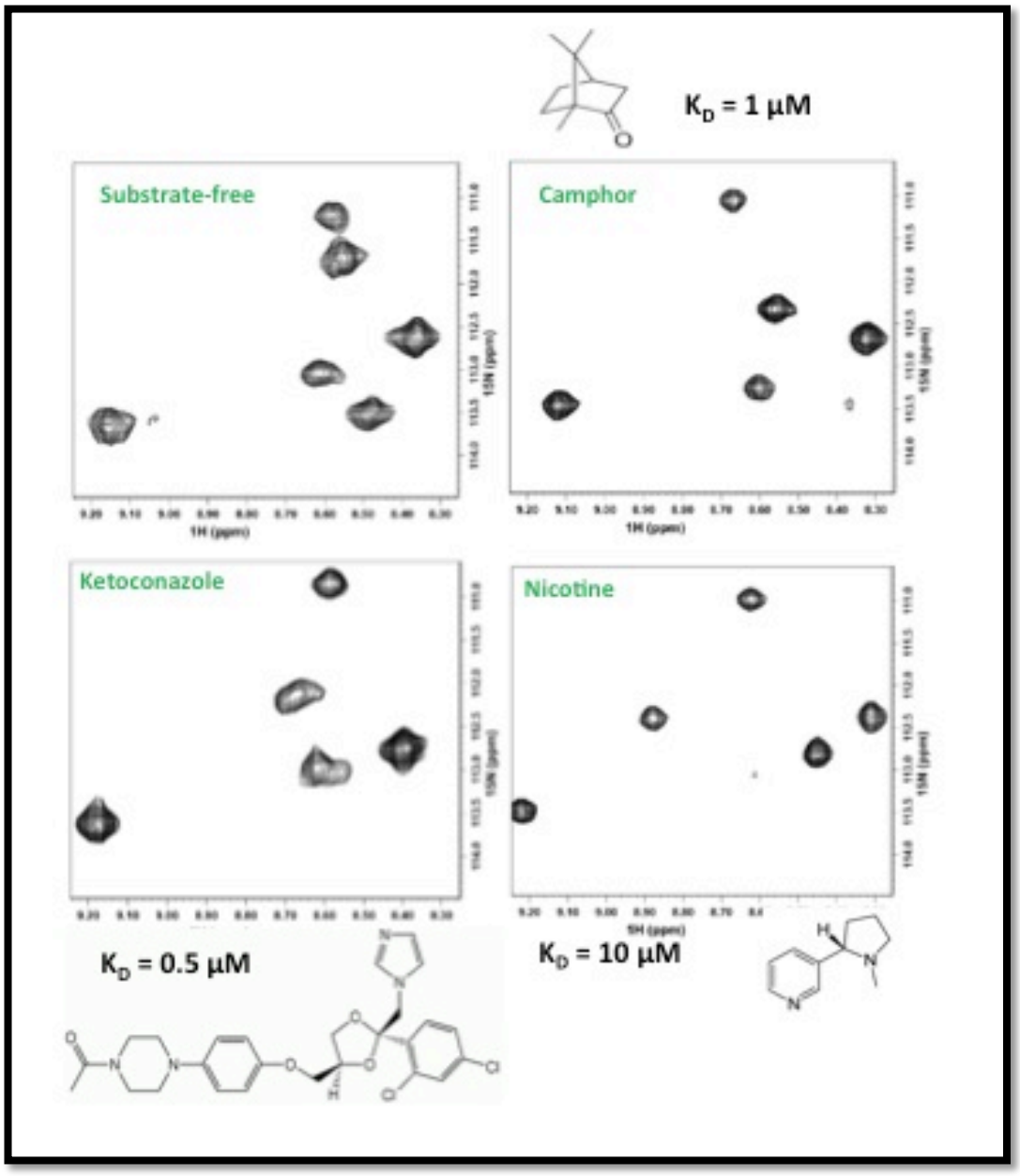

Figure 9: Differential dynamics of CYP101 upon binding to different ligands. Portions of 2D ^{1}H-^{15}N HSQC-TROSY NMR spectra of 15N-labeled CYP101 are compared with and without various ligands. Dynamic differences in terms of linewidths and peak characteristics are clearly observed amongst all spectra. The relevant affinities for all ligands (K_D values) to CYP101 are listed next to each spectrum and the structure of each ligand shown as well.

However, the spectrum of ligand-free CYP101 exhibited relatively more conformational heterogenity for the protein across much of the spectrum as evidenced by excessive line-broadening and splitting of single peaks into multiple peaks for several resonances in the spectrum. More than 80% of the peaks show line-broadening, with almost 40% of the peaks showing presence of multiple peaks. This is in contrast to the spectrum observed for the camphor-bound and nicotine-bound forms, which exhibited progressively less of these characteristics. The spectrum of nicotine-bound form contains peaks with the narrowest linewidths (almost half linewidth relative to the substrate-free form). This is observed for almost all the resonances in the spectrum. Interestingly, the spectrum of ketoconazole-bound form of CYP101 exhibits conformational heterogeneity similar to that of the substrate-free form with almost 80% of the peaks affected. This provides direct evidence that the various ligand-bound forms of CYP101 are experiencing conformational exchange on different timescales. This differential dynamics likely arises from the conformational selection of a different subset of conformations of each form of CYP101, in response to the ligand it binds. Such direct observation of differential dynamics in presence of various ligands has not been observed previously for a P450 by any experimental method.

It is interesting to note that the affinity of each ligand is very similar and within an order of magnitude of each other (Figure 9). Ketoconazole binds with the highest affinity, followed by camphor and then nicotine with the lowest affinity. Based on the affinity differences, it would be expected that the protein should lose most of its conformational mobility upon binding to a large ligand such as ketoconazole with such high affinity, presumably due to a multitude of potential binding contacts with the active site residues of CYP101. It is therefore surprising that the spectrum of ketoconazole-bound form exhibits similar, if not more, conformational heterogeneity compared to the substrate-free form. On the other hand, conformational heterogeneity and mobility seems to be lost upon weaker binding to the nicotine ligand, as indicated by the much narrower resonances in the nicotine-bound CYP101 spectrum. The binding of the

natural substrate camphor exhibits dynamic characteristics intermediate between that of ketoconazole and nicotine in line with its intermediate binding affinity.

A likely explanation for this surprising observation may come from the thermodynamic aspects of ligand binding, specifically a phenomenon known as entropy-enthalpy compensation. Entropy-enthalpy compensation in terms of ligand binding is defined as a change in enthalpy (ΔH), which is offset by a change in entropy (ΔS), resulting in no significant change to the overall free energy of binding (39). It is plausible that there is an additional entropic contribution in the binding of ketoconazole due to the increased conformational dynamics in the protein that is not present in the binding of nicotine. Crystal structure of CYP101 with nicotine in the active site shows that apart from direct ligation with the heme, nicotine makes only a few specific interactions with residues in the active site such as hydrogen bonding withTyr96 (20). The residue environment around nicotine in the active site is fairly ordered as evidenced by low B-factors for most of the residues in that region. Thus, binding of nicotine is very likely an enthalpically driven event. On the other hand, the increased conformational heterogeneity for substrate binding region residues upon binding ketoconazole indicates that the active site has not adjusted to fit the shape of ketoconazole entirely and therefore ketoconazole may retain considerable mobility in the active site, similar to the protein residues. This may preclude forming a multitude of contacts of CYP101 with ketoconazole, making a solely enthalpy-driven binding unlikely. A recent crystal structure of CYP3A4 with ketoconazole in the active site supports this premise, since very few contacts are observed between the protein and ketoconazole ligand in the active site, allowing much more freedom of movement for the ligand in the active site as well as higher B-factors for the surrounding residues (40). It is possible that this additional freedom may allow ketoconazole to adopt different orientations in the active site of CYP101 or bind to other sites on the surface of the protein, giving rise to the observed conformational heterogeneity. However, in CYP3A4, even though the active site is much larger than CYP101 and access to multiple orientations around the heme in 3A4 is more likely, ketoconazole is not observed to take multiple orientations (41). The

CYP101 active site is even smaller and may not allow the ketoconazole to adopt multiple orientations. We cannot rule out alternate binding modes for ketoconazole, however based on observations of single ligand binding in CYP3A4, this seems unlikely.

To account for the high observed affinity, one can propose that the loss in enthalpic interactions is compensated by an increase in the entropic contribution from the increased dynamics in the protein and ligand. It is possible that the increased entropy may also result from the release of disordered solvent molecules upon binding of ligands or may result from desolvation effects. However, our observation that the dynamic changes are propagated throughout the protein and not restricted only to the binding sites indicates that it is largely a protein-centric event as the protein adjusts its mobility differentially in response to the ligand binding. Further desolvation studies for the various ligands or osmotic stress experiments to monitor the release of solvent molecules however are clearly needed to test this hypothesis and rule out solvent-based contributions (43-47).

Such increased entropic effects upon binding have been observed previously in other protein systems. In a study done by Wand et al on calmodulin, it was observed that there was a significant change in the conformational entropy of the amino acids of the protein that bound to different target peptides (21). This indicates the possibility that in order to accommodate different substrates, parts of CYP101 may change their conformational dynamics. It is important to know the identity of these regions and the timescales of motions that are affected the most. In this project, we therefore seek to elucidate specific residues that undergo a change in their dynamics and their rate of dynamic motion using hydrogen-deuterium exchange (HDX) detected by NMR spectroscopy. We propose that CYP101 utilizes a dynamic change mechanism as a means to effect differential binding of ligands as part of an overall strategy involving entropy/enthalpy compensation. HDX experiments will allow us to characterize these dynamic changes on a global scale and provide insights into the types of slower motions that facilitate the changes observed in the NMR spectra of the various ligand-

bound forms. Furthermore, these experiments will set the stage for investigation of the full spectrum of dynamics on all timescales using CPMG exchange and NMR spin relaxation methods.

CHAPTER III
PREPARATION AND CHARACTERIZATION OF CYP101 IN PRESENCE OF VARIOUS LIGANDS

Cell Growth and Protein Expression: Competent BL-21 (DE3) *E.Coli* cells were transformed with a plasmid that encodes CYP101 with a fused His_6 tag. A single colony of BL-21 cells was used to inoculate 50 mL of sterile LB medium containing 50 µL of both kanamycin and chloramphenicol antibiotics, each at a stock concentration of 50 mg/mL. After the cells reached a certain optical density (OD_{600} ~ 0.6), they were transferred to sterile centrifuge tubes and centrifuged for 15 minutes at 6000 RPM and 4°C. The cell pellet was then resuspended in 1 liter of defined M9 media (M9) consisting of the following ingredients: 7.0g sodium phosphate dibasic anhydrous, 3.5g of potassium phosphate monobasic and 0.5g of sodium chloride/L, 1mL each of kanamycin and chloramphenicol antibiotics, 1g of ammonium chloride (N^{15} ammonium chloride can be used here in order to isotopically label the protein. Unless noted, N^{14} is used), 4g of dextrose, 70µL of 0.5 M iron(III) chloride, 100 µL of 1M calcium chloride, 1mL of trace metals (contents in g/L: 5 Na_2EDTA; 0.5 $FeCl_3$; 0.05 $ZnCL_2$; 0.01 $CuCl_2$; 0.01 $CoCl_2.6H_2O$; 0.01 H_3BO_3; 1.6 $MnCl_2.6H_2O$), 1mL of 1M magnesium sulfate and 25 µL of 2% thiamine.

The cells were grown in the defined media until they reached an optical density of 1 as measured at 600 nm, and then induced with 1 mL of 1M isopropyl β-D-thiogalactopyranoside (IPTG) solution for protein expression. Also, 35mg/L of delta-amino levulinic acid hydrochloride was added as heme precursor at the time of induction. Cells were harvested 4-6 hours after induction via centrifugation at 5000rpm, weighed and stored at -80°C until further use.

CYP101 Purification: Frozen cells were thawed and resuspended in 50mL of buffer A (2mM camphor, 50mM KPO_4, 50mM KCl) per 10g of cells. The resuspended solution was then cooled and sonicated for three cycles of 20 seconds each with one minute pauses between cycles on a Branson sonifier 250. This is done in order to lyse the cells and extract the protein from the cells. The cell lysate was centrifuged for 15

minutes at 10,000 RPM and 4°C in a SS-34 rotor, and the supernate was removed from the cell debris. The supernate was then run through a Co^{2+} metal affinity (Talon Metal Affinity Resin, Clontech Laboratories) column, which binds CYP101 due to the His_6 tag present on the protein. The bound protein was then washed with 20 times the column volume of buffer A and eluted with buffer A containing 125 mM imidazole solution. The eluted protein solution was then passed next through an anion exchange Q Sepharose fast flow column (GE Healthcare Life Sciences) and eluted with buffer A containing 250 mM KCl. The protein was then concentrated using a Millipore 10,000 MW cutoff centrifugal filter. Finally, the protein underwent another purification step involving size exclusion chromatography on a Amersham FPLC system. Confirmation of purity of individual fractions from the size exclusion chromatography was carried out using UV-Vis spectroscopy using an absorbance ratio of A_{391nm}/A_{280nm} where this absorbance ratio value of >1.4 was deemed to have protein with >95% purity. Concentration of the pure protein fractions was determined by following formula (Eq.2):

$$A_{391} = \varepsilon bc \qquad \text{Eq. 2}$$

A_{391} represents the absorbance value at 391nm, where the heme absorbs in the presence of camphor, ε is the extinction coefficient, which is 100 M^{-1} cm^{-1} for CYP101, b represents the path length of 1cm of the UV vis cuvette and c is the molar concentration of the protein. A characteristic UV-Vis scan (wavelength range 250-450nm) of the purified CYP101 protein bound to camphor is shown in Figure 10. The camphor bound CYP101 protein exhibits two major peaks in the UV vis spectrum in the selected wavelength range, The peak at about 280 nm arises from the aromatic residues present in the protein, while the 391 nm peak represents the absorbance of the heme group in the protein affected by the presence of camphor. This peak will shift depending on which ligand, if any, is bound in vicinity of the heme group.

Ligand Exchange: Since CYP101 is purified in buffer containing camphor, all purified protein samples are in camphor-bound form. In order to prepare CYP101 in a form to another ligand (substrate or inhibitor), the camphor-bound protein underwent a

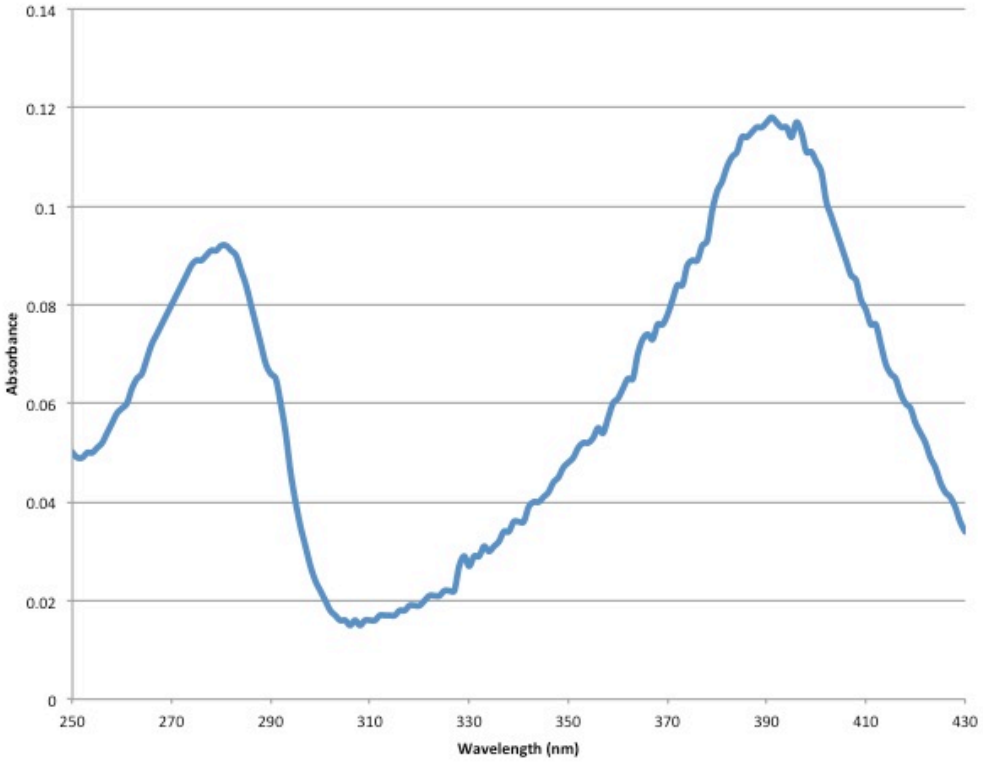

Figure 10: **UV Vis absorbance spectrum of purified camphor-bound CYP101.** The camphor-bound form shows a characteristic UV-Vis peak at 391 nm.

ligand exchange process. ~1mL of concentrated, purified camphor-bound protein was diluted with ~15mL of buffer without camphor and re-concentrated back to ~1mL to remove substrate via dialysis.This process was repeated 5 times to ensure maximum substrate removal, which is indicated by the color change from camphor bound (brown) to substrate-free form (red). Then in the final step, camphor-free buffer A containing either 2mM of nicotine solution or 2 mM of ketoconazole solution was added to the protein to allow for ligand binding. The protein was then re-concentrated to ~0.2 mM, and was used further for UV-Vis characterization and NMR studies.

Preparation of samples for NMR spectroscopy: The final NMR samples used for amide exchange experiments consisted of 0.2 mM of ^{15}N CYP101 in 50 mM potassium phosphate, 50 mM KCl, pH 7.4 buffer. The camphor-bound NMR samples used for amide exchange studies additionally contained 2 mM camphor. Ketoconazole and nicotine bound NMR samples similarly contained 2 mM ketoconazole and 2 mM nicotine respectively for acquisition of ^{1}H-^{15}N correlation spectra.

UV-Vis characterization of substrate-free and ligand-bound forms of CYP101: UV-Vis scans of substrate-free, nicotine-bound and ketoconazole-bound forms of CYP101 were obtained in a fashion similar to the camphor-bound form described above (Figure 11). Binding of the ligand was monitored by small shifts in wavelength of the absorbance peak at 391 nm observed in the camphor-bound form. In the substrate-free form, this peak shifts to 417 nm due to removal of camphor. Upon binding of the nicotine ligand, the wavelength of this peak shifts to approximately 420 nm, indicating the characteristic presence of CYP101 bound to an inhibitor. Upon binding of the inhibitor ketoconazole, however, the peak does not change its position relative to the substrate-free form but just gets broader. This is similar to what has been characterized previously in UV-Vis studies with these inhibitors (42). The substrate-free and various ligand-bound forms of ^{15}N labeled CYP101 were then prepared in a similar manner and used further for NMR studies.

^{15}N selective labeling of CYP101 by amino acid type: In order to facilitate further sequence-specific assignments for CYP101, protein samples selectively labeled with ^{15}N at backbone amides for the amino acid types Ala, Gly, Glu, Asp, Leu and Phe were prepared in the following manner – Cell growths and CYP101 protein expression were carried out exactly as described above except that each ^{15}N labeled amino acid was added along with 19 unlabeled amino acids in the following amounts (g/L) one hour prior to the time of induction: 1.0 S, 0.5 E, 0.4 D, 0.25 A, 0.25 G, 0.25 Q, 0.25 V, 0.25 L, 0.25 M, 0.25 T, 0.2 I, 0.2 N, 0.15 F, 0.15 Y, 0.15 K, 0.15 H, 0.15 R 0.125 C, 0.1 W, 0.0 P. The unlabeled equivalent was excluded for each individual labeled growth. The induction procedure was the same as described above. CYP101 expressing cells were harvested and the selectively labeled protein purified similar to the unlabeled or ^{15}N uniformly labeled protein. All samples were prepared in the camphor-bound form. Each purified selectively labeled protein sample was used further for NMR analysis.

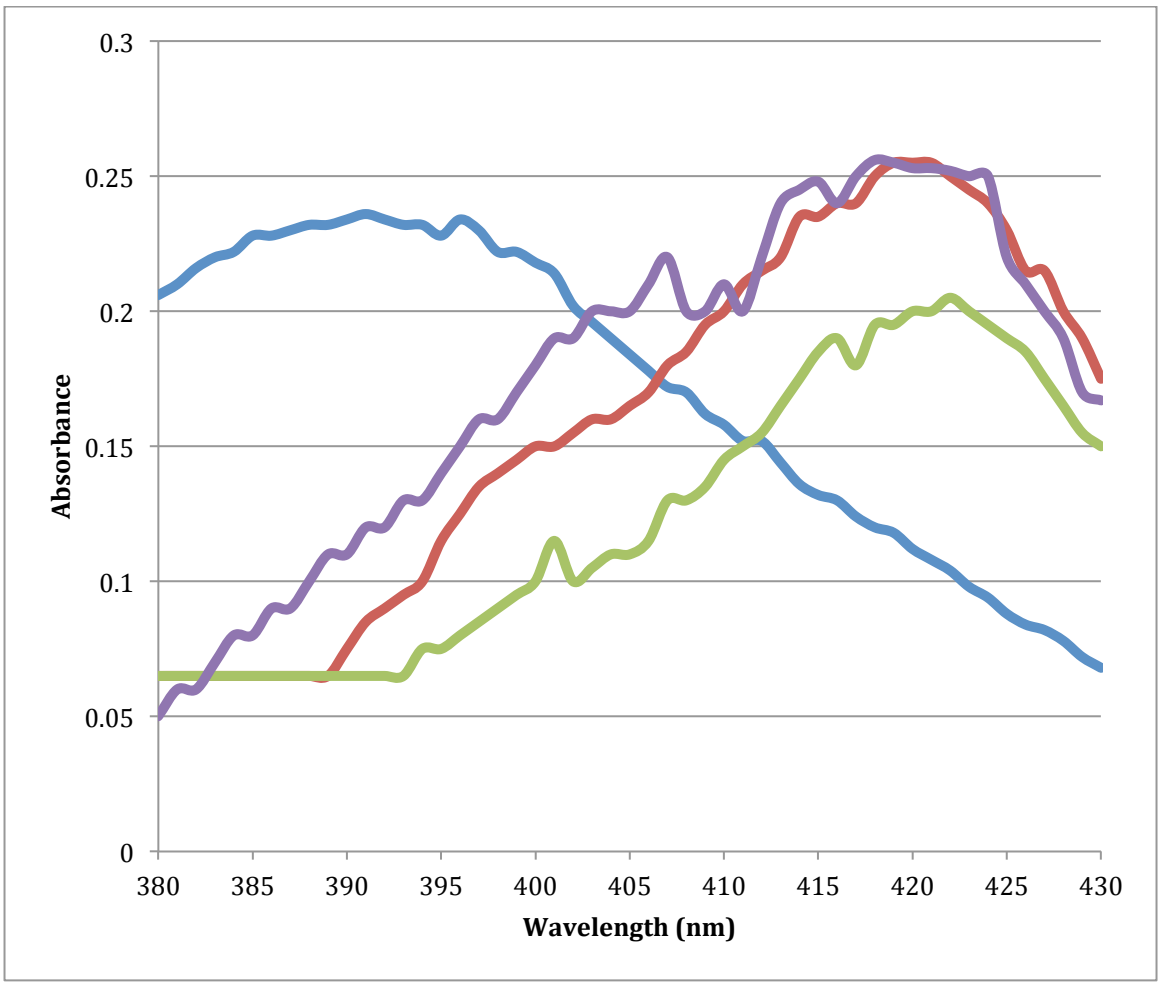

Figure 11: UV-Vis characterization of CYP101 in presence of different ligands. Characteristic UV-Vis peaks in the spectra for substrate-free CYP101 (red) and bound to various ligands [camphor-bound (blue), nicotine-bound (green) and ketoconazole-bound (purple) are shown.

CHAPTER IV
HYDROGEN-DEUTERIUM EXCHANGE NMR STUDIES ON CYP101

4.1 Collection of amide exchange data for CYP101 using NMR spectroscopy

Multiple techniques should be used in order to gain the most complete dynamic picture of an enzymatic system, since different processes occur on different timescales. The distinct motions at different timescales explain how the dynamics of the protein can influence differential binding. Essentially, all dynamic motions are responsible for protein dynamics. The fast bond vibrations build in to the slower motions, like a domino effect. These slower motions that facilitate differential substrate binding are the focus of this study. Hydrogen-deuterium exchange, amide exchange or HDX is a technique used to observe slower timescale dynamic motions, from millisecond-second, which provides an excellent way to observe larger domain motions and structural changes. The technique takes a hydrogenated protein devoid of all water, and replaces the exchangeable protons with deuterons by placing it in deuterated water (D_2O). The exchange of protons with the deuterons occurs in a time dependent manner, according to the accessibility of the exchangeable protons. The surface ones will exchange first, while those more buried will take a certain amount of time (Figure 12).

This exchange process can be monitored by 2D NMR spectroscopy, and the peaks corresponding to each protein residue can be seen to decrease in intensity if they are exchanging with the deuterons. There will come a point, when a peak is completely exchanged, it will disappear from the spectrum, since deuterium is not a spin half nucleus and cannot be detected by traditional NMR methods. Since HDX is a slow timescale technique, much information can be gained from this methodology. The rate of exchange can be calculated from the peak intensity (Equation 3) which can also give the dynamic timescale of the residues being investigated.

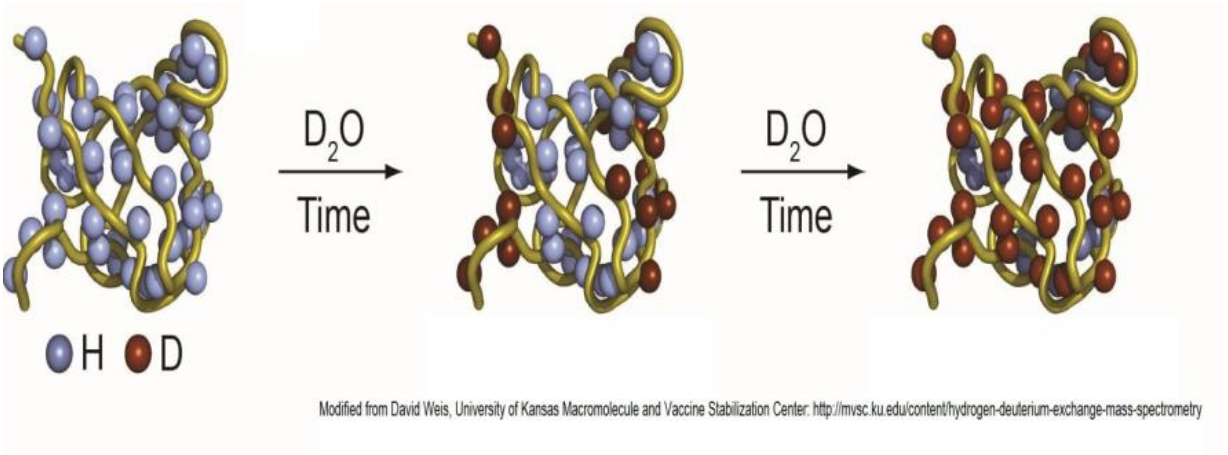

Figure 12: Mechanism of Hydrogen/Deuterium exchange (HDX). Exchangable hydrogens of proteins are exposed to D2O, and over time the deuterons will take the place of the protons. This process can be detected through NMR and Mass spectrometry. Modified from (43).

$$\ln I = \ln I_o - R_{ex}t \qquad \text{Eqn. 3}$$

where I is peak intensity, I_0 is the initial intensity, R_{ex} is the exchange rate and t is time period of the exchange (44). In addition, often the peak environment can be determined from this technique. For example, if the hydrogen being exchanged is a part of a hydrogen bond, the chemical shifts will be different than those hydrogens not involved in any interactions. The data can be analyzed from that perspective, and used to identify the presence of hydrogen bonds.

Amide exchange studies were carried out on CYP101 to compare the exchange rates between substrate-free and camphor-bound forms of CYP101 and characterize the regions experiencing slower timescale motions in CYP101. In order to carry out the exchange studies, CYP101 NMR samples were lyophilized overnight. Then, samples were hydrated with D_2O and immediately placed in a Varian 600 MHz spectrometer equipped with cold probe. Data collection started within the first 5 minutes of the D2O addition to allow for detection of fast exchanging resonances. ^{1}H-^{15}N 2D correlation spectra were collected with a traditional TROSY pulse sequence at 35° C. Each experiment consisted of 8 scans acquired with 1024 complex points in the direct dimension and 64 increments in the indirect dimension. The total acquisition time for each spectrum was approximately 28 minutes. A series of such spectra for substrate-free and camphor-bound forms of CYP101 were collected over a period of 39 hrs. The spectral acquisition was stopped at this time to prevent spurious data collection since CYP101 tends to become unstable and denature at 35^0 C after approximately 40 hours based on prior experience with other NMR experiments on CYP101. The resulting spectral data was processed using NMRPipe and SPARKY software (45, 46) to extract peak intensities from each dataset. Only simple exponential window functions were used during data processing with no baseline corrections. Peak intensities were measured only for peaks for which assignments were available and fitted to Equation 3 as a function of time to obtain amide exchange rates (Rex) for corresponding peaks

(Figures 13 and 14). Available amide exchange rates for both substrate-free and camphor-bound forms are listed in Table 1.

4.2 Comparison of amide exchange rates between substrate-free and camphor-bound forms of CYP101

Analysis of amide exchange data for substrate-free and camphor-bound forms reveals that only about 40% of resonances remain in the spectrum for each form after collection of the first spectrum corresponding to about 33 minutes of exposure to D_2O (Figure 15). The camphor-bound form retains a few more peaks in the spectrum compared to the substrate-free form, but not by much. This indicates that the surface-exposed residues exchange fast with the solvent in a similar fashion for both forms, which is expected from the overall similar structures for both forms of the protein. The differences in the exchange preference for a few residues more in the substrate-free form may reflect the exposure of more residues to the solvent due to the relatively open conformation sampled by the substrate-free form compared to the closed camphor-bound form. It is also observed that the camphor bound form has nicely defined peaks while the substrate free looks a little bit more disordered due to presence of additional line broadening relative to the camphor-bound form (Figure 13). While the faster exchange processes are similar for both forms, considerable differences are observed for the slower exchanging peaks. Over the various time points of data collection following the first time point, the peak intensities for both forms continue to decrease steadily in an exponential manner until after approximately 16 hrs, only about 20% of the peaks initially observed for camphor-bound form remain protected from the solvent, while less than 10% remain for the substrate-free form. These remaining peaks do not lose much intensity for the remainder of exchange time in the experiment. On this basis, we can classify the exchange process into 3 categories – 1) fast exchange (less than 33 min), 2) intermediate exchange (33 min to 16 hrs) and 3) slow exchange (16 hrs to 39 hrs). The differences in exchange rates for the peaks in these 3 categories will be discussed after the next section.

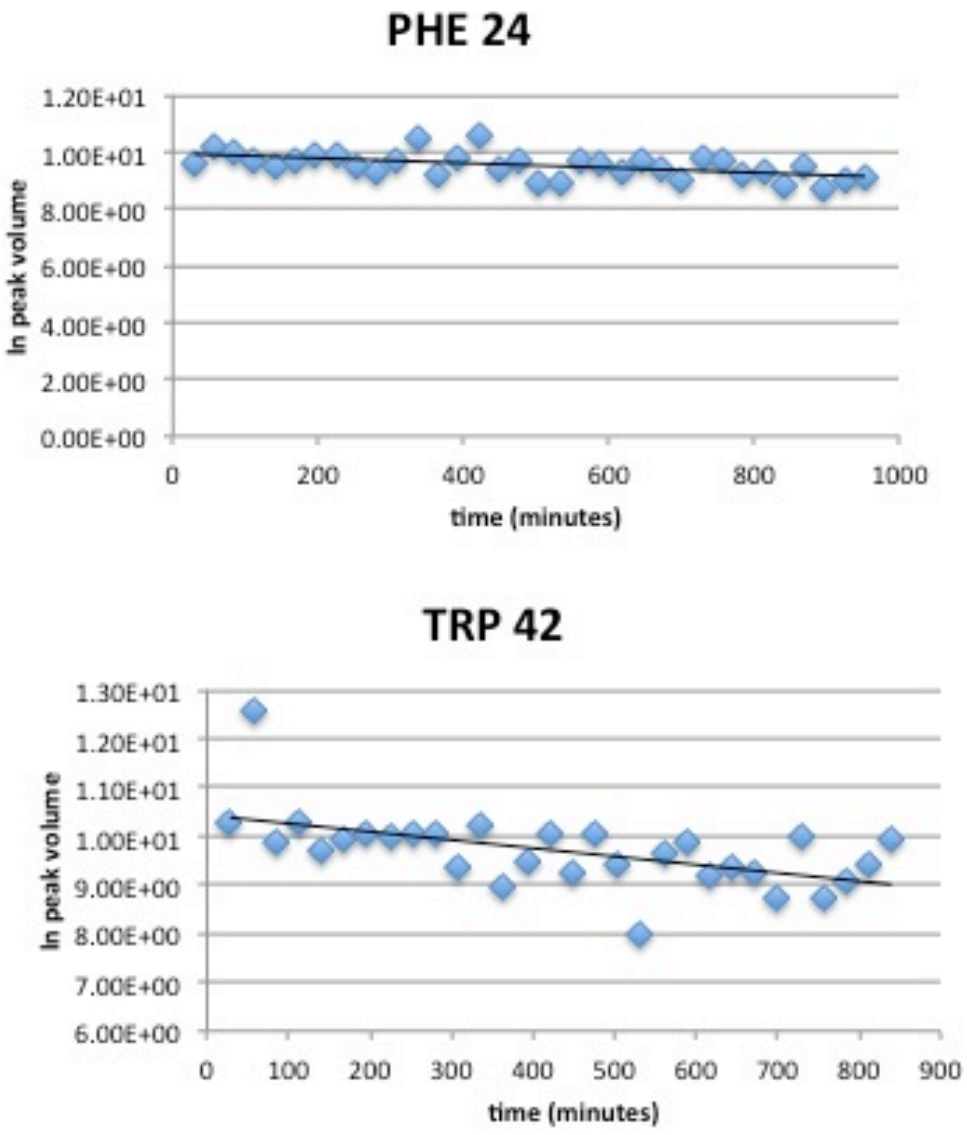

Figure 13: Logarithmic plots of peak intensities vs time for two example residues in camphor-bound CYP101.

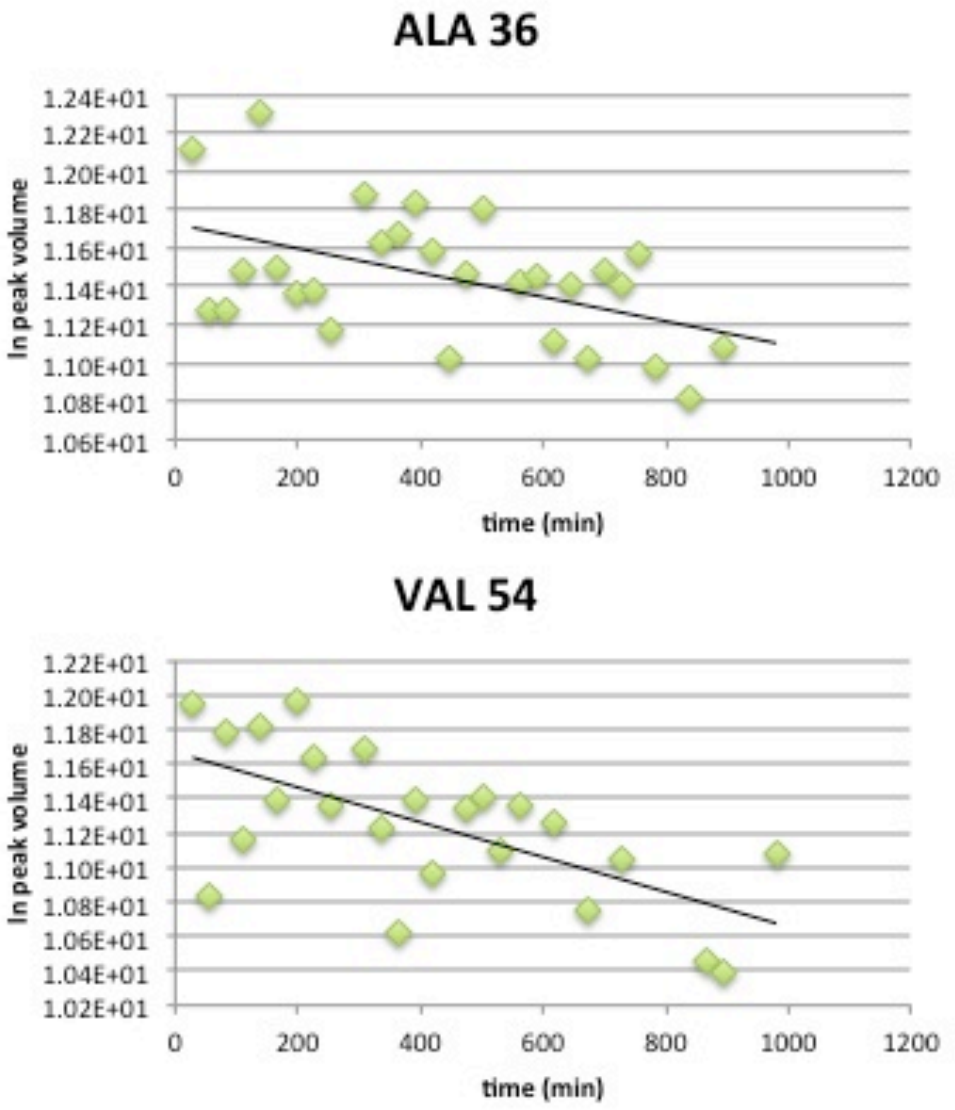

Figure 14: Logarithmic Plots of peak intensities vs. time for two example residues in substrate-free cyp101.

Table 1: H/D exchange rates measured by NMR for camphor-bound and substrate-free CYP101. Bold secondary structures are indicative of substrate binding regions. A (-) indicates no rates were measured for that residue. Error values were calculated from the fitting error to the logarithmic plots.

Camphor Bound			Substrate Free		
Residue	$K(s^{-1})$	Error	$K(s^{-1})$	Error	Secondary Structure
12-NH	Fast		fast		
18-NH	5.5×10^{-5}	3.3×10^{-6}	6.9×10^{-5}	1.3×10^{-6}	turn 1
24-NH	2.3×10^{-7}	0.8×10^{-7}	4.3×10^{-6}	2.8×10^{-7}	
36-NH	--		5.5×10^{-6}	1.2×10^{-6}	turn 3
37-NH	Fast		fast		turn 3
40-NH	8.7×10^{-6}	2.8×10^{-6}	fast		
42-NH	1.1×10^{-6}	4.6×10^{-7}	--		A Helix
47-NH	2.8×10^{-6}	1.1×10^{-6}	--		turn 4
50-NH	3.4×10^{-5}	1.4×10^{-6}	5.8×10^{-5}	3.4×10^{-6}	turn 4
52-NH	4.6×10^{-6}	1.7×10^{-6}	--		
54-NH	4.6×10^{-6}	1.7×10^{-6}	6.2×10^{-6}	2.4×10^{-6}	
55-NH	3.8×10^{-6}	1.2×10^{-6}	3.4×10^{-6}	1.4×10^{-6}	β1
59-NH	4.1×10^{-6}	1.6×10^{-6}	2×10^{-7}	4.5×10^{-6}	turn 5
60-NH	Fast		fast		turn 5
62-NH	2.8×10^{-7}	0.7×10^{-7}	--		
63-NH	--		7.8×10^{-5}	2.5×10^{-6}	β1
64-NH	4.3×10^{-6}	2.5×10^{-6}	5.9×10^{-6}	1.6×10^{-6}	β1
65-NH	6.7×10^{-7}	2.5×10^{-7}	--		β1
67-NH	5.3×10^{-7}	1.2×10^{-7}	6.2×10^{-6}	2.5×10^{-6}	B Helix
68-NH	Fast		fast		B Helix
70-NH	Fast		fast		B Helix
80-NH	6.4×10^{-7}	1.8×10^{-7}	5.1×10^{-5}	4.5×10^{-6}	
81-NH	4.8×10^{-7}	1.3×10^{-7}	2.5×10^{-6}	6.5×10^{-7}	
82-NH	5.6×10^{-6}	5.5×10^{-7}	fast		
83-NH	Fast		fast		**B' helix**

Continuation Table 1

Camphor Bound			Substrate Free		
Residue	$K(s^{-1})$	Error	$K(s^{-1})$	Error	Secondary Structure
87-NH	Fast		Fast		**B' helix**
88-NH	Fast		Fast		
90-NH	Fast		Fast		
91-NH	Fast		Fast		**B' helix**
92-NH	Fast		Fast		**B' helix**
93-NH	Fast		Fast		**B' helix**
94-NH	Fast		Fast		**B' helix**
95-NH	Fast		Fast		**B' helix**
96-NH	Fast		Fast		**B' helix**
98-NH	Fast		Fast		**B' helix**
101-NH	Fast		Fast		**B' helix**
102-NH	Fast		Fast		**B' helix**
104-NH	Fast		Fast		**B' helix**
107-NH	Fast		Fast		C helix
119-NH	3.3×10^{-5}	1.8×10^{-6}	Fast		C helix
123-NH	Fast		7.7×10^{-5}	2.2×10^{-6}	C helix
124-NH	Fast		Fast		C Helix
129-NH	Fast		2.3×10^{-7}	1.4×10^{-7}	
145-NH	--		--		β1
146-NH	6.7×10^{-5}	4.9×10^{-6}	7.3×10^{-5}	3.5×10^{-6}	β1
147-NH	--		0.9×10^{-7}	5.5×10^{-8}	β1
148-NH	3.2×10^{-7}	0.6×10^{-7}	1.5×10^{-7}	0.5×10^{-7}	β1
149-NH	Fast		6.2×10^{-7}	1.8×10^{-7}	β1
150-NH	6.7×10^{-6}	5.6×10^{-7}	6.4×10^{-6}	1.9×10^{-6}	β1

Continuation Table 1

Camphor Bound			Substrate Free		
Residue	$K(s^{-1})$	Error	$K(s^{-1})$	Error	Secondary Structure
151-NH	Fast		fast		E Helix
152-NH	Fast		5.6×10^{-6}	2.4×10^{-6}	E Helix
154-NH	Fast		fast		E Helix
161-NH	Fast		1.2×10^{-6}	5.6×10^{-7}	
164-NH	Fast		fast		
168-NH	8.9×10^{-5}	6.5×10^{-6}	8.1×10^{-5}	6.5×10^{-6}	
169-NH	4.8×10^{-5}	5.5×10^{-6}	--		
171-NH	Fast		fast		turn 8
172-NH	Fast		fast		
185-NH	6.4×10^{-6}	2.4×10^{-6}	5.2×10^{-6}	2.5×10^{-6}	**F Helix**
186-NH	3.8×10^{-7}	1.2×10^{-7}	4.6×10^{-6}	2.7×10^{-6}	
188-NH	Fast		fast		**FG Loop**
189-NH	Fast		fast		**FG Loop**
190-NH	Fast		fast		**FG Loop**
191-NH	--		--		**FG Loop**
192-NH	--		fast		**FG Loop**
193-NH	Fast		fast		**G Helix**
207-NH	4.6×10^{-6}	7.7×10^{-7}	--		G Helix
216-NH	Fast		fast		H Helix
217-NH	1.9×10^{-7}	0.4×10^{-7}	7.8×10^{-7}	2.2×10^{-7}	H Helix
226-NH	Fast		fast		H Helix
230-NH	Fast		fast		β2
243-NH	5.6×10^{-5}	6.5×10^{-6}	fast		**I Helix**
245-NH	6.3×10^{-6}	2.4×10^{-6}	--		**I Helix**
279-NH	Fast		fast		

Continuation Table 1

Camphor Bound			Substrate Free		
Residue	$K(s^{-1})$	Error	$K(s^{-1})$	Error	Secondary Structure
297-NH	--		--		**β3 Sheet**
305-NH	2.2×10^{-7}	7.6×10^{-8}	5.1×10^{-6}	1.4×10^{-6}	
306-NH	fast		fast		β4
307-NH	4.5×10^{-7}	6.5×10^{-8}	3.8×10^{-6}	7.6×10^{-7}	turn 11
308-NH	1.5×10^{-7}	2.2×10^{-8}	6.6×10^{-6}	2.4×10^{-6}	turn 11
309-NH	fast		9.1×10^{-5}	1.5×10^{-6}	turn 11
312-NH	1.8×10^{-7}	3.8×10^{-8}	7.3×10^{-6}	2.3×10^{-6}	β4
325-NH	fast		fast		
326-NH	3.8×10^{-6}	1.6×10^{-6}	fast		
329-NH	fast		fast		turn 12
330-NH	fast		fast		turn 12
334-NH	fast		fast		turn 13
338-NH	fast		fast		
339-NH	--		8.6×10^{-5}	7.6 x 10-^{6}	
340-NH	fast		fast		
376-NH	fast		fast		
384-NH	2×10^{-5}	4.2×10^{-6}	9.2×10^{-5}	6.5×10^{-6}	turn 16
389-NH	fast		fast		β5
390-NH	fast		fast		β5
391-NH	fast		6.8×10^{-6}	2.2×10^{-6}	β5
392-NH	3.7×10^{-7}	1.1×10^{-7}	fast		**β5 Sheet**
393-NH	fast		fast		**β5 Sheet**
394-NH	3.3×10^{-7}	0.5×10^{-7}	--		**β5 Sheet**
398-NH	4.4×10^{-7}	1.5×10^{-7}	fast		**β5 Sheet**
400-NH	fast		fast		
410-NH	fast		fast		
414-NH	fast		fast		

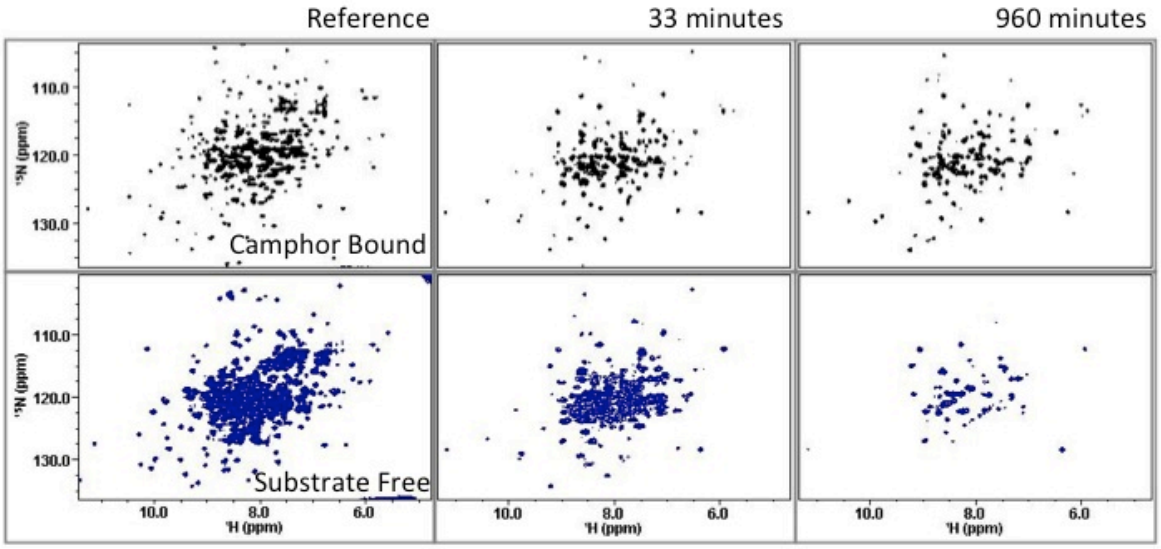

Figure 15: Comparison of HDX NMR data for camphor-bound and Substrate-free CYP101. Decrease in peak intensities of camphor bound and substrate free HDX CYP101 in 2D ^{1}H-^{15}N HSQC-TROSY NMR spectra collected at 35°C is shown over a period of time marked over each spectrum. The first panel shows a non-HDX TROSY NMR spectrum for ^{15}N CYP101 in the same conditions as the HDX spectra and was used as reference to identify residues undergoing exchange.

4.3 Additional assignments from ^{15}N selective labeling of CYP101

CYP101 has a total of 414 amino acids in its sequence, however assignments are available only for 107 residues in the oxidized form of the protein. Considerably more number of assignments are available for the reduced form of the protein (~290 residues). There are several reasons for this. The main reason is that oxidized CYP101 with its Fe^{+3} oxidation state of the heme iron is paramagnetic around room temperature, while reduced CYP101 with Fe^{+2} oxidation state of the heme iron is diamagnetic. This entails paramagnetic broadening of resonances up to a distance of 8 Å from the heme center in the oxidized form, which causes loss of resonances in the oxidized CYP101 TROSY spectrum and makes assignment of resonances even more difficult due to loss of sequential connectivities.

The traditional experiments used for protein backbone assignments are the three-dimensional multinuclear experiments such as HNCA, HNCO and HNCACB. These experiments rely on proteins that have been isotopically labeled with ^{15}N, ^{13}C and ^{2}H for detection. HNCO allows direct correlation of the N-H chemical shifts with the backbone carbonyl C=O chemical shifts of the previous residue. HNCA experiments not only detect the transfer of magnetization from the N-H bond of the residue of interest (i), to its own $C\alpha$,(i), but also the Cα of the previous residue (i-1). This allows sequential connectivity to be built between the residues, and allows walk-through of assignments through the protein sequence. However, HNCA and HNCO experiments by themselves are insufficient to provide unambiguous sequential assignments. HNCACB is also used to resolve ambiguities and obtain a more complete set of assignments, including partial side-chain assignments. This experiment correlates the Cα and Cβ chemical shifts of residues i and i-1, with the N-H (i) chemical shift. Subsequently, the chemical shift analysis of the residues allows them to be matched up and the connectivities formed between various residues. The chemical shifts between oxidized and reduced forms are distinct enough to prevent direct transfer of most assignments from reduced to oxidized CYP101. In addition, oxidized CYP101 resonances tend to be generally broader than

the reduced form and thus magnetization transfer does not occur very efficiently in 3D NMR experiments used for assignment. Thus, traditional 3D NMR experiments have been less effective in obtaining assignments for more than 107 in oxidized CYP101. During the assignment process using 3D NMR experiments such as HNCA and HNCACB, there are several small stretches of sequential connectivities observed in oxidized CYP101, but definite assignments cannot be made due to ambiguity. A way to overcome some of this ambiguity will be to identify the amino acids forming sequential connectivities by amino acid type and hope to find unique pairs in the amino acid sequence connecting two of the amino acid types, which will allow us to place them sequence-specifically in combination with other sequential connectivities and thus assign them. Selective labeling of amino acids within the protein is a simple way to accomplish the marking of resonances by amino acid type and therefore selective labeling of several amino acid types was undertaken in CYP101.

NMR samples of CYP101 selectively labeled with ^{15}N at the backbone amides were prepared for the following amino acid types – Ala, Gly, Leu, Phe, Asp and Glu using methods described in Chapter 3. The specific amino acid types were selected as a preliminary set for labeling based on the following criteria - their higher prevalence in the amino acid sequence of CYP101, their likely involvement in stretches of sequential connectivities in 3D experiments, easy identification due to their somewhat unique chemical shifts and their propensity to scramble the ^{15}N label to other amino acids (47). ^{1}H-^{15}N TROSY spectra for these selectively labeled samples were collected and matched up with a uniformly labeled spectrum to identify assigned and unassigned peaks in the uniformly labeled spectrum by amino acid type (Figures 16 and 17). The unassigned peaks now identified by amino acid type are now being reassessed in the 3D NMR data in the hopes that additional assignments can be made available for these residues in oxidized CYP101 in the short term.

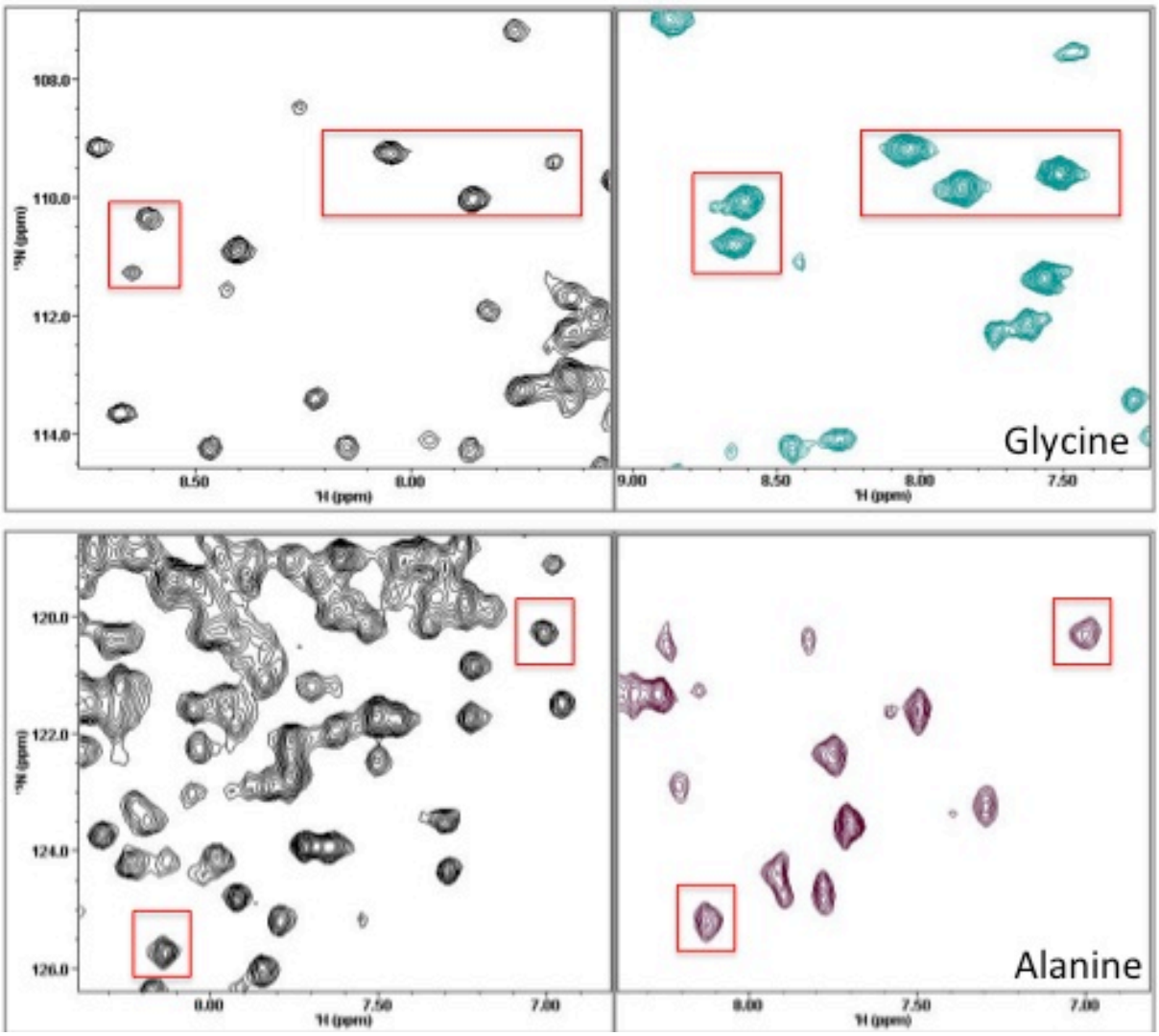

Figure 16: NMR spectra of CYP101 selectively labeled with ^{15}N Alanine and Glycine. Portions of NMR spectra of selectively labeled (burgundy for alanine, teal for glycine) and uniformly labeled (black) CYP101 are shown for comparison. Boxes in red highlight the corresponding peaks identified by amino acid type.

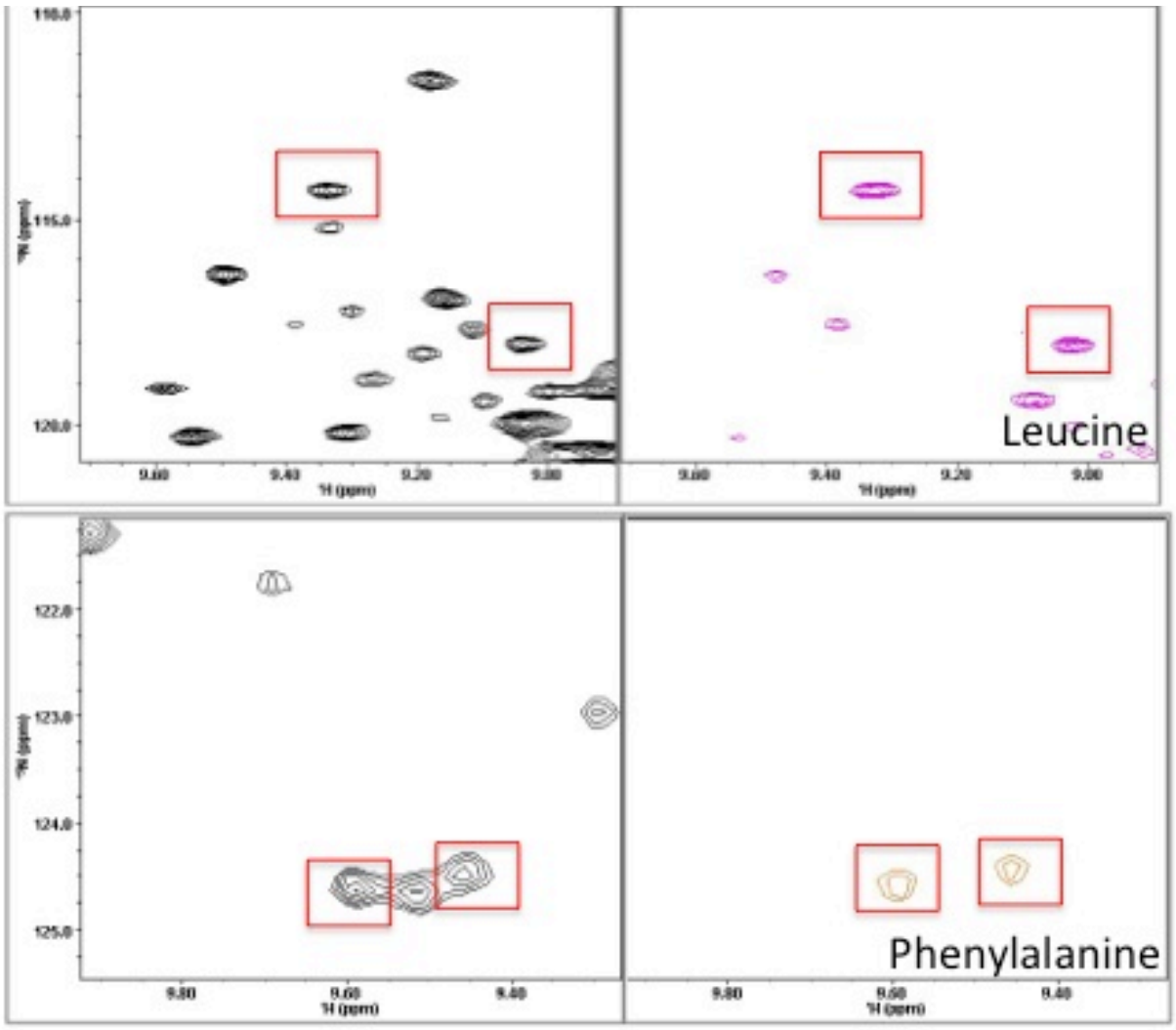

Figure 17: NMR spectra of CYP101 selectively labeled with ^{15}N Phenylalanine and Leucine. Portions of NMR spectra of selectively labeled (magenta for leuine, gold for phenylalanine) and uniformly labeled (black) CYP101 are shown for comparison. Boxes in red highlight the corresponding peaks identified by amino acid type.

The selective labeling of the preliminary set of amino acids was reasonably successful. For amino acids such as Phe and Ala, there was very minor scrambling of the ^{15}N label to other amino acids and thus the expected number of resonances (excluding the ones not observed due to paramagnetic broadening) were observed for these amino acids. Minor scrambling was observed for the Gly label to other amino acids such as Ser and Cys. On the other hand, major scrambling was noticed in the spectra for the Leu and Glu labels. Due to the inter-conversion of these amino acids with other amino acids as part of the metabolic pathways, some of the labeled nitrogens were incorporated in other amino acids, such as Ile and Val for Leu, while Glu scrambled to Gln and other unknown amino acids. The Asp labeled protein did not express in sufficient quantity to obtain a good quality NMR spectrum and ascertain its scrambling. Although scrambling posed a problem for unambiguous identification of amino acid type for several of the resonances in Leu and Glu selectively labeled samples as the number of peaks in the spectra did not correspond to the number of amino acids in the sequence, the peak intensities of the original label were generally higher and helped identify most of the possible peaks for that amino acid type.

The use of this methodology of using the higher intensity peaks for amino acid type identification is however not foolproof and the identities of these peaks need to be confirmed in subsequent analysis of 3D NMR data. However, the good matching of some of the higher intensity peaks with the previously assigned resonances of that amino acid type lends credence to this method and these peaks may serve as a good starting point for further 3D data analysis. Additional selective labeling of other amino acids will be conducted in future if these labels are unable to resolve ambiguities and provide more assignments for oxidized CYP101.

4.4 Differences in exchange rates between substrate-free and camphor-bound CYP101

Local dynamics in CYP101 were monitored using amide proton exchange rates for both substrate-free and camphor-bound forms. Overall, the H/D exchange in the substrate-free form was much faster than the camphor-bound form, almost by an order of magnitude in the majority of residues that could be analyzed. Significant differences are observed at the beginning of the B' helix, portions of the C helix as well as the I helix and the β5 sheet (Figures 18 and 19).

The B' helix and β5 sheet form part of the SRS, while the I helix is a key region in the active site. A large difference in exchange rates was seen for residue 243, which is on the N-terminal side of the I helix "kink", a region that is distorted and is known to change its conformation to provide a binding pocket for oxygen molecule in the reduced form. Residue 243 is still visible in the first few camphor-bound spectra acquired after D_2O addition, however is exchanged out completely even in the first acquired substrate-free spectrum. We have not been able to monitor exchange rates for other residues in the I helix as most of this helix is close to the paramagnetic heme center and its resonances are likely broadened out in oxidized CYP101 spectra. However, given the fact that this particular helix is known to respond to binding of substrates in almost all P450 structures by changing its conformation using the I helix "kink" to fill up the active site with the side-chains of several residues depending on the size and type of substrate, it is not surprising that we see a larger change in dynamics for this region between the substrate-free and camphor bound forms. It is likely that in the substrate-free form, the solvent has better access to portions of the helix, both due to lack of substrate and also due to increased dynamics of the substrate-free form. Our results indicate that the active site region may be in general more dynamic in the substrate-free form.

Another significant region that shows differential exchange rates in CYP101 is the loop between the B and B' helix, where residues 80-85 show considerably faster exchange rates in the substrate-free form compared to the camphor-bound form. This region is at the beginning of the B' helix, which is an important component of the SRS

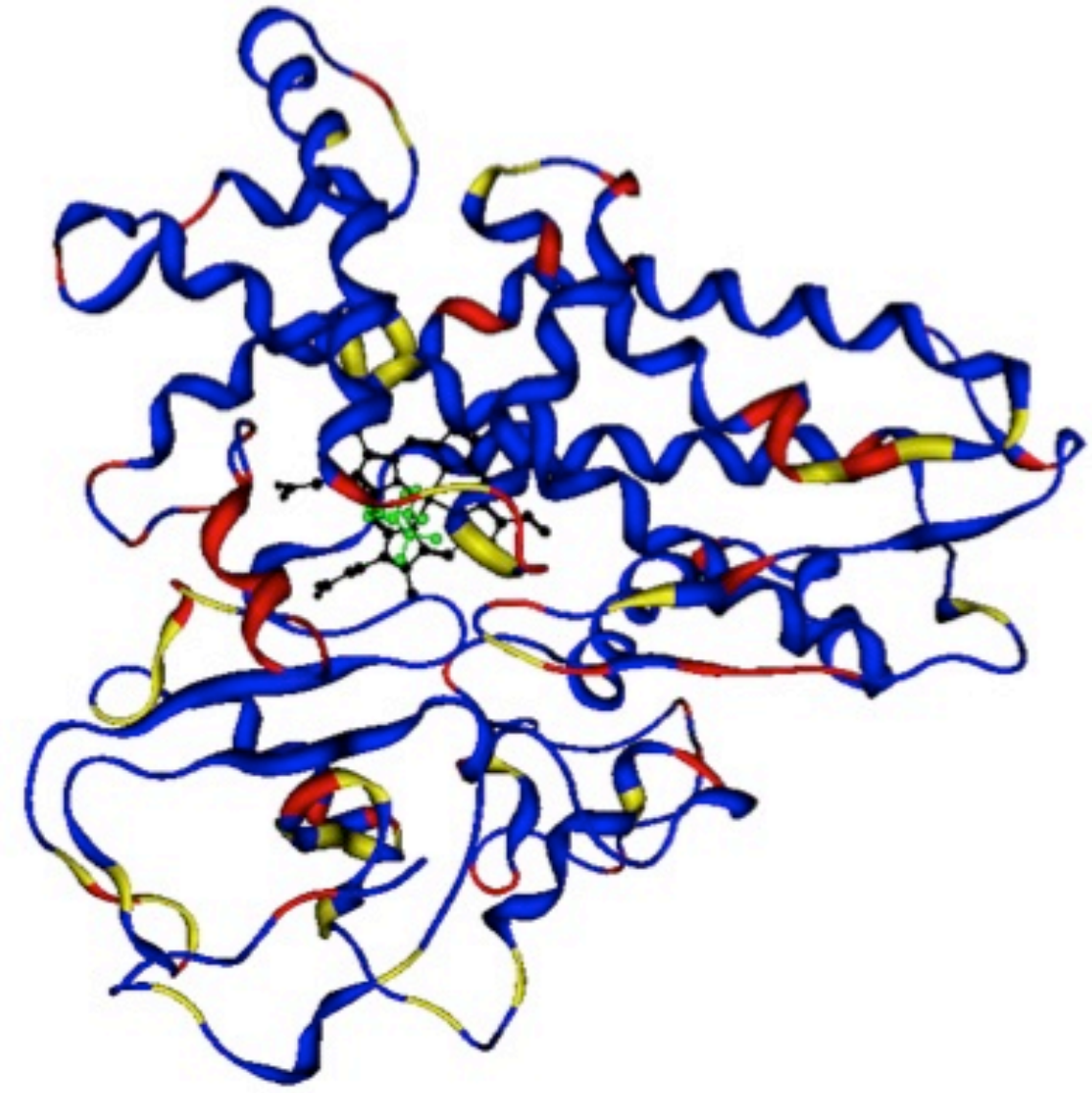

Figure 18: Exchange rates for various residues in camphor-bound CYP101 mapped onto the structure of CYP101. Residues with fast exchange rates (< 33 min) are shown in red, and intermediate and slow exchange rates (> 33 min) in yellow. Residues for which exchange rates could not be measured are colored blue.

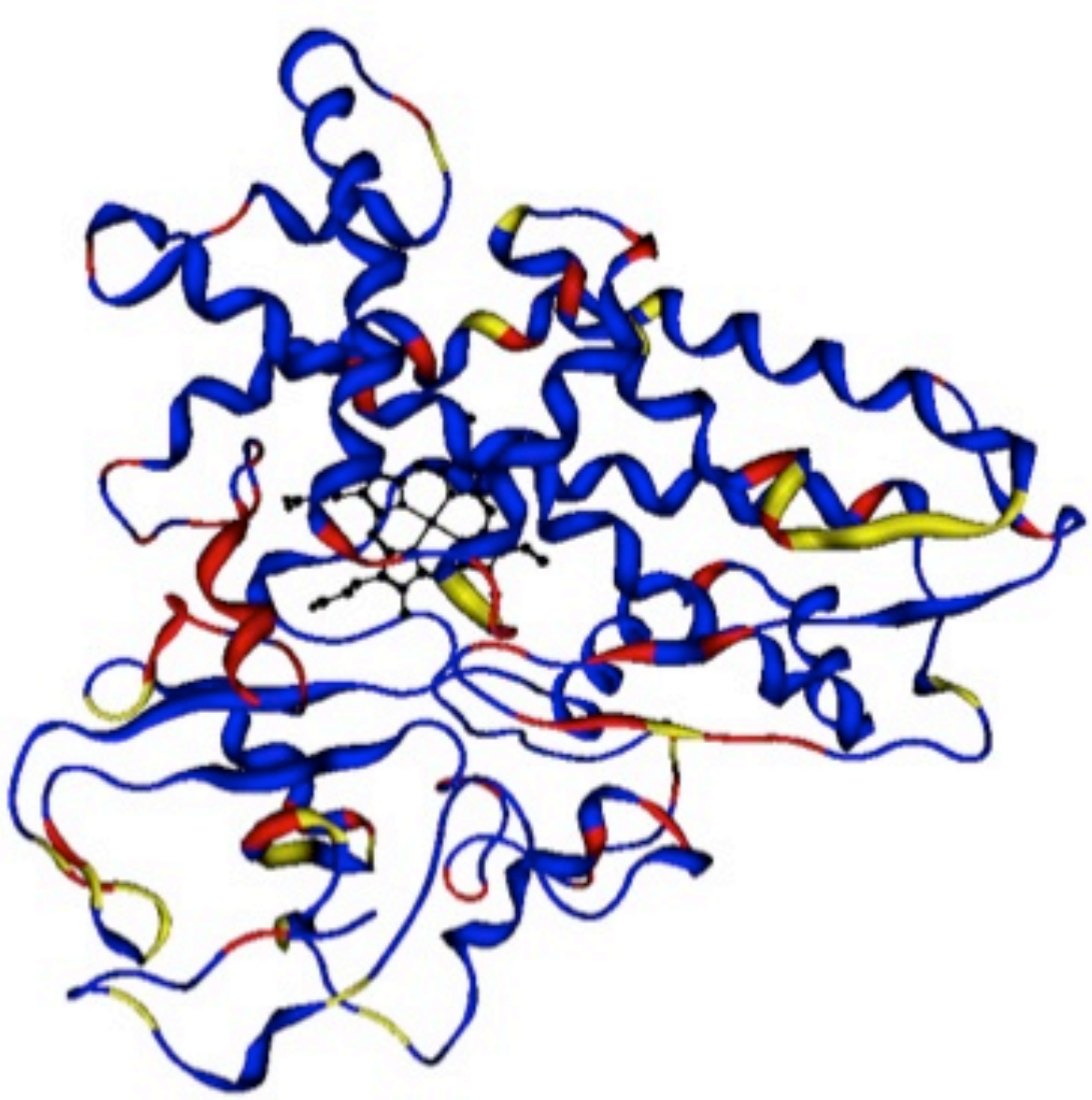

Figure 19: Exchange rates for various residues in substrate-free CYP101 mapped onto the structure of CYP101. Residues with fast exchange rates (< 33 min) are shown in red, and intermediate and slow exchange rates (> 33 min) in yellow. Residues for which exchange rates could not be measured are colored blue.

and undergoes a large conformational change between the substrate-free and camphor-bound forms as evidenced from the crystal structures. The B' helix is largely disordered in the substrate-free form and is likely sampling multiple conformations, which upon camphor binding becomes ordered due to formation of H-bond between Tyr96 side-chain and camphor. The B' helix exhibits fast exchange in both camphor-bound and substrate-free forms, most likely due to its surface location on CYP101, so it is not possible to figure out whether there are significant differences in the exchange rates for this helix in our experiments. However, the region leading upto the B' helix (residues 80-85) is more protected from the solvent and thus made it possible to measure the exchange rates in both forms. The much faster exchange rates in the substrate-free form for this region strongly suggests the more dynamic nature of this region and by inference also of the B' helix.

Exchange rates for the β5 sheet, which is also part of the SRS, similarly are different in the substrate-free and camphor-bound forms. The β5 sheet region is part of the conformational change that leads to the closed conformation from the open conformation upon binding of camphor, trapping the substrate inside the active site. Residues 392, 394 and 398 are part of the short β-turn that connects the two strands of the β5 sheet exchange very slowly in the camphor-bound form and increase their exchange rates by almost two orders of magnitude in the substrate-free form. This region makes several H-bonding, van der Waals and electrostatic contacts to both the C-terminal side of the I helix as well as the N-terminal region in the camphor-bound form. Both the N-terminal region of CYP101 and I helix are known to be fairly flexible regions. It is therefore possible that movement of the I helix region is coordinated with the movement of the β5 sheet to effect binding of camphor. In the substrate-free form, since there is no camphor present, the I helix is more dynamic as evidenced by the faster exchange rates for residue 243 and this motion disrupts some of its contacts with the β5 sheet, leading to the increased dynamics observed for β5 sheet in the substrate-free form as part of the open conformation. Increased B-factors have been observed for most regions in the SRS of substrate-free form (19), including the B' helix and β5 sheet

supporting the observation of increased dynamics in these regions in our experiments as well.

Other regions that show increased exchange rates in substrate-free form include residues in the C helix region and the β4 turn. The C helix is part of the proximal binding site for Pdx and mediates changes from the proximal binding site to the distal binding site. It is known from previous NMR studies with the Pdx-CYP101 complex that binding of Pdx on the proximal site causes large chemical shift perturbations and dynamic changes on the distal site, especially in the SRS regions. It was postulated that Pdx serves as an effector by changing the subset of conformations of the substrate access channel to prevent loss of substrate to unproductive monooxygenation. The involvement of C helix in this mechanism is critical as it likely transmits any dynamic changes to the distal site by using the modulation of its own dynamics. Thus it is sensitive to any substrate-related dynamic changes and is therefore not surprising that it shows a significant dynamic difference between the substrate-free and camphor-bound forms of CYP101. The β4 turn region acts like a hinge that connects parts of the proximal site region, namely the Cys binding loop and L helix. Thus, dynamic changes in the β4 turn residues are likely to mirror other regions in the proximal site.

There are also some local regions that follow a reverse trend of the camphor-bound form exchanging faster than the substrate-free form. For example, residues in the β1 sheet (148-150) and E helix (152 and 161) show much faster exchange rates in the camphor-bound form than the substrate form. Increases in exchange rates upon binding of camphor are also seen for residues 54 and 59 in the β1 turn region. However, the increase in exchange rates are not as high in magnitude compared to the normal trend of faster exchange rates in substrate-free form. It is not clear what the purpose of these increased exchange rates in the camphor-bound form is, other than they may just represent localized increase in fluctuations as a response to picking up the “slack” for decrease in dynamics elsewhere in the protein.

It is interesting to note that the regions showing the faster exchange rates in the substrate-free form versus the camphor-bound forms are similar to the regions showing the largest redox-dependent dynamic differences in oxidized versus reduced CYP101 (44). Parts of the B-B' helix loop show a decrease in dynamics upon reduction, just as it shows a decease in dynamics upon substrate binding. Regions including the β5 sheet also show similar trends between both forms. This raises the possibility that there are canonical regions in CYP101 that have inbuilt dynamic change capability. This may allow these regions to sample multiple conformations and allow selection of different subsets of conformation as a function of oxidation state and ligand binding. This notion can be tested via design of mutants targeting these regions, which would affect both binding and redox-dependent effects that can be monitored via NMR spectroscopy experiments sensitive to local dynamic fluctuations.

CHAPTER V
CONCLUSIONS AND FUTURE DIRECTIONS

The dynamics of a protein, particularly an enzyme, are an important determinant in its functionality. An enzyme uses its dynamic character, often as a means to facilitate promiscuity in binding substrates and control catalysis rates. A likely mechanism by which this promiscuity is accomplished is that of conformational selection or equilibrium dynamics. Cytochrome P450s are widely believed to employ this mechanism, where they use their conformational flexibility to modulate protein-substrate interactions to accommodate substrates of differing physic-chemical properties and product specificity. An in-depth study of conformational dynamics of P450s is therefore critical to fully understand the basis of substrate promiscuity in these enzymes, which will allow better design of drugs targeted against them in humans. In this work, the millisecond-second dynamic changes that occur upon binding of a substrate to a cytochrome P450 have been described. The model system of cytochrome P450cam or CYP101 was chosen for this purpose, as preliminary NMR studies showed that it exhibits differential dynamics throughout the protein upon binding of distinct ligands.

Hydrogen/deuterium exchange studies using NMR spectroscopy were carried out on CYP101 in presence and absence of a substrate to provide a residue-level picture of slow dynamic changes taking place in this protein upon binding of a substrate. Although previous NMR studies involving H/D exchange have looked at redox-dependent changes in CYP101 dynamics, this is the first study that looks at the effect of substrate binding on the slow timescale dynamics of this protein. Overall, it is found from this study that the substrate-free form of CYP101 exhibits more rapid amide proton exchange than the corresponding residues in camphor-bound form of CYP101. Specific residues in the substrate recognition sites and active site are found in particular to exhibit large differences in exchange rates between the two forms, pointing to the high conformational plasticity available to these regions in this protein. The dynamics of regions such as the β5 sheet and B' helix that are part of the substrate recognition sites (Figure 20) are sensitive to the presence of a substrate and motional differences on the

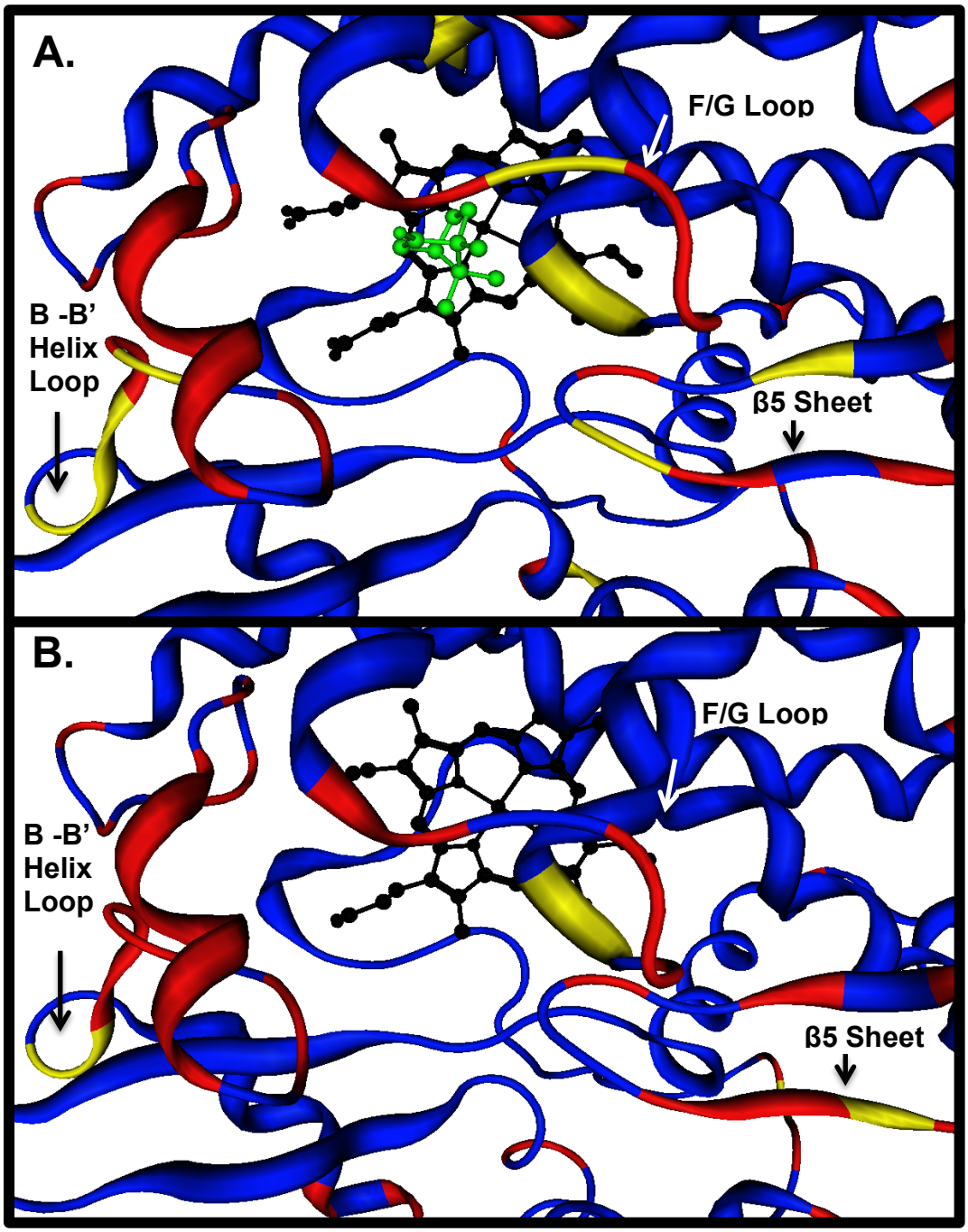

Figure 20: Comparison of exchange rates for selected regions in camphor-bound and substrate-free CYP101. Regions showing the largest differences between the two forms (Panel A: camphor-bound and Panel B: substrate-free) are highlighted by arrows.

millisecond timescale of several residues in this region suggest that they may be not only part of a gating mechanism allowing access to the active site, but also play an important role in the protein adapting to the differing requirements of substrate binding specificity.

It would be interesting to follow up on the studies here on the slower timescales with ones looking at motions on faster timescales in CYP101. Again, NMR spectroscopy is going to be a valuable technique in this regard due to its ability to characterize us-ms motions using CPMG relaxation dispersion experiments and on the ps-ns timescale using T_1/T_2 relaxation rate measurements. If similar regions are involved in modulating dynamic changes on the faster timescales, that would provide a mechanistic pathway for the origin of the slower motions and explain the high conformational plasticity of these regions. Such studies on the faster timescales are currently being undertaken for CYP101 in complex with various ligands. Our preliminary NMR studies with ligands such as nicotine and ketoconazole indicated even more dramatic dynamic differences in the proteins than observed in the presence of camphor. An entropy-enthalpy compensation mechanism was proposed to account for these dynamic differences as a function of ligand binding. This hypothesis certainly needs to be tested further with detailed thermodynamic and NMR studies on complexes of CYP101 with these ligands.

Neutron scattering is yet another technique that can be used to gain dynamic information on proteins on faster timescales. Neutron beams scattering off from proteins can undergo change in momentum and energy of the scattered neutrons relative to the incident neutrons leading to elastic or inelastic scattering. Traditionally, this technique has been used to study the properties of materials and the behavior of solutes in water (48). Neutron scattering, unlike neutron diffraction, does not gather information at atomic level resolution, but on a much larger scale and this information can be utilized for larger scale observations, such as global dynamic motions of a protein using inelastic or quasielastic measurements. Such measurements are typically carried out on hydrogenated proteins that are hydrated powders and can be used to track overall

mean-square displacements in the protein as a function of temperature, giving an overall picture of the internal dynamics of the protein. One can interpret this global picture in terms of individual local-level atomic motions by interpreting the scattering parameters in conjunction with molecular dynamics simulations, giving a comprehensive picture of the types of motions happening within a protein as it transitions through various conformational changes as part of its flexibility. Such methods have been applied previously to study global motions in CYP101 (49). Quasielastic neutron scattering was used to compare the global dynamics of substrate-free and camphor-bound forms of CYP101, where it was determined that substrate-free CYP101 exhibited larger amplitude motions than the camphor bound form, although their relaxation times were similar. Analysis of relevant MD trajectories identified similar regions exhibiting dynamic differences between substrate-free and camphor-bound forms and also on much faster timescales. However, the studies could not confirm the presence of any correlated motions between these regions.

The information about correlated motions can be garnered via another property of neutrons that involves coherent scattering, which occurs when the neutrons within the beam are in phase with one another and contribute to the signal, unlike incoherent scattering that is used to determine atomic level motions as a result of dephasing of neutrons due to their spin. Coherent dynamic neutron scattering, which arises from time-correlations in interatomic positions, can be used to derive detailed information on large scale collective motions. We have recently performed preliminary coherent neutron scattering experiments on CYP101 (50) that reveal the presence of in-phase collective dynamic modes running through the entire protein molecule on the ps-ns timescale and have found that these collective modes may be involved in the conformational gating mechanism in CYP101 to allow substrate access by involving similar regions in the substrate recognition sites as identified in this study. Since these studies were performed in the camphor-bound form, it will be intriguing to see whether the collective modes are also present in the substrate-free form and if so, whether the timescales are similar or different to that of the camphor-bound form. Such studies are

planned to be undertaken in the future and if successful have the potential to provide great insight into the mechanistic basis of the origin of conformational dynamics in flexible systems such as P450s.

REFERENCES

1. O'Brien PJ & Herschlag D (1999) Catalytic promiscuity and the evolution of new enzymatic activities. *Chemistry & Biology* 6(4):R91-R105.
2. Konrat R (2014) NMR contributions to structural dynamics studies of intrinsically disordered proteins. *Journal of Magnetic Resonance* 241(0):74-85.
3. Parigi G*, et al.* (2014) Long-Range Correlated Dynamics in Intrinsically Disordered Proteins. *Journal of the American Chemical Society* 136(46):16201-16209.
4. Aharoni A*, et al.* (2005) The 'evolvability' of promiscuous protein functions. *Nature Genetics* 37(1):73-76.
5. Bone R, Silen JL, & Agard DA (1989) STRUCTURAL PLASTICITY BROADENS THE SPECIFICITY OF AN ENGINEERED PROTEASE. *Nature* 339(6221):191-195.
6. Graham SE & Peterson JA (1999) How similar are P450s and what can their differences teach us? *Archives of Biochemistry and Biophysics* 369(1):24-29.
7. Guengerich FP (2008) Cytochrome P450 and Chemical Toxicology. *Chemical Research in Toxicology* 21(1):70-83.
8. Genade T (2003) Control Analysis of Adrenal Steroidogenesis. Masters of Science (University of Stellenbosch, http://tgenade.freeshell.org/thesis/thesisx.htm - tth_fIg3.3).
9. Foye WO (1981) *Principles of Medicinal Chemistry* (Lea & Febiger).
10. Scott EE*, et al.* (2003) An open conformation of mammalian cytochrome P450 2B4 at 1.6-Å resolution. *Proceedings of the National Academy of Sciences* 100(23):13196-13201.
11. Otyepka M, Berka K, & Anzenbacher P (2012) Is There a Relationship Between the Substrate Preferences and Structural Flexibility of Cytochromes P450? *Current Drug Metabolism* 13(2):130-142.
12. Cojocaru V, Winn PJ, & Wade RC (2007) The ins and outs of cytochrome P450s. *Biochimica et Biophysica Acta (BBA) - General Subjects* 1770(3):390-401.
13. Cojocaru VB-M, Kia; Sansom, Mark S. P. ; Wade, Rebecca C. (2011) Structure and Dynamics of the Membrane-Bound Cytochrome P450 2C9. *Plos Computational Biology* 7.
14. Omura T (2010) Structural diversity of cytochrome P450 enzyme system. *Journal of Biochemistry* 147(3):297-306.
15. Pochapsky TC, Kazanis S, & Dang M (2010) Conformational Plasticity and Structure/Function Relationships in Cytochromes P450. *Antioxid. Redox Signaling* 13(8):24.
16. LeLean JE, Moon N, Dunham WR, & Coon MJ (2000) EPR Spectrometry of Cytochrome P450 2B4: Effects of Mutations and Substrate Binding. *Biochemical and Biophysical Research Communications* 276(2):762-766.
17. Thielges MC, Chung JK, & Fayer MD (2011) Protein Dynamics in Cytochrome P450 Molecular Recognition and Substrate Specificity Using 2D IR Vibrational Echo Spectroscopy. *Journal of the American Chemical Society* 133(11):3995-4004.
18. Poulos TL, Finzel BC, & Howard AJ (1986) *Biochemistry* 25:5314.

19. Lee Y-T, Wilson RF, Rupniewski I, & Goodin DB (2010) P450cam Visits an Open Conformation in the Absence of Substrate. *Biochemistry* 49(16):3412-3419.
20. Strickler M*, et al.* (2003) Crystallographic Studies on the Complex Behavior of Nicotine Binding to P450cam (CYP101)†. *Biochemistry* 42(41):11943-11950.
21. Kern KH-WaD (2007) Dynamic Personalitites of Proteins. *Nature* 450.
22. Wand AJ (2001) Dynamic activation of protein function: A view emerging from NMR spectroscopy. *Nature Structural Biology* 8(11):926-931.
23. Daniel RMD, R.V.; Finney, J.L and Smith, J.C. (2003) The Role of Dynamics in Enzyme Activity. *Annual Review of Biophysics and Biomolecular Structure* 32:69-92.
24. Hoffmann A*, et al.* (2011) Quantifying heterogeneity and conformational dynamics from single molecule FRET of diffusing molecules: recurrence analysis of single particles (RASP). *Physical Chemistry Chemical Physics* 13(5):1857-1871.
25. Fenwick RB, van den Bedem H, Fraser JS, & Wright PE (2014) Integrated description of protein dynamics from room-temperature X-ray crystallography and NMR. *Proceedings of the National Academy of Sciences* 111(4):E445-E454.
26. Tenboer J*, et al.* (2014) Time-resolved serial crystallography captures high-resolution intermediates of photoactive yellow protein. *Science* 346(6214):1242-1246.
27. Kay LE (2005) NMR studies of protein structure and dynamics. *Journal of Magnetic Resonance* 173(2):193-207.
28. Mittermaier AK & Kay LE (2009) Observing biological dynamics at atomic resolution using NMR. *Trends in Biochemical Sciences* 34(12):601-611.
29. Hansen AL, Lundström P, Velyvis A, & Kay LE (2012) Quantifying Millisecond Exchange Dynamics in Proteins by CPMG Relaxation Dispersion NMR Using Side-Chain 1H Probes. *Journal of the American Chemical Society* 134(6):3178-3189.
30. Dempsey CE (2001) Hydrogen exchange in peptides and proteins using NMR spectroscopy. *Progress in Nuclear Magnetic Resonance Spectroscopy* 39(2):135-170.
31. Mandell JGF, A.M; Kominves, E.A. (1998) Measurement of Amide Hydrogen Exchange by MALDI-TOF Mass Spectrometry. *Analytical Chemistry* 70:3987-3995.
32. Stoll S*, et al.* (2012) Double electron–electron resonance shows cytochrome P450cam undergoes a conformational change in solution upon binding substrate. *Proceedings of the National Academy of Sciences* 109(32):12888-12893.
33. Lee Y-T, Glazer EC, Wilson RF, Stout CD, & Goodin DB (2011) Three Clusters of Conformational States in P450cam Reveal a Multistep Pathway for Closing of the Substrate Access Channel. *Biochemistry* 50(5):693-703.
34. Verras A & de Montellano PRO (2006) Protein dynamics and imidazole binding in cytochrome P450 enzymes. *Biochem. Soc. Trans.* 34:1170-1172.
35. Verras AA, A. and Ortiz de Montellano, P.R. (2006) Cytochrome P450 active site plasticity: attenuation of imidazole bindin in cytochrome p450cam by an L244 mutant. *Protein Eng. Des. Sel.* 19(11):491-496.

36. Ravichandran KG, Boddupalli SS, Hasermann CA, Peterson JA, & Deisenhofer J (1993) Crystal structure of hemoprotein domain of P450BM-3, a prototype for microsomal P450's. *Science* 261(5122):731-736.
37. Prasad S & Mitra S (2002) Role of Protein and Substrate Dynamics in Catalysis by Pseudomonas putida Cytochrome P450cam†. *Biochemistry* 41(49):14499-14508.
38. Lopes N (2013) Structural characterization of the Redox-Dependent differences in the Cytochrome P450cam-Putidaredoxin Complex using solution NMR spectroscopy. Masters of Science Masters (University of Tennessee, Knoxville, Knoxville, Tennessee).
39. Chodera JD & Mobley DL (2013) Entropy-Enthalpy Compensation: Role and Ramifications in Biomolecular Ligand Recognition and Design. *Annual Review of Biophysics* 42(1):121-142.
40. Ekroos M & Sjögren T (2006) Structural basis for ligand promiscuity in cytochrome P450 3A4. *Proceedings of the National Academy of Sciences* 103(37):13682-13687.
41. Greenblatt DJ*, et al.* (2011) Mechanism of cytochrome P450-3A inhibition by ketoconazole. *Journal of Pharmacy and Pharmacology* 63(2):214-221.
42. Hall EA & Bell SG (2015) The efficient and selective biocatalytic oxidation of norisoprenoid and aromatic substrates by CYP101B1 from Novosphingobium aromaticivorans DSM12444. *RSC Advances* 5(8):5762-5773.
43. Weiss D (2013) H/D Exchange. (http://weisgroup.ku.edu/hd-exchange).
44. Pochapsky SS, Dang M, OuYang B, Simorellis AK, & Pochapsky TC (2009) Redox-Dependent Dynamics in Cytochrome P450(cam). *Biochemistry* 48(20):4254-4261.
45. Goddard TGaK, D.G. (2015) Sparky University of California, San Francisco).
46. Delaglio F*, et al.* (1995) NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J Biomol NMR* 6(3):277-293.
47. Tanio MT, R; Tanaka, T. (2009) Amino acid-selective isotope labeling of proteins for NMR study: proteins secrered by Brevibacillus choshinensis. *Analytical Chemistry* 386:156-160.
48. Mamontov E & Ohl M (2013) Slow dynamics of water molecules in an aqueous solution of lithium chloride probed by neutron spin-echo. *Physical Chemistry Chemical Physics* 15(26):10732-10739.
49. Miao Y*, et al.* (2012) Coupled Flexibility Change in Cytochrome P450cam Substrate Binding Determined by Neutron Scattering, NMR, and Molecular Dynamics Simulation. *Biophysical Journal* 103(10):2167-2176.
50. Hong Lea (2015) Quantitatively Deriving the Large-Scale Collective Dynamics in Protein Molecules using Coherent Neutron Scattering.

VITA

Ana Virginia Bernal Gomez was born in Orlando, Florida, to Pedro Bernal and Rocio Gomez de Bernal. She attended Brookshire Elementary until fourth grade, whereupon her father took his whole family and moved to Santo Domingo, the capital of the small island nation of the Dominican Republic. Ana spent two years at Colegio Amador, and moved back to Orlando for seventh grade at Glenridge Middle School. She was in the marching band and orchestra at Winter Park High School, where she traveled to Memphis, TN for the Liberty Bowl and had her first taste of Tennessee. She went to Rollins College, where she received her degree in Biochemistry and Molecular Biology. During her senior year, she accepted her offer to attend the University of Tennessee to join the Biochemistry, Cellular and Molecular Biology department. Upon receiving her Master's degree, Ana hopes to continue living life to the fullest, and taking some chances now and then.