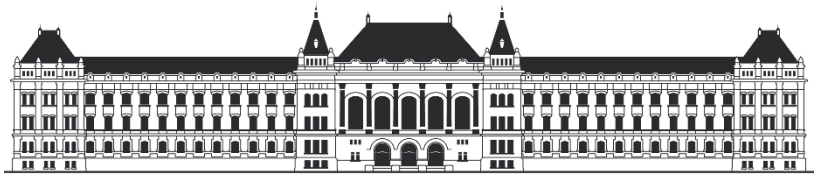

BUDAPESTI UNIVERSITY OF TECHNOLOGY AND ECONOMICS
FACULTY OF CHEMICAL TECHNOLOGY AND BIOTECHNOLOGY
GEORGE OLÁH PHD SCHOOL

Interactions in protein systems

Thesis book

Author: **Anikó Krezinger**

Supervisor: **Dr. Julianna Oláh**

Department of Inorganic and Analytical Chemistry

2017

1 Introduction

During my PhD studies I have investigated interactions in protein systems applying computational chemistry methods. I have investigated the changes in the hydrogen bond structure formed between *protein and solvent* water molecules as the effect of desalting. I have studied how a point mutation affects the *protein-protein* interactions between the subunits of the *Pf*CCT enzyme. I have also investigated *protein-ligand* interactions in the case of *Pf*CCT enzyme and cytidine diphosphocholine, as well as in the case of nitric-oxide and myoglobin, and in a study concentrating on the redox properties of heme containing proteins.

2 Applied methods

Modeling of protein molecules involves systems with a large number of atoms. In all cases, the choice of the applied method has to be determined by the problem, by the physical/ chemical quantity and property to be examined. Therefore, for structural investigations, homology modeling and Newtonian molecular dynamics simulations, for generating conformations Langevin molecular dynamics simulations were used in my PhD studies. Furthermore, energies were determined at the level of molecular mechanical, of DFT quantum mechanical and combined quantum mechanical/molecular mechanical (QM/MM) calculations. Molecular mechanical calculations and molecular dynamics simulations were performed using the CHARMM force field with one of the following programs: CHARMM, GROMACS or NAMD. Quantum mechanical (QM) and QM/MM calculations were carried out using various functionals, e.g. B3LYP, B3PW91, M06-2X, TPSSh and ωB97X-D functionals. The description of the so-called spin-forbidden reactions were carried out using special technique, the MECP method, which was designed for the determination of the geometries and energies of transition between different spin states.

3 Results

3.1 Hydration sphere structures of proteins

The appropriate spatial structure is essential for the activity of proteins. It is affected by both intramolecular interactions in the proteins and the intermolecular interactions formed with the solvent molecules, which is water in living cells. It has been well-established that the dominant conformational motions of proteins are profoundly affected by their hydration shell.[1,2] As a consequence, structural changes of the solvent should inevitably affect protein structure and function as well. Indeed, addition of compounds such as inorganic salts, organic molecules, acids or bases to the solution can perturb the structure of liquid water leading to the denaturation of the biomolecule.

In my PhD study, I have investigated the effect of salt concentration on the properties of the hydrogen bond network around a protein, and the perturbation effect of protein for water structure. I prepared three differently solvated protein systems, which differed in the concentration of salt (neutral, 0.5 M, 1.5 M, respectively), as well as systems without protein, but containing identical amount of salts were prepared (reference systems). Molecular dynamics simulations were carried out in order to generate conformations for each systems. The solvent water molecules were defined in hydration layers, then using statistical analysis hydrogen bond properties in and between the layers were investigated.

The statistical analysis showed that there are significant differences in the hydrogen bond properties (e.g. the number of formed hydrogen bonds and the strength of them) in the different solvent layers, from which the 2nd layer behaves as a transition layer between the bulk phase of the solvent and the protein surface. My results also showed, that water molecules over the hydrophilic and hydrophobic surface of the protein possess slightly different H-bonding properties, supporting the hypothesis of structural and dynamical heterogeneity of the water molecules over the protein surface. The protein affects the hydrogen bonded water network at least up to 4 layers, which is in accordance with recently reported sub-terahertz spectroscopic measurements. My calculations showed that as the salt concentration increases the number of the formed hydrogen bonds decreases, and reaches a similar value to the reference systems in the 3rd and the 4th solvation layers.

3.2 The effect of the R681H point mutation on the structure of *Pf*CCT dimer

Based on the 2014 year data of the World Health Organization the half of the humanity leave in area with high risk of malaria. The most harmful disease in human is infected by *Plasmodium falciparum* (*Pf*). Nowadays therapeutic drugs are available for the treatment of malaria, but due to the resistance to parasites it is necessary to develop the anti-malaria agents.[3] Targeting the lipid biosynthesis of the causative agent Plasmodium parasites is among the promising candidate antimalarial strategies. One of the most common membrane constituent is phosphatidylcholine (PC), which is synthetized via the so-called *Kennedy pathway*. The regulating and rate limiting step of the multistep synthesis is related to *CTP:phosphocholine cytidylyltransferase* (*CCT*) enzyme (Figure 1.). In laboratory experiments the R681 point mutant variant of the dimer *Pf*CCT enzyme had a thermo-sensitive phenotype.

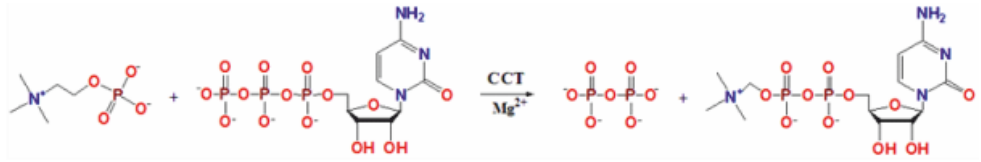

Figure 1. The catalyzed reaction by CCT enzyme

In my PhD study, I have investigated the structural and energetic differences of the wild type and the R681H mutant variants of the *Pf*CCT enzyme, in order to contribute to the explanation of the experimentally observed instability of the dimeric protein. I prepared homology models of the dimeric native and the thermos-sensitive point mutant enzyme variants, then using molecular dynamics simulations ensembles of configurations were generated. I determined the interaction energies between the subunits using the generated conformational assembles, and I calculated the surface of the interacting areas in both enzyme variants. My calculations suggest that the interacting network formed between the subunits in the ^{681}RYVD^{684} conserved motive is broken in the mutant enzyme variant, leading to the moving apart of the subunits and reduced interaction energy between them. Therefore, the interacting surface, mainly consisting of hydrophobic amino acid residues, decreases, resulting in the less favorable solvation of the mutant enzyme variants. These effects taken together with the entropy loss due to decrease of translational and rotational degrees of freedom more pronounced at higher temperature explain the experimentally observed thermos-sensitivity of the dimerization reaction.

3.3 The investigation of many-body interactions in *Pf*CCT enzyme

Non-covalent interactions, such as hydrogen bonds and salt bridges, play a crucial role in all biological systems.[4] Among others, they contribute to the three-dimensional structure of biomolecules, to ligand binding and to enzymatic catalysis. As such, *in silico* drug discovery and design methodologies depend to a large extent on proper description of these intermolecular forces. Cation-π interaction is a relatively newly recognized secondary molecular interaction, whereby a positively charged molecule or molecular moiety interacts with an aromatic ring.[5] The strength of a cation-π interaction is significant; it is comparable to hydrogen bonds and salt bridges.[6]

In my PhD study, I have investigated the cation-π interactions formed upon ligand binding in the *Pf*CCT enzyme (Figure 2.) in the wild type (*Pf*CCT MΔK^{WT}) and in two additional point mutant enzyme variants (*Pf*CCT MΔK^{W692Y} and *Pf*CCT MΔK^{W692F}).

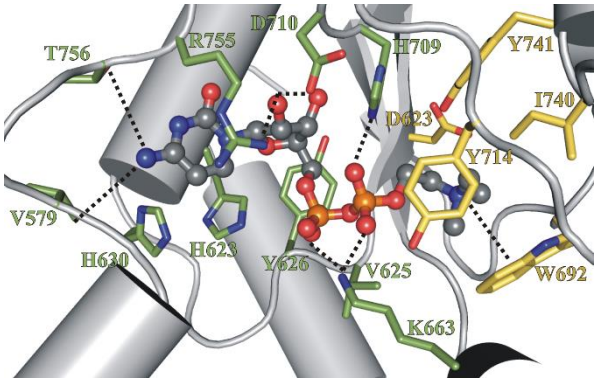

Figure 2. The ligand binding pocket of the *Pf*CCT enzyme (PMDB PM0078719[7])
The ligand molecule is represented by balls, the amino acid residues by sticks. Amino acid residues responsible for the orientation of the choline group of the ligand are colored yellow, and for the orientation of the cytidine-phosphate part are shown in olive, respectively. Interaction with reside involved in the point mutation (W692) is highlighted by dashed line.

Using molecular dynamics simulations, I generated conformational ensembles for the investigated three enzyme variants. I have successfully extended a method for the investigation of the individual contribution of the residues to the ligand binding, as a many-body interaction using QM/MM calculations. I concluded that in the mutant enzyme variants the cation-π interaction decreased between the ligand and the amino acid residue in the 692 position.
My results indicate that 2-body interaction energy terms overestimate the ligand biding energy, but this is compensated by the 3-body interaction energy terms in all cases.

3.4 Cytochrome P450-mediated metabolism of estrogens

Cytochrome P450 enzymes (CYPs) are one of the most versatile enzymes.[8] Their common feature is the heme group in the active site of the structures. The most of the CYP enzymes have monooxygenase activity and catalyze a wide range of reactions. In the last few decades, more and more information has been gathered that CYP enzymes may also bioactivate natural substances, such as hormones to toxic agents. The most striking example of that of estrogens. Despite the clear indication that estrogens exert carcinogenic activity,[9] the underlying mechanisms are still ill-understood. Various metabolic routes of estrogens are known in the human body. At least four CYP isoforms (CYP1A1, CYP1A2, CYP1B1 and CYP3A4) have been suggested to metabolize estrogens leading to various products. The chemical identity of products formed by CYPs are influenced by various factors, including ***(1)*** the active site structure of the enzymes, which ultimately determining the binding affinity and orientation of substrates in the active site, ***(2)*** isoform abundance and distribution in various tissues. In the case of estradiol and its related derivatives, the most prevalent metabolic routes are *2-, 4-hydroxylation* pathways of estrogens (Figure 3.).

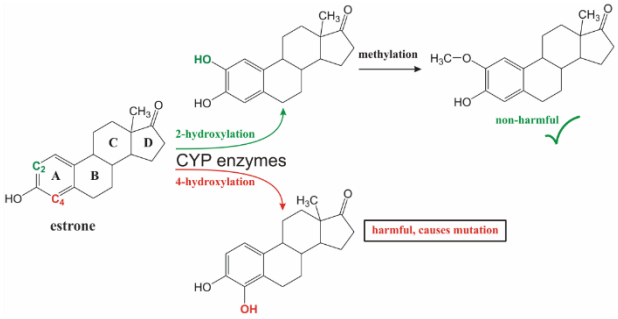

Figure 3. 2- and 4-hydroxylation pathways of estrogens catalyzed by cytochrome P450 enzymes represented through the example of estrone.

In my PhD study, I have investigated the binding of estrone to CYP enzyme variants using molecular docking technique, to reveal how the position of the ligand could influence the 2- vs 4-hydroxylation pathways of estrogen metabolism, in four CYP isoforms. My results suggest that CYP1A1 and CYP1A2 isoforms bind estrone in a way to facilitate the 2-hydroxilation pathway. Due to the large active site of CYP3A4 the possibility exists for metabolism at both sites. QM calculations on the 2- and 4-hydroxylation routes of models of estrone, equilin, and equilenin clearly indicated that C_4 becomes more nucleophilic with increasing conjugation between rings A and B and thus that it will be much more prone to be

oxidized by Compound I of CYP enzymes. This trend is also suggested by determining the activation energies of 2- and 4-hydroxylation pathways.

3.5 The reaction of nitric oxide and myoglobin

Small gas molecules, like nitric oxide and oxygen, are essential for higher animals. In order to provide a stationary concentration of oxygen, specific storage and transporter proteins evolved in nature. Myoglobin, a heme containing protein is primarily responsible for oxygen carrying in muscle,[10] but it is also capable of binding other diatomic molecules, such as nitric oxide (NO) and carbon monoxide (CO), or the cyanide ion (CN^-). Previous experiments indicate that myoglobin is responsible for the deactivation of nitric oxide by binding it to the red muscle tissue.[11]

Several studies have been carried out in order to give molecular or kinetic description of the binding modes of various ligands to myoglobin. However, a complete description of the whole process involving both the diffusion and chemical binding of nitric oxide to myoglobin, and the comparison of the rate constants for the two processes is still missing. Also, the rate-limiting step of the reaction is not known. In my PhD study, I have replaced this missing link.

Using molecular dynamics simulations, I investigated the migration pathways of nitric oxide from the solvent to the active center of myoglobin. I have identified four main migration pathways, which considerably differ in their length. I have calculated the reaction rate constant and compared it to the results from experiments and from gas phase calculations. I have also investigated the chemical binding of nitric oxide to the heme cofactor, especially the so-called spin-forbidden reactions in gas-phase, using QM calculations. I have determined the reaction mechanism by the identification of the resting state in the reactant and product states, and by the identification of the transition from one spin state to another, using special technique, the so-called *Minimum Energy Crossing Point* (*MECP*) method. I established that the reaction starts with a separated heme group in its quintet ground state and nitric oxide in the doublet state and at the end of the reaction a strongly-bound complex in the doublet spin state is formed. The reaction proceeds via the sextet and quartet adducts involving very small reaction barriers. As a consequence, the chemical binding of nitric oxide is very fast and diffusion controls the overall reaction.

3.6 Controlling the reactivity in heme containing enzymes

The heme cofactor is one of the most important prosthetic group in metalloproteins. Several enzymes contain in their active centrum, e.g. cytochrome P450 enzymes, peroxidases and globins. These proteins involved in many biological processes. One of the most remarkable feature of cytochrome P450 enzymes is their ability to activate inert C-H bonds through out a multistep catalytic cycle. During this cycle, an intermediate, the so-called *Compound I* („*CPD I*"), is responsible for the monooxygenation of the substrate.[12] Using experiments and computational techniques the reaction mechanism of monooxygenation was described,[13,14] therefore the generally accepted reaction mechanism is the so-called "*rebound mechanism*" (Figure 4.). During this, CPD I abstracts a hydrogen radical from the substrate resulting in a hydroxo complex, which loses OH radical to the alkane to produce the hydroxylated product. The energy necessary for the formation of the R-OH bond in the substrate is covered by the O-H bond formed in the hydroxo complex.

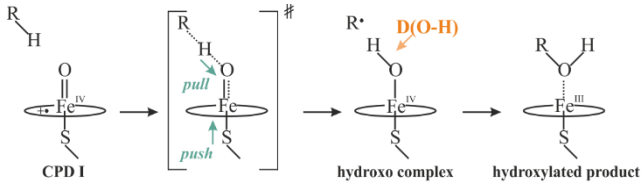

Figure 4. The reaction mechanism of monooxigenation

Recent researches related the hydrogen atom abstract capability of the Fe(IV)-complex to the electron donating capability of the proximal ligand („*L*").[15] The strength of the formed D(O-H) bond, using Hess's law in order to describe the formation/breaking of the O-H bond as an electron and proton transfer, was approximated by the following empirical form[16]:

$$D(O-H) = 23{,}06 \cdot \varepsilon^{0}_{CPD\ I} + 1{,}37 \cdot pKa_{CPD\ II} + constans \qquad (1)$$

where $\varepsilon^{0}_{CPD\ I}$ is the redox potential of the CPD I, and $pKa_{\ CPD\ II}$ is the *pKa* value of the form complex, the so-called *CPD II* complex (one electron reduced form of CPD I).

In my PhD study, I have investigated the monooxygenase activity of five heme containing enzymes. Using molecular dynamics simulations, I generated ensemble of conformations Applying QM/MM calculations I have shown that the redox activity of heme containing enzymes is effected by both the proximal ligand and the protein framework. The proximal framework serves as local control and is responsible for the largest energy contribution to the reaction, while the protein structure function as a global control via the fine tuning of the reaction.

4 Theses

I. I have established, that there are significant differences among the hydrogen bond network properties of the water molecules associated to the hydrophilic and hydrophobic surfaces of insulin, and the various solvent layers. The 2nd solvent layer is a transient layer between the bulk phase and protein surface. (**1st publication**)

II. Using molecular dynamics simulations, I have established, that as a consequence of the R681H point mutation in the *Plasmodium falciparum* CTP:phosphocholine cytidylyltransferase enzyme, the secondary interaction network between the subunits is broken, the subunits move away from each other, allowing their surface to become more accessible to the solvent, and the interaction energy between the dimer subunits is smaller than in the wild type enzyme. (**4th publication**)

III. I have established, that during cytidine diphosphate-choline binding by the *Plasmodium falciparum* CTP:phosphocholine cytidylyltransferase enzyme the most significant interactions among the choline group of the ligand and the surrounding amino acid residues are the 2- and 3-body interaction energy terms, and the 3-body interactions show an anti-cooperative effect with the 2-body terms. Furthermore, in the mutant enzyme variants the cation-π interaction decreased between the ligand and the amino acid residue in the 692 position. (**3rd publication**)

IV. I have established that the C_4 atom of estrone becomes more nucleophilic with increasing conjugation between rings A and B, and thus it is more prone to oxidation by CPD I of CYP enzymes. Its oxidation is facilitated by the orientation of the ligand in the investigated CYP enzymes (1A1, 1A2 and 3A4) when the A ring of estrone is bound close the heme cofactor. (**2nd publication**)

V. I have established, that nitric oxide can reach the active site of myoglobin using 4 migration pathways. The chemical binding is achieved through a series of spin-forbidden reactions involving small activation energies. Based on the calculated rate constants the mechanism of nitric oxide binding is diffusion controlled.

VI. I have established, that the monooxygenase activity of heme containing enzymes is controlled in two different ways: (1) the largest energy contribution to the reaction originates from the local control exerted by the proximal ligand, and (2) a smaller energy contribution to the reaction is due to the global control exerted by the protein framework surrounding the active site.

5 Applications

My goal was the investigation of various protein interactions. During my work, I did fundamental researches. With the help of my results, I supported earlier hypotheses and gave molecular explanations to phenomena unknown so far.

6 Publications[i]

Publications related to the PhD thesis:

1. **A. Lábas**, I. Bakó, J. Oláh Hydration sphere structure of proteins: A theoretical study, *Journal of Molecular Liquids*, **2017**. 238: 462-469 (**DOI:** 10.1016/j.molliq.2017.05.038)
   IF (2016): 3,648, cited: 0
2. **A. Lábas**, B. Krámos, J. Oláh Combined Docking and Quantum Chemical Study on CYP-Mediated Metabolism of Estrogens in Man, *Chemical Research in Toxicology*, 2016. 30(2):583 (**DOI:** 10.1021/acs.chemrestox.6b00330)
   IF (2016): 3,278, cited: 2
3. **A. Lábas**, B. Krámos, I. Bakó, J. Oláh Accurate modeling of cation-π interactions in enzymes: a case study on the CDPCho:phosphocholine cytidylyltransferase complex, *Structural Chemistry*, **2015**. 26: 1411 (**DOI:** 10.1007/s11224-015-0658-9)
   IF: 1,854, cited: 2
4. L. Marton, G.N. Nagy, O. Ozohanics, **A. Lábas**, B. Krámos, J. Oláh, K. Vékey, B.G. Vértessy Molecular Mechanism for the Thermo-Sensitive Phenotype of CHO-MT58 Cell Line Harbouring a Mutant CTP:Phosphocholine Cytidylyltransferase, *PLoS One*, 2015. 10(6):e0129632 (**DOI:** 10.1371/journal.pone.0129632)
   IF: 3,057, cited: 4

Other publications published in print:

5. I. Bakó, J. Oláh, **A. Lábas**, Sz. Bálint, L. Pusztai, MCB Funel Water-Formamide Mixtures: Topology of the Hydrogen-Bonded Network, *Journal of Molecular Liquids*, 2017. 228:25 (**DOI:** 10.1016/j.molliq.2016.10.052)
   IF (2016): 3,648, hivatkozások száma: 2
6. É. Gráczer, T. Szimler, A. Garamszegi, P.V. Konarev, **A. Lábas**, J. Oláh, A. Palló, D.I. Svergun, A. Merli, P. Závodszky, M.S. Weiss, M. Vas Dual Role of the Active Site Residues of Thermus thermophilus 3-Isopropylmalate Dehydrogenase: Chemical Catalysis and Domain Closure, *Biochemistry*, 2016. 55(3):560 (**DOI:** 10.1021/acs.biochem.5b00839)
   IF (2016): 2,938, hivatkozások száma: 1
7. **A. Lábas**, E. Szabó, L. Mones and M. Fuxreiter Optimization of reorganization energy drives evolution of the designed Kemp eliminase KE07, *BBA - Proteins and Proteomics*, 2013. 1834(5):908 (**DOI:** 10.1016/j.bbapap.2013.01.005)
   IF: 3,191, hivatkozások száma: 11

[i] The articles are published using my maiden name, Anikó Lábas.

Oral presentations

1. A. Lábas, J. Oláh and J. Harvey: MD, QM and QM/MM modeling of nitric oxide binding to myoglobin, 2015. CESTC2015 Banska Bystrica, **Slovakia** (Eng)
2. A. Lábas, J. Oláh and J. Harvey: MD, QM and QM/MM modeling of nitric oxide binding to myoglobin, 2015. EcostBio Meeting, Belgrade, **Serbia** (Eng)
3. A. Lábas et. al.: Interactions in protein systems, 2016. George Oláh PhD School - XIII. PhD. Conference, Budapest, **Hungary** (Hun)
4. A. Lábas, J. Oláh, J. Harvey: The investigation of the reaction of NO and myoglobin, 2015. Inorganic and Organometallic Working Committee Meeting, Pécs, **Hungary** (Hun)
5. A. Lábas et. al.: The effect of a point mutation in the structure and function of *Pf* CCT MΔK dimer, 2015. HAS Material and Molecular Structure Working Committee Meeting, Mátrafüred, **Hungary** (Hun)
6. A. Lábas, B. Krámos, J. Oláh: QM/MM study of cation-π interactions in enzymes, 2013. Inorganic and Organometallic Working Committee Meeting, Budapest, **Hungary** (Hun)
7. A. Lábas and E. Szabó: Molecular mechanism of enzymatic evolution – the Kemp eliminase, 2011. XXX. OTDK, Pécs, **Hungary** (Hun)
8. A. Lábas and E. Szabó: Molecular mechanism of enzymatic evolution – the Kemp eliminase, 2011. BME TDK, Budapest, **Hungary** (Hun)

Poster presentations

1. A. Lábas, J. Oláh, J. Harvey: Theoretical study of nitric oxide binding to myoglobin: a combined MD and QM study, 2016. Euro BIC 13, Budapest, **Hungary**
2. A. Lábas, B. Kármos, J. Oláh: QM/MM study of estrogen metabolism by CYPs to toxic agents, 2016. CTCC VII., Krakow, **Poland**
3. A. Lábas, B. Kármos, J. Oláh: Bioactivation of estrogens by CYPs to toxic agents, 2016. Girona Seminar, Girona, **Spain**
4. A. Lábas, J. Oláh, J. Harvey: The spin forbidden reaction of NO binding to myoglobin, 2016. **Virtual Winterschool** on Computational Chemistry
5. A. Lábas, J. Oláh, I. Bakó: Water networks in protein hydration shells: The molecular basis for the denaturation of insulin by sodium chloride, 2016. Athene's Chemistry, Budapest, **Hungary**
6. A. Lábas, J. Oláh, J. Harvey: MD and QM/MM modeling of nitric oxide binding to myoglobin, 2015. FEBS3+ Meeting, Portoroz, **Slovenia**
7. A. Lábas, B. Krámos, T. Szilvási, J. Oláh: QM/MM assessment of the proton and electron affinities of ligated heme iron centers in various enzyme families, 2015. Summer School Groningen, Groningen, **Netherlands**
8. A. Lábas, B. Krámos, T. Szilvási, J. Oláh: QM/MM assessment of the proton and electron affinities of ligated heme iron centers in various enzyme families, 2014. CESTC Conference, Nagybörzsöny, **Hungary**
9. A. Lábas, B. Krámos, J. Oláh: Quantum mechanical (QM) and combined quantum mechanics molecular mechanics (QM/MM) study of cation-π interactions, 2014. George Oláh Ph. D. School - XI. Ph. D. Conference, Budapest, **Hungary**
10. A. Lábas, B. Krámos, G.N. Nagy, B.G. Vértessy, J. Oláh: QM and QM/MM study of cation-π interactions, 2013. 9th European Conference of Computational Chemistry, Sopron, **Hungary**

7 References

1. Frauenfelder, H., Sligar, S. G. & Wolynes, P. G. The energy landscapes and motions of proteins. *Science* **254,** 1598–1603 (1991).

2. Austin, R. H., Beeson, K. W., Eisenstein, L., Frauenfelder, H. & Gunsalus, I. C. Dynamics of ligand binding to myoglobin. *Biochemistry* **14,** 5355–5373 (1975).

3. WHO. *World Malaria Report 2014*. *WHO* (World Health Organization, 2014).

4. Černý, J. & Hobza, P. Non-covalent interactions in biomacromolecules. *Phys. Chem. Chem. Phys.* **9,** 5291–5303 (2007).

5. Zacharias, N. & Dougherty, D. A. Cation-π interactions in ligand recognition and catalysis. *Trends Pharmacol. Sci.* **23,** 281–287 (2002).

6. Dougherty, D. A. The Cation−π Interaction. *Acc. Chem. Res.* **46,** 885–893 (2013).

7. Nagy, G. N. *et al.* Evolutionary and mechanistic insights into substrate and product accommodation of CTP:phosphocholine cytidylyltransferase from Plasmodium falciparum. *FEBS J.* **280,** 3132–3148 (2013).

8. Bernhardt, R. Cytochromes P450 as versatile biocatalysts. *J. Biotechnol.* **124,** 128–145 (2006).

9. Beral, V. & Collaborators, M. W. S. Breast cancer and hormone-replacement therapy in the Million Women Study. *Lancet* **362,** 419–427 (2003).

10. Ordway, G. A. & Garry, D. J. Myoglobin: an essential hemoprotein in striated muscle. *J. Exp. Biol.* **207,** 3441–3446 (2004).

11. Brunori, M. & al., et. Nitric oxide moves myoglobin centre stage. *Trends Biochem. Sci.* **26,** 209–210 (2001).

12. Shaik, S. *et al.* P450 Enzymes : Their Structure , Reactivity , and Selectivity s Modeled by QM / MM Calculations. *Chem. Rev.* **110,** 949–1017 (2010).

13. Groves, J. T. Key elements of the chemistry of cytochrome P-450: The oxygen rebound mechanism. *J. Chem. Educ.* **62,** 928–931 (1985).

14. Ortiz de Montellano, P. R. *Cytochrome P450 : structure, mechanism, and biochemistry*.

15. Groves, J. T. Enzymatic C–H bond activation: Using push to get pull. *Nat. Chem.* **6,** 89–91 (2014).

16. Green, M. T. C-H bond activation in heme proteins: the role of thiolate ligation in cytochrome P450. *Curr. Opin. Chem. Biol.* **13,** 84–88 (2009).