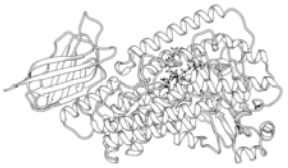
Understanding the Factors Governing the Reaction Specificity of Lipoxygenases A Theoretical Approach

Patricia Saura Martínez

Understanding the Factors Governing the Reaction Specificity of Lipoxygenases

A Theoretical Approach



Patricia Saura Martínez

Ph.D. Thesis
Ph.D. program in Chemistry

Supervisors:
Josep M. Lluch López
Àngels González Lafont

Departament de Química
Facultat de Ciències

2016

UNIVERSITAT AUTÒNOMA DE BARCELONA
DEPARTAMENT DE QUÍMICA
UNITAT DE QUÍMICA FÍSICA

JOSEP M. LLUCH LÓPEZ, Catedràtic del Departament de Química i membre de l'Institut de Biotecnologia i de Biomedicina i ÀNGELS GONZÁLEZ LAFONT, Catedràtica del Departament de Química i membre de l'Institut de Biotecnologia i de Biomedicina de la Universitat Autònoma de Barcelona, certifiquem que

Patricia Saura Martínez ha realitzat sota la nostra direcció, en el Departament de Química i l'Institut de Biotecnologia i de Biomedicina de la Universitat Autònoma de Barcelona, el treball d'investigació titulat:

Understanding the Factors Governing the Reaction Specificity of Lipoxygenases: A Theoretical Approach

que es presenta en aquesta memòria per optar al grau de Doctora en Química (amb Menció de Doctorat Internacional).
I perquè consti als efectes escaients, signem aquest certificat a Bellaterra a 18 de novembre de 2016:

Dr. Josep M. Lluch López

Dra. Àngels González Lafont

Contents

I Introduction v

1 Introduction 1

1.1 Overview 1

1.1.1 Classification 2

1.1.2 Naturally occurring substrates 3

1.1.3 Structure 4

1.2 Catalytic properties 8

1.2.1 Reaction mechanism 8

1.2.2 Kinetics 10

1.2.3 Factors determining lipoxygenases reaction specificity 11

1.3 Biological role of lipoxygenases 16

1.3.1 Metabolites 16

1.3.2 Pathological implications 17

1.3.3 Inhibitors 19

II Objectives 23

2 Objectives 25

III Theoretical Background 27

3 Theoretical background 29

3.1 Preamble 29

3.2 Molecular Mechanics 29
3.2.1 The Force Field energy expression 30
3.3 Quantum Mechanics 32
3.3.1 Wavefunction methods 32
3.3.2 Density Functional Theory 32
3.4 QM/MM Methods 34
3.4.1 Overview 35
3.4.2 QM/MM energy expression 35
3.4.3 The QM-MM boundary 38
3.5 Free Energy Calculations 40
3.5.1 Free Energy Perturbation 41
3.5.2 Umbrella Sampling 42
3.6 Molecular Dynamics 45
3.6.1 Equations of motion 45
3.6.2 NVT and NPT ensembles 46
3.6.3 Boundary conditions 47
3.7 Homology Modeling 49
3.8 Docking simulations 51
3.8.1 Search algorithms 51
3.8.2 Scoring functions 53
3.8.3 Validating docking results 54

IV Results and Discussion 55

4 15-LOX-1 57
4.1 Introduction 57
4.2 Methodology 59
4.3 Results 60
4.3.1 Selection of starting geometries 60
4.3.2 Potential Energy Profiles and Reaction Mechanism . . 62
4.3.3 Comparison between H_{13} and H_{10} abstraction processes 69

4.3.4 Study of the hydrogen abstraction step in LA 80
4.4 Conclusions 83

5 Leu597Ala-Ile663Ala 15-LOX-1 double mutant 87
5.1 Introduction 87
5.2 Methodology 87
5.2.1 Preparation of the *in silico* mutant 87
5.2.2 MD simulations 88
5.2.3 QM/MM calculations 89
5.3 Results and Discussion 90
5.3.1 Multiple binding modes of the AA:Leu597Ala-Ile663Ala/15-LOX-1 complexes 90
5.3.2 Reaction mechanism for the hydrogen abstraction step 95
5.3.3 Experimental results 101
5.3.4 Rotation of the peroxy moiety 101
5.3.5 Reduction of the peroxy radical and regeneration of the Fe(III)-OH$^-$ cofactor 105
5.4 Conclusions 107

6 Ile418Ala 15-LOX-1 mutant 109
6.1 Introduction 109
6.2 Methodology 110
6.3 Results and Discussion 110
6.3.1 Multiple binding modes of the AA:Ile418Ala/15-LOX-1 complexes 110
6.3.2 Study of the hydrogen-abstraction step 117
6.4 Conclusions 122

7 5-LOX 123
7.1 Introduction 123
7.1.1 Description of the X-ray structures 123
7.2 Methodology 125
7.2.1 Homology Modeling 126

7.2.2 Docking Simulations 128
7.2.3 *In silico* mutants 128
7.2.4 Molecular Dynamics Simulations 129
7.2.5 QM/MM Calculations 130
7.3 Results and Discussion 132
7.3.1 Homology Modeling 132
7.3.2 Generation of AA:Stable-5-LOX complexes 134
7.3.3 Reactivity study 148
7.3.4 Study of the stability of the complexes: long MD simulations 159
7.4 Conclusions 169

8 8*R*-LOX 173

8.1 Introduction 173
8.2 Methodology 175
8.2.1 Preparation of the initial coordinates of the AA:-8*R*-LOX complex 175
8.2.2 *In silico* mutagenesis 175
8.2.3 Molecular Dynamics simulations 175
8.2.4 QM/MM calculations 176
8.2.5 Free Energy Calculations 177
8.3 Results and Discussion 178
8.3.1 Hydrogen abstraction step 178
8.3.2 Oxygen insertion step 184
8.3.3 Study of the Gly427Ala mutant 190
8.4 Conclusions 195

V Conclusions 197

9 General conclusions 199

Part I

Introduction

Chapter 1

Introduction

1.1 Overview

Lipoxygenases(LOXs) constitute a family of non-heme, non-sulfur iron dioxygenases that catalyze the oxidation of lipid compounds.[1–3] These enzymes are widely distributed among nature, as they occur in animals, plants, fungi and bacteria. The first LOX discovered was the Soybean Lipoxygenase-1 (SLO-1) in 1932,[4] and for a long time it was believed that the presence of LOXs was restricted to the plant kingdom. In 1974 Hamberg and Samuelsson discovered a LOX variant in human platelets, which was named later as 12-LOX.[5] Nowadays the number of known isoforms belonging to this family has widely expanded; in the case of human, there are six functional genes that encode for lipoxygenases.[6]

Lipoxygenases produce the peroxidation of polyunsaturated fatty acids containing at least one 1,4-Z,Z-pentadiene (bisallylic methylene) units,[7–10] yielding a conjugated hydroperoxy derivative that is further metabolized into signaling compounds with particular roles in the different organisms, for example, green leaf volatiles and jasmonic acids in plants,[11–13] and lactones in microorganisms.[14,15] In mammalians the main substrate of LOXs is arachidonic acid (AA), the hydroperoxidation of which initiates the biosynthesis of lipid mediators that are crucial for human health, as they are implicated in a number of diseases: leukotrienes (pro-inflammatory lipid mediators) and lipoxins (anti-inflammatory lipid mediators), are involved in the inflammatory and immunity processes, which are supposed to be at the root of most illnesses.

1.1.1 Classification

Family of lipoxygenases enzymes is constituted of various isoforms, and there is not a universal classification system as there are various different aspects that can be taken into account (reactivity, tissular presence, evolutionary relatedness). One extended classification accounts for the positional specificity of the main substrate of LOXs. In the case of mammalian LOXs arachidonic acid (AA) is the most relevant substrate, and the different mammalian LOX isoenzymes catalyze its oxidation with high regio- and stereospecificity, in such a way that they are named 5-, 8-, 12-, and 15- according to the specific position of AA where the oxygen molecule is added (Figure 1.1).[3,10,16–19] This number is sometimes accompanied by the *S* or *R* denomination that specifies the product stereochemistry. As most structures oxidize fatty acids with *S*-stereochemistry, the "*S*-" is omitted, but *R*-LOX nomenclature is usually employed. In addition, the name of the isoform can be completed with the tissue where the enzyme is located. In that way mammalian 15-LOX can be subclassified into reticulocyte type 15-LOX (or 15-LOX-1) and epidermal type 15-LOX (or 15-LOX-2).

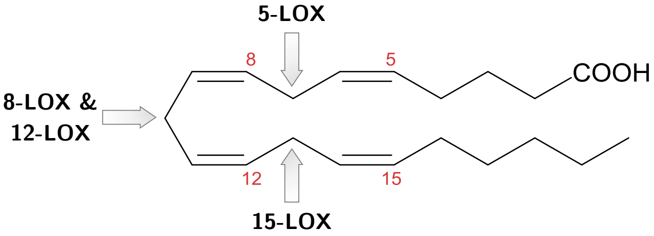


Figure 1.1: Classification system of mammalian LOXs according to the oxidation positions in arachidonic acid. Numbers in red indicate the carbon atom position where oxygen molecule is added to form the hydroperoxide derivative. The different starting positions for hydrogen abstraction are indicated by arrows, showing the LOXs isoform that acts over the corresponding position in each case.

This classification system is probably the most extended, but there are some authors that prefer alternative classification systems. The positional-specificity classification only considers the reaction specificity towards one substrate, arachidonic acid, but there are some mammalian LOXs for which this substrate is not the most important, or even more, there are other substrates (i.e. linoleic acid) with a different reaction specificity. Another aspect is that AA-position classification system does not take into account evolutionary aspects of the organisms. Different LOX isoforms that share a high degree of evolutionary relatedness may have different reaction specificity. That is the case of human gene ALOX15 (the nomenclature of the genes contains four letters and a number) which actually encodes a 15-LOX-1, and its homologous mouse gene alox15 which turns out to encode a 12-LOX in terms of reaction specificity.

Within this Thesis the positional-specificity classification system will be employed, as the main interest focuses on how the different isoforms yield different oxidation products starting from a common substrate, with an exquisite regio- and stereoselectivity. Such isoenzymes are roughly similar, and there are particular differences (specially in the active site of the protein) that are responsible of this specificty. This specificity would be required as the biological functions of the LOXs metabolites present huge disparity, sometimes producing opposite effects.

1.1.2 Naturally occurring substrates

Arachidonic acid (Figure 1.2) constitute the most important substrate in mammalian LOXs. Arachidonic acid or 5*Z*,8*Z*,11*Z*,14*Z*-eicosatetraenoic acid, is a 20 carbon atom polyunsaturated essential fatty acid from the ω-6 series that contains four unsaturations in positions 5, 8, 11 and 14. It is also named as acid (20:4), where the first digit stands for the total number of carbon atoms, and the second one for the number of double bonds.

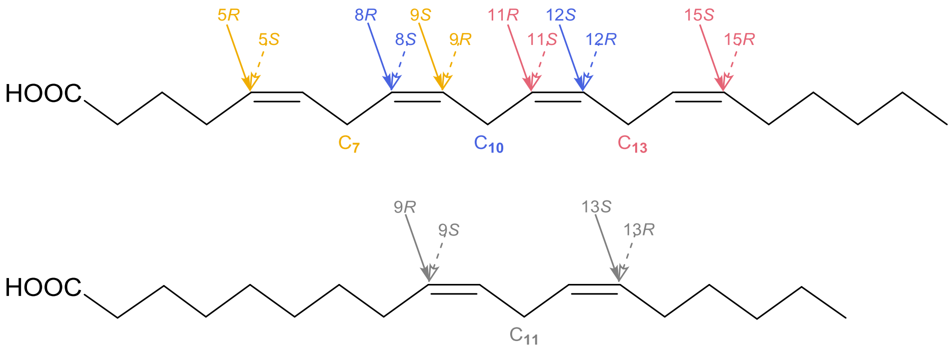


Figure 1.2: Representation of arachidonic acid and linoleic acid structures showing all possible oxidation positions. Oxidation positions are indicated in different colors according to the carbon atom from which a hydrogen is abstracted in the case of AA: C_7 positions are depicted in yellow, C_{10} positions in blue, and C_{13} positions in red. In the case of LA hydrogen atoms can be only abstracted from C_{11}, depicted in grey.

AA is the most important substrate in mammalian LOXs but it is not the only one. Most LOXs isoforms exhibits a wide substrate specificity, as there are a large number of fatty acid species that can be oxidized. All naturally occurring polyunsaturated fatty acids (PUFAs), including linoleic acid, linolenic acid and α-linolenic acid can be easily oxigenated by a LOX. Linoleic Acid (LA) is an essential fatty acid precursor of AA in animals, and it is the most important substrate for plant LOXs. LA belongs to the ω-6 series and contains two unsaturations in positions 9 and 12.

Other important natural fatty acids substrate are the ω-3 series eicosapentaenoic acid (EPA), and docosahexaenoic acid (DHA), that are present in marine organism (fish oil). These fatty acids serve as precursors of some anti-inflammatory lipid mediators (resolvins and protectins). Several LOXs isoforms can oxidize lipid esters (phospholipids, cholesterol esters, mono-, di- and triglycerids) that contain PUFAs, and even protein-lipid complex (membranes and lipoproteins) are LOXs substrates.[20–22]

1.1.3 Structure

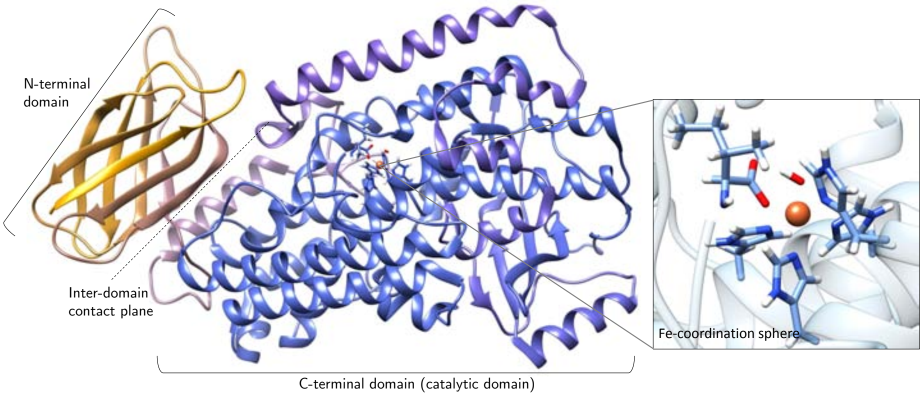


Figure 1.3: Representation of a structure of LOX (15-LOX-1). N-terminal domain is depicted in yellow: it is composed by eight anti-parallel β-strands that form a β-barrel. C-terminal domain is depicted in blue: it is mainly a helical domain, formed by 18 α-helices. Dashed line represents the interdomain contact plane. Box on the right shows a detail of the iron-coordination sphere.

Lipoxygenases display a single polipeptide chain that folds into two differentiated domains: a small N-terminal domain, and a large C-terminal domain which contains the active site of the protein. The N-terminal domain of all LOX isoforms consists of a β-barrel composed by 8 anti-parallel β-strands, and known as PLAT (Polycistin1 Lipoxygenase Alpha Toxin) domain. This domain was initially named as C_2-like domain (Coactosin-like protein) due to its resemblance to the calcium dependent membrane binding domain of phospholipases. Its function is conferring LOXs membrane binding capacity, especially in the case of 5-LOX.[23,24] Although there is a high degree of structural conservation in PLAT domain among LOXs, its sequence displays significantly less conservation than the catalytic domain.

The C-terminal domain is the largest domain, comprising about 80% of the total structure, and contains the substrate binding pocket and the catalytic non-heme iron. It is composed by around 20 α-helices that are interrupted once by a small subdomain of β-strands. The catalytic domain nucleus presents two long central helices, in which one of them adopts a π-helix conformation in part of its sequence. Four π-helical sections contain four of the five iron ligands. The two differentiated domains are separated by a loop region, and they keep contact by an interdomain plane.

This common core is found among all LOXs isoforms, but there are small differences, especially in the catalytic domain, that confers each isoforms its unique catalytic properties. Comparison among structures of rabbit 15-LOX-1, human 5-LOX, and coral 8*R*-LOX is given in detail below.

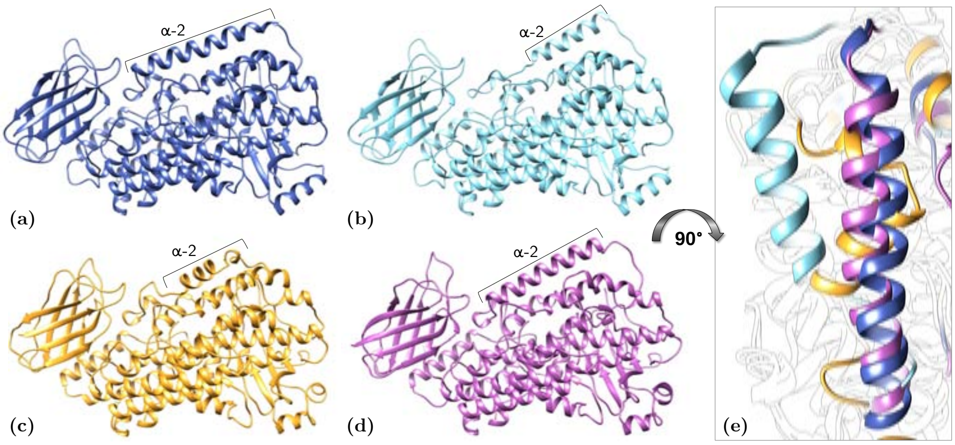


Figure 1.4: Representation of the complete structures of several lipoxygenases available in Protein Data Bank: a) Rabbit 15-LOX-1, subunit A, b) Rabbit 15-LOX-1, subunit B, c) Human 5-LOX, subunit A, d) Coral 8*R*-LOX, subunit A. e) A clear differentiative feature among the isoforms is found in α-2 helices structures: in 15-LOX-1 subunit A and 8*R*-LOX this is a long helix that covers the active site; in the case of 15-LOX-1 subunit B, crystallized with an inhibitor, there is a shortening and a 30° displacement of the helix; finally, in 5-LOX this helix is composed by two short helical segments and one turn of 90° between them.

1.1.3.1 15-Lipoxygense-1

Reticulocyte-type 15-LOX-1 has been one of the most studied mammalian LOXs during the past years. The crystallographic X-ray structure of the human 15-LOX-1 (15-hLOX-1) has still to be reported, but the corresponding X-ray structure of the analogous rabbit 15-LOX-1 was the first one to be resolved for a mammalian LOX;[25–27] human shares about 80% of sequence identitiy with rabbit.

Its catalytic domain is compound by 18 α-helices. The iron-coordination sphere contains four histidine residues (His361, His366, His541 and His545) and the carboxylate end of the terminal Ile663 residue. N-terminal domain contacts with this C-terminal domain by a contact plane of 1600 Å^2.[26] The most recent crystal structure of 15-LOX-1 (PDB ID 2P0M)[27] was solved in the presence of an inhibitor, RS7, and has been crystalized as a dimer in which only one of the subunits contains the inhibitor molecule (subunit B) whereas subunit A is a substrate-free form. The presence of this inhibitor, which is classified as a competitive inhibitor due to its resemblance to AA, induces conformational changes mainly in α-2 helix, that is displaced 30° with respect to the position in the inhibitor-free monomer, and smaller changes in α-18 helix (Figure 1.4 shows a comparison of the monomers and a detail of α-2 helices). The structure of the AA:15-LOX-1 Michaelis complex was previously determined in our group starting from the inhibitor-bound subunit,[28] and has been used as starting point for the work developed in this Thesis.

1.1.3.2 5-Lipoxygenase

The first X-ray structure solved of human 5-LOX does not correspond to the wild-type (WT) human 5-LOX, as it presents intrinsic instability what complicates the crystallization process. Thus, the so called human Stable-5-LOX (PDB code 3O8Y)[29] form of the enzyme that consists of a stabilized and soluble mutant of human 5-LOX has been solved. This specie lacks the putative membrane insertion residues so that the structure is solubilized, and also contains the triple mutation K653E-K654N-K655L, based on the sequence of coral 8*R*-LOX, that confers stability, but retains all the reactive features of the wild-type. Starting from this structure additional crystal structures have been solved based on an additional mutation that could lead to a phosphorylated mimic Ser663Asp (pdb codes 3V98 and 3V99) or the non-phosphorylated one Ser663Ala (pdb code 3V92). Among those structures, 3V99 consists of a ligand-bound form, in which one of the subunits contains the AA substrate (AA:Stable-5-LOX-Ser663Asp).[30]

The structure of 5-LOX presents some differences from the general scheme of LOXs structure. In the catalytic domain of human 5-LOX it is found an arched helix (which contains Lys414), that has been proposed to control the oxygen molecule access. There is also a clear difference in the folding of α-2 helix (Figure 1.4). The classical folding reported for 15-LOX-1 (ligand-free form) consists of a six to seven turns long helix that defines one edge of the active site. In the case of 5-LOX the region where α-2 should be placed is composed by two short perpendicular helical segments of three turns that are separated by a loop region. This unique α-2 arrangement causes a very different active site cavity structure in human 5-LOX. The rest of the secondary structure elements are preserved. The structure of the Fe-coordination sphere also differs from 15-LOX-1: there are three conserved Histidines (His367, His372, and His550), and the carboxylate end of terminal Ile673, but the fourth ligand is an Asparagine (Asn554).

There are also differences between the 5LOXs substrate-free and substrate-bound forms. While all the substrate free structures (with and

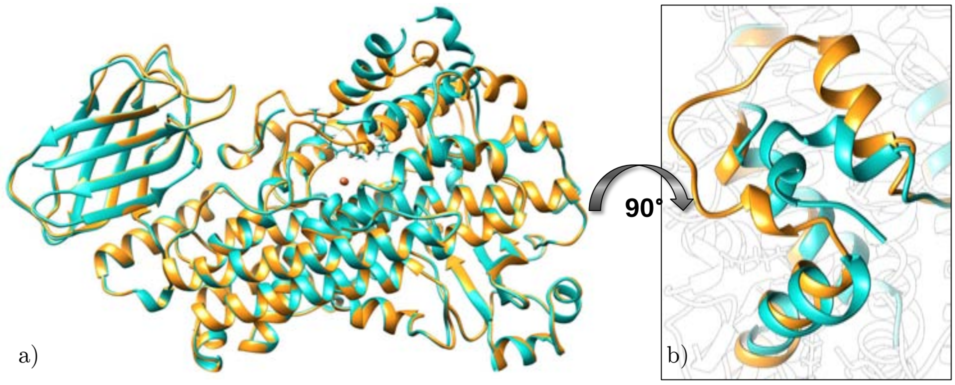


Figure 1.5: Representation of the structures of Stable-5-LOX (yellow) and the monomer A of Stable-5-LOX-S663D mutant (cyan), which contains the AA substrate. In a) a complete view of the superposition of both structures is shown; b) shows a more detailed top view of α-2 helices.

without the S663D mutation) are highly similar, the comparison with the substrate bound form shows an enzyme with substantial plasticity (Figure1.5). Between the unbound and bound system the enzyme has suffered a major conformational change, mainly in the region of α-2 helix (its end part) and the arched helix. In addition, a displacement of the 18 helix is observed, and other regions (some affecting α-2 helix) are disordered in the holoenzyme crystal structure but not in the apoenzyme one. However, the structure of the holoenzyme form is poorly solved, as it lacks some regions, has low resolution, and the coordinates of the substrate are not clear as the same authors warn. Therefore, the final structure of the substrate-bound form maybe differs from this X-ray structure.

1.1.3.3 8*R*-Lipoxygenase

LOXs from coral are larger than mammalian LOXs.[31] The catalytic domain of coral 8*R*-LOX shares the common scheme. Helix α-2 is a long helix of seven to eight turns, with a similar arrangement to 15-LOX-1 substrate-free (Figure 1.4). There is an arched helix that is kinked by two conserved amino-acids insertion in one helical turn, and lines the substrate binding site. The catalytic iron is coordinated to three Histidine residues (His384, His389 and His570), an Asparagine (Asn574) and the terminal Ile693. For this system there has been also solved a substrate free form (PDB ID 4QWE) and a substrate-bound form (PDB ID 4QWT), both of them crystalized as tetramers. In the case of structure 4QWT, only one of the monomers contains the substrate. When the two structures are compared there are not apparent differences, suggesting that the substrate entry in 8*R*-LOX does not induce major conformational changes.[32]

1.2 Catalytic properties

1.2.1 Reaction mechanism

The oxidation process of fatty acids performed by LOXs generally consists in three elementary steps: i) hydrogen abstraction , ii) oxygen insertion and iii) reduction of the peroxyl radical (Figure 1.6). The first step consists of a hydrogen abstraction from one bisallylic methylene of the fatty acid substrate by the enzyme Fe(III)-OH$^-$ cofactor to produce a pentadienyl radical and Fe(II)-OH$_2$ cofactor. This step has been proposed as the rate-limiting step, because a huge kinetic isotopic effect (KIE) was obtained for Soybean Lipoxygenase-1 (SLO-1) and attributed to a significant contribution of quantum tunneling.[33] In fact, SLO-1 has been reported as the biological system with the largest KIE. On the other hand, in SLO-1 it has been elucidated that the hydrogen abstraction process actually corresponds to a Proton Coupled Electron Transfer process (PCET),[34,35] being Fe(III) the electron acceptor and the hydroxyl group the proton acceptor.

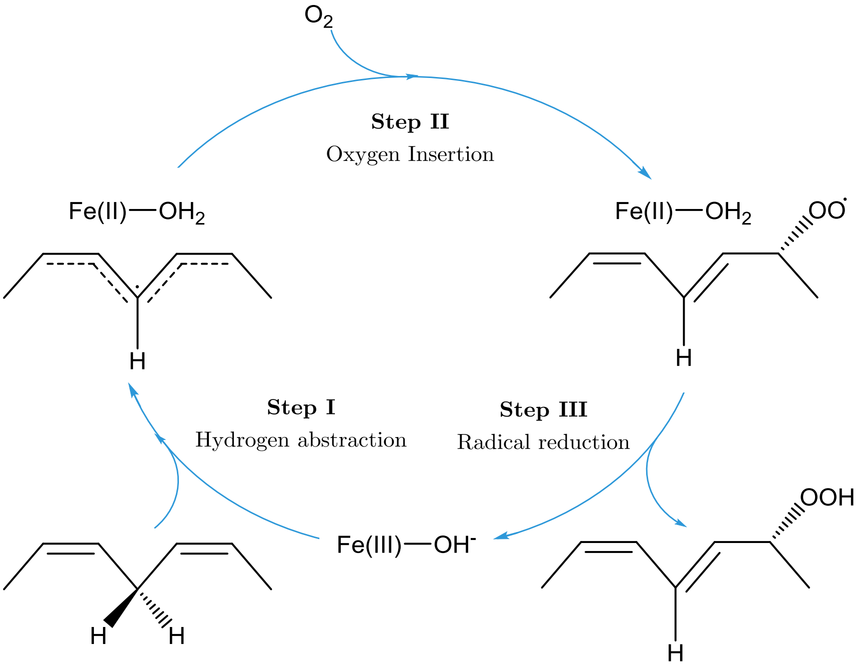


Figure 1.6: General steps of the reaction mechanism of lipoxygenases.

The first step also determines the regiospecificity of the global hydroperoxidation process, as it highly depends on which is the particular pentadiene group of the fatty acid that will transfer a hydrogen atom. There are three bisallylic methylene groups in AA in positions 7, 10 and 13, each containing two CâĂŞH bonds. These methylene groups consists of *sp*^{3} carbon atoms with two hydrogen atoms that could be abstracted. Then, there exists 12 possible oxidation products from AA (Figure 1.2). In an aqueous phase reaction all these possible hydroperoxidation products will be obtain in equal proportions, but in a LOX enzymatic system, only (or mostly) one of this products will be obtained. The structure of LA is more simple, as it only contains one bisallylic methylene group in position 11 (Figure 1.2). Then there are only two possible hydrogen atoms that can be transferred (pro-S and pro-R hydrogens in C_{11}) and four oxidation positions (S and R at positions 9 and 13). It is assumed that the nearness of the bisallylic methylene carbon atoms to the Fe(III)-OH$^-$ moiety highly determines the hydrogen atom to be abstracted. For a given fatty acid that closeness would greatly depend on the depth and width of the substrate binding pocket and the head/tail (carboxyl end or methyl end first, respectively) orientation of the incoming fatty acid. This way, each LOX isoenzyme would abstract a hydrogen atom from essentially a unique position.

The second reaction step consists of the attack of molecular oxygen over the pentadienyl radical yielding a peroxyl radical. Both processes (hydrogen abstraction and oxygen insertion) occur by opposite sides of the pentadienyl plane,[36,37] and it is generally accepted that the process is produced with *antarafacial* stereochemistry. However, to speak in terms of antarafacial or suprafacial, both hydrogen abstraction and oxygen insertion steps should be coupled, and molecular basis of such coupling have not been explored. For simplicity, the term *antarafacial* will be employed as a synonym of “opposite sides of the pentadienyl plane” in this Thesis (Figure 1.7).

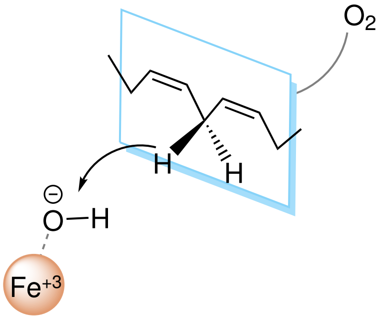


Figure 1.7: Hydrogen abstraction and oxygen insertion processes take place by opposite sides of the pentadienyl plane.

The third and last reaction step consists of the reduction of the peroxyl radical to yield the final hydroperoxy fatty acid and the regeneration of the Fe(III)-OH$^-$ cofactor of the enzyme. In this step a hydrogen atom needs to be transferred to the -OO$^\cdot$ group of the peroxyl substrate, and it is assumed that this atom proceeds from the Fe(II)-OH$_2$ cofactor, as the Fe(III)-OH$^-$ needs to be restored. In order to produce this hydrogen transfer, the peroxyl radical needs to evolve from an initial antarafacial configuration to a suprafacial arrangement to approach the -OO$^{\cdot}$ moiety to the enzyme cofactor. When the starting substrate is AA, the final product is n-H(p)ETE (hydro(per)oxyeicosatetraenoic acid), where n refers to the position of the hydroperoxy moiety. For LA, the product is n-H(p)ODE (hydro(per)oxyoctadecaenoic acid).

1.2.2 Kinetics

LOXs kinetics properties have been widely experimentally studied. The measured values of the kinetic parameters (K_{M}, k_{cat} and $k_{\mathrm{cat}}/K_{\mathrm{M}}$) and KIE, strongly depend on the assay conditions (pH, temperature, and concentration of other substances in the medium) but a summary of the comparison of these values among the different LOXs isoforms is given here in order to provide a better understanding of the differences in reaction and kinetic parameters.

15-LOX-1 presents k_{cat} and K_{M} values of 9.93 ± 0.28 s^{-1} and 3.7 ± 0.3 μM for H_{13} abstraction from AA.[38] In human 15-LOX-1 H/D KIEs for hydrogen abstraction from C_{13} and C_{10} have been experimentally determined by Holman and coworkers.[39] Both results turn out to be comparable (${}^{\mathrm{D}}k_{\mathrm{cat}} = 11.6 \pm 0.2$ s^{-1} for H_{13} abstraction and ${}^{\mathrm{D}}k_{\mathrm{cat}} = 8.5 \pm 4.0$ s^{-1} for H_{10} abstraction), as well as the corresponding ${}^{\mathrm{D}}k_{\mathrm{cat}}/K_{\mathrm{M}}$ values, which led the authors to suggest that these two hydrogen abstraction processes starting from different positions (C_{13} and C_{10}) follow a comparable mechanism, with probably very similar transition state structures. The same authors establish that this assumption is quite unexpected, as it implicates that a highly regiospecific enzyme accommodates two spatially separated carbon atoms centers in a similar way. Interestingly, there is an inversion of the regiospecificity in the reaction of human 15-LOX-1 with AA when the C_{13} position is dideuterated, being the predominant product the one derived from hydrogen abstraction from C_{10}, with a C_{13}:C_{10} ratio of 25:75,[40] whereas this C_{13}:C_{10} ratio is 89:11[38] in the protic enzyme.

Although AA is the most important substrate for mammalian LOXs, steady-state kinetic experiments show that 15-LOX-1 metabolizes LA better than AA, when comparted with 15-LOX-2 for which LA is a worse substrate. For instance, the values of the ratios $(k_{\mathrm{cat}}/K_{\mathrm{M}})^{\mathrm{AA}}/(k_{\mathrm{cat}}/K_{\mathrm{M}})^{\mathrm{LA}}$ and ${k_{\mathrm{cat}}}^{\mathrm{AA}}/{k_{\mathrm{cat}}}^{\mathrm{LA}}$ for human 15-LOX-1 are of 0.80 ± 0.1 and 0.68 ± 0.09 respectively, whereas for human 15-LOX-2 turn out to be 8.0 ± 1.0 and 5.3 ± 0.5, respectively.[39]

Kinetic parameters for human 5-LOX have also been provided.[29] WT human 5-LOX is an unstable system for which high flexibility and short half-life have been described. The Stable-5-LOX version somehow overcomes that problem, but still is less stable that other isoforms. K_{M} and k_{cat} for WT-5LOX and Stable-5-LOX have been reported to be similar, with values of 11 ± 0.02 μM and 0.30 ± 0.02 s^{-1} for WT-5LOX and 42 ± 11 μM and 0.43 ± 0.06 s^{-1} for Stable-5-LOX, respectively. This is a low rate constant when compared with other LOX isoforms, which could be attributed to its higher flexibility and lower half-life and stability. Conversely, coral 8*R*-LOX exhibits a high rate constant, with k_{cat}, K_{M}, and $k_{\mathrm{cat}}/K_{\mathrm{M}}$ values of 210 ± 23 s^{-1}, 30 ± 7 μM and 6.9 ± 1.7 s$^{-1}\mu$M^{-1}, respectively.[32]

1.2.3 Factors determining lipoxygenases reaction specificity

The LOX-catalyzed hydroperoxidation process starts with the positioning of the fatty acid substrate inside the active site cavity of the enzyme. The substrate binding-mode is an important factor that is probably related with the product reaction specificity, and that depends on the configuration of the active site cavity. In this respect, it is widely assumed that the proximity of the methylene group of the fatty acid substrate that transfers a hydrogen atom to the Fe(III)-OH$^-$ cofactor in the first reaction step is the key feature that determines the regiospecificity of the global hydroperoxidation process.[41–43] Thus, in principle, the specific configuration of the cavity and the nature of the substrate (flexibility) would allow that only one (or mostly one) bisallylic methylene unit could be close enough to the catalytic iron. Arachidonic acid, the main substrate in mammalian LOXs, contains three possible pentadiene units that can be reactive for a LOX, at positions 7, 10 and 13. This substrate possesses very high conformational flexibility, with more than 10^7 low energy conformations,[44] what makes challenging the search of bioactive conformations.

As indicated in Section 1.2.1, after hydrogen abstraction an oxygen molecule is added into the pentadienyl radical moiety. Oxygen can be inserted in two possible positions of the pentadienyl group, the $n + 2$ and the $n - 2$ positions, where n refers to the atom carbon number from which a hydrogen atom has been abstracted. In the case of 15-LOX-1, hydrogen abstraction is mostly produced from C_{13}, but also small amounts of the product that comes from abstraction from C_{10} are obtained. As it presents this dual regiospecificity, 15-LOX-1 is sometimes referred as 15/12-LOX. In rabbit 15-LOX-1 it has been determined that pro-*S* hydrogen abstraction from AA is produced with a product ratio of 97:3 (H_{13}/H_{10}),[40] and in the case of human 15-LOX-1 the ratio is 89:11.[45] When hydrogen abstraction takes place from C_{13}, oxygen is added in the $n + 2$ position, giving the oxidation product in position 15. Oxidation in the $n - 2$ position is less important, being the C_{15}:C_{11} product ratio of 85:4[38] and 85:1[46] for human and rabbit 15-LOX-1, respectively. Minor hydrogen abstraction from position H_{10} only leads to 12-HETE product formation, the corresponding $n + 2$ oxidation product.[38]

For wild-type human 5-LOX, hydrogen abstraction takes place essentially from C_7 in AA, with $> 98\%$ of 5-HETE formed (and sharing $< 2\%$ of 8-HETE oxidation product which arises from C_{10} hydrogen abstraction), being the S/R ratio $> 9{:}1$.[18] In coral 8*R*-LOX hydrogen abstraction is produced uniquely from H_{10}, being the oxidation pattern product formation $C_8{:}C_{12}$ of 97.9:2.1, and the $R{:}S$ ratio of 99:1 in the case of 8-HETE oxidation product.[32]

The general structure of the LOXs binding site cavity is a tubular hydrophobic pocket solvent-accessible from the protein surface. The shape of this cavity varies for the different isoforms; it is assumed that for most LOX isoforms this cavity presents a U-shape, which is the case of 8*R*-LOX,[31,32] but a different shape is found in the case of rabbit 15-LOX-1 (boot shape),[40] and in the case of human 5-LOX the cavity shape is more complex.[29]

The proximity of the bisallylic methylene group to the catalytic iron may greatly depend on the substrate binding cavity depth and width, and also on the orientation of the incoming substrate. Two possible orientations of the substrate are plausible: If the carboxylate end of AA enters first into the pocket, the resulting orientation is denominated head-first orientation, while when is the fatty acid methyl end who enters first into the cavity, the corresponding orientation is called tail-first orientation. In this sense, two possible hypothesis have been proposed to explain the substrate positioning:

1. The *orientation hypothesis*:[47] it is accepted that for a 15-LOX (15-LOX-1 or 15-LOX-2) AA enters the active site with a tail-first type orientation. This configuration allows C_{13} methylene group to be in a good arrangement with respect the catalytic iron to produce reactive structures in this position. Contrarily, if AA would access in the opposite orientation, head first, the C_7 methylene group would be the one to be close enough to the Fe(III)-OH^- cofactor. This could explain then the regioespecificity of 5-LOX.

2. The *spatial hypothesis*:[26,48] given an orientation (tail-first or head-first) of AA inside the LOX active site, it has been proposed that the volume of the substrate binding pocket is the factor tat determines LOX positional specificity. According to this spatial hypothesis then LOX regiospecificity depends on how deep the substrate slides into the binding pocket that, in turn, depends on the active site volume.

The orientation of the substrate inside the cavity is still a matter of discussion for many LOXs, although it is assumed that AA enters in its tail-first orientation in most isoforms: that would be the case of 15-LOXs, as indicated above, and 8*R*-LOX, for which a crystal structure of the

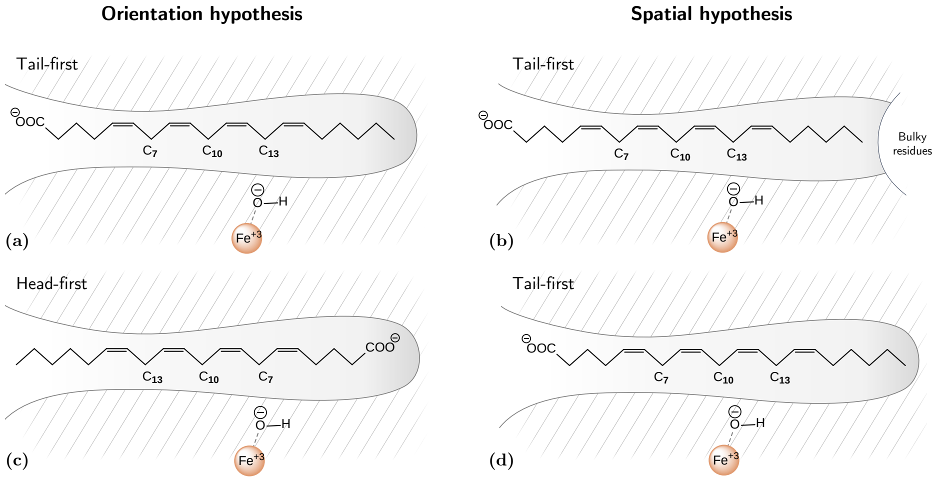


Figure 1.8: Schematic view of AA positioning of inside the active site cavity. For simplicity, the cavity has been represented as a straight tunnel. In a) and c) arachidonic acid enters the cavity with opposite orientations (tail first in a) and head first in c)) allowing different methylenes to be close to the Fe-OH$^-$ moiety (C_{13} in a) and C_7 in b)). In b) and d) the orientation of the incoming substrate is conserved but the structure of the cavity differs: in b) there are bulky residues at the bottom of the cavity which avoid the substrate to slide deeper, being C_{13} the closest methylene to the Fe-OH$^-$ cofactor; in d) the cavity is shallower and the substrate can get deeper, allowing methylene C_{10} to be closer to the catalytic iron. This scheme could not be applied to 8*R*-LOX, in which the cavity is a U-shape structure perpendicular to the 5-LOX and 15-LOX cavities.

AA:enzyme complex is available.[32] Tail-first orientation would allow the enzyme to act over both free and esterified fatty acids, as the esterified moiety is quite bulky to penetrate into the binding pocket. Also, head-first orientation would involve burying a negative charge, the carboxylate group, into a hydrophobic environment.[47,48] In the case of plant LOXs there is a positive Arginine residue at the bottom of the cavity that could favor this head-first orientation. In contrast, in rabbit 15-LOX-1 there is an Arginine (Arg403) but placed over the entry of the active site cavity that could interact with the fatty acid carboxylate end so favoring a tail-first orientation.[18] In the case of coral 8*R*-LOX there is also an Arginine in a similar position (Arg182) that also would permit this orientation.[32] Despite all that, it is possible that a different mechanism, or even both, may be used by an specific LOX isoform to achieve the desired reaction specificity. Published data suggest that for a given isoform, there could be an equilibrium between both orientations, but this steady state could be affected by substrate characteristics or reaction conditions.[37]

The spatial hypothesis has been extensively investigated for 12/15-LOX. In rabbit 15-LOX-1 it has been proposed that a collection of three residues placed at the bottom of the binding cavity (triad concept) is crucial in determining its reaction specificity.[40] This active site cavity bottom is aligned by Phe353, Ile418 and Ile419 in rabbit 15-LOX1. Alterations in the geometry of these residues side-chains lead to a subtle modification in the binding pocket volume, and therefore in the positioning of fatty acid substrates. If these strategic positions are occupied by smaller residues the substrate can slide deeper into the active site, allowing methylene C_{10} to occupy the equivalent position of C_{13} in the wild-type system.

For human 5-LOX the positioning of AA inside its active site is under lively debate. In fact, explaining how AA binds to human 5-LOX is even more challenging than for other LOX isoforms due to the unique configuration of its binding cavity. Not only the orientation, but also the substrate entry, are facts that are not still clear. In 5-LOX the active site cavity is not a simple channel as in other LOXs, and based on the crystallographic structures available two different ways for the substrate entering have been suggested: i) accessing from the residues Phe177 and Tyr181 (the so-called "FY-cork"), or, ii) accessing from Trp147, placed over the opposite end of this FY-cork.[29] Mutagenesis experiments transforming one of these residues into a smaller Alanine residue show an increase of the 5-HETE product formation, but when both residues together are mutated to Ala the concentration of this product slightly decreases with respect to the Stable-5LOX. The authors suggest that Phe177 may play a role in the placement of the α-2 helix, that lines the active site cavity.[49]

Regarding substrate positioning, there are some experimental evidences that could support the spatial hypothesis and a tail-first orientation in human 5-LOX: similarly to previous experiments in 15-LOX-1, Kühn and coworkers have generated site directed-mutant variants of the residues at the bottom of the cavity of human, murine and zebra-fish 5-LOX that transform small residues to bulkier ones and thus decrease the cavity volume.[18, 50, 51] In particular, in these experiments several combinations of mutations of residues F359, A424 and N425 to W359, I424 and M425 were carried out. Those mutations progressively transform a 5-lipoxygenating enzyme into a 15-lipoxygenating one. The explanation given is that reducing the volume cavity those mutants manage to hinder AA in a tail-first orientation for getting deeper into the active site so C_{13}, and not C_7, is closer to the cofactor.

Conversely, Newcomer and coworkers propose a head-first entry according to the orientation hypothesis.[30, 49] There exist some polar/positive resi-dues at the bottom of the cavity of human 5-LOX (His600, Asn425), that could be implicated in the interaction with the fatty acid carboxylate end. A head-first orientation has also been proposed for murine 8*S*-LOX.[52] Mutagenesis experiments in murine 8*S*-LOX suggest that a Histidine residue at the bottom of the active site (His602, which is the counterpart of His600 in human 5-LOX), interacts with the substrate carboxylate end leading to a head-first orientation. Other arguments that support the head-first orientation in 5-LOX propose that this orientation would be the only one capable to produce the 5-HETE product with S-stereochemistry as it is the $n-2$ oxidation product, in contraposition with 15S-HPETE, the 15-LOX product, which is obtained from an AA tail-first orientation and corresponds to the $n+2$ oxidation product. That is, the $n-2$ and $n+2$ oxidation products would present opposite stereochemistry starting from a common orientation (see Figure 1.2).

In the case of coral 8R-lipoxygenase it is quite clear that hydrogen abstraction comes essentially from a unique position, pro-S hydrogen in C_{10}, which leads to the oxygenation product in position 8 with R-stereospecificty. In the AA:8R-LOX complex, as mentioned above, Arg182 seems to harbor the substrate carboxylate end. Mutagenesis experiments altering Arg182 to an Ala have little impact in product specificity, but a dramatic change in the kinetic properties of the enzyme. Mutations at the bottom of the cavity, Ala589Ile and Ala620Ile, affected the specificity of the enzyme, leading to the formation of the 11R-product that arises from H_{13} abstraction.[32]

Other experiments have tried to explain the stereospecificity of the hydroperoxidation product, related with the second step of the reaction mechanism, the addition of the oxygen molecule. For several LOXs of known-reaction specificity, it has been demonstrated that many S-LOXs contain an Alanine residue in a critical position, whereas the counterpart R-LOXs contain a Glycine residue in the equivalent position.[3, 16, 53] Site-directed mutagenesis experiments changing Gly by Ala in human 12R-LOX lead to a major oxidation product with S-stereochemistry. The inverse strategy has been applied to human 15S-LOX-2 (Ala416Gly mutation) yielding 11R-HETE as the main product. Similar alterations have been observed in other species.[16, 54, 55] In rabbit 15-LOX-1, mutagenesis experiments lead to partial alterations in reaction specificity. Main product from Ala404Gly mutant is still 15S-HETE, but in a lesser proportion (75%) than in the wild type system, and the 11R-HETE ratio increases, but only in 20%. In coral 8R-LOX there is a Gly427 in this position, accordingly to its R-stereochemistry. Mutagenesis experiments transforming Gly427 into an Alanine produce almost exclusively the 12-HETE product with 98% of the S-stereoisomer.[16]

1.3 Biological role of lipoxygenases

1.3.1 Metabolites

Mammalian LOXs metabolites consist of lipid mediators of the eicosanoid cascade. This metabolic pathway starts with membrane-lipids hydrolysis by phospholipase-A_2, yielding polyunsaturated fatty acids. Lipoxygenases transform these fatty acids into the corresponding hydroperoxide derivatives which, under physiological conditions, are reduced to hydroxides that are precursors of lipid signaling compounds. Products derived from AA metabolism include leukotrienes and lipoxins (Figure 1.9).

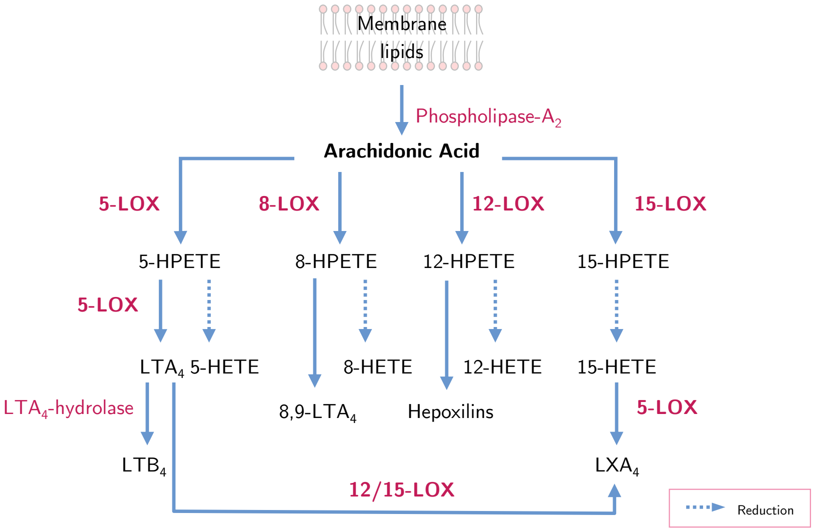


Figure 1.9: Arachidonic acid metabolic pathway in animal LOXs.

Leukotrienes (LT) are a class of signaling compounds promoters of inflammation. The name "leukotriene" was first introduced by Swedish biochemist Bengt Samuelsson in 1979,[56] and derives from "leukocyte", as they were first discovered in these kinds of cells, and "triene", as they present three conjugated double bonds. There are two major classes of LTs, non-peptide leukotrienes (LTA_4 and LTB_4), and peptide-leukotrienes (LTC_4, LTD_4 and LTE_4). In the former class, LTA_4 is produced by 5-LOX and consists of an unstable epoxide intermediate which yields LTB_4 by LTA_4-hydrolase. Leukotriene-B_4 is one of the most potent chemotactic agents known, and it is though as the most potent pro-inflammatory agent. It induces vascular permeability, netrophyl adhesion to the vascular wall and chemotaxis, which are functions involved in the inflammatory response. It also binds to receptors that stimulate inflammatory cells. Peptide-LTs are cysteinyl leukotrienes synthesized from LTA_4 via LTC_4-synthase. These compounds constitute spasmogenic agents and are involved in allergic pathologies.[9,57,58]

Traditionally resolution of inflammation was considered a passive process which was terminated by the decay of its signals with time, but during the past decade it has been found that termination of inflammation is an active process programmed since the starting of the inflammatory response, aimed to restore the normal tissue homeostasis. In this process other lipid mediators participate that are also products of LOXs metabolism. These compounds include lipoxins, resolvins, protectins, and maresins. **Lipoxins** (LXs) are anti-inflammatory and pro-resolving lipid mediators from AA metabolism. These compounds were first described by Samuelsson and coworkers in 1984,[59] and their name derives from "lipoxigenase interaction products", as they found that lipoxin synthesis requires the action of more than one LOX isoform. Its synthesis has been reported to be transcellular (Figure 1.10), i.e., different types of cells need to interact in order to synthetize this class of compounds.[60] The first step of the synthesis is catalyzed by 5-LOX, which is present in leukocytes, but for the second step a 12/15-LOX is required, which can be found for example in platelets. The inverse strategy also leads to the production of lipoxins: when AA is oxidized by a 12/15-LOX 15-HETE is obtained, which is further metabolized into LXA_4 by 5-LOX (see Figure 1.9). For this reason a promising strategy to fight inflammation would be transforming 5-LOX into a 12/15-LOX. Resolvins,[61] protectins,[62] and maresins[63] derived from ω-3 fatty acid metabolism (DHA and EPA) and also have protective and resolution functions.

Plant LOXs metabolites include jasmonic acids and green leaf volatiles.[11–13] Jasmonic acids, such as (+)-7-isojasmonic acid, are involved in programmed cellular death at wound sites of the plant, regulation of growth and development. These compounds are used as natural pesticides because they stimulate anti-pest defenses of plants without inhibiting growth.[64] Green leaf volatiles participate in ecological functions such as plant communication, defense against insects, reactive oxygen species removal, thermo-tolerance and environmental stress adaptation.[65] In the case of bacteria, LOXs principal metabolites are lactones, which have recently focus interest because its applications in the food industry as natural flavors.[10,15]

1.3.2 Pathological implications

Mammalian lipoxygenases display their biological function mainly via three mechanisms of action: synthesis of lipid mediators, modification of complex lipid-protein assemblies and modification of cellular redox state. The first of these mechanisms is related with the most well-studied role of lipoxygenases, as lipid mediators are related with inflammatory response and some related disorders and diseases.

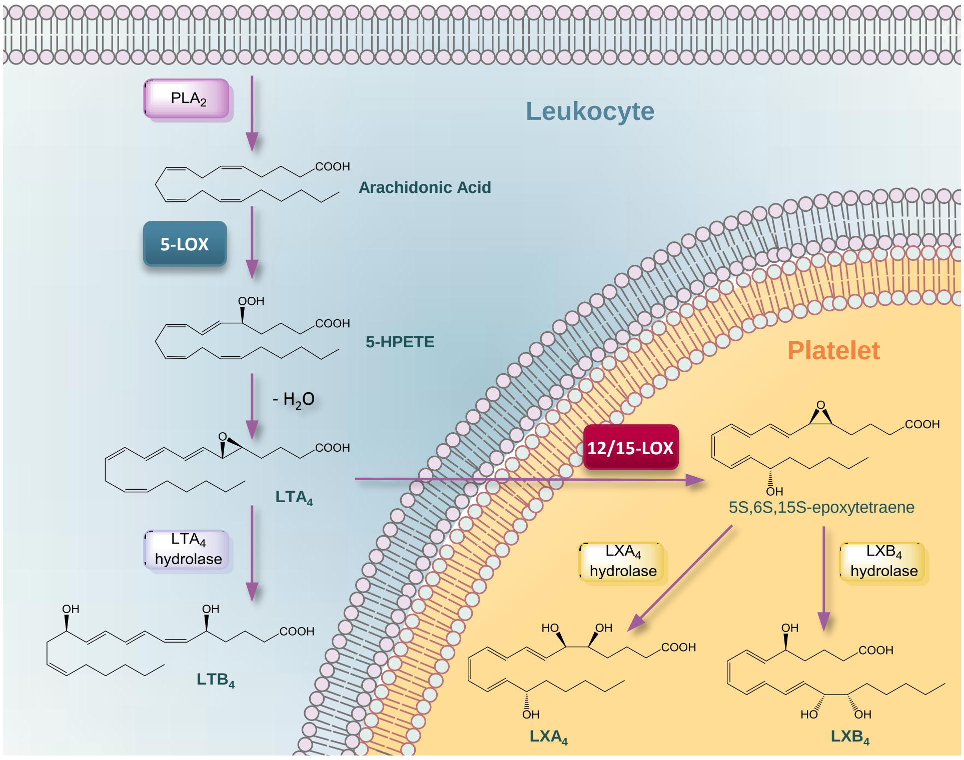


Figure 1.10: Schematic representation of transcellular synthesis of lipoxins. 5LOX, present in a leukocyte (left, in blue), initiates the synthesis of 5-HPETE from arachidonic acid, which is hydrolized into leukotrierne A4. LTA$_4$ can be further converted into LTB$_4$ by LTA$_4$-hydrolase into the leukocyte, or either can be transferred to a platelet (right, in yellow), that contains a 12/15-LOX, which transforms it into the 5*S*,6*S*,15*S*-epoxytetraene intermediate. This intermediate is further metabolized into lipoxin-A$_4$ and lipoxin-B$_4$ by LXA$_4$-hydrolase and LXB$_4$-hydrolase, respectively.

LOXs play a role in **cardiovascular diseases**. The activity of LOXs, particularly 15-LOX-1, is related with **atherosclerosis**.[66,67] This enzyme participates in the oxidation of LDL (Low Density Lipoproteins). Oxidized LDL presents pro-atherogenic properties because they trigger immune response and are rapidly phagocytosed by smooth muscle macrophages that are thus transformed into foam cells. These cells accumulate in the intravenous (subendothelial arterial) space producing fatty streaks, that are early atherosclerotic lesions.15-LOX-1 is not only capable of oxidizing LDL, but also HDL (High Density Lipoproteins), considering this effect also as pro-atherogenic. The role of 5-LOX has been also related with atherosclerosis, participating in the late stages of the vascular lesion by the synthesis of LTB$_4$.[68] 15-LOX-2 also participates in atherosclerosis, as it has been found expressed at high levels in late atherosclerotic lesions.[69] LOXs are also implicated in blood pressure regulation. 15-LOX-1 has also been related with blood pressure regulation and hypertension, as it has been involved in the regulation of the vascular tone.[70]

Many aspects of **carcinogenesis** such as tumor cell proliferation, metastasis, differentiation, apoptosis, as well as migration, invasion of carcinoma cells and angiogenesis, have been related with LOXs activities.[71] Unresolved chronic inflammation is a key process in tumor progression, and therefore, pro-resolving eicosanoids such as lipoxins are believed to exhibit anti-cancer activities.[72] Several investigations suggest a role of LOXs in colorectal,[73] breast[74] and prostate[75] cancer. Arachidonic acid metabolites 5-HETE and 12-HETE have been identified as pro-carcinogenic agents, as they induce cell proliferation, cell adhesion, metastasis and angiogenesis.[76,77] However, LA metabolite 13-HODE has turned out to be anti-carcinogenic, as it promotes apoptosis.[78] Expression of the different isoforms depends on the tumor. For instance, 5-LOX is expressed in the early stages of colorectal cancer, whereas 12-LOX has been identified in the late events such as metastasis and neoangiogenesis. The role of 15-LOX in colorectal cancer is still a matter of discussion.[58]

The activity of 5-LOX has been related with allergic inflammation, and plays an important role in the control of asthma.[79] Under asthmatic conditions activated immune cells produce a mixture of leukotrienes, which are the components of the slow-reacting substance of anaphylaxis. This mixture is composed mainly by peptide-LTs, which are more effective than histamine as bronchoconstrictors. Other allergic respiratory diseases such as allergic rhinitis[80] and bronchial hyperreactivity[81] involve the presence of leukotrienes. LOXs also play a role in some Central Nervous System and neurodegenerative diseases. Neurons are very sensitive to oxidative-damage, and LOXs are involved in oxidative-stress pathologies. In Alzheimer's both 5-LOX and 15-LOX have been implicated, but their precise roles are not clear; augmented levels of 12- and 15-HETE in cerebrospinal fluids of Alzheimer patients reveals a role of these isoforms.[82,83] Metabolic disorders have been also related with the action of LOXs isoforms: Type 1 and Type 2 diabetes, obesity and nonalcoholic fatty liver disease.[58]

1.3.3 Inhibitors

Because the involvement of LOXs in human diseases the design of LOXs inhibitors has become of major pharmacological interest. According to the mechanism of inhibition, LOXs inhibitors can be classified into five main categories: i) redox active inhibitors, ii) non-redox iron chelators, iii) substrate competitive inhibitors, iv) suicidal inhibitors and v) allosteric inhibitors.

Redox-active inhibitors include reducing agents that alter the oxidation state of the catalytic iron, turning its active ferric state into inactive ferrous state. Into this class there are many lipophilic compounds both natural plant-derived (e.g., caffeic acid, flavonoids, coumarins and several polyphenols) and synthetic compounds.[84,85] However, redox inhibitors present an important intrinsic problem as they interfere with other cellular redox reactions (e.g. the formation of methemoglobin[86]) and lack of selectivity. In general, non-redox inhibitors avoid toxicity associated with redox side-reactions, and therefore, the search of such class of inhibitors is a matter of interest.

Iron ligand inhibitors include compounds that act as iron-chelators without altering its oxidation state. This classification comprises hydroxamic acids or N-hydroxyurea derivatives, functional groups that are good iron-chelating agents and also possesses weak reducing properties. Inhibition is achieved by replacing one of the Fe-coordination sphere ligands (probably the hydroxyl/water molecule) by the inhibitor molecule itself. Chemical compounds with hydroxamic acid or N-hydroxyurea functional groups such as Zileuton,[87] (Figure 1.11) are good iron chelators. Zileuton is a 5-LOX inhibitor clinically approved as a treatment for asthma; in patients treated with this drug an improvement of the airway function and reduction of the inflammation of the respiratory system has been shown. However, Zileuton is not the first choice as it implicates side effects such as nausea and hepatic toxicity. Furthermore, only low therapeutic benefits have been reported for other inflammatory diseases such as rheumatoid arthritis and allergic rhinitis. As inhibitor Zileuton also presents disadvantages, as it has low potency and short half-life due to rapid metabolic breakdown. Despite, to date Zileuton is the only inhibitor that has been made to the clinics.

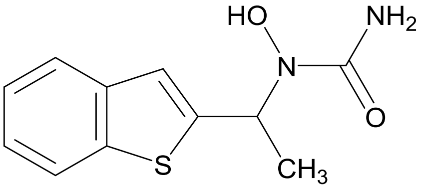


Figure 1.11: Chemical structure of Zileuton

A third class of inhibitors include those compounds which does not interact with the catalytic iron, but possess inhibitory activity as substrate competitors for protein binding. This class comprises structurally different molecules. Competitive inhibitors can also be used for crystallizing a LOX holoenzyme with a substrate mimic, as is the case of rabbit 15-LOX-1 crystallized with RS7.[27] Classification of LOXs inhibitors also includes suicidal inhibitors, which inactivate LOXs by irreversible binding, and allosteric inhibitors, that interact with a different site than the active site binding cavity.

The most pharmacological relevant isoform is 5-LOX. The strategies for acting against this system can be direct inhibition of the 5-LOX, and inhibition of the 5-LOX activating protein (FLAP),[84] a protein necessary for 5-LOX activation. Other promising strategy is transforming a 5-LOX into a 15-LOX system, as it will turn its inflammatory activity into an anti-inflammatory one by the synthesis of lipoxins,[30,50] as explained in the previous section. Since there are six functional LOXs isoforms in humans the design of an specific inhibitor for a given LOX is a challenge. Thus, the precise knowledge of the mechanism of interaction between a substrate and an specific isoform, particularly the knowledge of its regio- and stereospecificity, is fundamental prior to the design of a proper inhibitor with desirable properties.

Part II

Objectives

Chapter 2

Objectives

The goal of the Thesis presented here is to get a deep insight on the reaction mechanism of lipoxygenases (LOXs), putting a special focus on the subtle factors that govern their strict reaction specificity, especially in the reaction of the arachidonic acid substrate, the main substrate of mammalian (and thus human) LOXs.

The knowledge of the reaction specificity is crucial in the design of strategies to alter the function of these enzymes, in order to turn their normal inflammatory activity, involved in many diseases, into an anti-inflammatory one. A strategy is the use of inhibitors, but the design of such type of compounds requires an exhaustive knowledge of the reactivity and mechanism of their target receptor. An important characteristic of LOXs is that, starting from a common substrate (arachidonic acid, which is the central root of the eicosanoid metabolic pathway) each LOX-isoform would lead to a different product with different (even opposite) biological functions. How the same members of the family can produce such different reactions (and thus products) is an active discussion in the literature, and the work presented here is aimed to shed light to this point from a fundamental basis. A precise knowledge of the factors controlling the reaction specificity with the natural AA substrate is a necessary step prior to the rational design of specific inhibitors directed against each specific LOX isoform.

To this aim, three different members of the LOX family, mammalian 15-LOX-1, human 5-LOX, and coral 8*R*-LOX, have been the object of the study in this Thesis. The theoretical treatment of such complex systems requires the use of a set of different computational methods. In this sense, homology modeling and docking simulations will be employed to generate the different binding modes of the AA:LOX complex, and full atomistic MD simulations will be carried out to generate the different conformations of the fully solvated AA:LOX Michaelis complex and to analyze their stability.

Subsequent QM/MM calculations will be performed to study the reaction mechanism, with a treatment for the QM part at the DFT level to obtain a reliable description of the evolution of the electronic structure. A special focus will be given on the first reaction step, the hydrogen abstraction, which is the rate-limiting step of the whole reaction mechanism, and also determines the regio- and stereospecificity of the final hydroperoxydation product, as this depends on which particular hydrogen atom is abstracted. However, also the second step of the catalytic mechanism, the oxygen addition process, and how it accounts for the regio- and stereospecificity of the final products will be considered. Finally, different *in silico* mutants will also be studied to analyze the influence of the mutated residues in the reaction specificity.

In brief, the main objective of the theoretical work presented in this Thesis is to reveal the molecular origin of the regio- and stereospecificity of lipoxygenases withe the aim to rationalize the *in vitro* mutagenesis experiments carried out in these systems and also trying to predict the specificity of new mutants. As mentioned above, this is a first step in our contribution as theoretical chemists to the rational design of specific LOX inhibitors.

Part III

Theoretical Background

Chapter 3

Theoretical background

3.1 Preamble

During the last decades computational chemistry has been applied to understand chemical processes in biological systems. The constant increase in computer power has led to more accurate and rapid results obtained in systems of increasing size and complexity. Computational chemistry has become a powerful tool to understand enzyme catalysis, as it provides detailed molecular insights of the reaction mechanisms, and it can be used to make predictions. This highly-detailed point of view of the molecular basis of chemical processes that computational methods have shown to give, is not experimentally achievable today. This chapter provides a general overview of the foundations of the methodology employed through this Thesis. It first focuses on the expression of the potential energy function at different levels of theory (MM, QM and hybrid QM/MM), and the methods for calculating the free energy barriers of a reactive process. Next, Molecular Dynamics theory to study the evolution of the systems over time while exploring of the conformational space will be adressed. Finally, modeling techniques of homology modeling employed for the generation of the protein 3D structure, and docking simulations to obtain the substrate:protein Michaelis complex will be explained.

3.2 Molecular Mechanics

Many systems are extremely complex to be treated by quantum mechanics. Such systems are typically large, containing from thousand to millions of atoms, and still unaffordable for a complete QM description even with the computational power available today. In such cases a simpler description of the energy function but with some accuracy is required. Molecular mechanics (MM) uses classical mechanics to describe molecular systems. Each atom is treated as a classical particle with a determined mass, net charge, and radius (i.e. van der Waals radius). Thus, electrons and nuclei are not explicitly considered and chemical events, such as bond breaking and bond forming cannot be described in this way. The potential energy function constitutes the so-called force field. The description seems very simple, but the molecular mechanics energy function (the force field) must describe accurately the molecular events.

3.2.1 The Force Field energy expression

The classical potential energy function (the force field) has typically an expression like the following:

$$\begin{aligned} E_{\mathrm{MM}} = & \sum_{\text{bonds}} k_d (d - d_0)^2 + \sum_{\text{angles}} k_\theta (\theta - \theta_0)^2 + \sum_{\text{torsions}} k_\phi [1 + cos(n\phi + \delta)] \\ & + \sum_{\substack{\text{non-bonded} \\ \text{AB pairs}}} \left\{ 4\varepsilon_{\mathrm{AB}} \left[\left(\frac{\sigma_{\mathrm{AB}}}{r_{\mathrm{AB}}} \right)^{12} - \left(\frac{\sigma_{\mathrm{AB}}}{r_{\mathrm{AB}}} \right)^{6} \right] + \frac{1}{4\pi\varepsilon_0} \frac{q_{\mathrm{A}} q_{\mathrm{B}}}{r_{\mathrm{AB}}} \right\} \end{aligned} \tag{3.1}$$

In this expression there are two types of contributions, *bonded* and *non-bonded* terms. Bonded terms apply to a set of atoms connected trough a bond (1-2 terms), separated by two bonds (1-3 terms), and by three bonds (1-4 terms).

Typically, the bond (1-2 term) is treated as a spring with a reference distance equal to the experimental or calculated bond length, and usually the harmonic approximation is employed (Hooke's law) or even a Morse potential. In equation 3.1 d is the bond length, and d_0 is the reference distance. Near the reference distance the harmonic approximation works well, but for larger distortions of bonds the energy term tends to the infinite, so anharmonicity needs to be included.

For the treatment of angle bending (1-3 terms), the deviation of the current angle θ from the reference value θ_0 can be also treated with the Hooke's law. The torsional term implies the rotation of dihedral angles (1-4 terms), and are usually expressed as a cosine series expansion; ϕ is the dihedral angle, n is the multiplicity (minimum points in the function as ϕ rotates over 360°), and δ is the phase factor, which determines where the dihedral has its minimum value. A force field also accounts for the improper torsions or out-of-plane bending motions, which tend to restrain a set of atoms that form a π-system into the same plane.

Non-bonded terms accounts for the interaction among independent set of non-connected atoms. These interactions typically depend on an inverse power of the distance of the atoms involved. Non-bonded terms usually are electrostatic interactions and van der Waals interactions. In expression 3.1, van der Waals interactions are described by a Lennard-Jones 12-6 potential where r_{AB} is the distance between the interacting atoms, $\varepsilon_{\mathrm{AB}}$ is the potential well depth, and σ_{AB} is the finite distance where the non-bonded interaction between the uncharged particles is zero. This type of function is widely used as the r^{-12} repulsive contribution can be rapidly calculated by squaring the r^{-6} attractive contribution. Electrostatic interaction can be modeled by Coulomb's law, where q_{A} and q_{B} are the partial atomic charges, and ε_0 is the dielectric constant.

In equation 3.1 there also appear the terms which constitute the set of parameters (the force constants k_d, k_θ, k_ϕ, the atom charges, and the van der Waal radii), which are particular for each force field. The parametrization of a force field is probably the most difficult task. Parameters need to reproduce molecular properties in the correct environment, as for a given atom, the behavior not only depends on the chemical element itself, but also on the type of chemical bonding (i.e. the parameters for sp^2 and sp^3 aliphatic carbons must be different). A force field important feature that must fulfill is transferability, which implies that a given set of parameters can be used to model a series of related molecules instead of defining a new set for every molecule. This characteristic is very important when the force field is used to make predictions.

The exact analytical expression depends on the particular force field. Force fields parameters can be empirical, or derived form previous QM calculations of model molecules, and the functional form of the force field is always a compromise between accuracy and computational efficiency. Generally, force fields have internal features well developed for proteins and other biomolecules, and sometimes the parametrization of ligands is the most tricky part. The application to proteins is well developed since they are the focus of most problems. Other biomolecules of interest such as DNA, lipids ans sugars have specifically-derived set of parameters, implemented in most common force fields. Force fields widely used for biomolecular applications are CHARMM,[88] AMBER[89] and GROMOS.[90] All of them have been demonstrated to provide similar accuracy and good description of the molecular properties, and are especially derived for biomolecular polymers.

The CHARMM[88] force field has been used for the MM calculations in this Thesis. Additionally to the general expression of the MM energy (equation 3.1), CHARMM force field also incorporates some non-standard energy terms. The first one is the so-called Urey-Bradley term, which is a cross-term that accounts for angle bending interactions using a harmonic function of the distance between the 1-3 atoms. The other non-standard term is the CMAP, which is a cross-term of the protein backbone dihedral angles (ϕ and ψ), that improves conformational properties of protein backbones.

3.3 Quantum Mechanics

Quantum Mechanics (QM) treats the electronic structure of atoms explicitly, therefore is the tool to study chemical properties of a system such as reactivity. There are mainly two approaches in quantum mechanics, wave-function based methods, and electronic-density based methods. The former will be explained briefly, as the work developed here is based on the latter class of quantum mechanical methods.

3.3.1 Wavefunction methods

The starting point of quantum mechanics consist on solving the time-independent and non-relativistic SchrÃűdinger equation.

$$\hat{H}\Psi = E\Psi \tag{3.2}$$

The solution of the SchrÃűdinger equation for a system of N electrons is not exact. The first approach to this is the Hartree-Fock (H-F) approximation, which provides a simple description of electrons occupying orbitals. In Hartree-Fock approach, the wavefunction is expressed as a Slater determinant, which is the antisymmetric product of spin-orbitals:

$$\Psi(1,2,...,N) = (N!)^{-\frac{1}{2}} \sum_{q=1}^{N!} (-1)^{p_q} \hat{P}_q \{\chi_i(1)\chi_j(2)...\chi_k(N)\} \tag{3.3}$$

where N is the number of electrons, and χ_i are the spin-orbitals. $\hat{P}_q$ is the permutation operator, and p_q the number of transpositions required to obtain the permutation. This is a first approach, since H-F does not consider the electron correlation, and is insufficient for describing systems with unpaired electrons or breaking bonds. The development of *post*-Hartree-Fock methods permitted the inclusion of electron correlation. Such methods are Configuration Interaction, Coupled Cluster, and Perturbation Theory methods. These type of methods are computationally expensive for systems with a certain number of atoms.

3.3.2 Density Functional Theory

Density Functional Theory (DFT) is widely used in biomolecular applications due to its favorable computational effort-accuracy ratio. DFT bases on the idea that the energy of an electronic system can be defined in terms of its electron density, ρ so that for a N electron system $\rho(\vec{r})$ represents the electron density on a point, $\vec{r}$. Thus, the total energy E is considered as a functional of the electron density $E[\rho]$, so that for any ρ it corresponds a unique $E[\rho]$ value.

The main advantage of DFT methods against wavefunction methods is that, for a N electron system, the wavefunction depends on $4N$ coordinates, three spatial and one spin coordinates, (eq. 3.3) whereas the electron density only depends on three coordinates, no matter the number of electrons of the system. Thus, DFT methods are widely accepted for large complex systems.

Density Functional Theory is based on two theorems developed by Hohenberg and Kohn in 1964.[91] The first theorem establishes:

Theorem 3.3.1 (First Hohengerg-Kohn Theorem) *Any observable of a non-degenerate ground electronic state can be calculated, exactly in theory, from the electron density of the ground state. In other words, any observable can be written as a functional of the electron density of the ground state.*

Therefore a correspondence between the electron density and the total electronic energy is establish, which is given by:

$$E[\rho] = T[\rho] + V_{\mathrm{Ne}}[\rho] + V_{\mathrm{ee}}[\rho](+V_{\mathrm{NN}}) \tag{3.4}$$

where $T[\rho]$ corresponds to the electronic kinetic energy, $V_{\mathrm{Ne}}[\rho]$ is the electron-nucleus interaction potential energy, $V_{\mathrm{ee}}[\rho]$ is the electron-electron interaction potential energy and V_{NN} is the nucleus-nucleus potential interaction energy (constant for a given nuclear configuration and can be omitted). Second Hohenberg-Kohn's theorem sets that:

Theorem 3.3.2 (Second Hohengerg-Kohn Theorem) *The electron density of a non-degenerate electronic ground state can be calculated, exactly in principle, determining the density that minimizes the energy of the ground state.*

That is, for a trial density $\tilde{\rho}(\vec{r})$ it is fulfilled:

$$E_0 \leq E_{\mathrm{v}}[\tilde{\rho}(\vec{r})] \tag{3.5}$$

However, the exact form of neither $T[\rho]$ nor $V_{\mathrm{ee}}[\rho]$ is known. In 1965 Kohn and Sham[92] proposed a method to calculate the kinetic energy from ρ. This method is based on a fictional system of N non-interacting electrons described by monoelectronic wavefunctions, such that it presents the same electron density as an interacting electrons system. Thus, the energy functional is re-defined as:

$$E_{\mathrm{v}}[\rho] = T_{\mathrm{s}}[\rho] + \int v_{\mathrm{n}}(\vec{r})\rho(\vec{r})d\vec{r} + J[\rho] + E_{\mathrm{xc}}[\rho] \tag{3.6}$$

where $T_{\mathrm{s}}[\rho]$ is the non-interacting electron kinetic energy, $v_{\mathrm{n}}(\vec{r})$ is the external potential, and $J[\rho]$ is the Coulomb energy. The main problem of DFT theory comes from the term $E_{\mathrm{xc}}[\rho]$, the so-called correlation-exchange energy, for which the exact expression is unknown. This functional contains the exchange contribution of the electrons, associated with the interaction of electrons with the same spin, and the electronic correlation contribution, which is associated with opposite spin electrons. Different approaches calculate this term (or the correlation-exchange by particle, $\varepsilon_{\mathrm{xc}}$) leading to the different DFT known methods. From less to more accurate, these approximations are:

- *Local Density Approximation* (LDA): $\varepsilon_{\mathrm{xc}}$ functional only depends on the electronic density. Correlation and exchange contributions are treated separately.

- *Local Spin Density Approximation* (LSDA): used for open-shell systems, α and β spin contributions are treated separately.

- *Non-local Corrections or Generalized Gradient Approximations* (GGA): this approach introduces the electronic density gradients, $\vec{\nabla}\rho$, in the description of the correlation and exchange effects. Therefore, it not only takes into account the density in every point, but also the density variation around each point.

- *Meta-GGA Approximation*: besides the terms included in GGA approximation, meta-GGA also includes the kinetic energy density, $\tau(\vec{r})$.

On the other hand, by *Hybrid Methods* the exchange term is calculated as a suitable combination of exact H-F exchange and the density functional exchange contributions.

3.4 QM/MM Methods

Quantum Mechanics/Molecular Mechanics (QM/MM) methodology constitutes the state of the art for the study of chemical processes in complex systems, such as enzyme catalysis.[93] QM/MM era begins in 1972 with Warshel and Karplus paper,[94] who introduced the QM/MM concept and presented a methodology with some of the features that today are essential. The first application to an enzyme reaction was due to Warshel and Levitt in 1976.[95]

However, this approach did not begin to be widely used until 1990, when Field, Bash and Karplus described in detail a combination of semiempirical QM methods with the CHARMM force field and carefully evaluated the precision and accuracy of the QM/MM treatment in comparison with *ab initio* and experimental data.[96] During the past decade several works have recorded the development of QM/MM methods. Nowadays the QM/MM methodology is a valuable tool not only for biomolecular systems,[97] but also for inorganic/organometalic systems, the solid phase, and explicit solvent.

3.4.1 Overview

In a QM/MM treatment, the system ($\mathbb{S}$) is divided into two regions, the inner region ($\mathbb{I}$) treated quantum-mechanically, and the outer region ($\mathbb{O}$) described by a force field. The inner region is referred as the QM region, and the outer region as the MM region. The system total energy cannot simply be described as the summation of the subsystem energies as there exists (strong) QM-MM interactions, therefore, coupling terms need to be considered. Special cautions are required in the QM-MM boundary region. During a QM/MM calculation, the QM/MM partition (Figure 3.1) can either be fixed or variable.

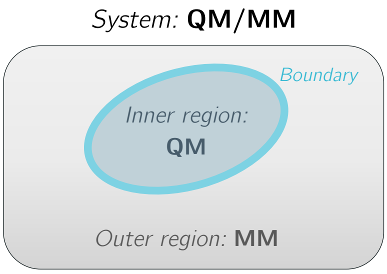


Figure 3.1: Partition of the complete system in outer and inner subsystems. Boundary is represented by a blue thick line.

3.4.2 QM/MM energy expression

QM/MM schemes can be classified according to the total energy expression. This classification leads to two different approaches, additive and subtractive, which are the opposite schemes.

A *subtractive scheme* considers a full MM calculation in the whole system (i), and QM (ii) and MM (iii) calculations in the inner subsystem, so the final total QM/MM energy is obtained summing (i) and (ii) and subtracting (iii) to avoid multiple counting:

$$E^{\text{sub}}_{\text{QM/MM}}(\mathbb{S}) = E_{\text{MM}}(\mathbb{S}) + E_{\text{QM}}(\mathbb{I}+\mathbb{L}) - E_{\text{MM}}(\mathbb{I}+\mathbb{L}) \tag{3.7}$$

Subscripts indicate the type of calculation, and parenthesis the system in which is performed. As is expressed, equation 3.7 is fulfilled for a link-atoms ($\mathbb{L}$) scheme, in which calculations are carried out in an inner capped system ($\mathbb{I}+\mathbb{L}$).

This scheme can be thought as a complete MM treatment where a certain region of space has been cut out to be treated with QM method. Its main advantage resides in its simplicity, as it does not require coupling terms and QM and MM standard procedures can be directly applied. It also corrects link atoms artifacts as long as the MM force terms in which link atoms are involved reproduce reasonably the QM potential. However, it presents some disadvantages: MM parameters are required for treating the inner subsystem as MM, and QM-MM coupling is managed at MM level. The latter causes problems specially with electrostatic interactions.

The opposite approach are *additive schemes*, which presents the following energy expression:

$$E^{\text{add}}_{\text{QM/MM}}(\mathbb{S}) = E_{\text{MM}}(\mathbb{O}) + E_{\text{QM}}(\mathbb{I}+\mathbb{L}) + E_{\text{QM-MM}}(\mathbb{I},\mathbb{O}) \tag{3.8}$$

Unlike the subtractive scheme, in an additive scheme the MM calculation is only performed in the outer subsystem. Moreover, an explicit coupling term $E_{\text{QM-MM}}(\mathbb{I}\ ,\mathbb{O})$ gathers the subsystems interaction terms. Inner caped subsystem $\mathbb{I}+\mathbb{L}$ is treated at the QM level as in a subtractive scheme. Currently most schemes use the additive approach.

The exact expression of the QM-MM coupling term leads to a particular QM/MM method. This term includes the force field interactions (bonded, electrostatic and van der Waals) among QM and MM atoms:

$$E_{\text{QM-MM}}(\mathbb{I},\mathbb{O}) = E^{\text{bond}}_{\text{QM/MM}} + E^{\text{vdW}}_{\text{QM/MM}} + E^{\text{elect}}_{\text{QM/MM}} \tag{3.9}$$

Coupling term contributions are detailed below.

3.4.2.1 Electrostatic QM-MM interaction

Electrostatic coupling between QM charge density and MM region charge model can be managed at different complexity levels. This coupling or *embedding* is characterized by the mutual polarization extent that both subsystems reciprocally exert. According to this criterion it is classified in mechanical, electrostatic, and polarized embedding.[98,99]

In the *mechanical embedding* scheme electrostatic QM-MM interactions are treated at the same level as electrostatic MM-MM interactions, simply applying the point-charge model of the MM method (normally rigid point-charges, but dipoles can be used as well) to the QM region. Conceptually this method is very simple and computationally efficient, but important disadvantages are inherent to it: QM density is not affected (polarized) by its electrostatic environment, problems associated with the potential energy surface when the QM charge distribution changes (as it occurs in a chemical reaction), or problems in the derivation of proper MM charges.

These difficulties are overcome by an *electrostatic or electronic embedding* scheme, in which QM calculations are carried out in the presence of the MM charges, e.g., incorporating these point-charges as monoelectronic terms in the QM Hamiltonian:

$$\hat{H}^{\text{el}}_{\text{QM-MM}} = -\sum_{i}^{N}\sum_{J\in\mathbb{O}}^{L}\frac{q_J}{|\mathbf{r}_i - \mathbf{R}_J|} + \sum_{\alpha\in\mathbb{I}+\mathbb{L}}^{M}\sum_{J\in\mathbb{O}}^{L}\frac{q_J Q_\alpha}{|\mathbf{R}_\alpha - \mathbf{R}_J|} \tag{3.10}$$

where q_J are the punctual MM charges located in $\mathbf{R}_J$, Q_α are the QM nuclear charges in $\mathbf{R}_\alpha$, and $\mathbf{r}_i$ designate the electrons positions. Indexes i, J, and α expands to the N electrons, L punctual charges, and M QM nuclei, respectively.

In an electrostatic embedding scheme the electronic structure of the inner region can be adapted to the changes produced in its environment and it is automatically polarized by it. In this case electrostatic QM-MM interaction is treated at the QM level, ensuring a more advanced and accurate description but requiring a higher computational demand. The QM-MM boundary requires especial care when point charges are close to the QM electronic density as it may cause overpolarization. This problem becomes important when the boundary crosses covalent bonds. The electrostatic embedding scheme is the most used for biomolecular applications.

Polarized embedding schemes include a flexible MM charges model which is indeed polarized by QM charge distribution. Into this scheme there are two possible approaches:

- a polarizable charge model in the MM region which is polarized by the QM electric field but does not act over the QM density
- a self-consistent formulation including the MM polarizable model into the QM Hamiltonian, allowing mutual polarization.

An accurate description of the electrostatic forces derived from the QM subsystem is essential for realistic modeling of biomolecular systems. Computationally speaking, it is difficult to include all the electrostatic contributions; however, simple electrostatic QM/MM cuts cause problems due to the long range nature of Coulombic interactions.[100,101] Trustful and efficient treatment of such long-range interactions is well establish in classical Molecular Dynamics (MD) but only recently has won adepts in the QM/MM world.

3.4.2.2 Other non-bonded and bonded QM-MM interactions

Coupling QM-MM term also contains van der Waals and bonded contributions whose treatment is simpler as are managed at pure MM level (eq. 3.1) regardless the QM/MM (additive or subtractive) scheme. Other bonded interactions (bond stretching, angle bending, torsion, etc.) can be described by a standard parameters set as well, which could be also completed with additional parameters.

3.4.3 The QM-MM boundary

3.4.3.1 General description of boundary schemes

Figure 3.2 shows an schematic representation of a covalent bond cut by the QM-MM boundary:

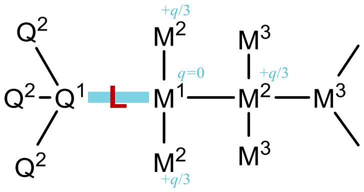


Figure 3.2: Representation of the atoms present in the QM-MM boundary (represented by a blue thick line) treated with a link atoms (red) scheme. Partial charges show a charge-shift model scheme: original MM charge q is removed from M^1 and eventually distributed between M^2 atoms. Additional charge pairs are placed in the proximity of M^3 atoms to restore the original M^1-M^2 dipoles (not shown).

Directly linked QM and MM atoms are designed as Q^1 and M^1, respectively, and constitute the so-called boundary or limit atoms. First MM layer, that is, the layer directly bound to M^1 is labeled as M^2, second MM layer, separated from M^1 by two atoms is labeled as M^3, and so on. The same labeling is applied to the QM atoms, designed by Q.

Simpler solution of defining QM-MM boundary consists obviously in defining QM and MM subsystems in such a way the boundary does not cross covalent bonds. However, in most cases cutting covalent bonds by the QM-MM boundary is unavoidable. Three considerations must be taken into account in such cases: i) the dangling bond from QM atom Q^1 must be saturated as QM region cannot be truncated (i.e. treating the bond as in an homolitic or heterolitic breakage); ii) for electronic or polarized embedding overpolarization caused by MM charges over QM density must be prevented, especially if link atoms are used; iii) bonded MM terms implicating both subsystem atoms must be chosen in order to avoid double counting of interactions. The general purpose seeks a well balanced description of QM-MM interactions in subsystems limit. Main boundary schemes are:

- *Link atoms scheme*, which introduce an additional atomic center L (normally a hydrogen atom) that does not belong to the real system. Link atoms are covalently bonded to Q^1 and saturate its free valence.
- *Localized orbitals schemes*, which place hybrid orbitals over the boundary keeping some of them frozen. Orbitals saturate QM region replacing the capped bond. Within these methods are included Local Self-Consistent Field (LSCF), Frozen Orbital method, Generalized Hybrid Orbital (GHO) and effective fragment potential (EFP).

QM-MM boundary must be placed as distant as possible of the reactive region, although this increases the computational cost, as it defines the QM size. On the other hand, QM atoms involved in bond breakage and formation should be implicated in no bonded coupling scheme. As dihedral terms extend to more than two inner region bonds (depending on the boundary scheme) such atoms should be placed at least three bonds from the frontier. The cut bond should not be polar and not be involved in conjugated interactions. A good choice are C-C aliphatic bonds, whereas amide partial double bond is less desirable.

3.4.3.2 Link atoms scheme

By a link atoms scheme[96, 102] the free valence originated in QM/MM boundary is saturated by an additional atom covalently bonded to Q^1 (see Figure 3.2), which is a hydrogen atom in most cases. QM calculations are then carried out in an electronically saturated system consisting of the sum of the inner system plus the link atoms, $\mathbb{I} + \mathbb{L}$. Q^1-M^1 bond is described at the MM level. The use of link atoms scheme is widely accepted mainly due to its simplicity and easy implementation, but there are some problems associated with the introduction of an additional atomic center: the introduction of additional degrees of freedom associated with each link atom, overpolarization of the QM density in electronic/polarized embedding schemes as link atoms are spatially close to the MM boundary atoms (typically at 0.5 Å), and finally, link atoms are chemical- and electronically different of the group that replace.

Despite these problems link atoms scheme is widely used and several type of schemes in this way have been developed.[97] Additional degrees of freedom can be removed by the use of restraints,[103] modeling link atom position as a function of Q^1 and M^1 positions; link atom L is placed along Q^1-M^1 bond, and Q^1-L is related with Q^1-M^1 by a scaling factor. On the other hand, overpolarization is always present in some extent when a single point charge interacts with a polarizable charge distribution, and gets stronger the closer the QM density point charge and the more flexible density. This problem is less severe when small basis sets centered on the atoms are employed in QM calculation. In the context of link atoms scheme different methods are proposed to reduce overpolarization:

- Removing monoelectronic integrals associated with link atoms.
- Removing MM point charges in the bond region from Hamiltonian.
- Redistribution/displacement of the point charges in the bond region (Figure 3.2). The schemes that preserve charge and often dipole in the boundary region, which remove the overpolarization caused by M^1, overcome the main deficiencies of those schemes that remove the charge. One widely used is the charge-shift scheme introduced by Sherwood and co-workers.[104,105]
- “Blur" charges close to the QM region replacing them by charge distribution (i.e. Gaussian functions).

Hybrid orbital approaches[106–109] are more fundamental from a theoretical point of view as they provide a QM level description of the boundaries. However, they are technically more complex, due at least to orthogonality limitations required to prevent frozen and active orbital mixture in the SCF treatment. Moreover, localized orbitals themselves and/or the specific parameter set related must be previously determined by calculations on a model system. These parameters are not transferable in general and may need modification when changing MM force field, QM method, or basis sets. In general link atoms scheme is employed in the same extent as localized orbitals. Both techniques offer reasonable accuracy when they are carefully used.

3.5 Free Energy Calculations

Free energy is the measure of the driving force of a chemical reaction. As bonds are broken and formed in a chemical process, QM theory is required for the calculation of free energy barriers. Calculation over the potential energy surface provide the internal energy at zero Kelvin, $T = 0K$, and Helmholtz free energy is given by:

$$A = -k_{\mathrm{B}}T \ln Q \tag{3.11}$$

where Q is the canonical partition function (NVT ensemble). For NPT ensemble, Gibbs free energy is obtained. In practice, condensed phase systems present very similar ΔA and ΔG values. The calculation of free energy differences along a geometric parameter such as the reaction coordinate, ξ, permits the computation of free energy barriers for chemical processes. Such type of calculations require sampling the system over its degrees of freedom but ξ to account for thermal and entropic contributions. In practice, in computer simulations the probability density $P(\xi)$ is computed rather than the partition function, which according to the ergodic hypothesis, are equal assuming infinite sampling. Then, the Potential of Mean Force (the free energy along the reaction coordinate), is calculated by:

$$W(\xi) = -k_{\mathrm{B}}T \ln P(\xi) \tag{3.12}$$

To compute the PMF, the main problem is that sampling in high energy regions (e.g. transition state) is poorly done, and to overcome this difficulty there are two principal methods: Free Energy Perturbation[110] and Umbrella Sampling.[111,112]

3.5.1 Free Energy Perturbation

Free Energy Perturbation (FEP)[110] is a method that relates the free energy difference between an initial state (reference) and a final state (target). The term “perturbation" is misleading, since no perturbation theory is applied here. The original FEP method can be extended to QM/MM calculations by applying the sampling only in the MM part, while the QM part is kept frozen.[113–115] This is advantageous from the computational point fo view, since QM calculations are more demanding. However, there exists the evident disadvantage that the phase space is restricted to the degrees of freedom of the MM part as the QM part is fixed. The entropy change of the QM part is not sampled, but might be estimated from a harmonic approximation. Also, the density of QM part is estimated by electrostatic potential charges (ESP) in its interaction with the MM part during the sampling.

A FEP calculation requires a calculation of the potential energy profile, calculation of the energy of perturbation, ΔE_{pert}, and sampling of ΔE_{pert}. Sampling is preformed through MD. QM/MM FEP during a reaction coordinate is calculated by defining a set of windows along this reaction coordinate. The free energy difference between two consecutive windows is calculated. One particular structure from window i is perturbed forward to window $i+1$ by:

$$\Delta E_{\text{pert}}^{i\to i+1} = \underbrace{E_{\text{QM/MM}}(\vec{r}_{\text{QM}}^{\,i+1}, \vec{r}_{\text{MM}}^{\,i})}_{\text{perturbed}} - \underbrace{E_{\text{QM/MM}}(\vec{r}_{\text{QM}}^{\,i}, \vec{r}_{\text{MM}}^{\,i})}_{\text{unperturbed}} \tag{3.13}$$

The perturbed energy is calculated with the MM positions of the current window i and the QM positions of the next window $i+1$, in a forward perturbation. In a backward perturbation window $i+1$ is perturbed to i with an analogous expression (interchanging i and $i+1$ in Equation 3.13).

For each perturbation term (forward and backward) between two consecutive windows a complete sampling is calculated, in order to obtain the free energy difference between the windows, which is given by:

$$\Delta A^{i\to i+1} = \Delta E_{\text{QM}}^{i\to i+1} + \Delta A_{\text{QM/MM}}^{i\to i+1} \tag{3.14}$$

where $\Delta E_{\text{QM}}^{i\to i+1}$ is the difference of the QM energies of window i and $i+1$. The change in free energy due to QM/MM interactions and the MM part are incorporated in the $\Delta A_{\text{QM/MM}}^{i\to i+1}$ term, which is obtained through sampling:

$$\Delta A_{\text{QM/MM}}^{i\to i+1} = -\frac{1}{\beta}\ln\langle\exp(-\beta\Delta E_{\text{pert}}^{i\to i+1})\rangle_{\text{MM},i} \tag{3.15}$$

Thus, for a window i the energy is sampled over the MM part, which moves according to the forces from the QM part of i. Then the average is taken over the MM coordinates, since the QM part is frozen. As mentioned above, the entropic effects of the QM part can be taken into account by calculating the harmonic frequencies of the QM atoms and evaluating the difference $\Delta A_{\text{QM}} - \Delta E_{\text{QM}}$ by statistical-thermodynamic methods for the stationary points. Also, the electrostatic interaction between QM and MM regions is calculated for a given electron density of the frozen QM part instead of evaluating the full SCF iterations in each MD step.

3.5.2 Umbrella Sampling

The idea behind the Umbrella Sampling (US) method[111,112] is to add a biasing potential in order to restrain the system to a given volume of the phase space. Doing so, permits the simulation to sample high energy regions that otherwise (with an ordinary simulation) are never sampled. This is specially important when the simulation is performed over a reaction coordinate in order to compute the potential energy profile and sample regions near the transition state, which are hardly sampled due to their higher energy. This profile is calculated from a series of biased simulations (windows) each one in a particular value of the reaction coordinate. A probability distribution is obtained for each window, and the final profile is compute “unbiasing" the result.

$$E^b(r) = E^u(r) + \omega_i(\xi) \tag{3.16}$$

Typical biasing potential is a harmonic potential which pulls the simulation on each particular window i to their corresponding reference value of the reaction coordinate, ξ_i^{ref}:

$$\omega_i(\xi) = \frac{k}{2}(\xi - \xi_i^{\text{ref}})^2 \tag{3.17}$$

The biasing potential ω_i is added to the energy, and only depends on the reaction coordinate. A probability distribution is then obtained for each window, containing the biasing term. The unbiased free energy is obtained once the distribution is unbiased. And then the profile is computed with one of the several methods developed to analyze US. Two of these methods will be explained here.

In practice, for a given reaction coordinate, the potential energy profile is first constructed, and the different geometries along a reaction coordinate can be considered as the starting point of the US simulations (i.e. windows). Then a QM/MM MD simulation (in this case QM/MM formalism is required, as a chemical process takes place), is run in each window, but with the corresponding biasing potential, that pulls the system to the value of the reaction coordinate where the window is set. The rest of the system explores the conformational space allowed by the biasing potential during the MD simulation (QM/MM MD).

The unbiased probability has to be recovered from the biased one. This leads to the probability of a particular window, and in the end all the windows are combined to obtain the free energy profile. There are several ways of computing the free energy profile. The most extended is the Weight Histogram Analysis method (WHAM).[116] Very recently, a new method based on a Markov model formalism was developed for the same purpose, the so-called Dynamic Histogram Analysis Method (DHAM).[117]

3.5.2.1 Weighted Histogram Analysis Method

The weighted histogram analysis method[116] is based on dividing the reaction coordinate ξ into discrete bins (a histogram), in each one of these there is a relative probability of observing the state of interest.

WHAM computes global probability distribution by a weighted sum of the individual windows.

$$P^u(\xi) = \sum_{i}^{\text{windows}} p_i(\xi) P_i^u(\xi) \tag{3.18}$$

The weights p_i are subject to the condition $\sum_i p_i = 1$, and they are calculated by minimizing the statistical error on the weights assigned to each particular window P_i^u. These two conditions give rise to the following expressions.

$$P^u(\xi) = \frac{\sum_i^{\text{windows}} n_i(\xi)}{\sum_i^{\text{windows}} N_i \exp[(F_i - \omega_i(\xi))/k_\text{B}T]} \tag{3.19}$$

$$\exp(-F_i/k_\text{B}T) = \int P^u(\xi) \exp[-\omega_i(\xi)/k_\text{B}T] d\xi \tag{3.20}$$

where $n_i(\xi)$ is the number of counts in histogram bin associated with ξ, and N_i is the number of steps sampled in window i. Since $P^u(\xi)$ and F_i are unknown, equations 3.19 and 3.20 are solved iteratively to self-consistency, although the convergence can be slow. F_i is the shift introduced in the calculation of the unbiased probability from the biased distributions in US, and is defined as $F_i = -k_\text{B}T \ln \langle [-\omega_i(\xi)/k_\text{B}T] \rangle$.

WHAM is a widely-used method to compute the potential of mean force but it presents some drawbacks, mostly derived from the incorrect sampling of the windows that can lead to an incorrect free energy barrier, especially when the sampling does not reach equilibrium. In some simulations complete equilibration within a window may be never achieved. These disadvantages can be overcome by the use of DHAM method.

3.5.2.2 Dynamic Histogram Analysis Method

Dynamic histogram analysis method[117] is a procedure for unbiasing the US simulations based on Markov States theory. Similarly to WHAM, the reaction coordinate ξ is divided into small bins in each biasing window k, which are considered as the microstates for a Markov State Model treatment. Then the different trajectories in each window are projected into the bins along the reaction coordinate. For each simulation k, a transition count matrix $T^{(k)}$ is constructed, in which each element of the matrix, $T_{ji}^{(k)}$ denotes the integer number that accounts for the transitions from bin i into bin j in k. Then the unbiased Markov transition probability matrix between bins M_{ji} is constructed using the maximum likelihood approach.

The Markov matrix contains both energetic and dynamic information for the motion along the reaction coordinate. Thus, it can be also used to model the projected dynamics along the reaction coordinate and estimate barrier-crossing rates in addition to the potential of mean force. The eigenvector corresponding to the first eigenvalue ($\lambda_1 = 1$) of the transition matrix M_{ji} lead to the unbiased probability density of each bin, $P(\xi)$, and the second largest eigenvalue provides the relaxation times at different lag times, and the rate constant is calculated as the inverse of the average of these relaxation times.

DHAM takes into account the dynamics and time sequence, and does not require the equilibrium is reached in each window, which is advantageous with respect to WHAM. The local transition probabilities of DHAM are well converged even when the global equilibrium populations are not fully sampled. DHAM method alleviates the problem of not reaching local equilibrium in US simulation (which causes errors in WHAM) by explicitly accounting for the time dependence. At the limit of uncorrelated equilibrium distribution at long time steps, DHAM yields the equations of WHAM. Therefore, DHAM can be considered as a generalization of WHAM, in which reaching equilibrium is no longer necessary.

3.6 Molecular Dynamics

Molecular Dynamics (MD) provides the theory to study the evolution of molecular systems during time series.MD first appearing dates back to the late 1950's, when Alder and Wainwright studied the interaction of hard spheres to explain liquids behavior.[118] It was not until 1977 when the first simulation of a protein, bovine pancreatic trypsin inhibitor, was carried out by McCammon, Gelin and Karplus.[119] Today the complexity of the systems that can be simulated by MD is continuously expanded thanks to the evolution of the computational resources.

The recent development of Graphic Processor Units (GPUs),[120] has led to an increase in simulation time of several folds. Modern GPUs contain hundreds of arithmetic units that can be harnessed to give tremendous acceleration for numerically intensive calculations. In the case of MM MD typically simulations can handle up to 1,000,000 atoms.[121] Today, MD calculations reach simulations times of hundreds of nanoseconds, which only few years ago were considered unaffordable.

3.6.1 Equations of motion

In MD simulations the evolution of a molecular system over the time is obtained. In classical MD, the movement of the atoms of the system is calculated according to the Newtonian equations of motion:

$$-\frac{dE}{d\vec{r}_i} = m_i \frac{d^2\vec{r}_i}{dt^2}, \quad i = 1, 2...N \tag{3.21}$$

where m_i is the mass of atom i, $\vec{r}_i$ its position, and E is the potential energy function (in classic MD is the force field energy expression, see eq. 3.1), depending on the position of all the N atoms of the system. Newton's equations (Eq. 3.21) of a N particle system do not have analytical solution. Then the equations are solved by numerical methods, which are based on *Finite Difference Methods*: time step is discretized during the integration, and the force is considered constant in each time step; the forces are then determined in the next step, leading to new positions and velocities. All the algorithms for integrating the equations of motion are based on truncated Taylor expansions of the positions and dynamic properties (velocities, accelerations, ...).

There are different algorithms to integrate the equations of motion. The most common algorithm used for integrating Newton's differential equations is the Verlet algorithm.[122] The next step $t + \Delta t$ (where Δt is the time step) is computed taking into account the previous step and the current step, but velocities do not appear explicitly in its expression. Its variant Velocity-Verlet integrator[123] provides positions, velocities, and accelerations at the same time without compromising precision. Other variant of Verlet integrator is the "leap-frog" method.[124] It also is advantageous over the Verlet method as it includes velocities explicitly and does not require the subtraction of large numbers that would introduce important numeric errors.

The choice of a correct time step is crucial during an MD simulation. Very short time steps provide more accuracy, but also slow-down the MD calculations. On the other hand, a too large time step can lead to instabilities in the integrator. NAMD software incorporates a multiple-time-stepping method that improves integration efficiency by computing slower-varying forces less frequently than faster-varying ones.[125] In a typical MD calculation some bonds and angles are constrained to satisfy specific values during the simulation. The most used method is the SHAKE algorithm,[126] that fixes bond lengths and angles involving hydrogen reducing the number of high-frequency motions.

3.6.2 NVT and NPT ensembles

MD simulations provide microscopical properties as the phase space of the system is sampled. The idea of obtaining a macroscopic property from a simulation is based on the *ergodic hypothesis*, which establishes that a given system visits all its accessible states in an infinite simulation, and this is equivalent to infinite simulations of a finite time. Thus from an MD simulation some macroscopic magnitudes can be derived. This is interesting for simulating experimental conditions, as typical experiments are performed at constant pressure and temperature, thus NPT (isothermal-isobaric) or NVT (canonical) ensembles are employed in MD simulations, especially for condensed-phase systems as proteins.

The ensemble distribution of the system (NVT, NPT) has to be correctly generated by the integrator in use. For this purpose, the system is coupled to a reservoir, with a coupling either deterministic or stochastic. Stochastic couplings are easier to implement, and provide dynamic stability by their friction terms (MD software as NAMD[125] incorporate stochastic coupling). One of this couplings is given by Langevin equation (Eq. 3.22), which generates a Boltzmann distribution for the NVT ensemble, by introducing additional damping and random forces. Temperature is also controlled by this method.

$$m_i\vec{a}_i = F(\vec{r}_i) - \gamma\vec{v}_i + R(t) \tag{3.22}$$

where m_i is the particle mass, $\vec{a}_i$ its acceleration, $F(\vec{r}_i)$ is the force, $\vec{r}_i$ the position, γ the friction coefficient, and $R(t)$ a Gaussian random force. Other used reservoirs are NosÃľ-Hoover thermostat[127,128] and Berendsen thermostat and barostat.[129]

3.6.3 Boundary conditions

During MD simulations of condensed phase systems the one under study must be treated as if it was located in a bulk solution. The boundaries of the actual system must then be able to reproduce such effect. There are to main approaches, periodic and non-periodic boundary conditions.

3.6.3.1 Periodic Boundary Conditions.

During a simulation under periodic boundary conditions (PBC) the system is surrounded by identical images in all possible directions, in order to simulate an infinite environment for the solvated system. The particles experience forces as they were in a bulky solution. The coordinates of the images are identical to the system coordinates except by adding or subtracting multiples of the box sides. When a particle leaves the system it is immediately replaced by an image particle entering by the opposite site (wrapping). The interaction of the current system with its surrounding images simulates the interaction of the system within a bulky medium.

3.6.3.2 Non-periodic boundary conditions

There are some specific cases when it is not necessary to simulate an infinite environment. That is the case when studying local changes onto a system, for example, when the interest focuses on an specific change into the active site of a protein. In those cases the environment is considered as a droplet. To avoid the inherent problems of working in a vacuum boundary, some explicit solvent molecules can be added creating a skin of solvent molecules

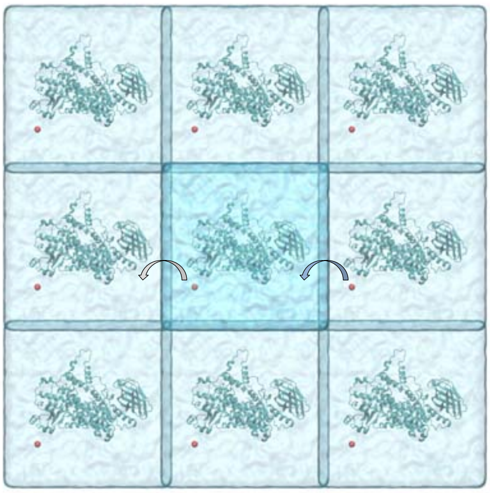


Figure 3.3: An illustrative representation of periodic boundary conditions (boxes are only shown in the plane of the page for clarity). A water molecule is highlighted as a red sphere. When the molecule leaves the system, it is immediately replaced by a molecule of an image. The system itself (central box) is highlighted in a darker color.

surrounding the system. This approach avoids dealing with huge numbers of atoms that are common in PBC.

When the interest is focused only in a specific part of the system, it can be divided into two differentiated regions: the reaction zone, where the simulation is fully performed, and the reservoir zone, surrounding the former region, which is kept frozen during the simulation (or restrained), creating a shell. This is the *stochastic boundary conditions* approach.

3.6.3.3 Particle Mesh Ewald.

The electrostatic interactions of a simulation under PBC are normally treated by Particle Mesh Ewald (PME) method. PME is a fast method to compute Ewald summation, which describes long-range electrostatic interactions for a spatially limited system under PBC.[130] In such system, this summation is infinite, as it involves the charge of the N particles of the system, plus the infinite number of images; however, for a charge-neutral system is conditionally convergent (the value depends on the order in which its terms are considered (truncated)). Ewald sum expression is given by:

$$V = \frac{1}{2} \sum_{|\mathbf{n}|=0}{}' \sum_{i=1}^{N} \sum_{j=1}^{N} \frac{q_i q_j}{4\pi\varepsilon_0 |\mathbf{r}_{ij} + \mathbf{n}|} \tag{3.23}$$

where $\mathbf{n}$ is the position vector of a given point in the cubic lattice ($\mathbf{n} = (n_xL, n_yL, n_zL)$, being L the length of the box and n_x, n_y and n_z integers), and the prime on the first summation indicates that the series do not include the interaction $i = j$ for $\mathbf{n} = 0$, that is, in the primary box.

Ewald summation is then a procedure to calculate the electrostatic interactions in an infinite lattice, which is the case of a system under PBC. To avoid slow and conditional convergence, the traditional summation is transformed into two different sums, which are modulated by a parameter, function, $f(r)$, in order to have one term acting for short-range interaction (the real space, the actual system) and other term acting over long-range interaction (the reciprocal space, image boxes).

$$\frac{1}{r} = \underbrace{\frac{f(r)}{r}}_{\text{real}} + \underbrace{\frac{1-f(r)}{r}}_{\text{reciprocal}} \tag{3.24}$$

The convergence of this summation is extremely slow, and further treatments try to overcome this drawback. One improvement consists on transforming this summation into two series of more rapid convergence. Without entering into details, this is the foundation of the PME method, which uses a fast Fourier transform. Traditional Ewald sum scales with N^2, but incorporating the PME algorithm reduces it to $N \ln N$. The basic idea is that PME computes the k-sum for fixed positions on a grid instead of atoms positions: atomic charges are distributed over the grid, potential and forces are computed, and finally are transformed back into the atomic positions.[131]

3.7 Homology Modeling

Protein three-dimensional structure determination by experimental techniques is not always successful although a huge knowledge of protein sequences is available. This has led to the development of computational methods to determine the protein structure by its sequence, which can be divided into two main groups: *ab initio* methods[132] and homology modeling.[133] The first method computes the protein structure by its folding prediction from its physico-chemical properties; this method is not affordable nowadays for most problems as it requires tremendous computational resources and has thus only been tested for tiny proteins. The latter is based on the previous knowledge of the 3D structure of one or more proteins related with the protein of unknown structure, whose sequence is known. The enormous difficulty of protein folding structure prediction has led homology modeling to be a widely used choice.

Homology modeling technique is based on two major observations. First, protein tertiary structure is better conserved than amino acid sequence;[134, 135] thus, even proteins that have diverged appreciably in sequence but still share detectable similarity will also share common structural properties, particularly the overall fold. Second, the structure of a protein is uniquely determined by its aminoacid sequence,[136] thus, knowing this sequence the 3D structure can be derived, at least in theory. Nowadays, thanks to the growth of the protein structures registered in Protein Data Bank, a precise size limit to the relationship between primary sequence and structure has been established. As the percentage of identical residues and the number of aligned residues increase in two aligned sequences, a region for which homology modeling is suitable can be defined.[137]

The general procedure of homology modelling can be summarized in seven steps,[133] although depending on the particular case, not all the steps are necessary. For example, if a crystal structure is incomplete, and a different crystal structure of the same protein is available, the template search step can be ignored. So, during a standard homology modelling procedure the following steps are performed: i) template recognitizon and alignment, ii) alignment correction, iii) backbone generation, iv) loop modeling, v) side chains modeling, vi) model optimization and vii) model validation.

The generation of the backbone coordinates (atoms C, C_α, N, O) is the simplest part, since normally the template coordinates can be directly used. Loop modeling can be a tricky step because loop conformations are difficult to predict due to the nature of the loops (in most cases solvent-exposed and with contacts in the crystal structure), and there is not a unique way to carry out loop modeling. There are two main approaches:

- *Knowledge-based*: consists of searching in PDB for known loops with endpoints that match the residues between which the loop will be constructed. Most modeling programs support this feature.[138]
- *Energy-based*: this approach is an *ab initio* fold prediction. An energy function evaluates the quality of the loop. This function is minimized through Monte Carlo or Molecular Dynamics.[138]

For short loops (5-8 residues) the predicted structure is reasonable.

Side-chain modeling requires finding the most stable side-chain orientation. Conformations of residues side-chains (rotamers), are similar in structurally similar proteins: they present a χ angle (torsion around C_α-C_β) with a similar value. In similar proteins, rotamers and contacts are conserved. When two sequences differ, with less than 35% of sequence identity, rotamers may differ up to 45%.[139] To facilitate the search, the most probable rotamer values are collected in rotamer libraries.

Finally the generated structure is minimized normally by MM minimizations or simulated annealing. The final structure then is validated by a series of checks that measure the quality of the generated coordinates (side-chain quality, overall 3D structure, etc).

3.8 Docking simulations

Docking is a tool for predicting the binding mode and affinity of two molecules, the ligand, which is typically a small molecule as a substrate or drug, and the receptor, typically a large molecule like a protein.[140] Docking procedure can be rationalized as the search for the precise ligand conformations and orientations, usually referred as “pose", inside the receptor binding site. Docking methodology acts trough two basic tools, the *search algorithm*, which explores the conformations of the ligand inside the receptor, and the *scoring* or *fitness function*, which evaluates the binding affinities of the conformations generated in order to find the best binding mode.

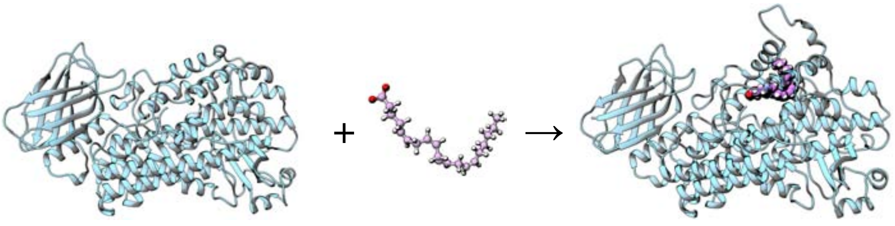


Figure 3.4: Basic idea of docking procedure.

3.8.1 Search algorithms

The *search algorithm* is a procedure for exploring the conformational landscape of the ligand to find the best poses within the receptor binding site. A good search algorithm must consider enthalpic and entropic effects, but in most approaches inclusion of entropy is not straightforward. Search algorithms are classified into systematic, deterministic and stochastic.[141]

- Systematic algorithms explore all the degrees of freedom of the ligand (translational, rotational, and vibrational) during the search, generating all the possible ligand binding combinations.[142,143]
- Stochastic algorithms randomly change all ligands degrees of freedom at each step. There are several types of stochastic algorithms, among them Monte Carlo, Evolutionary Algorithms (which will be explained in detail below), Tabu Search,[144] and Swarm Optimization.[145]
- Deterministic methods generates a solution that depends on the initial conditions. Methods based on MD simulations are an example.[146]

One widely used search procedure are the evolutionary or genetic algorithms (GA),[147,148] included in many docking software codes as GOLD (Genetic Optimization for Ligand Docking).[149] This procedure mimics biological evolution by manipulating a collection of data structures called "chromosome", which contains a set of values describing translation, orientation, and conformation of the ligand, and encode a possible protein-ligand interaction. Randomly, a starting population (the first generation) of chromosomes is produced, and the GA applies two genetic operators, crossover (a combination of chromosomes) and mutation (a random perturbation introduced in the chromosome), resulting in a new set of chromosomes, which will replace the least-fit members of the forming set, until a final population of optimal individuals assessed by a predefined fitness function is obtained.

During the search procedure the flexibility of the molecules involved (ligand and receptor) can be included only in the ligand, for which normally the maximum flexibility is selected, or both in ligand and receptor (protein). The inclusion of protein flexibility during the search is a challenge due to the large number of degrees of freedom. Most strategies consider partial flexibility of the receptor. Upon ligand binding, the receptor experiences a conformational change, and it is now widely accepted that proteins exists as an ensemble of conformational states rather than a unique state.[150]

Diverse strategies have been developed to deal with the problem of protein flexibility,[151] and they can be classified into five main types: soft docking, side-chain flexibility, molecular relaxation, ensemble docking, and collective degrees of freedom.

- Soft docking accounts for conformational changes by softening the repulsive term of the Lennard-Jones potential, thus allowing small overlaps between protein and ligand atoms. The main advantage of this method is the speed, but it should be employed only to consider local motions.

- Side-chain flexibility, as its name indicates, consists of allowing the side chain of selected residues within the binding site to move by exploring their torsional degrees of freedom (rotamers). To avoid combinatorial explosion, the use of pre-defined rotamer libraries that predict discrete and low energy conformations of side-chain residues is one of the most used schemes, including several variations.

- By molecular relaxation the ligand-protein complex generated during the search in a previous rigid approach is minimized by several methods as energy minimization, Monte Carlo, or Molecular Dynamics.

- Ensemble docking considers the receptor flexibility by docking the ligand to an ensemble of protein conformations instead a single conformation.

- Finally, methods based on collective degrees of freedom consider the full flexibility of the receptor, the global flexibility, by using collective variables representing dominant motion modes obtained from the original high-dimensional representation of the protein motion. Its principal disadvantage is that the degrees of freedom are not the native ones.

3.8.2 Scoring functions

Scoring functions evaluate the quality of the different ligand poses generated in a docking calculation. A good scoring function must be able to distinguish between the true binding modes and other alternative modes. There are three main classes of scoring functions: force-field based, empirical, and knowledge-based.

Force-field based scoring functions[152] are derived from a classical force field, and have the form of a sum of different energy terms. These scoring functions present the limitation of the absence of solvation and entropic effects. Knowledge-based scoring functions[153] collect the data of the interacting atom pairs in known ligand:protein complexes with available tridimensional structure, and these data are converted in numerical values based on the occurrence of such pairwise interactions. Their main limitation is the dependence on the number of structures available.

Empirical scoring functions approximate binding energies by the sum of individual uncorrelated terms that are derived from experimentally observed binding energies of several ligand:receptor complexes by regression. The main advantage of these functions is the computational cost, but present the disadvantage of the requirement of a parametrization process. One example is ChemScore function,[154] implemented in many docking packages as GOLD. This function assigns a value to the docking pose taking into account an estimation of the binding free energy, $\Delta G_{\text{binding}}$, and two terms that introduce a clash penalty for close contacts (P_{clash}), and internal torsion terms against poor internal conformations, P_{int}. Additionally, covalent and constraint scores can be included.

$$\text{ChemScore} = \Delta G_{\text{binding}} + P_{\text{clash}} + c_{\text{int}} P_{\text{int}} + (c_{\text{cov}} P_{\text{cov}} + P_{\text{constr}}) \tag{3.25}$$

The change in free energy upon ligand binding is estimated by the following expression:

$$\Delta G_{\text{binding}} = \Delta G_0 + \Delta G_{\text{H bond}} + \Delta G_{\text{metal}} + \Delta G_{\text{lipo}} + \Delta G_{\text{rot}} \tag{3.26}$$

Each component of equation 3.26 depends of the magnitude of a particular contribution to free energy (e.q. hydrogen bonds, lipophilic interactions), and a scale factor determined by regression.

The limitations of each individual scoring method can be overcome by a consensus scoring function,[155] combining different scores to compensate errors from individual scoring functions. However, the use of such functions must be done with care, as correlation of the different scoring functions can lead to error amplification.

3.8.3 Validating docking results

There is not a unique way to determine which solution of the ensemble of solutions generated by the docking is the best. Sometimes the answer can be very clear, and one of the solutions is clearly most stable or frequent, but generally the answer is not straightforward. The myriad of poses generated by the search algorithm can be classified according to a clustering process, and most populated clusters generally contain the biologically active binding mode.

Despite all the considerations that can be taken into account when finding the binding mode of a molecule onto a protein by docking simulations, one should not stop here, as docking methodology is based on many approximations, and further studies with the generated complex should be performed in order to validate that binding mode. This problem (the binding mode) is very complex and to date, there is not a feasible, universal tool to address it.

Part IV

Results and Discussion

Chapter 4

15-LOX-1

4.1 Introduction

Reticulocyte-type 15-LOX-1 has been one of the most studied mammalian lipoxygenases in recent years. The X-ray structure of human 15-LOX-1 has still to be reported, but the corresponding crystallographic structure for rabbit 15-LOX-1 was the first solved for a mammalian LOX.[26,27] It has to be noted that rabbit 15-LOX-1 has approximately 80% sequence identity with human 15-LOX-1.[156]

Most of the products generated by the reaction of human 15-LOX-1 with AA come from the initial hydrogen abstraction from the carbon C_{13} (with, for instance, a C_{13}:C_{10} ratio of 10:1[157] or 89:11[38]). The same happens for rabbit 15-LOX-1, with a C_{13}:C_{10} ratio of 97:3.[40] Interestingly, a dramatic inversion in the regiospecificity of the reaction of human 15-LOX-1 with AA occurs when the methylene group of AA at position 13 is dideuterated, with the products derived from hydrogen abstraction from C_{10} being now predominant by a ratio C_{13}:C_{10} of 25:75. In addition, it has been experimentally determined that the H/D kinetic isotope effects (KIEs) for the hydrogen abstraction steps from C_{13} and C_{10} in human 15-LOX-1 turn out to be comparable (${}^{\mathrm{D}}k_{\mathrm{cat}} = 11.6 \pm 2.0$ and ${}^{\mathrm{D}}k_{\mathrm{cat}} = 8.5 \pm 4.0$, respectively). The corresponding ${}^{\mathrm{D}}k_{\mathrm{cat}}/K_{\mathrm{M}}$ values are also comparable for the two carbon centers. From these results Holman and coworkers[38,39] suggested that the hydrogen abstraction at these two structurally different positions (C_{13} and C_{10}) follows a comparable mechanism with, probably, very similar transition state structures. This conclusion sounded quite odd and unexpected for the same authors because a highly regioespecific enzyme like human 15-LOX-1 should be however able to spatially accommodate those two disparate carbon positions in a similar way.

On the other hand, in a previous work in our group,[28] a study combining protein-ligand docking, molecular dynamics simulations, and *in silico* mutagenesis was performed to investigate the binding modes of AA in the active site of rabbit 15-LOX-1. In that work, simulations gave similar average values for the C_{13}/H_{13}-OH–Fe(III) and C_{10}/H_{10}-OH–Fe(III) distances, in such a way that those carbon atoms are not really in such structurally different locations in a great number of configurations of the rabbit 15-LOX-1:AA Michaelis complex. Even more, the distances corresponding to C_{10} tend to be somewhat smaller than the ones in the case of C_{13}, and so, not only both hydrogen abstractions seem to be possible, but hydrogen abstraction from C_{10} could be guessed as the most favorable, contrarily to the experimental evidence. Thus, how does human 15-LOX-1 (and rabbit 15-LOX-1) manage to achieve such an exquisite regioespecificity?

Other important aspect of LOXs is their substrate specificity. The most common substrate for plant LOXs is LA, whereas animal LOXs preferentially react with AA. For instance, only AA, but not LA, is a substrate of human 5-LOX and 12-LOX. However, mammalian reticulocyte 15-LOX-1 and epithelial 15-LOX-2 can catalyze both LA and AA oxidations. There are though discrepancies in the literature about the substrate specificity for 15-LOXs, steady-state kinetic experiments suggest that human 15-LOX-1 metabolizes LA better than AA, whereas human 15-LOX-2 metabolizes LA very poorly.[158] The question that must be answered to understand substrate specificity in 15-LOXs is how these enzymes can accommodate and recognize for catalysis two substrates that are different in length and are also different in the degree of unsaturations (four double bonds in AA versus two double bonds in LA).

The aim of this Chapter is to get a deep insight on the mechanism of the hydrogen abstraction reaction from AA and LA catalyzed by 15-LOX-1, which can serve as a model of human 15-LOX-1. A special emphasis has been put on shedding light on the subtle factors that govern the strict regioespecificity of this first step of the hydroperoxidation of AA in spite of the fact that, in principle, both the hydrogen abstractions from C_{13} and from C_{10} would seem to be plausible. To this aim, quantum mechanics/molecular mechanics (QM/MM) calculations on the entire solvated rabbit 15-LOX-1:AA complex have been carried out, employing the density functional theory (DFT) to treat the QM region in order to obtain a reliable enough description of the evolution of the electronic structure of the system along the reaction and of the chemical properties depending on it. For simplicity, hereafter rabbit 15-LOX-1 will be referred only as 15-LOX-1.

4.2 Methodology

All QM/MM calculations have been performed with the modular program package ChemShell,[105] using TURBOMOLE[159] for all of the DFT calculations. The MM calculations were carried out by the DL-POLY[160] module in ChemShell, using the CHARMM22[161,162] (for the protein atoms) and CHARMM27[163,164] (for the lipid moiety included in the MM region) force fields. An electronic embedding scheme[98] has been adopted in all calculations, and a hydrogen link atoms scheme has been employed to treat the QM/MM boundary using the charge shift model. No cutoffs were introduced for the nonbonding MM and QM/MM interactions.[165]

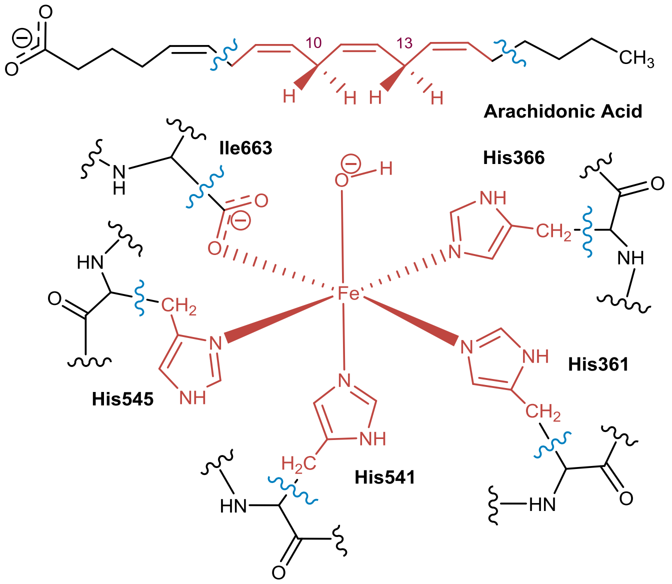


Figure 4.1: Representation of the QM/MM boundary. QM atoms are depicted in red, and link atoms are represented by blue wavy lines.

QM/MM optimizations were carried out employing the limited-memory Broyden-Fletcher-Goldfard-Shanno (L-BFGS) algorithm[166,167] in the case of energy minimizations, and the microiterative optimizer, combining both the partitioned rational function optimizer (P-RFO)[168,169] and the L-BFGS, was used during the transition state searches. All of these algorithms are implemented in the HDLCopt (Hybrid Delocalized Internal Coordinate Scheme)[170] ChemShell module. We have taken the natural population analysis (NPA)[171] atomic populations evaluated from the spin density as implemented in TURBOMOLE. The VMD program[172] has been used to generate the pictures of molecules presented here.

In all calculations the QM subsystem has been described by the B3LYP hybrid functional, and the 6-31G(d) Pople basis set[173] has been used for all atoms except for Fe, which was described by the LANL2DZ basis set.[174] The QM/MM partition used in the reaction of AA is shown in Figure 4.1. This QM region includes 74 atoms (link atoms not included): 24 atoms of lipid substrate, 11 atoms of each of the 4 His residues in the Fe coordination sphere (His361, His366, His541, and His545), 3 atoms of the terminal Ile663 in the coordination sphere, and the Fe(III)-OH^- cofactor, able to accept the hydrogen from AA. Seven link atoms were used, five along the bonds Cα-QM atom of the five residues in the Fe coordination sphere and two bonded to the aliphatic carbons of the lipid substrate (placed between C_6-C_7 and C_{16}-C_{17}). In the case of LA, 29 atoms of the lipid substrate were included within the QM region, with the same atoms for the Fe-coordination sphere (and same number of link atoms) as AA, making a total of 79 QM atoms.

QM/MM free energy perturbation calculations (QM/MM-FEP) were performed with the qmmmfep module of Chem-Shell.[114] According to this methodology, the QM part is frozen at each window of the reaction coordinate, and only the computationally less demanding MM part is sampled in the FEP calculation. The MM-MD sampling at each window consisted of an equilibration period of 20 ps followed by FEP sampling for a minimum of 20 ps. Forward and backward profiles were performed to check for convergence. All the MD simulations were done in a canonical (NVT) ensemble using the Berendsen thermostat. A time step of 1 fs was used. QM/MM electrostatic interactions during the MM-MD sampling at each window were calculated by replacing the frozen QM density by electrostatic potential (ESP) adjusted charges.

All the structures used in this study have been taken from previous works developed within the group,[28, 175] in which a series of classical molecular dynamics (MD) trajectories under periodic boundary conditions of the AA:15-LOX-1 and LA:15-LOX-1 Michaelis complexes were performed. Considering the snapshots of these trajectories, a selection criterion has been applied to finally choose the reactive structures studied in this work.

4.3 Results

4.3.1 Selection of starting geometries

As mentioned above, a selection criterion has been applied to filter the structures generated from previously generated long classical MDs trajectories, in order to select those structures whose geometric characteristics have been considered adequate to initiate the reactive process. This criterion is based on the initial distance between the hydrogen atom that would be abstracted, pro-S H_{13} and/or pro-S H_{10} (hereafter referred only as H_{13} and H_{10}, respectively), and the oxygen atom of the hydrogen acceptor, i.e., the oxygen of the Fe(III)-OH^- cofactor. This distance has been established to be ≤ 3.0 Å for the reaction to take place. Furthermore, the distance from the corresponding carbon atom, C_{13} or C_{10}, to that oxygen atom had to be greater than the distance from H_{13} and H_{10}, respectively, thus ensuring that the corresponding C-H bond is properly oriented for hydrogen abstraction.

Following application of this selection criterion, a set of structures was filtered from the rest that do not fulfill the required characteristics. The selection process has been repeated three times: applying the selection criterion only to H_{13}, only to H_{10}, or to both at once; therefore, we have studied three types of structures. It has to be noticed that the number of structures with H_{10} in the proximity of the Fe(III)-OH^- cofactor is significantly larger that those with H_{13}, as was already reported in the previous work of our group.[28]

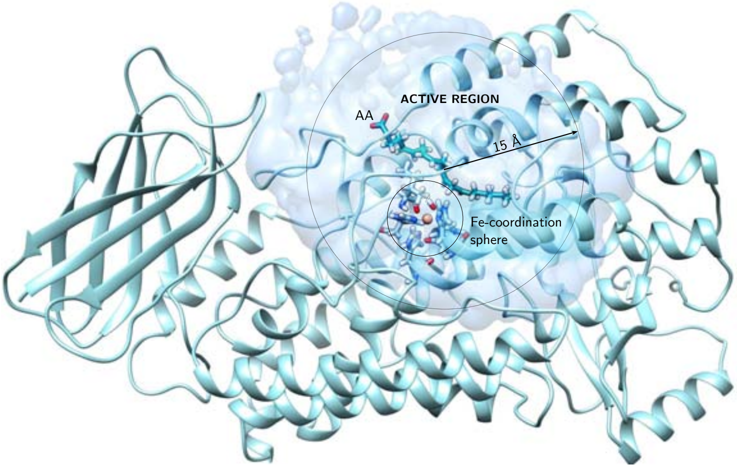


Figure 4.2: Selection of water molecules solvating the active site cavity. The light blue transparent surface represents a water volume selected choosing all water molecules within 15 Å from AA molecule (shown here in sticks model). This solvation model ensures that the active site cavity is fully solvated, and also avoids the problems associated with large number of atoms.

The geometries selected in this way have been further manipulated to deal with the calculation programs. In this sense, as the system was originally fully solvated into an orthorhombic water box containing near 75000 atoms, all the water molecules outside a 15 Å radius volume centered on the AA molecule have been removed, reducing the size of the system to near 12000 atoms. This has been done in order to reduce the computational cost of handling a large number of atoms during the QM/MM calculations, but ensuring that the active site region is completely solvated. Furthermore, all residues and water molecules included in a 15 Å radius sphere centered on the C_{11} of the AA molecule, along with the complete AA, were included in the active region during the optimization processes, allowed to move freely, while the remaining atoms were kept frozen. The active region comprises nearly 2000 atoms.

Finally, the geometries of 17 structures obtained according to the filtering process described above were optimized to obtain the reactive minima, whose main characteristics are shown in Tables 4.1 and 4.2. Those reactive minima have been used as starting points to build the QM/MM potential energy profiles for the hydrogen abstractions of H_{13} (structures I to IX), H_{10} (structures XII to XVII), or both of them (structures X and XI) of the AA by the oxygen atom of the Fe(III)-OH^- cofactor. It can be seen that the set of these reactive minima covers a wide region of the configurational space. For instance, the dispersion of geometries is such that the distance from the hydrogen atom to be transferred to the acceptor oxygen atom ranges from 2.47 Å to 3.78 Å for structures I to IX (H_{13}), and from 2.67 Å to 3.16 Å for structures XII to XVII (H_{10}).

The QM part used (described above) was the same for all investigated structures (Figure 4.1). The remaining atoms comprise the MM part of the system. Table 3 shows the total number of atoms of each of the structures in question, indicating the number of active atoms. These numbers differ for the different structures because the number of selected water molecules may vary. A complete view of one of the AA:15-LOX-1 Michaelis complexes studied (in particular corresponds to structure I, see above) is pictured in Figure 4.2.

4.3.2 Potential Energy Profiles and Reaction Mechanism

This section describes the procedure used for the study of the reactivity of all structures (I to XVII) and the corresponding results. For the sake of brevity, we essentially focus on one case corresponding to a suitable structure for H_{13} abstraction (structure I), one appropriate structure for H_{10} abstraction (structure XII), and a suitable third one for both reactive processes (structure XI). Figures 4.3 a) and 4.3 b) show the active site of the reactant geometries of structures I and XII, respectively.

The optimized geometries corresponding to the reactant minima have been used as the starting point to construct the potential energy profile along the reaction coordinate. This variable, the reaction coordinate (z), is defined as the difference between the distances of the breaking bond (C_{13}-H_{13} or C_{10}-H_{10} for H_{13} or H_{10} abstractions, respectively) and the forming bond (H_{13}-O or H_{10}-O, respectively):

$$z = r_{\text{break}} - r_{\text{form}} \tag{4.1}$$

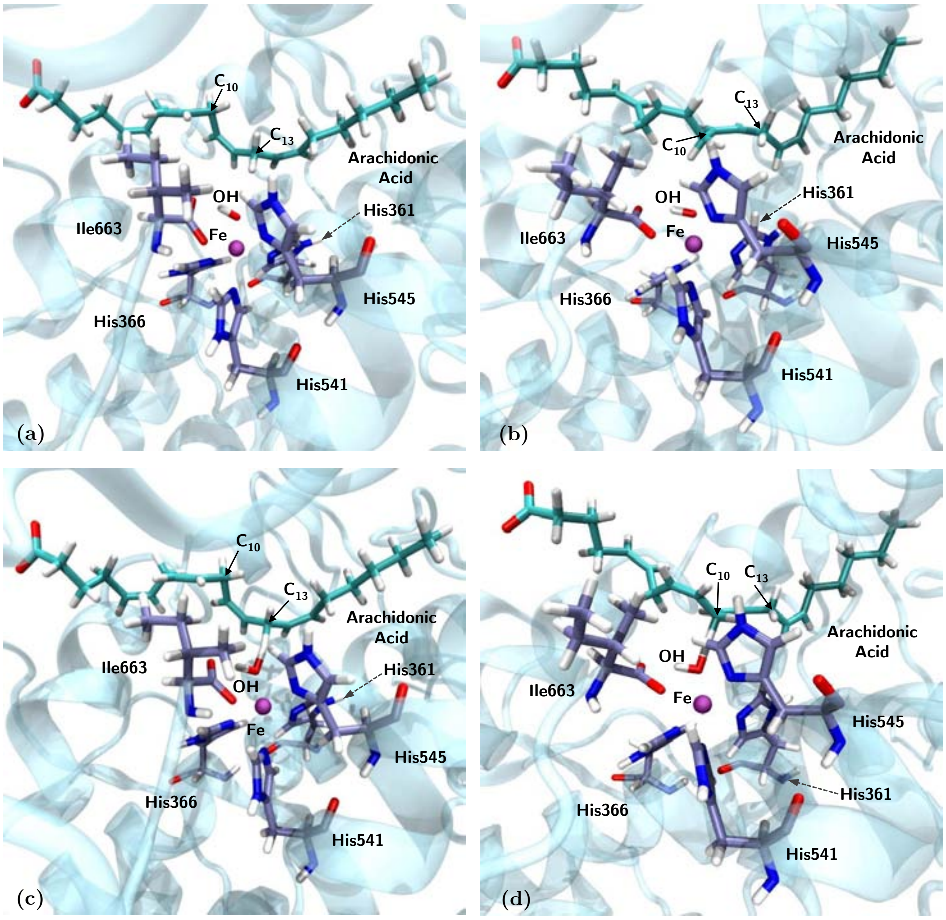


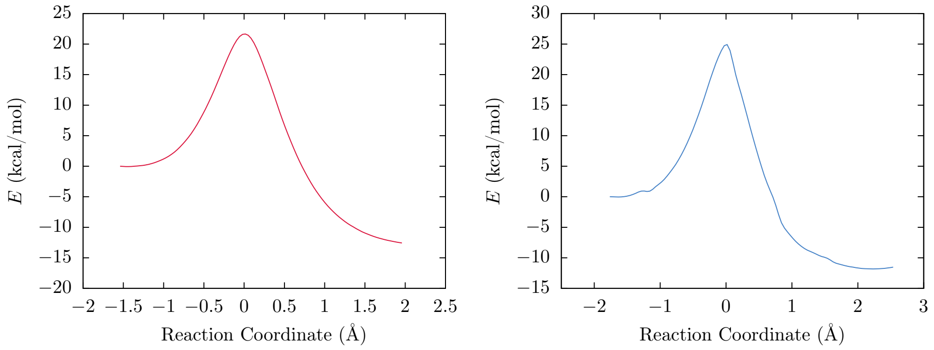
Figure 4.3: Representation of two characteristics structures (Structure I, representative of H_{13} abstraction and structure XII, representative of H_{10} abstraction) showing the minima and TSs.

The potential energy profiles have been constructed by a series of optimizations over the active part of the system in the presence of harmonic restraints on the z coordinate in the form:

$$V = \frac{k}{2}(z - z^0)^2 \tag{4.2}$$

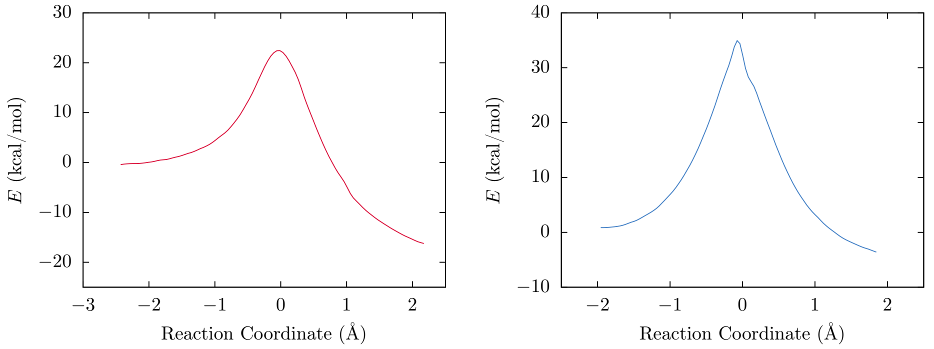
where k is the force constant, which has been set to a value of 3.0 a.u. (hartree/Bohr^2), and z^0 is the reference value for the reaction coordinate at each energy minimization calculation, which increase with a step size of 0.1 Å. Figure 4.4 represents the potential energy profile of some selected structures.

The maximum energy point of each potential energy profile has been selected for a direct transition state (TS) search. TS structures have been



(a) H_{13} abstraction of structure I.

(b) H_{10} abstraction of structure XII



(c) H_{13} abstraction of structure XI.

(d) H_{10} abstraction of structure XI

Figure 4.4: Potential energy profiles for the hydrogen abstraction step in selected structures.

optimized using the microiterative procedure explained in the Methodology section. A small core containing six atoms has been selected for the explicit Hessian matrix calculation with the P-RFO algorithm, including the Fe-OH^- cofactor and the bisallylic methylene involved in the hydrogen transfer in each case (13 or 10), while the rest of the active part of the system has been minimized by the L-BFGS procedure.In each case a saddle point type structure has been optimized, and the frequency calculations show that this structure has one single imaginary frequency (Table 4.3), and the normal mode corresponds to the investigated reaction. Table 4.1 shows that the TS structures present very similar distances between H_{13} and its donor atom (C_{13}-H_{13}), and between H_{13} and its acceptor atom (O-H_{13}) (1.34 Å and 1.33 Å, respectively) and therefore, as for the H transfer, they are highly symmetrical. To reach the transition state the donor and acceptor atoms need to approach each other roughly 1 Å (from 3.64 Å to 2.65 Å), with the C_{13}-H_{13}-O angle becoming fairly linear (167.2°), and the angle H_{13}-O-H already adopting nearly the value corresponding to the water molecule finally formed (see Figure 4.3 c)). In all the studied structures a transition state structure was found following this procedure (for H_{10} abstraction the core atoms are C_{10}, pro-S H_{10}, pro-R H_{10}, and the Fe-OH$^-$ cofactor; Figure 4.3d represents a TS structure of the H_{10} abstraction process.

Table 4.1: Main geometric parameters of reactants and TSs of the structures selected for H_{13} abstraction. Distances are given in Å and angles in degrees.

Structure	**d(H$_{13}$-OH)**		**d(C$_{13}$-H$_{13}$)**		**d(C$_{13}$-OH)**		**$\angle$(C$_{13}$-H$_{13}$-O)**		**$\angle$(H$_{13}$-O-H)**	
	R	**TS**	**R**	**TS**	**R**	**TS**	**R**	**TS**	**R**	**TS**
I	1.10	1.34	2.64	1.33	3.64	2.65	149.0	167.2	78.9	98.3
II	1.11	1.31	3.64	1.36	4.02	2.65	102.0	166.9	57.3	99.2
III	1.10	1.27	2.96	1.40	3.80	2.66	133.5	171.4	70.2	97.4
IV	1.10	1.31	3.27	1.35	3.92	2.65	118.7	173.9	76.2	99.8
V	1.11	1.32	3.78	1.33	4.09	2.64	98.2	167.5	50.5	95.9
VI	1.10	1.35	2.47	1.31	3.50	2.64	155.6	166.6	87.4	100.4
VII	1.10	1.34	2.75	1.33	3.72	2.65	146.6	165.8	83.4	99.2
VIII	1.10	1.31	3.17	1.33	3.91	2.62	124.9	165.2	116.2	99.6
IX	1.10	1.30	3.09	1.36	3.85	2.63	126.2	163.2	113.6	99.7
X	1.09	1.32	4.02	1.35	4.94	2.66	142.9	172.6	119.9	105.1
XI	1.10	1.30	3.83	1.34	4.24	2.62	104.5	165.7	99.5	99.9

Table 4.2: Main geometric parameters of reactants and TSs of the structures selected for H_{10} abstraction. Distances are given in Å and angles in degrees.

Structure	**d(H$_{10}$-OH)**		**d(C$_{10}$-H$_{10}$)**		**d(C$_{10}$-OH)**		**$\angle$(C$_{10}$-H$_{10}$-O)**		**$\angle$(H$_{10}$-O-H)**	
	R	**TS**	**R**	**TS**	**R**	**TS**	**R**	**TS**	**R**	**TS**
XII	1.10	1.34	2.86	1.33	3.78	2.67	141.0	175.5	78.0	99.5
XIII	1.10	1.34	2.71	1.30	3.76	2.64	161.7	174.8	99.8	101.7
XIV	1.10	1.31	3.08	1.36	3.94	2.67	135.4	175.0	68.7	97.4
XV	1.10	1.34	2.67	1.30	3.73	2.64	161.6	172.0	94.0	102.5
XVI	1.10	1.34	3.16	1.33	3.86	2.65	122.5	171.0	59.9	97.4
XVII	1.10	1.33	3.14	1.30	3.82	2.63	121.3	170.5	34.5	90.3
I	1.10	1.30	5.83	1.39	5.85	2.69	85.7	173.8	55.9	91.0
X	1.10	1.32	3.74	1.37	4.15	2.69	104.5	175.5	68.6	99.7
XI	1.10	1.28	3.75	1.38	4.08	2.66	99.7	174.6	37.9	93.4

For the H_{13} abstraction in structure I, the potential energy barrier turns out to be 21.6 kcal/mol and the reaction is exoergic by 12.9 kcal/mol (see Table 4.3). In this structure, the reactant is located at $z = -1.54$ Å, the transition state (TS) at $z = 0.01$ Å , and the product at $z = 2.33$ Å. The NPA atomic populations evaluated from the spin density (SD), which indicate the atomic excess of α spin, along the potential energy profile have been calculated as indicated in the Methodology section. The evolution of the SDs on C_{11}, C_{12}, C_{13}, C_{14} and C_{15} of the pentadienyl system, H_{13}, and the Fe atom of the active site, is shown in Figure 4.5a.

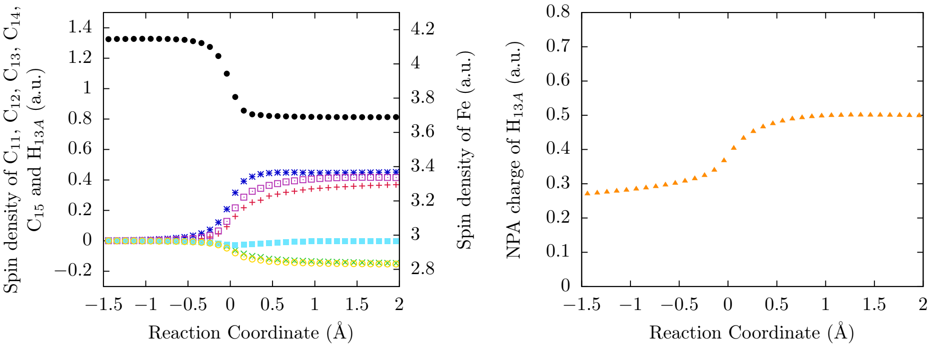
The SD on Fe decreases from 4.1 a.u. to 3.7 a.u., as corresponds to the transition of a Fe(III) sextet configuration, to a Fe(II) quintet configuration. As H_{13} is being transferred, a growing SD appears on C_{13}, which is

Table 4.3: Potential energy barriers, $\Delta E^{\ddagger}_{\mathrm{Total}}$, reaction energies, $\Delta E^{R}_{\mathrm{Total}}$, and TS imaginary frequencies, $\bar{\nu}_{\mathrm{TS}}$, of H_{13} and H_{10} abstraction processes. Energies are given in kcal/mol, and frequencies in cm^{-1}. For potential energy barriers, $E_{(\mathrm{QM,MM})}$ and $E_{(\mathrm{MM,QM})}$ are also shown (see text below).

Structure	Abstraction	$\Delta E^{\ddagger}_{\mathrm{QM,MM}}$	$\Delta E^{\ddagger}_{\mathrm{MM,QM}}$	$\Delta E^{\ddagger}_{\mathrm{Total}}$	$\Delta E^{R}_{\mathrm{Total}}$	$\bar{\nu}_{\mathrm{TS}}$
I	H_{13}	21.5	0.1	21.6	-12.9	1773 i
II	H_{13}	18.2	4.1	22.3	-16.5	1682 i
III	H_{13}	20.3	-0.7	19.6	-10.2	1312 i
IV	H_{13}	22.5	-0.3	22.2	-11.8	1648 i
V	H_{13}	22.8	1.9	24.7	-16.5	1771 i
VI	H_{13}	21.8	-3.2	18.6	-15.0	1885 i
VII	H_{13}	22.4	1.0	23.4	-14.6	1808 i
VIII	H_{13}	13.7	5.2	18.9	-21.7	1708 i
IX	H_{13}	17.5	3.1	20.6	-15.1	1627 i
X	H_{13}	24.2	6.2	30.4	-10.2	1697 i
XI	H_{13}	20.6	1.9	22.5	-17.2	1642 i
XII	H_{10}	22.4	1.7	24.1	-11.8	1812 i
XIII	H_{10}	22.0	1.6	23.6	-7.6	1774 i
XIV	H_{10}	17.4	5.3	22.7	-9.5	1639 i
XV	H_{10}	18.6	5.7	24.3	-12.5	1803 i
XVI	H_{10}	26.0	-0.2	25.8	-16.0	1848 i
XVII	H_{10}	33.2	-1.8	31.4	-9.3	1888 i
I	H_{10}	32.8	7.8	40.6	-8.8	1528 i
X	H_{10}	20.2	8.9	29.1	-9.1	1681 i
XI	H_{10}	22.2	8.7	30.9	-11.5	1437 i

becoming delocalized over C_{13}, C_{11} and C_{15} as the reaction progresses, going from 0 a.u. to around 0.4 a.u. on each one of these atoms. The latter two, therefore, will be the hit points of molecular oxygen during the next step of the hydroperoxidation process.The SDs on C_{12} and C_{14} go from 0 a.u. to -0.15 a.u. during the reactive process. The summation of the spin density on the complete pentadienyl group results in 0.9 a.u., indicating the presence of one unpaired electron. In turn, the SD on H_{13} remains virtually zero during the process, suggesting that this atom is rather a proton than a hydrogen atom bearing its electron. The representation of the SD isosurfaces, as shown in Figure 4.6, provides a good pictorial view of those changes.

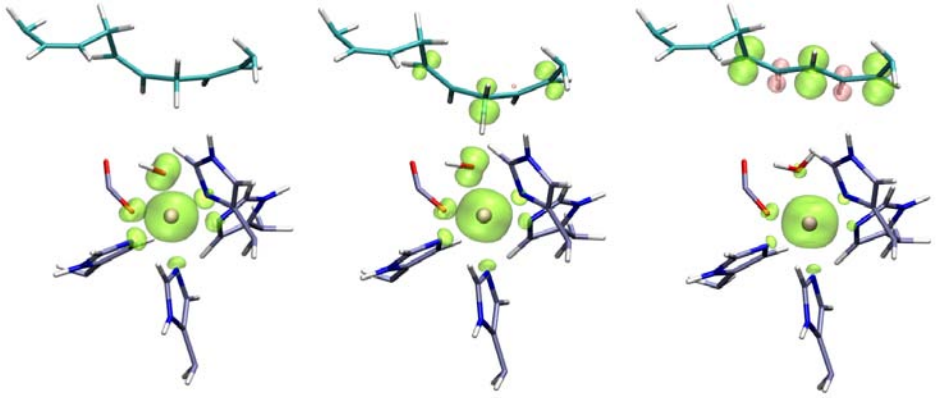
On the other hand, the NPA net charges on the H_{13} atom during the reaction (Figure 4.5b) are positive, and go from nearly 0.3 a.u. to 0.5 a.u., confirming that this atom is in fact a proton. All these analyses of the evolution of the electronic structure along the transfer show that the hydrogen abstraction is actually a Proton Coupled Electron Transfer (PCET) process,[34,176–178] in which the electron and proton are transferred in a concerted way to different acceptors: the proton is transferred to the OH^{-} group



(a) Spin density. **(b)** NPA charge.

Figure 4.5: (a) Spin densities (SDs) on atoms C_{11} (red), C_{12} (green), C_{13} (blue), C_{14} (yellow), C_{15} (magenta), H_{13} (cyan) and Fe (black), and (b) NPA charges on H_{13}, during H_{13} abstraction on structure I.

oxygen to produce water, while the electron is transferred from the C_{11}-C_{15} pentadiene group of AA to the Fe(III) to give Fe (II).



(a) Reactant **(b)** Transition state. **(c)** Product.

Figure 4.6: .Representation of SD surfaces. Green surface corresponds to α SD, and pink surface corresponds to β SD. Only QM atoms are shown. These pictures corresponds to structure I.

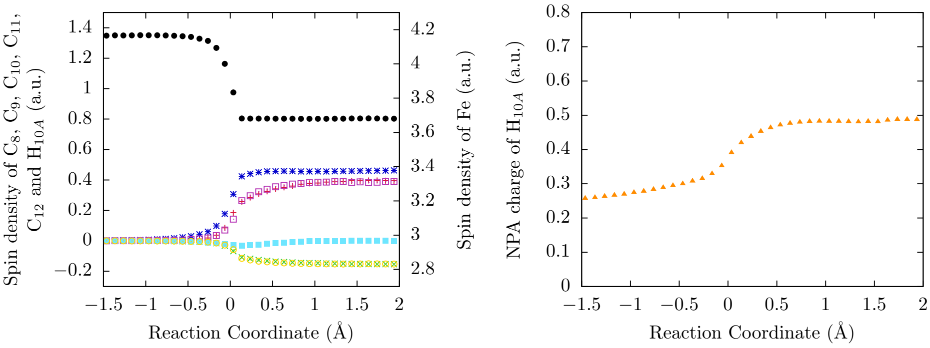
As for the evolution of the electronic structure and the geometry of the corresponding transition state structure, all the H_{13} abstraction processes analyzed (for structures I to XI) behave as the one in structure I. Very interestingly the geometries of all the TS structures around the shifting H_{13} are very similar (see Table 4.1), in spite of the existence of a wide dispersion of structures corresponding to reactants, as mentioned above. It has to be emphasized that, within the set of the 11 TS structures, the differences between the larger and the smaller values of the C_{13}-H_{13}, O-H_{13} and C_{13}-O distances, and the $\angle C_{13}$-H_{13}-O and $\angle H_{13}$-O-H angles are as small as 0.08

Å, 0.09 Å, 0.04 Å, 10.7° and 9.2°, respectively. These values turn out to be extraordinarily lower than the corresponding values within the set of the 11 structures I to XI, which are 0.02 Å, 1.55 Å, 1.44 Å, 57.4° and 69.4°, respectively. On the other hand, what actually makes a difference among the H_{13} abstractions is the height of the potential energy barriers, which range from 18.6 kcal/mol (structure VI) to 30.4 kcal/mol (structure X), and the reaction energies, ranging from -21.7 kcal/mol (structure VIII) to - 10.2 kcal/mol (structures III and X). These results suggest that the differences observed in the potential energy barriers are primarily due to the differences among the structures of the reactive minima rather than the transition states ones.

Following the same procedure described above, the analysis of the H_{10} abstraction on structure XII will be now presented. From the geometrical point of view the scenario is roughly identical to the one for abstraction of H_{13} on structure I. The transition state structure (see Table 4.2 and Figure 4.3) has distances between H_{10} and its donor atom (C_{10}-H_{10}), and between H_{10} and its acceptor atom (O-H_{10}) very similar (1.34 Å and 1.33 Å, respectively), and to reach the transition state the donor and acceptor atoms need to approach each other a bit more than 1 Å (from 3.78 Å to 2.67 Å), with the $\angle C_{10}$-H_{10}-O angle becoming almost linear, with a value of 175.5°, and the angle $\angle H_{10}$-O-H already adopting nearly the value corresponding to the water molecule finally formed (see Figure 4.3).

The potential energy barrier of H_{10} abstraction in structure XII is 24.1 kcal/mol (more than 2 kcal/mol higher than for abstraction of H_{13} on structure I), and the reaction is exoergic by 11.8 kcal/mol. The reactant is located at z = 1.76 Å, the TS at z = 0.01 Å, and the product at z = 2.24 Å. On the other hand, changing the corresponding atoms, the evolution of the NPA atomic populations evaluated from the spin density matches the H_{13} case (Figure 4.7a). Initially, an α SD appears on C_{10}, and delocalizes over C_8, C_{10}, and C_{12} as the H_{10} transfer progresses. The latter two carbon atoms will be the hit points of the molecular oxygen in the following step of the mechanism. The H_{10} atom behaves as a proton: its SD remains virtually zero and its NPA net charge is increasingly positive during the transfer. Therefore, the H_{10} abstraction is also a PCET process, with the electron being transferred from the C_8-C_{12} pentadiene group of AA (see Figure 4.7a). So, at this point it is clear that both H_{13} and H_{10} abstraction processes occur essentially with the same reaction mechanism.

Analyzing the rest of H_{10} abstractions we see that this process is feasible provided that the initial conditions in the reactive complexes are appropriate. Again, the initial noticeable dispersion in the configurational space within the set of the eight structures X to XVII, measured as the differences between the largest and the smallest values of the C_{10}-H_{10}, O-H_{10} and C_{10}-O distances, and the $\angle C_{10}$-H_{10}-O and $\angle H_{10}$-O-H angles (0.00 Å, 1.08



(a) Spin density. **(b)** NPA charge.

Figure 4.7: (a) Spin densities (SDs) on atoms C_8 (red), C_9 (green), C_{10} (blue), C_{11} (yellow), C_{12} (magenta), H_{10} (cyan) and Fe (black), and (b) NPA charges on H_{10}, during H_{10} abstraction on structure XII.

Å, 0.42 Å, 62.0° and 65.3°, respectively), is highly reduced within the set of the eight transition state structures (see Table 4.2), which geometries around the jumping H_{10} match quite well one another (the corresponding values of those differences are 0.06 Å, 0.08 Å, 0.06 Å, 5.0° and 12.2°, respectively). Thus, the important difference among the transition state structures lies on the height of the potential energy barriers (Table 4.3), which range from 22.7 kcal/mol (structure XIV) to 31.4 kcal/mol (structure XVII), an interval of energies somewhat higher than the one for the H_{13} abstraction.

Likewise, there exists a certain dispersion on the values of reaction energies, ranging from -16 kcal/mol (structure XVI) to -7.6 kcal/mol (structure XIII), an interval somewhat less exoergic than in the case of the H_{13} abstractions. In Tables 4.2 and 4.3 we have also included the results corresponding to the H_{10} abstraction in structure I, which, as previously mentioned, is suitable for an H_{13} abstraction, but not for an H_{10} abstraction because the initial O-H_{10} and C_{10}-O distances are as large as 5.83 Å and 5.85 Å, respectively. Very interestingly, even in this case the transition state structure converges to exactly the same zone of the configurational space as the other H_{10} abstractions, although, indeed, the potential energy barrier appears completely out of range (40.6 kcal/mol).

4.3.3 Comparison between H_{13} and H_{10} abstraction processes

At this point, both H_{13} and H_{10} abstraction processes can be compared. It is evident that only those reactive complexes that have a methylene hydrogen of AA close to the acceptor oxygen atom of the Fe(III)-OH^- cofactor are able to allow a thermally feasible hydrogen transfer. In the case of 15-LOX-1 only the reactive complexes with H_{13} or/and H_{10} near the acceptor oxygen atom are thermally accessible. As a consequence, only the H_{13} or H_{10} abstractions are possible. Our results have shown that, starting from the convenient reactive structures, the reaction mechanisms for both H_{13} and H_{10} abstractions are identical, with transition state structures that involve matching geometries around the shifting hydrogen. So, we have confirmed the suggestion by Holman and coworkers about the similarity of the mechanisms and transition states for both abstractions, fact that explains the comparable experimental ${}^{D}k_{\mathrm{cat}}$ and ${}^{D}k_{\mathrm{cat}}/K_{\mathrm{M}}$ values for the two carbon centers in human human 15-LOX-1.

The main difference between the H_{13} and H_{10} abstractions lies on the respective potential energy barriers. Figure 4.8 displays the potential energy barriers for the 20 hydrogen abstractions studied in this chapter as a function of the corresponding H-O (H_{13}-O or H_{10}-O) distance in each reactive structure.

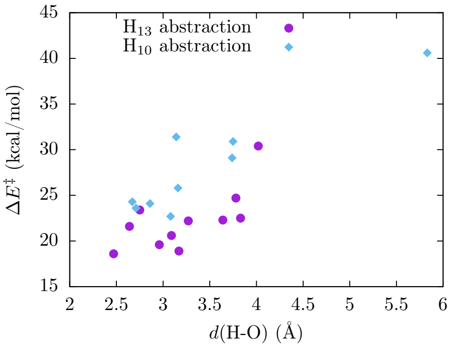


Figure 4.8: Potential energy barriers versus initial H-O distance for H_{13} and H_{10} abstractions

Firstly, it can be seen that in some way the barriers tend to increase as the H-O distance becomes larger, although it is not possible to establish a precise correlation between the values of the potential energy barriers and the initial H-O distances. It is clear that, being equivalent TS state structures, the barrier height depends on the geometry of the reactive structures, although many additional parameters define it. Note the huge value of 40 kcal/mol for the H_{10} abstraction in structure I, which is not suitable for that abstraction (the O-H_{10} distance is 5.83 Å), but for the H_{13} abstraction (the O-H_{13} distance is 2.64 Å). The difference in the energy barriers is the reason why most of the products generated by the reaction of 15-LOX-1 with AA comes from the hydrogen abstraction from C_{13}, in spite of the fact that both H_{13} and H_{10} are spatially accessible for the abstraction.

So far we have seen that both processes occur with essentially the same reaction mechanism, and in both cases there are structures with the transferred hydrogen atom close enough to the oxygen acceptor. However, H_{13} abstraction process is more favorable than H_{10} abstraction. So here arises the question of why apparently H_{13}/H_{10} identical abstraction mechanisms and very similar transition state structures occur with somewhat different potential energy barriers, in this way preserving the exquisite regiospecificity of 15-LOX-1. To make a more detailed comparison between the potential energy barriers, the total QM/MM energy has been decomposed as follows:

$$E_{\text{Total}} = E_{(\text{QM,MM})} + E_{(\text{MM,QM})} \tag{4.3}$$

where $E_{(\text{QM,MM})}$ is the sum of the QM energy and the electrostatic interaction between the QM and MM subsystems, and $E_{(\text{MM,QM})}$ is the sum of the MM energy and the bonded and van der Waals interactions between the QM and MM subsystems. Results of this decomposition are listed in Table 4.3. Two facts can be highlighted. First, similar or even almost identical potential energy barriers can be decomposed in different ways. In this sense, for structures II, IV and XI, in which H_{13} abstraction occurs, the corresponding potential energy barriers are of 22.3, 22.2 and 22.5 kcal/mol, respectively, but they can be decomposed into 18.2, 22.5 and 20.6 kcal/mol $E_{(\text{QM,MM})}$ components, and 4.1, -0.3 and 1.9 kcal/mol for the $E_{(\text{MM,QM})}$ components, respectively. This result is a direct consequence of the diversity of the initial reactant structures and the similarity of the TS ones: reactant structures will have to evolve in a different way to reach the corresponding TS structure.

The second fact is that, when the $E_{(\text{QM,MM})}$ component is under 20 kcal/mol in a few cases, the $E_{(\text{MM,QM})}$ component turns out to be large. This is specially true for the H_{10} abstraction. In some structures (e.g VIII, XIV, XV) when the QM part contributes with the lowest values, the MM part conversely contributes with the highest (which could be due to either the MM part itself, or to the non-electrostatic interaction with the QM part). This indicates that, in some initial structures, the better the QM part is ready to react, the worse prepared is the corresponding MM part.

To get a deeper insight on the differences between the H_{13}/H_{10} abstractions we will focus now on structure XI (see Figure 4.9), where the C_{13}-O and C_{10}-O distances (3.83 Å and 3.75 Å, respectively) are practically equal, which seems ready to allow any of the two possible hydrogen transfers. The existence of such type of structures shows that a highly regiospecific enzyme like 15-LOX-1 (or human 15-LOX-1) is able to spatially accommodate the two carbon positions (C_{13} and C_{10}), even at once, in a similar way with respect to the acceptor Fe(III)-OH^- moiety. However, the potential energy barrier for the H_{10} abstraction turns out to be more than 8 kcal/mol higher than the corresponding to the H_{13} abstraction (see Table 4.3), the difference mostly arising from the $E_{(\text{MM,QM})}$ component (6.8 kcal/mol). Likewise, the H_{13} abstraction is 5.7 kcal/mol more exoergic than the H_{10} abstraction.

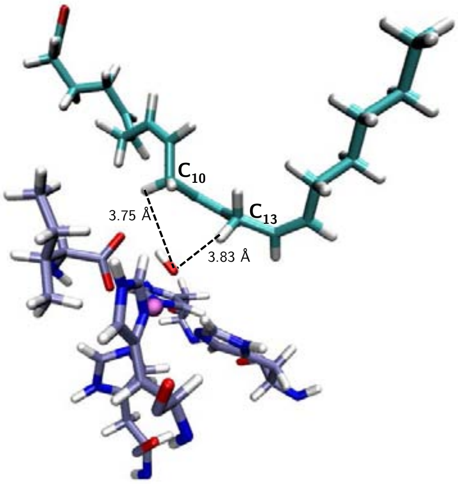


Figure 4.9: Reactant of structure XI. Only AA and Fe coordination sphere are shown. Distances $d(H_{13}\text{-}OH)$ and $d(H_{10}\text{-}O)$ are indicated.

As the H_{13} hydrogen transfer progresses in a given reactive structure, C_{13} tends to an sp^2 hybridization and the moiety corresponding to the nascent AA-pentadienyl radical approaches to planarity by conjugation in a π system consisting of 5 electrons (the unpaired electron originally at C_{13} and 4 π electrons of the two adjacent double bonds) delocalized over 5 carbon atoms (from C_{11} to C_{15}). Exactly the same happens for the H_{10} abstraction, now being C_{10} the atom tending to an sp^2 hybridization, and being from C_8 to C_{12} the carbon atoms forming the planar AA-pentadienyl radical. The larger the geometrical change from an initially non-planar structure to a planar pentadienyl and, specially, the greater steric hindrance to that motion, the higher the contribution to the potential energy barrier.

Figure 4.10 shows the geometric evolution of the active site of structure XI along the hydrogen transfer, using a color scale to visualize the progress from the reactant to the product. As for the H_{13} abstraction, the dihedral angles C_{11}-C_{12}-C_{13}-C_{14} and C_{12}-C_{13}-C_{14}-C_{15} at structure XI are 188.5° and 122.8°, respectively, whereas they become 165.7° and 176.7°, respectively, at the radical product. This change involves a small motion of AA (see Figure 4.10 (a)), which does not practically affect the region of the C_8-C_9 double bond, and that occurs in a region of the 15-LOX-1 cavity without significant steric hindrance to the motion of AA. However, the dihedral angles C_8-C_9-C_{10}-C_{11} and C_9-C_{10}-C_{11}-C_{12} at structure XI are 105.7° and 109.9°, respectively, whereas they become -0.2° and 174.8°, respectively, at the radical product after the H_{10} abstraction. This change produces a broad motion of AA (see Figure 4.10 (b)), fundamentally in the region of the C_8-C_9 double bond (from C_7 to C_{10}), where the side chains of Leu597 and Ile663

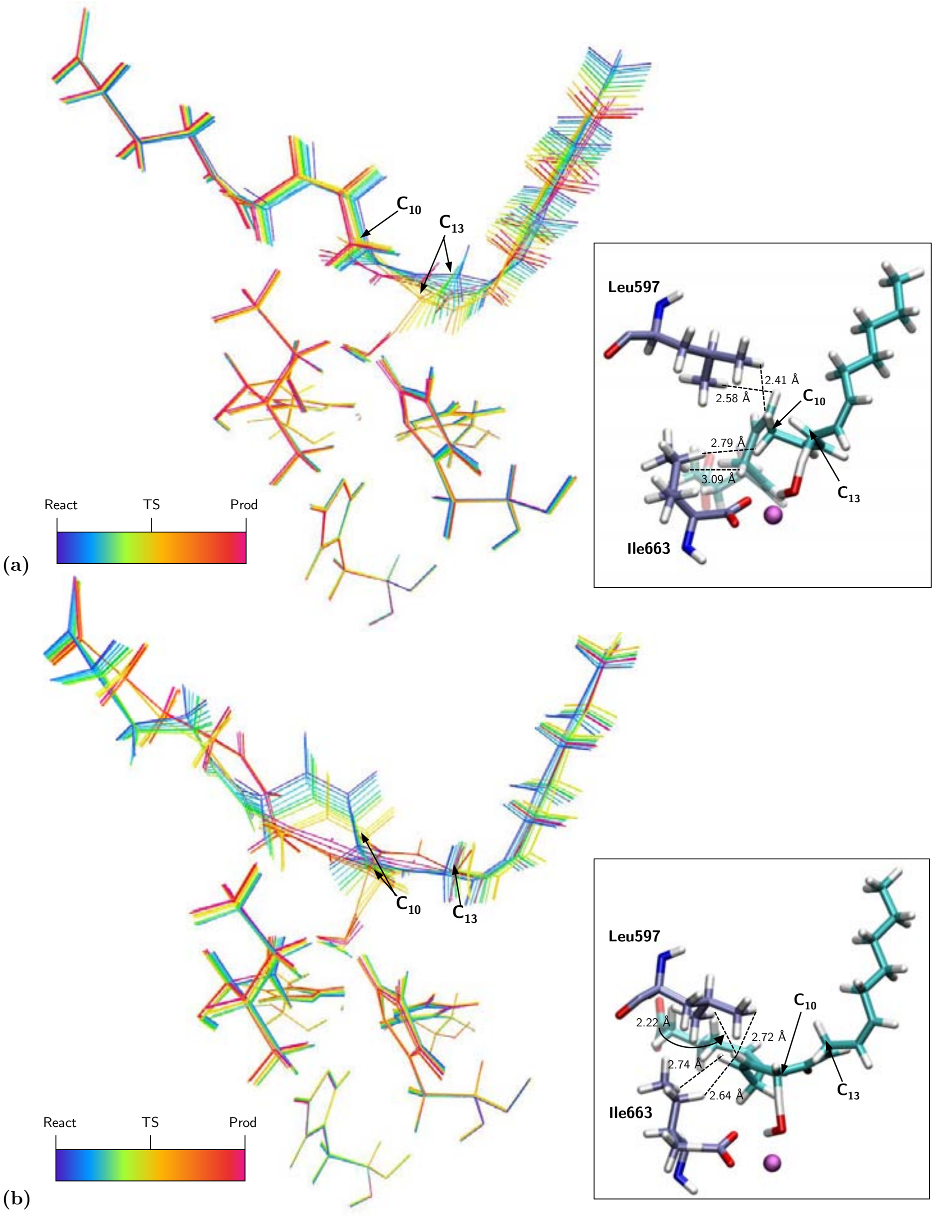


Figure 4.10: Representation of the structures along the reaction coordinates for H_{13} (a) and H_{10} (b) abstractions in structure XI. Color scale is used going from blue (reactants) to red (products). Boxes on the right show the closest distances from Ile663 and Leu597 to arachidonic acid in the transition state structures.

(specially the ethyl group attached to its β-carbon) significantly hinder this evolution. The change of the dihedral angle C_8-C_9-C_{10}-C_{11} pushes the region of the C_8-C_9 double bond against Leu597 and Ile663.

For instance, it can be seen in the inset of Figure 4.10 (b) that the two hydrogen atoms, H_8 and H_9, in the C_8-C_9 double bond become very close to some hydrogen atoms of Leu597 and Ile663 (as much close as 2.22 Å) at the H_{10} transition state structure. In contrast, the closest hydrogen atoms of AA, H_7, H_9 and H_{10}, to the hydrogen atoms of Leu597 and Ile663 practically do not move and keep somewhat farther along the H_{13} abstraction (see the transition state structure in the inset of Figure 4.10 (a)).

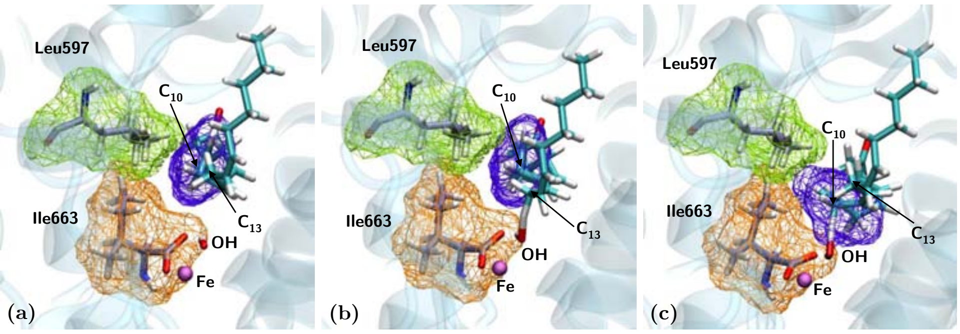


Figure 4.11: Representation of the van der Waals surfaces of Ile663 (orange), Leu597 (green) and the AA atoms between C_7 and C_{10} (blue). All the pictures correspond to structure XI: (a) reactant, (b) TS of H_{13} abstraction, (c) TS of H_{10} abstraction.

The compression along the hydrogen transfer of the C_8-C_9 double bond region of AA against Leu597 and Ile663 when the H_{10} abstraction (but not for the H_{13} abstraction) occurs can be better seen in Figure 4.11, where the van der Waals surfaces of these critical regions are represented for the initial reactive structure XI, the H_{13} abstraction transition state structure, and the H_{10} abstraction transition state structure. It is worth noting that the van der Waals region around the C_8-C_9 double of AA at the H_{13} transition state structure (Figure 4.11 (b)) matches the one at the reactant structure XI (Figure 4.11 (a)), whereas the mobile van der Waals region around the C_8-C_9 double of AA at the H_{10} transition state structure (Figure 4.11 (c)) is about to penetrate the van der Waals regions of both Leu597 and Ile663, and evolves just to this limit. The energy penalty of the interpenetration of two van der Waals regions is quite huge. To avoid that, the necessary evolution of the dihedral angles C_8-C_9-C_{10}-C_{11} and C_9-C_{10}-C_{11}-C_{12} requires the intense motion of the AA head, as seen in Figure 4.10 (b). This is the cause of the higher value of the $E_{(\mathrm{MM,QM})}$ (Table 4.3) for the H_{10} abstraction in comparison to the H_{13} abstraction in structure XI.

A more detailed analysis of the different energy terms clarifies this point. The van der Waals contributions of Leu597 to the potential energy barriers for the H_{13} and H_{10} abstractions are 0.20 and 0.17 kcal/mol, respectively. Likewise, the respective values for Ile663 are 0.19 and -0.05 kcal/mol. All of those values turn out to be very small, therefore corroborating that the van der Waals regions in the case of the H_{10} abstraction stop just before collision. This fact forces the distortion of the AA head to allow the change of the dihedral angles C_8-C_9-C_{10}-C_{11} and C_9-C_{10}-C_{11}-C_{12} to occur, as the system overcomes other geometrical changes to avoid the collision. The distortion pays an energy penalty, but lower than the one that would be associated with the interpenetration of the van der Waals surfaces. Therefore, the motion of the C_1-C_6 head of AA implies an electrostatic contribution of 7.5 kcal/mol to the H_{10} potential energy barrier, but just of -0.75 kcal/mol for the H_{13} case.

This disfavorable electrostatic contribution to the H_{10} abstraction energy barrier is due in part to the weakening of the Arg403-carboxylate of AA interaction. Figure 4.12 represents Arg403 interacting with the COO^- group in reactant, TS of H_{13} abstraction, and TS of H_{10} abstraction, showing the distances between the interacting atoms, and their relative orientation. For H_{10} abstraction, the distortion of the Arg403-COO interaction costs 2.3 kcal/mol, in contrast with the case of H_{13} (-0.14 kcal/mol). In addition, the contribution to the energy barrier of the deformation of the C_1-C_6 bond angles is 2.6 kcal/mol for H_{10}, but only -0.43 kcal/mol for H_{13}. To summarize, we conclude that the subtle steric hindrance by Leu597 and Ile663 is mainly responsible for the difference between the H_{13}/H_{10} potential energy barriers and reaction energies in structure XI. This steric hindrance, as usual, does not produce an increase of the energy associated with the van der Waals terms, but provokes additional motions in other parts of the system, just to circumvent the huge penalty that should be paid for the interpenetration of two van der Waals regions.

In general, the H_{10} abstraction is disfavored in comparison with the H_{13} abstraction due to the steric hindrance by Leu597 or Ile663 or both at once. In some structures even Ile173 contributes to that hindrance. This gives rise to somewhat higher potential energy barriers for the H_{10} abstraction than for the H_{13} abstraction, and so explaining the strict regiospecificity of the first step of the hydroperoxidation of AA by 15-LOX-1 and human 15-LOX-1, despite the fact that both the hydrogen abstractions from C_{13} and from C_{10} would seem feasible. As for Ile663, in all the 9 structures selected as suitable for the H_{13} abstraction the ethyl group bonded to the β-carbon of Ile663 is the one that faces the C_8-C_9 double bond region of AA (Figure 4.13 (a) shows the example of structure I). Conversely, only structures selected for the H_{10} abstraction where the shorter methyl group of that β-carbon (structures XII to XVI) is the one oriented towards the C_8-C_9 double bond region of AA give rise to the H_{10} abstraction with an energy barrier smaller than 26 kcal/mol (see structure XII in Figure 4.13 (b). The hindrance by the methyl group is indeed lower than by the ethyl group, but the effect is still significant.

In order to compare both reactive processes and also with experi-

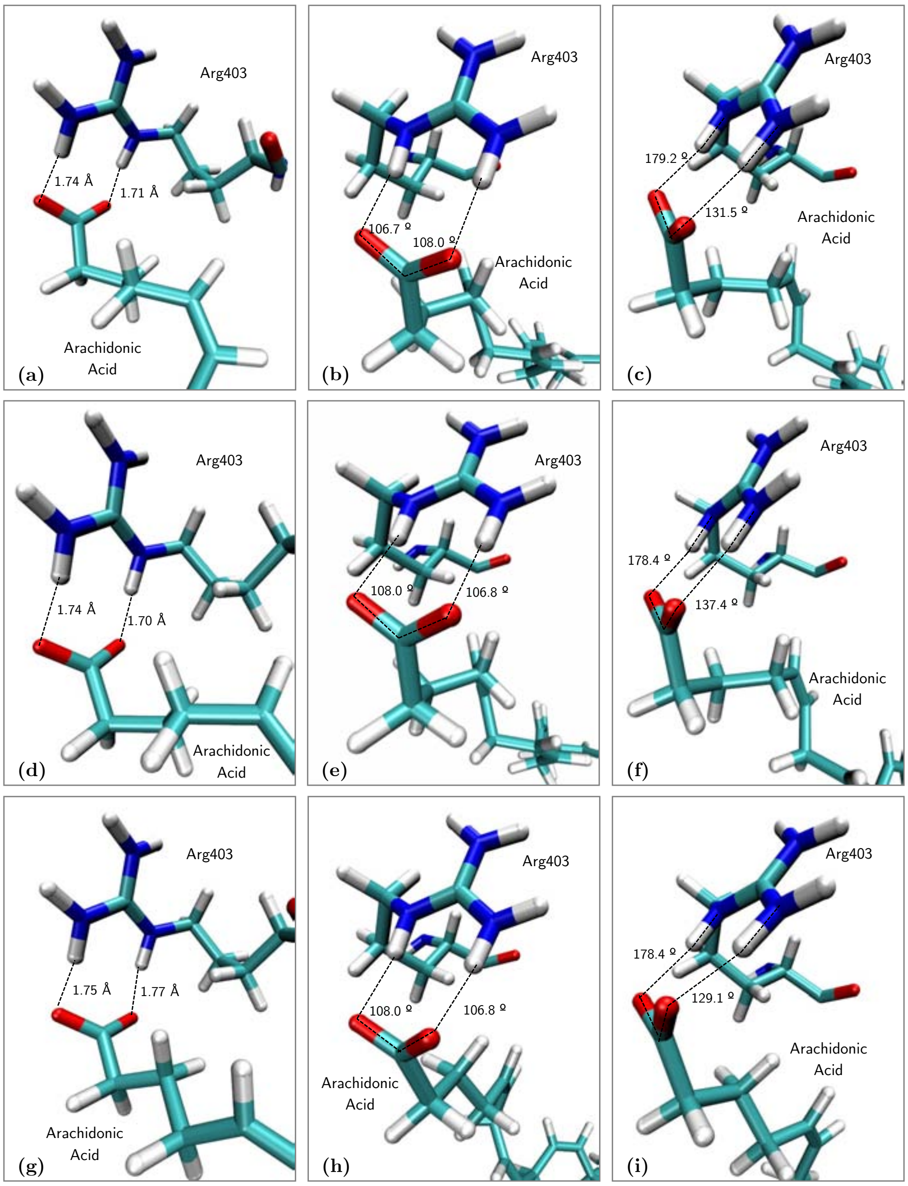


Figure 4.12: Representation of Arachidonic Acid carboxylate end and Arg403 of structure XI in: (a), (b) and (c) the reactant of structure XI; (d), (e), (f) TS of H_{13} abstraction and, (g), (h), (i) TS of H_{10} abstraction, showing: (a), (d), (g) O_{AA}-H_{R403} distances; (b), (e), (h) $C_{1,AA}$-O_{AA}-H_{R403} angles; (c), (f), (i) $C_{1,AA}$-O_{AA}-H_{R403}- N_{R403} dihedrals.

mental results a magnitude like a kind of average potential energy barrier should be provided. It is worth mentioning that the high dispersion in the $\Delta E^{\ddagger}_{\mathrm{Total}}$ observed for each reactive process (H_{13} and H_{10} abstraction) which is a direct consequence of the high dispersion in the reactant minima geometries, as the corresponding TSs present a similar structure in both processes.

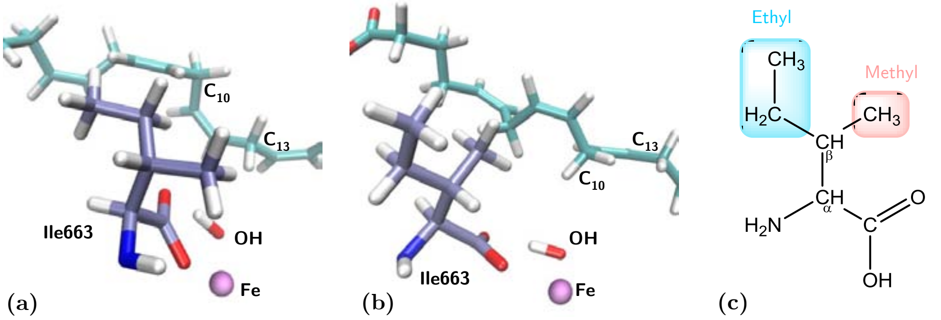


Figure 4.13: Representation of the AA substrate and the terminal Ile663 in the reactants of structure I (a) and structure XII (b). (c) chemical structure of Isoleucine, highlighting the ethyl (blue) and methyl (groups) of its side chain. It is observed that in structure I, the ethyl group of Ile663 side chain faces the AA C_8-C_9 double bond, whereas in structure XII is the methyl group of Ile663. The position of these two groups can be alternated by the rotation around the Cα-Cβ bond of Isoleucine.

Thus, assuming that each structure reacts with a local rate constant that fulfills the conventional transition state theory and that free energy and potential energy barriers present similar values,[179] the overall rate constant could be calculated by an exponential average using the equation:[180, 181]

$$\Delta E^{\ddagger}_{\text{Av}} = -RT \ln \left\{ \frac{1}{n} \sum_{i=1}^{n} \exp\left(\frac{-\Delta E^{\ddagger}_{i}}{RT} \right) \right\} \tag{4.4}$$

where R is the gas constant, T is the absolute temperature, $\Delta E^{\ddagger}_{i}$ is the potential energy barrier for each structure (Table 4.3) and n is the total number of structures. Using Eq. 4.4 the average potential energy barrier for H_{13} abstraction is of 19.6 kcal/mol, whereas for H_{10} abstraction the average potential energy barrier results in 23.6 kcal/mol at $T = 300$ K. Thus, there is a difference of 4.0 kcal/mol, which agrees well with the value of 2.1 kcal/mol derived from the experimental result of the C_{13}:C_{10} ratio of 97:3 for this system, employing the conventional transition state theory and assuming kinetic control.[1]

The accuracy of the calculated potential energy barriers has been assessed by means of single point energy calculations to examine the sensitivity of the results to the choice of functional and basis set. The 40 new potential energy barriers for H_{13} and H_{10} abstractions are given in Table 4.4. It can be observed that only small changes of the potential energy values are

---

[1] According to the conventional Transition State Theory, the rate constant is related with the phenomenological free energy by: $k(T) = \frac{k_B T}{h} K^0 e^{-\Delta G^{\ddagger 0}/RT}$. Assuming kinetic control, the ratio of the rate constants for H_{13} and H_{10} abstraction should be the same as the ratio for the product formation, 97:3. Thus, it results in a difference of 2.1 kcal/mol at $T = 300$ K.

Table 4.4: Potential energy barriers at different levels of theory for H_{13} and H_{10} abstractions, and ZPE correction at B3LYP/[LANL2DZ(Fe), 6-31G(d)(rest)] level. All values are given in kcal/mol. All the structures were optimized at the B3LYP/[LANL2DZ(Fe), 6-31G(d)(rest)] level.

Structure	Abstraction	B3LYP/B1[a]	B3LYP-D2[b]/B1	B3LYP/B2[c]	TPSSH/B1	ZPE-corrected
I	H_{13}	21.6	20.7	21.5	21.4	18.1
II	H_{13}	22.3	19.2	23.3	22.3	18.7
III	H_{13}	19.6	16.3	21.9	20.3	16.0
IV	H_{13}	22.2	19.0	23.4	23.0	18.7
V	H_{13}	24.7	22.7	25.9	24.6	21.4
VI	H_{13}	18.6	17.2	18.7	18.5	14.8
VII	H_{13}	23.4	22.1	23.9	23.5	19.8
VIII	H_{13}	18.9	16.4	18.9	19.3	15.3
IX	H_{13}	20.6	17.8	20.7	20.4	17.1
X	H_{13}	30.4	24.6	32.2	31.2	26.5
XI	H_{13}	22.5	19.3	23.1	22.9	19.1
XII	H_{10}	24.1	21.6	24.9	24.0	20.3
XIII	H_{10}	23.6	21.4	23.6	23.4	19.8
XIV	H_{10}	22.7	21.4	23.4	22.7	18.9
XV	H_{10}	24.3	22.5	25.3	24.6	20.4
XVI	H_{10}	25.8	23.6	26.2	25.6	22.0
XVII	H_{10}	31.4	31.3	32.0	31.5	27.3
I	H_{10}	40.6	39.2	42.2	40.9	36.9
X	H_{10}	29.1	26.8	30.3	29.3	25.5
XI	H_{10}	30.9	29.1	31.7	31.1	27.7

[a] B1: LANL2DZ(Fe), 6-31G(d)(rest) basis set
[b] Grimme-type D2 energy dispersion correction
[c] B2: LANL2DZ(Fe), Ahlrichs-vtz(rest) basis set

obtained by using the Ahlrichs-vtz triple-ζ basis set[182] and the TPSSH functional.[183] Thus, the difference of the exponential average potential energy barriers (calculated according to equation 4.4) for H_{13} versus H_{10} abstraction are 4.5 kcal/mol with the Ahlrichs-vtz basis set, and 3.9 kcal/mol with the TPSSH functional, in good agreement with the double-zeta B3LYP result given above. As the inclusion of dispersion energy in DFT/MM calculations in enzyme-catalyzed reactions has been shown to be important, we have also recomputed the 20 potential energy barriers by means of single-point energy calculations using the B3LYP-D2 functional with Grimme-type corrections[184] (see Table 4.4).The inclusion of dispersion effects somewhat decreases the potential energy barriers of H_{13} and H_{10} abstractions but without a significant change in their relative barriers being the dispersion-corrected average potential energy difference of 4.7 kcal/mol. In Table 4.4 we have also included the zero-point energy (ZPE) corrected B3LYP/[LANDZ2D(Fe), 6-31G(d) (rest)] barriers for all the H_{13} and H_{10} abstractions. It can be observed that the inclusion of the ZPE lowers the H-abstraction barriers but it does not modify the exponential average energy difference of 4.0 kcal/mol favoring H_{13} abstraction versus H_{10} abstraction.

4.3.3.1 QM/MM free energy perturbation calculations

To estimate the importance of the thermal and entropic effects in the hydrogen abstraction reaction pathway, the free energy along the reaction coordinate has been evaluated by means of the QM/MM-FEP method for the hydrogen abstraction step from AA acid. According to this methodology, the QM part is frozen at each window of the reaction coordinate (in total 28 FEP windows were used), and only the computationally less demanding MM part is sampled. In this way, the contributions to the free energy of the environment due to the MM part and the QM/MM interactions are included. As indicated in the Methodology section, the density of the fixed QM part is not only frozen but approximated by ESP charges when calculating its interaction with the MM part according to ChemShell protocol.

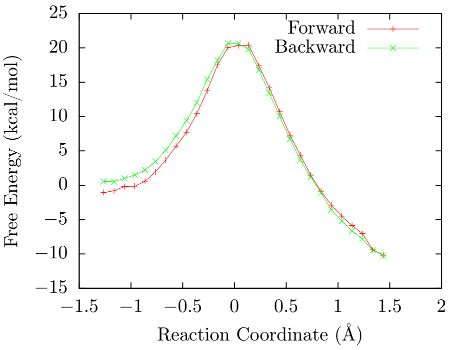


Figure 4.14: Forward and backward free energy profiles along the H_{13} abstraction in structure VI.

The forward and backward free energy profiles are shown in Figure 4.14. This profile corresponds to the H_{13} abstraction process in structure VI, which occurs with a potential energy barrier of 18.6 kcal/mol (see Table 4.3). The data for the forward and backward directions agree quite well with only minor hysteresis effects. The forward free energy barrier turns out to be 21.5 kcal/mol, and the backward free energy barrier is of 20.1 kcal/mol, giving an average free energy barrier of 20.8 kcal/mol. The difference with respect to the corresponding potential energy barrier is of 2.2 kcal/mol due to the sampling of the environment, which indicates that the thermal and entropic contributions are small. Therefore, the size of these contributions is not big enough to significantly modify the main findings based on the considerations of the potential energy barriers.

4.3.4 Study of the hydrogen abstraction step in LA

As indicated in the Introduction Section, AA is the most common substrate in mammalian LOXs. However, there are mammalian LOXs that present dual substrate specificity and 15-LOX-1 (and also 15-LOX-2) can also catalyze the peroxidation of LA. For comparative purposes, the theoretical study of the hydrogen abstraction from LA has been also included in this Chapter, in the case of rabbit 15-LOX-1. Note that just H_{11} can be abstracted from LA, as C_{11} is the only bisallylic methylene unit in this substrate.

4.3.4.1 Potential Energy Profiles

The structures of the LA:15-LOX-1 Michaelis complex were taken from MD simulations done previously in our research group.[175] The selection criterion to filter out the pre-catalytic complexes was similar to the one employed for AA (see above): $d(H_{11}\text{-OH}) < 3.0$ Å and $d(C_{11}\text{-OH}) > d(H_{11}\text{-OH})$. Next, following the same procedure as for AA, the selected structures were prepared to perform the hydrogen abstraction step. The corresponding potential energy barriers, and geometric parameters of the reactant minima and TS structures, are listed in Table 4.5.

Table 4.5: Potential energy barriers (in kcal/mol) and main geometric parameters (distances in Å and angles in degrees) of reactants and TSs of the H_{11} abstraction in LA.

Structure	$\Delta E^\ddagger$	$d(H_{11}\text{-OH})$		$d(C_{11}\text{-}H_{11})$		$d(C_{11}\text{-OH})$		$\angle(C_{11}\text{-}H_{11}\text{-O})$		$\angle(H_{11}\text{-O-H})$	
		R	**TS**	**R**	**TS**	**R**	**TS**	**R**	**TS**	**R**	**TS**
XVIII	14.6	2.25	1.40	1.10	1.28	3.32	2.65	162.5	163.5	94.4	99.4
XIX	21.3	3.65	1.41	1.10	1.28	3.97	2.68	98.7	169.2	41.2	91.7
XX	23.1	3.69	1.39	1.10	1.28	3.88	2.67	92.1	172.2	48.2	93.3
XXI	23.2	3.25	1.38	1.09	1.29	3.65	2.65	102.9	167.2	29.0	88.5
XXII	26.0	2.76	1.37	1.10	1.31	3.68	2.68	141.1	171.4	90.6	100.3
XXIII	13.6	2.19	1.39	1.10	1.27	3.24	2.64	157.9	164.7	92.3	98.6
XXIV	26.4	2.93	1.36	1.10	1.32	3.59	2.67	124.6	171.2	89.1	100.6
XXV	28.8	2.83	1.37	1.10	1.32	3.77	2.69	143.5	171.5	94.0	100.2
XXVI	23.0	2.78	1.39	1.10	1.29	3.71	2.68	143.0	174.3	87.7	98.7
XXVII	15.9	2.29	1.38	1.10	1.30	3.32	2.65	153.9	164.0	97.5	99.3
XXVIII	15.4	2.12	1.34	1.10	1.31	3.19	2.62	163.2	165.3	92.4	99.2
XXIV	16.0	2.33	1.37	1.10	1.29	3.35	2.64	153.7	163.9	95.5	98.9
XXX	19.1	2.73	1.40	1.10	1.28	3.62	2.67	138.2	169.2	51.2	94.2

It can be observed that, like in the case of AA, there is a great dispersion of potential energy barriers in the reaction with LA. For instance, the lowest potential energy barrier, which corresponds to structure XXIII, has a value of 13.6 kcal/mol, whereas structure XXV presents a potential energy barrier of 28.8 kcal/mol. Again, the differences observed in the potential energy barriers can be attributed to the differences in the reactant minima structures rather than to the TS ensembles (see distances and angles in Table 4.5). Figure 4.15 shows the reactant and TS geometries of structure XVIII. Also, as has been highlighted before, there is not a precise correlation between the H_{11}-OH distances and the energy barriers when those distances are short enough. However, the lowest potential energy barriers correspond to those profiles initiated from reactive structures with the shorter H_{11}-OH distances (structures with such distance < 2.5 Å present an energy barrier < 16 kcal/mol).

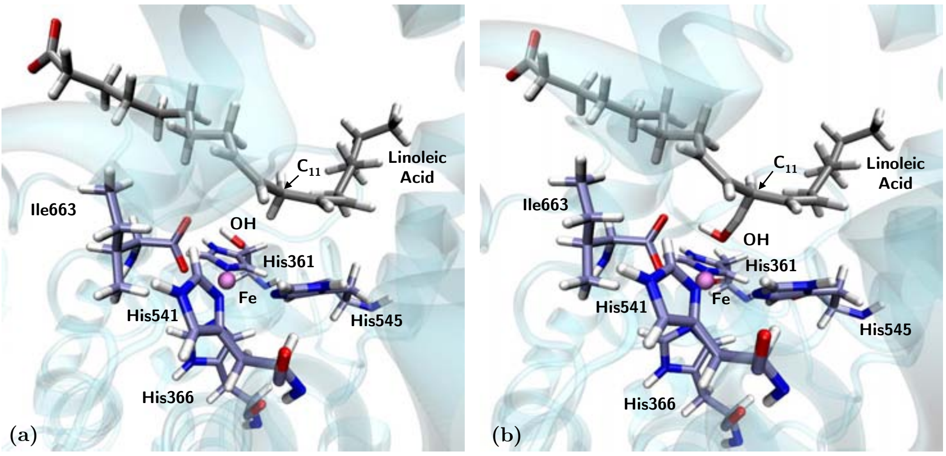


Figure 4.15: Representation of the reactant minimum (a) and TS structure (b) in structure XVIII.

Using equation 4.4 the exponential average energy barrier has been calculated for the hydrogen abstraction from LA. At $T = 300$ K, it results in 15.0 kcal/mol, which is 4.6 kcal/mol lower than the corresponding value of the H_{13} abstraction of AA (19.6 kcal/mol). Thus, this result agrees with the faster peroxidation of LA versus AA by 15-LOX-1, taking into account that the hydrogen abstraction step is the rate-limiting one.

4.3.4.2 QM/MM free energy perturbation calculations

Like in the case of the AA reaction, to estimate the contributions of the thermal and entropic effects in the hydrogen abstraction reaction pathway, the free energy along the reaction coordinate has been evaluated for the hydrogen abstraction step from LA acid using the QM/MM-FEP method. A total of 60 FEP windows were sampled along the H_{11} abstraction reaction coordinate.

The forward and backward free energy profiles are shown in Figure 4.16 for the H_{11} abstraction pathway in structure XVIII, which occurs with a potential energy barrier of 14.6 kcal/mol. Again, the data for the forward and backward directions agree quite well, with only minor hysteresis effects.

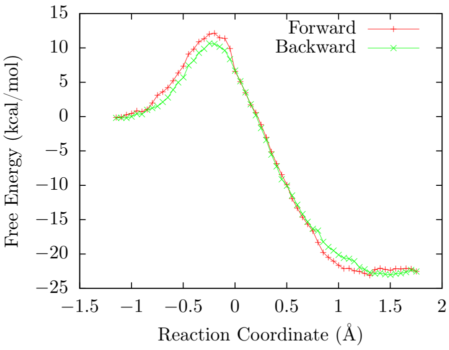


Figure 4.16: Forward and backward free energy profiles along the H_{11} abstraction from LA.

The forward free energy barrier is of 12.1 kcal/mol, and the backward free energy barrier of 10.8 kcal/mol; then, an average free energy barrier of 11.4 kcal/mol is obtained. Compared with the QM/MM potential energy barrier, this free energy barrier gives a contribution of -3.2 kcal/mol to the barrier due to the sampling of the environment. _

4.3.4.3 Comparison between AA and LA binding modes

In previous studies in our research group,[28,175] it was demonstrated that 15-LOX-1 presents an active site cavity with sufficient volume to accommodate the longer AA substrate as well as the shorter LA, and still have both interacting with Arg403. The comparison of the QM/MM minimum energy structures of AA:15-LOX-1 and LA:15-LOX-1 complexes can give more molecular insights on the way both fatty acid substrates bind into the active site for catalysis. Figure 4.17 shows the overlay of the reactant minimum of structure XVIII, with two different minima of the AA:15-LOX-1 complex (structures I and II).

It can be observed that C_{11} of LA is located around the same position as C_{13} of AA or even deeper into the active site, although LA is two carbon atoms shorter than AA. In this similar location, the bisallylic methylene group at C_{11} of LA is not surrounded by any nearby residue side chain, like occurs to the bisallylic methylene group at C_{13} in AA. In this structural comparison, it has to be taken into account that, besides the different number of carbon atoms of the fatty acid chain, AA has three double bonds within the carbon chain between C_1 and C_{13}, whereas LA only has one double bond between C_1 and C_{11}. As the number of double bonds increases between C_1 and C_n, the C_1-C_n distances tend to be shorter. For instance, the C_1-C_{11} distance of structure XVIII of 11.5 Å, is nearly the same as the C_1-C_{13} distance for AA in structure I (Figure 4.17a). Moreover, the par-

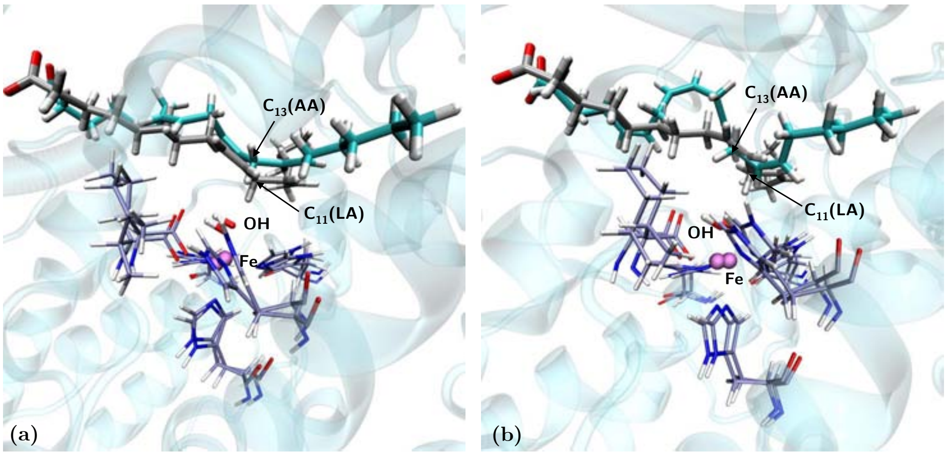


Figure 4.17: Overlay between the AA:15-LOX-1 (AA depicted in cyan) and the LA:15-LOX-1 (LA depicted in gray) complexes. (a) shows the overlay between the reactant minima of structures I (of the AA:15-LOX-1 complex) and XVIII (of the LA:15-LOX-1 complex), and (b) shows the overlay between the reactant minima of structures II (AA:15-LOX-1) and XVIII. Labels of the Fe-coordination protein residues are not shown for clarity.

ticular orientation of the three double bonds of AA in structure II (Figure 4.17b) reduces the C_1-C_{13} distance down to 10.4 Å. These structural considerations explain how 15-LOX-1 can accommodate the different reactive bisallylic methylene groups of the two substrates in a similar location, able to initiate the reactive process.

4.4 Conclusions

In the work presented in this Chapter quantum mechanics/molecular mechanics methods have been employed to study the mechanism of the hydrogen abstraction reaction from arachidonic acid (AA) catalyzed by 15-LOX-1, an enzyme that has approximately 80% sequence identity with human 15-LOX-1. These enzymes exhibit a very high regiospecificity, in such a way that most of the products arise from the initial hydrogen abstraction from the carbon C_{13} of AA.(17-19) It has been assumed that this regiospecificity of the different mammalian lipoxygenase isoenzymes is due to the depth and width of the substrate binding pocket and the head/tail orientation of the incoming AA, in such a way that each lipoxygenase enzyme would abstract a hydrogen atom from essentially a unique position, the only one that is able to be close to the oxygen acceptor atom of the Fe(III)-OH$^-$ cofactor. However, unexpected experimental results for human 15-LOX-1 have suggested that the hydrogen abstractions from two structurally different positions (C_{13} and C_{10}) follow a comparable mechanism(18-20) with similar transition state structures, which does not seem very compatible with the known exquisite regiospecificity of this enzyme.

The results presented here show a wide dispersion of geometries of the AA:15-LOX-1 Michaelis complexes able to transfer a hydrogen atom to the oxygen acceptor of the Fe(III)-OH^- cofactor. There are structures capable to transfer H_{13}, H_{10}, or even both, so in principle, both processes are feasible. Very interestingly, transition state structures for H_{13} and H_{10} abstraction present very similar geometry around the shifting hydrogen, thus, the dispersion of energy barriers can be attributed to the dispersion of the initial reactant structures. Even more, when comparing the H_{13} and H_{10} abstractions, we see that the reaction mechanisms for both are identical (proton-coupled electron transfer processes), with transition state structures that match their geometries around the shifting hydrogen. This confirms the suggestion by Holman and co-workers(18-20) about the similarity of the mechanisms and transition states for both abstractions.

The exponential average potential energy barriers we have obtained are 19.6 kcal/mol for the H_{13} abstraction and 23.6 kcal/mol for the H_{10} abstraction. This difference of 4.0 kcal/mol agrees well with the value of 2.1 kcal/mol derived from the experimental C_{13}:C_{10} ratio of 97:3 for 15-LOX-1. Steric hindrance by residues Leu597 and Ile663 turns out to be mainly responsible for that difference. As for the second residue, the catalytic iron is coordinated by the main-chain carboxylate of the C terminal Ile663, which contributes to determine the electronic density on this iron atom and so regulates its redox potential. In addition, that carboxylate forms a hydrogen bond with the hydrogen atom of the OH^- group of the Fe(III)-OH^- cofactor, so helping the right positioning of that OH^-. However, any of the amino acids with aliphatic side-chain groups (Gly, Ala, ...) could have been adequate to play exactly the same role. Why then has nature chosen just the bulkiest of all them? This is because the large side chains of Ile663 and Leu597 significantly hinder the geometric evolution that AA needs to experience during the H_{10} abstraction but have no role when H_{13} is transferred.

Therefore, we conclude that, despite the fact that AA can perfectly accommodate both H_{13} and H_{10} to be transferred, the subtle steric hindrance by Leu597 and Ile663 of 15-LOX-1 turns out to be the main cause for the difference between the H_{13}/H_{10} potential energy barriers and, as a consequence, for the strict regiospecificity exhibited by the enzyme 15-lipoxygenase-1. It has to be emphasized that these residues are conserved in the mammalian lipoxygenase isoenzymes, where they are probably also responsible for discerning between the two hydrogen atoms of AA that can be positioned to be transferred. The crucial physiological role of all these enzymes would not be possible without the regiospecificity induced by such a subtle effect.

Finally, the calculations of the hydrogen abstraction process in LA show that the hydroperoxidation of LA by 15-lOX-1 is a faster chemical reaction than the H_{13} abstraction from AA, in accordance with the dual substrate specificity of 15-LOX-1, but with LA being the preferential substrate of this LOX.

The results presented in this Chapter have been published in papers:

- *ACS Catalysis*, vol. 4, no. 12, pp. 4351-4363, 2014
- *Journal of Physical Chemistry B*, vol. 8, no 20, pp. 1950-1960, 2015

Chapter 5

Leu597Ala-Ile663Ala 15-LOX-1 double mutant

5.1 Introduction

In the previous Chapter it was proposed that the presence of bulky residues in the active site cavity, especially residues Leu597 and Ile663, could play a role in controlling 15-LOX-1 regiospecificity. These residues are placed in the proximity of the substrate and may hinder the geometric evolution of the AA substrate from reactant to product in the H_{10} abstraction step, but not in the H_{13} abstraction one. Thus, a possibility for increasing the hydrogen abstraction from C_{10} could be the mutation of the bulky residues Leu597 and Ile663 to smaller ones, so avoiding their steric hindrance. To test this hypothesis the study of the hydrogen abstraction step of H_{13} and H_{10} in AA by the double mutant Leu597Ala-Ile663Ala has been carried out here.

5.2 Methodology

5.2.1 Preparation of the *in silico* mutant

The *in silico* double mutant Leu597Ala-Ile663Ala has been prepared from one of the AA:15-LOX-1 wild-type (WT) complex structures studied in the previous Chapter. Particularly, the selected structure corresponds to the reactant minimum of structure XI (see Figure 4.9), that was one of the structures initially prepared to react with both pro-S H_{10} and H_{13} of AA in the hydrogen abstraction process. The double mutant has been generated by changing the residues Leu597 and Ile663 to Alanine with the Mutator Plugin of the VMD program.[172]

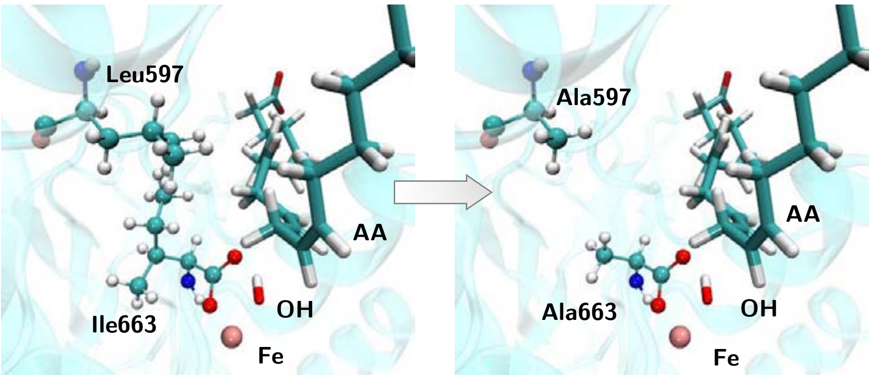


Figure 5.1: Representation of the starting structure of the AA:WT/15-LOX-1 complex (left) used to generate the double mutant Leu597Ala-Ile663Ala (right).

5.2.2 MD simulations

The molecular configurations of the Michaelis complex of the AA:Leu597Ala-Ile663Ala/15-LOX-1 mutant have been generated by means of MD simulations. For that purpose the following protocol has been applied. First, the substrate:protein complex has been fully solvated in an orthorhombic box of preequilibrated TIP3P water molecules, with dimensions of 125 Å x 79 Å x 85 Å (Figure 5.2). Next, the total charge of the system has been neutralized by adding six sodium cations. The resulting system contains nearly 75500 atoms, of which about 10600 of them belong to the protein.

The prepared system has been submitted to 1500 MM minimization steps with harmonic restrains applied to the protein and the substrate, followed by 1500 MM minimization steps without restrains (in both cases with the steepest descent algorithm) to avoid bad contacts and then, MD simulations under periodic boundary conditions (PBC) were started. The system was gradually heated from 25 K to 300 K during 80 ps, followed by an equilibration step of 200 ps at 300 K. During the heating steps, harmonic restraints were applied to the protein heavy atoms, which were gradually released, while the ligand and water molecules were kept free of restraints. Finally, a production step of 15 ns was performed by running three independent trajectories of 5 ns of production (with different assignment of the initial atomic velocities).

The quilibration step was done at constant volume and temperature (NVT), and production steps were done at constant pressure and temperature (NPT) using the extended system constant pressure and the Hoover constant temperature algorithms.[186] A time step of 2 fs was used for the production steps. All of the bonds and angles involving hydrogen atoms were constrained by the SHAKE algorithm.[126] PBCs were built with the

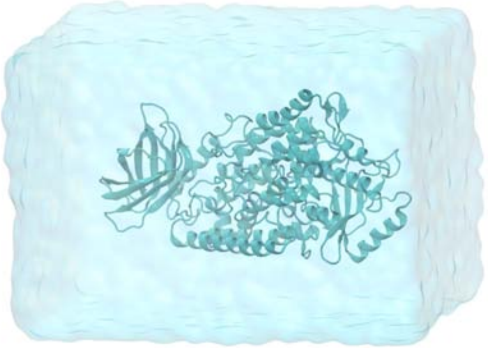


Figure 5.2: Representation of the complete system, substrate:protein complex inside the waterbox orthoromic cell unit: the AA:Leu597Ala-Ile663Ala/15-LOX-1 complex is represented by solid cartoon, and the water box is represented by a transparent surface.

CRYSTAL facility of CHARMM,[88, 126] using an orthorhombic cell unit. For long-range electrostatic, the Particle Mesh Ewald (PME) method was used.

All MD simulations were run with the program CHARMM version c35-b1. The force field topology and parameters for AA were derived from the CHARMM27[163, 164] ones, and CHARMM22[161, 162] force field was used for protein atoms except for Fe and its first coordination sphere, for which the parameter specifically derived by Saam *et al.*[187] were used. Structural analyses of the simulations were performed with the CORREL module of CHARMM.

5.2.3 QM/MM calculations

The methodology of the QM/MM calculations was the same that the used previously for the WT system. The QM/MM calculations have been carried out with the modular program package ChemShell,[105] with TURBOMOLE[159] for the DFT calculations, and the CHARMM22 and CHARMM27 force fields for the MM calculations using the DL-POLY module of ChemShell. The interaction between the QM and the MM subsystems has been treated by an electronic embedding scheme,[98] and a link atom scheme has been adopted to describe the QM/MM boundary using the charge shift model. No cutoffs were introduced for the nonbonding MM and QM/MM interactions.[165]

QM/MM optimizations were performed with the HDLCopt module, employing the L-BFGS algorithm[166, 167] for energy minimizations, and the microiterative optimizer, combining both P-RFO[168, 169] and L-BFGS, during the transition state searches. In all, around 12000 atoms are included in the QM/MM calculations, of which around 2000 are moved during the optimization process.

In all calculations the QM subsystem has been described by the B3LYP functional[188] and the 6-31G(d) basis set[173] has been used for C, H, O and N atoms, and the LANL2DZ basis set[174] for Fe. The QM region, which contains 74 atoms (link atoms not included), consist of 24 atoms of the lipid substrate, 11 atoms of each of the four His residues in the Fe coordination sphere (His361, His366, His541 and His545), 3 atoms of the terminal residue 663 in the coordination sphere, and the Fe(III)-OH^- cofactor, able to accept the hydrogen from AA. Seven link atoms were used, five along the bonds Cα-QM atom of the five residues in the Fe coordination sphere, and two bonded to the aliphatic carbons of the lipid substrate (placed between C_6-C_7 and C_{16}-C_{17}).

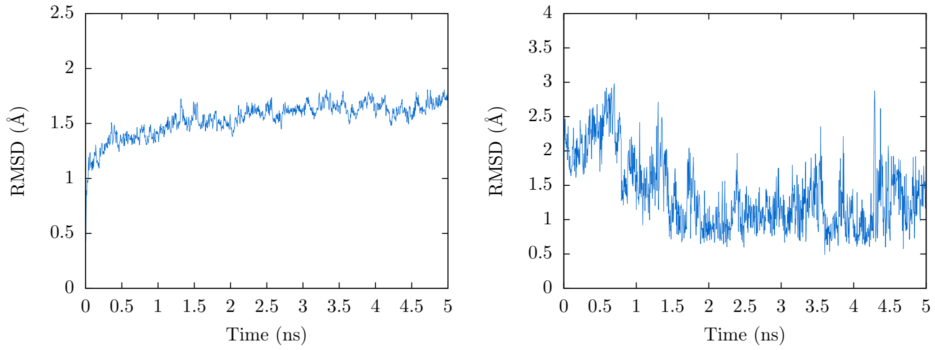
5.3 Results and Discussion

5.3.1 Multiple binding modes of the AA:Leu597Ala-Ile663Ala/15-LOX-1 complexes

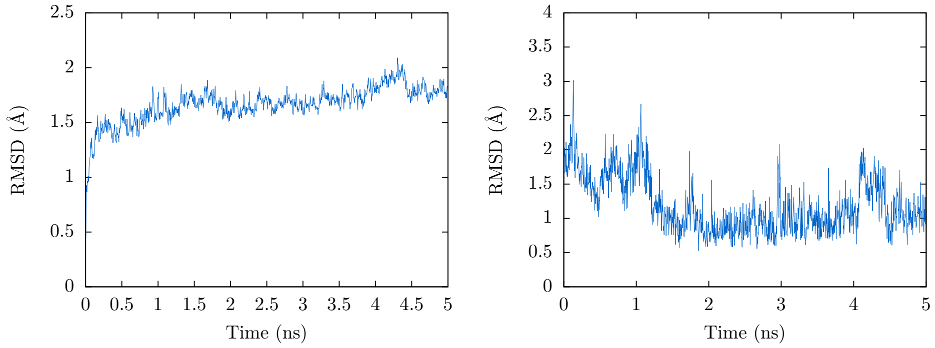
The exploration of the conformational landscape of the AA binding modes into the Leu597Ala-Ile663Ala mutant has been performed by running three non-correlated 5 ns simulations, each one initiated with a slightly different protocol (i.e. assigning different initial velocities to the atoms). Those three simulations are referred here as trajectory 1 (TRJ1), trajectory 2 (TRJ2) and trajectory 3 (TRJ3). The MD protocol applied here ensures a wider exploration of the conformational space. Three non-correlated trajectories allow to explore different separated regions of the conformational landscape of the substrate:protein complex which could not be reached through a single long trajectory. However, it is necessary to ensure that the separated trajectories are long enough in order to equilibrate the system.

The Root Mean Square Deviations (RMSDs) of the protein atoms (excluding hydrogens) along TRJ1, TRJ2 and TRJ3, with the reference to the initial frame of each one of the 5-ns production trajectories, are shown in Figure 5.3. It can be observed that the three RMSDs approximately level off so indicating that the protein in the AA:Leu597Ala-Ile663Ala/15-LOX-1 Michaelis complex has relaxed to a more stable structure and equilibration might be assumed. In this sense, the calculated average RMSD values of 1.55 ± 0.15 Å, 1.66 ± 0.16 Å and 1.67 ± 0.19 Å for TRJ1, TRJ2 and TRJ3, respectively, reflect this stability of the overall protein structure.

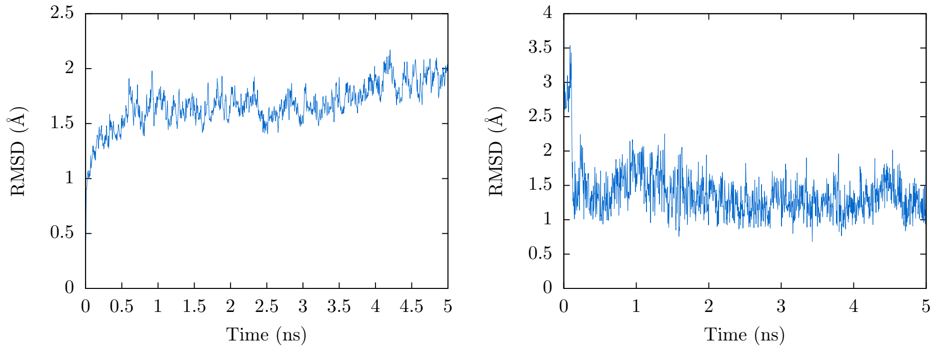
The RMSDs of the AA heavy atoms with respect to the substrate average structure again along the production periods of TRJ1, TRJ2 and TRJ3 trajectories are also shown in Figure 5.3. It can be observed that AA experiments significant rearrangements inside the enzyme mutant active site



(a) Protein RMSD in TRJ1. **(b)** AA RMSD in TRJ1.



(c) Protein RMSD in TRJ2. **(d)** AA RMSD in TRJ2.



(e) Protein RMSD in TRJ3. **(f)** AA RMSD in Trj3.

Figure 5.3: Protein and AA RMSD during the MD production step. Protein RMSD has been calculated with respect to the starting coordinates of the production step, whereas AA RMSD has been calculated with respect the average of the production step coordinates.

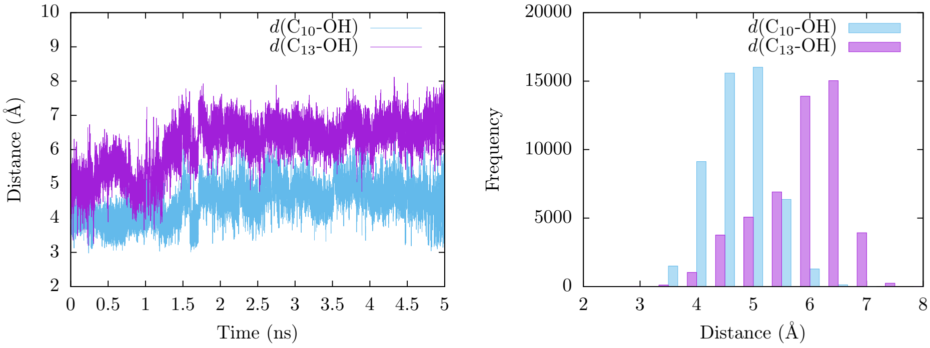
in accordance with the large conformational variability of the AA:Leu597Ala-Ile663Ala/15-LOX-1 Michaelis complex, also observed in the simulations of the AA:WT/15-LOX-1 complex. Mostly, those substrate motions consist in rather rigid internal rotations and vibrations of the unsaturated AA aliphatic carbon-chain that fluctuates around an average bounding mode inside the large active site of the double mutant 15-LOX-1. In particular, the high RMSD values at the beginning of TRJ1 correspond to a set of folded conformations of AA in which the substrate aliphatic carbon-chain has turned into itself with the AA tail in such a way to fill the room left by the size reduction of the residue side chains in the Leu597Ala-Ile663Ala double mutant. Later on along the same trajectory, AA fluctuates for a longer period of time around more extended conformations (see Figure 5.5 below). These results highlight that in the double mutant Michaelis complex there is enough room and side chain flexibility in the cavity to allow some conformational variability in AA, especially for the positioning of the conjugated double bonds and the methyl end.

In addition, distances of the atoms of the methylene groups of interest (C_{10} and C_{13}) to the oxygen atom of the Fe(III)-OH^- cofactor, have been monitored during the trajectories. In Table 5.1 the average distances $d(C_{10}\text{-OH})$, $d(H_{10}\text{-OH})$, $d(C_{13}\text{-OH})$, and $d(H_{13}\text{-OH})$ (where H_{10} and H_{13} are the pro-S hydrogens), measured over the production steps, are provided. The results show that the $d(C_{10}/H_{10}\text{-OH})$ average distances are smaller than the $d(C_{13}/H_{13}\text{-OH})$ ones. In principle, this result would be in agreement with the behavior as a 12-lipoxygenating enzyme (i.e. preference for H_{10} abstraction and oxygenation at C_{12}) of the AA:Leu597Ala-Ile663Ala/15-LOX-1 mutant because C_{10}/H_{10} are in average closer than C_{13}/H_{13} to the Fe(III)-OH^- cofactor that initiates the H-abstraction process from the corresponding methylene group of AA.

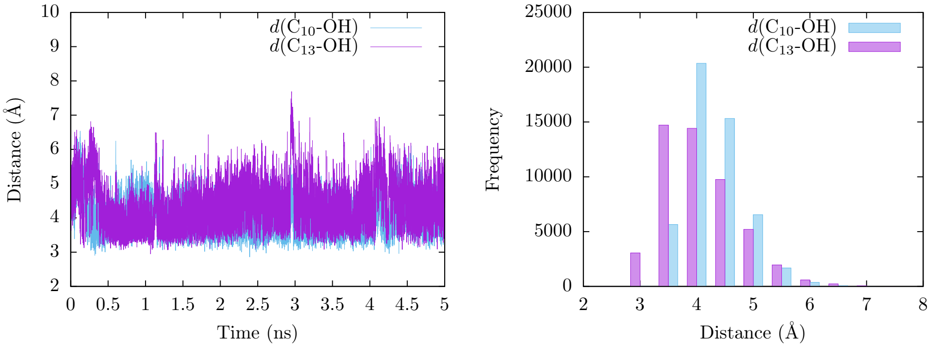
Table 5.1: Average distances of interest during the production steps of the MD simulations. Distances are given in Å along with the root mean square deviation included in parentheses.

Geometric parameter	TRJ1	TRJ2	TRJ3
$d(C_{13}\text{-OH})$	6.16 (0.73)	4.34 (0.66)	5.31 (0.77)
$d(C_{10}\text{-OH})$	4.46 (0.53)	4.04 (0.49)	4.18 (0.48)
$d(H_{13}\text{-OH})$	6.23 (0.74)	3.71 (1.00)	5.58 (0.77)
$d(H_{10}\text{-OH})$	4.75 (0.75)	3.65 (0.95)	4.76 (0.71)

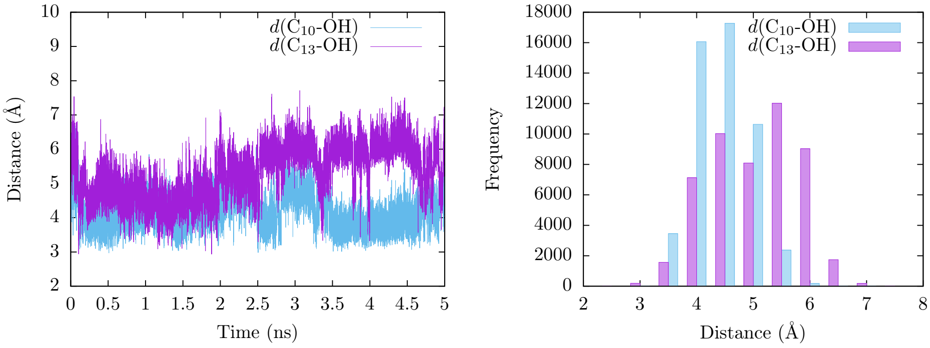
However, these average distances, extracted from MD simulations of an enzyme:substrate Michaelis complex, must be analyzed with great care. In fact, in that previous study on the AA binding to WT 15-LOX-1,[28] the $d(C_{10}/H_{10}\text{-OH})$ average distances obtained from MD trajectories were comparable to the $d(C_{13}/H_{13}\text{-OH})$ ones, in this case in apparent contradiction with the experimental regioselectivity (preference for H_{13} abstraction and oxygenation at C_{15}) of WT 15-LOX-1. But as has been demonstrated in the previous Chapter, the distance criterion of the reactive atoms to the Fe(III)-OH- cofactor is not the unique factor to take into account.



(a) Evolution of distances in TRJ1. **(b)** Histogram of distances in TRJ1.



(c) Evolution of distances in TRJ2. **(d)** Histogram of distances in TRJ2.



(e) Evolution of distances in TRJ3. **(f)** Histogram of distances in TRJ3.

Figure 5.4: Evolution and histogram of distances C_{13}-OH (purple) and C_{10}-OH (blue) during the production steps in Trajectory1, Trajectory2 and Trajectory3 of MD simulations in mutant Leu597Ala-Ile663Ala 15-LOX-1.

In the AA:Leu597Ala-Ile663Ala/15-LOX-1 mutant though, the d(C_{10}-/H_{10}-OH) average distances are in between 0.5 to 2.3 Å shorter than in the AA:WT/15-LOX-1 Michaelis complex, showing that the reduction in the side chains of Leu597 and Ile663 by means of the double mutation especially favor the approach of the C_{10}/H_{10} group of AA to the Fe(III)-OH^- cofactor. The C_{13}/H_{13} group of AA only approaches clearly to the cofactor in one trajectory of the double mutant simulation, TRJ2 (see Figure 5.4 (c) and (d), in which, however, the C_{10}/H_{10} is close enough as well. Therefore C_{13}/H_{13} remains in average farther away from the Fe(III)-OH^- cofactor in the AA:Leu597Ala-Ile663Ala/15-LOX-1 Michaelis complex than in the AA:WT/15-LOX-1 one. This result indicates that, globally, the *in silico* simulation of the Leu597Ala-Ile663Ala 15-LOX-1 mutant explores for a longer period of time than the WT 15-LOX-1 simulation configurations of AA bound into the active site with the C_{10}-methylene group better exposed for the H-abstraction reaction than the pentadiene centered at C_{13}. This is a first relevant result indicating that the double mutation Leu597Ala-Ile663Ala might favor the change in regioselectivity of a 15-LOX enzyme to a 12-LOX one.

However, we have to take into account as already mentioned in the study of AA:WT/15-LOX-1 complex, that the d(C_{10}/H_{10}-OH) and d(C_{13}/H_{13}-OH) distances may not be the most relevant molecular factor for explaining LOX regioselectivity. In WT 15-LOX-1 as well as in the double mutant studied here, the position of C_{10} and C_{13} of AA with respect to the OH- molecule might be very similar (as occurred for example with structure XI of the previous Chapter (Figure 1.9) or rather different, depending not only on the precise positioning of AA but also in the internal angles/dihedrals of the AA carbon skeleton.

Figure 5.5 shows four snapshots of the MD simulation of the AA:Leu597Ala-Ile663Ala/15-LOX-1 Michaelis complex. In Figure 5.5 (a-b) two structures of the mutant are shown, in which AA is prone for H_{10}-abstraction as H_{10} is closer to OH- than H_{13}. In Figure 5.5 b, AA adopts the folded configuration mentioned above in which the H_{13} atom is as far as 5.65 Å from the oxygen atom of OH- group, whereas H_{10} remains at short distances (2.89 Å). In contrast, in Figure 5.5 c it is shown a structure of the mutant, in which AA is prone for H_{13}-abstraction as H_{13} is much closer to OH- than H_{10}. Finally, Figure 5.5 d is an example of a configuration in which both H-abstractions might be plausible as both hydrogen atoms are located at quite similar short distances from the OH- acceptor group.

In comparing these four MD snapshots it can be appreciated the variability in the orientation of the different pentadiene groups of the AA carbon chain and also the different locations of the AA tail inside the enlarged hydrophobic cavity of the Leu597Ala-Ile663Ala mutant. In Figure 5.5 (e-g) an overlay between the active sites of the WT and the mutant AA:15-LOX-1 complexes is depicted to compare the relative orientation of the AA substrate with respect to Leu597/Ile663 or Ala597/Ala663 side chains. The van der Waals surfaces corresponding to the side chains of the two mutated residues clearly show (Figure 5.5 f) the extra available space for the positioning of AA within the cavity, in contrast to the close proximity of the Leu597/Ile663 side chains to the substrate in the WT enzyme (Figure 5.5 g). In these configurations van der Waals contacts between the residues side chains and the AA carbon-chain are established.

To proceed further with the analysis of the molecular factors that determine the regioselectivity of lipoxygenases, a selection criterion has been adopted to filter the so-called reactive structures from the global ensemble of configurations of the Michaelis complex to initiate the study of hydrogen abstraction step. This selection criterion was based on the following geometrical conditions: $d(\mathrm{H}_{13}\text{-OH})$ or $d(\mathrm{H}_{10}\text{-OH}) \leq 3.0$ Å and $d(\mathrm{C}_{13}\text{-OH}) > d(\mathrm{H}_{13}\text{-OH})$ or $d(\mathrm{C}_{10}\text{-OH}) > d(\mathrm{H}_{10}\text{-OH})$, thus ensuring that the corresponding C-H bond is properly oriented for hydrogen abstraction. In this way, two groups of reactive structures were extracted from each MD trajectory: structures suitable for H_{13}-abstraction and those suitable for H_{10}-abstraction. Next, a clustering analysis was carried out within those two groups of reactive structures based on the RMSD of AA heavy atoms with a cutoff of 1 Å (Figure5.6 shows the cluster of structures suitable for H_{10} abstraction). The most populated clusters represent the preferred binding modes of AA in the active site of the AA:15-LOX-1 mutant Michaelis complex corresponding to reactive structures, that is, ready to initiate the H-abstraction of H_{10} or H_{13}.

5.3.2 Reaction mechanism for the hydrogen abstraction step

Several representative structures (numbered from I to XXII) from the most populated clusters of reactive structures were geometrically optimized and the minimum energy structures corresponding to the reactants for the Leu597-Ala-Ile663Ala mutants of AA:15-LOX-1 complex were located on the QM/MM potential energy surface. From those minima, the potential energy profiles for the H_{10}-abstractions (structures I to XI in Table 5.2), and H_{13}-abstractions (structures XII to XXII in Table 5.3) were calculated, and the corresponding transition state (TS) structures were located following the methodological protocol detailed in the Methods section.

In Table 5.2 the main geometric parameters corresponding to the stationary points (minima and TS structures) of the calculated H-abstraction energy profiles for H_{10} abstraction are given. The $d(\mathrm{H}_{10}\text{-OH})$ and $d(\mathrm{C}_{10}\text{-OH})$ distance values at the reactants minima for structures I to XI range from 2.19 to 3.68 Å, and from 3.29 to 3.86 Å, respectively, whereas in these

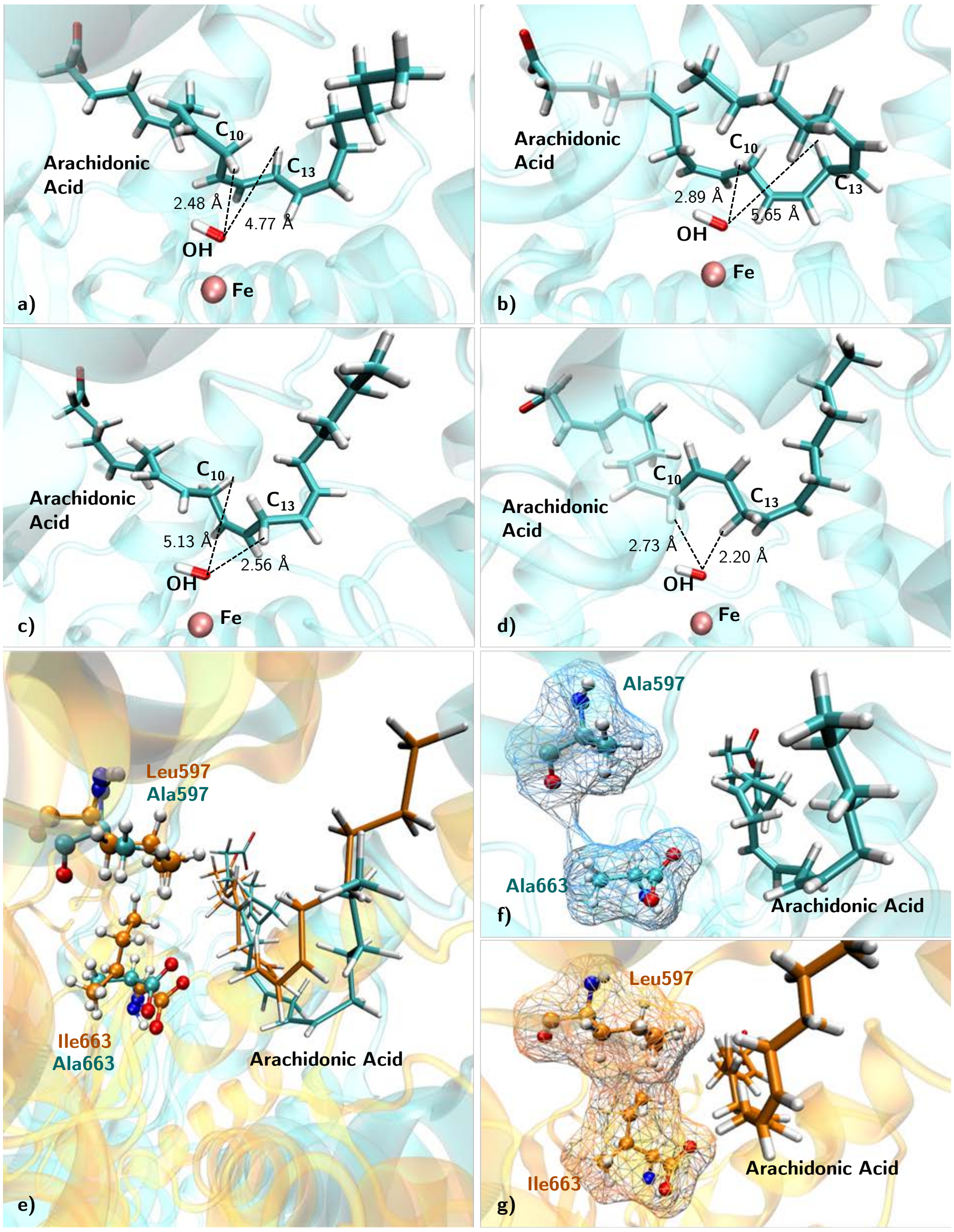


Figure 5.5: Images a) to d): Representation of the active site of the AA:Leu597Ala-Ile663Ala/15-LOX-1 mutant showing the arachidonic acid molecule and the Fe(III)-OH^- cofactor in different snapshots of the MD: a) snapshot 1; b) snapshot 2; c) snapshot 3; d) snapshot 4. Images e) to g): Comparison of the wild type (orange) and mutant (blue): e) superposition of WT and mutant 15-LOX-1, showing the arachidonic acid and the residues 597 and 663, f) representation of the van der Waals surfaces of Ala597 and Ala663 in the mutant, g) representation of the van der Waals surfaces of Leu597 and Ile663 in the WT enzyme.

(a) Trajectory 1. **(b)** Trajectory 2. **(c)** Trajectory 3.

Figure 5.6: Representation of the most populated clusters of the structures that fulfill the selection criterion for the H_{10} abstraction process. Color code from more to less populated clusters: blue, red, cyan, purple, yellow.

Table 5.2: Main geometric parameters of the reactant and TS of the structures selected for H_{10} abstraction. Distances are given in Å and angles in degrees.

Structure	**d(H_{10}-OH)**		**d(C_{10}-H_{10})**		**d(C_{10}-OH)**		**$\angle$(C_{10}-H_{10}-O)**		**$\angle$(H_{10}-O-H)**	
	R	**TS**	**R**	**TS**	**R**	**TS**	**R**	**TS**	**R**	**TS**
I	2.89	1.35	1.10	1.30	3.60	2.63	122.1	163.7	71.3	93.3
II	3.03	1.32	1.09	1.32	3.71	2.61	120.3	161.1	84.5	96.3
III	2.59	1.37	1.10	1.29	3.61	2.65	153.9	172.8	98.9	103.9
IV	2.60	1.39	1.10	1.27	3.53	2.66	140.9	174.4	80.2	98.7
V	2.65	1.37	1.10	1.28	3.51	2.63	135.2	167.2	73.1	98.3
VI	2.19	1.36	1.10	1.29	3.29	2.63	177.5	169.5	90.8	101.2
VII	2.58	1.36	1.10	1.28	3.56	2.63	147.8	173.1	83.9	99.2
VIII	3.00	1.35	1.10	1.29	3.55	2.63	111.6	171.7	82.7	99.9
IX	2.58	1.32	1.10	1.30	3.32	2.60	124.3	166.9	86.4	97.7
X	3.46	1.32	1.10	1.31	3.65	2.60	91.4	161.1	91.6	96.3
XI	3.68	1.33	1.10	1.30	3.86	2.62	91.1	166.5	75.1	94.3

structures the smallest d(H_{13}-OH) and d(C_{13}-OH) distances are 4.88 and 4.49 Å, respectively. However, as explained above, there are many other molecular configurations of the mutant in which also C_{13} is located at reactive distances. Structures XII to XXII, represented in Table 5.3, are representative of such set of configurations in which, in fact, H_{13} is closer than H_{10} to the OH-group. This is the reason why those structures have been selected to initiate the H_{13}-abstraction potential energy profiles. The d(H_{13}-OH) and d(C_{13}-OH) distance values at the reactants minima for structures XII to XXII range from 2.33 to 3.04 Å and from 3.27 to 3.93 Å, respectively, whereas the smallest d(H_{10}-OH) and d(C_{10}-OH) distances are as short as 2.53 and 3.59 Å, respectively, for these structures. A similar dispersion of geometries of the reactants minima was also present in the AA:WT/15-LOX-1 complex (see Tables 1.1 and 1.2 in the previous Chapter). It is worth pointing out, though, that the double mutation increases the d(H_{10} -OH) distance range (from 2.67 Å to 3.16 Å in WT 15-LOX-1 and from 2.19 to 3.68 Å in mutant 15-LOX-1), probably due to the extra space available: C_{10}/H_{10} can get closer to the

Fe(III)-OH^- cofactor but there are also reactive minima at longer distances.

The values of the ∠(C-H-O) angle, also given in Tables 5.2 and 5.3, indicate the degree of linearity of the H-abstraction process. At reactant structures there is a relevant dispersion of ∠C-H-O angles, all of them being far from linearity for H_{10}-abstraction, as well as for H_{13}-abstraction structures, with the exception of structure VI (H_{10}) and structure XIV (H_{13}) that have ∠(C-H_{10}-O) and ∠(C-H_{13}-O) angles of 177.5° and 177.3°, respectively. On the other hand, as in AA:WT/15-LOX-1, the H_{10}- and H_{13}-abstraction TS structures are more alike around the transferring hydrogen atom than the corresponding reactants minima. That is, the dispersion of the d(C_{10}/H_{10} -OH) and d(C_{13}/H_{13} -OH) distance values is smaller at the TSs than at reactants. Moreover, the three atoms involved in the H-abstraction process are almost collinear (the ∠(C-H-O) angle ranges from 161° to 174°) at all the TS structures, and also the ∠(H-O-H) angles are more similar among the TS structures than at reactants.

Table 5.3: Main geometric parameters of the reactant and TS of the structures selected for H_{13} abstraction. Distances are given in Å and angles in degrees.

Structure	**d(H_{13}-OH)**		**d(C_{13}-H_{13})**		**d(C_{13}-OH)**		**∠(C_{13}-H_{13}-O)**		**∠(H_{13}-O-H)**	
	R	**TS**	**R**	**TS**	**R**	**TS**	**R**	**TS**	**R**	**TS**
XII	2.9	1.27	1.10	1.34	3.79	2.6	137.9	169.7	91.5	102.0
XIII	2.9	1.3	1.09	1.31	3.78	2.6	137.4	169.2	94.2	103.8
XIV	2.55	1.29	1.10	1.34	3.64	2.62	177.3	173.9	101.8	102.2
XV	2.79	1.32	1.10	1.31	3.78	2.61	149.3	165.7	93.5	99.0
XVI	2.89	1.29	1.10	1.33	3.86	2.60	148.4	168.7	98.7	106.0
XVII	3.04	1.26	1.10	1.35	3.93	2.59	138.2	167.1	91.9	103.3
XVIII	2.39	1.29	1.10	1.33	3.43	2.62	156.6	172.5	110.5	103.9
XIX	2.48	1.3	1.10	1.31	3.44	2.61	145.6	173.0	99.2	103.0
XX	2.33	1.32	1.10	1.30	3.27	2.62	142.2	169.1	112.6	102.2
XXI	2.88	1.31	1.10	1.32	3.84	2.63	145.3	172.6	124.2	105.0
XXII	2.43	1.32	1.10	1.31	3.48	2.62	158.6	172.0	109.4	102.7

In this respect, our results demonstrate again that the H_{13}/H_{10}-abstraction processes present identical mechanisms and very similar TS structures, so corroborating the suggestion made by Holman and co-workers.[38,39] Therefore, the conformational variability of AA with respect to the OH^- group around the C_{10}/H_{10} or C_{13}/H_{13} bonds is greater among the set of reactive structures before the H-transfer takes place than among the set of transition state structures when the shifting hydrogen is already transferring.

The energy penalty for the H-abstraction process must depend somehow on these geometrical differences among the productive AA:15-LOX-1 mutant structures, however, the correlation with a unique geometrical parameter is not straightforward. That is, shorter C_{13}/H_{13}-OH distances do not always involve lower barriers. When the difference between d(H_{13}-OH) and d(H_{10}-OH) distances at one given minimum energy structure is rather large (i.e. more than 1 Å), the simulations have shown a correlation with the potential energy barrier: longer distance means higher barrier. However, our simulations also have shown that there exists a wide dispersion of geometries of the AA:Leu597Ala-Ile663Ala/15-LOX-1 Michaelis complex, as well as of the AA:WT/15-LOX-1 one, in a good disposition to transfer H_{13} or H_{10}, or even both, to the acceptor oxygen atom, being the difference in the d(H_{13}-OH) and d(H_{10}-OH) distances rather small. In those cases, the calculations on the AA:WT/15-LOX-1 system revealed that there is not a clear correlation with the H-abstraction barriers heights.

Table 5.4: Potential energy barriers $\Delta E^{\ddagger}$ (kcal/mol), and exponential average potential energy barriers $\Delta E^{\ddagger}_{\text{average}}$ (kcal/mol) for H_{10}- and H_{13}-abstractions

Structure	Abstraction	$\Delta E^{\ddagger}$	Structure	Abstraction	$\Delta E^{\ddagger}$
I	H_{10}	25.47	XII	H_{13}	25.41
II	H_{10}	24.85	XIII	H_{13}	22.81
III	H_{10}	15.46	XIV	H_{13}	23.6
IV	H_{10}	12.64	XV	H_{13}	19.49
V	H_{10}	15.80	XVI	H_{13}	23.76
VI	H_{10}	12.84	XVII	H_{13}	24.95
VII	H_{10}	15.96	XVIII	H_{13}	20.22
VIII	H_{10}	16.08	XIX	H_{13}	17.46
IX	H_{10}	18.54	XX	H_{13}	17.51
X	H_{10}	21.03	XXI	H_{13}	20.88
XI	H_{10}	18.52	XXII	H_{13}	17.43
	$\Delta E^{\ddagger}_{\text{average}} =$	13.70		$\Delta E^{\ddagger}_{\text{average}} =$	18.20

For the double mutant of the AA:15-LOX-1 complex studied here, the potential energy barriers along with the exponential average potential energy barriers[180,181,189] are given for the H_{10}-abstraction (I-XI structures) and the H_{13}-abstraction (XII-XXII structures) processes in Table 5.4. It can be observed that the two structures (IV and VI) that present the lowest barriers for H_{10}-abstraction (12.6 and 12.8 kcal/mol, respectively), have quite different d(H_{10}-OH) distances at the reactant minima (2.60 and 2.19 Å, respectively). Structure XI, that has the greatest d(H_{10}-OH) distance (3.68 Å) among the selected reactive structures, does not present the highest potential energy barrier. Moreover, two structures with the same d(H_{10}-OH) distance of 2.58 Å (structures VII and IX) present a difference of 2.6 kcal/mol in the potential energy barrier heights. This confirms in the AA:Leu597Ala-Ile663Ala/15-LOX-1 mutant, as we already concluded for the AA:WT/15-LOX-1 complex, that the d(H_{13}-OH) and d(H_{10}-OH) distances are not the unique factor, and not even the most important in many cases, contributing to the energetic penalty of the rate-limiting H-abstraction step.

When the H_{10}-abstraction energy profiles (structures I to XI), calculated here in the AA:Leu597Ala-Ile663Ala/15-LOX-1 mutant complex, are compared with the corresponding energy profiles in the AA:WT/15-LOX-1 system (see Table 4.3 in Chapter 4) the most relevant difference is that much lower potential energy barriers are obtained for the AA:15-LOX-1 double mutant. Potential energy barriers for H_{10}-abstraction as low as 12.6 and 12.8 kcal/mol, and some others in between 15.5 and 16.1 kcal/mol are obtained in the mutant (Table 5.4), whereas in AA:WT/15-LOX-1 the lowest H_{10}-abstraction potential energy barrier had a value of 22.7 kcal/mol (Table 1.3 in previous Chapter). In contrast, the H_{13}-abstraction potential energy barriers (structures XII-XXII) do not experiment such a big change in the AA:15-LOX-1 double mutant, even though the barriers obtained for the mutant move to a somewhat lower energetic range than those obtained for AA:WT/15-LOX-1 (which presented a lowest H_{13}-abstraction potential energy barrier of 18.6 kcal/mol). An estimation of the exponential average potential energy barriers for H_{10} and H_{13}-abstractions has been calculated from the individual potential energy barriers given in Table 5.4 using the exponential average (equation 4.4).[180, 181, 189] For the mutant, the obtained exponential average potential energy barriers are 13.7 kcal/mol for H_{10}-abstraction versus 18.2 kcal/mol for H_{13}-abstraction at $T = 300$ K, clearly indicating that H_{10}-abstraction is preferred.

For AA:WT/15-LOX-1 it was shown that the regioselectivity of this enzyme, that is, the C_{13}:C_{10} ratio of 97:3, is mainly due to the difference between H_{13}- versus H_{10}-abstraction potential energy barriers, being the difference between the exponential average potential energy barriers of 4 kcal/mol favoring H_{13}-abstraction. In the AA:Leu597Ala-Ile663Ala/15-LOX-1 mutant H_{10} abstraction is favored versus H_{13} by 4.5 kcal/mol. So this *in silico* mutagenesis experiment seems to suggest a way to convert a 15-lipoxygenating enzyme into a 12-lipoxygenating one. In the previous study on AA:WT/15-LOX-1, it was concluded that the steric hindrance imposed by the side chains of residues Leu597 and Ile663 was an important molecular factor controlling regioselectivity when H_{13}- versus H_{10}-abstraction processes were compared. The present study confirms the importance of those two large conserved hydrophobic residues in the active site of 15-LOX-1. In their presence the geometric evolution that AA needs to do along the H_{10} abstraction is significantly hindered, but has no role when H_{13} is transferred. When those two bulky residues are mutated to smaller ones, H_{10} abstraction is not impeded anymore and the regioselectivity of the hydrogen abstraction from AA is inverted.

5.3.3 Experimental results

The work presented here has been carried out in collaboration with the group of Prof. Hartmut Kühn, who designed an experiment to test our hypothesis of the effect of Leu597Ala-Ile663Ala mutation in the reaction specificity. In this experiment, different 15-LOX-1 variants were tested, the WT 15-LOX-1, the single mutants Leu597Ala and Ile663Ala, and the double mutant Leu597Ala-Ile663Ala. In each case, the product formation pattern was identified by means of HPLC (High-Performance Liquid Chromatography). However, in all cases 15-HETE was identified as the major product whereas 12-HETE only contributed a small share of the product mixture: for the single mutants Leu597Ala and Ile663Ala, the % of 15-HETE formed was 94.75 and 88.82 %, respectively, and for the double mutant Leu597Ala-Ile663Ala, the % of 15-HETE was 96.33 %.

These results indicate that the regioselectivity of AA oxygenation is not impacted by the introduction of small residues at the Leu597 and Ile663 positions. The *in vitro* results apparently contradict the *in silico* calculations. However, it is important to note that experiments determine the net regioselectivity of the overall oxidation process, whereas the theoretical study is limited to the hydrogen abstraction step, because it is assumed (and it happens for most systems) that this step governs the global regioselectivity. The apparent disagreement between both results suggests that this might not be the case.

5.3.4 Rotation of the peroxy moiety

To solve this obvious discrepancy additional *in silico* experiments have been carried out. The entire LOX reaction involves three consecutive elementary reactions: hydrogen abstraction, addition of molecular oxygen to the pentadienyl radical, and back-hydrogen transfer to the peroxy radical. The calculations presented in the previous section just refer to the first step, whereas the experimental results correspond to the analysis of the products of the complete reaction.

Under normoxic conditions[190] the addition of molecular oxygen addition to carbon-centered radical is faster than the initial hydrogen abstraction. In particular, in previous works in our group it has been shown[191–193] that the molecular oxygen addition to pentadienyl radicals in AA or linoleic acid are quite faster than the previous hydrogen transfer. Much less is known about the third step. Because the molecular oxygen attacks the pentadienyl radical antarafacially to the leaving hydrogen atom (that is, in the opposite side of iron atom), a rotation of the peroxy moiety to a suprafacial arrangement is required before the back-hydrogen transfer can occur.[194] The peroxy rotation involved in the third reaction step has been theoretically studied in this section.

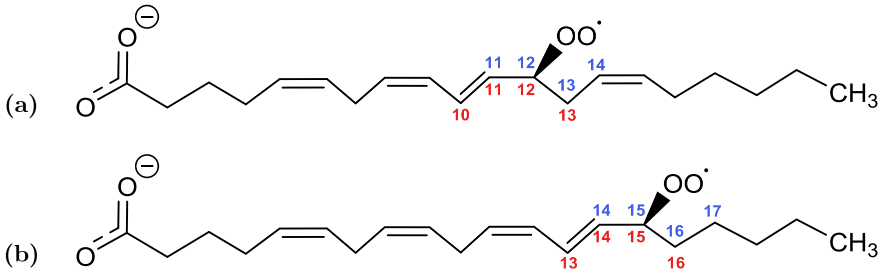
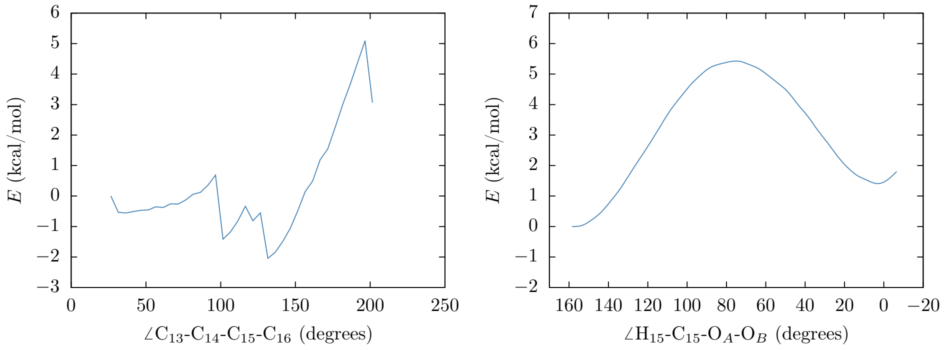


Figure 5.7: Possible dihedral angles involved in the peroxy moiety rotation in a) 12-peroxyl radical and b) 15-peroxyl radical.

Only structures involving low hydrogen abstraction barriers will be able to go on towards the molecular oxygen addition. Then, structure III, with a hydrogen abstraction energy barrier of 15.5 kcal/mol for the H_{10} abstraction process, and structure XV, with an energy barrier of 19.5 kcal/mol for the H_{13} abstraction, have been selected. To generate the corresponding peroxy radicals, the -OO· moiety has been placed “by hand" in the corresponding carbon atom of the optimized pentadienyl radicals (C_{15} in the H_{13} abstraction product and C_{12} in the H_{10} abstraction product), in an antarafacial arrangement with respect to the Fe center. The resulting structures have been optimized at the QM/MM level, yielding the peroxy radical, with the -OO· moiety bonded to an sp^3 carbon atom, and two conjugated double bonds in the remaining pentadienyl system. In each case, the QM part contains the pentadienyl moiety, the -OO· group, and the Fe and its coordination sphere.

To reach the suprafacial arrangement of the -OO· group the reaction coordinate has been defined as the rotation around one of the C-C bonds that contains this -OO· group. To describe this movement a dihedral angle of the carbon atoms centered on the bond around which the rotation is performed has been defined. As there are two possible C-C bonds that contain the carbon atoms of interest, it is possible to define two dihedral angles in every case: for the peroxy radical in C_{12}, dihedrals $\angle C_{10}\text{-}C_{11}\text{-}C_{12}\text{-}C_{13}$ and $\angle C_{11}\text{-}C_{12}\text{-}C_{13}\text{-}C_{14}$ have been tested; likewise, for C_{15}, dihedrals $\angle C_{13}\text{-}C_{14}\text{-}C_{15}\text{-}C_{16}$ and $\angle C_{14}\text{-}C_{15}\text{-}C_{16}\text{-}C_{17}$ were selected. Figure 5.7 represents the chemical structures of the 12- and 15-peroxy radical products showing the two possible dihedral angles in each case.

For each dihedral angle it is also possible to define two rotation directions, clockwise and counterclockwise. Therefore, for every system studied, four potential energy profiles have been calculated. In the case of the peroxy radical in C_{12}, none of the four reaction coordinates is able to produce the complete rotation of the -OO·, from an antarafacial to a suprafacial arrangement. This is a consequence of the nature of the substrate, which is very flexible and contains a large number of atoms that need to accommodate as the process goes on. When the evolution of the substrate along each reaction coordinate is visualized, it is observed that the -OO· moiety can hardly move forward. Instead, the rest of the atoms of AA are the ones that reorganize when the reaction coordinate progresses. The analysis of the geometry of structure III before the hydrogen transfer (see Figure 5.5 b)) permits to understand why. It is observed that the AA tail is not oriented towards the cavity end, but turned into itself pointing to the AA carboxylate end. This arrangement is repeated along different structures drawn from the MD simulations. These structures can be grouped in two main clusters (red and yellow), that are shown in Figure 5.6a.



(a) Clockwise rotation of $\angle C_{13}$-C_{14}-C_{15}-C_{16}. **(b)** Rotation of $\angle H_{15}$-C_{15}-O_A-O_B.

Figure 5.8: Potential energy profiles for the processes involved in the rearrangement of the peroxy moiety from an antarafacial configuration to a suprafacial configuration.

Structures with this arrangement are precisely the ones that yield the lower potential energy barriers for the H_{10} abstraction process, with values under 16 kcal/mol. However, when the substrate is positioned in this fashion, the evolution of the peroxy radical from an antarafacial to a suprafacial configuration is greatly hindered. The AA tail can accommodate in this position because of the extra space available in the active site gained with the mutations to smaller residues, but in this situation the rotation along the C-C bonds is extremely hindered by the AA substrate itself: the AA tail has very little space to allow the required rotation movement. There are indeed other different conformations of AA. However, for these other conformations the potential energy barrier for the hydrogen abstraction is higher than the more favored H_{13} ones, thus they are not responsible of an alteration in the regiospecificity. On the contrary, when the H_{10} abstraction step is favored, the evolution of the peroxy radical from antarafacial to suprafacial is highly impeded, so the regiospecificity is not globally affected.

In the case of the peroxy radical in C_{15} the reaction coordinate

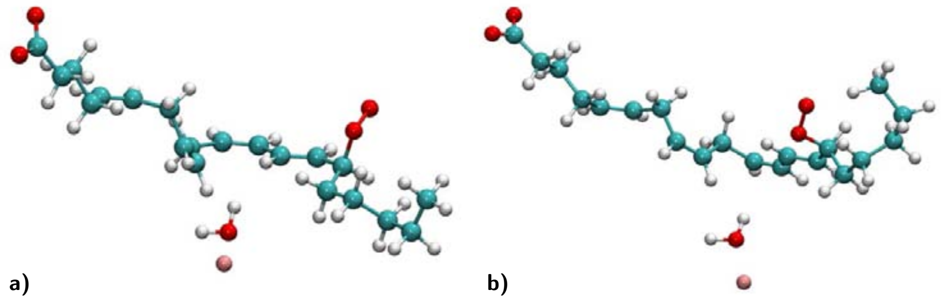


Figure 5.9: Minima corresponding to a) the antarafacial and b) the suprafacial form of the 15-peroxy radical structure. Only the most relevant atoms are shown.

corresponding to the clockwise rotation of the dihedral angle ∠C_{13}-C_{14}-C_{15}-C_{16} leads to a suprafacial arrangement of the -OO· group with the lowest energy barrier (Figure 5.8 a)). This profile does not exhibit a smooth curve, but the difference between the lowest and the highest energy points is around 5 kcal/mol, which is lower than the hydrogen abstraction potential energy barrier. Therefore, hydrogen abstraction is still the rate limiting step in this case. The suprafacial energy minimum (202 degrees, 3.1 kcal/mol above the antarafacial structure) has been optimized leading to the structure shown in Figure 5.9b (now the dihedral angle turns out to be 130.4 degrees), when it is compared with the corresponding antarafacial structure (Figure 5.9a). The change in the dihedral angle value when the optimization is carried out occurs because there are other regions of the substrate that can be relaxed.

Thus, it has been shown that for the oxidation product in C_{15}, the rotation pathway to evolve from an antarafacial to a suprafacial conformation is not hindered. When the structure studied here is look closely, it is observed that the AA tail is positioned towards the end of the active site cavity (Figure 5.9), contrarily to the oxidation product in C_{12}, in which the AA is twisted to the opposite side as in Figure 5.5b. When the AA skeleton is extended occupying the total length of the active site cavity, the rotation movement that AA undergoes to approach the -OO· group to the Fe is not hindered, allowing this process to occur with low energetic cost.

After the substrate has evolved from an antarafacial to a suprafacial arrangement, more conformational changes are needed before it is ready to undertake the third step of the reaction mechanism. The outermost oxygen atom of the -OO· moiety is not directed towards the catalytic iron, and an additional rotation movement is required to approach this oxygen to one of the hydrogens of the Fe(II)-OH_2 group. To this aim we have defined the dihedral angle ∠H_{15}-C_{15}-O_A-O_B as the reaction coordinate. The corresponding potential energy profile exhibits a smooth curve in this case (see Figure 5.8b). The pathway for the rotation is not hindered as there are not any residues in the proximity of the oxygen atoms of the -OO· moiety. This process has a low energetic cost, with a potential energy barrier of about 5 kcal/mol to get to the first minimum, where the outer oxygen atom (O_B) is already quite close to the hydrogen atom that will be back transferred. This minimum has been optimized, yielding the structure shown in Figure 5.10.

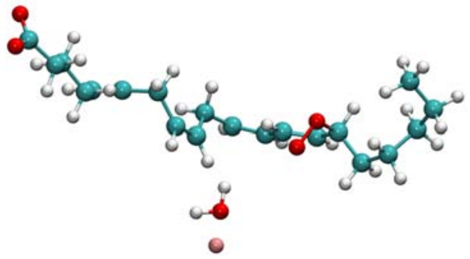


Figure 5.10: Peroxy radical structure ready for the back-hydrogen transfer from the Fe(II)-OH_2 cofactor to the -OO· moiety. The outermost oxygen atom of the -OO· group is pointing to one of the hydrogen atoms of the cofactor after the rotation of the group around the dihedral angle $\angle H_{15}$-C_{15}-O_A-O_B

5.3.5 Reduction of the peroxy radical and regeneration of the Fe(III)-OH- cofactor

The structure shown in Figure 5.10 is ready to initiate the third step of the reaction mechanism, the reduction of the peroxyl radical to the final hydroperoxide, and the regeneration of the Fe(III)-OH- cofactor. One of the hydrogens of Fe(II)-OH_2 can be transferred to the -OO· moiety. The reaction coordinate for this process has been defined as the difference between the breaking bond (O-H of the OH_2 in the cofactor) and the forming bond (O-H of the final -OOH group). The corresponding potential energy profile is shown in Figure 5.11.

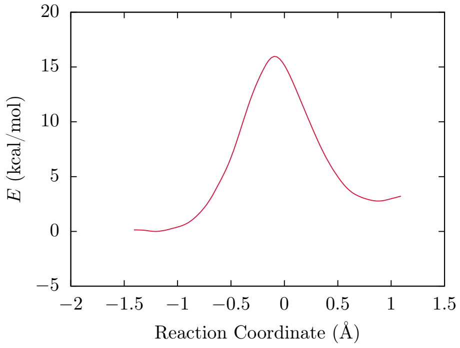


Figure 5.11: Potential energy profile for hydrogen transfer from the H_2O molecule of the Fe(II)-OH_2 cofactor to the -OO· moiety of the peroxy radical.

The process proceeds with a potential energy barrier of 16.05 kcal/mol. A transition state structure has been localized using a small core of 8 atoms: Fe(II)-OH_2 cofactor, the -OO· moiety, C_{15} and H_{15}. This barrier is lower than the potential energy barrier obtained for the hydrogen abstraction in this structure, which is of 19.5 kcal/mol. Then, the third step is not the rate limiting step. The process is endoergic, with a reaction energy of 2.78 kcal/mol. Figure 5.12 shows the structure of the stationary points of this reaction.

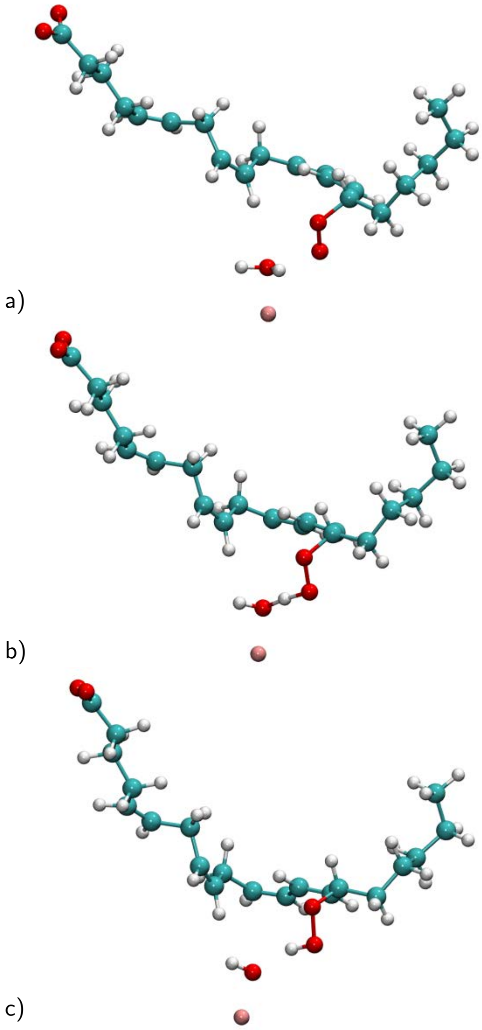
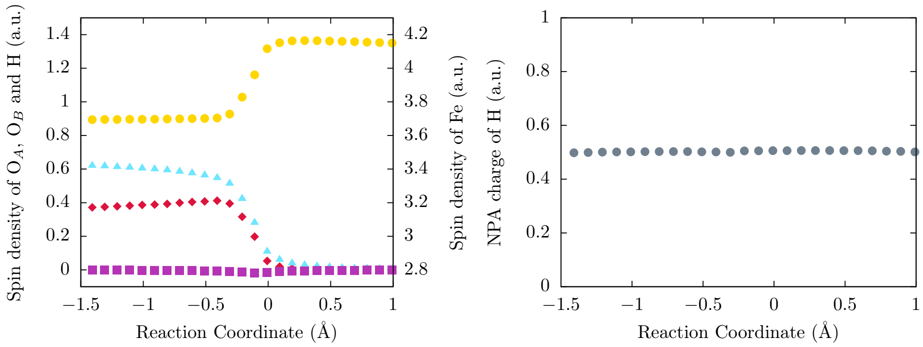


Figure 5.12: Representations of a) reactant, b) transition-state and c) product structures of the hydrogen transfer from the Fe(II)-OH_2 group to the peroxy radical. Only the substrate and the cofactor atoms are shown for simplicity.

For this back-hydrogen transfer step a PCET (Proton coupled electron transfer) mechanism takes place. The evolution of the spin density (expressed as the excess of α spin) during the hydrogen transfer has been evaluated. Figure 5.13 shows the spin densities of Fe, the oxygen atoms of the -OO· moiety, and the hydrogen transferred from the Fe(II)-OH_2 cofactor along the reaction coordinate.



(a) Spin density. **(b)** NPA charge.

Figure 5.13: Evolution of the spin densities on atoms Fe (yellow), O_A (red), O_B (cyan) and H (purple) and NPA charges on H (grey) during the hydrogen atom transfer from the Fe(II)-OH_2 cofactor to the -OO· moiety in the substrate.

The spin density of Fe increases from 3.7 to 4.2 a.u., as it would correspond to the transition from a Fe(II) quintet configuration to a Fe(III) sextet configuration. Spin densities of the oxygen atoms from the -OO· moiety decrease from 1 a.u. (0.3 a.u. on O_A and 0.7 a.u. on O_B) in the peroxy radical, which indicates the presence of one unpaired electron, to zero in the hydroperoxy product, showing that this group has accepted another electron that reduces the radical. The spin density of the hydrogen atom transferred remains virtually zero during the process, suggesting that this atom corresponds to a proton. When the atomic charge over this atom is analyzed, this result is confirmed: the NPA (natural population analysis)[171] net charge has a value near 0.5 a.u. during the reactive process. The electron that reduces the radical proceeds from the catalytic iron, which is oxidized from Fe(II) to Fe(III). So finally, the catalytic iron recovers its Fe(III) oxidation state and the OH- group is regenerated; then the Fe(III)-OH- cofactor is ready to initiate a new catalytic cycle.

5.4 Conclusions

In this chapter a work combining molecular dynamics simulations plus quantum mechanics/molecular mechanics calculations has been carried out to show that the double mutant Leu597Ala-Ile663Ala of rabbit 15-LOX-1 mostly catalyzes the hydrogen abstraction reaction from C_{10} of arachidonic acid versus hydrogen abstraction from C_{13}. The results confirm that both bulky residues, Leu597 and Ile663, drastically hinder the geometric motion that arachidonic acid needs to undergo to abstract a hydrogen from C_{10}, this way driving the hydrogen abstraction regioespecificity in wild type enzymes towards abstraction of a hydrogen from C_{13}. This hindrance does not exist in the double mutant. So, this *in silico* mutagenesis experiment apparently suggests a way to convert a 15-lipoxygenating enzyme to a 12-lipoxygenating one. However, site-directed mutagenesis together with HPLC analyses show that 15-HETE is the major product for both wild type and the double mutant enzyme, this way keeping invariant the regioselectivity of the whole hydroperoxidation process (15-lipoxygenating).

That apparent discrepancy between theoretical and experimental results can be understood if one realizes that the calculations only refer to the first step of the reaction, whereas the experimental results are based on the analysis of the products of the whole hydroperoxidation process. In particular, the key point lies on the conformational reorganization that the system must undergo to make it possible the back-hydrogen transfer (third step). After the molecular oxygen addition, the -OO· group, either at C_{12} (after H_{10} abstraction) or C_{15} (after H_{13} abstraction), is set in an antarafacial arrangement with respect to the Fe center, from where the back-hydrogen is not possible. Our QM/MM calculations show that the rotation of the -OO· group towards a suprafacial arrangement is quite easy in the case of the peroxy radical in C_{15}, but it is highly impeded and cannot be performed when the peroxy radical is in C_{12}.

To summarize, mutation of the bulky residues Leu597 and Ile663 to shorter ones creates an extra space inside the active site where arachidonic acid can accommodate in such a way that the hydrogen abstraction reaction from C_{10} is faster than from C_{13}. However, the peroxy radicals so formed at C_{12} are unproductive because those arachidonic acid configurations that favored the H_{10} abstraction prevent by steric hindrance the required rotation to a suprafacial arrangement to lead to the final hydroperoxy product. Then, the global regioselectivity of the reaction turns out to be the same for wild type and the double mutant Leu597Ala-Ile663Ala of rabbit 15-LOX-1, as has been experimentally demonstrated. Thus, it has been presented here an example of hydroperoxidation of arachidonic acid catalyzed by lipoxygenases where the regioselectivy of the initial hydrogen transfer reaction (the step generally assumed to be the rate determining one) and the whole regioselectrivity do not coincide.

The results presented in this Chapter have been published in paper:

- *ChemPhysChem*, vol. 17, no. 20, pp 1439-7641, 2016

Chapter 6

Ile418Ala 15-LOX-1 mutant

6.1 Introduction

To get a deeper insight into the factors governing the reaction specificity of 15-LOX-1, another strategy to transform it into a 12-lipoxygenating enzyme has been studied in this Thesis. According to the spatial hypothesis eplained in Section 1.2.3 in the Indtroduction Chapter, in 15-LOXs there are three critical amino acids at the bottom of the cavity ("triad concept") that position the substrate fatty acid in agreement with the enzyme dual regiospecificity. In rabbit 15-LOX-1 these three residues are Ile418, Phe359 and Ile593.[40] Experimentally, multiple mutagenesis studies (23, 24, 27, 29, 40) on a number of mammalian 15-LOXs have previously confirmed the triad concept, and an Ala-scan of the triad determinants of rabbit 15-LOX-1 indicated that a single Ile418Ala exchange (referred also as the major Sloane position) is sufficient to completely convert the reaction specificity of this enzyme from

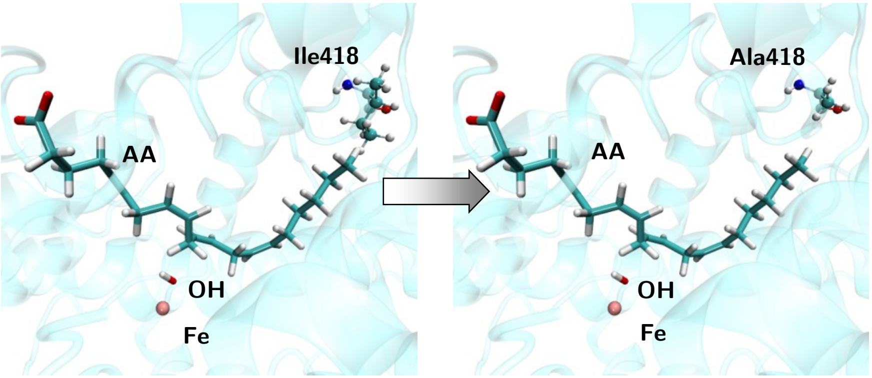


Figure 6.1: Representation of the starting structure (left) employed to generate the mutant Ile418Ala (right).

15- to 12-lipoxygenation (23). This *in vitro* conversion from a 15- to a 12-lipoxygenating LOX by means of the Ile418Ala mutation has been explained according to the spatial hypothesis with the AA bound in tail-first orientation (Figure 6.1.

In this Thesis the study of *in silico* Ile418Ala mutation has been carried out for rabbit 15-LOX-1 to gain further molecular insights into this specificity shift induced by mutations. The theoretical work presented in this Chapter was developed in collaboration the experimental group of Prof. Kühn. Moreover, they propose that this single residue is involved in an evolutionary switch, as lower mammals (mice, rats, pigs) express 12-lipoxygenating 15-LOX-1 orthologs but higher developed primates including human express 15-lipoxygenating isoenzymes. Rabbit, which expresses 15-LOX-1 and presents an Isoleucine in this position, is an exception.

6.2 Methodology

Following the same procedure explained in the Methodology section of the previous Chapter, the single mutant Ile418Ala has been generated from a previous structure of the Michaelis complex of the WT system changing Ile418 to an Alanine by the mutator plugin of VMD.[172] As done before, the starting coordinates of WT structure correspond to the reactant minimum of the structure XI, which is prepared to react with both H_{10} and H_{13} in the hydrogen abstraction process.

Next, the molecular configurations of the Ile418Ala mutant AA:15-LOX-1 complex were generated by MD simulations. The MD protocol applied for the Ile418Ala mutant was the same for the Leu597Ala-Ile663Ala mutant. Thus, three independent trajectories of 5 ns of production runs were carried out for the Ile418Ala mutant initiated with a slightly independent protocol, giving a total of 15 ns of production run. Likewise, the QM/MM calculations setup for the reactivity study of the catalytically productive AA:Ile418Ala/15-LOX-1 complexes was the same as in the previous Chapter.

6.3 Results and Discussion

6.3.1 Multiple binding modes of the AA:Ile418Ala/15-LOX-1 complexes

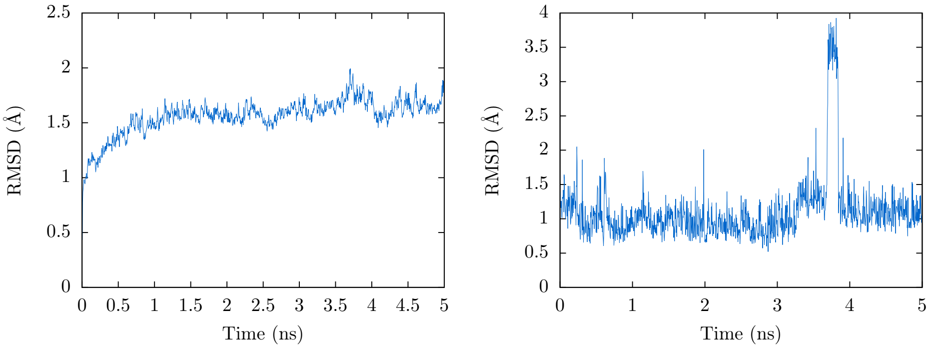
As has been previously done for the Leu597Ala-Ile663Ala mutant of rabbit 15-LOX-1, the three non-correlated simulations carried out to explore the conformational landscape of the AA:Ile418Ala/15-LOX-1 mutant system are referred here as trajectory 1 (TRJ1), trajectory 2 (TRJ2) and trajectory 3 (TRJ3). The RMSD of the protein backbone with respect to the initial structures, and the RMSD of the 20 carbon atoms of AA with respect to the substrate average structure along TRJ1, TRJ2 and TRJ3 are shown in Figure 6.2.

It can be observed that the three protein RMSDs with respect to the initial structure level off at a value around 1.5 Å, so indicating that the protein in the AA:Ile418Ala/15-LOX-1 mutant systems has relaxed to a more stable structure. The average RMSDs values for AA along TRJ1 and TRJ2 are very similar: 1.10 (0.46) Å and 1.13 (0.45) Å, respectively (standard deviation in parentheses). In both trajectories, however, there is a peak in the RMSDs evolution reflecting a clear reorientation of the AA molecule during a certain period of the simulation. The average RMSD for AA along TRJ3 is greater (1.78 (0.56) Å) because the substrate changes its configuration more frequently in this third simulation.

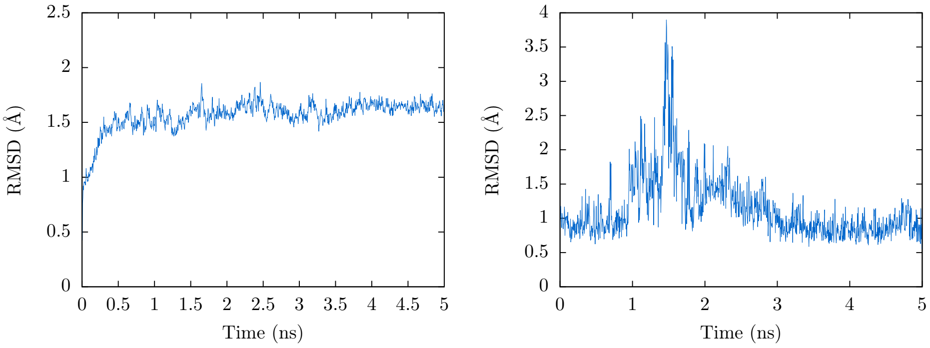
Table 6.1: Average distances of interest during the production steps of the MD simulations of the Ile418Ala mutant. Distances are given in Å along with the standard deviation included in parentheses. Only pro-*S* hydrogen are shown here.

Geometric parameter	**TRJ1**	**TRJ2**	**TRJ3**
$d(C_{13}\text{-OH})$	5.82 (0.61)	6.09 (0.68)	4.53 (1.05)
$d(C_{10}\text{-OH})$	5.28 (0.63)	3.84 (0.53)	4.11 (0.47)
$d(H_{13}\text{-OH})$	6.20 (0.71)	6.02 (0.77)	4.43 (1.32)
$d(H_{10}\text{-OH})$	4.84 (0.85)	3.21 (0.92)	4.43 (0.82)

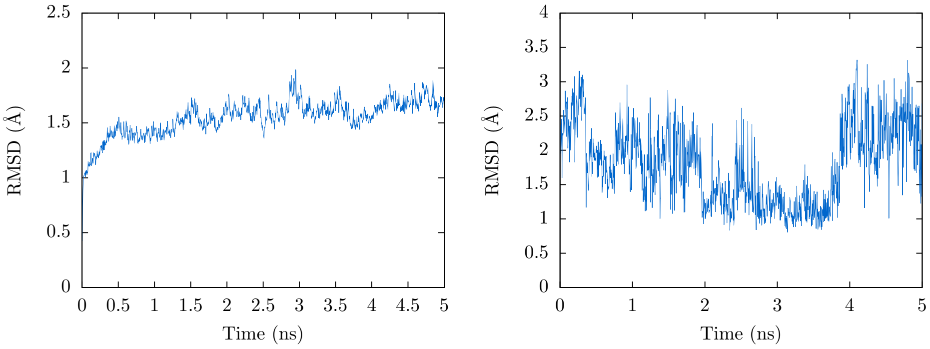
In Table 6.1 the average distances $d(C_{10}\text{-OH})$, $d(H_{10}\text{-OH})$, $d(C_{13}\text{-OH})$, and $d(H_{13}\text{-OH})$ are given for the three trajectories. The hydrogens H_{13} and H_{10} referred here are the pro-*S* hydrogens. The results for TRJ1 and TRJ2 indicate that the $d(C_{13}\text{-OH})$ average distances are greater than the $d(C_{10}\text{-OH})$ ones. Similarly, the $d(H_{13}\text{-OH})$ average distances are greater than the corresponding $d(H_{10}\text{-OH})$ average distances. Moreover, H_{10} seems to be, in average, better oriented for being abstracted than H_{13} in TRJ1 and TRJ2 as the $d(H_{10}\text{-OH})$ values are smaller than the $d(C_{10}\text{-OH})$ ones, whereas for the same two trajectories, $d(H_{13}\text{-OH})$ is greater or nearly equal than $d(C_{13}\text{-OH})$. In principle, this result would be in agreement with the behavior as a 12-lipoxygenating enzyme, (i.e. preference for H_{10} abstraction and oxygenation at C_{12}) of the AA:Ile418Ala/15-LOX-1 mutant observed experimentally. However, these results of a geometric parameter (average distances) extracted from the MD simulation of the substrate:protein Michaelis complex must be analyzed with great care. In fact, the results for TRJ3 seem to contradict the outcome of the two other trajectories because the $d(C_{13}$-



(a) Protein RMSD in TRJ1. **(b)** AA RMSD in TRJ1.



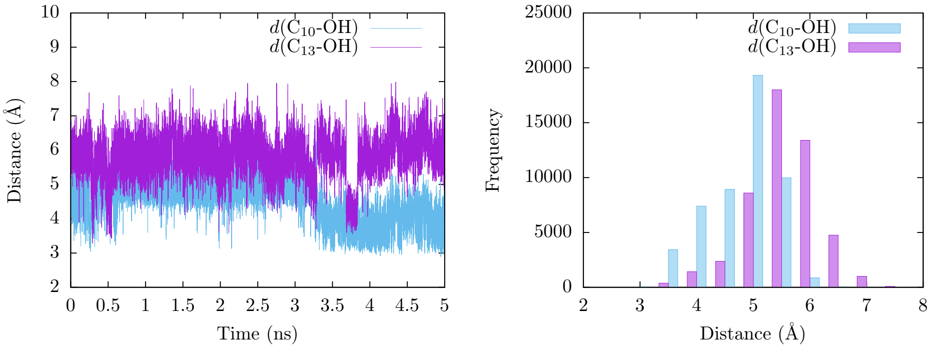
(c) Protein RMSD in TRJ2. **(d)** AA RMSD in TRJ2.



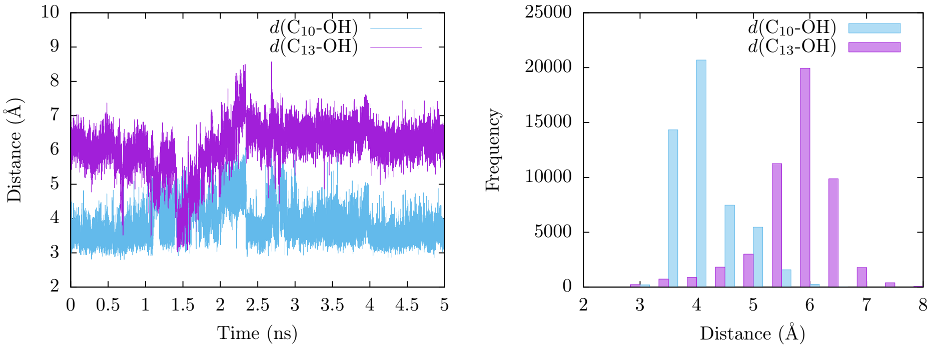
(e) Protein RMSD in TRJ3. **(f)** AA RMSD in TRJ3.

Figure 6.2: Protein and AA RMSDs during the MD production steps. Protein RMSD has been calculated with respect to the starting coordinates of the production step (= last coordinates of the equilibration step), whereas AA RMSD has been calculated with respect the average of the production step coordinates.

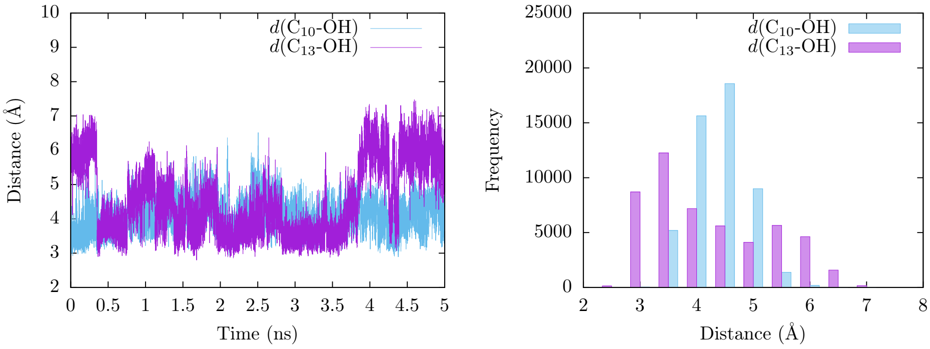
OH) average distance is still somewhat greater than the $d(C_{10}\text{-OH})$ average distance, but the $d(H_{13}\text{-OH})$ and $d(H_{10}\text{-OH})$ average distances turn out to be equal.



(a) Evolution of distances in TRJ1. **(b)** Histogram of distances in TRJ1.



(c) Evolution of distances in TRJ2. **(d)** Histogram of distances in TRJ2.



(e) Evolution of distances in TRJ3. **(f)** Histogram of distances in TRJ3.

Figure 6.3: Evolution of distances of interest and histograms during the production step for the mutant Ile418Ala.

A more detailed information is provided by the evolution of the distances along the MD trajectories. Figure 6.3 shows the evolution of $d(C_{10}\text{-OH})$ and $d(C_{13}\text{-OH})$ along the three trajectories, along with the histogram charts of the values explored by these two geometric parameters during the trajectories. Again, it is observed that C_{10} methylene is clearly closer to the Fe(III)-OH cofactor than C_{13} in TRJ1 and TRJ2. In TRJ2 C_{13} approaches to the oxygen atom at $\sim$1.5 ns, but it quickly moves away to a more distant position that keeps during the rest of the trajectory. In TRJ3, C_{13} gets to a close distance to the oxygen atom of the cofactor, which is similar to the C_{10} distance, during more than 3 ns, but at the end of the simulation C_{13} moves further away whereas C_{10} approaches.

It has to be noticed that in the previous study on the AA binding to wild type 15-LOX-1,[28] the $d(C_{10}/H_{10}\text{-OH})$ average distances obtained from MD trajectories were comparable to or even shorter than the $d(C_{13}/H_{13}\text{-OH})$ ones, in apparent contradiction to the experimental preference for H_{13} abstraction of WT 15-LOX-1. In that study, it was concluded then that the similar average values obtained for the $d(C_{10}/H_{10}\text{-OH})$ and $d(C_{13}/H_{13}\text{-OH})$ distances in the simulations of the AA:WT/15-LOX-1 complex reveal that those two carbon atoms are not really in structurally different locations in a great number of the generated conformations for AA at the active site of 15-LOX-1. In fact, both carbon atoms can be simultaneously located close enough to the OH- molecule for the corresponding H-abstraction processes being plausible. The same conclusions are applicable for the AA:Ile418Ala/15-LOX-1 complex, although a global structural change can be observed in comparing the average distances of AA:WT/15-LOX-1 and mutant AA:Ile418Ala/15-LOX-1 complexes. The relevant result is that the difference in average distances $[d(C_{13}\text{-OH}) - d(C_{10}\text{-OH})]$ moves to clearly greater values in the Michaelis complex of the Ile418Ala mutant than in WT 15-LOX-1. That is, C_{13}/H_{13} are farther away from the OH group with respect to C_{10}/H_{10} in the AA:Ile418Ala/15-LOX-1 mutant than in AA:WT/15-LOX-1.

To get deeper insights into the structural alterations in AA positioning provoked by the Ile418Ala mutation, additional geometric parameters measured from the MD simulations are reported here. As the mutated residue, Ala418, is located at the end of the active site cavity, the hydrophobic interaction of the AA tail, the methyl end, with Ala418 has been characterized in terms of distances of the closest atoms. Table 6.2 shows the distances of AA C_{20} to the carbon atoms of the protein backbone in residue 418 in both mutant and WT systems. The results corresponding to the MD simulations of the AA:WT/15-LOX-1 have been taken from reference.[28]

TRJ1 and TRJ2 of the mutant Ile418Ala clearly exhibit shorter $d(C_{20}\text{-C}\alpha\text{-Ala/Ile418})$ and $d(C_{20}\text{-CO-Ala/Ile418})$ values than the corresponding to the three WT trajectories. Mutant TRJ3, which presents the largest distance values of the three Ile418Ala simulations, shows similar distances (but still somewhat smaller) to the results of WT TRJ2 and TRJ3. These results indicate that the I418 side-chain prevents the approximation of the AA tail to the residue backbone, whereas in the Ile418Ala mutant, the smaller methyl side-chain facilitates this approach. Therefore, it seems that AA is shifted to a deeper position inside the binding pocket as suggested by the spatial hypothesis for explaining the enzyme reaction specificity.

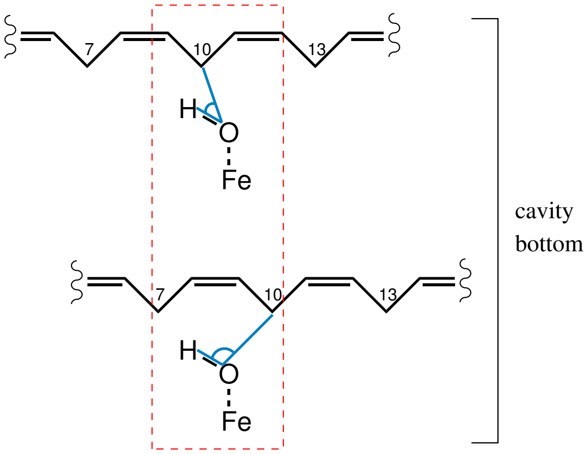


Figure 6.4: Measurement of C_{10}-O-H angle. The migration of AA towards the bottom of the cavity in the Ile418Ala mutant (bottom of the Figure) provokes an increase of the C_{10}-O-H value, with respect to the WT (top), the position of the Fe-coordination environment is not affected by the mutation.

Another geometric parameter measured to corroborate the spatial hypothesis were the average C_{10}-O-H and C_{13}-O-H angles: greater values indicate a deeper location into the active site cavity for the measured carbon (see Figure 6.4). The results are shown in Table 6.3. Again, these results are compared with the results measured from AA:WT/15-LOX-1 MD simulations[28] (Table 6.3).

Here we observe that the $\angle(C_{13}\text{-O-H})$ and $\angle(C_{10}\text{-O-H})$ values cor-

Table 6.2: Average distances of AA C_{20} to residue Ile418 (WT) and Ala418 (mutant) during the production steps of the MD simulations. Distances are given in Å along with the standard deviation included in parentheses.

Trajectory	$d(C_{20}\text{-C}\alpha\text{-Ala/Ile418})$	$d(C_{20}\text{-CO-Ala/Ile418})$
WT 1	13.42 (1.11)	13.20 (1.20)
WT 2	7.45 (0.90)	7.61 (0.95)
WT 3	7.75 (0.76)	7.57 (0.77)
Mutant 1	5.23 (1.02)	5.03 (1.05)
Mutant 2	5.15 (0.39)	5.33 (0.61)
Mutant 3	7.15 (1.81)	7.14 (1.93)

Table 6.3: Average angles for WT and mutant Ile418Ala MD production trajectories. Angles are given in degrees, and SD values are in parenthesis.

Trajectory	**$\angle$(C_{13}-O-H)**	**$\angle$(C_{10}-O-H)**
WT 1	87.6 (11.1)	61.1 (13.3)
WT 2	86.1 (21.3)	56.3 (19.6)
WT 3	66.6 (12.9)	46.6 (10.2)
Mutant 1	102.5 (8.2)	75.6 (9.7)
Mutant 2	101.1 (8.0)	80.2 (12.5)
Mutant 3	103.7 (11.6)	69.2 (15.9)

relate with the position of these carbons in the active site: greater values indicate that these carbon atoms are located deeper in the active site (see Figure 6.4). Accordingly, the $\angle$(C_{13}-O-H) angles in WT 15-LOX-1 as well as in the mutant are greater than the $\angle$(C_{10}-O-H) values. More significantly, the $\angle$(C_{13}-O-H) and $\angle$(C_{10}-O-H) values in the mutant are greater than in WT 15-LOX-1. It is also interesting that angles $\angle$(C_{10}-O-H) in the mutant exhibit similar values to the angles $\angle$(C_{13}-O-H) in the WT.

Taken together, the results presented here demonstrate that AA migrates deeper into the cavity of the mutant enzyme for the global set of AA structures. The new configuration of the mutant allows C_{10} to occupy the position of C_{13} in the WT, which could be the reason for the inversion of the regiospecificity. This displacement is due to the increase of free space at the bottom of the cavity left by the Ile418Ala mutation. All these structural analyses give a preliminar idea of the alterations of the binding modes of AA caused by the Ile418Ala mutation. However, as revealed the reactivity study of AA:WT/15-LOX-1,(Chapter 4) the analysis of the molecular factors that determine regioselectivity must be done over the set of reactive structures ready to initiate the H-abstraction process. The selection criterion adopted here to filter that set of reactive structures is the same as has been applied in the previous studies: $d(H_{13}\text{-OH})$ or $d(H_{10}\text{-OH}) \leq 3.0$ Å and $d(C_{13}\text{-OH}) > d(H_{13}\text{-OH})$ or $d(C_{10}\text{-OH}) > d(H_{10}\text{-OH})$, respectively, thus ensuring that the corresponding C-H bond is properly oriented for hydrogen abstraction. In this way, two groups of reactive structures were extracted from the set of MD trajectories: structures suitable for H_{13}-abstraction and those suitable for H_{10}-abstraction.

Next, a clustering analysis was carried out within those two groups based on the RMSD of AA heavy atoms. Figure 6.6 shows the clusters of the set of structures suitable for H_{10} abstraction. In Figure 6.5 a representative reactive structure for H_{10}-abstraction of the AA:Ile418Ala/15-LOX-1 mutant complex has been depicted overlaid with one of the reactive configurations for H_{13}-abstraction of the WT system. In these structures, it can be observed the displacement of AA towards the cavity bottom in the mutant enzyme as well as the approach of the AA methyl end to the backbone of Ala418 in agreement with the spatial hypothesis.

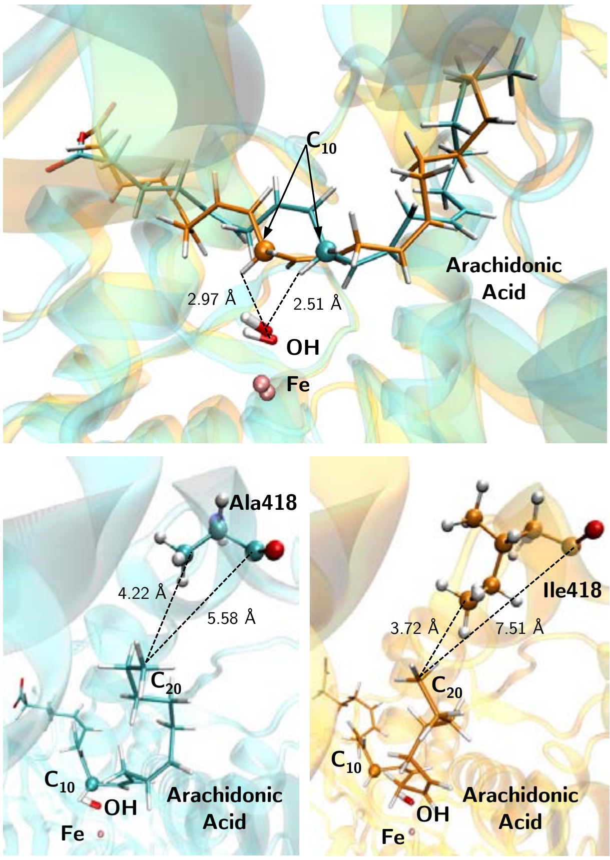
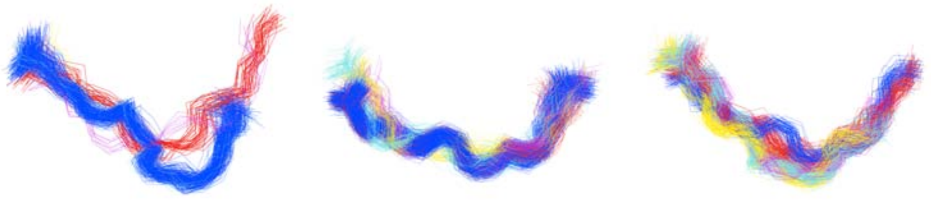


Figure 6.5: Overlay of AA:WT/15-LOX-1 (orange) and AA:Ile418Ala/15-LOX-1 (blue) active sites from two snapshots of the MD trajectories of both *in silico* models. The location of the arachidonic acid molecule with respect to the OH group is shown. AA, OH and Fe are represented in sticks and C_{10}, Ala 418 and Ile418 are represented in ball and sticks.

6.3.2 Study of the hydrogen-abstraction step

Representative structures from the most populated clusters were then optimized and the corresponding reactant minima were located on the QM/MM potential energy surface. From those minima, the potential energy profiles for the H_{10}-abstractions (structures I to X), and H_{13}-abstractions (structures XI to XX) were calculated, and the corresponding transition state structures were located. Geometric parameters of interest of the optimized reactant minima and TS structures are given for the structures selected for H_{10} and H_{13} abstractions in Tables 6.4 and 6.5, respectively. The d(H_{10}-



(a) Trajectory 1. **(b)** Trajectory 2. **(c)** Trajectory 3.

Figure 6.6: Representation of the most populated clusters of the structures that fulfill the selection criterion for the H_{10} abstraction process.

OH) and $d(C_{10}\text{-OH})$ distance values at the reactant minima for structures I to X range from 2.43 to 3.17 Å and from 3.28 to 3.67 Å, respectively. In these structures the smallest $d(H_{13}\text{-OH})$ and $d(C_{13}\text{-OH})$ distances are 5.21 and 4.39 Å, respectively. For structures XI to XX, C_{13} is located at reactive distances and H_{13} is closer than H_{10} to the OH group. On the other hand, as in AA:WT/15-LOX-1, the H_{10}- and H_{13}-abstraction TS structures are more alike around the transferring hydrogen atom than the corresponding reactant minima (for H_{10}-abstraction TSs, $d(H_{10}\text{-OH})$ and $d(C_{10}\text{-OH})$ range from 1.29 to 1.37 Å and from 1.26 to 1.35 Å, respectively; the corresponding $d(H_{13}\text{-OH})$ and $d(C_{13}\text{-OH})$ for H_{13}-abstraction TSs range from 1.32 to 1.36 Å and from 1.29 to 1.31 Å, respectively). Moreover, the three atoms involved in the H-abstraction process are almost collinear (the $\angle$(C-H-O) ranges from 165° to 176°) at all the TS structures.

Table 6.4: Main geometric parameters of reactants and TSs of the structures selected for H_{10} abstraction. Distances are given in Å and angles in degrees.

Structure	**$d(H_{10}\text{-OH})$**		**$d(C_{10}\text{-}H_{10})$**		**$d(C_{10}\text{-OH})$**		**$\angle(C\text{-}H_{10}\text{-}O)$**		**$\angle(H_{10}\text{-O-H})$**	
	R	**TS**	**R**	**TS**	**R**	**TS**	**R**	**TS**	**R**	**TS**
I	2.79	1.36	1.10	1.28	3.44	2.62	117.8	165.2	81.5	94.5
II	2.67	1.36	1.10	1.28	3.51	2.62	133.2	166.7	74.3	94.1
III	2.70	1.35	1.10	1.28	3.28	2.63	112.4	176.2	75.4	100.5
IV	2.63	1.37	1.10	1.26	3.29	2.63	117.7	173.3	81.5	99.4
V	3.17	1.36	1.10	1.28	3.47	2.63	96.6	173.7	62.0	97.8
VI	2.57	1.30	1.09	1.35	3.66	2.64	172.0	172.0	98.2	103.7
VII	2.60	1.30	1.09	1.34	3.67	2.63	162.9	173.4	89.1	103.0
VIII	2.58	1.29	1.10	1.33	3.63	2.61	162.0	172.3	91.9	100.9
IX	2.43	1.31	1.10	1.30	3.52	2.61	171.6	176.4	87.5	97.8
X	2.54	1.30	1.10	1.32	3.59	2.62	160.0	174.7	95.4	97.2

In this respect, our results demonstrate again that the H_{13}/H_{10} âĂŞabstraction processes present identical mechanisms and very similar transition state structures, and that the abstraction barrier heights must depend

Table 6.5: Main geometric parameters of reactants and TSs of the structures selected for H_{13} abstraction. Distances are given in Å and angles in degrees.

Structure	**d(H_{13}-OH)**		**d(C_{13}-H_{13})**		**d(C_{13}-OH)**		**∠(C-H_{13}-O)**		**∠(H_{13}-O-H)**	
	R	**TS**	**R**	**TS**	**R**	**TS**	**R**	**TS**	**R**	**TS**
XI	2.47	1.33	1.10	1.31	3.42	2.63	145.0	172.1	108.8	105.0
XII	2.98	1.33	1.10	1.31	3.91	2.64	143.5	175.7	99.9	105.7
XIII	2.22	1.34	1.10	1.30	3.31	2.63	171.0	172.3	101.4	103.0
XIV	2.46	1.36	1.10	1.30	3.53	2.65	161.6	169.7	111.0	105.6
XV	2.64	1.36	1.10	1.29	3.73	2.65	175.2	174.1	106.2	104.6
XVI	2.38	1.32	1.10	1.32	3.45	2.64	162.3	173.6	120.4	107.1
XVII	2.41	1.34	1.10	1.30	3.50	2.64	171.0	175.5	102.4	105.3
XVIII	2.75	1.35	1.10	1.30	3.83	2.64	169.1	171.9	107.1	105.3
XIX	2.40	1.34	1.10	1.31	3.48	2.65	166.2	174.4	112.3	107.4
XX	2.41	1.32	1.10	1.31	3.48	2.63	162.8	172.6	96.1	103.0

on the geometry of the structures at the reactant Michaelis complex. Among those geometric parameters, the d(H_{13}-OH) and d(H_{10}-OH) distances are obviously very important. So, when the difference between d(H_{13}-OH) and d(H_{10}-OH) distances at one given structure is large (more than 1 Å), there is a clear correlation with the potential energy barrier: longer distance means higher barrier. However, in those reactive structures where the difference in the d(H_{13}-OH) and d(H_{10}-OH) distances is rather small, the calculations on the AA:WT/15-LOX-1 system revealed that there is not a clear correlation with the H-abstraction barriers heights (see Figure 4.8 in Chapter 4), and the same is observed here for AA:Ile418Ala/15-LOX-1 (compare the Tables of geometric parameters and potential energy barriers).

In Table 6.6 the potential energy barriers are given for the H_{10}-abstraction (I-X structures) and the H_{13}-abstraction (XI-XX structures) processes. It can be observed that the lowest barrier for H_{10}-abstraction of 12.8 kcal/mol corresponds to structure IV, for which the d(H_{10}-OH) distance is 2.63 Å, whereas for structure VI, which has a d(H_{10}-OH) distance of 2.57 Å, the potential energy barrier raises up to 29.9 kcal/mol being the highest one. This confirms in the Ile418Ala mutant that the d(H_{13}-OH) and d(H_{10}-OH) distances are not the unique factor contributing to the energetic penalty of the rate-limiting H-abstraction step. Therefore, these structural differences at the reactants are not the main factor that causes the change in regioselectivity in the mutant enzyme. As for the H_{10} abstractions (structures I to X), the most significant energetic difference with respect to the AA:WT/15-LOX-1 results is that much lower potential energy barriers are obtained in the AA:Ile418Ala/15-LOX-1 mutant. Potential energy barriers for H_{10}-abstraction as low as 12.8 and 14.6 kcal/mol, and some others in between 16.8 and 17.8 kcal/mol, are obtained (Table 6.6) whereas in AA:WT/15-LOX-1 the lowest H_{10}-abstraction potential energy barrier had a value of 22.7 kcal/mol (see Table in Chapter). In contrast, the H_{13}-abstraction potential energy barriers (structures XI-XX) do not experiment such a big change in the AA:Ile418Ala/15-LOX-1 mutant, even though they move to a somewhat lower energetic range than in the AA:WT/15-LOX-1 complex.

An estimation of the exponential average potential energy barriers for H_{10}- and H_{13}- abstractions can be calculated from the individual potential energy barriers of Table 6.6 by using equation 4.4 of the exponential average.[180, 181, 189] For the Ile418Ala mutant the obtained average potential energy barriers are 14.1 kcal/mol for H_{10}-abstraction versus 16.5 kcal/mol for H_{13}-abstraction at $T = 300$K, clearly suggesting that H_{10}-abstraction is preferred.

Table 6.6: Potential energy barriers $\Delta E^{\ddagger}$ (kcal/mol), and exponential average potential energy barriers $\Delta E^{\ddagger}_{\text{average}}$ (kcal/mol) for H_{10}- and H_{13}-abstractions in the Ile418Ala mutant.

Structure	**Abstraction**	$\Delta E^{\ddagger}$	**Structure**	**Abstraction**	$\Delta E^{\ddagger}$
I	H_{10}	16.80	XI	H_{13}	21.19
II	H_{10}	17.50	XII	H_{13}	21.76
III	H_{10}	14.63	XIII	H_{13}	15.31
IV	H_{10}	12.76	XIV	H_{13}	16.28
V	H_{10}	17.79	XV	H_{13}	19.74
VI	H_{10}	29.93	XVI	H_{13}	18.81
VII	H_{10}	25.35	XVII	H_{13}	17.00
VIII	H_{10}	26.79	XVIII	H_{13}	17.92
IX	H_{10}	20.62	XIX	H_{13}	20.23
X	H_{10}	23.96	XX	H_{13}	17.62
	$\Delta E^{\ddagger}_{\text{average}} =$	14.10		$\Delta E^{\ddagger}_{\text{average}} =$	16.50

The difference between the average energy barrier for H_{10} in comparison with the average energy barrier for H_{13} (14.1 vs. 16.5 kcal/mol) is more than enough to explain the C_{13}:C_{10} hydrogen ratio of 8/92 reported in mutagenesis experiments. As a matter of fact this experimental ratio would correspond to a difference of 1.5 kcal/mol at $T = 300$ K. Our obtained difference of 2.4 kcal/mol is quite reasonable taking into account the complexity of the problem. Inside the active site of 15-LOX-1, AA can adopt an extremely huge number of relative positions. This huge dispersion of geometries is due to the flexibility of the enzyme residues and, especially, to the enormous amount of possible positions (including torsions) of AA due to its high flexibility (AA flexibility is explained in the Introduction chapter). Most of these structures would react too slowly to have any weight in the reaction. Some (in fact many) of them (the reactive structures) can react to yield the products, although with a wide range of rates. The measured rate of the reaction is the result of the contribution of all the reactive structures, with a global rate that can be obtained as a weighted average of the rates corresponding to the distinct reactive structures. Assuming kinetic control, the C_{13}:C_{10} hydrogen abstraction ratio depends on the ratio between the corresponding rate constants, that is, roughly on the difference between the corresponding energy barriers. These energy barriers depend on many factors, not only on a unique geometric parameter (for instance, the distance from the abstracted hydrogen to the iron-bound hydroxyl).

In rabbit 15-LOX-1 the Ile418Ala mutation alters the position of arachidonic acid at the active site of 15-LOX-1 in such a way that the fatty acid substrate slides in deeper into the substrate binding pocket, which is consistent with the triad concept. This can be clearly seen as an example in Figure 6.5a. From a molecular point of view, this actually means that, again, a huge amount of AA positions exists. A lot of new AA positions (new structures) with the AA deeper into the active site are now accessible with the Ile418Ala mutation. However, there are different ways the tail of AA can penetrate in the active site, and not all of them involve a closer approach of H_{10} than H_{13} to the iron-bound hydroxyl (this is the case of Trajectory 3). However, among the amount of new AA structures now possible in the Ile418Ala mutant, many reactive structures appear where the corresponding H_{10} energy barriers are lower than the energy barriers corresponding to the H_{13} reactive structures (see Table 6.6). These reactive structures that favor the H_{10} abstraction did not exist in WT, and they appear in the Ile418Ala mutant just because AA slides in deeper into the substrate binding pocket due to the augment of free space at the bottom of the cavity (see clusters in Figure 6.6).

According to the results previously presented in this thesis, in wild-type rabbit 15-LOX-1 it was possible to find reactive structures suitable for H_{13} abstraction, others suitable for H_{10} abstraction, and even others appropriate for both reactive processes, according to the distance criterion applied, even when distances are not the unique (nor even the main) factor affecting the potential energy barrier. As the previous results of the Leu597Ala-Ile663Ala mutation confirm, it is the steric hindrance of these (and even other) bulky residues what provokes the difference observed in the potential energy barriers for the hydrogen abstraction process (in both WT and the double mutant). In the case of the Ile418Ala mutation, these bulky residues still play their role in hindering the evolution of hydrogen abstraction process, but this no longer happens to C_{10}, as it now occupies the former position of C_{13} as a consequence of the cavity bottom residue mutation.

Thus, the origin of the lower H_{10} energy barriers is that the motion of AA towards the bottom of the active site in the Ile418Ala mutant allows the appearance of reactive structures where C_{10} is in the region occupied by C_{13} in wild type (and with a similar conformation and environment), in such a way that the H_{10} abstraction in the Ile418Ala mutant can occur with similar easiness as H_{13} in WT (without steric disturbance). This produces lower energy barriers for H_{10} abstraction and drastically alters the positional specificity. That is, our calculations confirm at molecular level that the single Ile418Ala mutation converts a 15-lipoxygenating enzyme in a 12-lipoxygenating one as has been demonstrated experimentally by Prof. Kühn.

6.4 Conclusions

In this Chapter a study of the *in silico* mutant Ile418Ala combining MD simulations and QM/MM calculations has been carried out. The results suggest that AA adopt new types of conformations inside the active site where the distance distributions of C_{10} and C_{13} with respect to the oxygen atom of the Fe(III)-OH cofactor, move to lower values with respect to the WT system in the case of C_{10}, but C_{13} still presents a distance close enough to initiate the reactive process of H_{13} abstraction. This situation in which both positions are similarly close to the catalytic iron was also reflected in the MD simulations of the WT 15-LOX-1 that have been previously carried out in our group. However, the QM/MM potential energy barriers clearly reveal an inversion in the reaction specificity, as H_{10} abstraction is now more favored in the mutant that H_{13} abstraction: the difference between the average energy barrier for H_{10} in comparison with the average energy barrier for H_{13} (14.1 vs. 16.5 kcal/mol) is more than enough to explain the C_{13}:C_{10} hydrogen ratio of 8/92.

As explained by the triad concept, the substrate moves towards the end of the active site cavity as a consequence of the mutation, that enlarges the cavity channel. Traditionally, it was thought that the mutation provoked that C_{10} moved to closer distances than C_{13} causing the alteration in the regioselectivity. However, as has been demonstrated, distances are not the unique factor. More precisely, both carbon atoms C_{10} and C_{13} are close enough to the catalytic iron in terms of distances, and the mutation Ile418Ala provokes that C_{10} now occupies the position of C_{13} in WT. What happens here is that C_{13} position is less hindered during its pathway of hydrogen abstraction in the WT system than its counterpart C_{10}. The mutation itself allows now C_{10} occupy this less hindered position.

The results of this Chapter have been published in paper:

- *PNAS*, vol. 113, no. 30, pp. E4266-E4275, 2016

Chapter 7

5-LOX

7.1 Introduction

Human 5-LOX (5-LOX) is the most important member of the lipoxygenase family. Its importance lies in its anti-inflammatory activity: when it works alone, 5-LOX transforms AA into pro-inflammatory leukotrienes, but when it works in combination with 12/15-LOX, 5-LOX can lead to the synthesis of anti-inflammatory lipoxins by a transcellular process.[60,195,196] Thus, the conversion of 5-LOX into a 15-lipoxygenating enzyme has been proposed as an appealing strategy for the induction of lipoxins synthesis and therefore, for turning an inflammatory activity into an anti-inflammatory one.[18,50,51] So, the understanding of the factors that control the catalytic properties of LOXs and, particularly, their regio- and stereoselectivities, is of primary relevance for biomedicinal and biotechnological purposes. The work presented in this Chapter consists in the first computational study of the binding modes and reactivity of human 5-LOX with its natural substrate, arachidonic acid.

7.1.1 Description of the X-ray structures

As has been mentioned in the Introduction Chapter, there are several crystallographic structures of human 5-LOX including a substrate bound form.[29,30] These structures have been obtained for the so-called human Stable-5-LOX form of the enzyme, that consists in a stabilized and soluble mutant of h5-LOX.[29] A summary of the available X-ray structures along with their description is given in Table 7.1.

Only one of those structures (pdb code 3V99[30]) represents a substrate-bound form of 5-LOX (AA:Stable-5-LOX-Ser663Asp). Unfortunately, this crystal structure is poorly solved and lacks some protein regions. While all the substrate-free structures are highly similar (protein backbone RMSD =

Table 7.1: Description of the human 5-LOX crystallographic structures available in the Protein Data Bank.

PDB code	Description
3O8Y	Stable human 5-LOX
3V98	Stable human Ser663Asp 5-LOX
3V92	Stable human Ser663Val 5-LOX
3V99	Stable human Ser663Asp 5-LOX. Subunit A has been crystallized with the AA substrate

0.3 Å), the comparison with the substrate-bound form shows an enzyme with substantial plasticity. Between the unbound and bound systems the enzyme has suffered a major conformational change, mainly in the region of α-2 helix. Complex slow motions are therefore involved in the entire AA:LOX binding process. Hereafter, the substrate-free forms will be referred as apoenzyme or apo form, whereas the substrate-bound form will be referred as holoenzyme or holo form.

Based on the crystallographic data available of 5-LOX, two hypotheses have been proposed to explain the substrate access to the cavity: (i) access from the residues Phe177-Tyr181 (FY-cork), or, (ii) access from Trp147 which is at the opposite end (Figure 7.1a).[29] The way the substrate enters into the binding site will not be discussed here, but the way it is placed once it has entered. So, from here on head/tail-first orientation refers to the relative position of AA carboxylate end with respect to Lys409, which is supposed to play a similar role as Arg403 in rabbit 15-LOX,[28,197] in such a way that tail-first orientation means that AA carboxylate end points to the cavity end where Lys409 is placed, whereas head first means the opposite orientation.

In structure 3V99-A the substrate appears to be bound in a head-first orientation with its carboxylate group oriented towards the bottom of the cavity (at the opposite end of Lys409). However, as the authors themselves state,[30] this orientation has been modeled in a rather approximate manner since several of the X-ray coordinates of the protein are missing in its close vicinity (flexible loops from residues 172 to 176, 190 to 198, 294 to 299, 611 to 613, α helix from residues 414 to 429, and terminal Ile673 (Figure 7.1b). For this reason, the density of the carboxylate group of the substrate is not clearly defined, and as the authors caution about, the substrate orientation cannot be unambiguously determined.

The problem of the substrate orientation inside the active site cavity of 5-LOX is still a matter of discussion, as explained in the Introduction

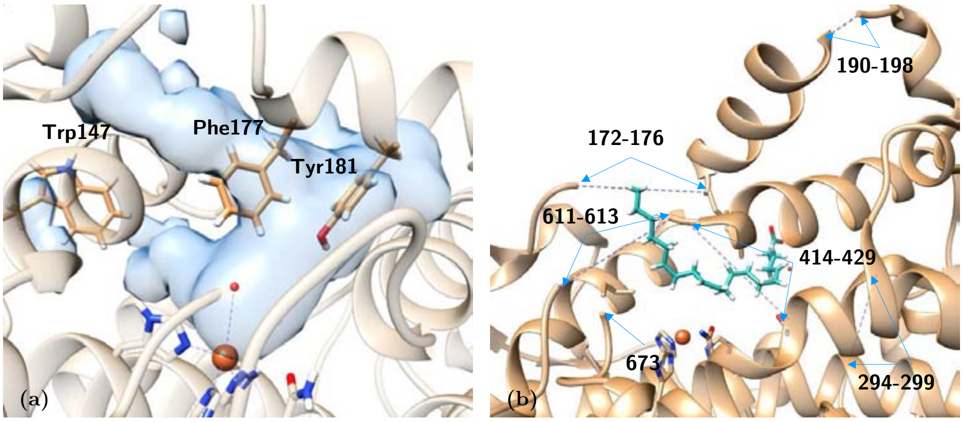


Figure 7.1: (a) Active site of structure 3O8Y, showing the residues proposed for substrate access into the binding site: residue Trp147 in one end, and residues Phe177 and Tyr181 (FY-cork) at the opposite end. (b) Representation of structure 3V99 (subunit A), showing the missing fragments with dashed lines; the AA binding mode inferred from the X-ray diffraction data is also shown.

Chapter. To date, there is not a clear experimental evidence that addresses this issue. The work presented here is aimed to determine the possible binding modes of the AA:5-LOX Michaelis complex consistent with its catalytic properties, that is, preference for pro-*S* H_7 abstraction. Due to the complexity of this system (large dimension, substrate of large flexibility, multiple abstraction sites, etcâĂę) a multiscale approach combining Homology Modeling, protein-ligand Dockings, Molecular Dynamics simulations and QM/MM calculations has been set up and applied. In concrete, a homology modeling study has first been performed to generate a model of the holo form of human Stable-5-LOX that completes the missing regions of the crystallographic structure (pdb code 3V99). Then a series of intensive protein-ligand docking experiments have been performed to generate AA:Stable-5-LOX complexes both in the holo model and the apo structure (pdb code 3O8Y). As AA is a very flexible ligand, the stability and conformational fluctuations of the AA:Stable-5-LOX complexes have been studied by Molecular Dynamics simulations. *In silico* models of the mutants involving a change of the cavity volume[50] have also been tested in order to determine the effect of these mutations in the positioning of AA in the binding site. Finally, the study of the hydrogen abstraction step in the different binding modes has been performed by QM/MM calculations.

7.2 Methodology

As mentioned above, to generate the final and complete model of the AA:5-LOX complex a multiscale approach combining homology modelling, protein-ligand docking and molecular dynamics simulations has been developed here. Finally, the reactivity studies have been performed by QM/MM calculations.

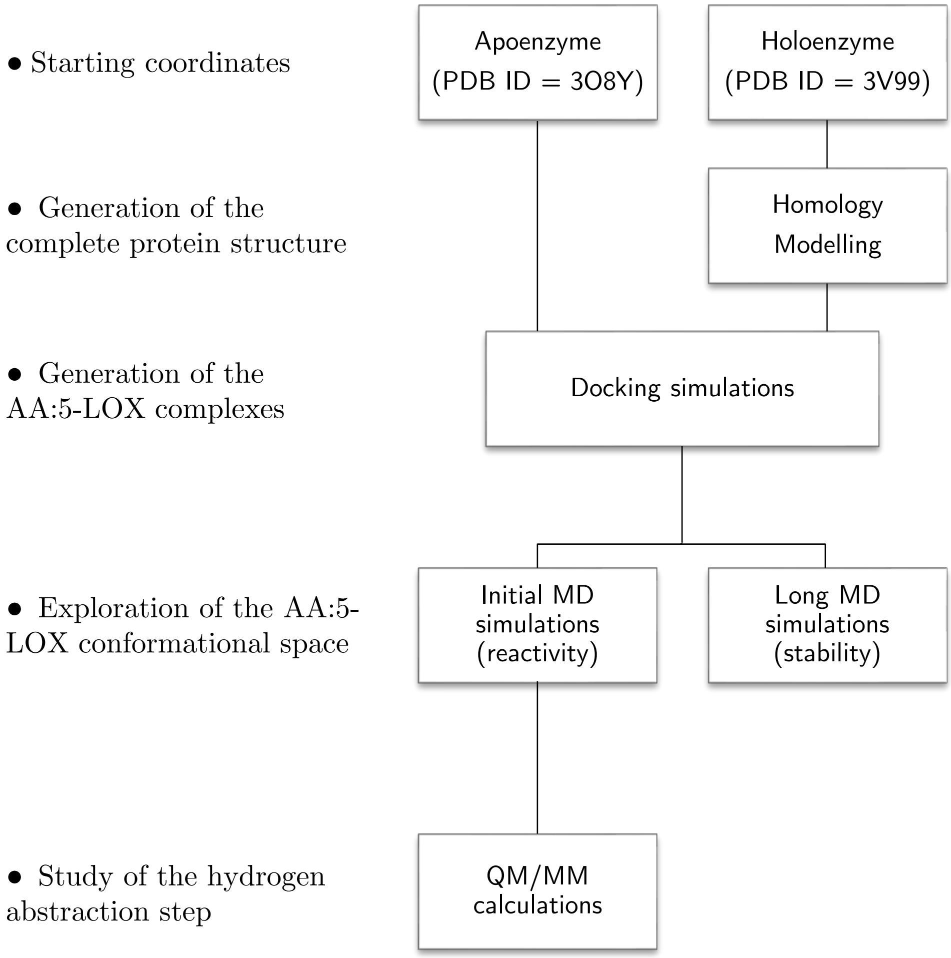


Figure 7.2: Scheme of the multiscale methodological procedure.

7.2.1 Homology Modeling

The first step of this study has consisted in building a valid and complete 3D model of the holo form of human Stable-5-LOX. All the calculations were carried out using the program Modeller version 9.12.[198] The strategy employed combines both loop modeling and structural overlapping.

As the purpose of this stage of the study is to complete the missing regions of the crystallographic structure of the holo form of the enzyme (pdb code 3V99) (i.e. we dispose of part of the structure, and we do not need to generate the complete 3D structure from the primary sequence), no templates from analogous enzymes were used to construct the model. It is important to notice that the 3V99 structure owns two different subunits (A and B), that both miss major structural pieces and only subunit A displays electron density associated with the presence of the substrate (Table 7.2), so, the same structure of subunit A of the pdb 3V99 was used as template. The choice of this template structure can be justified by the fact that it is actually the only structure of a holoenzyme available.

Table 7.2: Missing fragments in crystal structure 3V99. Indexes of the missing residues and its type of secondary structure or other structural features are provided.

Subunit	Missing segment	Characteristics
3V99-A	172-176	Loop
	190-198	Loop
	294-299	Loop
	414-429	α-helix
	611-613	Loop
	673	C-terminal Ile
3V99-B	172-176	Loop
	194-209	Loop
	414-429	Loop

Regarding all the loop regions that lack less than 9 residues, the loop refinement procedure built in Modeller was used. The non-default option “very slow” was used to ensure an exhaustive level of conformational sampling and to maintain an affordable time of calculation. The remaining region that comprises 414 to 429 residues presents a far wider gap and overcomes the limitation in the accuracy of the loop refinement algorithm provided by Modeller.[198] Therefore, this region of the model has been built using structural matching with one the apo structures of the same protein, particularly structure 3V92.[29]

Last but not least, the initial structure misses one of the coordinating residues of the iron, the terminal Ile673 which coordinates the Fe by its terminal -COO group. Therefore, a complete coordination sphere of the metal was modeled applying distance restraints during the modeling. Precisely, distances have been restrained to a value of 2.04 Å between the Fe and its coordinating atoms. Asn554 residue, which is not reported among the Fe-coordinating residues in the original paper of Stable 5-LOX crystallographic structure, has been included anyway in the coordination sphere as this residue appears in the proximity of the Fe and has been found in other LOX participating in the metal coordination environment.

In total, the Modeller runs were constituted by the generation of 20 models followed by 20 steps of loop refinement. The full models were generated with a degree of refinement of Modeller (option “slow") that includes a thorough Molecular Dynamics and Simulated Annealing procedure as part of the model optimization. Then, the 20 runs of loop modeling process used a more extensive refinement procedure (option “slow large"). Finally, the most encouraging structures were minimized at MM level with a CHARMM full-atom force field, and the quality of the resulting models were assessed with ERRAT,[199] VERIFY-3D,[200] and the PROCHECK Ramachandran plot.[201]

7.2.2 Docking Simulations

Docking simulations were run with the program GOLD5.2.[149] Two different receptors were used, the subunit A of structure 3O8Y, corresponding to the apoenzyme, and the final model of the holoenzyme generated by homology modeling as described above. The option of GOLD to deal with the interaction of metal ions of metalloenzymes with organic ligands was active during the docking simulations although limited at the search of hexacoordinated geometries of the metal. Prior to docking, the protonation states of the titrable residues were determined using the PROPKA server.[202] Simulations were carried out for both rigid and flexible side chains receptors in all cases, except for the mutants of the holoenzyme (see below) for which only rigid receptors were used.

To ensure a complete exploration of the conformational space for AA, the maximum efficiency of the genetic algorithm was selected in the docking setup. The binding site cavity of the receptor consisted in a 19 Å radius sphere centered in the Fe atom. The fitness function employed was the Chemscore function.[154] The parameters of the genetic algorithm were set to the default values of the program. In case of considering local rearrangements of the LOX binding site, different rotameric schemes were tested selecting the rotamers from the GOLD library.[203]

Finally, the solutions were analyzed on the basis of the consistency with catalytic processes. As such, the full benches of docking solutions of each experiment were analyzed throughout a clustering process as well as filtered in function of the distances between the reactive oxygen of the Fe(III)-OH$^-$ cofactor (and the other cofactor, Fe(III)-OH$^-$) and the different unsaturated carbons of the substrate that could be involved in the mechanism (C_7, C_{10} and C_{13}).

7.2.3 *In silico* mutants

The different *in silico* mutants have been generated by changing the corresponding residues with the Mutator Plugin of the VMD program.[172] In the case of the F359W mutation, it has been observed that the side-chain of

Tryptophan is too large to fit in the cavity left by the original Phe359 residue, and its structure is distorted when it is mutated by simple substitution of the side-chain of Phenylalanine. So, for all mutants containing this mutation, the position of the side chain of W359 was optimized by CHARMM Molecular Mechanics (MM) optimization.[88, 161, 162]

7.2.4 Molecular Dynamics Simulations

7.2.4.1 Initial MD simulations: reactivity studies

In order to generate molecular configurations for each of the three AA:Stable-5-LOX Michaelis complexes considered in this work for subsequent QM/MM calculations, and with a special focus on the distances between the potential reacting groups (i.e., the oxygen of the OH^- group and C_7, C_{10} or C_{13} of AA), the following protocol was applied. The system was fully solvated in a cubic box of preequilibrated TIP3P water molecules, with dimensions of 115 Å X 115 Å X 115 Å. The total charge of the system was neutralized by adding 22 sodium cations. The resulting system contained nearly 143000 atoms, of which about 10600 belong to the protein.

The system was submitted to 1500 MM minimization steps with harmonic restrains applied to the protein and the substrate, followed by 1500 steps without restrains (in both cases with the steepest descent algorithm) to avoid bad contacts and then, MD simulations under periodic boundary conditions (PBC) were started. The system was gradually heated from 25 K to 300 K during 80 ps, followed by an equilibration step of 400 ps at 300 K at constant volume (NVT). During the heating steps, harmonic restraints were applied to the protein, which were gradually released, while the ligand (and water molecules) was kept free of restraints. Finally, a production step of 20 ns was carried out in order to relax the system and to generate a set of snapshots ready for the hydrogen abstraction reactions. Production steps were carried out at constant pressure and temperature (NPT) using the extended system constant pressure and the Hoover constant temperature algorithms.[186]

A time step of 2 fs was used for the production steps. All of the bonds and angles involving hydrogen atoms were constrained by the SHAKE algorithm.[126] PBCs were built with the CRYSTAL facility of CHARMM. For long-range electrostatics, the PME method was used. MD simulations were run with CHARMM version c35-b1.[88] The force field topology and parameters for AA were derived from the CHARMM27 ones.[163, 164] CHARMM22[161, 162] force field was used for protein atoms but Fe and its first coordination sphere, for which the parameters specifically derived by Saam *et al.* were used,[187] updating the parameters for the coordination with Asn554.

Structural analyses of the simulations were performed with the CORREL module of CHARMM.

7.2.4.2 Long MD simulations: stability of the AA:5-LOX complexes

Long molecular dynamics simulations with production runs of 100 ns were run for each of the three AA:Stable-5-LOX complexes considered in this work. The details of the calculations were the same as described above, but with smaller solvation boxes of sizes of 118 Å x 93 Å x 73 Å for the holoenzyme complex, and 118 Å x 80 Å x 71 Å for the apoenzyme complex. The differences in box sizes originates in the different conformation of the apo and holo complexes, as the holoenyme displays a more extended conformation requiring a larger solvation box. The size of the *current* box has been reduced compared with the previous cubic box due to the limitations in computational cost and time of the more-demanding 100 ns simulation.

7.2.5 QM/MM Calculations

Using the trajectories of the Michaelis complex obtained by MD simulations, a set of snapshots have been selected to study the first step of the reaction mechanism, the hydrogen abstraction step, by QM/MM calculations. All QM/MM calculations have been performed with the modular program package ChemShell,[105] using TURBOMOLE[159] for all of the DFT calculations. The MM calculations were carried out by the DL-POLY module[160] in ChemShell, using the CHARMM22[161,162] and CHARMM27[163,164] (for the lipid moiety included in the MM region) force fields. An electronic embedding scheme[98] has been adopted in all calculations, and a hydrogen link atoms scheme has been employed to treat the QM/MM boundary using the charge shift model. No cutoffs were introduced for the nonbonding MM and QM/MM interactions.[165] QM/MM optimizations were carried out employing the L-BFGS algorithm[166,167] in the case of energy minimizations, and the microiterative optimizer, combining both P-RFO[168,169] and L-BFGS, was used during the transition state structure searches. All of these algorithms are implemented in the HDLCopt[170] ChemShell module. For the QM subsystem the hybrid functional B3LYP was used with the 6-31G(d) Pople basis set[173] for all atoms except Fe, which was described by the LANL2DZ basis set.[174] QM region includes 64 atoms (link atoms not included): 11 atoms of each of the 3 His residues in the Fe coordination sphere (His367, His372, His550), 8 atoms of Asn554, 3 atoms of the terminal Ile673 in the coordination sphere, and the Fe(III)âĂŤOH cofactor, able to accept the hydrogen from AA. For every case, a QM region including the pentadienyl group around the methylene of interest, and the adjacent C (and its corresponding hydrogens) atoms was chosen, containing a total of 17 QM atoms of the lipid substrate. Seven link atoms were used, five along the bonds CαâĂŞQM atom of the five residues in the Fe coordination sphere and two in the AA chain. Figure 7.3 shows the QM/MM partition.

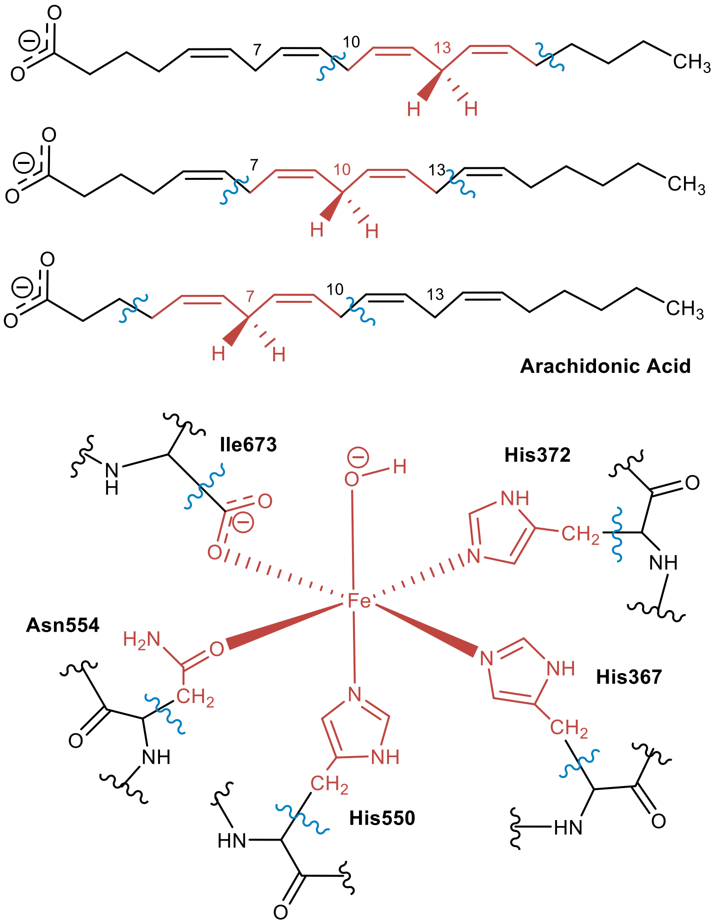


Figure 7.3: QM/MM partition. QM atoms are depicted in red, and link atoms are represented by wavy blue lines. For each process a different partition of AA has been used, and the top of the figure shows the three options employed, for H_{13} abstraction (upper), H_{10} abstraction (middle) and H_7 abstraction (lower).

7.3 Results and Discussion

7.3.1 Homology Modeling

7.3.1.1 Construction of the holoenzyme model

The use of the only available X-ray structure of a holoenzyme to generate a binding mode of AA into human-5-LOX compatible with hydrogen atom abstraction at C_7 (the one leading to 5-lipoxygenation), presented the problem that this structure lacks the coordinates of some regions. Thus, homology modeling techniques have been employed to build these missing regions. A total of 400 3D models of the holo form of Stable-5-LOX were produced in this part of the study accounting both for modeling and loop refinement processes. The best models were first analyzed based on the energies provided by Modeller, and the 3 lowest energy solutions were finally selected, as those values were not discriminative enough. For each of these models, the 3 lowest energy loop models were also selected which lead to a total of 9 model candidates that were further refined and analyzed.

First, a CHARMM MM optimization of the side chains was performed in order to reduce possible constraints (bad contacts) remaining from the modeling process. Then, the quality of the 9 minimized models was evaluated by ERRAT (accounting for the quality of residues side-chains), VERIFY-3D (quality of the global 3D structure), and the PROCHECK (which evaluates the Ramachandran plot). In all cases, the final values of each geometry checker were higher than 90% and suggest that all candidates are potentially good models for human-Stable-5-LOX (Table 7.3). The final model was selected using a consensus between both energetic and structural qualities.

Table 7.3: Homology models quality evaluation. Model indexes shown in Table (first column) are those provided by the Modeller calculation. The best solution (seventh row), selected by a consensus among the three checkers, is highlighted in bold.

Model	ERRAT	VERIFY-3D	PROCHECK
20003	90.35	92.79	91.50
30003	89.98	93.71	91.70
170003	90.32	92.18	91.50
80007	93.23	93.56	92.00
110007	93.20	93.87	92.00
120007	93.21	93.71	92.00
20019	**94.46**	**94.48**	**91.50**
100019	93.99	92.79	92.20
190019	94.45	91.41	92.20

According to these results model 7 (index 20019 according to Modeller nomenclature) is the best solution and thus, it has been selected to perform the following up studies. Figure 7.4 shows this final model.

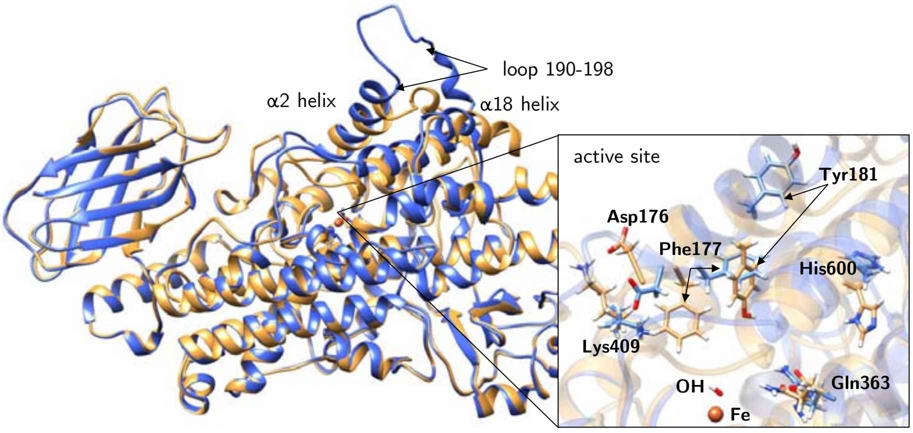


Figure 7.4: a) Superposition of the holoenzyme model generated by homology modeling (blue) and the apoenzyme crystal structure (ochre), highlighting the principal differences between the two structures. The box on the left shows a detail of the active sites of both structures, highlighting the residues that present more significant differences in their side-chain orientations.

7.3.1.2 Comparison between apoenzyme and holoenzyme structures

The most remarkable difference between the Stable-5-LOX-apoenzyme crystal structure and the Stable-5-LOX-holoenzyme model generated by homology modeling (hereafter the model will be referred as the holoenzyme for simplicity) is the position of the α-2 helix. In the structure of the 5-LOX apoenzyme, α-2 helix comprises residues 176-193 and presents a geometry that is unique in 5-LOX with respect to other LOX isoforms such as 15-LOX and 8R-LOX. In the structure of the 5-LOX holoenzyme, α-2 connects with an unfolded loop (residues 190-198) that displays a very wide solvent-exposed conformation. In the corresponding apo form, this loop has packed and folded toward the cavity. A detailed representation of the loop structure generated by the homology modeling process is shown in Figure 7.5.

There are also significant differences in side chain orientation of some residues of the active site (see Figure 7.4), principally in Asp176, Phe177, Tyr181, Leu607 (*add to the picture*), Lys409, Gln363, Arg592 and His600. The cases of Lys409 and Phe177 stand out. Lys409 is a positively charge residue that could establish a favoring electrostatic interaction with the negative carboxylate group of the fatty acid substrate. Lys409 forms a salt bridge with Asp176 and both residues are pointing away from the active site in the

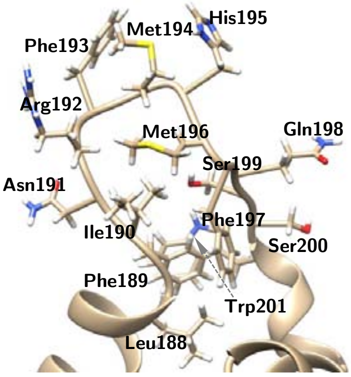


Figure 7.5: Detail of loop 190-198 showing residues side chains in the holoenzyme model.

apoenzyme; therefore, Lys409 cannot tie the carboxylate end of AA inside the binding pocket. However, in the holoenzyme this salt bridge is placed towards the active site allowing Lys409 to establish a possible interaction with AA inside the cavity. Phe177 and Tyr181, which belong to α-2 helix, form a “cork" that blocks the cavity entrance in the apoenzyme, but in the holoenzyme their positioning is in a different orientation that would allow the entrance of the substrate to the binding pocket. Another important difference between the apo and holo forms, is an apparently slight repositioning of α-18 helix, that can affect somehow the active site shape and volume. Arg596 and His600, which belong to this helix, present different side-chain conformations in apoenzyme and holoenzyme.

7.3.2 Generation of AA:Stable-5-LOX complexes

As has been previously commented, in order to generate the AA:Stable-5-LOX complexes compatible with catalysis the use of the holoenzyme structure (with the missing parts modeled as explained above) seemed to be the most reliable choice, but on the other hand, and despite the substantially packed binding site of the apoenzyme structure, the possibility of this conformation to bare some activity cannot easily be ruled out. Thus, docking calculations were performed in both apoenzyme and holoenzyme structures. In both cases two different approaches were employed, using both rigid and flexible side chain approaches in order to account for small conformational changes that could improve the quality of the docking solutions.

7.3.2.1 Rigid holoenzyme

First, docking simulations were performed on holoenzyme structure with an initial approximation of a rigid receptor and, consequently, with identical geometry of the model generated by homology modelling. For this approach, a total of 200 runs of the genetic algorithm docking program GOLD were performed. The scoring values of the resulting solutions ranged from 34.30 to 29.01 Chemscore units. However, they display an apparently wide geometric variability and were therefore further analyzed under filtering processes. The first one was conducted with a clustering protocol and using an RMSD (of the substrate heavy atoms) with a cutoff of 5.88 Å. This leads to 7 main clusters (see Figure 7.6). The scoring function value of each representative model of the clusters is provided in Table 7.4.

Table 7.4: Scoring function (in ChemScore units) of the representative models and cluster population (with RMSD = 5.88 Å) of the docking simulations of the rigid holoenzyme receptor.

Cluster number	Score	Population
1	34.30	16
2	34.26	53
3	33.97	107
4	33.25	3
5	32.34	18
6	30.41	2
7	30.32	1

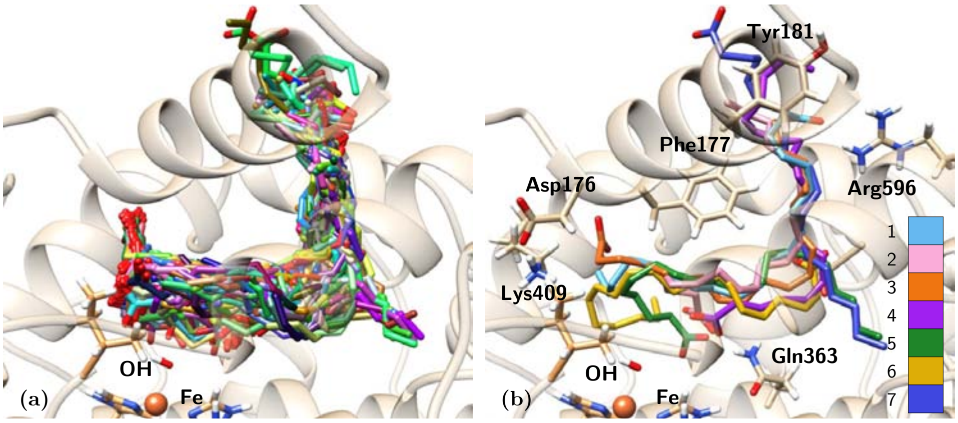


Figure 7.6: (a) Solutions of the docking simulation of the holoenzyme model with rigid side chains; (b) Representative models of clustering with RMSD = 5.88 Å, showing residues of interest. The color scale represented in (b) assigns a different color to the representative element of each cluster.

The binding modes of AA were determined by the presence of polar and hydrogen-bond donor residues of the protein in H-bond distances to the oxygen atoms of the AA carboxylate. Five residues fulfill this condition: Lys409, Gln363, Asn425, Arg596, and Phe177 (its backbone). The most populated cluster (cluster 3) includes solutions with AA carboxylate end bound to Lys409. In total, there are 116 solutions with H-bond to Lys409, and they exhibit a tail-first orientation. Solutions with AA carboxylate end close to Gln363 are less numerous, 34 solutions, and display the AA tail turned outside the Fe(III)-OH$^-$ cofactor, pointing to Arg596 (Figure 7.6).

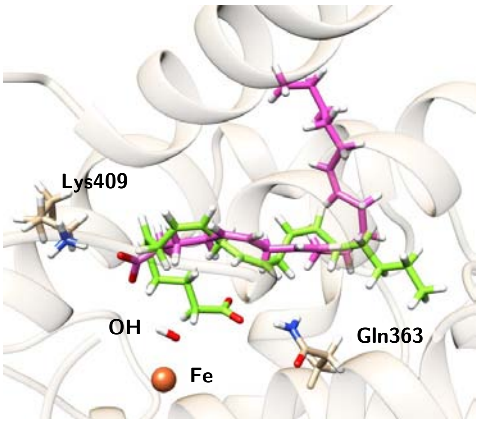


Figure 7.7: Representation of the lowest energy solutions displaying d(C$_7$-OH) < 5 Å obtained from docking simulations of rigid holoenzyme model: in tail-first orientation (pink), and head-first orientation with the AA carboxylate end close to Gln363 (green).

Next, to better analyze the dockings into the context of pre-reactive orientations, the results were filtered in function of the distances between the reactive groups of the substrate and the Fe(III)-OH$^-$ moiety. The analyses of the distances from carbons C$_7$, C$_{10}$ and C$_{13}$ of the AA to the oxygen of the Fe(III)-OH$^-$ cofactor show that there is a reasonable number of solutions with d(C$_7$-OH) < 5 Å (29 solutions), whereas there are few solutions with d(C$_{10}$-OH) or d(C$_{13}$-OH) < 5 Å (4 solutions in both cases). Based purely on a distance criterion, these results are clearly consistent with the regiospecificity of 5-LOX. Thus, this holoenzyme structure could in principle correspond to the catalytically productive form of 5-LOX. Importantly, solutions with d(C$_7$-OH) < 5 Å belong to clusters 3 and 5, and have residues Lys409, Phe177(backbone) and Gln363 in the vicinity of the AA carboxylate end. The best energy solution with d(C$_7$-OH) < 5 Å which belongs to cluster 3, has a tail-first orientation with the AA carboxylate close to Lys409 (Figure 7.7) and Phe177. On the contrary, the best energy solution with the AA carboxylate end close to Gln363 and d(C$_7$-OH) < 5 Å belongs to cluster 5 and is energetically less favorable by 1.39 Chemscore units. Hence, these calculations indicate (this first approach) that tail-first orientation is preferred from a catalytic perspective. Figure 7.7 shows these two docking poses.

7.3.2.2 Flexible side-chains holoenzyme

An improved exploration of the conformational space of the AA:5-LOX complexes has been performed by the use of rotameric schemes. Rotamers allow the model to overcome possible low energy local arrangements of the crystal structure to gain activity and predict binding modes of the substrate that otherwise could not be explored. Residues Asp176, Phe177, Asn180, Gln363, Lys409, Asn425, Arg596, Leu607 and Ile673 were set as flexible using GOLD rotamer library. In this case, a number of 500 runs of the genetic algorithm docking program GOLD were performed to better account for the large number of degrees of freedom of the system and the vast cavity dimension of the binding site of 5-LOX. The scoring function results ranged from 33.03 to 26.41 in Chemscore units, thus indicating that no improvement of binding energies are predicted. The solutions were clustered with an RMSD of 10.30 Å, giving 8 clusters. This increase of the RMSD value illustrates the wide conformational variability that offers the formation of the complex between AA and the binding site of the enzyme (see Figure 7.8 a)). The scoring function value for the representative model of each cluster is shown in Table 7.5.

Table 7.5: Scoring function (in ChemScore units) of the representative models and cluster population (with RMSD = 10.30 Å) of the docking simulations of the holoenzyme receptor with flexible side chains

Cluster number	Score	Population
1	33.03	23
2	31.47	349
3	31.32	34
4	31.15	20
5	30.99	50
6	30.17	7
7	29.73	16
8	26.41	1

Using such a large number of rotamer residues provokes the generation of non-productive solutions in which the substrate falls away from the binding site (as shown in Figure 7.8), but allows better exploration of productive configurations inside the cavity. Only solutions with C_7 close enough for catalysis are analyzed.

Positive residues in H-bond distance to the AA carboxylate group are Lys409 and the backbone of Phe177, Gln363, Gln557, Asn425, Gln611, Asn613 and Arg596. There are about 300 solutions in the vicinity of Gln611, but these solutions correspond to the non-productive ones mentioned in the previous paragraph. In those poses the substrate is very far from the cat-

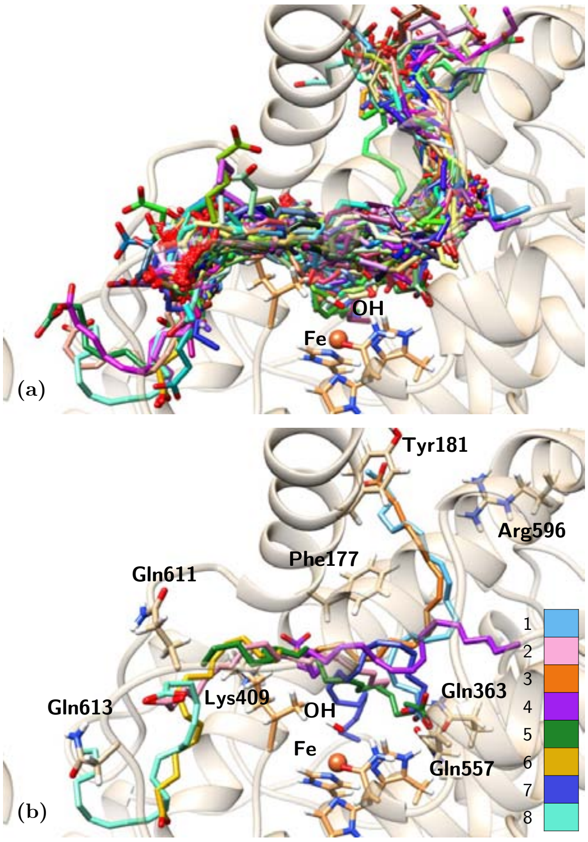


Figure 7.8: (a) Solutions of the docking simulation of the holoenzyme model with flexible side chains; (b) Representative models of clustering with RMSD = 10.33 Å. The color scale represented in (b) assigns a different color to the representative element of each cluster.

alytic iron and for this reason they are not considered in further analyses. Solutions bonded to Lys409 and with hydrophobic contacts with Phe177 (15 solutions) exhibit a tail-first type orientation, while solutions interacting with Gln363 and Gln557, which is spatially close to Gln363 (12 solutions), present a head-first type orientation. Flexible side-chains allow the AA methyl end to be located inside the binding cavity when the carboxylate end is bound to Gln363 (Gln557) (see Figure 7.8b), allowing head-first orientation positioning, in contrast to the case of the rigid holoenzyme (see Figure 7.7), in which solutions H-bonded to Gln363 had the AA tail pointing outward to the cavity.

Analysis of distances to the methylenes of interest shows that there are 47 solutions with C_7 close to the oxygen of the Fe(III)-OH$^-$ cofactor. In this group there are both solutions with head-first and tail-first orientation with good values of the scoring function. Figure 7.9 shows one representative example of each type of orientation. Solution 34, which belongs to cluster

1, is interacting with Lys409 (tail first) and has a $d(C_7\text{-OH})$ of 3.74 Å. Solution 12, with the opposite orientation (head first), has a $d(C_7\text{-OH})$ of 4.25 Å. These two structures will be used to study the regiospecificity of the products expected for these two different binding orientations (tail/head first) in the holoenzyme structure.

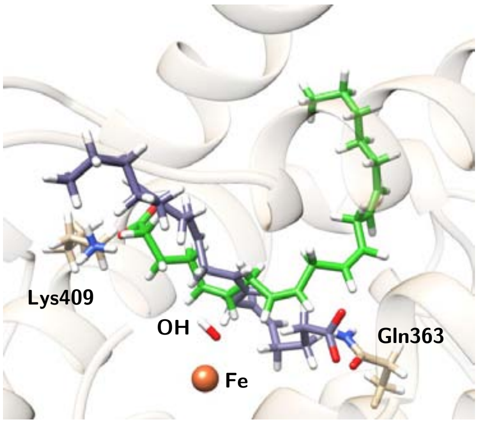


Figure 7.9: Representation of solutions 12 (head first, purple-gray) and 34 (tail first, green) of docking simulations of flexible side-chains holoenzyme model.

In Figure 7.10 the cavity lining residues for the selected head-tail AA orientation in the holoenzyme are indicated. Several invariant amino acids that have been shown to interact with AA in other LOXs are used in 5-LOX, although there are also some specific 5-LOX residues defining this binding site. A detailed comparison of 5-LOX with other LOX isoforms is provided by the alignment of cavity lining residues. To perform this comparison of 5-LOX with other LOX isoforms, the degree of conservation of several active site aminoacids has been provided by the results of Multiple Sequence Alignment (MSA) performed with the CLUSTAL-Ω server. Invariant or highly conserved, and 5-LOX specific residues are specified in the alignment chart (Figure 7.11). Note that human 5-LOX is compared with other members of the LOX family with different reaction specificity, and also belonging to different organisms. It is interesting that Lys409 is specific of 5-LOX (present in human and mouse 5-LOXs), and aligns with Arg403 of rabbit and human 15-LOX-1, which also fulfills this function of anchoring the substrate carboxylate end.

7.3.2.3 *In silico* mutants

To better confront our modeling to experimental data on 5-LOX, we intended to rationalize mutagenetic works previously published.[18,50] Mutations were first done on residues Ala424, Asn425 and Phe359, which are placed at the bottom of the active site cavity and have been proposed to influence 5-

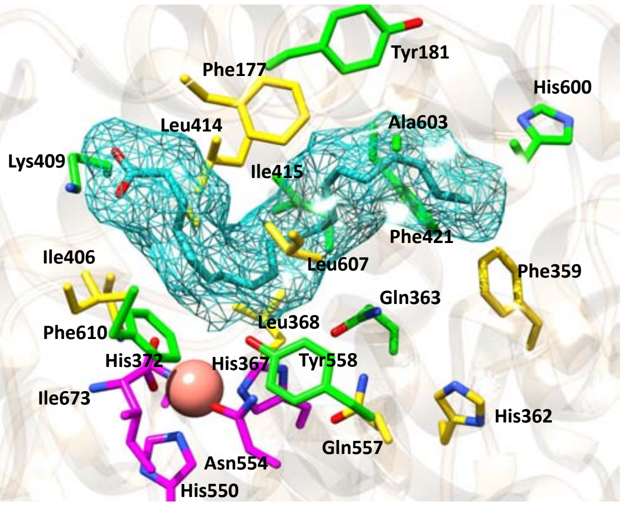


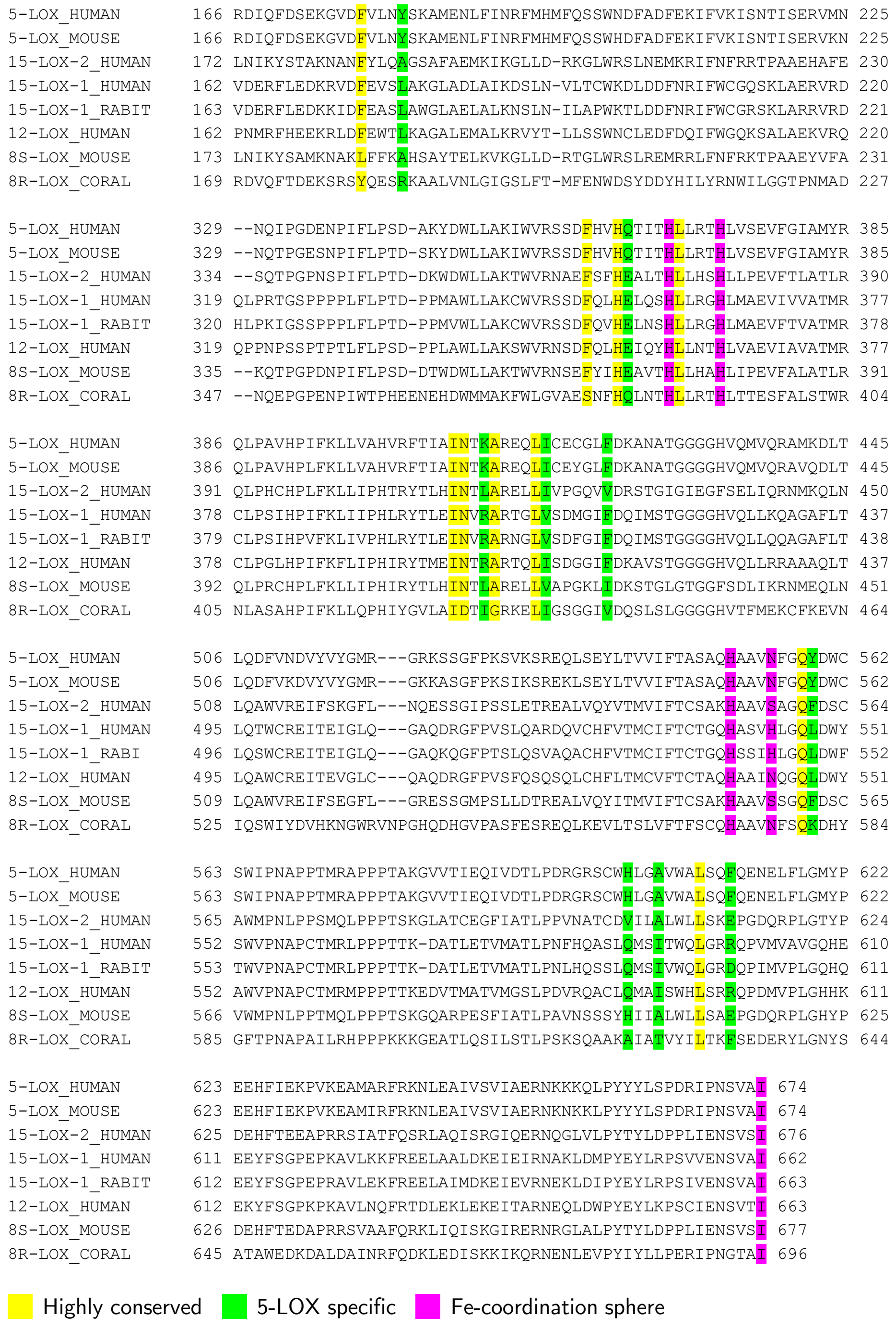
Figure 7.10: Representation of the lining cavity residues in the holo tail-first AA:5-LOX complex. Highly conserved residues are depicted in yellow, 5-LOX specific residues in green, and the Fe-coordinating residues in magenta (see also the MSA chart).

LOX regioselectivity by decreasing the volume of the binding active site. In particular, the Phe359Trp-Ala424Ile-Asn425Met triple mutant of human 5-LOX has been reported to yield a 42% of 15-lipoxygenation (hydrogen abstraction at C_{13}) and 53% of 8-lipoxygenation (hydrogen abstraction at C_{10}).

On the other hand, in the above docking simulations residues Lys409 and Gln363 seemed to play a role in AA carboxylate end binding; thus, we have performed additional docking simulations for several mutations of those residues in the holoenzyme. Table 7.6 shows the results for all the tested mutants. For each mutant, 200 runs of the genetic algorithm docking program GOLD were performed. Based on the previous docking simulations in which the introduction of flexibility to the amino acids did not substantially change the chemical interpretation of the resulting solutions, in all this virtual mutagenetic study only rigid receptors were used.

The difference in the active site cavity volumes between Stable-5-LOX and triple F359W-A424I-N425M mutant has been estimated by measuring the volume of the involved residues 359, 424 and 425, resulting in a decrease of 40.1 Å^3 by effect of the mutation (see Figure 7.12).

When the mutation implicates residues F359W-A424I-N425M the number of docking solutions interacting with Lys409 decreases (from 116 in WT to 5 in the triple mutant) as the number of mutated residues of the bottom of the cavity increases. Solutions bound to Lys409 seem to need a larger volume of the cavity to accommodate the AA tail. The three residues (359, 424 and 425) together block the open channel where the Arg596 is



```
5-LOX_HUMAN       166 RDIQFDSEKGVDFVLNYSKAMENLFINRFMHMFQSSWNDFADFEKIFVKISNTISERVMN 225
5-LOX_MOUSE       166 RDIQFDSEKGVDFVLNYSKAMENLFINRFMHMFQSSWHDFADFEKIFVKISNTISERVKN 225
15-LOX-2_HUMAN    172 LNIKYSTAKNANFYLQAGSAFAEMKIKGLLD-RKGLWRSLNEMKRIFNFRRTPAAEHAFE 230
15-LOX-1_HUMAN    162 VDERFLEDKRVDFEVSLAKGLADLAIKDSLN-VLTCWKDLDDFNRIFWCGQSKLAERVRD 220
15-LOX-1_RABIT    163 VDERFLEDKKIDFEASLAWGLAELALKNSLN-ILAPWKTLDDFNRIFWCGRSKLARRVRD 221
12-LOX_HUMAN      162 PNMRFHEEKRLDFEWTLKAGALEMALKRVYT-LLSSWNCLEDFDQIFWGQKSALAEKVRQ 220
8S-LOX_MOUSE      173 LNIKYSAMKNAKLFFKAHSAYTELKVKGLLD-RTGLWRSLREMRRLFNFRKTPAAEYVFA 231
8R-LOX_CORAL      169 RDVQFTDEKSRSYQESRKAALVNLGIGSLFT-MFENWDSYDDYHILYRNWILGGTPNMAD 227

5-LOX_HUMAN       329 --NQIPGDENPIFLPSD-AKYDWLLAKIWVRSSDFHVHQTITHLLRTHLVSEVFGIAMYR 385
5-LOX_MOUSE       329 --NQTPGESNPIFLPTD-SKYDWLLAKIWVRSSDFHVHQTITHLLRTHLVSEVFGIAMYR 385
15-LOX-2_HUMAN    334 --SQTPGPNSPIFLPTD-DKWDWLLAKTWVRNAEFSFHEALTHLLHSHLLPEVFTLATLR 390
15-LOX-1_HUMAN    319 QLPRTGSPPPPLFLPTD-PPMAWLLAKCWVRSSDFQLHELQSHLLRGHLMAEVIVVATMR 377
15-LOX-1_RABIT    320 HLPKIGSSPPPLFLPTD-PPMVWLLAKCWVRSSDFQVHELNSHLLRGHLMAEVFTVATMR 378
12-LOX_HUMAN      319 QPPNPSSPTPTLFLPSD-PPLAWLLAKSWVRNSDFQLHEIQYHLLNTHLVAEVIAVATMR 377
8S-LOX_MOUSE      335 --KQTPGPDNPIFLPSD-DTWDWLLAKTWVRNSEFYIHEAVTHLLHAHLIPEVFALATLR 391
8R-LOX_CORAL      347 --NQEPGPENPIWTPHEENEHDWMMAKFWLGVAESNFHQLNTHLLRTHLTTESFALSTWR 404

5-LOX_HUMAN       386 QLPAVHPIFKLLVAHVRFTIAINTKAREQLICECGLFDKANATGGGGHVQMVQRAMKDLT 445
5-LOX_MOUSE       386 QLPAVHPLFKLLVAHVRFTIAINTKAREQLICEYGLFDKANATGGGGHVQMVQRAVQDLT 445
15-LOX-2_HUMAN    391 QLPHCHPLFKLLIPHTRYTLHINTLARELLIVPGQVVDRSTGIGIEGFSELIQRNMKQLN 450
15-LOX-1_HUMAN    378 CLPSIHPIFKLIIPHLRYTLEINVRARTGLVSDMGIFDQIMSTGGGGHVQLLKQAGAFLT 437
15-LOX-1_RABIT    379 CLPSIHPVFKLIVPHLRYTLEINVRARNGLVSDFGIFDQIMSTGGGGHVQLLQQAGAFLT 438
12-LOX_HUMAN      378 CLPGLHPIFKFLIPHIRYTMEINTRARTQLISDGGIFDKAVSTGGGGHVQLLRRAAAQLT 437
8S-LOX_MOUSE      392 QLPRCHPLFKLLIPHIRYTLHINTLARELLVAPGKLIDKSTGLGTGGFSDLIKRNMEQLN 451
8R-LOX_CORAL      405 NLASAHPIFKLLQPHIYGVLAIDTIGRKELIGSGGIVDQSLSLGGGGHVTFMEKCFKEVN 464

5-LOX_HUMAN       506 LQDFVNDVYVYGMR---GRKSSGFPKSVKSREQLSEYLTVVIFTASAQHAAVNFGQYDWC 562
5-LOX_MOUSE       506 LQDFVKDVYVYGMR---GKKASGFPKSIKSREKLSEYLTVVIFTASAQHAAVNFGQYDWC 562
15-LOX-2_HUMAN    508 LQAWVREIFSKGFL---NQESSGIPSSLETREALVQYVTMVIFTCSAKHAAVSAGQFDSC 564
15-LOX-1_HUMAN    495 LQTWCREITEIGLQ---GAQDRGFPVSLQARDQVCHFVTMCIFTCTGQHASVHLGQLDWY 551
15-LOX-1_RABI     496 LQSWCREITEIGLQ---GAQKQGFPTSLQSVAQACHFVTMCIFTCTGQHSSIHLGQLDWF 552
12-LOX_HUMAN      495 LQAWCREITEVGLC---QAQDRGFPVSFQSQSQLCHFLTMCVFTCTAQHAAINQGQLDWY 551
8S-LOX_MOUSE      509 LQAWVREIFSEGFL---GRESSGMPSLLDTREALVQYITMVIFTCSAKHAAVSSGQFDSC 565
8R-LOX_CORAL      525 IQSWIYDVHKNGWRVNPGHQDHGVPASFESREQLKEVLTSLVFTFSCQHAAVNFSQKDHY 584

5-LOX_HUMAN       563 SWIPNAPPTMRAPPPTAKGVVTIEQIVDTLPDRGRSCWHLGAVWALSQFQENELFLGMYP 622
5-LOX_MOUSE       563 SWIPNAPPTMRAPPPTAKGVVTIEQIVDTLPDRGRSCWHLGAVWALSQFQENELFLGMYP 622
15-LOX-2_HUMAN    565 AWMPNLPPSMQLPPPTSKGLATCEGFIATLPPVNATCDVILALWLLSKEPGDQRPLGTYP 624
15-LOX-1_HUMAN    552 SWVPNAPCTMRLPPPTTK-DATLETVMATLPNFHQASLQMSITWQLGRRQPVMVAVGQHE 610
15-LOX-1_RABIT    553 TWVPNAPCTMRLPPPTTK-DATLETVMATLPNLHQSSLQMSIVWQLGRDQPIMVPLGQHQ 611
12-LOX_HUMAN      552 AWVPNAPCTMRMPPPTTKEDVTMATVMGSLPDVRQACLQMAISWHLSRRQPDMVPLGHHK 611
8S-LOX_MOUSE      566 VWMPNLPPTMQLPPPTSKGQARPESFIATLPAVNSSSYHIIALWLLSAEPGDQRPLGHYP 625
8R-LOX_CORAL      585 GFTPNAPAILRHPPPKKKGEATLQSILSTLPSKSQAAKAIATVYILTKFSEDERYLGNYS 644

5-LOX_HUMAN       623 EEHFIEKPVKEAMARFRKNLEAIVSVIAERNKKKQLPYYYLSPDRIPNSVAI 674
5-LOX_MOUSE       623 EEHFIEKPVKEAMIRFRKNLEAIVSVIAERNKNKKLPYYYLSPDRIPNSVAI 674
15-LOX-2_HUMAN    625 DEHFTEEAPRRSIATFQSRLAQISRGIQERNQGLVLPYTYLDPPLIENSVSI 676
15-LOX-1_HUMAN    611 EEYFSGPEPKAVLKKFREELAALDKEIEIRNAKLDMPYEYLRPSVVENSVAI 662
15-LOX-1_RABIT    612 EEYFSGPEPRAVLEKFREELAIMDKEIEVRNEKLDIPYEYLRPSIVENSVAI 663
12-LOX_HUMAN      612 EKYFSGPKPKAVLNQFRTDLEKLEKEITARNEQLDWPYEYLKPSCIENSVTI 663
8S-LOX_MOUSE      626 DEHFTEDAPRRSVAAFQRKLIQISKGIRERNRGLALPYTYLDPPLIENSVSI 677
8R-LOX_CORAL      645 ATAWEDKDALDAINRFQDKLEDISKKIKQRNENLEVPYIYLLPERIPNGTAI 696
```

🟨 Highly conserved 🟩 5-LOX specific 🟪 Fe-coordination sphere

Figure 7.11: MSA of selected LOX isoforms.

placed (Figure 7.13). In Stable-5-LOX this channel can accommodate the AA tail in those solutions with the carboxylate group bound to Lys409.

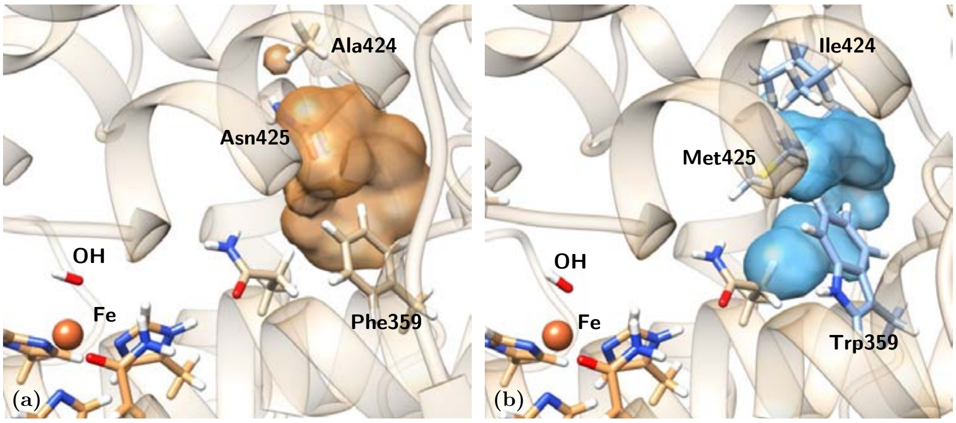


Figure 7.12: Representation of the volume of a) residues F359, A424 and N425 in Stable-5-LOX (orange); and b) residues W359, I424 and M425 in triple mutant (blue).

Table 7.6: Number of docking solutions interacting to indicated residues in Stable-5-LOX and mutants of holoenzyme receptor.

System	F177 (bb)	N180	W359	Q363	K409	R411	N425	R596
F359W	24	5	-	33	111	-	1	17
F359W-N425M	28	3	-	46	106	-	-	11
F359W-A424I-N425M	7	1	3	170	5	8	-	-
K409L	93	9	-	28	-	-	12	45
Q363L	32	13	-	-	102	-	5	24
Stable-5-LOX	27	-	-	34	116	-	2	13

When the channel is blocked, the AA tail is twisted inside the binding cavity, and this alteration of the cavity provokes that some solutions appear close to Arg411. This residue, which is spatially close to Lys409 (Figure 7.13 b), does not present any interaction with AA carboxylate in the docking solutions of the Stable-5-LOX. On the other hand, the number of docking solutions interacting with Gln363 increases (from 34 to 170, Table 7.6) as the cavity volume decreases by effect of the triple mutation.

Table 7.7: Number of solutions with C_7, C_{10} and C_{13} close to the OH^- group in Stable-5-LOX and mutants of holoenzyme receptor.

System	$d(C_7\text{-OH}) < 5$ Å	$d(C_{10}\text{-OH}) < 5$ Å	$d(C_{13}\text{-OH}) < 5$ Å
F359W	25	3	5
F359W-N425M	35	4	0
F359W-A424I-N425M	126	103	71
K409L	41	8	15
Q363L	39	12	8
Stable-5-LOX	29	4	4

As the number of mutated residues increases, solutions with $d(C_7\text{-OH}) < 5$ Å do so (see Table 7.7). Likewise, many solutions have more than

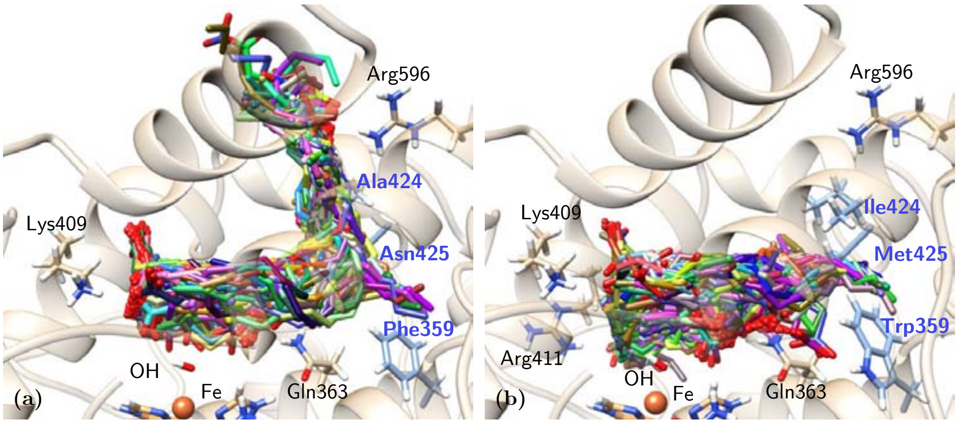


Figure 7.13: Docking solutions of (a) Stable-5-LOX holoenzyme; and (b) F359W-A424I-N425M-Stable-5-LOX mutant holoenzyme. The residues involved in the mutation (both in the wild-type and mutant images) have been highlighted in blue.

one methylene close to the OH^- group. This may be due to the new situation of AA inside the active site: because of the lower volume of the cavity, it can adopt more twisted poses surrounding the OH^- group, allowing several methylenes to approach this group. These results could support the "spatial hypothesis"[17,48] mentioned in the Introduction (with the tail-first orientation being responsible of 5-lipoxygenation), although the reason why a reduction of the 5-LOX active-site volume leads to 15-lipoxygenation is that the head-first orientation becomes the favored one in the mutant. As expected, the docking solutions interacting with Lys409 or Gln363 disappear in mutants K409L or Q363L, respectively, which also lead to more variabilily in the carbon atom suitably positioned for hydrogen abstraction.

7.3.2.4 Rigid apoenzyme

Following the same procedure for the apoenzyme structure, an initial approximation of a rigid receptor and, consequently, with identical geometry of the crystal structure was adopted in first place. For this approach, a total of 200 docking solutions were generated. In this docking run, solutions presented scores very close in energy and ranged from 32.29 to 23.92 Chemscore units, which correspond to solutions with relatively good predicted binding affinities, although the lowest energy solution is slightly worse than in the previous case of the holo form (ca. 32 .vs. 34 Chemscore units).

The solutions were clustered with an RMSD of 2.96 Å, giving 7 clusters (Figure 7.14). The RMSD value selected here to group the solutions into the same number of clusters (7 clusters) is lower than in the previous case, with rigid holoenzyme receptor (RMSD = 5.88 Å), which indicates that the dispersion of AA conformations is smaller here, even though the geometric variability is still large. This is consistent with the different cavity shapes between apoenzyme and holoenzyme. The scores of the representative models and the population of each cluster are shown in Table 7.8 (the models are ranked according to their relative score).

Table 7.8: Scoring function (in ChemScore units) of the representative models and cluster population (with RMSD = 2.96 Å) of the docking simulations of the rigid apoenzyme receptor

Cluster number	Score	Population
1	32.29	35
2	32.18	29
3	31.83	22
4	31.25	15
5	30.78	7
6	30.15	87
7	28.24	5

Next, the solutions were classified according to the presence of protein residues in H-bond distance to the AA carboxylate end, which were Ala424 (the backbone), Asn425, His600, Tyr181 and His372. The distances between the carbon C_7, C_{10} and C_{13} of AA and the oxygen atom of the OH^- group were measured for all the docking solutions. Only 3 solutions display $d(C_7\text{-OH}) < 5$ Å and 2 solutions $d(C_{10}\text{-OH}) < 5$ Å, but 39 solutions with $d(C_{13}\text{-OH}) < 5$ Å were found, with the AA carboxylate end close to His600. Based on the actual knowledge on lipoxygenase reaction, only the three solutions with the C_7 atom below the distance threshold of 5 Å would lead to the correct product of the reaction according to the 5-LOX regiospecificity. It is clear that such poses represent an extremely marginal set of docking orientations and it is reasonable to assume that in principle, the X-ray apoenzyme structure would not, or only residually, behave as a productive form of 5-LOX.

A more detailed analysis of the distance between the bisallylic methylene groups and the OH^- in the apoenzyme showed that in this group of solutions where AA carboxylate group is interacting with His600, C_{13} is closer to the OH^- group, whereas C_7 is far away. The presence of solutions interacting with His600 is interesting, since this residue has been proposed experimentally as the anchoring residue of the carboxylate end of the fatty acid substrate.[49] Although the molecular details by which His600 affects AA positioning are not known, it has been related with the polar/charged nature of the residue and a potential head-first binding of AA.

Interestingly, holoenzyme docking simulations did not yield any bind-

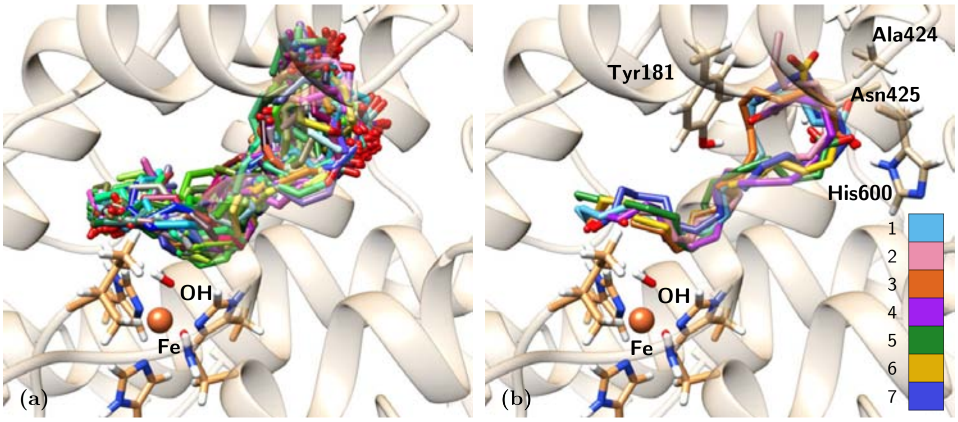


Figure 7.14: (a) Solutions of the docking simulation of the apoenzyme structure with rigid side chains; (b) Representative models of clustering with RMSD = 2.96 Å, showing residues of interest. The color scale represented in (b) assigns a different color to the representative element of each cluster.

ing mode with the AA carboxylate group interacting with His600. This is probably a consequence of the different orientations displayed by His600 in the apoenzyme (inward, pointing to the catalytic iron) and the holoenzyme (outward, pointing to the end of the cavity). Notice that this result is not a consequence of the homology modeling simulation, as His600 does not belong to the missing regions of structure 3V99 (see Table 7.2). The different orientation is already found among the different crystallographic structures (Figure 7.15).

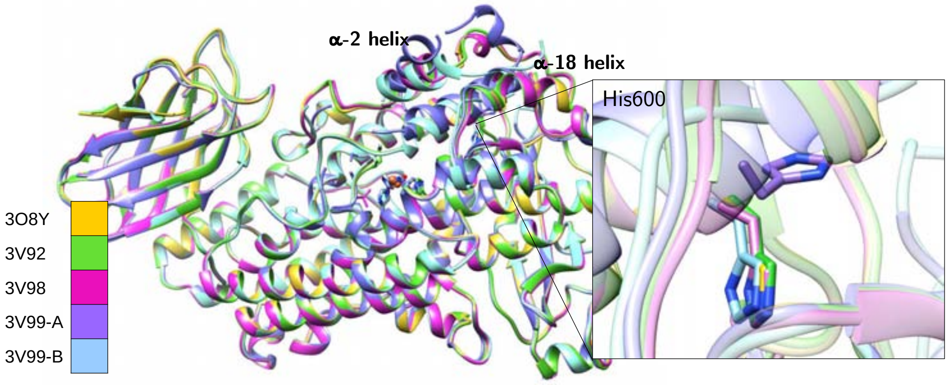


Figure 7.15: Overlay of 5-LOX crystallographic structures (coloring code on the left) showing a detail of His600 (box on the right). In structures 3O8Y, 3V92 and 3V98, only one subunit is depicted as in these cases the two monomers are almost identical. In structure 3V99 both subunits A and B are depicted as they exhibit differences. The orientation of His600 in structure 3V99-A (outward) clearly differs from the remaining structures (inward).

As structure 3V99 is a phosphorylation mimicking mutant Stable-

Ser663Asp, in order to discard that the difference in His600 orientation is a consequence of the phosphorylation mimicking mutation a comparison of the holoenzyme system with the “phosphorylated" apoenzyme structure (pdb entry 3V98) is also reported here. It can be observed that the orientation of His600 in 3V98 is consistent with the Stable-5-LOX (pdb entry 3O8Y), as well as the “non-phosphorylated" mutant Ser663Ala (pdb entry 3V92), keeping the difference with subunit 3V99-A. Thus, we are induced to consider that the conformational changes in the holoenzyme structure and, particularly, in His600 orientation, are a consequence of the substrate entry rather than the phosphorylation mimicking mutation.

7.3.2.5 Flexible side-chain apoenzyme

An improved exploration of the conformational space of the AA:5-LOX complexes has been performed by the use of several rotameric schemes. Rotamers allow the model to overcome possible low energy local arrangements of the crystal structure to gain activity and predict binding modes of the substrate that otherwise could not be explored. Thus, these rotameric approaches could allow determining if the apoenzyme structure could be transformed into a productive structure of 5-LOX by rotation of key residues. Focus has been given to Asp176, Phe177, Tyr181, and Lys409, as these are residues of the binding site of the enzyme that have different orientations in the apoenzyme and holoenzyme as mentioned in the previous Section (see Figure 7.4). A run of 200 docking solutions and using the GOLD rotamer library for all these four residues was performed in a first place. The scores of the representative models and the population of each cluster are shown in Table 7.9.

Table 7.9: Scoring function (in ChemScore units) of the representative models and cluster population (with RMSD = 3.62 Å) of the docking simulations of the apoenzyme receptor with flexible side-chains (F177, Y181, E176, K409) provided by the GOLD rotamer library.

Cluster number	Score	Population
1	33.53	39
2	30.56	60
3	30.38	48
4	29.45	21
5	27.32	28
6	22.86	1
7	21.94	3

The former Table refers to a simulation allowing the flexibility of the four residues F177, Y181, E176 and K409, simultaneously, but other combination of rotamers of these residues have been also tested. However, these docking solutions still resemble those obtained in the previous rigid receptor scheme, indicating that conformational changes to achieve the productive structure are beyond the change of these residues side-chain orientations. Table 7.10 shows the different combinations of rotamers performed to release the binding site constraints.

Table 7.10: Number of solutions with C_7, C_{10} and C_{13} close to the OH^- group, for the different docking calculations with rotamers in apoenzyme. Calculation 1 corresponds to the initial rigid approximation, 2 to rotameric flexibility afforded to residues F177 and K409 and 3 to residues F177, Y181, E176 and K409.

Experiment	$d(C_7\text{-OH}) < 5$ Å	$d(C_{10}\text{-OH}) < 5$ Å	$d(C_{13}\text{-OH}) < 5$ Å
1	3	2	39
2	9	3	73
3	17	-	26

Overall, the increase of the number of flexible side chains in the docking leads to a slight increase of the number of solutions with $d(C_7\text{-OH}) < 5$ Å (up to 17, see Table 7.10). However, this is apparently a consequence of very subtle relaxations of the active site cavity upon binding since the final position of the residues that have been allowed with flexibility remains nearly unchanged with respect to the X-ray position (Figure 7.16). The highest number of catalytically consistent solutions are still those with $d(C_{13}\text{-OH}) < 5$ Å, reinforcing the idea that the product that the apoenzyme would form is not the main product of human 5-LOX (just assuming distance criterion, and just considering the docking simulations results). It seems that rotamers cannot yield to the productive structure of 5-LOX. The positioning that cork residues exhibit in the holoenzyme model is never achieved in the apoenzyme by side chain rotamers (see Figure 7.16, thus, the enzyme needs to experience further structural changes to reach activity.

Up to this point it seems clear that starting from the apoenyme receptor, it is not possible to obtain a substrate:protein complex bearing a 5-LOX activity, in terms of the distances from the bisallylic methylenes to the catalytic iron. Most solutions show an interaction of the AA carboxylate end with His600. Solution 11 has been selected from the docking calculations into the rigid apoenzyme receptor (Figure 7.17); this solution has a similar score to the most stable solution (30.78 ChemScore units in solution 11 and 32.29 in solution 1), but its C_{13}-OH, C_{10}-OH and C_7-OH distances are smaller than the corresponding ones in the most stable solution; actually, it corresponds to the most stable solution with $d(C_{13}\text{-OH}) < 5$ Å, and it presents a $d(C_7\text{-OH})$ distance of 7.39 Å. Still, as in the AA dockings to the holoenzyme structure no solutions show this interaction, we have selected one of these AA:apoenzyme solutions to study this AA-His600 interaction.

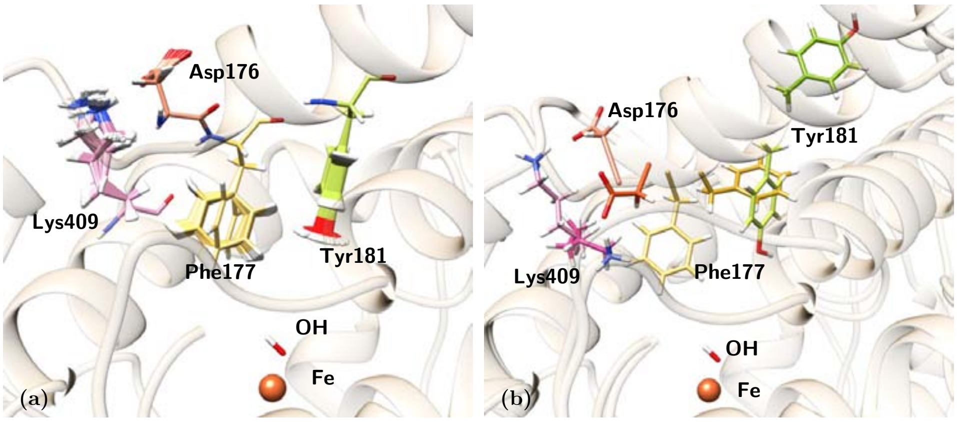


Figure 7.16: (a) Position of the side chains of Asp176 (orange), Phe177 (yellow), Tyr181 (green) and Lys409 (pink) generated by the GOLD rotamers library during the docking simulations of apoenzyme with flexible side chains; (b) superposition of apo and holo forms, showing the former residues (the same colors as in (a)); holo residues are drawn in darker colors.

7.3.3 Reactivity study

To determine if any of the binding modes obtained by the docking simulations is consistent with the reactivity of the 5-LOX system, a study of the reactivity combining MD simulations plus QM/MM calculations has been carried out.

7.3.3.1 Molecular Dynamics Simulations

A first MD study aimed at the study of the reactivity and thus, at the generation of the pre-catalytic complexes from the different binding modes obtained in the docking simulations presented above has been carried out. Solutions 12 and 34 of the flexible side-chains holoenzyme docking simulation, which are representative of head-first (with the substrate carboxylate group interacting with Gln363) and tail-first orientations (with Lys409), respectively, and solution 11 from the rigid apoenzyme docking simulation, which is also a representative member of a head-first orientation and presents the AA carboxylate end interacting with His600, were selected to perform this study. The goal of this preliminary MD simulation study is to determine if any of these substrate:protein complexes and their corresponding binding modes is consistent with a 5-LOX reaction specificity (i.e. preference of abstraction of the pro-S hydrogen from C_7).

As has been stated in the previous section, a first indicator of the consistency of the substrate:protein complexes with the regiospecificity of 5-LOX is a measure of the distances from the bisalyllic methylenes (both C and H distances) to the oxygen of the Fe(III)-OH$^-$ cofactor. However, as

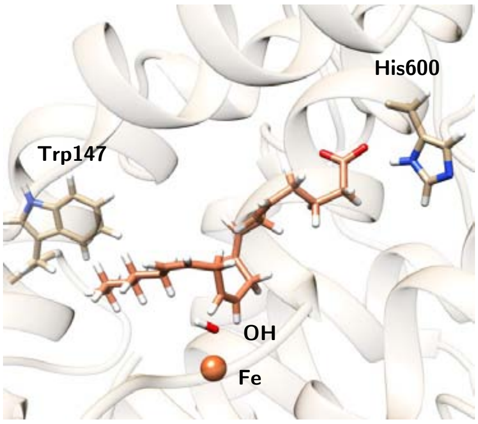


Figure 7.17: Representation of solution 11 of the docking simulations using the rigid apoenzyme receptor.

has been demonstrated in the previous chapters, the distance criteria alone may not be enough to extract conclusions about the reactivity and thus a calculation of the energy barriers for hydrogen transfer is required. Basically, although the initial distance of the atom that is going to be abstracted has a crucial influence in the potential energy barrier, it is not the unique factor and other considerations have to be taken into account, for example, how difficult is the evolution from a non-planar structure in the reactant to a planar structure in the abstraction product. Nevertheless, when one hydrogen is significantly further away than another one, it can be expected that the latter will be preferably abstracted. Therefore, the distance distribution of the three bisallylic methylene groups to the oxygen of the Fe(III)-OH$^-$ cofactor have been monitored during the MD simulations, and finally a set of snapshots has been selected to calculate the hydrogen abstraction paths and potential energy barriers. Once each set was generated, the feasibility of the different hydrogen abstractions for each one of the three possible AA binding modes was studied by QM/MM calculations.

MD simulations of crystal structure Along with the binding modes generated in this work, the pose that AA exhibits in crystal structure 3V99-A was also studied by MD simulations. Since in this X-ray structure some protein coordinates are incomplete, the coordinates of the holoenzyme model have been used instead. Then, crystal structure 3V99-A and the holoenzyme model were superimposed, and AA coordinates were placed into the holoenzyme model to obtain a complete structure of the Michaelis complex.

In the crystal structure pose, initially the AA molecule does not seem to be interacting to any positive residue of the protein, and its carboxylate end is placed at the bottom of the active site cavity (head-first orientation). During the MD simulation the carboxylate end moves towards the region of Arg596 and Asn425, and retains this position during the 10 ns production run trajectory that was carried out for this complex. . The average distance (standard deviation) from C_7 to the oxygen of the OH^- group is of 13.44 (1.10) Å, whereas $d(C_{10}\text{-OH})$ and $d(C_{13}\text{-OH})$ are of 10.92 (1.29) Å and 8.78 (1.33) Å, respectively. Thus, there are not structures with any bisallylic methylene close enough to the Fe(III)-OH^- cofactor to consider that model as a 5-lipoxygenating catalytically productive structure. Therefore it is clear that starting from this complex it will be difficult to obtain structures with any bisallylic methylene close enough to the Fe(III)-OH^- cofactor to consider this model a catalytically productive one. Therefore longer MD simulations have not been performed for this complex, as at short production steps the substrate already exhibits a behavior that is non-compatible with the expected reaction specificity. In the same way, the QM/MM calculations have been carried out on the other three complexes and the results are presented below.

Tail first AA:holo complex (solution 34) The first complex studied corresponds to a tail-first binding mode with the AA carboxylate interacting with Lys409 (Figure 7.9, green). During the 20 ns of simulation of solution 34, different configurations of the AA:Stable-5-LOX Michaelis complex were explored. The histograms of the distances from C_7, C_{10} and C_{13} to the oxygen of the OH^- group are presented in Figure 7.18.

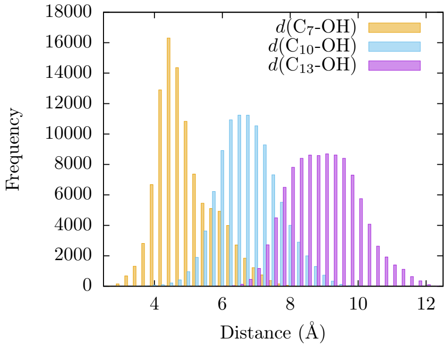


Figure 7.18: Histogram of distances d(C_7-OH) (yellow), d(C_{10}-OH) (blue) and d(C_{13}-OH) (purple), during the 20 ns MD production step of holo solution 34 (tail first).

Figure 7.18 shows that $d(C_7\text{-OH})$ presents a narrow distribution centered at 4.5 Å, whereas the distribution of distances $d(C_{10}\text{-OH})$ and $d(C_{13}\text{-OH})$ are broader and centered at longer distances around 7.0 Å and 9.0 Å, respectively. These results suggest that a tail-first binding mode with the AA carboxylate interacting with Lys409 gives conformations in which the C_7 methylene is the only one close to the hydroxyl group.

Head-first AA:holo complex (solution 12) The second binding mode studied corresponds to a head-first orientation with the AA carboxylate initially interacting with Gln363 (Figure 7.9, gray). The MD simulations for this orientation resulted in the histograms of the distances from C_7, C_{10} and C_{13} to the oxygen of the OH^- group along presented in Figure 7.19 . The position of the AA carboxylate end remained close to Gln363, indicating that this orientation is in principle also possible.

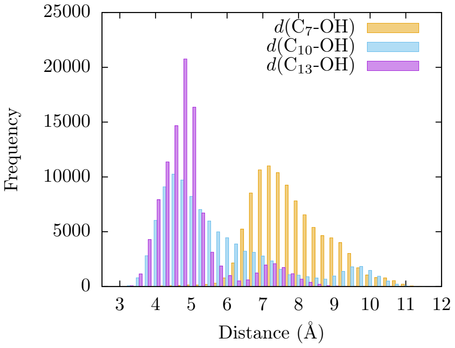


Figure 7.19: Histogram of distances $d(C_7\text{-OH})$ (yellow), $d(C_{10}\text{-OH})$ (blue) and $d(C_{13}\text{-OH})$ (purple), during the 20 ns MD production step of holo solution 12 (head first).

MD simulations of this solution 12 (head first) give broader distributions of all distances than those obtained for the tail-first orientation. $d(C_{10}\text{-OH})$ and $d(C_{13}\text{-OH})$ present lower values than $d(C_7\text{-OH})$, which has its maximum population at $\sim$ 7.5 Å. AA remains with the carboxylate end bound to Gln363 (and also Gln557, which is spatially close to Gln363), but its backbone adopts twisted conformations in which C_{10} and C_{13} can be simultaneously close to the hydroxyl group, while C_7 is clearly further away from the OH^- group. Thus, these MD simulations suggest that this head-first model does not provide structures consistent with the 5-lipoxygenating function of this enzyme with the AA substrate. Still, the hydrogen abstraction step was also studied in snapshots drawn from this simulation.

Head-first AA:apo complex (solution 11) Up to this point, the results can be summarized in that the proposed tail-first orientation is consistent with the experimental 5-LOX regio- and stereoselectivity, whereas the head-first orientation adopted in the holo structure would give a 15/11-LOX or 12/8-LOX regioselectivity, according to the distance criterion of the bisallylic methylenes. However, and as has been already mentioned in the Introduction and the docking simulation Section, recent mutagenesis experiments suggest a role of His600 in substrate binding. In these experiments, it has been shown that the H600A mutant displays less than 10% of Stable-5-LOX activity and a affinity for the substrate. Moreover, no detectable catalytic activity was observed for the H600V mutant. These results lead the authors to suggest that His600 is essential for proper AA positioning in the 5-LOX active site to produce the 5-HPETE product, and also that the presence of this amino acid deep in the cavity may be favoring the head-first substrate orientation.

The hypothesis of a head-first orientation in human 5-LOX is not new, and had been proposed in the literature in previous studies arguing that the stereochemistry of the final hydroperoxidation product, 5*S*-HPETE, could only be jutified with a head-first orientation in contraposition with 15*S*-HPETE, the 15LOX product, which is obtained from an AA tail-first orientation. (see Introduction chapter for details). A head-first orientation has also been proposed for murine 8*S*-LOX. However, it is important to note that this interpretation of the experimental observations is not necessary right, as it does not consider all the conformational changes that AA can undergo once placed inside the active site (methylene carbons consist on sp^3 carbons that can rotate alternating the side of the plane that faces the Fe(III)-OH$^-$ group), which in principle could yield both *R*- and *S*-HPETE products, even if starting from a given AA orientation.

Nevertheless, as explained in section 3.3, in order to study a possible head-first orientation bearing a AA-His600 interaction, the apoenzyme X-ray structure has been used to generate the coordinates of complexes presenting such contact. The docking pose solution 11 (head first) was selected and prepared for an MD simulation following the same procedure as for the holoenzyme complexes. During the simulation time studied the position of the substrate remains stable. The histogram of the distances from C_7, C_{10} and C_{13} to the acceptor OH$^-$ is shown in Figure 7.20.

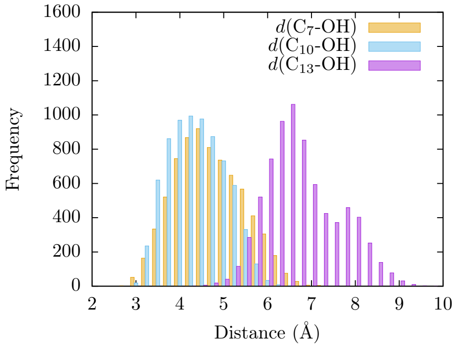


Figure 7.20: Histogram of distances d(C_7-OH) (yellow), d(C_{10}-OH) (blue) and d(C_{13}-OH) (purple), during the 20 ns MD production step of apo solution 11 (head first).

A similar distance distribution can be observed for all the distances, being $d(C_{10}\text{-OH})$ the smallest of them, which is centered around 4 Å, followed very closely by $d(C_7\text{-OH})$ which is very similar and centered around 4.5 Å; $d(C_{13}\text{-OH})$ is quite greater with a distribution centered at 6.5 Å. In this situation the behavior differs from the case of the holoenzyme in which for a tail-first orientation (solution 34) the distance distribution is clearly more discriminating, with the $d(\mathrm{C}_7\text{-OH})$ distance being the shortest one (Figure 7.18), whereas for the head-first orientation (solution 12), the distances distribution clearly differentiates $d(\mathrm{C}_7\text{-OH})$ (the greatest value) from $d(\mathrm{C}_{10}\text{-OH})$ and $d(\mathrm{C}_{13}\text{-OH})$, which are similar (Figure 7.19). According to a distance criterion, in solution 11 of the apoenzyme both C_7 and C_{10} could be suitable for transferring a hydrogen atom to the OH^- group. It is also worth mentioning that along the MD simulation of solution 11, the aromatic side chains of the two FY-cork residues (Tyr181 and Phe177) remain as in the apoenzyme crystal structure. That is, once in the cavity AA carboxylate end can interact with His600 without the need of ejecting or repositioning the FY-cork.[49]

7.3.3.2 QM/MM potential energy profiles for the hydrogen abstraction reactions in the AA:5-LOX complex

Reactivity of the holoenzyme complexes To study the reaction mechanism of the hydrogen abstraction step the following procedure has been employed. First, from the previous MD simulations of the holoenzyme solution 34 (with tail-first orientation), and solution 12 (with head-first orientation), four snapshots were selected in order to obtain the initial reactant structures that initiate the reactive process. These snapshots were representative of each MD simulation, with a set of distances $d(\mathrm{C}_7\text{-OH})$, $d(\mathrm{C}_{10}\text{-OH})$ and $d(\mathrm{C}_{13}\text{-OH})$ very close to the center of the distances distributions of every case (see Table 7.11 and Figures 7.18 and 7.19).

As the complete system (substrate:protein complex + water solvation box) derived from the MD simulations comprises a huge number of atoms (150 000 atoms), the selected snapshots were modified to reduce the size of the system. In this way, water molecules outside a 15 Å radius volume centered on the AA molecule were removed, giving a total system size of 12 000 atoms. All residues and water molecules within a 15 Å sphere centered on AA C_{11}, including the complete AA molecule, were set as the active region of the optimization processes, whereas the remaining region of the system was kept frozen. Next, a full QM/MM optimization was performed to obtain the pre-reactive minima.

Table 7.11 shows the main geometric parameters of the different pre-reactive minima obtained from the optimization of the four selected snapshots. Snapshots I (Figure 7.21a) and II correspond to a tail-first orientation, whereas snapshots III (Figure 7.21b) and IV present a head-first orientation. In each case the closest hydrogen atom of the corresponding methylene group has been the abstracted atom (showed in Table 7.11)[1] . Additionally, in snapshots I and III not only the most suitable hydrogen but also some of the remaining hydrogens have been abstracted for comparative purposes. Note that in snapshots I and II, H_7 pro-*S* (which should be the more reactive for 5-LOX specificity) is not only close enough to be transferred to the OH^- group, but in a good orientation as $d(C_7\text{-OH})$ is clearly greater than $d(H_7\text{-OH})$.

Table 7.11: Main geometric parameters of the pre-reactive minima obtained from the optimization of the four selected snapshots. Subscript X stands for 7, 10 or 13. Distances are given in Å.

Snapshot	H_X	Orientation	$d(C_X\text{-OH})$	$d(H_{X,A}\text{-OH})$	$d(H_{X,B}\text{-OH})$
I	H_7 pro-*S*	Tail First	3.41	4.32	2.60
I	H_{10B}	Tail First	5.42	5.89	5.08
I	H_{13B}	Tail First	6.72	7.05	5.67
II	H_7 pro-*S*	Tail First	3.49	4.32	2.55
III	H_{7B}	Head First	5.31	5.15	4.85
III	H_{13A}	Head First	5.26	4.16	4.18
IV	H_{10A}	Head First	5.01	4.74	5.90

Table 7.12: Initial $d(\text{H-OH})$ and potential energy barriers corresponding to the indicated hydrogen abstractions for the different pre-reactive structures obtained from optimization of the four selected snapshots.

Snapshot	H_X	Orientation	Anchoring residue	$d(H_X\text{-OH})$ (Å)	$\Delta E^\ddagger$ (kcal/mol)
I	H_7 proS	Tail First	Lys409	2.60	18.1
I	H_{10B}	Tail First	Lys409	5.08	25.5
I	H_{13B}	Tail First	Lys409	5.67	50.7
II	H_7 proS	Tail First	Lys409	2.55	18.1
III	H_{7B}	Head First	Gln363	4.85	36.2
III	H_{13A}	Head First	Gln363	4.16	19.8
IV	H_{10A}	Head First	Gln363	4.74	25.0

The optimized geometry corresponding to each pre-reactive structure has been used as a starting point to build the potential energy profile along the reaction coordinate for the hydrogen abstraction process. Note that for snapshot I abstraction of H_{10B} and H_{13B} were also considered in case they could still react with reasonable barrier heights even though they are significantly further away. The reaction coordinate, z, is defined as the difference between the distance of the breaking bond (C_7-H_7, C_{10}-H_{10} or C_{13}âĂŞH_{13}

---

[1] Only the stereochemistry of the hydrogen preferentially abstracted, H_7 pro-*S* is given, as the stereochemistry of the minor products is not known. In this latter case, the two hydrogens of a methylene group will be referred as H_A and H_B.

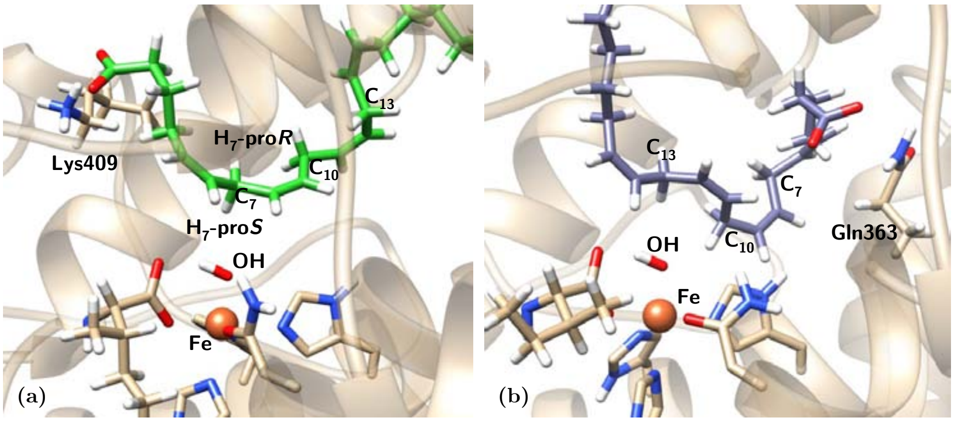


Figure 7.21: Representation of two characteristic reactant minima showing AA and the Fe coordination sphere: a) structure obtained from optimization of snapshot I (tail-first); b) structure obtained from optimization of snapshot III (head-first).

for H_7, H_{10} and H_{13} abstractions, respectively) and the forming bond (H_7-O, H_{10}-O and H_{13}-O, respectively). To construct the potential energy profile, a series of optimizations of the mobile part of the system have been performed in the presence of harmonic restrictions on the reaction coordinate as $V = \frac{k}{2}(z - z^0)^2$, where k is the force constant, which has been set equal to 3.0 hartree/Bohr^2, and z^0 is the reference value for the reaction coordinate z at each energy minimization calculation, which increases with a step size of 0.1 Å.

The highest energy point of each profile was selected as starting point for a direct transition state search, using the microiterative searching procedure as mentioned before. A small core containing 6 atoms (C, pro-S and pro-R hydrogens, and Fe(III)-OH$^-$ cofactor) was selected for the P-RFO algorithm. In every case a transition state structure could be located. The potential energy barriers for all the processes are listed in Table 7.12. Values of additional geometric parameters corresponding to the pre-reactive minima and the transition state structures are given in Table 7.13.

As can be observed in Table 7.12 the wide range of potential energy barrier values, which range from 18.06 to 50.73 kcal/mol, are highlighted. This situation is directly related with the wide dispersion of geometries in terms of the initial distance between the hydrogen that will be transferred and the accepting oxygen during the reactive process, with values ranging from 2.55 to 5.67 Å(see Table 7.12). Distances values are directly determined by how the AA is placed in the active site of the enzyme. Importantly, the calculated potential energy barriers for H_7 abstraction (18.06 kcal/mol) are in qualitative agreement with experimental data on lipoxygenases. Contrary to this, the H_{10} and H_{13} abstraction barriers (7.4 and 32.6 kcal/mol higher than

Table 7.13: Main geometric parameters of the pre-reactive minima and TS's of holoenzyme structures. Subscript X stands for 7, 10 or 13. Distance are given in Å and angles in degrees.

Snapshot	H_X	$d(C_X\text{-}O)$		$d(C_X\text{-}H_X)$		$d(H_X\text{-}O)$		$\angle(C_X\text{-}O\text{-}H)$		$\angle(H_X\text{-}O\text{-}H)$	
		R	TS	R	TS	R	TS	R	TS	R	TS
I	H_{7B}	3.41	2.65	1.10	1.29	2.60	1.37	29.2	59.6	70.3	89.5
I	H_{10B}	5.42	2.62	1.10	1.29	5.08	1.34	60.0	68.0	88.9	81.1
I	H_{13B}	6.72	2.66	1.11	1.33	5.67	1.33	68.4	61.7	100.9	100.3
II	H_{7B}	3.49	2.65	1.10	1.30	2.55	1.35	36.0	59.3	77.4	93.3
III	H_{7B}	5.31	2.66	1.10	1.34	4.85	1.34	92.2	71.0	131.7	140.4
III	H_{13A}	4.18	2.61	1.10	1.28	4.16	1.34	83.6	171.2	90.7	97.4
IV	H_{10A}	5.01	2.63	1.10	1.28	4.74	1.35	98.0	176.3	113.0	107.7

that for H_7 abstraction, respectively), indicate that these are not plausible processes.

Thus, taking all these results together, it is shown that this tail-first binding mode, in which AA is interacting with Lys409, is compatible with 5-LOX regiospecificity: not only C_7 is the closest methylene group, but also it is close enough to the OH^- group, while C_{10} and C_{13} are considerably further away. The calculated energy barriers confirm that these pre-reactive minima would be favorable for H_7 abstraction, and unfavorable for H_{10} and H_{13} abstraction processes. In addition, the abstracted H_7 turns out to be the pro-*S* one (see Figure 7.21a), in good agreement with the experimental stereoselectivity.

On the other hand, with a head-first type orientation, MD simulations have not provided with structures with H_7 in a good arrangement to initiate the reactive process, whereas H_{13} (and in a lesser extent H_{10}), were close enough for the reaction to take place. Again, the potential energy barriers support this fact: with a head-first orientation, the most favorable process is the H_{13} abstraction, with a potential energy barrier of 19.84 kcal/mol, whereas H_{10} abstraction is 5.2 kcal/mol higher and, most importantly, H_7 abstraction is 16.36 kcal/mol higher.

At this point, it has to be underlined that, according to Figure 7.19, the probability of finding a head-first orientation snapshot with a C_7-OH distance as short as in the pre-reactive structure optimized from snapshot III, turns out to be extremely low. Even in this way the barrier for the H_7 abstraction would be over 36 kcal/mol. Thus, it can be concluded that H_7 abstraction does not occur when AA is in the head-first orientation. Therefore, the proposed tail-first orientation complex is consistent with the experimental 5-LOX regio- and stereoselectivity, whereas a head-first orientation will yield a 15/11-LOX or 12/8-LOX regioselectivity, in terms of hydrogen abstraction.

Reactivity of the apoenzyme complex As seen in the MD section, in the case of the apoenzyme:substrate complex, distance distributions of C_7 and C_{10} were very similar, thus indicating that in principle this system would not discriminate between hydrogen abstractions from C_7 and C_{10}, from a distance criterion point of view. The subsequent QM/MM study is aimed to determine if C_7, C_{10} or even both positions, are reactive for the hydrogen abstraction process, and then, which regioespecificity is consistent within this substrate binding mode.

Then, a series of snapshots from the MD simulation of solution 11 in the apoenzyme, with a set of distances consistent with the distribution reported in the previous section (Figure 7.20, have been selected as the starting structures for the study of the first step of the reaction mechanism. The four selected snapshots were optimized giving the corresponding pre-reactive minima, whose characteristics are shown in Table 7.14. Values of additional geometric parameters corresponding to the pre-reactive minima and the transition state structures are given in Table 7.15. In each case the closest hydrogen atom of the corresponding methylene group has been the abstracted atom.

Table 7.14: Main geometric parameters of the pre-reactive minima obtained from the optimization of the four selected snapshots. Subscript X stands for 7 or 10.

Snapshots	H-abstracted	$d(C_X\text{-OH})$ (Å)	$d(H_X\text{,A-OH})$ (Å)	$d(H_X\text{,B-OH})$ (Å)
V	H_7 proS	3.52	3.96	2.49
VI	H_{10B}	3.5	4.15	2.44
VII	H_7 proS	4.19	4.41	3.13
VIII	H_{10A}	3.63	2.68	4.13

Table 7.15: Main geometric parameters of the pre-reactive minima and TS's of apoenzyme structures. Subscript X stands for 7, 10 or 13. Distance are given in Å and angles in degrees.

Snapshot	H_X	$d(C_X\text{-O})$		$d(C_X\text{-}H_X)$		$d(H_X\text{-O})$		$\angle(C_X\text{-O-H})$		$\angle(H_X\text{-O-H})$	
		R	TS	R	TS	R	TS	R	TS	R	TS
V	H_7 proS	3.52	2.66	1.09	1.30	2.49	1.36	58.8	55.5	104.2	89.6
VI	H_{10}B	3.50	2.62	1.10	1.31	2.44	1.32	47.2	63.9	85.6	90.8
VII	H_7 proS	4.19	2.63	1.10	1.33	3.13	1.32	71.4	67.5	102.6	81.4
VII	H_{10}A	3.63	2.62	1.09	1.31	2.68	1.31	143.8	173.3	118.8	105.5

The structure of the reactant minimum of snapshot VI is represented in Figure 7.22. This snapshot presents H_{10} close enough and in good orientation to the Fe(III)-OH^- cofactor.

In the apoenzyme structure the potential energy barriers range from 17.35 kcal/mol to 36.00 kcal/mol (Table 7.16), but the set of initial distances

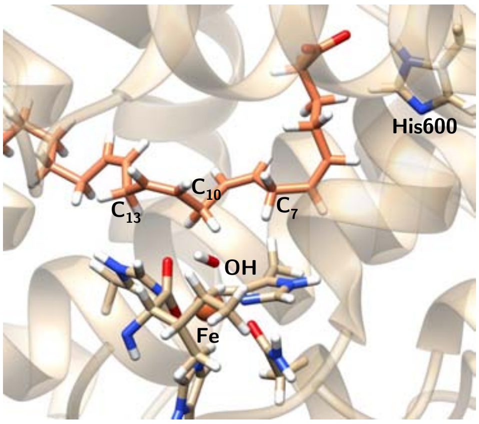


Figure 7.22: Representation of a characteristic reactant minimum of the apo complex obtained from optimization of snapshot VI showing the AA substrate and the Fe coordination sphere.

Table 7.16: Initial *d*(H-OH) and potential energy barriers corresponding to the indicated hydrogen abstractions in the different pre-reactive structures obtained from optimization of the four selected snapshots.

Structure	H-abstracted	Orientation	Anchoring residue	*d*(Habs.-OH) (Å)	ÎŤEâĂą(kcal/mol)
V	H_7 proS	Head First	His600	2.49	18.2
VI	H_{10B}	Head First	His600	2.44	17.3
VII	H_7 proS	Head First	His600	3.13	23.4
VIII	H_{10A}	Head First	His600	2.68	36.0

exhibits a narrower distribution than in the case of the holoenzyme structure, ranging from 2.44 to 3.13 Å. Thus, a wide dispersion of energy barriers comes from a narrow dispersion of initial geometries in terms of the set of atoms that constitute the reactive part of the system (the pentadienyl system defined by the bisallylic methylene and the Fe(III)-OH^- cofactor). Within these structures there are some which favor a H_{10} abstraction process, but there are also structures with an initial configuration favorable for a pro-*S* H_7 abstraction. This binding mode of AA in principle will be consistent with 5-LOX specificity, but also with 8/12-LOX specificity.

In fact, 5-LOX is known to attack both carbon atoms, but for slightly different substrates: C_7 is attacked in the first step to generate 5-HPETE and, subsequently, the hydrogen on C_{10} is abstracted from the first-oxidation product in the generation of leukotriene LTA_4. As the substrates are different, though, it is not clear whether their binding orientation should be the same or not. Still, a model which puts both carbons proximal to the catalytic machinery, given the geometry of the bisallylic methylens, might indeed be plausible. Actually, in the previous studies on 15-LOX-1 (see previous Chapters), it has been shown that AA can adopt conformations in the active site of the enzyme in which both C_{10} and C_{13} are at reactive distances of the OH^- group. In that case, abstraction of H_{13} is favored by $\sim$ 4 kcal/mol (in agreement with the experimental ratio of 97:3 for 15/12 products in 15-LOX) probably due to steric hindrance for H_{10} abstraction, which is not necessarily the case here.

Taken all the results together, to this point it can be established that both apo and holo structures can accommodate the AA substrate into the active site while having C_7 at reactive distances and with considerably low potential energy barriers for the hydrogen abstraction process of pro-*S* H_7. Particularly, for the holoenzyme structure the binding mode compatible with these characteristics presents the AA in a tail-first orientation with the AA carboxylate end interacting with Lys409, and the AA methyl tail occupying a region that is consistent with the mutagenesis experiments of Kühn and coworkers; the corresponding potential energy barriers show that only H_7 process is feasible, being H_{10} and H_{13} unfavorable, which agrees with 5-LOX regio- and stereospecificity. On the other hand, for the apoenzyme structure a head-first orientation binding mode has been identified with the AA carboxylate end interacting with His600 and the methyl tail pointing towards the region of Trp147; in this case the potential energy barriers indicate that both H_7 and H_{10} abstractions are possible, the latter one being scarcely observed experimentally.

The experimental observation that the only holoenzyme X-ray structure available with the AA substrate differs from the apoenzyme ones strongly suggests that the protein suffers conformational changes upon substrate binding, and that the final pre-catalytic complex will be different from the apoenzyme structure. Unfortunately, the incomplete X-ray coordinates of the holoenzyme structure do not permit to unambiguously determine the AA orientation. Furthermore, all the methylene groups of the AA structure are too far to be considered as a catalytically productive complex. Thus, this holoenzyme crystal structure may not be the catalytically active one, and thus the homology model generated here.

7.3.4 Study of the stability of the complexes: long MD simulations

Up to this point, a reactivity study of the generated binding modes has been provided. However, as explained before, even when it has been possible to find AA:protein complexes compatible with 5-LOX catalysis, it has not been possible to discriminate if any of these structures is the final productive structure of 5-LOX. Taking all the considerations explained here, a final study of the obtained AA:Stable-5-LOX complexes has been carried out by

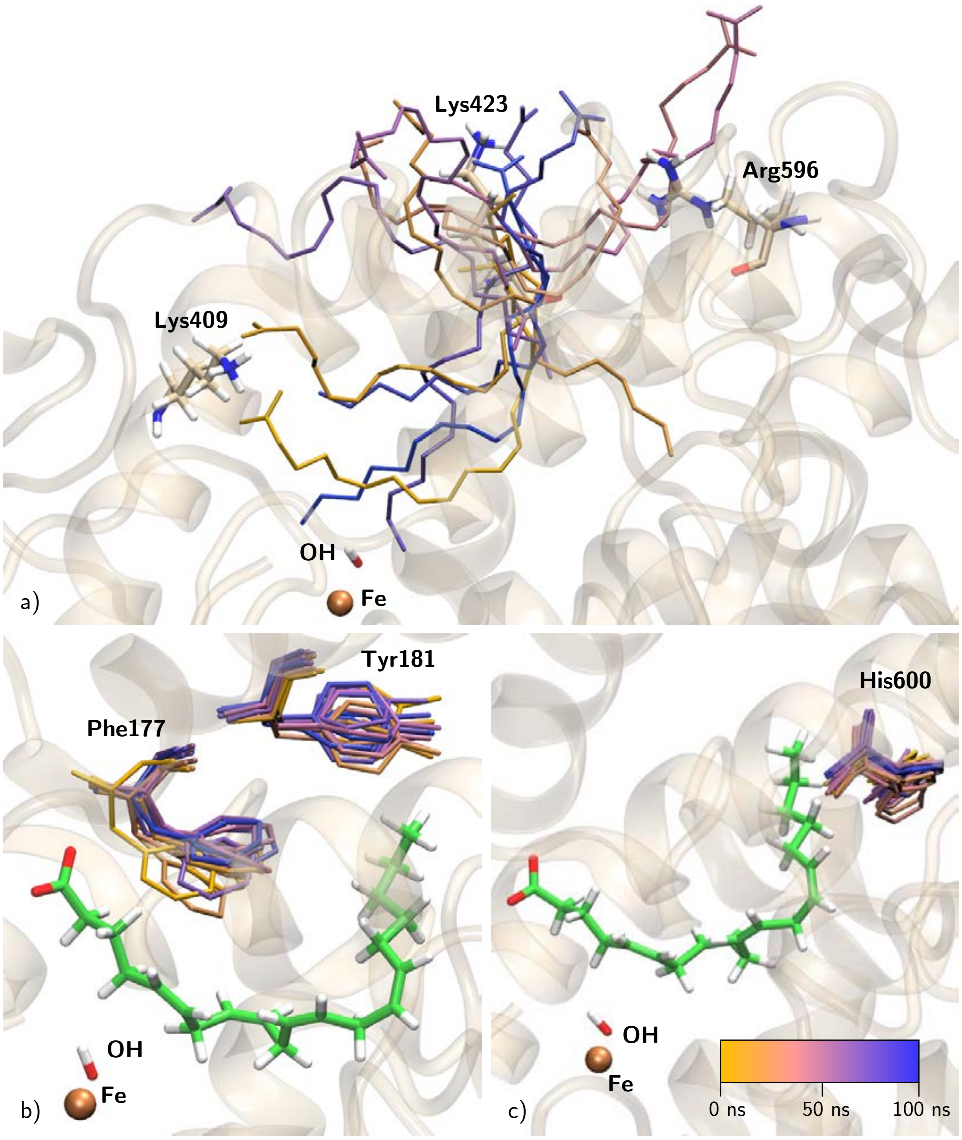
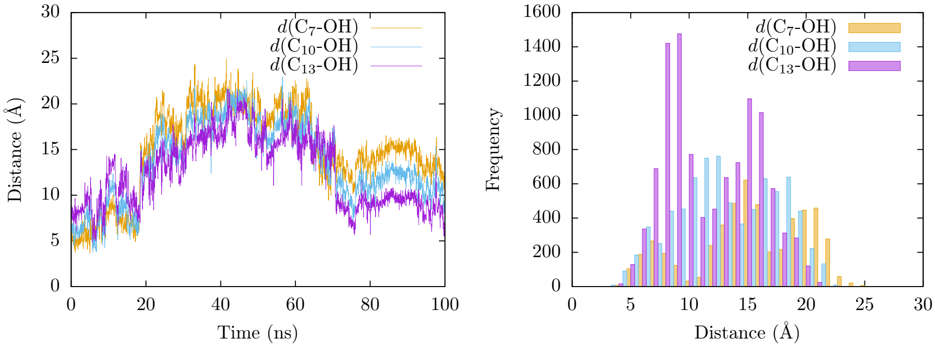


Figure 7.23: Representation of the evolution of the system during the 100ns production step. a) Evolution of AA. The enzyme is only shown in its initial position, as a reference. Initially, the AA carboxylate end is interacting with Lys409 (yellow structures). As the simulation progresses, this interaction is lost, and the substrate leaves the active site but it remains at the protein surface. Later on, the AA carboxylate end interacts with Lys423 and Arg596, which are placed at the other end of the active site cavity. Finally, the AA tail enters the cavity through the opposite end, giving a reverse orientation to the starting one (structures in blue); b) evolution of the FY-cork (Phe177-Tyr181), and c) evolution of His600, superimposed with one snapshot of AA (green).

means of long MD simulations in order to get new insights of their structural behavior along the trajectory. Obviously, the way AA is positioned in the active site will affect the geometric changes that the enzyme experiences and vice versa (substrate-induced conformational changes). This MD study is aimed to estimate the long-term stability and conformational changes that has not been observed in the previous ones. Therefore, MD simulations of 100 ns production runs were carried out for each of the complexes studied in the previous sections: solution 34 (holo, tail first), solution 12 (holo, head first) and solution 11(apo, head first).

7.3.4.1 Tail first AA:holo complex (solution 34)

For solution 34 (tail first) of docking to holoenzyme, the complex remains relatively stable during the first 18 ns after which AA abandons the active site. During this 18 ns, the AA carboxylate interacts with Lys409 and the distance order (with respect to the Fe(III)-OH$^-$ cofactor) $C_7 < C_{10} < C_{13}$ is maintained (see Figure 7.24). After this point, the substrate abandons the cavity, but it remains at the protein surface and, later on, the AA carboxylate interacts with Lys423 and Arg596, which are placed at the other end of the active site cavity. The evolution of the substrate along the 100 ns trajectory is depicted in Figure 7.23a.



(a) Evolution of distances. **(b)** Histogram of distances.

Figure 7.24: Evolution of the distances $d(C_7\text{-OH})$, $d(C_{10}\text{-OH})$ and $d(C_{13}\text{-OH})$ along the 100 ns of production step of the holo tail-first complex.

Figure 7.25 shows the distances of the AA carboxylate group to some polar/positive protein residues during the 100 ns trajectory. As commented, AA interacts with Lys409 during the first 18 ns and then, after abandoning the active site, it contacts alternatively with Lys423 and Arg596. At the end, during the last 25 ns, AA loses its interaction with Arg596 and it establishes a interaction with Lys423 that lasts till the end of the simualation. Finally, the AA tail enters the cavity through the opposite end, giving a reverse orientation to the starting one and with the bisallylic carbons far from the OH$^-$ group, but essentially maintaining the same configuration of its backbone.

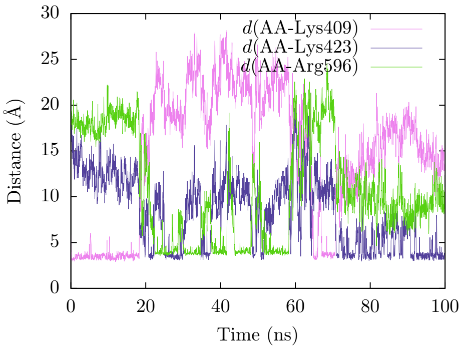
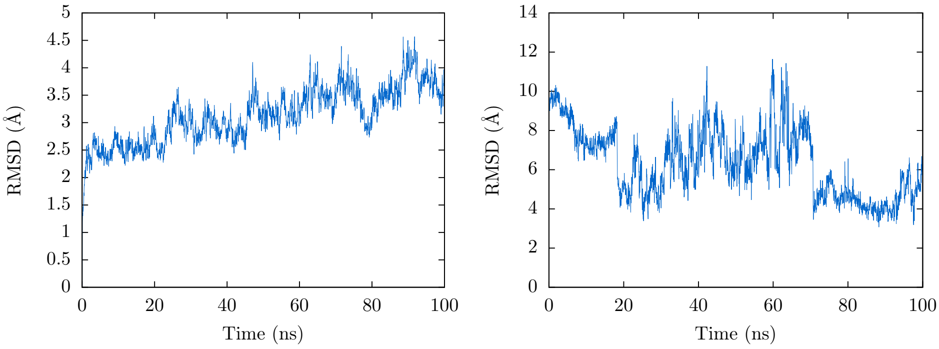


Figure 7.25: Evolution of the distance of the AA carboxylate end to polar/positive residues Lys409, Lys423 and Arg596 during the 100 ns production step of the holoenzyme tail-first simulation.

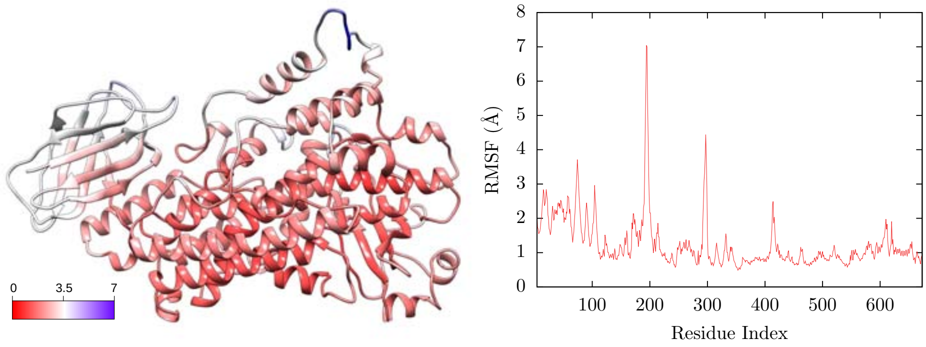
The Phe177+Tyr181 pair stays in a similar conformation as in the crystallographic holoenzyme structure (Figure 7.23b), which gives AA access toward the solvent from inside the binding cavity. Also, His600 orientation (Figure 7.23c) remains nearly unchanged during the course of the simulation, keeping its outward position.



(a) Protein RMSD. **(b)** AA RMSD.

Figure 7.26: Protein and AA heavy atoms RMSDs during the 100 ns production step of the AA tail-first:holo 5-LOX complex .

The active site cavity of the holoenzyme structure obtained here by homology modeling, which basically is a consequence of the structure of the original incomplete crystallographic coordinates, presents a wider and more opened active site cavity than the apoenzyme structure, and also exhibits great plasticity. A wide access channel is precisely at the cavity end where Lys409 is placed, thus this fact could explain why AA leaves the cavity, apparently, so easily. The RMSD of the protein backbone with respect to the initial structure, and the RMSD of the 20 carbon atoms of AA with respect to the substrate average structure is represented in Figure 7.26. The protein



(a) Protein structure.

(b) RMSF of protein Cα.

Figure 7.27: Protein RMSF along the 100 ns production step, showing: (a) protein structure colored according to RMSF values (from 0 (red) to 7 (blue)); (b) RMSF of protein Cα versus protein residue index.

RMSD continuously increases during the simulation, reaching a value near 4 Å at the end, which indicates that the system is not well stabilized and also agrees with the high flexibility of 5-LOX reported in the literature.[29] The RMSF (Root Mean Square Fluctuation) of the protein Cα atoms during the 100 ns trajectory is represented in Figure 7.27. The RMSF presents several peaks, specially standing out a sharp peak in between residues 190-200, which are precisely the residues of the unfolded loop modeled by homology modeling that cannot be seen in the crystallographic coordinates, corroborating the high flexibility of the system.

7.3.4.2 Head first AA:holo complex (solution 12)

For solution 12 (head first) of docking to holoenzyme, AA remains in the active site cavity for the whole 100 ns (see Figure 7.28a), and the bisallylic carbon C_7 is quite far from the OH^- group (Figure 7.29). In Figure 7.30 the evolution of the distances from the AA carboxylate end to polar/positive residues is presented. Initially, the AA carboxylate end is interacting with Gln363 and Gln557. At the end, it is positioned in the part of the cavity surrounded by Gln363, Gln557, Asn424, Ala425, Phe359, His600 and, notably, Tyr181. Tyr181 has changed conformation and is now pointing into the cavity (as in the apo structure) and is making a hydrogen bond with the AA carboxylate (in the last snapshot, the d(Tyr181(OH)-AA(COO)) distance if of 1.75 Å). Phe177 has also changed conformation and points into the cavity providing, all together, a lid that seems to prevent AA from escaping. During the simulation, the hydrophobic tail moves from its initial position to the part of the active site delimited by Trp147, occupying the same zone as described for the solution 11 of the apoenzyme dockings (head first). In

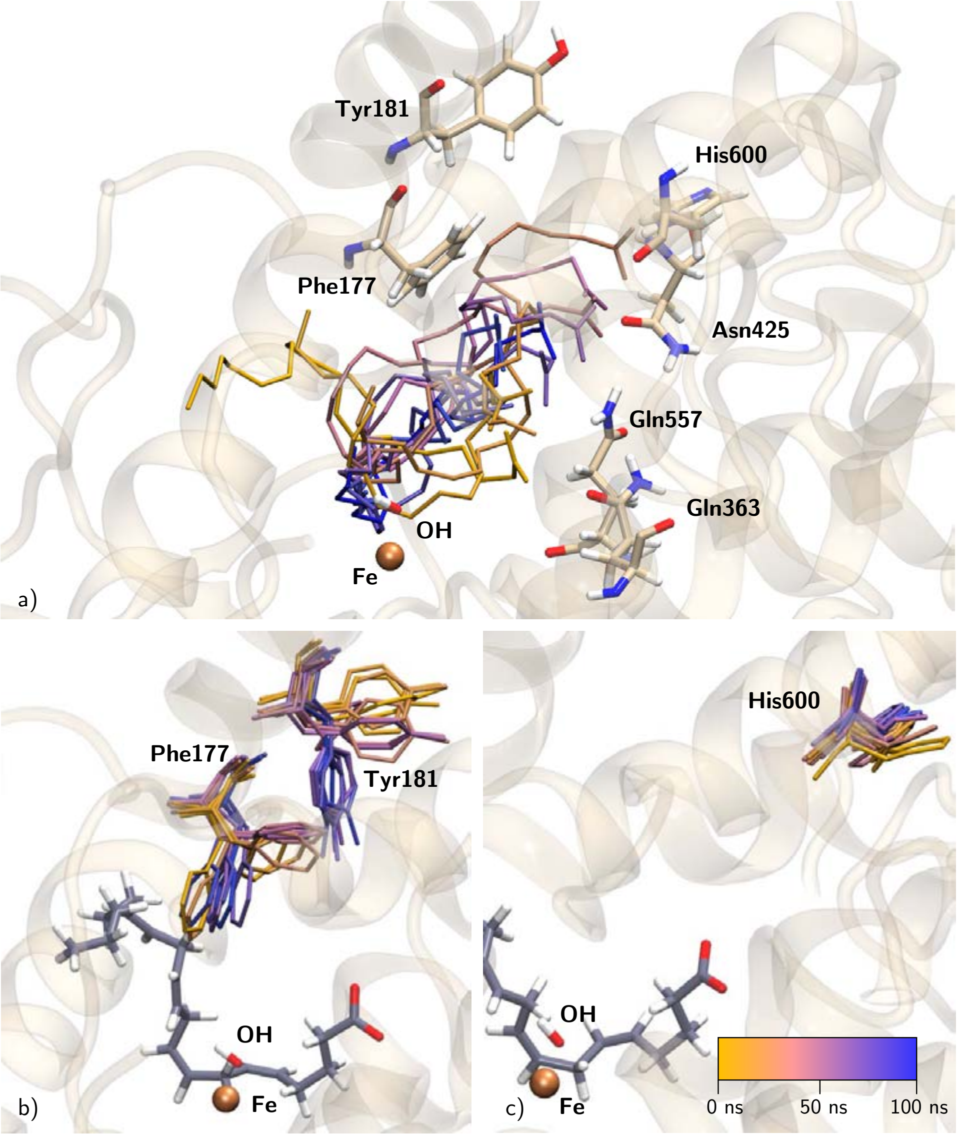
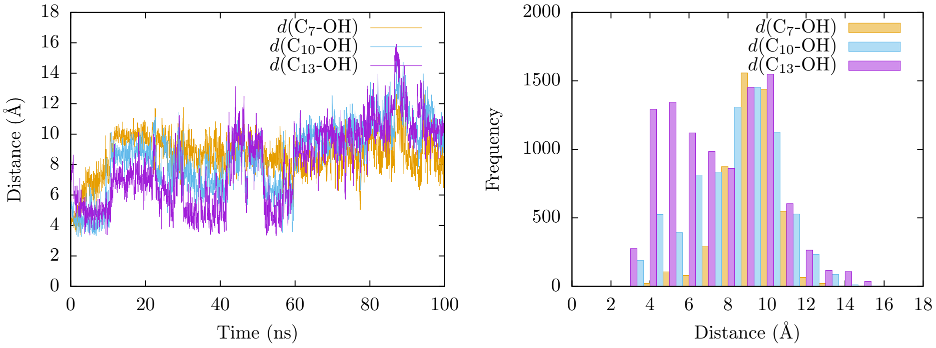


Figure 7.28: Representation of the evolution of the system during the 100 ns production step. a) Evolution of AA. The enzyme is only shown in its initial position. Initially, the AA carboxylate end is interacting with Gln363 and Gln557 in a head-first orientation. This orientation remains during the simulation, but the AA carboxylate end interacts with several positive/polar residues, among His600, Asn425, Gln363 and Gln557; b) evolution of the FY-cork (Phe177-Tyr181), and c) evolution of His600, superimposed with one snapshot of AA (purple-gray).

this case, His600 also maintains its outward orientation during the 100 ns trajectory (Figure 7.28c).

The protein RMSD also tends to values near 4 Å (Figure 7.31a),



(a) Evolution of distances. **(b)** Histogram of distances.

Figure 7.29: Evolution of the distances $d(C_7\text{-}OH)$, $d(C_{10}\text{-}OH)$ and $d(C_{13}\text{-}OH)$ along the 100 ns of production step of holo head-first complex.

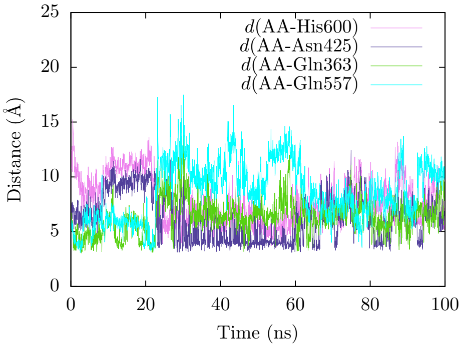
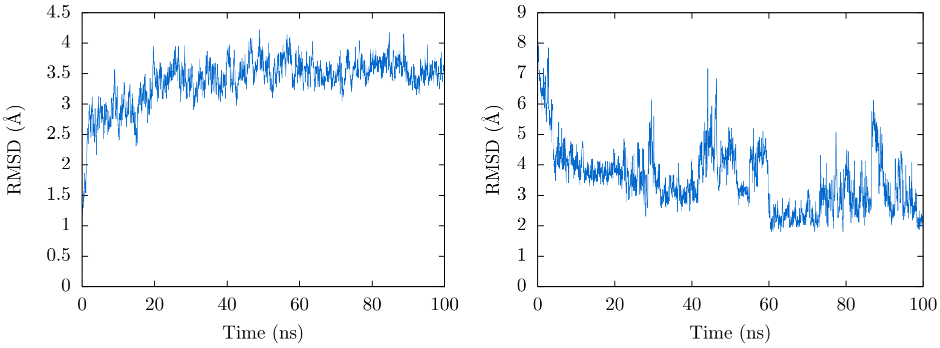


Figure 7.30: Evolution of the distance of the AA carboxylate end to polar/positive residues His600, Asn425, Gln363 and Gln557 during the 100 ns production steps of the holoenzyme head-first simulation.



(a) Protein RMSD. **(b)** AA RMSD.

Figure 7.31: Protein and AA heavy atoms RMSDs during the 100 ns production step of the AA head-first:holo 5-LOX complex .

as in the previous case of the holoenzyme tail-first complex (Figure 7.26a). However, the AA RMSD (Figure 7.31b) shows lower values (Figure 7.26b), as in this case, the substrate stays inside the binding cavity for the whole 100 ns trajectory.

7.3.4.3 Head-first AA:apo complex (solution 11)

For solution 11 (head first) of the docking to apoenzyme, the simulation shows less structural changes (Figure 7.32). Distances distribution for C_7, C_{10} and C_{13} remain qualitatively the same as the one shown during the 20 ns trajectory (compare Figures 7.20 and 7.33), with both C_7 and C_{10} being close to the OH^- group (here C_7 more populated at shorter distances). The AA carboxylate starts interacting with His600 and Asn425 and, after 45 ns a slight shift moves it to establish a hydrogen bond with Tyr181, as observed for solution 12 (see the distances of the AA carboxylate end to close residues in Figure 7.35). His600 has rotated and it is now in the conformation observed for the holoenzyme (Figure 7.32c). Phe177 and Tyr181 stay in a conformation similar to that observed in the crystal structure (Figure 7.32b). Interestingly, during the 100 ns trajectories, solution 12 and 11 (both head first) converge to be localized in the same region of the active site, although the central part of AA is bent down towards the Fe(III)-OH^- cofactor for solution 11, whereas it is bent towards the top (surface) of the cavity for solution 12, which brings the bisallylic carbons away from the Fe(III)-OH^- cofactor.

The apoenzyme structure clearly exhibits less flexibility than the holoenzyme (both head and tail first) during the 100 ns production step. The RMSD for the heavy protein atoms in the apoenzyme reaches its maximum value at $\sim$ 2.5 Å, in contraposition to the holoenzyme simulations, which tend to a protein RMSD of near 4 Å in both cases (tail and head first). There are also remarkably differences in the RMSF functions (Figure 7.36 shows the RMSFs of all AA:5-LOX complexes). The apo form not only does not present the high RMSF peak around residue $\sim$ 200 (the solvent-exposed loop), but it also exhibits higher values at other residues, whereas in the holoenzyme this loop shows a clear maximum.

In a recent study of mutagenesis experiments, it has been proposed that the Phe177/Tyr181 pair does not act as a cork as previously proposed, but as a twist-and-pour closure. This implies that the conformation of these two residues in the holo crystal structure will not be compatible with the pre-catalytic geometry. The results obtained here from the MD simulations suggest that this FY pair probably needs to close the active site cavity during the catalytic process, even if it cannot be concluded that the instability of the holo tail-first is due to the opening of this FY pair. Thus, these 100 ns

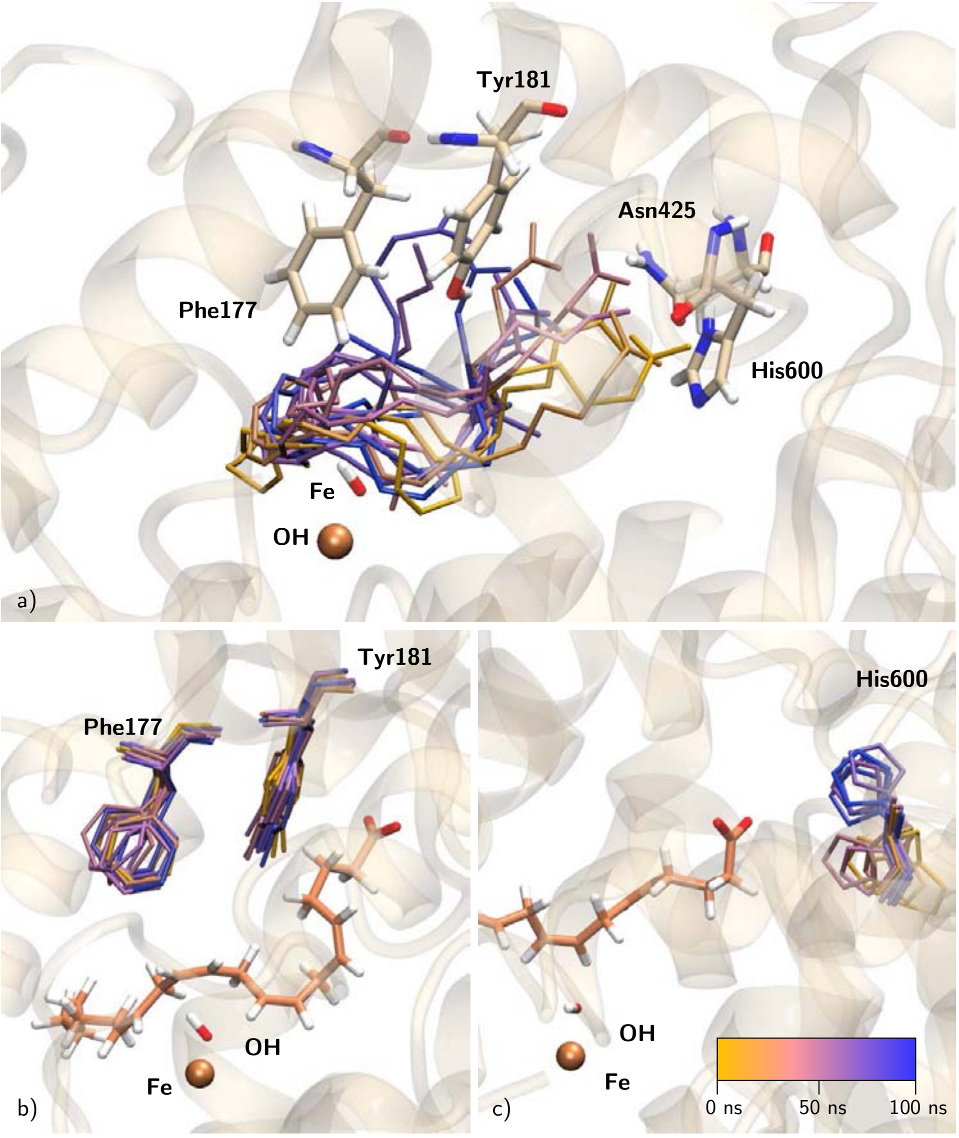
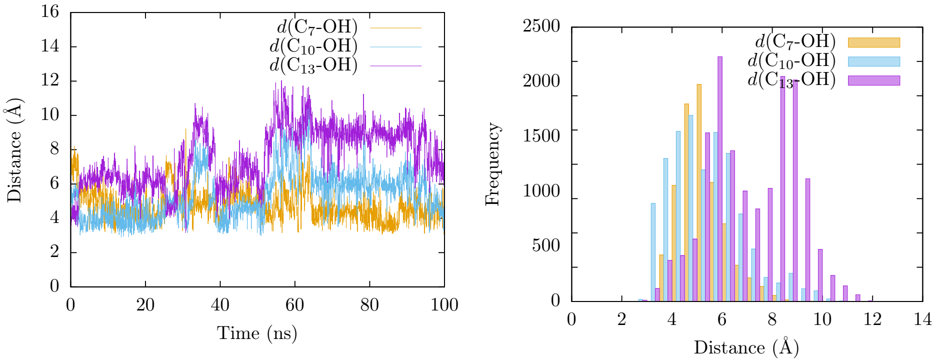


Figure 7.32: Representation of the evolution of the system during the 100 ns production step. a) Evolution of AA. The enzyme is only shown in its initial position. Initially, AA carboxylate end is interacting with His600 and Asn425. After 45 ns, AA head moves towards Tyr181; b) evolution of the FY-cork (Phe177-Tyr181), and c) evolution of His600, superimposed with one snapshot of AA (orange).

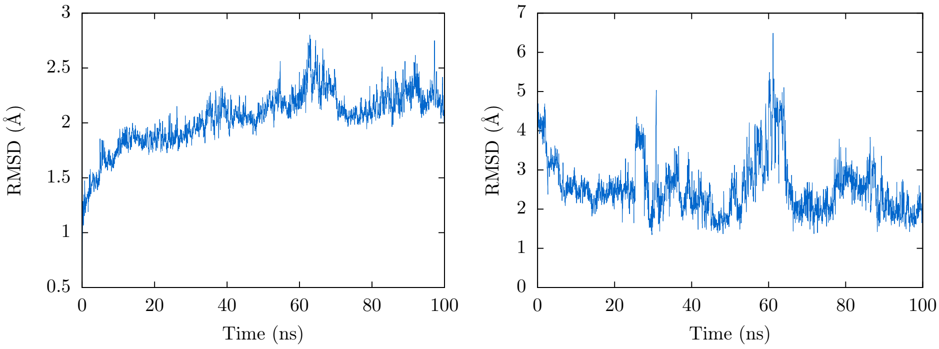
simulations seem to support the idea proposed by Newcomer and coworkers that the Phe177+Tyr181 pair would be âĂIJclosingâĂİ the active site during the chemical reaction. Moreover, in a head-first orientation Tyr181 is the residue that interacts with the AA carboxylate. However, in the case of the tail-first orientation, also studied in this work, it cannot be concluded if the



(a) Evolution of distances in apo head-first (is there other orientation?.

(b) Histogram of distances in apo head-first.

Figure 7.33: Evolution of distances d(C_7-OH), d(C_{10}-OH) and d(C_{13}-OH) along the 100 ns of production step of apo head-first.



(a) Protein RMSD.

(b) AA RMSD.

Figure 7.34: Protein and AA heavy atoms RMSDs during the 100 ns of production step of the AA head-first:apo 5-LOX complex .

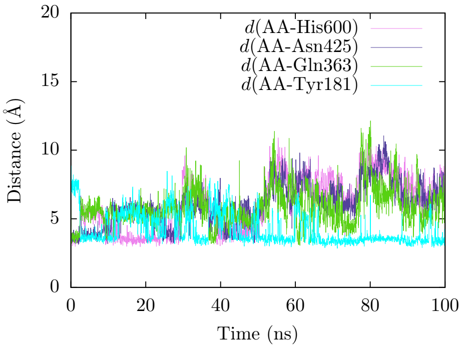


Figure 7.35: Evolution of the distance of the AA carboxylate end to polar/positive residues His600, Asn425, Gln363 and Tyr181 during the 100 ns production steps of the holoenzyme tail-first simulation.

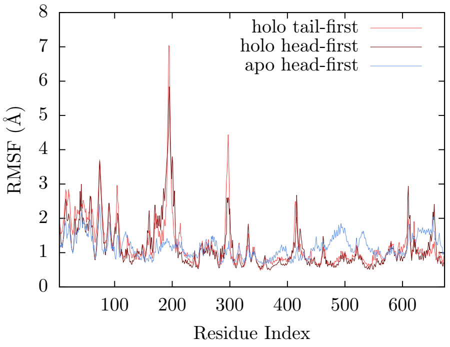


Figure 7.36: RMSFs of protein Cα versus residue index for holo tail first (red), holo head first (brown) and apo head first (blue) complexes along the 100 ns trajectories.

AA:Stable-5-LOX complex is not stable due to the binding mode or as a consequence of the FY-cork being open. On the other hand, His600 seems to adopt the conformation observed in the holo crystal structure and it is not directly interacting with the AA carboxylate, which may be indicating that the effect observed experimentally upon mutation of this residue is due to a change in shape and/or volume of the binding cavity (direct or by affecting other residues), as has been proposed by Kühn and coworkers for the F359W-A424I-N425M triple mutant.[18,50]

7.4 Conclusions

In this Chapter a study of human 5-LOX has been performed using a multiscale approach combining homology modeling, protein-ligand docking, MD simulations and QM/MM calculations. The study is based on the crystallographic data available, using one of the structures of the apo form (3O8Y), and the only one available structure of the holo form (3V99), which lacks some coordinates and has been completed by homology modeling. For both apo and holo forms, the most probable AA binding modes have been determined, as well as the regio- and stereoselectivity expected, and the residence time of AA (i.e. the AA stability), for each binding mode.

The results presented here show that there are two possible binding modes in the holoenzyme: (a) a tail-first type orientation, with the AA carboxylate end interacting with Lys409 and, (b) a head-first type orientation in which the AA carboxylate group interacts with Gln363 (and Gln557). With a tail-first orientation, C_7 is the closest methylene group, whereas C_{10} and C_{13} are further away. In this case, the abstraction of the H_7 pro-S is clearly the most favorable process. On the contrary, with a head-first orientation, C_{10} and C_{13} are much closer to the OH$^-$ group than C_7, with H_{13} abstraction the most favorable process and H_7 abstraction being not possible. Thus, AA can be positioned in the holoenzyme with both head- and tail-first orientations, but only tail first exhibits a 5-lipoxygenating activity, whereas head first yields 15/11-LOX or 12/8-LOX specificity, in terms of hydrogen abstraction. These results are in agreement with the docking simulations of the in silico mutants: when the residues of the bottom of the active site cavity are mutated to bulkier residues, Lys409 no longer can bind AA (tail-first orientation) as the cavity volume decreases, but Gln363 can still bind the AA carboxylate end (head-first orientation) and also gives positions with C_{13} and C_{10} close to the Fe(III)-OH$^-$ cofactor. Thus, AA can be positioned in the active site with both head-first and tail-first orientations, but only tail-first orientation gives structures reactive by C_7.

On the other hand, for the apoenzyme a head-first orientation binding mode has been identified, with the AA carboxylate end interacting with His600, and its methyl tail pointing towards Trp147. His600 is positioned outward in the holoenzyme, whereas in the apoenzyme is directed inward to the Fe(III)-OH$^-$ cofactor. As a consequence, AA cannot interact with His600 in the holoenzyme. In this binding mode of the apoenzyme, both C_7 and C_{10} are close enough to the catalytic iron being C_{13} quite further, and QM/MM calculations predict that both H_7 pro-S and H_{10} abstraction processes are feasible. Thus, we could suppose that the reactive binding mode of AA inside human 5-LOX could be a head-first orientation with the carboxylate end interaction with His600, as proposed by an experimental model.[49] However, at the current state of the experimental knowledge, it would seem somewhat unlikely for three reasons: Firstly, AA should have entered into the active site maintaining the inward His600 orientation corresponding to the apoenzyme, in contraposition to the His600 outward orientation of the available crystallographic structure of the holoenzyme. Second, that orientation would lead to both the pro-S H_7 and H_{10} abstractions, what it is not experimentally observed. Third, in a recent study by Wu et al.,[204] a large molecule known as a competitive inhibitor of 5-LOX, was docked into the binding pocket of the apoenzyme crystal structure. The negative correlation coefficient obtained between the experimental and predicted IC_{50} made the authors conclude that this structure cannot be used for drug screening before the conformational flexibility of 5-LOX is further analyzed and included in the model.

Finally, long MD simulations have shown that only the head-first binding modes (both holo and apo forms), remain relatively stable during the 100 ns, whereas in the tail-first binding mode, AA loses its interaction with Lys409 and abandons the cavity at 18 ns, entering later through the opposite side and resulting in a reverse orientation which inverts the methylene distances order. In the holoenzyme head-first binding mode, AA starts interacting with Gln363 and Gln557, but it ends up at the part of the cavity surrounded by Gln363, Gln557, Asn425, His600 and Tyr181, and with C_7 far away from the catalytic iron. In the apoenzyme AA carboxylate end evolves to form a hydrogen bond with Tyr181, whereas His600 rotates from the inward to the outward orientation corresponding to the holoenzyme.

Summarizing, based on the crystallographic data, three possible binding modes of AA to 5-LOX have been identified, but only one of them would clearly exhibit predominant 5-lipoxygenating activity. However, long MD simulations show that this binding mode is not stable, at least with the holoenzyme crystalographic structure available. So, it can be concluded, that the holoenzyme currently available at the PDB is not good enough to reproduce the actual conformation of the protein when the catalytic process occurs.

The results of this Chapter have been published in paper:

- *Physical Chemistry Chemical Physics*, vol. 18, no. 113, pp. 23017-23035, 2016

Chapter 8

8*R*-LOX

8.1 Introduction

So far, the factors governing the reaction specificity of the first step of LOXs reaction mechanism have been studied for 15-LOX-1 and 5-LOX. These two isoforms are *S*-lipoxygenating enzymes, as they generate the final hydroperoxidation product with *S*-stereochemistry. To complete our knowledge of the regio- and stereospecificity of LOXs reaction, the determinants of the stereocontrol will be also considered. The number of *S*-LOXs in nature is greater than their counterparts *R*-LOXs, and much less is known about the latter ones. Coral 8*R*-LOX, that generates 8*R*-HETE, is one of the *R*-LOXs family members whose crystallographic structure has been solved. The crystal structure of the ligand-free form of 8*R*-LOX was reported a few years ago (PDB ID 2FNQ)[31] and, more recently, a complete structure of the AA:8*R*-LOX complex was released by the same authors (PDB ID 4QWT).[32] In this structure, unlike the case of human 5-LOX, the substrate coordinates are perfectly determined, and a tail-first substrate orientation with the AA carboxylate end interacting with Arg182 (Figure 8.1) is observed.

Experimentally, it has been observed that the hydrogen abstraction step comes essentially from a unique position, pro-*R* H_{10}.[32] In the crystallographic structure of the AA-bound form the substrate is well-positioned to yield the hydrogen abstraction from C_{10} (Figure 8.1). What is more interesting, though, is the second reaction step, the oxygen insertion. It has been proposed in the literature that a single residue can direct the stereochemistry of the hydroperoxidation product.[3,16,53] In this sense, it has been found that in many *S*-LOXs a conserved Alanine of the active site is located over the pentadienyl group of the substrate, whereas in many *R*-LOXs this residue turns out to be a Glycine. This Gly/Ala "switch" can direct the incoming oxygen molecule either to the $n-2$ or $n+2$ position of the pentadienyl moiety (where n is the hydrogen abstraction position), which results in hydroperoxidation products with opposite stereochemistry.

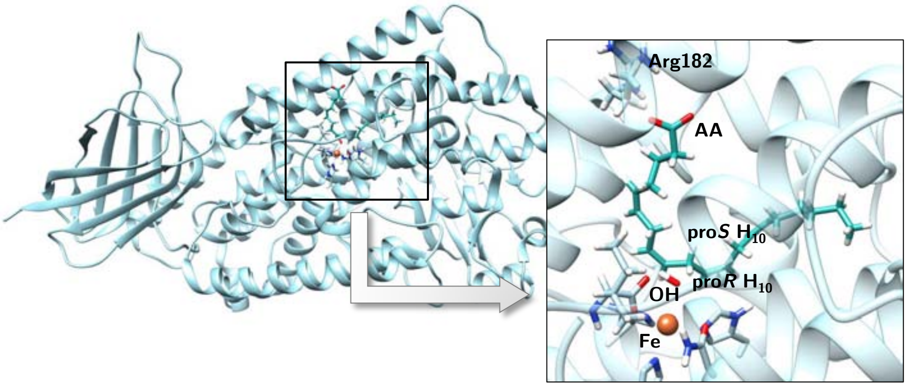


Figure 8.1: Representation of subunit C of structure 4QWT. The box on the left shows a detailed view of the active site, showing the AA substrate, the Fe-coordination sphere, and Arg182.

The existence of this "switch" has been suggested based on mutagenesis experiments.[16] For instance, the mutation of Gly427 in 8R-LOX to an Ala changes dramatically the product specificity, resulting in almost exclusively 12-HETE with 98% of the S-stereoisomer. Similar alterations are also observed in another R-LOX, human 12R-LOX, in which the Gly441Ala mutant produces the 8-HETE S-stereoisomer. In the same way, when the opposite mutation Ala → Gly is performed in a S-LOX, the corresponding mutant generates the R-HETE product. This alteration has been observed in mouse 8S-LOX and human 15S-LOX-2 whose mutants led to the 12R-HETE and 11R-HETE products, respectively. To explain this change in the enzyme stereoselectivity, it has been proposed that the Alanine residue in 15S-LOX-2, for instance, shields the $n-2$ (C_{11}) position favoring the oxidation at the $n+2$ (C_{15}) position in S-LOXs, whereas in 8R-LOXs the Glycine does not block the $n-2$ (C_8) position favoring the oxidation at this site. So, when Ala is mutated to Gly in 15S-LOX-2, oxidation at position $n-2$ is not impeded anymore leading to the 11R-HETE product, whereas in the Gly427Ala 8R-LOX mutant position $n-2$ is blocked and oxygen enters at the $n+2$ position giving the 12S-HETE product. Moreover, the presence of one or another residue in this position may suppose further structural changes in the active site, and the complete molecular basis is not well-understood yet. In coral 8R-LOX, in particular, it has been suggested that the invariant Leu431, which is spatially close to Gly427, may be involved in controlling the oxygen access, but this supposition is based only in crystallographic data.

So, the main interest of this Chapter is focused on the second step of the reaction mechanism, the oxygen insertion step. This second step has been extensively studied in our research group[191–193,205] in rabbit 15-LOX-

1, a *S*-LOX. In this Thesis coral WT 8*R*-LOX and its Gly427Ala mutant have been selected to explore the mechanistic basis of the stereocontrol of LOXs hydroperoxidation process by means of analyzing at molecular level the oxygen insertion step. In addition, the hydrogen abstraction reaction (first step of the catalytic mechanism) in coral WT 8*R*-LOX has also been studied.

8.2 Methodology

8.2.1 Preparation of the initial coordinates of the AA:-8*R*-LOX complex

Crystal structure 4QWT of 8*R*-LOX contains the AA substrate in subunit C. This subunit including the substrate has been selected and prepared to perform the study of the reactivity of this system. The crystallographic coordinates lack a small region (indexes 307-315) in all the four subunits. This loop region has been reconstructed in subunit C by the “loop refinement” facility of program Modeller9.2.[198]

8.2.2 *In silico* mutagenesis

The starting structure employed to generate the *in silico* mutant Gly427Ala corresponds to the hydrogen abstraction product of the WT 8*R*-LOX (see results below), as this mutation has been proposed to affect only the second reaction step. The mutant was generated by changing the residue Gly427 to Ala with the Mutator Plugin of the VMD program.[172]

8.2.3 Molecular Dynamics simulations

All hydrogen coordinates have been generated with the program PROPKA version 3,[206, 207] using a standard pH = 7 for the tritable residues. The system has been solvated in a cubic box of pre-equilibrated TIP3P water molecules of dimensions 125 Å x 125 Å x 125 Å. Then the system was neutralized with K^+ ions, and an additional KCl salt concentration was added to reproduce the experimental ionic strength of a 0.15 M solution. The resulting system contains nearly 184000 atoms, about 11000 of which belonging to the protein.

MD simulations were run with the program NAMD2.9.[125] The CHARMM36 force field[208] was used for the protein and lipid atoms, and the Fe coordination sphere was described by an update of the parameters specifically derived by Saam *et. al.*.[187] First, the system was sent to 10000 initial energy minimization steps using the Conjugate Gradient algorithm. Then, MD simulations using PBC were carried out. An equilibration step of 1 ns at constant temperature and pressure (300 K and 1 atm) was run . Langevin dynamics was used to control temperature and the Nosé-Hoover method[209] with Langevin piston pressure control[210] was employed to control fluctuations in the barostat. Harmonic restraints to the protein backbone (with a force constant of 1.0 kcal mol^{-1} $\mathrm{Å}^{-2}$) and side chains (with a force constant of 0.1 kcal mol^{-1} $\mathrm{Å}^{-2}$) were applied by means of the definition of a set of collective coordinates,[211] while the remaining system was kept free of restraints. Next, a production step of 100 ns was within the same (N,P,T) ensemble. A time step of 2 fs was used along the whole MD trajectory. All bonds and angles containing hydrogen atoms were constrained by the SHAKE algorithm.[212] Long-range electrostatic interactions were treated by the PME method.[213] A cutoff of 12 Å was used to treat the non-bonding interactions.

8.2.3.1 MD simulations of the Gly427Ala mutant

For the Gly427ALa mutant the same protocol as described in the previous section was applied, except that, for the production step, a total of 50 ns were run.

8.2.4 QM/MM calculations

The program Q-CHEM (version 4.2)[214, 215] with the CHARMM interface[88] was used. The complete system derived from the MD simulations was trimmed to a 20 Å radius sphere centered on the AA molecule, containing 2000 atoms. The QM region was described for all the atoms at the B3LYP-D3/6-31G(d,p) level.[216, 217] The MM region was described by the CHARMM36 force field.[208] For the study of the hydrogen abstraction step (with a sextet multiplicity) the QM region was comprised by 55 atoms (Figure 8.2a): 8 atoms of each imidazole ring of the three Histidines, the side chain of Asn574, the terminal carboxylate group of Ile693, the pentadienyl group around C_{10} of the lipid substrate, and the Fe(III)-OH^- cofactor. A full electrostatic embedding scheme has been employed and seven link atoms were used for the QM/MM boundary. No cutoffs were introduced for the non-bonding MM and QM/MM interactions.

For the oxygen insertion step (with a doublet multiplicity because the Fe atom is included in the MM region), a smaller QM region was employed, containing the oxygen molecule and the pentadienyl radical centered on C_{10}, making a total of 18 QM atoms (Figure 8.2b). In this case, two link atoms were employed in the QM/MM boundary.

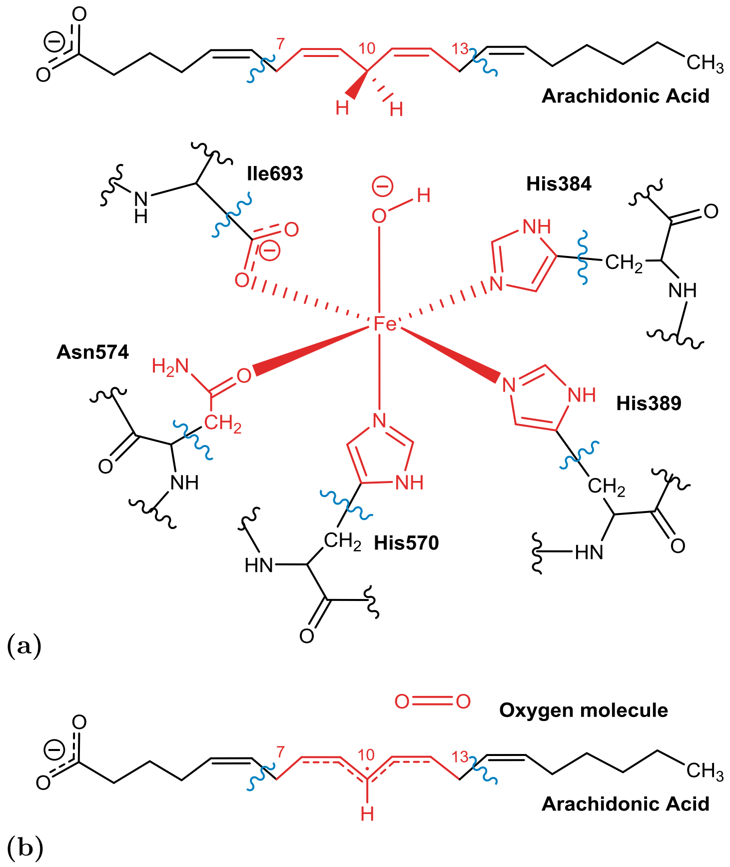


Figure 8.2: Representation of the QM/MM boundary employed in the hydrogen abstraction step (a), and in the oxygen insertion step (b). QM atoms are depicted in red, and link atoms are represented by wavy blue lines.

8.2.5 Free Energy Calculations

The Free Energy Profile has been calculated using the Umbrella Sampling (US) method. For the hydrogen abstraction step, a total of 19 windows separated by 0.2 Å along the reaction coordinate were selected, comprising the entire range of the reaction coordinate. In the same way, 13 windows were selected to study the oxygen insertion step. For each window, an initial step 0.2 ps of classic MD was performed, followed by 3 ps of QM(B3LYP-D3/6-31G(d,p))/MM dynamics. Each simulation was restrained into the corresponding particular window with a force constant of 250.0 kcal mol^{-1} Å^{-2}). The free energy profile was computed employing two different algorithms, the binless implementation of the weighted histogram analysis method (WHAM),[116,218] and a high-precision implementation of the dynamic histogram analysis method (DHAM)[117] to calculate both the free energy profile and the kinetic rate constant of the reaction.

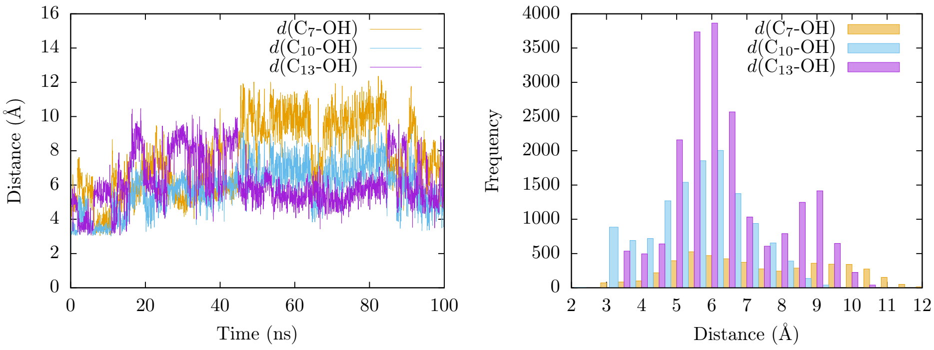
8.3 Results and Discussion

8.3.1 Hydrogen abstraction step

8.3.1.1 MD simulations

The starting structure for the MD simulations was derived from the crystal structure 4QWT, subunit C, crystalized with the substrate in the active site. In this structure the substrate is positioned in a tail-first orientation with AA carboxylate end interacting with Arg182. Although this crystal structure could be considered as a good starting point to perform the pro-*R* H_{10} abstraction process, as the C_{10} methylene group is very close to the catalytic iron, an structure derived from the MD simulations has been selected for this purpose, in order to account for the relaxation of the system at the given temperature.

In this sense, the first criterion to select a starting structure in good agreement with the regioselectivity of 8*R*-LOX is based on the distances of the three reactive bisallylic methylenes to the oxygen atom of the Fe(III)-OH^- cofactor. These distances have been monitored along the MD simulation giving the results presented in Figure 8.3. The average of the $d(C_7\text{-OH})$, $d(C_{10}\text{-OH})$ and $d(C_{13}\text{-OH})$ distances (standard deviation in parenthesis), are of 7.57 (2.09) Å, 5.47 (1.33) Å and 6.16 (1.50) Å, respectively. Therefore, C_{10} is in average closer to the catalytic iron than C_7 and C_{13} as it would be expected for a 8-LOX reaction specificity. In Figure 8.3a the evolution of the three distances is represented. During the first 15 ns, C_{10} is clearly closer to the Fe(III)-OH^- cofactor. After 15 ns all the three bisallylic methylenes get further, but $d(C_{10}\text{-OH})$ remains as the shortest distance till 45 ns. Then an inversion of the distances order occurs and $d(C_{13}\text{-OH})$ becomes the shortest



(a) Evolution of distances. **(b)** Histogram of distances.

Figure 8.3: Evolution of distances $d(C_7\text{-OH})$, $d(C_{10}\text{-OH})$ and $d(C_{13}\text{-OH})$ along the 100 ns of production step.

distance till 85 ns, after what, $d(\mathrm{C}_{10}\text{-OH})$ becomes again the closest. Thus, fluctuations of the distances take place during the MD production trajectory, but a good number of conformations where C_{10} is in a better disposition with respect the oxygen of the OH^- are generated. Figure 8.4 shows the first and last frame of the MD production step.

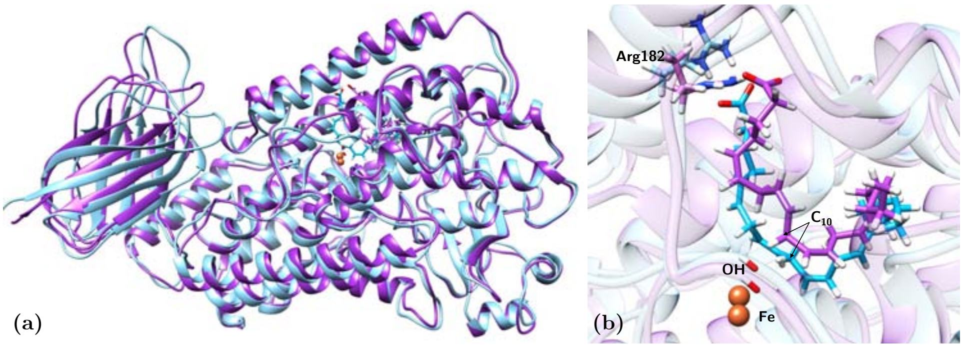
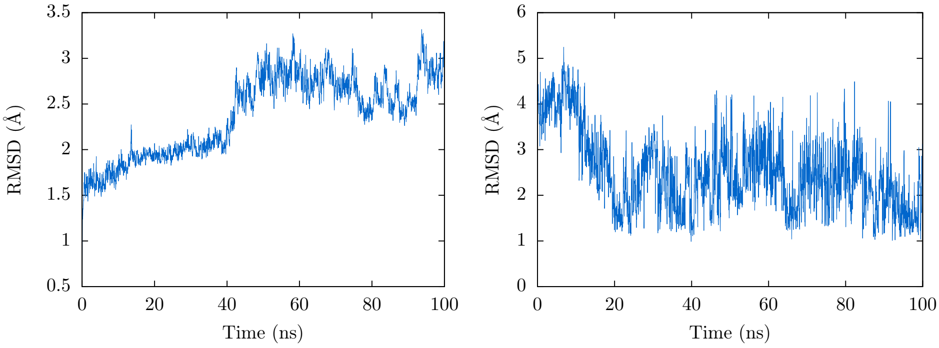


Figure 8.4: Overlay of the first (blue) and last (purple) frame of the MD production step showing the complete protein structure (a) and the active site cavity (b).

The RMSD of the protein heavy atoms with respect to the initial frame of the production step, and the RMSD of the AA heavy atoms with respect to the average structure are represented in Figure 8.5. A jump the in protein RMSD is observed at 45 ns, which mainly corresponds to a fluctuation in the N-terminal domain. From this point, the protein RMSD stabilizes. In the case of the AA RMSD it is observed a larger value at the beginning of the simulation after which the RMSD also stabilizes.



(a) Protein RMSD. **(b)** AA RMSD.

Figure 8.5: Protein and AA heavy atoms RMSDs during the 100-ns production step.

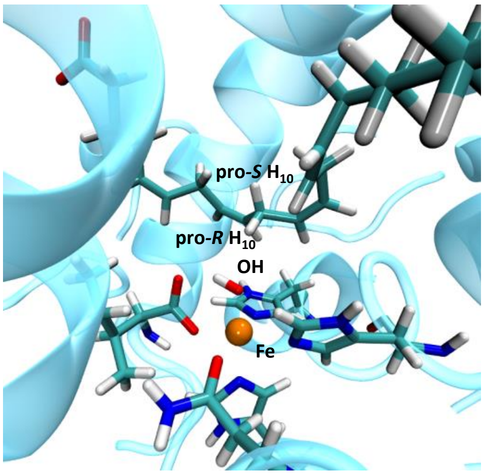


Figure 8.6: Selected snapshot form the MD simulation to perform the reactivity study. This snapshot is found during the first steps of the MD production runs, where the distances distribution of the bisallylic methylenes is in agreement with the regiospecificity of 8*R*-LOX in terms of hydrogen abstraction.

8.3.1.2 QM/MM calculations

The structure selected for the subsequent calculations (Figure 8.6) has been selected from the first 45 ns of the MD production step according to the distances criterion . Geometric parameters of interest of this structure are given in Table 8.1.

Table 8.1: Geometric distances (in Å) of the selected snapshot from MD simulation and the resulting structure optimized at the QM/MM level.

Structure	***d*(C_7-OH)**	***d*(C_{13}-OH)**	***d*(C_{10}-OH)**	***d*(pro-*R* H_{10}-OH)**	***d*(pro-*S* H_{10}-OH)**
MD snapshot	4.93	5.74	3.41	2.75	4.40
Reactant minimum	4.96	5.57	3.22	2.44	4.14

It is observed that the reactant structure fulfills the conditions of 8-lipoxygenation specificity in terms of hydrogen abstraction: the closest carbon atom to the oxygen of Fe(III)-OH^- cofactor is clearly C_{10}; pro-*R* H_{10} is closer than C_{10}, presenting thus a good orientation towards the reactive OH^- to initiate the reactive process, and also, pro-*S* H_{10} is further than pro-*R* H_{10} (and C_{10} as well), thus being not able to compete with pro-*R* H_{10} for the reaction.

The optimized geometry of the reactant minimum has been used as the starting point to construct the potential energy profile along the reaction coordinate. The reaction coordinate, z, is defined as the difference between the distance of the breaking bond (C_{10}-H_{10}) and the forming bond (H_{10}-O), being H_{10} the pro-R hydrogen. To construct the potential energy profile, a series of optimizations of the mobile part of the system have been performed in the presence of harmonic restrictions on the reaction coordinate, which increases with a step size of 0.2 Å. The process has been repeated forward and backward until convergence to avoid hysteresis. The resulting potential energy profile is depicted in Figure 8.7.

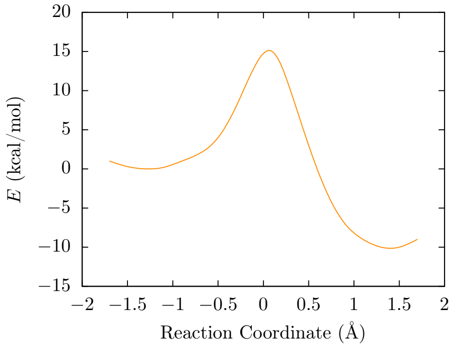
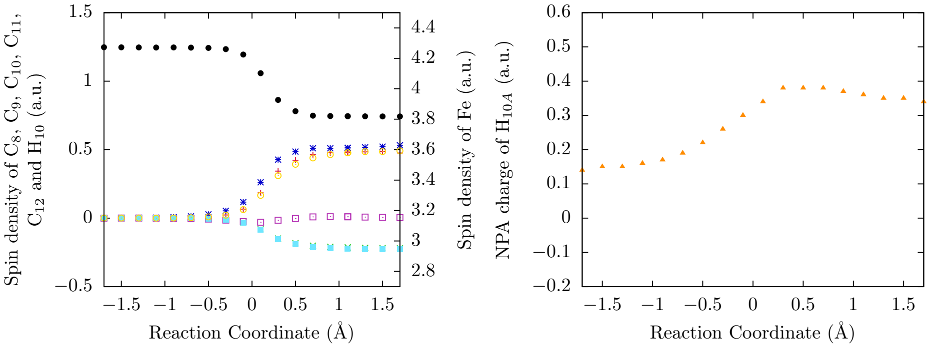


Figure 8.7: Potential energy profile for pro-R H_{10} abstraction. The converge profile is achieved for a potential energy barrier of ∼ 15 kcal/mol.

Along this profile, reactants are located approximately at $z = -1.3$ Å, and products at $z = 1.3$ Å. The transition state (TS) structure is located at $z = 0.0$ Å, which indicates that the TS is highly symmetrical with respect to the movement of the transferred atom along the reaction coordinate. The potential energy barrier for pro-R H_{10} abstraction is around 15.0 kcal/mol. This is a low barrier compared with previous studies of hydrogen abstraction processes in similar systems (19.6 kcal/mol in 15-LOX-1 (see Chapter 3) and 18.0 kcal/mol in 15-LOX-2 for pro-S H_{13} abstraction). This result is in agreement with experiment, since the catalytic rate constant for 8R-LOX is higher than for other mammalian LOXs (see Introduction Chapter for details). The reaction is exoergic by 10.0 kcal/mol.

8.3.1.3 Spin densities

Atomic populations from spin density, which indicate the excess of α spin, have been evaluated during the reactive process for atoms C_8, C_9, C_{10}, C_{11} and C_{12} of the pentadienyl system, the abstracted hydrogen pro-R H_{10} , and the catalytic iron (Figure 8.8). Initially, the spin density on the mentioned AA atoms is zero, and as the reaction progresses, an increasing spin density appears on C_{10} and delocalizes over C_{10}, C_8 and C_{12}, going from 0 a.u. to 0.5 a.u. over each of these atoms. At the same time, the spin density on C_9 and C_{11} changes from 0 to -0.2 a.u. The sum of spin densities over these carbon atoms yields approximately 1.0 a.u. which corresponds to one unpaired electron delocalized over the pentadienyl radical system. The spin density on Fe decreases from 4.1 to 3.7 au, as corresponds to the transition of a Fe(III) sextet configuration to a Fe(II) quintet configuration. Spin density on pro-*R* H_{10} remains virtually zero during the reactive process, indicating that this atom correspond to a proton rather than a hydrogen atom bearing an electron. Furthermore, this atom is positively charged during the reaction process, whit a charge changing from 0.15 to 0.38 a.u. Thus, this process corresponds to a Proton Coupled Electron Transfer (PCET) process, in which the proton and the electron are transferred to different acceptors: the proton is transferred to the OH^- group oxygen to produce water, whereas the electron is transferred from the C_8-C_{12} pentadiene group of AA to the Fe(III) to form Fe(II).



(a) Spin density. **(b)** NPA change.

Figure 8.8: (a) Spin densities on atoms C_8 (red), C_9 (green), C_{10} (blue), C_{11} (yellow), C_{12} (cyan), pro-*R* H_{10} (magenta) and Fe (black), and (b) NPA chages on pro-*R* H_{10} during the hydrogen abstraction step.

8.3.1.4 Free energy calculations

To account for the thermal and entropic contributions the free energy profiles were calculated using the Umbrella Sampling (US) method. Biased QM/MM molecular dynamics simulations were performed for 57 ps distributed in 19 umbrella windows at the B3LYP-D3/6-31G(d,p) level using the same QM region as previously defined (Figure 8.2a). Here we have used the recently developed DHAM method to calculated both the free energies as well as the kinetic rate corresponding to the catalytic reaction. We used a new high precision Matlab implementation of the DHAM method to numerically obtain accurate free energy profiles for this reaction with a high free energy barrier, in which the largest eigenvalues of the Markov matrix are almost degenerate. The histograms of the different US windows are represented in Figure 8.9, where it is observed that there is good overlapping among the windows.

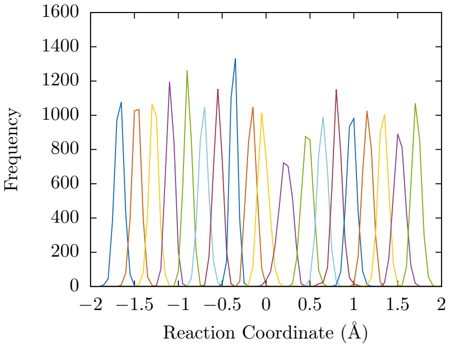


Figure 8.9: Histogram of the 19 US windows sampled in the hydrogen abstraction process.

The Markov matrices corresponding to the unbiased transition probabilities of the system have been calculated at ten different lag times (1, 2, 3... 10 fs) and three different number of bins (600, 800 and 1000), and their spectral decomposition have been determined. The slowest relaxation rate corresponding to the second largest eigenvalue provided an average rate constant of 35 s^{-1} with a considerable standard deviation of 30 s^{-1}, calculated using the different conditions (lag times and number of bins) above. The free energy barrier using the DHAM method was found to be 16.0 kcal/mol (Figure 8.10), very similar to the potential energy barrier (Figure 8.7). The DHAM calculated free energy compares well with the phenomenological free energy barrier of 14.4 kcal/mol at $T = 300$ K ($k_{\text{cat}} = 210$ s^{-1}). The unidimensional free energy profile using the traditional WHAM method has been also calculated for comparison with DHAM (Figure 8.10). A free energy barrier of 17.7 kcal/mol is obtained in this case.

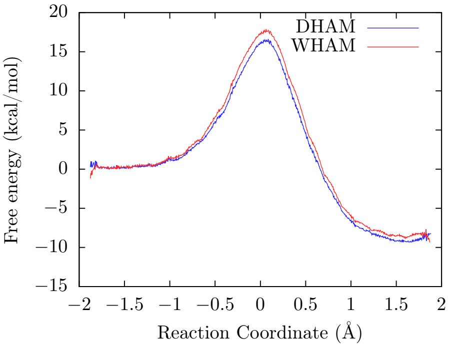


Figure 8.10: Free energy profiles (red WHAM and blue DHAM) calculated at the B3LYP-D3/6-31G(d,p) level of theory for the hydrogen abstraction step.

At this point, it can be noted that, according to canonical variational transition-state theory,[219, 220] the free energy barrier (ΔG^{CVT}) is related to the rate constant by means of the equation

$$k(T) = \Gamma \frac{k_\mathrm{B}T}{h} e^{-\Delta G^\mathrm{CVT}/RT} \tag{8.1}$$

where k is the rate constant, T is the temperature, k_B is BoltzmannâĂŹs constant, h is PlanckâĂŹs constant, R is the gas constant, and Γ is the quasiclassical transmission factor that corrects the rate constant for dynamical recrossing through the dividing surface corresponding to the canonical variational transition (CVT) state (that is, where the generalized transition-state free energy of activation is maximum). The frequency factor $k_\mathrm{B}T/h$ equals 6.3×10^{12} s^{-1} at $T = 300$ K, but the calculation of Γ is not easy. The advantage of using the DHAM method is that, at least at an approximate level, the barrier-crossing times can be estimated directly from the global analysis of local Umbrella Sampling trajectories, so that the effect of the dynamical recrossing is incorporated. Thus, the entire pre-exponential factor in Eq. 8.1 can be obtained substituting in that equation the rate constant and the free energy barrier calculated by the DHAM method, leading to a value of approximately 14.2×10^{12} s^{-1} at 300 K. This value does not indicate a significant contribution of dynamical recrossing at the CVT dividing surface in this case. Such a conclusion is in agreement with the results obtained in a previous study carried out in our research group on the H-abstraction process in 15-LOX-2.[221]

8.3.2 Oxygen insertion step

8.3.2.1 Initial position for oxygen attack

To start the study of the second reaction step, the oxygen addition to the AA pentadienyl moiety, the structure corresponding to the product minimum of pro-R H_{10} hydrogen abstraction step (see previous section) has been fully optimized at the QM/MM level. The resulting structure consists of a radical delocalized over the C_8-C_9-C_{10}-C_{11}-C_{12} pentadienyl system (see spin density in Figure 8.8). The oxygen positioning procedure has been performed over this structure. In order to avoid favoring certain initial positions, several oxygen molecules were initially placed randomly taking into account the Cartesian axes (Figure 8.11) and making a total of 18 oxygen molecules.

Figure 8.12 shows the initial poses of the oxygen molecules of Figure 8.11 along the AA substrate. The oxygen molecules have been placed pointing to the target carbon atom with an initial distance from the closest oxygen atom to the carbon atom of 3.5 Å. Two different sets of structures were generated, those prepared for the oxygen attack at C_8, and those corresponding to the attack at C_{12}, making a total of 36 starting structures.

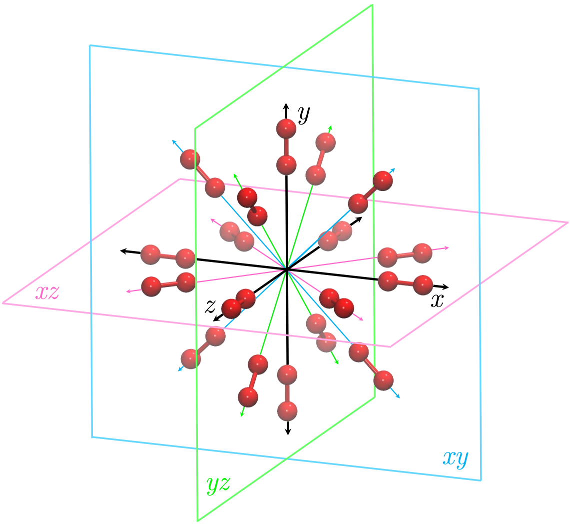
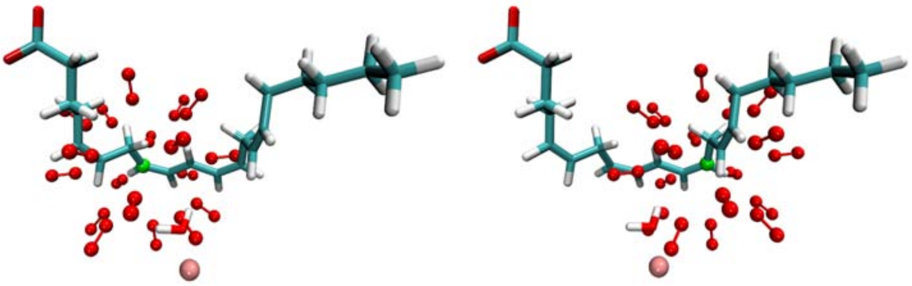


Figure 8.11: Starting positions of the oxygen molecules. These molecules are placed along the *x*, *y*, and *z* Cartesian axes and along the bisector axes contained in the *xy*, *xz* and *yz* planes, being the target carbon (C_8 and C_{12}) atom at the origin. A total of 18 oxygen molecules were placed around each of the two reactive carbon atoms as shown in the Figure.

For each structure a single-point energy calculation has been performed in order to filter out the most stable positions. The results show that there are two positions that clearly exhibit higher stability (one corresponding to the placement around C_8 and the other around C_{12}), and it stands out that in both positions the oxygen molecules are similarly placed (Figure 8.13), being somehow above C_{10}. So, in principle, oxygen addition to both C_8 and C_{12} could be feasible. It is also interesting that the two most stable oxygen poses have a relative orientation with respect to the Fe(III)-OH^- cofactor of antarafacial stereochemistry, which is in good agreement with



(a) O_2 molecules around C_8. **(b)** O_2 molecules around C_{12}.

Figure 8.12: Initial oxygen positioning around C_8 (a) and C_{12} (b). In each case, 18 O_2 molecules have been positioned with a distance of 3.5 Å to the target carbon atom, represented as a green sphere.

the known experimental data. These single-point energy calculations were performed with the complete QM part, including the coordination sphere.

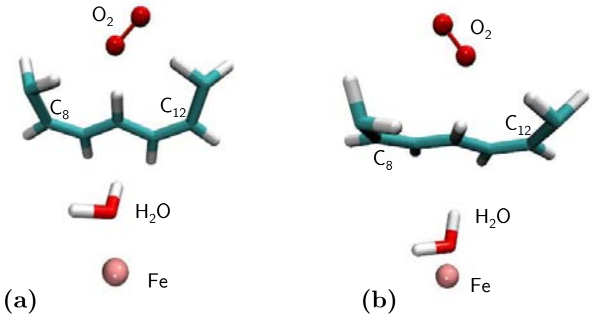


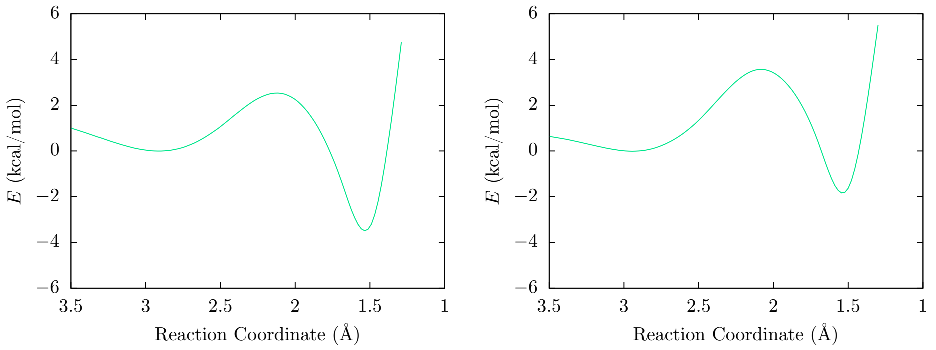
Figure 8.13: Representation of the two most stable initial positions for the oxygen attack at (a) C_8 and (b) C_{12}.

8.3.2.2 Potential energy profiles for the oxygen insertion step

Following the same procedure as that for the hydrogen abstraction step, the potential energy profiles for the oxygen attack have been constructed The reaction coordinate for the second reaction step, z_2, has been defined as the distance between the closest oxygen atom of the incoming O_2 molecule and the target carbon atom (C_8 or C_{12}) of the pentadienyl radical system. This coordinate has been scanned forward and backward until convergence to avoid hysteresis. A step size of 0.2 Å has been defined. The potential energy profiles for the insertion at C_8 and C_{12} are shown in Figure 8.14 As the oxygen attack process only involves the oxygen molecule and the AA radical substrate, a small QM region containing only the pentadienyl moiety and the oxygen molecule (18 atoms), and excluding the Fe coordination environment has been selected (Figure 8.2b). The QM region has been described at B3LYP-D3/6-31G(d,p) level of theory.

For oxygen attack at C_8 reactants lay approximately at $z_2 = 3.0$ Å, products at $z_2 = 1.5$ Å, and TS at $z_2 = 2.2$ Å. The potential energy barrier for this process is approximately 2.5 kcal/mol, and the reaction is exoergic by 3.6 kcal/mol. In the case of the oxygen addition at C_{12}, reactants are located at $z_2 = 2.9$ Å, products at $z_2 = 1.5$ Å, and TS at $z_2 = 2.1$ Å. In this case, the potential energy barrier is of 3.6 kcal/mol, and the reaction is exoergic by 1.6 kcal/mol. Thus, there is a difference of 1.1 kcal/mol in potential energy barriers, being the insertion at C_8 the most favorable process, which is in agreement with the experimental data. Atomic populations from spin density for oxygen insertion step are also provided (Figure 8.15).

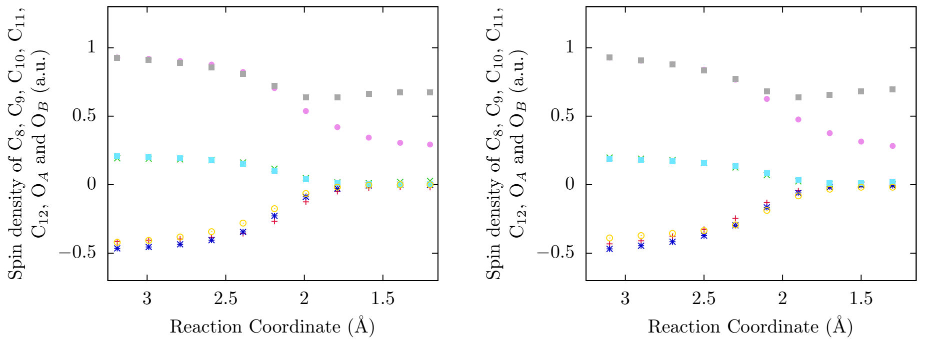
In both cases, spin density on C_8 and C_{12} increased from -0.5 to



(a) O_2 insertion at C_8.

(b) O_2 insertion at C_{12}.

Figure 8.14: Potential energy profiles for the oxygen insertion at C_8 (a) and C_{12} (b).



(a) O_2 insertion at C_8.

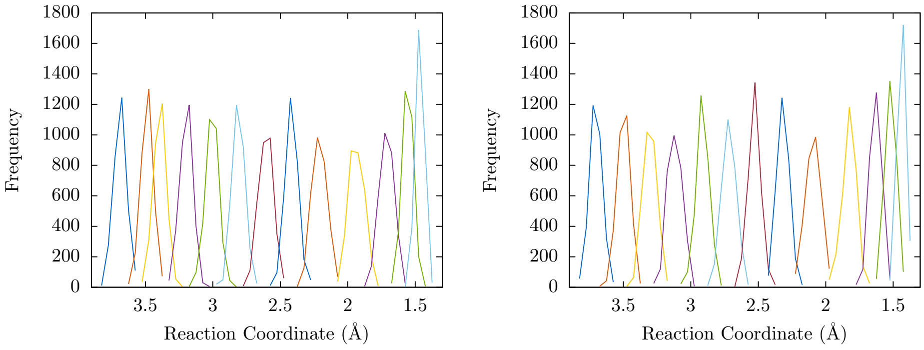
(b) O_2 insertion at C_{12}.

Figure 8.15: Evolution of spin density on atoms C_8 (red), C_9 (green), C_{10} (blue), C_{11} (yellow), C_{12} (cyan), O_A (pink) and O_B (gray) during the oxygen insertion at C_8 (a) and C_{12} (b).

0 a.u., and on the oxygen atoms decreased from 1 a.u. in both atoms to 0.7 a.u in the case of O_B, and 0.3 in O_A, which is the oxygen participating in the peroxide bond. Therefore, at the end of the process, the unpaired electron initially delocalized over the pentadienyl system delocalizes on the two oxygen atoms of the -OO· group, as the sum of their spin densities yields 1 a.u. (0.3 + 0.7).

8.3.2.3 Free energy profiles

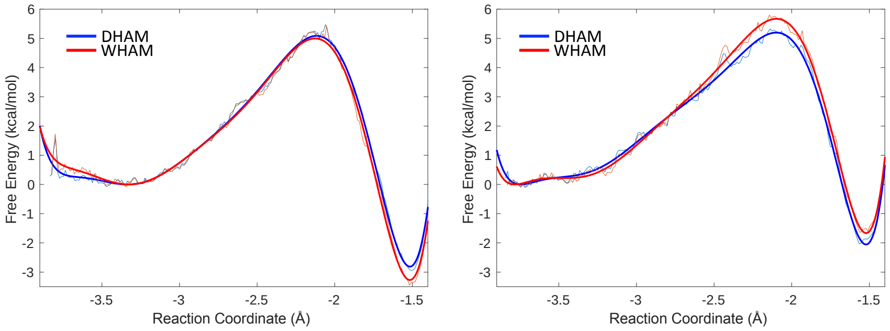
Following the same procedure as that for the hydrogen abstraction step, the free energy profiles have been computed for the oxygen insertion step using the US method. A total of 39 ps distributed in 13 umbrella windows were run by QM/MM MD simulations at the B3LYP-D3/6-31G(d,p) level and employing the small QM region (Figure 8.2b) defined for the reaction coordinate of the oxygen insertion step. To compute the free energy profiles both DHAM and WHAM methods were used. The histograms of the US windows are represented in Figure 8.16, showing a good overlapping.



(a) O_2 insertion at C_8.   **(b)** O_2 insertion at C_{12}.

Figure 8.16: Histograms of the 13 US windows of the oxygen insertion at C_8 (a) and C_{12} (b).

All the free energy curves have been fitted to a 9th degree polynomial function, and the reported free energy barriers have been calculated from the maximum (TS region) and minimum of the reactant valley of the fitted curve. The free energy barriers (Table 8.2) resulted in 5.1 and 5.2 kcal/mol for oxygen insertion at C_8 and C_{12}, respectively, by the WHAM method, but they turned out to be 5.0 and 5.7 kcal/mol for attack at C_8 and C_{12}, respectively, by the DHAM method. The difference of 0.7 kcal/mol (DHAM) is small when compared with the potential energy barriers, but is in agreement with the experimental regio- and stereospecificity trends of the second reaction step.



(a) O_2 insertion at C_8 **(b)** O_2 insertion at C_{12}

Figure 8.17: Free energy profiles for the oxygen insertion step at C_8 (a) and C_{12} (b) in the WT system.

Table 8.2: Summary of the potential energy ($\Delta E^{\ddagger}$) and free energy barriers calculated with WHAM and DHAM (all energies in kcal/mol) of the oxygen insertion at C_8 and C_{12} in the WT and the Gly427Ala mutant.

Reaction position	**WT**			**Mutant**		
	$\Delta E^{\ddagger}$	Free energy		$\Delta E^{\ddagger}$	Free energy	
		WHAM	DHAM		WHAM	DHAM
C_8	2.5	5.1	5.0	4.1	5.8	6.0
C_{12}	3.6	5.2	5.7	1.3	5.1	4.7

8.3.2.4 Discussion

During the approach of the oxygen molecule towards the target carbon atoms, protein residues in the closeness of the O_2 molecule and the carbon atoms have been evaluated. Residue Leu385, which is close to C_8 but even closer to C_{12}, seems to hinder the path of the oxygen towards C_{12}, favoring the attack at C_8. Leu385 is spatially close to Gly427, the residue proposed to control the stereochemistry of the reaction. Both residues are placed over the plane of the pentadienyl system . It is observed that there are no residues with an atom at less than 3 Å of C_8, but Leu385 is at less than this distance of C_{12}. A van dee Waals surface representation of the protein residues close to the pentadienyl system shows that there is something like a channel in the vicinity of C_8 (Figure 8.18). This channel, which is enclosed by Gly427 and Leu385 is connected with the region where the most stable initial oxygen positions (Figure 8.13) were obtained.

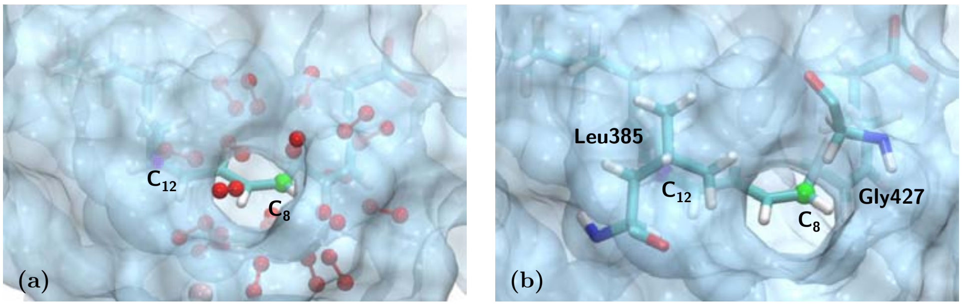


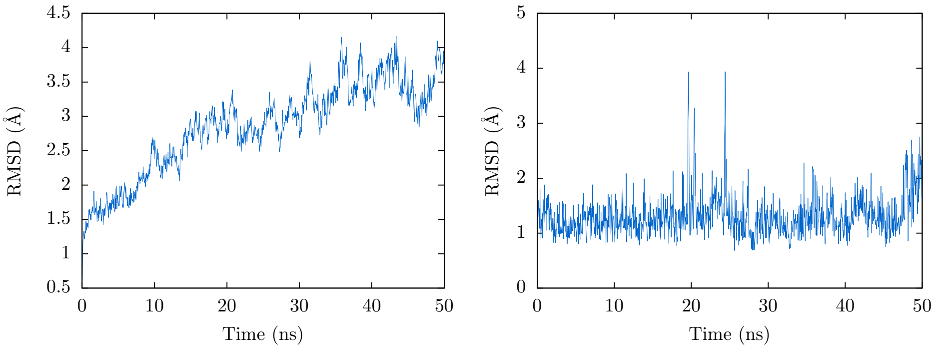
Figure 8.18: Van der Waals surface representation of protein residues close to the AA pentadienyl system (at < 5 Å). In (a), the representation has been superimposed with the initial position of the oxygen molecules. In (b), residues Leu385 and Gly427 are depicted. C_8 and C_{12} are highlighted by a green and violet sphere, respectively.

As can be seen in Figure 8.18, the channel proposed here is located closer to C_8 than to C_{12}, and this could explain the regioselectivity of the oxygen insertion step. Also, the position of Leu385, seems to be hindering the reaction path towards C_{12}.

8.3.3 Study of the Gly427Ala mutant

8.3.3.1 MD simulations

The mutant species Gly427Ala was prepared from the optimized structure of the hydrogen abstraction product of the WT system. For the mutant structure, some MD simulation steps have been performed in order to re-accommodate the AA radical in the new environment. To run the MD simulations the complete system (protein + waterbox) was recovered from the trimmed system employed in previous QM/MM calculations. The parameters for the radical AA were derived from a previous work using the charges obtained for the radical product in the corresponding QM/MM calculation.

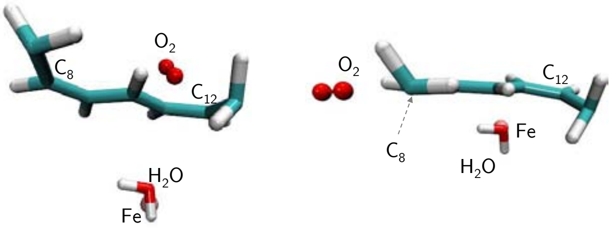


(a) Protein RMSD. **(b)** AA RMSD.

Figure 8.19: Protein and AA (radical) heavy atoms RMSDs during the 50-ns production step of the Gly427Ala mutant.

The protein and AA RMSDs corresponding to their heavy atoms, calculated with respect to the initial and average structures of the MD production trajectory, respectively, are represented in Figure 8.19. The protein RMSD increases up to 3.5 Å. AA RMSD stabilizes at 1.5 Å during all the production step, indicating that the fluctuations of the radical inside the binding cavity are small.

An structure from this MD simulation of the AA:Gly427Ala-8*R*-LOX system has been selected to perform the second reaction step. During the MD simulation process, a redistribution of the residues in contact with the mutated Ala427 takes place, leading to a new configuration of the active site cavity. To find the initial position of the oxygen molecule, the same procedure as that described in the previous section has been carried out. Thus, a total of 36 oxygen poses (18 at C_8 and 18 at C_{12}) have been generated. Single-point energy calculations led to the selection of the two oxygen positions shown in Figure 8.20. Like in the WT system, the most stable position to attack C_{12} in the mutant presents the oxygen molecule in an antarafacial configuration with respect the Fe(II)-OH_2 cofactor (Figure 8.20a)). The second position (Figure 8.20b)), prepared to react with C_8, presents an oxygen molecule almost coplanar to the pentadienyl radical moiety; however, as the oxygen approaches to C_8 in the reaction coordinate, it yields an antarafacial configuration (see below). Thus, these two positions have been selected as the starting point of the reaction coordinate for oxygen attack at C_{12} and C_8, respectively.

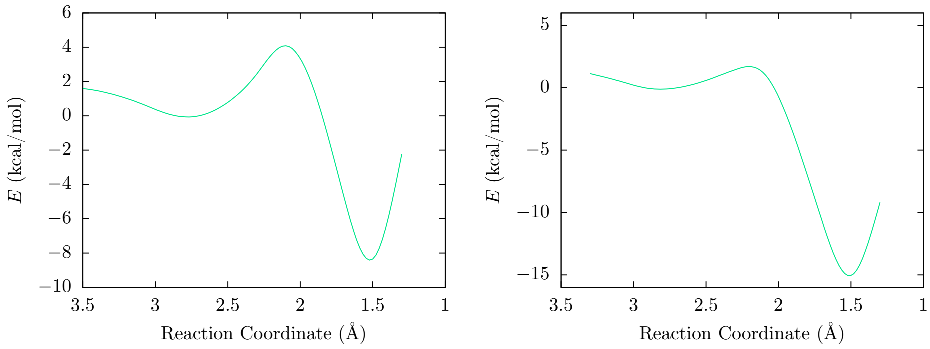


(a) **(b)**

Figure 8.20: Representation of the two most stable initial positions for the oxygen attack at (a) C_8 and (b) C_{12} in the Gly427Ala mutant.

8.3.3.2 Potential energy profiles

The reaction coordinate settings for the mutant have been the same as those for the WT system (Section 8.3.2.2). The potential energy profiles corresponding to the oxygen addition at C_8 and C_{12} of the two most stable positions are represented in Figure 8.21.



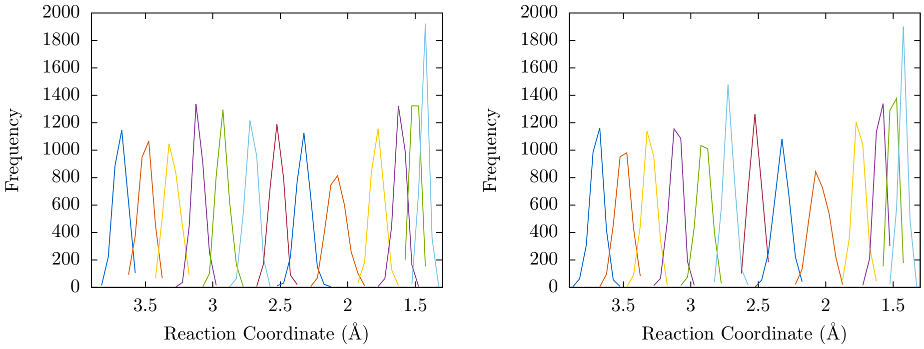
(a) O_2 insertion at C_8. **(b)** O_2 insertion at C_{12}.

Figure 8.21: Potential energy profiles for the oxygen insertion at C_8 (a) and C_{12} (b) in the mutant Gly427Ala.

Oxygen insertion in C_8 occurs with a potential energy barrier of approximately 4.1 kcal/mol and is exoergic by 8.4 kcal/mol. Reactants are located at $z_2 = 2.7$ Å, products at $z_2 = 1.5$ Å, and the TS at $z_2 = 2.1$ Å. In the case of C_{12}, the potential energy barrier is of 1.3 kcal/mol, and the reaction energy is of -7.1 kcal/mol. Reactants lie at $z_2 = 2.7$ Å, products at $z_2 = 1.5$ Å, and the TS at $z_2 = 2.1$ Å. Therefore, a dramatic inversion of the reaction specificity has been produced, being the oxygenation process at C_{12} favored over C_8 by 2.8 kcal/mol, in good agreement with experimental results.

8.3.3.3 Free energy profiles

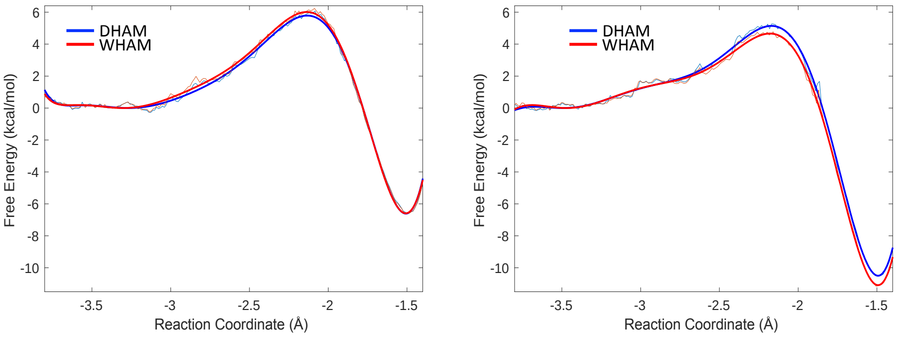
The free energy profiles for the oxygen insertion step in the Gly427Ala mutant were calculated with the same procedure as that for the WT system (Section 7.3.2.4). The US windows histograms are represented in Figure 8.22, and the free energy profiles are shown in Figure 8.23.



(a) O_2 insertion at C_8.   **(b)** O_2 insertion at C_{12}.

Figure 8.22: Histograms of the 13 US windows of the oxygen insertion at C_8 (a) and C_{12} (b) in the Gly427Ala mutant.

For the oxygen insertion at C_8 the free energy barrier results in 5.8 (WHAM) and 6.0 (DHAM) kcal/mol, whereas for the oxygen addition at C_{12}, the free energy barrier turns out to be 5.1 (WHAM) and 4.7 (DHAM) kcal/mol (see Table 8.2). Again, like in the case of the WT system, this difference in the free energy barrier for both processes (attack at C_8 and C_{12}) decreases with respect to the difference in the potential energy barriers (which is of 2.8 kcal/mol favoring C_{12}), being now of 1.3 kcal/mol (DHAM) favoring also the attack at C_{12}. In any case, the inversion of the reaction specificity in the mutant is also obtained with the free energy calculations.



(a) O_2 insertion at C_8 **(b)** O_2 insertion at C_{12}

Figure 8.23: Free energy profiles for the oxygen insertion step at C_8 (a) and C_{12} (b) in the Gly427Ala mutant.

8.3.3.4 Discussion

To analyze the details of the changes produced due to the introduction of a methyl moiety by the mutation of Gly427 into an Alanine, a first representation of the van der Waals protein surface around the AA substrate has been depicted in Figure 8.25. This representation corresponds to the starting structure used to generate the Gly427Ala mutant and prior to the MD simulation. For this reason, it presents the same geometry as the WT system studied in the previous section (except for the mutated residue).

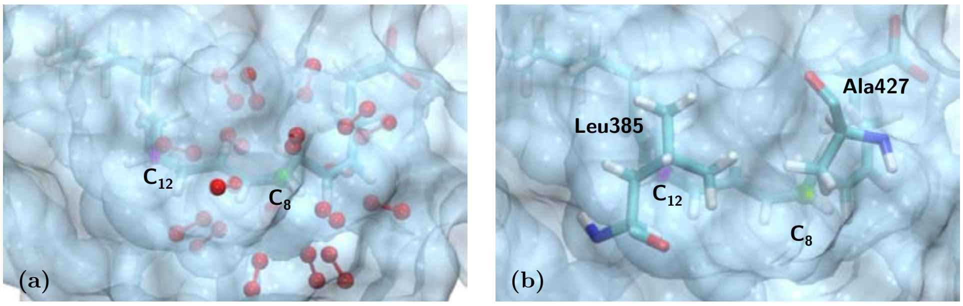


Figure 8.24: Van der waals surface representation of protein residues close to the AA pentadienyl system (at < 5 Å) in the starting structure of mutant Gly427Ala. In (a), the representation has been superimposed with the initial position of the oxygen molecules used in the WT system. In (b), residues Leu385 and Ala427 are depicted. C_8 and C_{12} are highlighted by a green and violet sphere, respectively.

During the MD simulation, the active site cavity reorganizes as a consequence of the mutation. This reorganization leads to a relocation of the oxygen access channel of the WT system (Figure 8.25), in which it was favoring the oxygen entrance towards C_8 (see Figure 8.18). Then, a new and different channel appears which is enclosed by Leu385, Ala427, Ile437 and

Val384. This channel is somehow directed towards C_{12}, which could explain the change in the reaction specificity. Also, the placement of Ala427 pushes Leu385 and affects its surroundings, causing the aperture of the new channel.

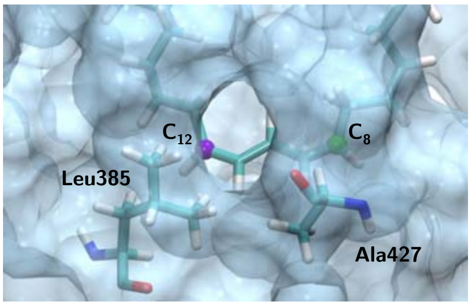


Figure 8.25: Van der Waals surface representation of protein residues close to the AA pentadienyl system (at < 5 Å) in the selected MD snapshot of mutant Gly427Ala used to the study of the oxygen addition step. Residues Leu385 and the mutated Ala427 are depicted. C_8 and C_{12} are highlighted by a green and violet sphere, respectively.

In an experimental work by Newcomer and co-workers[32] it has been proposed that is Leu431 rather than Leu385 the residue involved in the configuration of the oxygen access channel. In this crystallographic structure, this residue is indeed in a closer contact with Gly427 than its counterpart Leu385, and it is also closer to the AA substrate. However, in the work presented in this Chapter, Leu431 moves to a further position during the MD simulation, whereas Leu385 remains closer to the AA substrate (Figure 8.26). Both residues are conserved, so it seems that the presence of a Leucine in this key position could be involved in the regio- and stereo-control of the oxygenation process.

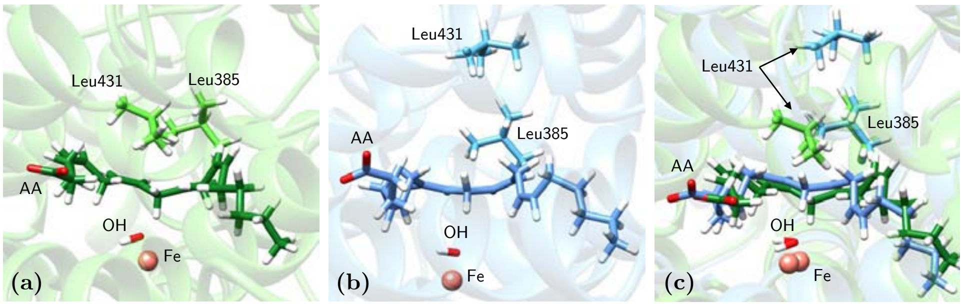


Figure 8.26: Representation of the crystal structure ((a), green) and the selected snapshot from the MD simulation of WT system prior to the hydrogen abstraction step ((b), blue), showing the arachidonic acid substrate and residues Leu431 and Leu385; (c) shows an overlay of (a) and (b).

8.4 Conclusions

In this Chapter a work combining MD simulations, QM/MM calculations, and umbrella sampling free energy simulations has been performed to study the reaction mechanism of the hydrogen abstraction process in WT coral 8*R*-LOX, and the oxygen insertion step in both WT and the *in silico* mutant Gly427Ala of 8*R*-LOX. The structure of 8*R*-LOX used here was taken from a crystal structure containing the AA substrate inside the active site cavity, which exhibits a tail-first orientation with its carboxylate end interacting with Arg182.

The MD simulations show that C_{10} is in average the closest carbon atom to the catalytic iron, which is in agreement with the regioselectivity of 8*R*-LOX. During the whole simulation, the substrate retains its tail-first orientation and its interaction with the Arg182 residue.

The potential energy profile for the hydrogen abstraction step of pro-*R* H_{10} occurs with a barrier of 15 kcal/mol. This barrier height is lower than in other LOX isoforms (see previous Chapters), which agrees with the experimental result of a higher catalytic rate constant ($k_{\text{cat}} = 210\ \text{s}^{-1}$) in 8*R*-LOX. The free energy barrier calculated by the DHAM method is of 16.0 kcal/mol, very close to the potential energy barrier, which suggest that the entropic and thermal contributions for this system are small. The calculations also provided the catalytic rate constant, and showed that the recrossing is negligible at the canonical variational transition state. The free energy barrier with the traditional WHAM method is larger than with DHAM, of 17.7 kcal/mol.

The second reaction step, the oxygen insertion, has been studied for the two possible positions starting from the H_{10} abstraction step: C_8 (which should be the most favored position in an 8-LOX), and C_{12}. The potential energy barriers for the two processes resulted in 2.5 kcal/mol and 3.6 kcal/mol, respectively. Thus, the oxygen addition is favored in C_8 by 1.1 kcal/mol, which agrees with the experimental regioselectivity. Free energy barriers resulted in 5.0 and 5.7 kcal/mol by the DHAM method, also showing that the addition at C_8 is more favored. Analyzing the configuration of the active site cavity it is observed a channel that is located over the pentadienyl moiety of the AA substrate, and closer to C_8 than to C_{12}. The channel is enclosed by Gly427 and Leu385. When the Gly427 is mutated to an Alanine, the configuration of the channel changes, being now redirectered towards C_{12}: Ala427 pushes Leu385 altering the configuration of the channel and blocking the region over C_8. Now the potential energy barriers result in 4.1 kcal/mol for C_8 addition, and 1.5 kcal/mol for insertion at C_{12}, and free energy barriers (DHAM) results now in 6.0 kcal/mol for C_8 and 4.7 kcal/mol for C_{12} additions, respectively, thus the enzyme regio- and stereospecificity is inverted. This result confirms the proposal of the Gly/Ala "switch" by Coffa and Brash.[16]

Finally, Newcomer *et al.*[32] proposed that was Leu431 the residue involved in the oxygen access channel. However, our results show that Leu431 moves away from the AA substrate whereas Leu385 remains closer. Interestingly, both Leu385 and Leu431 are highly conserved Leucines. The presence of such type of residues (bulky and hydrophobic), conserved among many LOX isoforms, could be key in controlling the regio- and stereoselectivity of the different steps of the reaction mechanism. The strategic position of the different Leucines (and Isoleucines) in the active site sterically hinders the substrate regions that do not lead to the final hydroperoxidation product with the desired regio- and stereospecificity.

Part V

Conclusions

Chapter 9

General conclusions

In the work presented in this Thesis, as required, homology modeling, docking and molecular dynamics simulations and quantum mechanics/molecular mechanics calculations of the potential and free energy along the reaction coordinate, have been employed to study the catalytic mechanism of three different LOXs: 15-LOX-1, 5-LOX and 8*R*-LOX. Special focus on the hydrogen abstraction step from the arachidonic acid substrate has been given to understand the subtle molecular factors that govern the high regio- and stereospecificty present in this reaction. The main conclusions can be summarized as follows:

1. There exists a wide dispersion in the potential energy barriers of the hydrogen abstraction process. The differences in the potential energy barriers arise from the difference in the reactant minima geometries rather than from the TSs ones, as it has been observed that all the TS structures obtained from the different hydrogen abstraction processes are alike.

2. The evolution of the spin densities has shown that the hydrogen abstraction step is actually a proton coupled electron transfer process, as the proton and the electron of the shifting hydrogen atom are transferred to different acceptors: the electron goes to the Fe(III) yielding Fe(II), and the proton to the OH^- group yielding a H_2O molecule.

3. Distances of the transferring hydrogen atom to the oxygen acceptor in the Fe(III)-OH^- cofactor are not the unique factor that determines the ease of the process. There is not a correlation between this geometric parameter and the barrier height: different structures with very similar H-O distance, can lead to very different potential energy barriers and vice versa.

4. How hindered is the geometrical evolution of the unsaturated fatty acid from the non-planar reactant structure to the planar product, dramatically affects the energy barrier in the hydrogen abstraction step. Bulky aliphatic residues are placed in strategic positions that hinder the reaction of the bisallylic methylene groups of the substrate that do not lead to the main product.

5. The reaction of 15-LOX-1 with LA is faster than the reaction with AA as indicated by the calculated potential/free energy barriers. Thus, LA is the preferential substrate for this isoform in accordance with experimental results. The two substrates can be accommodated into the 15-LOX-1 active site with their reactive bisallylic methylene groups in good position to initiate the reactive process.

6. The potential energy barrier for the hydrogen abstraction from AA differs among the different LOXs. In 8*R*-LOX, this process is faster than in 15-LOX-1 and 5-LOX, which agrees with the experimentally reported higher k_{cat} for 8*R*-LOX.

7. Different strategies that alter different regions of the active site cavity (cavity bottom, the region close to Fe(III)-OH$^-$) transforming bulky residues into smaller ones by means of *in silico* mutagenesis have shown to change the reaction specificity of 15-LOX-1 in terms of the hydrogen abstraction step. The new configuration of the active site cavity allows new conformations of the substrate that lead to products with different reaction specificity in agreement with *in vitro* mutagenesis experiments. Certain mutations can even change which is the step of the reaction mechanism defining the overall oxidation regioselectivity, usually assumed to be the hydrogen abstraction step.

8. The presence of bulky residues is also involved in the oxygen insertion step. *In silico* mutagenesis in 8*R*-LOX has shown that the Gly/Ala switch reported in the literature, that changes the regio- and stereospecificity of the oxygen insertion step, is connected with bulky Leucines surrounding the substrate pentadienyl radical. This Gly-Ala mutation changes the configuration of these residues and alters the oxygen access channel.

9. For most LOXs isoforms the AA substrate exhibits a tail-first orientation (15-LOX-1 and 8*R*-LOX), but for human 5-LOX, both tail- and head-first orientations could be in principle possible, and pre-reactive AA:5-LOX complexes are generated in those two orientations. So, the final productive structure of 5-LOX, is still unknown. This is probably due to the problems associated with its intrinsic flexibility, which not only made its crystallization process unfeasible (all the available

crystal structures correspond to mutants), but also complicates the determination of the binding modes.

In summary, the LOXs active site cavity contains different bulky aliphatic residues that affect substrate positioning and are also involved in the control of the reaction specificity of the different steps of the global reaction mechanism. By means of these residues a particular set of conformations of the substrate inside the active site that lead to the catalytically active binding modes are selected. Moreover, those bulky residues also hinder the non-desired reaction paths depending on the LOX isoform, so being the molecular basis of LOXs reaction specificity. For 15-LOX-1 and 8*R*-LOX it is clear which are the binding modes compatible with their catalytic properties, and we have been able to modify their reaction specificity by *in silico* mutagenesis. The same cannot be concluded for 5-LOX, for which three different binding modes were in principle possible, with one of them presenting predominant 5-lipoxygenating activity but, at the same time, being not stable, at least within the crystallographic holoenzyme structure used. So, the question of the catalytically productive binding mode in 5-LOX still remains open although our work provides a lot of molecular information that might be useful in future studies.

Bibliography

[1] A. R. Brash, "Lipoxygenases: Occurrence, functions, catalysis, and acquisition of substrate," *Journal of Biological Chemistry*, vol. 274, no. 34, pp. 23679–23682, 1999.

[2] H. Kühn, J. Saam, S. Eibach, H. G. Holzhutter, I. Ivanov, and M. Walther, "Structural biology of mammalian lipoxygenasese enzymatic consequences of targeted alterations of the protein structure," *Biochemical and Biophysical Research Communications*, vol. 338, no. 1, pp. 93–101, 2005.

[3] C. Schneider, D. A. Pratt, N. A. Porter, and A. R. Brash, "Control of oxygenation in lipoxygenase and cyclooxygenase catalysis," *Chemistry "&" Biology*, vol. 14, no. 5, pp. 473–488, 2007.

[4] E. Andre and K. W. Hou, "The presence of a lipid oxidase in soybean Glycine soya," *Acad. Sci. (Paris)*, vol. 194, pp. 645–647, 1932.

[5] M. Hamberg and S. .B, "Prostaglandin endoperoxides - novel transformations of arachidonic-acid in human platelets," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 71, no. 9, pp. 3400–3404, 1974.

[6] C. Funk, X. Chen, E. Johnson, and L. Zhao, "Lipoxygenase genes and their targeted disruption," *Prostaglandins & Other Lipid Mediators*, vol. 68-9, no. SI, pp. 303–312, 2002.

[7] A. Catalano and A. Procopio, "New aspects on the role of lipoxygenases in cancer progression," *Histology and Histopathology*, vol. 20, no. 3, pp. 969–975, 2005.

[8] A. D. Dobrian, D. C. Lieb, B. K. Cole, D. A. Taylor-Fishwick, S. K. Chakrabarti, and J. L. Nadler, "Functional and pathological roles of the 12- and 15-lipoxygenases," *Progress in Lipid Research*, vol. 50, no. 1, pp. 115–131, 2011.

[9] J. Z. Haeggström and C. D. Funk, "Lipoxygenase and Leukotriene Pathways: Biochemistry, Biology, and Roles in Disease," *Chemical Reviews*, vol. 111, no. 10, pp. 5866–5898, 2011.

[10] Y. C. Joo and D. K. Oh, "Lipoxygenases: Potential starting biocatalysts for the synthesis of signaling compounds," *Biotechnology Advances*, vol. 30, no. 6, pp. 1524–1532, 2012.

[11] S. Allmann, R. Halitschke, R. C. Schuurink, and I. T. Baldwin, "Oxylipin channelling in Nicotiana attenuata: lipoxygenase 2 supplies substrates for green leaf volatile production," *Plant Cell and Environment*, vol. 33, no. 12, pp. 2028–2040, 2010.

[12] G. von Merey, N. Veyrat, G. Mahuku, R. L. Valdez, T. C. J. Turlings, and M. D'Alessandro, "Dispensing synthetic green leaf volatiles in maize fields increases the release of sesquiterpenes by the plants, but has little effect on the attraction of pest and beneficial insects," *Phytochemistry*, vol. 72, no. 14-15, pp. 1838–1847, 2011.

[13] J. S. Yuan, S. J. Himanen, J. K. Holopainen, F. Chen, and C. N. Stewart, "Smelling global climate change: mitigation of function for plant volatile organic compounds," *Trends in Ecology & Evolution*, vol. 24, no. 6, pp. 323–331, 2009.

[14] Y. Q. Guo, C. L. Feng, H. L. Song, Z. Y. Wang, Q. Ren, and R. Wang, "Effect of pox3 gene disruption using self-cloning crf1 cassette in yarrowia lipolytica on the gamma-decalactone production," *World Journal of Microbiology & Biotechnology*, vol. 27, no. 12, pp. 2807–2812, 2011.

[15] F. C. Huang and W. Schwab, "Cloning and characterization of a 9-lipoxygenase gene induced by pathogen attack from nicotiana benthamiana for biotechnological application," *Bmc Biotechnology*, vol. 11, 2011.

[16] G. Coffa and A. R. Brash, "A single active site residue directs oxygenation stereospecificity in lipoxygenases: Stereocontrol is linked to the position of oxygenation," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 101, no. 44, pp. 15579–15584, 2004.

[17] I. Ivanov, D. Heydeck, K. Hofheinz, J. Roffeis, V. B. O'Donnell, H. Kühn, and M. Walther, "Molecular enzymology of lipoxygenases," *Archives of Biochemistry and Biophysics*, vol. 503, no. 2, pp. 161–174, 2010.

[18] K. Schwarz, M. Walther, M. Anton, C. Gerth, I. Feussner, and H. Kühn, "Structural basis for lipoxygenase specificity - conversion of the human leukocyte 5-lipoxygenase to a 15-lipoxygenating enzyme species by site-directed mutagenesis," *Journal of Biological Chemistry*, vol. 276, no. 1, pp. 773–779, 2001.

[19] R. Vogel, C. Jansen, J. Roffeis, P. Reddanna, P. Forsell, H. E. Claesson, H. Kühn, and M. Walther, "Applicability of the triad concept for the positional specificity of mammalian lipoxygenases," *Journal of Biological Chemistry*, vol. 285, no. 8, pp. 5369–5376, 2010.

[20] T. Schewe, W. Halangk, C. Hiebsch, and S. M. Rapoport, "Lipoxygenase in rabbit reticulocytes which attacks phospholipids and intact mitochondria," *Febs Letters*, vol. 60, no. 1, pp. 149–152, 1975.

[21] Y. Takahashi, W. C. Glasgow, H. Suzuki, Y. Taketani, S. Yamamoto, M. Anton, H. Kühn, and A. R. Brash, "Investigation of the oxygenation of phospholipids by the porcine leukocyte and human platelet arachidonate 12-lipoxygenases," *European Journal of Biochemistry*, vol. 218, no. 1, pp. 165–171, 1993.

[22] J. M. Upston, J. Neuzil, P. K. Witting, R. Alleva, and R. Stocker, "Oxidation of free fatty acids in low density lipoprotein by 15-lipoxygenase stimulates nonenzymic, alpha-tocopherol-mediated peroxidation of cholesteryl esters," *Journal of Biological Chemistry*, vol. 272, no. 48, pp. 30067–30074, 1997.

[23] T. Hammarberg, P. Provost, B. Persson, and O. Radmark, "The N-terminal domain of 5-lipoxygenase binds calcium and mediates calcium stimulation of enzyme activity," *Journal of Biological Chemistry*, vol. 275, no. 49, pp. 38787–38793, 2000.

[24] M. Hammel, M. Walther, R. Prassl, and H. Kuhn, "Structural flexibility of the N-terminal beta-barrel domain of 15-lipoxygenase-1 probed by small angle X-ray scattering. Functional consequences for activity regulation and membrane binding," *Journal of Molecular Biology*, vol. 343, no. 4, pp. 917–929, 2004.

[25] D. L. Sloane, M. F. Browner, Z. Dauter, K. Wilson, R. J. Fletterick, and E. Sigal, "Purification and crystallization of 15-lipoxygenase from rabbit reticulocytes," *Biochemical and Biophysical Research Communications*, vol. 173, no. 2, pp. 507–513, 1990.

[26] S. A. Gillmor, A. Villasenor, R. Fletterick, E. Sigal, and M. F. Browner, "The structure of mammalian 15-lipoxygenase reveals similarity to the lipases and the determinants of substrate specificity," *Nature Structural Biology*, vol. 4, no. 12, pp. 1003–1009, 1997.

[27] J. Choi, J. K. Chon, S. Kim, and W. Shin, "Conformational flexibility in mammalian 15S-lipoxygenase: Reinterpretation of the crystallographic data," *Proteins-Structure Function and Bioinformatics*, vol. 70, no. 3, pp. 1023–1032, 2008.

[28] L. Toledo, L. Masgrau, J. D. Maréchal, J. M. Lluch, and À. González-Lafont, "Insights into the mechanism of binding of arachidonic acid to mammalian 15-lipoxygenases," *Journal of Physical Chemistry B*, vol. 114, no. 20, pp. 7037–7046, 2010.

[29] N. C. Gilbert, S. G. Bartlett, M. T. Waight, D. B. Neau, W. E. Boeglin, A. R. Brash, and M. E. Newcomer, "The structure of human 5-lipoxygenase," *Science*, vol. 331, no. 6014, pp. 217–219, 2011.

[30] N. C. Gilbert, Z. Rui, D. B. Neau, M. T. Waight, S. G. Bartlett, W. E. Boeglin, A. R. Brash, and M. E. Newcomer, "Conversion of human 5-lipoxygenase to a 15-lipoxygenase by a point mutation to mimic phosphorylation at serine-663," *Faseb Journal*, vol. 26, no. 8, pp. 3222–3229, 2012.

[31] D. B. Neau, N. C. Gilbert, S. G. Bartlett, W. Boeglin, A. R. Brash, and M. E. Newcomer, "The 1.85 Å structure of an 8R-lipoxygenase suggests a general model for lipoxygenase product specificity," *Biochemistry*, vol. 48, no. 33, pp. 7906–7915, 2009.

[32] D. B. Neau, G. Bender, W. E. Boeglin, S. G. Bartlett, A. R. Brash, and M. E. Newcomer, "Crystal structure of a lipoxygenase in complex with substrate. The arachidonic acid-binding site of 8R-lipoxygenase," *Journal of Biological Chemistry*, vol. 289, no. 46, pp. 31905–31913, 2014.

[33] M. H. Glickman and J. P. Klinman, “Nature of Rate-Limiting Steps in the Soybean Lipoxygenase-1 Reaction,” *Biochemistry*, vol. 34, no. 43, pp. 14077–14092, 1995.

[34] E. Hatcher, A. V. Soudackov, and S. Hammes-Schiffer, “Proton-coupled electron transfer in soybean lipoxygenase,” *Journal of the American Chemical Society*, vol. 126, no. 18, pp. 5763–5775, 2004.

[35] N. Lehnert and E. I. Solomon, “Density-functional investigation on the mechanism of h-atom abstraction by lipoxygenase,” *Journal of Biological Inorganic Chemistry*, vol. 8, no. 3, pp. 294–305, 2003.

[36] K. W. Rickert and J. P. Klinman, “Nature of hydrogen transfer in soybean lipoxygenase 1: Separation of primary and secondary isotope effects,” *Biochemistry*, vol. 38, no. 38, pp. 12218–12228, 1999.

[37] R. L. Maas and A. R. Brash, “Evidence for a lipoxygenase mechanism in the biosynthesis of epoxide and dihydroxy leukotrienes from 15(s)-hydroperoxyicosatetraenoic acid by human-platelets and porcine leukocytes,” *Proceedings of the National Academy of Sciences of the United States of America-Biological Sciences*, vol. 80, no. 10, pp. 2884–2888, 1983.

[38] C. Jacquot, A. T. Wecksler, C. M. McGinley, E. N. Segraves, T. R. Holman, and W. A. van der Donk, “Isotope sensitive branching and kinetic isotope effects in the reaction of deuterated arachidonic acids with human 12- and 15-lipoxygenases,” *Biochemistry*, vol. 47, no. 27, pp. 7295–7303, 2008.

[39] A. T. Wecksler, C. Jacquot, W. A. van der Donk, and T. R. Holman, “Mechanistic investigations of human reticulocyte 15-and platelet 12-lipoxygenases with arachidonic acid,” *Biochemistry*, vol. 48, no. 26, pp. 6259–6267, 2009.

[40] S. Borngraber, M. Browner, S. Gillmor, C. Gerth, M. Anton, R. Fletterick, and H. Kühn, “Shape and specificity in mammalian 15-lipoxygenase active site - The functional, interplay of sequence determinants for the reaction specificity,” *Journal of Biological Chemistry*, vol. 274, no. 52, pp. 37345–37350, 1999.

[41] M. E. Newcomer and A. R. Brash, “The structural basis for specificity in lipoxygenase catalysis,” *Protein Science*, vol. 24, no. 3, pp. 298–309, 2015.

[42] I. Ivanov, H. Kühn, and D. Heydeck, “Structural and functional biology of arachidonic acid 15-lipoxygenase-1 (alox15),” *Gene*, vol. 573, no. 1, pp. 1–32, 2015.

[43] H. Kühn, “Structural basis for the positional specificity of lipoxygenases,” *Prostaglandins & Other Lipid Mediators*, vol. 62, no. 3, pp. 255–270, 2000.

[44] P. Gund and T. Y. Shen, “Model for prostaglandin synthetase cyclooxygenation site and its inhibition by antiinflammatory arylacetic acids,” *Journal of Medicinal Chemistry*, vol. 20, no. 9, pp. 1146–1152, 1977.

[45] Q. F. Gan, M. F. Browner, D. L. Sloane, and E. Sigal, “Defining the arachidonic acid binding site of human 15-lipoxygenase - Molecular modeling and mutagenesis,” *Journal of Biological Chemistry*, vol. 271, no. 41, pp. 25412–25418, 1996.

[46] C. Jansen, K. Hofheinz, R. Vogel, J. Roffeis, M. Anton, P. Reddanna, H. Kuhn, and M. Walther, “Stereocontrol of Arachidonic Acid Oxygenation by Vertebrate Lipoxygenases-Newly Cloned Zebrafish Lipoxygenase1 does not Follow the Ala-versus-Gly Concept,” *Journal of Biological Chemistry*, vol. 286, no. 43, pp. 37804–37812, 2011.

[47] S. T. Prigge, B. J. Gaffney, and L. M. Amzel, “Relation between positional specificity and chirality in mammalian lipoxygenases,” *Nature Structural Biology*, vol. 5, no. 3, pp. 178–179, 1998.

[48] M. F. Browner, S. A. Gillmor, and R. Fletterick, “Burying a charge,” *Nature Structural Biology*, vol. 5, no. 3, pp. 179–179, 1998.

[49] S. Mitra, S. G. Bartlett, and M. E. Newcomer, “Identification of the substrate access portal of 5-lipoxygenase,” *Biochemistry*, vol. 54, no. 41, pp. 6333–6342, 2015.

[50] K. Hofheinz, K. R. Kakularam, S. Adel, M. Anton, A. Polymarasetty, P. Reddanna, H. Kühn, and T. Horn, “Conversion of pro-inflammatory murine alox5 into an anti-inflammatory 15s-lipoxygenating enzyme by multiple mutations of sequence determinants,” *Archives of Biochemistry and Biophysics*, vol. 530, no. 1, pp. 40–47, 2013.

[51] S. Adel, D. Heydeck, H. Kühn, and C. Ufer, “The lipoxygenase pathway in zebrafish. expression and characterization of zebrafish ALOX5 and comparison with its human ortholog,” *Biochimica et Biophysica Acta (BBA) - Molecular and Cell Biology of Lipids*, vol. 1861, no. 1, pp. 1 – 11, 2016.

[52] M. Jisaka, R. B. Kim, W. E. Boeglin, and A. R. Brash, “Identification of amino acid determinants of the positional specificity of mouse 8s-lipoxygenase and human 15s-lipoxygenase-2,” *Journal of Biological Chemistry*, vol. 275, no. 2, pp. 1287–1293, 2000.

[53] G. Coffa, C. Schneider, and A. R. Brash, “A comprehensive model of positional and stereo control in lipoxygenases,” *Biochemical and Biophysical Research Communications*, vol. 338, no. 1, pp. 87–92, 2005.

[54] W. E. Boeglin, A. Itoh, Y. X. Zheng, G. Coffa, G. A. Howe, and A. R. Brash, “Investigation of Substrate Binding and Product Stereochemistry Issues in Two Linoleate 9-Lipoxygenases,” *Lipids*, vol. 43, no. 11, pp. 979–987, 2008.

[55] G. Coffa, A. N. Imber, B. C. Maguire, G. Laxmikanthan, C. Schneider, B. J. Gaffney, and A. R. Brash, “On the relationships of substrate orientation, hydrogen abstraction, and product stereochemistry in single and double dioxygenations by soybean lipoxygenase-1 and its Ala542Gly mutant,” *Journal of Biological Chemistry*, vol. 280, no. 46, pp. 38756–38766, 2005.

[56] B. Samuelsson, P. Borgeat, S. Hammarstrom, and R. Murphy, “Introduction of a nomenclature - Leukotrienes,” *Prostaglandins*, vol. 17, no. 6, pp. 785–787, 1979.

[57] J. Haeggström and A. Wetterholm, “Enzymes and receptors in the leukotriene cascade,” *Cellular and Molecular Life Sciences*, vol. 59, no. 5, pp. 742–753, 2002.

[58] H. Kühn, S. Banthiya, and K. van Leyen, “Mammalian lipoxygenases and their biological relevance,” *Biochimica et Biophysica Acta - Molecular and Cell Biology of Lipids*, vol. 1851, no. 4, SI, pp. 308–330, 2015.

[59] C. Serhan, M. Hamberg, and B. Samuelsson, “Lipoxins - Novel series of biologically-active compounds formed from arachidonic-acid in human-leukotrienes,” *Proceedings of the National Academy of Sciences of the United States of America - Biological Sciences*, vol. 81, no. 17, pp. 5335–5339, 1984.

[60] G. Folco and R. C. Murphy, “Eicosanoid transcellular biosynthesis: From cell-cell interactions to in vivo tissue responses,” *Pharmacological Reviews*, vol. 58, no. 3, pp. 375–388, 2006.

[61] M. Spite, J. Claria, and C. N. Serhan, “Resolvins, specialized proresolving lipid mediators, and their potential roles in metabolic diseases,” *Cell Metabolism*, vol. 19, no. 1, pp. 21–36, 2014.

[62] J. Belkner, R. Wiesner, H. Kühn, and V. Lankin, “The oxygenation of cholesterol esters by the reticulocyte lipoxygenase,” *FEBS Letters*, vol. 279, no. 1, pp. 110–114, 1991.

[63] C. N. Serhan, J. Dalli, S. Karamnov, A. Choi, C.-K. Park, Z.-Z. Xu, R.-R. Ji, M. Zhu, and N. A. Petasis, “Macrophage proresolving mediator maresin 1 stimulates tissue regeneration and controls pain,” *FASEB Journal*, vol. 26, no. 4, pp. 1755–1765, 2012.

[64] G. A. Howe and G. Jander, *Plant immunity to insect herbivores*, vol. 59 of *Annual Review of Plant Biology*, pp. 41–66. Palo Alto: Annual Reviews, 2008.

[65] M. Buchhaupt, J. C. Guder, M. M. W. Etschmann, and J. Schrader, “Synthesis of green note aroma compounds by biotransformation of fatty acids using yeast cells coexpressing lipoxygenase and hydroperoxide lyase,” *Applied Microbiology and Biotechnology*, vol. 93, no. 1, pp. 159–168, 2012.

[66] L. Zhao and C. Funk, “Lipoxygenase pathways in atherogenesis,” *Trends in Cardiovascular Medicine*, vol. 14, no. 5, pp. 191–195, 2004.

[67] J. Wittwer and M. Hersberger, “The two faces of the 15-lipoxygenase in atherosclerosis,” *Prostaglandins Leukotrienes and Essential Fatty Acids*, vol. 77, no. 2, pp. 67–77, 2007.

[68] D. Poeckel and C. D. Funk, “The 5-lipoxygenase/leukotriene pathway in preclinical models of cardiovascular disease,” *Cardiovascular Research*, vol. 86, no. 2, SI, 2010.

[69] S. J. A. Wuest, T. Horn, J. Marti-Jaun, H. Kühn, and M. Hersberger, “Association of polymorphisms in the ALOX15B gene with coronary artery disease,” *Clinical Biochemistry*, vol. 47, no. 6, 2014.

[70] Y. Chawengsub, K. M. Gauthier, and W. B. Campbell, “Role of arachidonic acid lipoxygenase metabolites in the regulation of vascular tone,” *American Journal of Physiology-Heart and Circulatory Physiology*, vol. 297, no. 2, pp. H495–H507, 2009.

[71] D. Wang and R. N. Dubois, “Eicosanoids and cancer,” *Nature Reviews Cancer*, vol. 10, no. 3, pp. 181–193, 2010.

[72] A. Ryan and C. Godson, “Lipoxins: regulators of resolution,” *Current Opinion in Pharmacology*, vol. 10, no. 2, pp. 166–172, 2010.

[73] H. Kamitani, M. Geller, and T. Eling, “Expression of 15-lipoxygenase by human colorectal carcinoma caco-2 cells during apoptosis and cell differentiation,” *Journal of Biological Chemistry*, vol. 273, no. 34, pp. 21569–21577, 1998.

[74] N. Reddy, A. Everhart, T. Eling, and W. Glasgow, “Characterization of a 15-lipoxygenase in human breast carcinoma BT-20 cells: Stimulation of 13-HODE formation by TGF(alpha)/EGF,” *Biochemical and Biophysical Research Communications*, vol. 231, no. 1, pp. 111–116, 1997.

[75] S. A. Spindler, F. H. Sarkar, W. A. Sakr, M. L. Blackburn, A. W. Bull, M. LaGattuta, and R. G. Reddy, “Production of 13-hydroxyoctadecadienoic acid (13-hode) by prostate tumors and cell lines,” *Biochemical and Biophysical Research Communications*, vol. 239, no. 3, pp. 775–781, 1997.

[76] R. DuBois, “Leukotriene A(4) signaling, inflammation, and cancer,” *Journal of the National Cancer Institute*, vol. 95, no. 14, pp. 1028–1029, 2003.

[77] D. Nie, G. Hillman, T. Geddes, K. Tang, C. Pierson, D. Grignon, and K. Honn, “Platelet-type 12-lipoxygenase in a human prostate carcinoma stimulates angiogenesis and tumor growth,” *Cancer Research*, vol. 58, no. 18, pp. 4047–4051, 1998.

[78] I. Shureiqi, W. Jiang, X. Zuo, Y. Wu, J. Stimmel, L. Leesnitzer, J. Morris, H. Fan, S. Fischer, and S. Lippman, “The 15-lipoxygenase-1 product 13-S-hydroxyoctadecadienoic acid down-regulates PPAR-delta to induce apoptosis in colorectal cancer cells,” *Proceedings of the National Academy of Sciences of the United States of America*, vol. 100, no. 17, pp. 9968–9973, 2003.

[79] R. K. Singh, R. Tandon, S. G. Dastidar, and A. Ray, “A review on leukotrienes and their receptors with reference to asthma,” *Journal of Asthma*, vol. 50, no. 9, pp. 922–931, 2013.

[80] B. Cobanoglu, E. Toskala, A. Ural, and C. Cingi, “Role of Leukotriene Antagonists and Antihistamines in the Treatment of Allergic Rhinitis,” *Current Allergy and Asthma Reports*, vol. 13, no. 2, pp. 203–208, 2013.

[81] A. Leff, “Role of leukotrienes in bronchial hyperresponsiveness and cellular responses in airways,” *American Journal of Respiratory and Critical Care Medicine*, vol. 161, no. 2, S, pp. S125–S132, 2000.

[82] G. A. Czapski, K. Czubowicz, J. B. Strosznajder, and R. P. Strosznajder, “The Lipoxygenases: Their Regulation and Implication in Alzheimer’s Disease,” *Neurochemical Research*, vol. 41, no. 1-2, SI, pp. 243–257, 2016.

[83] D. Pratico, V. Zhukareva, Y. Yao, K. Uryu, C. Funk, J. Lawson, J. Trojanowski, and V. Lee, “12/15-Lipoxygenase is increased in Alzheimer’s disease - Possible involvement in brain oxidative stress,” *American Journal of Pathology*, vol. 164, no. 5, pp. 1655–1662, 2004.

[84] B. Hofmann and D. Steinhilber, “5-Lipoxygenase inhibitors: a review of recent patents (2010-2012),” *Expert Opinion on Therapeutic Patents*, vol. 23, no. 7, pp. 895–909, 2013.

[85] H. Sadeghian and A. Jabbari, "15-Lipoxygenase inhibitors: a patent review," *Expert Opinion on Therapeutic Patents*, vol. 26, no. 1, pp. 65–88, 2016.

[86] R. McMillan and E. Walker, "Design of therapeutically effective 5-lipoxygenase inhibitors," *Trends in Pharmacological Sciences*, vol. 13, no. 8, pp. 323–330, 1992.

[87] A. Rossi, C. Pergola, A. Koeberle, M. Hoffmann, F. Dehm, P. Bramanti, S. Cuzzocrea, O. Werz, and L. Sautebin, "The 5-lipoxygenase inhibitor, zileuton, suppresses prostaglandin biosynthesis by inhibition of arachidonic acid release in macrophages," *British Journal of Pharmacology*, vol. 161, no. 3, pp. 555–570, 2010.

[88] B. R. Brooks, C. L. Brooks, A. D. Mackerell, L. Nilsson, R. J. Petrella, B. Roux, Y. Won, G. Archontis, C. Bartels, S. Boresch, A. Caflisch, L. Caves, Q. Cui, A. R. Dinner, M. Feig, S. Fischer, J. Gao, M. Hodoscek, W. Im, K. Kuczera, T. Lazaridis, J. Ma, V. Ovchinnikov, E. Paci, R. W. Pastor, C. B. Post, J. Z. Pu, M. Schaefer, B. Tidor, R. M. Venable, H. L. Woodcock, X. Wu, W. Yang, D. M. York, and M. Karplus, "CHARMM: The biomolecular simulation program," *Journal of Computational Chemistry*, vol. 30, no. 10, pp. 1545–1614, 2009.

[89] W. D. Cornell, P. Cieplak, C. I. Bayly, I. R. Gould, K. M. Merz, D. M. Ferguson, D. C. Spellmeyer, T. Fox, J. W. Caldwell, and P. A. Kollman, "A 2nd generation force-field for the simulation of proteins, nucleic-acids, and organic molecules," *Journal of the American Chemical Society*, vol. 117, no. 19, pp. 5179–5197, 1995.

[90] W. R. P. Scott, P. H. Hunenberger, I. G. Tironi, A. E. Mark, S. R. Billeter, J. Fennen, A. E. Torda, T. Huber, P. Kruger, and W. F. van Gunsteren, "The gromos biomolecular simulation program package," *Journal of Physical Chemistry A*, vol. 103, no. 19, pp. 3596–3607, 1999.

[91] P. Hohenberg and W. Kohn, "Inhomogeneous electron gas," *Physical review*, vol. 136, no. 3B, p. B864, 1964.

[92] W. Kohn and L. J. Sham, "Self-consistent equations including exchange and correlation effects," *Physical review*, vol. 140, no. 4A, p. A1133, 1965.

[93] S. F. Sousa, P. A. Fernandes, and M. J. Ramos, "Computational enzymatic catalysis - clarifying enzymatic mechanisms with the help of computers," *Phys. Chem. Chem. Phys.*, vol. 14, pp. 12431–12441, 2012.

[94] A. Warshel and M. Karplus, "Calculation of ground and excited state potential surfaces of conjugated molecules. i. formulation and parametrization," *Journal of the American Chemical Society*, vol. 94, no. 16, pp. 5612–5625, 1972.

[95] A. Warshel and M. Levitt, "Theoretical studies of enzymic reactions - Dielectric, electrostatic and steric stabilization of carbonium-ion in reaction of lysozyme," *Journal of Molecular Biology*, vol. 103, no. 2, pp. 227–249, 1976.

[96] M. J. Field, P. A. Bash, and M. Karplus, "A combined quantum-mechanical and molecular mechanical potential for molecular-dynamics simulations," *Journal of Computational Chemistry*, vol. 11, no. 6, pp. 700–733, 1990.

[97] H. M. Senn and W. Thiel, “Qm/mm studies of enzymes,” *Current Opinion in Chemical Biology*, vol. 11, no. 2, pp. 182–187, 2007.

[98] D. Bakowies and W. Thiel, “Hybrid Models for Combined Quantum Mechanical and Molecular Mechanical Approaches,” *The Journal of Physical Chemistry*, vol. 100, no. 25, pp. 10580–10594, 1996.

[99] I. Antes and W. Thiel, “Adjusted connection atoms for combined quantum mechanical and molecular mechanical methods,” *Journal of Physical Chemistry A*, vol. 103, no. 46, pp. 9290–9295, 1999.

[100] K. Nam, J. Gao, and D. M. York, “An efficient linear-scaling ewald method for long-range electrostatic interactions in combined qm/mm calculations,” *Journal of Chemical Theory and Computation*, vol. 1, no. 1, pp. 2–13, 2005.

[101] P. Schaefer, D. Riccardi, and Q. Cui, “Reliable treatment of electrostatics in combined qm/mm simulation of macromolecules,” *The Journal of Chemical Physics*, vol. 123, no. 1, 2005.

[102] U. C. Singh and P. A. Kollman, “A combined abinitio quantum-mechanical and molecular mechanical method for carrying out simulations on complex molecular-systems - applications to the ch3cl + cl- exchange-reaction and gas-phase protonation of polyethers,” *Journal of Computational Chemistry*, vol. 7, no. 6, pp. 718–730, 1986.

[103] F. Maseras and K. Morokuma, “Imomm - a new integrated ab-initio plus molecular mechanics geometry optimization scheme of equilibrium structures and transition-states,” *Journal of Computational Chemistry*, vol. 16, no. 9, pp. 1170–1179, 1995.

[104] P. Sherwood, A. H. de Vries, S. J. Collins, S. P. Greatbanks, N. A. Burton, M. A. Vincent, and I. H. Hillier, “Computer simulation of zeolite structure and reactivity using embedded cluster methods,” *Faraday Discussions*, vol. 106, pp. 79–92, 1997.

[105] P. Sherwood, A. H. de Vries, M. F. Guest, G. Schreckenbach, C. R. A. Catlow, S. A. French, A. A. Sokol, S. T. Bromley, W. Thiel, A. J. Turner, S. Billeter, F. Terstegen, S. Thiel, J. Kendrick, S. C. Rogers, J. Casci, M. Watson, F. King, E. Karlsen, M. Sjovoll, A. Fahmi, A. Schafer, and C. Lennartz, “QUASI: A general purpose implementation of the qm/mm approach and its application to problems in catalysis,” *Journal of Molecular Structure-Theochem*, vol. 632, pp. 1–28, 2003.

[106] R. B. Murphy, D. M. Philipp, and R. A. Friesner, “Frozen orbital qm/mm methods for density functional theory,” *Chemical Physics Letters*, vol. 321, no. 1-2, pp. 113–120, 2000.

[107] D. M. Philipp and R. A. Friesner, “Mixed ab initio qm/mm modeling using frozen orbitals and tests with alanine dipeptide and tetrapeptide,” *Journal of Computational Chemistry*, vol. 20, no. 14, pp. 1468–1494, 1999.

[108] M. Garcia-Viloca and J. L. Gao, “Generalized hybrid orbital for the treatment of boundary atoms in combined quantum mechanical and molecular mechanical calculations using the semiempirical parameterized model 3 method,” *Theoretical Chemistry Accounts*, vol. 111, no. 2-6, pp. 280–286, 2004.

[109] J. L. Gao, P. Amara, C. Alhambra, and M. J. Field, “A generalized hybrid orbital (gho) method for the treatment of boundary atoms in combined qm/mm calculations,” *Journal of Physical Chemistry A*, vol. 102, no. 24, pp. 4714–4721, 1998.

[110] R. Zwanzig, “High-temperature equation of state by a perturbation method .1. Nonpolar gases,” *Journal of Chemical Physics*, vol. 22, no. 8, pp. 1420–1426, 1954.

[111] G. Torrie and J. Valleau, “Monte-carlo free-energy estimates using non-boltzmann sampling - application to subcritical lennard-jones fluid,” *Chemical Physics Letters*, vol. 28, no. 4, pp. 578–581, 1974.

[112] G. Torrie and J. Valleau, “Non-physical sampling distributions in monte-carlo free-energy estimation - umbrella sampling,” *Journal of Computational Physics*, vol. 23, no. 2, pp. 187–199, 1977.

[113] Y. Zhang, H. Liu, and W. Yang, “Free energy calculation on enzyme reactions with an efficient iterative procedure to determine minimum energy paths on a combined ab initio QM/MM potential energy surface,” *Journal of Chemical Physics*, vol. 112, no. 8, pp. 3483–3492, 2000.

[114] J. Kastner, H. Senn, S. Thiel, N. Otte, and W. Thiel, “QM/MM free-energy perturbation compared to thermodynamic integration and umbrella sampling: Application to an enzymatic reaction,” *Journal of Chemical Theory and Computation*, vol. 2, no. 2, pp. 452–461, 2006.

[115] H. M. Senn, J. Kaestner, J. Breidung, and W. Thiel, “Finite-temperature effects in enzymatic reactions - Insights from QM/MM free-energy simulations,” *Canadian Journal of Chemistry*, vol. 87, no. 10, pp. 1322–1337, 2009.

[116] S. Kumar, J. M. Rosenberg, D. Bouzida, R. H. Swendsen, and P. A. Kollman, “The weighted histogram analysis method for free-energy calculations on biomolecules. I. The method,” *Journal of Computational Chemistry*, vol. 13, no. 8, pp. 1011–1021, 1992.

[117] E. Rosta and G. Hummer, “Free energies from dynamic weighted histogram analysis using unbiased markov state model,” *Journal of Chemical Theory and Computation*, vol. 11, no. 1, pp. 276–285, 2015.

[118] B. Alder and T. Wainwright, “Phase transition for a hard sphere system,” *Journal of Chemical Physics*, vol. 27, no. 5, pp. 1208–1209, 1957.

[119] J. McCammon, B. Gelin, and M. Karplus, “Dynamics of folded proteins,” *Nature*, vol. 267, no. 5612, pp. 585–590, 1977.

[120] J. E. Stone, D. J. Hardy, I. S. Ufimtsev, and K. Schulten, “GPU-accelerated molecular modeling coming of age,” *Journal of Molecular Graphics and Modelling*, vol. 29, no. 2, pp. 116 – 125, 2010.

[121] “GPU accelerated NAMD: NAMD benckmark datasets.” http://www.nvidia.com/object/gpu-accelerated-applications-namd-benchmarks.html.

[122] L. Verlet, “Computer "experiments" on classical fluids. i. thermodynamical properties of lennard-jones molecules,” *PhysPhysical Review*, vol. 159, pp. 98–103, 1967.

[123] W. C. Swope, H. C. Andersen, P. H. Berens, and K. R. Wilson, "A computer simulation method for the calculation of equilibrium constants for the formation of physical clusters of molecules: Application to small water clusters," *The Journal of Chemical Physics*, vol. 76, no. 1, pp. 637–649, 1982.

[124] R. W. Hockney, ""the potential calculation and some applications"," *Methods in Computational Physics*, vol. 9, pp. 135–211, 1970.

[125] J. C. Phillips, R. Braun, W. Wang, J. Gumbart, E. Tajkhorshid, E. Villa, C. Chipot, R. D. Skeel, L. KalÃĭ, and K. Schulten, "Scalable molecular dynamics with NAMD," *Journal of Computational Chemistry*, vol. 26, no. 16, pp. 1781–1802, 2005.

[126] J. P. Ryckaert, G. Ciccotti, and H. J. C. Berendsen, "Numerical-integration of cartesian equations of motion of a system with constraints - molecular-dynamics of n-alkanes," *Journal of Computational Physics*, vol. 23, no. 3, pp. 327–341, 1977.

[127] W. G. Hoover, "Canonical dynamics - equilibrium phase-space distributions," *Physical Review A*, vol. 31, no. 3, pp. 1695–1697, 1985.

[128] S. Nosé, "A unified formulation of the constant temperature molecular dynamics methods," *The Journal of chemical physics*, vol. 81, no. 1, pp. 511–519, 1984.

[129] H. J. Berendsen, J. v. Postma, W. F. van Gunsteren, A. DiNola, and J. Haak, "Molecular dynamics with coupling to an external bath," *The Journal of chemical physics*, vol. 81, no. 8, pp. 3684–3690, 1984.

[130] S. E. Feller, R. W. Pastor, A. Rojnuckarin, S. Bogusz, and B. R. Brooks, "Effect of electrostatic force truncation on interfacial and transport properties of water," *The Journal of Physical Chemistry*, vol. 100, no. 42, pp. 17011–17020, 1996.

[131] U. Essmann, L. Perera, M. L. Berkowitz, T. Darden, H. Lee, and L. G. Pedersen, "A smooth particle mesh ewald method," *The Journal of chemical physics*, vol. 103, no. 19, pp. 8577–8593, 1995.

[132] D. Hai-You, J. Ya, and Z. Yang, "Protein structure prediction," *Acta Physica Sinica*, vol. 65, no. 17, 2016.

[133] E. Krieger, S. B. Nabuurs, and G. Vriend, "Homology modeling," *Methods of biochemical analysis*, vol. 44, pp. 509–524, 2003.

[134] C. Chothia and A. M. Lesk, "The relation between the divergence of sequence and structure in proteins.," *The EMBO journal*, vol. 5, no. 4, p. 823, 1986.

[135] C. Sander and R. Schneider, "Database of homology-derived protein structures and the structural meaning of sequence alignment," *Proteins: Structure, Function, and Bioinformatics*, vol. 9, no. 1, pp. 56–68, 1991.

[136] C. J. Epstein, R. F. Goldberger, and C. B. Anfinsen, "The genetic control of tertiary protein structure: studies with model systems," in *Cold Spring Harbor symposia on quantitative biology*, vol. 28, pp. 439–449, 1963.

[137] B. Rost, "Twilight zone of protein sequence alignments," *Protein engineering*, vol. 12, no. 2, pp. 85–94, 1999.

[138] A. Fiser, R. K. G. Do, and A. Šali, “Modeling of loops in protein structures,” *Protein science*, vol. 9, no. 09, pp. 1753–1773, 2000.

[139] R. Sánchez and A. Sali, “Large-scale protein structure modeling of the saccharomyces cerevisiae genome,” *Proceedings of the National Academy of Sciences*, vol. 95, no. 23, pp. 13597–13602, 1998.

[140] S. F. Sousa, P. A. Fernandes, and M. J. Ramos, “ProteinâĂŞligand docking: Current status and future challenges,” *Proteins: Structure, Function, and Bioinformatics*, vol. 65, no. 1, pp. 15–26, 2006.

[141] N. Brooijmans and I. D. Kuntz, “Molecular recognition and docking algorithms,” *Annual Review of Biophysics and Biomolecular Structure*, vol. 32, no. 1, pp. 335–373, 2003.

[142] S.-Y. Huang and X. Zou, “Advances and challenges in protein-ligand docking,” *International Journal of Molecular Sciences*, vol. 11, no. 8, p. 3016, 2010.

[143] T. J. Ewing, S. Makino, A. G. Skillman, and I. D. Kuntz, “Dock 4.0: Search strategies for automated molecular docking of flexible molecule databases,” *Journal of Computer-Aided Molecular Design*, vol. 15, no. 5, pp. 411–428, 2001.

[144] C. A. Baxter, C. W. Murray, D. E. Clark, D. R. Westhead, and M. D. Eldridge, “Flexible docking using tabu search and an empirical estimate of binding affinity,” *Proteins: Structure, Function, and Bioinformatics*, vol. 33, no. 3, pp. 367–382, 1998.

[145] S. Janson, D. Merkle, and M. Middendorf, “Molecular docking with multi-objective particle swarm optimization,” *Applied Soft Computing*, vol. 8, no. 1, pp. 666 – 675, 2008.

[146] G. Wu, D. H. Robertson, C. L. Brooks, and M. Vieth, “Detailed analysis of grid-based molecular docking: A case study of CDOCKER-A CHARMm-based MD docking algorithm,” *Journal of Computational Chemistry*, vol. 24, no. 13, pp. 1549–1562, 2003.

[147] D. E. Clark and D. R. Westhead, “Evolutionary algorithms in computer-aided molecular design,” *Journal of Computer-Aided Molecular Design*, vol. 10, no. 4, pp. 337–358, 1996.

[148] D. Clark, “Evolutionary algorithms in molecular design. methods and principles in medicinal chemistry,” 2000.

[149] M. L. Verdonk, J. C. Cole, M. J. Hartshorn, C. W. Murray, and R. D. Taylor, “Improved protein-ligand docking using GOLD,” *Proteins-Structure Function and Genetics*, vol. 52, no. 4, pp. 609–623, 2003.

[150] J.-P. Changeux and S. Edelstein, “Conformational selection or induced fit? 50 years of debate resolved,” *F1000 Biol Rep*, vol. 3, p. 19, 2011.

[151] M. L. Teodoro and L. E. Kavraki, “Conformational flexibility models for the receptor in structure based drug design,” *Current pharmaceutical design*, vol. 9, no. 20, pp. 1635–1648, 2003.

[152] C. Bissantz, G. Folkers, and D. Rognan, “Protein-based virtual screening of chemical databases. 1. evaluation of different docking/scoring combinations,” *Journal of medicinal chemistry*, vol. 43, no. 25, pp. 4759–4767, 2000.

[153] A. Wallqvist and D. Covell, “Docking enzyme-inhibitor complexes using a preference-based free-energy surface,” *Proteins: Structure, Function, and Bioinformatics*, vol. 25, no. 4, pp. 403–419, 1996.

[154] M. D. Eldridge, C. W. Murray, T. R. Auton, G. V. Paolini, and R. P. Mee, “Empirical scoring functions: I. the development of a fast empirical scoring function to estimate the binding affinity of ligands in receptor complexes,” *Journal of Computer-Aided Molecular Design*, vol. 11, no. 5, pp. 425–445, 1997.

[155] P. S. Charifson, J. J. Corkery, M. A. Murcko, and W. P. Walters, “Consensus scoring: A method for obtaining improved hit rates from docking databases of three-dimensional structures into proteins,” *Journal of medicinal chemistry*, vol. 42, no. 25, pp. 5100–5109, 1999.

[156] G. Rai, V. Kenyon, A. Jadhav, L. Schultz, M. Armstrong, J. B. Jameson, E. Hoobler, W. Leister, A. Simeonov, T. R. Holman, and D. J. Maloney, “Discovery of potent and selective inhibitors of human reticulocyte 15-lipoxygenase-1,” *Journal of Medicinal Chemistry*, vol. 53, no. 20, pp. 7392–7404, 2010.

[157] Q. F. Gan, M. F. Browner, D. L. Sloane, and E. Sigal, “Defining the arachidonic acid binding site of human 15-lipoxygenase - Molecular modeling and mutagenesis,” *Journal of Biological Chemistry*, vol. 271, no. 41, pp. 25412–25418, 1996.

[158] A. T. Wecksler, V. Kenyon, J. D. Deschamps, and T. R. Holman

[159] R. Ahlrichs, M. Bar, M. Haser, H. Horn, and C. Kolmel, “Electronic structure calculations on workstation computers: The program system Turbomole,” *Chemical Physics Letters*, vol. 162, no. 3, pp. 165–169, 1989.

[160] W. Smith and T. R. Forester, “Dl_poly_2.0: A general-purpose parallel molecular dynamics simulation package,” *Journal of Molecular Graphics*, vol. 14, no. 3, pp. 136–141, 1996.

[161] A. D. MacKerell, D. Bashford, M. Bellott, R. L. Dunbrack, J. D. Evanseck, M. J. Field, S. Fischer, J. Gao, H. Guo, S. Ha, D. Joseph-McCarthy, L. Kuchnir, K. Kuczera, F. T. K. Lau, C. Mattos, S. Michnick, T. Ngo, D. T. Nguyen, B. Prodhom, W. E. Reiher, B. Roux, M. Schlenkrich, J. C. Smith, R. Stote, J. Straub, M. Watanabe, J. Wiorkiewicz-Kuczera, D. Yin, and M. Karplus, “All-atom empirical potential for molecular modeling and dynamics studies of proteins,” *Journal of Physical Chemistry B*, vol. 102, no. 18, pp. 3586–3616, 1998.

[162] A. D. MacKerell, M. Feig, and C. L. Brooks, “Extending the treatment of backbone energetics in protein force fields: Limitations of gas-phase quantum mechanics in reproducing protein conformational distributions in molecular dynamics simulations,” *Journal of Computational Chemistry*, vol. 25, no. 11, pp. 1400–1415, 2004.

[163] S. E. Feller, K. Gawrisch, and A. D. MacKerell, "Polyunsaturated Fatty Acids in Lipid Bilayers: Intrinsic and Environmental Contributions to Their Unique Physical Properties," *Journal of the American Chemical Society*, vol. 124, no. 2, pp. 318–326, 2001.

[164] S. E. Feller and A. D. MacKerell, "An Improved Empirical Potential Energy Function for Molecular Simulations of Phospholipids," *The Journal of Physical Chemistry B*, vol. 104, no. 31, pp. 7510–7515, 2000.

[165] A. H. de Vries, P. Sherwood, S. J. Collins, A. M. Rigby, M. Rigutto, and G. J. Kramer, "Zeolite Structure and Reactivity by Combined Quantum-Chemical-Classical Calculations," *The Journal of Physical Chemistry B*, vol. 103, no. 29, pp. 6133–6141, 1999.

[166] D. C. Liu and J. Nocedal, "On the limited memory bfgs method for large-scale optimization," *Mathematical Programming*, vol. 45, no. 3, pp. 503–528, 1989.

[167] J. Nocedal, "Updating quasi-newton matrices with limited storage," *Mathematics of Computation*, vol. 35, no. 151, pp. 773–782, 1980.

[168] J. Baker, "An algorithm for the location of transition states," *Journal of Computational Chemistry*, vol. 7, no. 4, pp. 385–395, 1986.

[169] A. Banerjee, N. Adams, J. Simons, and R. Shepard, "Search for stationary points on surfaces," *Journal of Physical Chemistry*, vol. 89, no. 1, pp. 52–57, 1985.

[170] S. R. Billeter, A. J. Turner, and W. Thiel, "Linear scaling geometry optimisation and transition state search in hybrid delocalised internal coordinates," *Physical Chemistry Chemical Physics*, vol. 2, no. 10, pp. 2177–2186, 2000.

[171] A. E. Reed, R. B. Weinstock, and F. Weinhold, "Natural-population analysis," *Journal of Chemical Physics*, vol. 83, no. 2, pp. 735–746, 1985.

[172] W. Humphrey, A. Dalke, and K. Schulten, "Vmd: Visual molecular dynamics," *Journal of Molecular Graphics & Modelling*, vol. 14, no. 1, pp. 33–38, 1996.

[173] Harihara.Pc and J. A. Pople, "Influence of polarization functions on molecular-orbital hydrogenation energies," *Theoretica Chimica Acta*, vol. 28, no. 3, pp. 213–222, 1973.

[174] P. J. Hay and W. R. Wadt, "*Ab initio* effective core potentials for molecular calculations - Potentials for the transition-metal atoms Sc to Hg," *Journal of Chemical Physics*, vol. 82, no. 1, pp. 270–283, 1985.

[175] L. Toledo, L. Masgrau, J. M. Lluch, and À. González-Lafont, "Substrate binding to mammalian 15-lipoxygenase," *Journal of Computer-Aided Molecular Design*, vol. 25, no. 9, p. 825, 2011.

[176] E. Hatcher, A. V. Soudackov, and S. Hammes-Schiffer, "Proton-coupled electron transfer in soybean lipoxygenase: Dynamical behavior and temperature dependence of kinetic isotope effects," *Journal of the American Chemical Society*, vol. 129, no. 1, pp. 187–196, 2007.

[177] M. H. V. Huynh and T. J. Meyer, "Proton-coupled electron transfer," *Chemical Reviews*, vol. 107, no. 11, pp. 5004–5064, 2007.

[178] D. R. Weinberg, C. J. Gagliardi, J. F. Hull, C. F. Murphy, C. A. Kent, B. C. Westlake, A. Paul, D. H. Ess, D. G. McCafferty, and T. J. Meyer, “Proton-coupled electron transfer,” *Chemical Reviews*, vol. 112, no. 7, pp. 4016–4093, 2012.

[179] F. Claeyssens, J. N. Harvey, F. R. Manby, R. A. Mata, A. J. Mulholland, K. E. Ranaghan, M. SchÃijtz, S. Thiel, W. Thiel, and H.-J. Werner, “High-Accuracy Computation of Reaction Barriers in Enzymes,” *Angewandte Chemie International Edition*, vol. 45, no. 41, pp. 6856–6859, 2006.

[180] R. Lonsdale, J. N. Harvey, and A. J. Mulholland, “Compound i reactivity defines alkene oxidation selectivity in cytochrome p450cam,” *Journal of Physical Chemistry B*, vol. 114, no. 2, pp. 1156–1162, 2010.

[181] R. Lonsdale, K. T. Houghton, J. Zurek, C. M. Bathelt, N. Foloppe, M. J. de Groot, J. N. Harvey, and A. J. Mulholland, “Quantum mechanics/molecular mechanics modeling of regioselectivity of drug metabolism in cytochrome p450 2c9,” *Journal of the American Chemical Society*, vol. 135, no. 21, pp. 8001–8015, 2013.

[182] A. Schafer, H. Horn, and R. Ahlrichs, “Fully optimized contracted gaussian-basis sets for atoms li to kr,” *Journal of Chemical Physics*, vol. 97, no. 4, pp. 2571–2577, 1992.

[183] J. Tao, J. Perdew, V. Staroverov, and G. Scuseria, “Climbing the density functional ladder: Nonempirical meta-generalized gradient approximation designed for molecules and solids,” *Physical Review Letters*, vol. 91, no. 14, 2003.

[184] S. Grimme, “Semiempirical GGA-type density functional constructed with a long-range dispersion correction,” *Journal of Computational Chemistry*, vol. 27, no. 15, pp. 1787–1799, 2006.

[185] P. Saura, R. Suardíaz, L. Masgrau, J. M. Lluch, and À. González-Lafont, “Unraveling How Enzymes Can Use Bulky Residues To Drive Site-Selective C–H Activation: The Case of Mammalian Lipoxygenases Catalyzing Arachidonic Acid Oxidation,” *ACS Catalysis*, pp. 4351–4363, 2014.

[186] W. G. Hoover, “Canonical dynamics - Equilibrium phase-space distributions,” *Physical Review A*, vol. 31, no. 3, pp. 1695–1697, 1985.

[187] J. Saam, I. Ivanov, M. Walther, H. G. Holzhutter, and H. Kühn, “Molecular dioxygen enters the active site of 12/15-lipoxygenase via dynamic oxygen access channels,” *Proceedings of the National Academy of Sciences of the United States of America*, vol. 104, no. 33, pp. 13319–13324, 2007.

[188] A. D. Becke, “Density-functional thermochemistry .3. the role of exact exchange,” *Journal of Chemical Physics*, vol. 98, no. 7, pp. 5648–5652, 1993.

[189] A. M. Cooper and J. Kastner, “Averaging techniques for reaction barriers in qm/mm simulations,” *Chemphyschem*, vol. 15, no. 15, pp. 3264–3269, 2014.

[190] H. Y. Yin, L. B. Xu, and N. A. Porter, “Free radical lipid peroxidation: Mechanisms and analysis,” *Chemical Reviews*, vol. 111, no. 10, pp. 5944–5972, 2011. ISI Document Delivery No.: 834YS Times Cited: 172 Cited Reference Count: 234 Yin, Huiyong Xu, Libin Porter, Ned A. Amer chemical soc Washington.

[191] R. Suardíaz, L. Masgrau, J. M. Lluch, and A. González-Lafont, "An insight into the regiospecificity of linoleic acid peroxidation catalyzed by mammalian 15-lipoxygenases," *Journal of Physical Chemistry B*, vol. 117, no. 14, pp. 3747–3754, 2013.

[192] R. Suardíaz, L. Masgrau, J. M. Lluch, and A. González-Lafont, "On the regio- and stereospecificity of arachidonic acid peroxidation catalyzed by mammalian 15-lypoxygenases: A combined molecular dynamics and qm/mm study," *Chemphyschem*, vol. 14, no. 16, pp. 3777–3787, 2013.

[193] R. Suardíaz, L. Masgrau, J. M. Lluch, and A. González-Lafont, "Regio- and Stereospecificity in the Oxygenation of Arachidonic Acid Catalyzed by Leu597 Mutants of Rabbit 15-Lipoxygenase: A QM/MM Study," *Chemphyschem*, vol. 15, no. 11, pp. 2303–2310, 2014.

[194] E. A. C. Bushnell, R. Jamil, and J. W. Gauld, "Gaining insight into the chemistry of lipoxygenases: a computational investigation into the catalytic mechanism of (8R)-lipoxygenase," *Journal of Biological Inorganic Chemistry*, vol. 18, no. 3, pp. 343–355, 2013.

[195] O. Radmark and B. Samuelsson, "Regulation of the activity of 5-lipoxygenase, a key enzyme in leukotriene biosynthesis," *Biochemical and Biophysical Research Communications*, vol. 396, no. 1, pp. 105–110, 2010.

[196] C. N. Serhan, N. Chiang, and T. E. Van Dyke, "Resolving inflammation: dual anti-inflammatory and pro-resolution lipid mediators," *Nature Reviews Immunology*, vol. 8, no. 5, pp. 349–361, 2008.

[197] A. Di Venere, T. Horn, S. Stehling, G. Mei, L. Masgrau, A. González-Lafont, H. Kühn, and I. Ivanov, "Role of arg403 for thermostability and catalytic activity of rabbit 12/15-lipoxygenase," *Biochimica Et Biophysica Acta-Molecular and Cell Biology of Lipids*, vol. 1831, no. 6, pp. 1079–1088, 2013.

[198] B. Webb and A. Sali, "Comparative Protein Structure Modeling Using MODELLER," *Current protocols in bioinformatics*, vol. 47, pp. 5.6.1–5.6.32, 2014. 25199792.

[199] C. Colovos and T. O. Yeates, "Verification of protein structures - Patterns of nonbonded atomic interactions," *Protein Science*, vol. 2, no. 9, pp. 1511–1519, 1993.

[200] D. Eisenberg, R. Luthy, and J. U. Bowie, *VERIFY3D: Assessment of protein models with three-dimensional profiles*, vol. 277 of *Methods in Enzymology*, pp. 396–404. San Diego: Elsevier Academic Press Inc, 1997.

[201] R. A. Laskowski, M. W. Macarthur, D. S. Moss, and J. M. Thornton, "PROCHECK - A program to check the stereochemical quality of protein structures," *Journal of Applied Crystallography*, vol. 26, pp. 283–291, 1993.

[202] H. Li, A. D. Robertson, and J. H. Jensen, "Very fast empirical prediction and rationalization of protein pk(a) values," *Proteins-Structure Function and Bioinformatics*, vol. 61, no. 4, pp. 704–721, 2005.

[203] S. C. Lovell, J. M. Word, J. S. Richardson, and D. C. Richardson, “The penultimate rotamer library,” *Proteins-Structure Function and Genetics*, vol. 40, no. 3, pp. 389–408, 2000.

[204] Y. R. Wu, C. He, Y. Gao, S. He, Y. Liu, and L. H. Lai, “Dynamic modeling of human 5-lipoxygenase-inhibitor interactions helps to discover novel inhibitors,” *Journal of Medicinal Chemistry*, vol. 55, no. 6, pp. 2597–2605, 2012.

[205] R. Suardíaz, L. Masgrau, A. González-Lafont, and J. M. Lluch, “On the regiospecificity of hydroperoxidation of fatty acids by mammalian lipoxygenases,” *Febs Journal*, vol. 280, pp. 94–94, 2013.

[206] C. R. Søndergaard, M. H. M. Olsson, M. Rostkowski, and J. H. Jensen, “Improved treatment of ligands and coupling effects in empirical calculation and rationalization of pka values,” *Journal of Chemical Theory and Computation*, vol. 7, no. 7, pp. 2284–2295, 2011.

[207] M. H. M. Olsson, C. R. Søndergaard, M. Rostkowski, and J. H. Jensen, “Propka3: Consistent treatment of internal and surface residues in empirical pka predictions,” *Journal of Chemical Theory and Computation*, vol. 7, no. 2, pp. 525–537, 2011.

[208] J. Huang and A. D. MacKerell, “CHARMM36 all-atom additive protein force field: Validation based on comparison to NMR data,” *Journal of Computational Chemistry*, vol. 34, no. 25, pp. 2135–2145, 2013.

[209] G. J. Martyna, D. J. Tobias, and M. L. Klein, “Constant pressure molecular dynamics algorithms,” *The Journal of Chemical Physics*, vol. 101, no. 5, pp. 4177–4189, 1994.

[210] S. E. Feller, Y. Zhang, R. W. Pastor, and B. R. Brooks, “Constant pressure molecular dynamics simulation: The langevin piston method,” *The Journal of Chemical Physics*, vol. 103, no. 11, pp. 4613–4621, 1995.

[211] Giacomo Fiorin and Michael L. Klein and Jérôme Hénin, “Using collective variables to drive molecular dynamics simulations,” *Molecular Physics*, vol. 111, no. 22-23, pp. 3345–3362, 2013.

[212] H. C. Andersen, “Rattle: A “velocity” version of the shake algorithm for molecular dynamics calculations,” *Journal of Computational Physics*, vol. 52, no. 1, pp. 24 – 34, 1983.

[213] T. Darden, D. York, and L. Pedersen, “Particle mesh Ewald: An Nlog(N) method for Ewald sums in large systems,” *The Journal of Chemical Physics*, vol. 98, no. 12, pp. 10089–10092, 1993.

[214] Y. Shao, Z. Gan, E. Epifanovsky, A. T. Gilbert, M. Wormit, J. Kussmann, A. W. Lange, A. Behn, J. Deng, X. Feng, D. Ghosh, M. Goldey, P. R. Horn, *et al.*, “Advances in molecular quantum chemistry contained in the q-chem 4 program package,” *Molecular Physics*, vol. 113, no. 2, pp. 184–215, 2015.

[215] Y. Shao, L. F. Molnar, Y. Jung, J. Kussmann, C. Ochsenfeld, S. T. Brown, A. T. Gilbert, L. V. Slipchenko, S. V. Levchenko, D. P. O'Neill, R. A. DiStasio Jr, *et al.*, “Advances in methods and algorithms in a modern quantum chemistry program package,” *Phys. Chem. Chem. Phys.*, vol. 8, pp. 3172–3191, 2006.

[216] S. Grimme, J. Antony, S. Ehrlich, and H. Krieg, "A consistent and accurate ab initio parametrization of density functional dispersion correction (dft-d) for the 94 elements h-pu," *The Journal of Chemical Physics*, vol. 132, no. 15, 2010.

[217] S. Grimme, S. Ehrlich, and L. Goerigk, "Effect of the damping function in dispersion corrected density functional theory," *Journal of Computational Chemistry*, vol. 32, no. 7, pp. 1456–1465, 2011.

[218] E. Rosta, M. Nowotny, W. Yang, and G. Hummer, "Catalytic mechanism of rna backbone cleavage by ribonuclease h from quantum mechanics/molecular mechanics simulations," *Journal of the American Chemical Society*, vol. 133, no. 23, pp. 8934–8941, 2011.

[219] J. Pu, J. Gao, and D. G. Truhlar, "Multidimensional tunneling, recrossing, and the transmission coefficient for enzymatic reactions," *Chemical Reviews*, vol. 106, no. 8, pp. 3140–3169, 2006.

[220] A. Fernández-Ramos, J. A. Miller, S. J. Klippenstein, and D. G. Truhlar, "Modeling the kinetics of bimolecular reactions," *Chemical Reviews*, vol. 106, no. 11, pp. 4518–4584, 2006.

[221] R. Suardíaz, P. G. Jambrina, L. Masgrau, A. González-Lafont, E. Rosta, and J. M. Lluch, "Understanding the mechanism of the hydrogen abstraction from arachidonic acid catalyzed by the human enzyme 15-lipoxygenase-2. A quantum mechanics/molecular mechanics free energy simulation," *Journal of Chemical Theory and Computation*, vol. 12, no. 4, pp. 2079–2090, 2016.