Biophysical insights into dietary antioxidants as cyclooxygenase modulators

A thesis submitted in fulfilment of the requirements for the degree of
Master of Science

Julia Liang

B.Sc. (Hons.) Pathology, University of Melbourne

School of Science

College of Science, Engineering and Health

RMIT University

February 2018

Declaration

I certify that except where due acknowledgement has been made, the work is that of the author alone; the work has not been submitted previously, in whole or in part, to qualify for any other academic award; the content of the thesis/project is the result of work which has been carried out since the official commencement date of the approved research program; any editorial work, paid or unpaid, carried out by a third party is acknowledged; and, ethics procedures and guidelines have been followed.

I acknowledge the support I have received for my research through the provision of an Australian Government Research Training Program Scholarship.

<div align="right">

Julia Liang

27 February 2018

</div>

Acknowledgements

Publication list

Bonvino NP, **Liang J**, McCord ED, Zafiris E, Benetti N, Ray NB, Hung A, Boskou D, Karagiannis TC, *OliveNet™: A Comprehensive Library of Compounds from Olea europaea*. Database, 2018. 2018: p.bay106-bay106.

Bonvino NP, Ververis K, **Liang J**, Benetti N, Zafiris E, Hung A, Ray N.B, Boskou D, Karagiannis TC, *Utilisation of a new comprehensive database to investigate anti-inflammatory effects of olive phenolics*. J Nutr Intermed Metab, 2017. 8: p.105.

List of abbreviations

1OL	1-oleyltyrosol
AA	Arachidonic acid
AD	Alzheimer's disease
ADMET	Absorption, distribution, metabolism, excretion, and toxicity
APBS	Adaptive Poisson-Boltzmann Solver
BBB	Blood brain barrier
COM	Centre of mass
COX	Cyclooxygenase
CP	Communication pathways
CT	Commute time
CYP	Cytochrome P450
DMPC	1,2-dimyristoyl-sn-glycero-3-phosphocholine
DOPC	1,2-dioleyl-sn-glycero-3-phosphocholine
DPPC	Dipalmitoylphosphatidycholine
EADG	Elenolic acid diglucoside
ED	Essential dynamics
EGF	Epidermal growth factor
EML	Essential Medicines List
EVOO	Extra virgin olive oil
FLP	Flurbiprofen
GI	Gastrointestinal
GROMACS	GROningen Machine for Chemical Simulation
H33342	Hoescht 33342
hERG	Human Ether-a-go-go Related Gene
IDS	Independent dynamic segment
iNOS	Inducible nitric oxide synthase
LFA	Local Feature Analysis
LG2	Ligstroside derivative 2
MBD	Membrane binding domain
MD	Molecular dynamics
MDP	Molecular dynamics parameters
MMHTE	Methyl malate-β-hydoxytyrosol ester
MM-PBSA	Molecular Mechanics-Poisson Boltzmann Surface Area
MONETA	Modular NETwork Analysis
NCI	National Computational Infrastructure
NPT	Constant number of particles, pressure, and temperature
NSAIDs	Non-steroidal anti-inflammatory drugs
NVT	Constant number of particles, volume, and temperature
OLEO	Oleocanthal

OPLS	Optimized Potentials for Liquid Simulations
PAMPA	Parallel artificial membrane permeability assay
PBC	Periodic boundary condition
PCA	Principal components analysis
PDB	Protein Data Bank
PGE_2	Prostaglandin E2
PGH_2	Prostaglandin endoperoxidase H_2
PGHS	Prostaglandin endoperoxide H synthase
PME	Particle-Mesh Ewald
PMF	Potential of mean force
QM/MM	Quantum mechanical/molecular mechanical
QMPLD	Quantum mechanics-polarized ligand docking
QSAR	Quantitative structure-activity relationships
RMSD	Root mean square deviation
RMSF	Root mean square fluctuation
ROS	Reactive oxygen species
SAL	Salicylate
SAR	Structure-activity relationships
SASA	Solvent accessible surface area
SAV	Solvent accessible volume
SMD	Steered molecular dynamics
TXA_2	Thromboxane A_2
VLSCI	Victorian Life Sciences Computation Initiative
VMD	Visual molecular dynamics
WCA	Weeks-Chandler-Andersen
WHAM	Weighted Histogram Analysis Method
XP	Extra precision

Table of contents

List of figures

List of tables

Abstract

Non-steroidal anti-inflammatory drugs (NSAIDs) are among the most widely used therapeutic agents around the world, commonly used to reduce pain. These work by targeting cyclooxygenase (COX) enzymes, which are responsible for the production of inflammatory mediators. There are adverse effects with the use of NSAIDs, including gastrointestinal bleeding, renal disease, and cardiovascular effects. Hence, there has been a rise in the development of alternatives to traditional NSAIDs. Olive oil is a main component of the Mediterranean diet, and is reputable as part of a healthy lifestyle. Phenolic compounds derived from *Olea europaea* contribute to the antioxidant, anti-microbial, and anti-inflammatory properties of extra virgin olive oil. However, specific mechanisms of action are not yet clear. A previous study found that oleocanthal (OLEO), a phenolic compound derived from the olive, had similar effects to ibuprofen, a commonly used NSAID. There are a multitude of additional compounds in the olive that have yet to be investigated. In this project, it was sought to identify potential olive derived compounds with the ability to inhibit COX enzymes to be used in anti-inflammatory therapeutics. The mechanisms of COX inhibition were also studied using *in silico* approaches.

Following a literature review on COX proteins and olive compounds in Chapter 1, a description of computational theory surrounding the *in silico* methods employed in this thesis are presented in Chapter 2. In Chapter 3, a comprehensive literature search was performed to create a library of olive compounds, focussing on the class of phenolics for the purpose of this project.

The structure of human COX-1 was constructed using homology modelling methods in Chapter 4, followed by virtual screening of the olive phenolic library using molecular docking to determine the COX inhibitory potential of all identified ligands. From the docking study, it was determined that 1-oleyltyrosol (1OL) and ligstroside derivative 2 (LG2) demonstrated the greatest binding affinity to both COX-1 and COX-2. Further screening of the compound library was performed by analysing their biological availability in Chapter 5. From examination of absorption, distribution, metabolism, excretion, and toxicity (ADMET) properties of the library, a novel phenolic compound—methyl malate-β-hydoxytyrosol ester

(MMHTE)—was found to both fulfil ADMET criteria and demonstrate strong binding to COX-1 and COX-2. These phenolic compounds were selected for further analysis using molecular dynamics simulations. To complement the ADMET data, a preliminary study on membrane permeability was performed. This was conducted using steered MD simulations of these compounds through a 1,2-dioleyl-sn-glycero-3-phosphocholine (DOPC) lipid bilayer, followed by umbrella sampling simulations of OLEO and MMHTE to estimate the free energy of membrane permeation.

Chapter 6 presents a detailed study on the mechanisms of COX inhibition by these selected compounds using MD simulations. Classical MD simulations were carried out on COX-1 and COX-2 complexed with 1OL, LG2, OLEO, MMHTE, as well as their native ligands that were present in the crystal structure. The stability and backbone fluctuation of these complexes were determined. Protein dynamics were examined using essential dynamics methods and network analysis, which identified that the N-terminal epidermal growth factor-like domain and membrane bound domains of COX-1 and -2 exhibited altered motions when ligands were bound. Distinct dynamical modules were identified, as well as the finding that COX-2 inter-residue communications were more sensitive to ligand binding compared to COX-1. The residue contributions to binding free energy were computed using Molecular Mechanics-Poisson Boltzmann Surface Area (MM-PPBSA) methods.

Through this research, novel olive phenolic compounds were identified which may possess COX inhibitory properties. Future work may provide additional details of the mechanism of COX inhibition, as well as the synthesis of these novel compounds for *in vitro* and *in vivo* validation. Furthermore, it may be demonstrated that olive-derived compounds present a possible avenue for the development of more effective and safe therapeutics in inflammation, as well as provide mechanisms for the anti-inflammatory effects of low dosage dietary COX inhibitors.

1. Introduction

The cyclooxygenase (COX) enzymes are involved in a wide range of physiologic and pathologic processes, and are the target of non-steroidal anti-inflammatory drugs (NSAIDs). In this chapter, a brief introduction describing the role of COX enzymes is discussed in Section 1.1. The structure and function of the protein is described, as well as its mechanism of action and a brief discussion of the history regarding inhibitors of this protein. COX inhibitors are discussed in further detail in Section 1.2. In Section 1.3, an introduction into olives and their potent compounds are presented, as well as a discussion of their COX inhibitory potential. Section 1.4 describes the role of computational methods in drug discovery, particularly in COX inhibitory drugs. The hypothesis and aims of this project are presented in Section 1.5.

1.1 Role of cyclooxygenase enzymes in inflammation

1.1.1 *Structure and function of cyclooxygenase enzymes*

COX enzymes, also known as prostaglandin endoperoxide H synthases, are responsible for the production of mediators that drive the inflammatory process. This involves the production of prostaglandins, which are synthesised from COX catalysed pathways. NSAIDs are among the most widely used therapeutics globally, and are well characterised in their ability to inhibit COX isoforms COX-1 and COX-2.

In general, COX-1 is constitutively expressed and present in nearly all tissues, whereas COX-2 is induced during inflammation [1]. COX-1 has a role in the production of prostaglandins involved in various physiological functions, such as the maintenance of renal function and mucosal production in the gastrointestinal (GI) tract [2-4]. Conversely, COX-2 expression is induced by cytokines and inflammatory stimuli, facilitating the development of pain, inflammation, and fever, as well as being implicated in some cancers [4]. A third isoform, COX-3, is a variant of COX-1 expressed in the brain and heart, and possesses cyclooxygenase activity 80% lower than that of COX-1 [5]. COX-3 has been suggested to contribute to the synthesis of prostanoids in inflammation; however its function is not clearly understood in humans [5-7].

COX enzymes are membrane-bound in the endoplasmic reticulum and are comprised of three independent folding units: an epidermal growth factor (EGF) domain, a membrane-binding domain (MBD), and large globular enzymatic domain [8]. There are two adjacent but distinct active sites for peroxidase and COX activity, with the COX active site lined by a long, hydrophobic channel [8].

COX-1 and COX-2 isoenzymes are highly conserved and are structurally homologous, sharing a 60% sequence identity [9]. COX-1 and COX-2 contain 576 and 587 amino acids, respectively. The enzymes are almost identical sequentially, with the exception of an insertion found in COX-2 after THR 106 which has no equivalent residue in COX-1, and an 18 amino acid insertion after the C-terminus of COX-2 [9]. Thus, the main differences reside in four areas: first, differing lengths of signal peptides are present in both isoforms; second, the presence of an additional 18 amino acid insertion in COX-2 after the C-terminus; third, differences in the sequence between the isoforms at MBD; and finally, COX-1 is glycosylated at three sites, while COX-2 is variably glycosylated at two to four sites [9-11].

1.1.2 *Prostaglandin production by COX enzymes*

Vital to the inflammatory process, prostanoids are a subclass of eicosanoids which are signalling molecules involved in the oxidation of polyunsaturated fatty acids. These consist of prostaglandins, thromboxanes, and prostacyclin. The synthesis of prostanoids occurs in three steps: 1) mobilisation of a fatty acid substrate, usually arachidonic acid (AA) from membrane phospholipids catalysed by phospholipase A2; 2) transformation of AA to prostaglandin endoperoxidase H_2 (PGH_2) through a prostaglandin endoperoxide H synthase (PGHS); 3) conversion of PGH_2 to specific prostanoids via synthases and specific isomerases [12].

Prostaglandins are hormone-like factors mediating autocrine and paracrine signalling in many physiological and pathological processes. They are implicated in diseases such as cancer, inflammation, cardiovascular disease, and hypertension [13]. Generally, prostaglandins produced by COX-1 play a role in physiological homeostasis, while those produced by COX-2 contribute to inflammatory effects.

Prostaglandin E2 (PGE$_2$) is one of the main prostaglandins mediating the inflammatory process. It dilates small blood vessels, leading to redness and heat as the increased vascular permeability causes swelling of tissues [6, 14]. Pain is induced through a sensitising action on the peripheral terminals of sensory nerves [6, 14]. PGE$_2$ also acts on neurons, contributing to systemic responses of inflammation such as fatigue, fever, and pain hypersensitivity [6, 13, 14].

As well as inflammatory responses PGE$_2$ has been shown to influence a range of other biological processes. The diversity in the effects of PGE$_2$ can be attributed in part to four E-type prostaglandin receptors (EP1-4) which are heterogeneously coupled to intracellular signal transduction pathways [13]. It has been suggested that PGE$_2$ may have an immunomodulatory effect at multiple levels within the immune system, since prostaglandin receptors have been found to be present on major subsets of cells involved in the adaptive immune response, including T, B, and dendritic cells, [13, 15]. Furthermore, PGE$_2$ has been observed to have multiple and apparently opposing effects as both a dilator and constrictor in vascular smooth muscle [13, 16, 17].

In the kidney, PGE$_2$ production is critical for renal function by maintaining renal blood flow and glomerular filtration rate during physiological stress [13, 18]. The transport of salt and water is modulated in the distal tubule, stimulating the release of renin [18]. PGE$_2$ has also shown to contribute to the maintenance of blood pressure, especially in a high salt setting [13, 19].

1.1.3 Development of anti-inflammatory drugs

Salicylates have been used for pain relief since ancient times. Salicylate-rich plants were described to have the ability to alleviate pain. Stone tablets from the Assyrians of the Sumerian period noted the use of willow leaves, and in ancient Egypt, the Ebers papyrus dating back 3,500 years referred to the application of myrtle leaves for rheumatic pains [20, 21]. Hippocrates recommended the use of willow bark around 400 BC; this became part of Western medicine throughout classical antiquity and the Middle Ages [21]. Plants containing salicylates were also used for therapeutic purposes throughout China and other parts of Asia [21].

It was not until 1763 that the first scientific descriptions of the beneficial effects of willow bark were given. In a letter to the Royal Society, the Reverend Edward Stone of Chipping Norton in Oxfordshire, England described the successful treatment of 50 patients with fever with doses of 1 dram (1.8g) of powdered willow bark [22]. As chemical techniques developed in the 18th and 19th centuries, scientists were able to characterise the compounds extracted from the willow bark. In 1828, Johann Büchner isolated and named salicin, meaning "willow" in Latin [20]. A year later, Henri Leroux isolated a pure crystalline form, and salicyclic acid was generated by Raffaele Piria in 1838 [20, 23]. Physicians began prescribing the purified compounds, with Thomas Maclagan describing the beneficial effects for patients with rheumatism in *The Lancet* in 1876 [20, 24].

Investigating a less irritating replacement for salicyclic acid, Felix Hoffman, a German chemist working for the Bayer Company, modified the structure of salicyclic acid to obtain acetylsalicyclic acid in 1897 [20, 23]. The drug was named aspirin, and the first mass marketing of a pharmaceutical agent occurred as Bayer spread information about aspirin to 30,000 doctors [23]. In 1904, the powdered form of aspirin was replaced with a tablet, and the drug became a part of everyday life [23].

From the 1950s onwards, other drugs sharing anti-inflammatory and analgesic effects were discovered, such as ibuprofen and naproxen, and called 'non-steroidal, anti-inflammatory drugs' (NSAIDs) [21]. In 1971, John Vane first described the mechanism of action of aspirin and aspirin-like drugs, demonstrating that the decreased production of prostaglandins was achieved by the inhibition of COX enzymes [25]. Later, the COX enzymes were found to exist in two forms: COX-1 in 1976, and COX-2 in 1991 [4]. NSAIDs remain among the most widely used therapeutic agents globally, used by approximately 30 million people daily and accounting for approximately 60% of the over-the-counter analgesic market in the US [3, 26, 27].

1.2 COX inhibitors

NSAIDs describe a diverse class of drugs that function by competitively inhibiting cyclooxygenase enzymes to produce analgesic and antipyretic effects. NSAIDs are widely used around the world, and are a recommended therapy for patients with osteoarthritis and rheumatoid arthritis [3, 28-30].

1.2.1 Complications associated with NSAID usage

While NSAIDs are commonly used for pain relief, there are important adverse side effects associated with long-term use. These include GI complications, renal toxicity, exacerbated hypertension, and cardiovascular events [3, 31]. The use of these drugs is further complicated by the fact that many patients with chronic conditions requiring NSAIDs, particularly the elderly, have comorbidities that increase the risk of complications [32].

Traditional NSAIDs have been found to be associated with GI complications, with the risk of upper GI complications increasing by 3- to 5-fold [33]. COX-1 derived PGE_2 plays a protective role in maintaining the integrity of the gastric mucosa, and adverse effects of the GI tract are associated with inhibition of prostaglandins by non-selective NSAIDs [13, 34]. NSAID usage has previously been suggested as one of the most common causes of gastric injury in the United States, with 1998 estimates stating that approximately 100,000 people were hospitalised annually due to GI complications [35, 36]. The mortality rate is reported to be approximately 5% [36, 37].

The use of NSAIDs is associated with an increased risk of cardiovascular complications, such as myocardial infarction, heart failure, and hypertension [38]. The mechanism for this action has been suggested to affect COX inhibition by the imbalance of COX-2-regulated production of pro-aggregatory thromboxane in platelets, and anti-aggregatory prostaglandin I_2 [3, 39].

1.2.2 Selective COX inhibitors

Since COX-2 is induced by inflammatory stimuli, it was suggested that the anti-inflammatory actions of NSAIDs are due to COX-2 inhibition, while the unwanted side effects are due to inhibition of COX-1. In the early 2000's, COX-2 selective NSAIDs were seen as an alternative to traditional NSAIDs, having a similar efficacy and enhanced GI tolerability [3]. However, it was later found with the use of COX-2 selective drugs, there was a significant increase in the risk of vascular events [40].

The Vioxx GI Outcomes Research (VIGOR) study compared the efficacy and adverse effects of rofecoxib, a selective COX-2 inhibitor, with non-selective naproxen in patients with rheumatoid arthritis [41]. There was a four-fold increase in the incidence of myocardial

infarction in the rofecoxib group compared to naproxen [41]. From the Adenomatous Polyp Prevention on Vioxx (APPROVe) study, it was further confirmed that the use of rofecoxib in patients with a history of colorectal adenomas was associated with an increased cardiovascular risk of almost two-fold compared to the placebo [42]. Following these studies, rofecoxib was voluntarily withdrawn from the market in 2004 [3]. Another selective COX-2 inhibitor, valdecoxib was subsequently withdrawn in 2005 [43]. As a consequence, there was a reduction in the number of prescriptions for all NSAIDs in the following years [3, 43, 44].

Listed on 74 national Essential Medicines Lists (EMLs), diclofenac is the most commonly used NSAID worldwide, having a market share close to that of the next three most popular drugs combined [45]. Diclofenac and exoricoxib are considered "high risk" for cardiovascular complications, and these two drugs combined account for one-third of the market across 15 countries with no difference across high- and low-income countries [45]. A study by McGettigan et al. suggested that diclofenac be removed from EMLs since it carried a cardiovascular risk similar to that of rofecoxib, which was withdrawn from global markets due to cardiovascular toxicity [45]. Despite naproxen having the lowest cardiovascular risk, 51 of the countries listing diclofenac on their EML did not list naproxen, and it is only listed in 27 out of 86 national EMLs published or updated since 2007 [45]. Meta-analyses of published studies have found that there are no significant differences in effectiveness of pain relief between NSAIDs at standard doses [3].
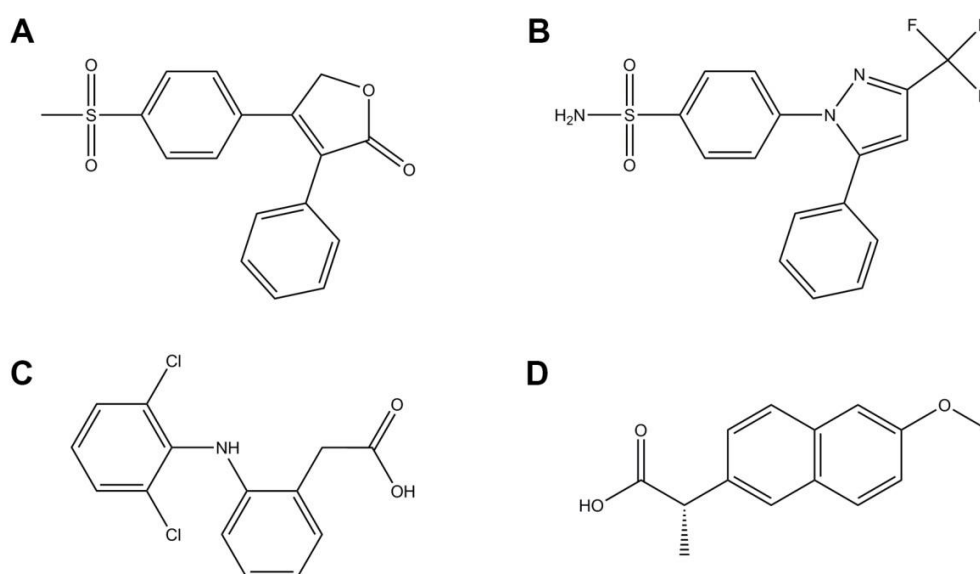
Figure 1.1: Chemical structures of COX-2 selective inhibitors rofecoxib (A) and celecoxib (B), and non-selective NSAIDs diclofenac (C) and naproxen (D).

Through these findings, the main assumption became that increased COX-1 selectivity was associated with increased risk of GI complications, while increased COX-2 selectivity was associated with increased risk of cardiovascular toxicity [46]. However, further evidence suggests that risks associated with NSAID usage are more complex, with COX selectivity alone being insufficient to determine risk. The increased cardiovascular risk can be attributed to all NSAIDs, both COX-2 selective and non-selective [47]. This could be due to both types of drugs reducing the amount of prostaglandin I_2 (PGI_2), which is an anti-thrombotic and anti-platelet hormone, thereby increasing platelet activity [46]. The action of NSAIDs, rather than COX-2 selective inhibitors, could be another mechanism for increased cardiovascular risk. This mechanism reduces levels of thromboxane A_2 (TXA_2), a metabolite produced by COX-2 thereby possibly indicating reduced platelet activation [46]. Both NSAIDs and COX-2 selective drugs have been shown to dose-dependently increase blood pressure, adding to the risk of thrombosis [46]. Furthermore, the Prospective Randomized Evaluation of Celecoxib Integrated Safety versus Ibuprofen or Naproxen (PRECISION) trial found that celecoxib, a COX-2 selective inhibitor, was not associated with increased cardiovascular events compared to non-selective NSAIDs ibuprofen and naproxen [48]. This thesis proposes that alternative approaches to COX inhibition should be considered, rather than allowing the COX-selectivity of a drug to define its side effects [46].

While the usage of NSAIDs is associated with adverse side effects as mentioned above, low dosages have been known to confer health benefits. In particular, aspirin has been shown to be beneficial for cardiovascular health when administered in low doses [49]. With an increasing intake of aspirin or ibuprofen, a significant decline in the risk of the four major types of cancer: colon, breast, lung, and prostate cancer [50]. Reductions in the risk of oesophageal, stomach, and ovarian cancers were also observed, with protective effects more apparent after five or more years of use [50]. These results were attributed to the role of COX-2 overexpression and increased prostaglandin synthesis in carcinogenesis and metastasis [50, 51]. NSAID usage has also been shown to reduce the secretion of amyloid-beta peptides implicated in Alzheimer's disease [52].

1.3 <u>Olive-derived compounds and their anti-inflammatory action</u>

Olive oil, obtained from the fruits or drupes of *Olea europaea*, is a key component of the Mediterranean diet, and has been used for centuries as a preventative and therapeutic commodity. The Mediterranean diet was first identified in the Seven Countries Study as having potential to increase longevity [53]. A diet rich in extra virgin olive oil (EVOO) has been shown to benefit health through the prevention and reduction of hypertension, cardiovascular risk, oxidative stress, obesity, type 2 diabetes, cancer, and inflammatory processes [54].

EVOO is produced by direct pressing or centrifugation of the olives [55]. Many of the health benefits of EVOO are attributed to its high phenolic compound content, consisting of antioxidant, anti-inflammatory, anti-cancer, and antimicrobial properties. EVOO contains approximately 10 times more simple phenolics compared to refined olive oil [56]. There are a wide range of phenolic compounds present in the olive, with the major classes being simple phenols, secoiridoids and polyphenols [57]. Oleuropein and hydroxytyrosol have been identified as being among the major phenolic compounds identified in olive cultivars, and are the subjects of investigation for their pharmacological effects [58, 59].

Olive phenolic compounds possess potent antioxidant activities, scavenging free radicals and removing reactive oxygen species (ROS) to reduce oxidative stress within cells. A range of diseases depends on the imbalance in ROS, such as the oxidation of low density lipoproteins in the formation of atherosclerotic lesions [60]. Oleuropein and olive leaf extract have been demonstrated to act as a skin protectant, preventing ultraviolet B radiation-induced damage and carcinogenesis in mice, as well as reducing MMP-2, MMP-9, MMP-13, VEGF, and COX-2 in the skin [61, 62]. Some olive phenolic compounds have strong anti-inflammatory effects both *in vitro* and *in vivo*, such as through the inhibition of platelet aggregation [63], and by reducing the production of inflammatory mediators in monocytes [64].
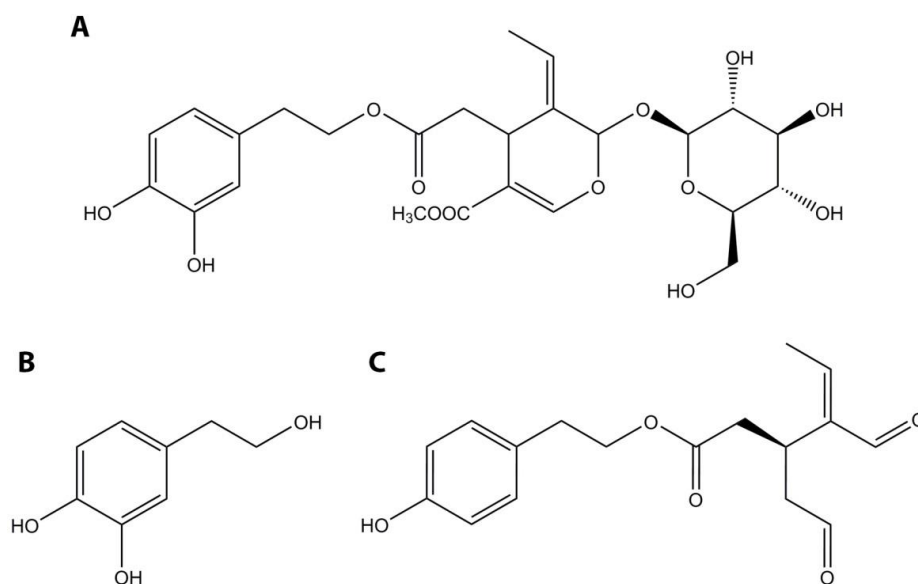
Figure 1.2: Chemical structure of olive phenolic compounds oleuropein (A), hydroxytyrosol (B), and oleocanthal (C).

In the search for naturally derived compounds with pharmacological properties, oleocanthal has recently become a compound of interest. Anti-cancer properties have been observed, demonstrating anti-proliferative effects and encouraging cell apoptosis in various human cancer cell lines [65-67]. Oleocanthal has also been investigated in neurodegeneration, for its anti-aggregation activities on tau protein implicated in Alzheimer's disease [68]. *In vitro* studies by Li et al. showed that oleocanthal is able to inhibit the polymerisation of tau protein through a covalent mechanism [69]. In the landmark study by Beauchamp et al., oleocanthal was found to share similar pharmacological activity to the non-steroidal anti-inflammatory drug ibuprofen [70]. Like ibuprofen, oleocanthal was shown to non-selectively inhibit COX-1 and COX-2 [70]. Oleocanthal is responsible for the distinct throat irritation of EVOO (oleo- for olive, -canth- for sting, -al for aldehyde), a property that Beauchamp et al. noted was similar to the ingestion of ibuprofen solutions [70, 71].

With a daily ingestion of 50 g of EVOO containing 200 μg per ml of oleocanthal of which 60-90% is absorbed, Beauchamp et al. noted that this would correspond to an intake of up to 9 mg per day, equivalent to approximately 10% of the ibuprofen dosage for pain relief—a low dosage [70]. The discovery of the COX inhibitory potential of oleocanthal provided a link in the mechanism of health benefits attributed to the Mediterranean diet. It is therefore plausible that low, chronic consumption of naturally occurring COX inhibitors, such as those

derived from the olive, may reduce inflammation over time, and hence contribute to the reduced development of chronic inflammatory disease.

1.4 Computational studies in drug discovery

Experimental methods for the determination of the ability of a drug to affect a target include enzyme activity assays, which can be determined by absorbance or fluorescence [72]. Since there are a large number of phenolic compounds in the olive, it would not be feasible to test these experimentally. Therefore, *in silico* methods were utilised in this thesis. This included molecular docking and dynamic simulations, which has become an evolving area of research proving to be useful in drug discovery methods. Molecular modelling methods have become a powerful tool for building, visualising, and analysing processes on a molecular level. With an ever increasing amount of biological data becoming available, computational methods have become an important step in drug design.

Previous computational studies have elucidated the structure and function of the COX proteins in detail. A study by Nina et al. used computational approaches to the study association of COX-1 with the membrane bilayer, and provided insight into the mechanism proposed in 1994 of the enzyme being anchored via one leaflet of the bilayer [73]. An advanced computational technique—metadynamics simulations—was employed by Limongelli et al. to simulate the dissociation of a selective inhibitor in both COX-1 and COX-2 to find a previously unreported binding mode in COX-2 [74]. This aided in explaining the time-dependent inhibitory behaviour of COX-2 selective inhibitors [74]. Also utilising advanced computational approaches, Lei at al. employed Born-Oppenheimer molecular dynamics simulations with *ab initio* quantum mechanical/molecular mechanical potential and umbrella sampling to study the detailed mechanism of the biochemical action of aspirin [75].

Molecular docking involves predicting the binding affinity and orientation of putative inhibitors in a selected macromolecular target. These methods have led to two general classes of selective COX-2 inhibitors, diarylheterocycles and methanesulfonanilides [76]. Selective COX-2 inhibitors mainly fall into the diarylheterocycle family, characterised by two vicinal (adjacent) aryl groups attached to an unsaturated central ring, such as in rofecoxib and celecoxib (Figure 1.1 A and B, respectively). A computational method that has become increasingly popular in drug design is the use of quantitative structure-activity relationships (QSAR). Using structural features of known active and inactive compounds, putative

pharmacophore models are generated for potential inhibitors. In conjunction with molecular docking methods, this technique has been utilised in the development of novel COX inhibitors [77-80].

However, there are limitations that must be taken into account in docking studies. Since there is a high computational cost in considering the flexibility of both the receptor and ligand, a common approach was the treatment of the protein as a rigid structure [81-83]. More recently, the flexibility of the receptor is able to be modelled by allowing the movement of side chains in programs such as AutoDock 4.0 [84]. Other challenges include the sampling accuracy of the docking algorithm and choice of scoring function [85]. A common approach to determine the accuracy of the docking program is to compare predicted docking modes with experimentally determined ones [86].

While molecular dynamics simulation can be utilised to describe more biologically relevant systems, it is necessary to also consider the limitations of this technique. These mainly occur due to the size of the system and length of the possible simulation, as requirements for computational resources increase with these factors. For example, some biological processes such as protein folding can occur on microsecond timescales, while classical simulations are typically run on nanosecond timescales [87]. Generally, molecular dynamics simulations are extremely useful for providing insight into time-dependent fluctuations and conformational changes in systems that are useful in understanding their functions, especially in drug-target binding [88].

Despite the challenges, computational tools have almost become routine, and many drugs developed in part by structure-based drug design methods are in late-stage clinical trials or have reached the market [81, 88]. *In silico* methods are a complementary technique to traditional experiments, reducing the amount of work required and gaining insight into interactions at an atomistic level.

1.5 Hypothesis and aims

We hypothesise that olive derived compounds provide a viable basis for the development of therapeutics in inflammatory processes. Therefore, the specific aims of this project are to:

- Identify appropriate candidate compounds as inhibitors of target proteins in inflammation
- Examine *in silico* mechanisms of inhibition, involving protein-ligand complexes implicated in inflammatory pathways through molecular simulations

2. Computational theory

Computational techniques are a useful tool that allows examination of atomic level interactions that occur between proteins and drug targets. This chapter provides an overview of computational methods used to investigate biological systems, particularly those of protein and ligand drug targets. Basic theory around these techniques will be discussed, as well as a brief description of the methods used in this thesis.

Following development of the olive compound library, computational methods utilised include: homology modelling to construct human COX-1, followed by docking to propose the ligands most likely to bind strongly to the COX enzymes amongst the olive library. Absorption, distribution, metabolism, excretion, and toxicity (ADMET) analysis was performed on the selected ligands to further narrow the candidates, and molecular dynamics (MD) simulations on the top ligands to delve into molecular mechanisms of action. This involved essential dynamics, network analysis, and molecular mechanics Poisson-Boltzmann surface area (MM-PBSA) to study protein dynamics and residue-level interactions from the MD trajectories. MD was also performed on the top ligands to examine their membrane permeation process as an adjunct to the ADMET property analysis.
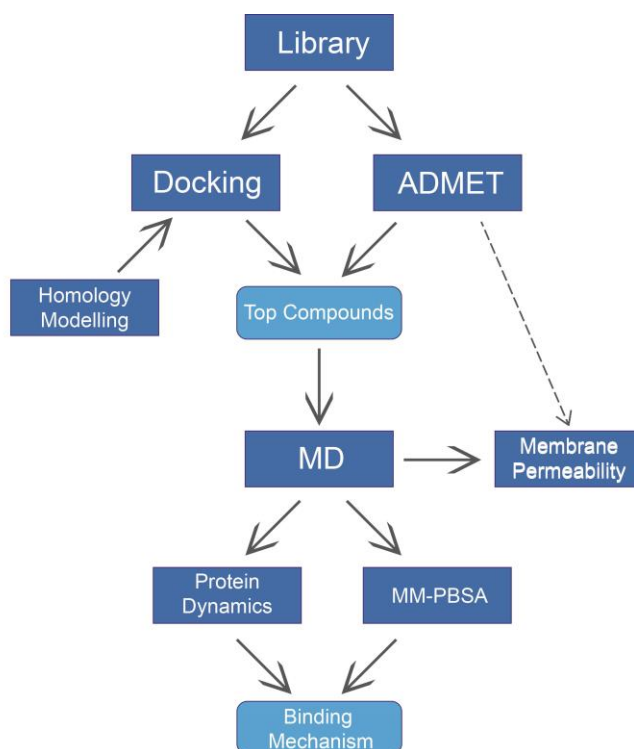
Figure 2.1: Graphical summary of methods used

2.1 Molecular docking

In drug discovery research, the screening of a large number of compounds is often required for a particular protein target. Molecular docking is one of the tools employed that can enhance the rate of this screening. Initially, molecular docking was used for the prediction and reproduction of protein-ligand complexes [89-92]. This subsequently led to the development of molecular docking as a technique in drug discovery for the identification and optimisation of lead compounds, in many cases by the screening of databases. The development of combinatorial chemistry has led to the application of molecular docking in aiding the design of libraries to pre-screen compound databases *in silico* [89].

There are two main parts to docking programs: the search of the configurational and conformational degrees and freedom, and the scoring or evaluation function [89]. Force fields refer to parameters that can be used to describe the energy of a protein, and are utilised in molecular docking methods. The potential energy is often defined using a molecular mechanics force field, and the search algorithm thoroughly searches the conformational space of the ligand based on potential energy of its intra and inter-molecular interactions to determine the global energy minimum. In rigid docking, different positions for the ligand in the active site of the receptor are explored using translational and rotational degrees of freedom [89]. For flexible ligand docking which was performed in this thesis, the exploration of torsional degrees of freedom of the ligand was added to the process [89]. Scoring functions assess both the steric and chemical complementarity between the ligand and receptor [89].

In Chapter 4, the molecular docking calculations performed utilised the quantum mechanics-polarized ligand docking (QPLD) protocol implemented in the Schrodinger Suite [93, 94]. In this method, the docking was carried out with the extra precision (XP) docking protocol of Glide, which was previously demonstrated to consistently predict ligand binding poses with 90.0% accuracy [95, 96]. It has also been shown that Glide is able to predict experimental poses of ligands with RMSDs ranging from 1.5 to 2 Å [96, 97]. In a 2004 study evaluating a dataset of 150 protein-ligand complexes, it was found that Glide was able to correctly identify the crystallographic pose within 2.0 Å in 61% of the cases, compared to other docking programs: 48% with GOLD and 45% with ICM [96, 98].

Glide works by performing a systematic search of the conformational, orientation, and positional space of the docked ligand using the OPLS-AA (Optimized Potentials for Liquid

Simulations) force field [99]. Monte Carlo sampling is then used to further refine the best possible conformation [96, 100]. The scoring function of Glide is based on an empirically based function by Eldridge et al. [101].

The QPLD docking algorithm incorporates quantum mechanical/molecular mechanical (QM/MM) calculations to further take into the account electric charges in protein-ligand docking. In this method, only ligands are treated as the quantum region, as fixed charges of ligands obtained from force field parameterization are replaced with QM/MM calculations in the protein environment [94]. The protein is treated with classical molecular mechanics defined by the OPLS force field.

2.2 ADMET prediction

Drug discovery is a complex process involving disease selection, target identification, hit discovery, lead optimisation, pre-clinical, and clinical trials [102]. Estimation of permeation and excretion rates of drugs is of key importance. In 1991, approximately 40% of attrition for all candidate drugs was related to adverse pharmacokinetic and bioavailability results [103]. Experimental measures of permeability such as the parallel artificial membrane permeability assay (PAMPA) and cell based Caco-2 assay have assisted in reducing drug attrition rates [104-107]. The development of combinatorial chemistry and high-throughput screening has significantly increased the need for initial data on absorption, distribution, metabolism, excretion, and toxicity (ADMET) of compounds. Hence, *in silico* approaches have been developed to accelerate the drug discovery process. In Chapter 5, the pharmacokinetic and bioavailable properties of olive phenolic compounds were estimated using BIOVIA Discovery Studio (San Diego, US). Based on experimental assay results, training sets are used to develop linear response models with quantitative structure-property relationships (QSAR), which can be used to make predictions on molecules of interest [108].

Physiochemical properties have a major impact on the pharmacokinetic and metabolic fate of drugs in the body. Lipophilicity is a key parameter that can determine the membrane permeability, which is related to drug absorption, distribution, and route of drug clearance in the body. One of the main measures of lipophilicity is the partition coefficient in an octanol/water system, frequently expressed as log P [108]. Other properties commonly

calculated include solubility, acid strength (pKa), hydrogen bonding capacity, and permeability [108].

As well as membrane permeability, biological events can affect drug absorption such as the influence of drug transporters and metabolism. Models can be based on descriptors such as log P or polar surface area, which describes the hydrogen-bonding potential of a compound [109]. It is important to consider the plasma protein binding ability of a drug, since they are able to bind to a variety of particles in the blood, including red blood cells, platelets, and proteins such as albumin [110]. Binding to these factors would render the drug unable to pass through membranes and act on their intended drug target [110]. Blood brain barrier (BBB) penetration profiles may also be considered when assessing drugs to determine its ability to target or avoid the brain [108]. The metabolism of drugs primarily occurs in the liver. The most important enzymes involved in this process are the cytochrome P450s (CYPs), particularly CYP3A4, CYP2D6, CYP2C9 and CYP2C19 in humans [108]. QSAR models have also been used to predict hepatotoxicity, which is a major factor in the high attrition rate of drugs [111, 112].

To gain further insight into the permeability process of olive compounds, molecular dynamics simulations were utilised to examine the permeation mechanism of selected compounds, and to obtain a qualitative description of the forces and energies involved in permeation through a simple model lipid bilayer. This will be discussed in section 2.4.4 below.

2.3 Computational methods for studying protein structure

Computational techniques are being increasingly used to complement experiments. With increasing computational power and improvement in theoretical algorithms, computational modelling is being frequently applied to study the structure and function of biological macromolecules on a physical basis that is not accessible using traditional experimental techniques. These methods allow for the examination of interactions that dictate the behaviour of nanoscale systems, allowing behaviour to be observed at an atomic level. This section provides an overview of the computational methodologies commonly used to investigate biological systems, such as the interaction between drugs and protein targets.

2.3.1 Homology modelling

Obtaining a reliable 3-dimensional structure of macromolecules is crucial for the understanding of protein-ligand complex formation. While a lack of knowledge of 3D structures has previously impeded efforts in determining the binding specificities of ligands with protein targets, the increase in modelling software and growing number of known protein structures is rendering homology modelling an increasingly popular and reliable method for obtaining 3D coordinates of proteins [113]. Homology modelling techniques are based on the fact that evolutionarily related proteins share a similar structure [114]. In proteins, the structural conformation is more highly conserved than its amino acid sequence, with small or medium changes in a sequence usually having little impact on the 3D structure [115]. Example applications of homology modelling include hypotheses in drug design, ligand binding site, substrate specificity, and function annotation [113]. Homology modelling can also be used to provide starting models for solving structures from X-ray crystallography, NMR, and electron microscopy [116].

Protein structures obtained using homology modelling methods have been successfully used in docking [117-119]. However, it is worth noting that the reliability of docking results is heavily dependent on the quality of the homology model. A study by McGovern and Shoichet compared the quality of docking results when ligand bound or apo crystallographic structures, or homology modelled structures were used as templates [120]. While ligand bound and apo crystallographic structures tended to yield more successful docking calculations, homology models produced enrichment factors of ten or better in eight of the ten systems studied [120]. Where modelled conformations showed poorer discrimination between ligands and decoys, it was often attributed to changes in the geometry of the binding site. This study highlighted that the quality of docking results can be influenced by slight conformational changes, emphasising the importance of selecting appropriate templates for modelling proteins. In general, it was found that homology models constructed with a template of high sequence similarity often performed better [120]. As well as sequence identity, crystal structures of a higher quality should be selected [121].

In this thesis, a lack of experimental structures for human COX-1 rendered it necessary to construct a predicted model using homology modelling methods. This technique uses one or more existing experimental structures similar to the protein of interest to serve as a template structure, such that the protein model of a target sequence can be constructed. This involves 4

main steps: template identification, target-template sequence alignment, model building, and model evaluation. Modeller 9.16 was used to construct the model [122].

2.3.1.1 Template identification

The initial step of homology modelling involves the identification of a suitable template for the query sequence. The target sequence, the structure to be modelled, is used as a query to search a structural database to identify a suitable template sequence that has an experimentally determined structure. Structural databases such as the Protein Data Bank (PDB) are searched for an appropriate sequence [123].

2.3.1.2 Target-template sequence alignment

A determinant of the quality of a model is the identity between template and target sequences. A target-template sequence identity of between 30 and 40%, sequence alignment can be performed using BLAST, FASTA, and SSEARCH sequence alignment methods [124, 125]. A pairwise alignment is carried out between the target and template sequences to obtain an optimal alignments [126]. This can identify the most conserved and variable regions among the two sequences, which can provide information about their evolutionary relationships.

When the sequence similarity is between the range of 10 to 30%, or the "twilight zone", profile sequence methods can be used to increase the accuracy of the sequence alignment [127, 128]. Multiple sequence alignment can be performed using a number of related sequences, which allows the derivation of position-specific scoring matrices [129]. This can be performed using PSI-BLAST, which implements a heuristic search algorithm for short motifs [130].

2.3.1.3 Model building

Following sequence alignment, model building can be carried out. This utilises atomic and residual information extracted from the aligned sequences, and focusses on four aspects: backbone construction, side chain modelling, loop modelling, and model optimisation [126].

In this study, Modeller v9.16 was used to build the homology model of human COX-1 in Chapter 4, which was subsequently used for molecular dynamics simulations in Chapter 6 [122]. Modeller initially generates restraints on the structure of the target sequence using its alignment to related protein structures as a guide, derived from assuming similar corresponding distances between aligned residues in the template and target structures [131]. These restraints usually incorporate stereochemical restraints on bond lengths, bond angles,

dihedral angles, and non-bonded atom-atom contacts from a molecular mechanics force field [131, 132]. The resulting model is derived by minimisation of violations from all restraints [131].

2.3.1.4 Model evaluation

After a model has been built, it is important that the model be checked for possible errors. A sequence identity of greater than 30% is generally a good indicator of the expected accuracy of the model [131]. If the sequence identity falls below 30%, the model evaluation methods may be useful to determine whether errors have occurred [115].

Tools such as PROCHECK and WHATCHECK assess the stereochemistry of a model—such as bonds, bond angles, dihedral angles, and non-bonded atom-atom distances—to determine whether a model satisfies the restraints used to calculate it [133, 134]. Ramachandran plots are a technique used to visualise the distributions of φ and ψ torsional angles in the protein structure, which are parameters important for protein folding [135]. In proteins, φ angles generally remain within the range of -60° to -150°, while ψ angles are between -60° and 120° [135, 136].

To assess the template alignment, or whether a correct template was used, tools such as ProSA and Verify3D can be used [137, 138]. These tools calculate a pseudo energy profile of a model, reporting the energy for each position in the model such that peaks in the profile frequently correspond to errors [131].

2.3.2 Molecular mechanics and dynamics

Molecular modelling is a major field describing a range of techniques that are applied to model the behaviour of molecules. The advantage of molecular modelling lies in the ability to explore the changes in a controlled system of atomic detail over timescales reaching femto, micro, or milliseconds, at various environmental conditions which cannot be performed using conventional experiments.

One of the main approaches in studying the dynamics of biological systems is molecular mechanics-based classical MD. In this thesis, MD simulations were utilised in Chapter 6 to investigate the mechanism of COX complexed with olive ligands. Classical MD simulations were performed for COX-1 and COX-2 with each of the phenolic ligands selected from

docking and ADMET procedures. Simulations were carried out using the GROMACS software package [139].

2.3.2.1 *Force fields and potential energy function*

Force fields are a set of rules or parameters that dictate the properties of each molecule within the simulated system. These calculate the potential energy of a system of atoms, such as in a biological system, as well as incorporating geometric and energy-related parameters to study structure-activity relationships and mimic experimental observations [140-142]. While force fields are complex equations, the simplicity of its representation of molecular features ensures that energy and force calculations are fast, even for large systems [142].

Force field parameters descriptors for atoms include the atomic mass, partial charges, van der Waals radius, bond angle, bond length, and dihedral angles. Different force fields vary in their parameterisation, rendering then not necessarily interchangeable. Additionally, not all force fields allow representation of all molecule types [142]. Despite this, simulations performing using modern force-fields are normally equivalent [142, 143]. The four most commonly used force fields in structural biology are OPLS, CHARMM, AMBER, and GROMOS [141, 143-147].

Force field descriptions involve potential energy functions due to bonded (E_{bonded}) and non-bonded ($E_{nonbonded}$) interactions between atoms of a system. Bonded interactions describe bond stretching, valence angle bending, and the rotation of dihedral angles. Non-bonded interactions are described by electrostatic and van der Waals interactions.

$$E_{bonded} = \sum_{bonds} K_b \, (b - b_0)^2 + \sum_{angles} K_\theta \, (\theta - \theta_0)^2 + \sum_{dihedrals} K_\chi \, (1 + \cos(n\chi - \sigma))$$

(2.1)

and

$$E_{nonbonded} = \sum_{\substack{nonbonded \\ pairs \\ i \; j}} \varepsilon_{ij} \left(\left[\left(\frac{R_{min,ij}}{r_{ij}} \right)^{12} - 2 \left(\frac{R_{min,ij}}{r_{ij}} \right)^6 \right] + \frac{q_i q_j}{r_{ij}} \right)$$

(2.2)

The total potential energy function is described with:

$$E_{total} = E_{bonded} + E_{nonbonded} + E_{other} \qquad (2.3)$$

Where E_{bonded} represents the total energy contribution from bonded interactions, $E_{nonbonded}$ from non-bonded interactions, and E_{other} describing other terms that may be specific to a force field [148].

The first term in Equation 2.1 resembles Hooke's law for the potential energy of a spring, describing the stretching of bonds in a quadratic form. A sum for all bonded pairs of atoms are represented, where b is the bond length, and K_b and b_0 describe the stiffness and equilibrium length, respectively. The second term describes the bending of the angle (θ) that is formed between triplets of consecutively bonded atoms, with K_θ and θ_0 describing the stiffness and equilibrium geometry of the angle, respectively. The third term is a quadratic equation describing the energetics associated with the rotation of the dihedral angle formed by quadruplets of consecutively bonded atoms. A cosine function is used due to the periodic nature of dihedral rotations, where χ is the value of the dihedral, K_χ forms the energetic parameter determining the barrier height, n is the periodicity, and σ is the phase.

The non-bonded or intermolecular interactions are described by Equation 2.2. The first term is the van der Waals term, representing the attractive and repulsive nature of the atoms. ε_{ij} is the strength of the Lennard-Jones interaction and $R_{min,ij}$ defines the distance at which the Lennard-jones energy is at a minimum, with both these terms based on the types of the two interacting atoms i and j. The interatomic distance is defined with r_{ij}, while q_i and q_j are the parameters describing the effective charges on atoms i and j. These charges are partial atomic charges, representing the overall charge distribution of a molecule.

Differences between commonly used force fields are included into the term E_{other} accordingly. For example, CHARMM includes an additional Urey-Bradley angle term providing additional degrees of freedom for an accurate reproduction of vibrational spectra during parameterisation. Other differences involve the various scaling constants applied to Lennard-Jones and Coulomb interactions between atom pairs applied by each force field [149].

To complete the force field parameters, the potential energy functions discussed above must be accompanied with a set of parameters describing the energetic and geometric properties of interacting particles. The optimisation of force field parameters is important to ensure the simulations replicate experimental data. Data required may involve experimental spectroscopic, thermodynamic, crystallographic, or quantum mechanical methods.

Another aspect of force fields that must be considered in bimolecular simulations is the treatment of solvation. The selection of force field may indirectly determine the water model that is to be used, since most force fields have been developed in conjunction with a specific water model. For example, the TIP3P water model is developed with AMBER, CHARMM, and OPLS force fields, while GROMOS uses the SPC and F3C water models.

2.3.2.2 Equations of motion

In MD, atoms and bonds are considered as solid balls that are connected by springs which oscillate within an optimal distance. MD is based on numerical integration of Newton's law of motion, $F = ma$. Through this, accelerations and velocities are calculated, updating the position of atoms to produce a trajectory that describes their progress over time. The equation of motion is described with:

$$\frac{F_i}{m_i} = \frac{d^2 r_i}{dt^2}$$

(2.4)

where F_i is the force acting on each atom i, m_i is the mass, and r_i is the position. Given the positions and velocities at time t, the idea is to find the positions and velocities at time $t + \delta t$, where δt is the time interval or time step between two snapshots of the simulation. In order to avoid instability, the time step must be smaller the fastest movement in the system, but large enough to avoid unnecessary computation. This step is the major bottleneck of the simulation procedure [142]. Usually, the time step is between 1 and 2 fs for atomistic simulations.

To calculate the positions and velocities at time $t + \delta t$, Equation 2.1 is solved using an integration scheme. Commonly used is the Verlet algorithm, which requires the knowledge of current positions, $r(t)$; acceleration $a(t)$; and the position from the previous step $r(t - \delta t)$ [150]. The position of the next step can be found with:

$$r(t + \delta t) = 2r(t) - r(t - \delta t) + dt^2 a(t) \tag{2.5}$$

While the velocities do not appear in the Verlet integration algorithm, they can be calculated by using $r(t - \delta t)$ and $r(t + \delta t)$ to perform a Taylor series expansion about $r(t)$ to give:

$$v(t) = \frac{r(t + \delta t) - r(t - \delta t)}{2\delta t} \tag{2.6}$$

The advantage of this method lies in that it is straightforward, and is computationally more conservative. However, its precision is compromised, and the lack of an explicit velocity term complicates the calculation of velocities. To address this issue, the original Verlet algorithm was improved on over the years, producing the leap-frog integration technique. In this method, the following relationships are utilised:

$$v\left(t + \tfrac{1}{2}\delta t\right) = v\left(t - \tfrac{1}{2}\delta t\right) + \delta t a(t) \tag{2.7}$$

$$r(t + \delta t) = r(t) + \delta t v\left(t + \tfrac{1}{2}\delta t\right) \tag{2.8}$$

The velocities $v\left(t + \tfrac{1}{2}\delta t\right)$ are initially calculated with velocities at time $t - \tfrac{1}{2}\delta t$ and accelerations at time t as shown in Equation 2.7. Positions $r(t + \delta t)$ are subsequently deduced from the velocities with positions at time $r(t)$ with Equation 2.8. Velocities are computed mid-step with:

$$v(t) = \tfrac{1}{2}\left[v\left(t + \tfrac{1}{2}\delta t\right) + v\left(t - \tfrac{1}{2}\delta t\right)\right] \tag{2.6}$$

Thus, the velocity of the particle is calculated at $t + \tfrac{1}{2}\delta t$ such that it "leaps-over" its position. This enables the velocities to be explicitly calculated, however they are not calculated at the same time as their positions.

2.3.2.3 Periodic boundary conditions

The use of period boundary conditions enables a simulation to be performed using a relatively small number of particles, such that the forces experienced by the particles were as if they were in bulk fluid. For example, a cubic box of particles is replicated in all directions to produce a periodic array. During the simulation, if a particle were to leave the box it would be replaced by an image particle from the opposite side, ensuring a constant number of

particles present within the central box [151]. Depending on the system to be studied, there are several types of space-filling boxes that may be more computationally efficient. Examples include the cubic unit cell, rhombic dodecahedron, and truncated octahedron.

2.3.2.4 Non-bonded interactions

One of the most computationally expensive parts of the simulation is the calculation of non-bonded energies and forces. A common approach is the minimum image convention, which considers the nearest image of each particle for the calculation of short-range non-bonded interaction terms [152]. Another method commonly used is applying the non-bonded cut-off, or potential truncation method. This ignores interactions between pairs of atoms that are further apart than the cut-off, since the greatest contribution of forces comes from neighbouring particles. However, this method can become computationally expensive and erroneous as the size of the system increases.

To address some of these issues, Ewald sums treat the system as infinitely periodic and calculates interaction energies and forces between the cut-off length with the Ewald algorithm [153]. While original Ewald sums are computationally expensive, further improvements have been made such as the Particle-Mesh Ewald (PME) method [154]. This enables rigorous treatment of long-range electrostatics with greater computational efficiency [155]

2.3.3 Thermodynamic ensembles

During MD simulations, it is necessary to control the temperature and pressure of the system. This may be due to factors such as solute drift and friction of particles within the unit cell that may introduce thermodynamic errors. Thus, it may be desired for simulations to be performed in different ensembles, such as the canonical (NVT) and isothermal-isobaric ensemble (NPT), such that simulation data may be more relevant to experimental data.

2.3.3.1 Temperature coupling

Commonly used methods for controlling the temperature of the system include the Berendsen and Nosé-Hoover schemes [156-158]. The Berendsen temperature coupling scheme replicates weak coupling with first-order kinetics to an external heat bath. Deviations from a given temperature T_0 are corrected with:

$$\frac{dT}{dt} = \frac{T_0 - T}{\tau}$$

(2.7)

where a deviation in temperature decays exponentially with a time constant τ. The velocities of each particle is scaled at every time step with a time-dependent factor λ:

$$\lambda = \left[1 + \frac{\Delta t}{\tau_T} \left\{ \frac{T_0}{T\left(t - \frac{\Delta t}{2}\right)} - 1 \right\} \right]^2$$

(2.8)

where τ_T is the scaling time constant, which is close to the temperature coupling time constant. The change in temperature is less than the scaling energy, as the kinetic energy change is partially redistributed between the kinetic energy and potential energy. The Berendsen temperature coupling method is efficient for stabilising a system at a target temperature.

The Nosé-Hoover temperature coupling algorithm improves descriptions for canonical ensemble simulations. This approach was first proposed by Nosé, and later modified by Hoover [157, 158]. The Nosé-Hoover method is an extension of Berendsen, introducing a thermal reservoir in addition to a friction constant ξ. Thus, the equation of motion of particles from Equation 2.4 is rewritten as:

$$\frac{d^2 \boldsymbol{r}_i}{dt^2} = \frac{\boldsymbol{F}_i}{m_i} - \xi \frac{d\boldsymbol{r}_i}{dt}$$

(2.9)

where the motion for ξ is given by:

$$\frac{d\xi}{dt} = \frac{1}{Q} (T - T_0)$$

(2.10)

The reference temperature is denoted by T_0, while T is the current instantaneous temperature of the system, and Q is the coupling strength. The main difference between the Berendsen and Nosé-Hoover approaches is that the Berendsen algorithm more rapidly reaches the target temperature and relaxes the system, while the Nosé-Hoover method takes a longer amount of time and oscillates around the required ensemble.

2.3.3.2 Pressure coupling

Pressure coupling may also be necessary to correctly represent a molecular system. Two of the most common approaches are the Berendsen barometer and the Parrinello-Rahman algorithm [156, 159]. These methods can be combined with previously described temperature coupling methods. In the Berendsen pressure coupling algorithm, the change in pressure towards a given reference temperature $\boldsymbol{P}_0$ is:

$$\frac{d\boldsymbol{P}}{dt} = \frac{\boldsymbol{P}_0 - \boldsymbol{P}}{\tau_p} \tag{2.11}$$

With this method, the box length is rescaled at each time step by a scaling matrix μ_0 as the pressure moves the system towards P_0, and each element of μ_0 is given as:

$$\boldsymbol{\mu_0} = 1 - \beta \frac{\Delta t}{3\tau_p} [\boldsymbol{P}_0 - \boldsymbol{P}] \tag{2.12}$$

where β is the isothermal compressibility of the system, equivalent to 4.6×10^{-5} bar^{-1} for water at 1 atm and 300 K.

The Parrinello-Rahman algorithm extends the Nosé-Hoover temperature coupling algorithm to provide a more accurate description of pressure coupling within the system. Box vectors are given by the matrix $\boldsymbol{b}$:

$$\frac{d\boldsymbol{b}^2}{dt^2} = V\boldsymbol{W}^{-1}\boldsymbol{b}'^{-1}\left(\boldsymbol{P} - \boldsymbol{P}_{ref}\right) \tag{2.13}$$

where V represents the volume of the box, and $\boldsymbol{W}$ representing a matrix parameter determining coupling strength. $\boldsymbol{P}$ and $\boldsymbol{P}_{ref}$ denote the current and reference pressures, respectively. The equations of motion are also altered accordingly.

$$\frac{d^2\boldsymbol{r}_i}{dt^2} = \frac{\boldsymbol{F}_i}{m_i} - \boldsymbol{M}\frac{d\boldsymbol{r}_i}{dt} \tag{2.14}$$

$$\boldsymbol{M} = \boldsymbol{b}^{-1}\left[\boldsymbol{b}\frac{d\boldsymbol{b}'}{dt} + \frac{d\boldsymbol{b}}{dt}\boldsymbol{b}'\right]\boldsymbol{b}'^{-1} \tag{2.15}$$

A combination of the Parrinello-Rahman algorithm with the Nosé-Hoover ensemble allows an accurate isothermal-isobaric ensemble to be generated, with a disadvantage of being slower to produce a desired ensemble. The Berendsen coupling method has the advantage of faster kinetics, with a possible disadvantage of reduced accuracy.

2.4 Analyses of MDS trajectories

MD simulations generate a wealth of data. This section describes the applications of trajectory data generated from the MD simulations that were utilised in this thesis. In Chapter 6 following classical MD simulations, the motions of the COX proteins were explored using essential dynamics, and the protein communication network examined using MONETA. The individual residue contributions to binding of these olive ligands were studied using MM-PBSA techniques. Also utilising molecular dynamics methodologies, the permeation of olive ligands through the lipid membrane were studied as a complement to ADMET data in Chapter 5. This involved performing steered molecular dynamics simulations, followed by umbrella sampling. These techniques will be discussed in this section.

2.4.1   Essential dynamics

Some features of protein function can only be understood when dynamic properties are considered. For example, the diffusion of small substrates through haem-dependent enzymes depends on the transient appearance of channels in the protein structure [142, 160, 161]. Knowledge of dynamics may be necessary to describe motions involved in relevant conformational states of functional importance. Extracting information from a trajectory where a range of conformations are explored can lead to a better understanding of the biological function of proteins. Hence, principal components analysis (PCA) has commonly become employed to determine the most important motions in proteins.

PCA is a multivariate statistical technique that reduces the number of dimensions required to describe protein dynamics by filtering observed motions from the largest to smallest spatial scales [162-165]. PCA extracts important elements of data using either a covariance or correlation matrix, which is constructed from atomic coordinates describing atomic displacements that comprise a trajectory [166]. An eigenvalue decomposition is performed on the matrix, producing a complete set of eigenvectors and corresponding eigenvalues that describe a portion of the motion, with larger eigenvalues denoting motions of a larger spatial scale [163]. Applying PCA to a protein trajectory is termed 'essential dynamics', since the most essential motions are extracted from a set of sampled conformations [167]. In Chapter 6,

essential dynamics was examined using g_covar and g_anaeig tools implemented in GROMACS.

2.4.2 Network analysis

Allostery is one of the ways in which protein function is modified. Allosteric events perturb the protein structure, propagating energetic strain that can lead to the shape and properties of binding sites to change. Using conformational ensembles computed from MD trajectories, data was analysed and extracted to determine residue interactions and dynamical correlations between residues or domains. This was performed using the MONETA program package [168]. MONETA was developed to identify clusters of locally coupled residues, and chains of non-covalently connected residues displaying concerted motions at long range.

Similar to essential dynamics mentioned above, the identification of clusters of locally coupled residues involves PCA. These regions, called independent dynamic segments (IDSs), are clusters of residues in which highly concerted atomic fluctuations are observed independent from the rest of the protein. To identify the IDSs, the statistical technique Local Feature Analysis (LFA) is adapted for analysing fluctuations in atomic coordinates in MD simulations [169, 170]. Originally developed for image analysis, LFA extracts local outputs of reduced dimensionality from global PCA modes.

Communication pathways or chains of residues with high communication propensities are also identified with this analysis are also identified. These are defined based on a measure of communication propensity [171]. The communication between two residues is estimated as fast when their commute time $CT(i,j)$ is small. The commute time is expressed as a variance of inter-residue distance over the trajectory [172]:

$$CT(i,j) = \langle \left(d_{ij} - \bar{d}_{ij} \right)^2 \rangle \tag{2.16}$$

where $d_{ij} = |r_i - r_j|$ represents the distance between Cα atoms of residues i and j.

2.4.3 MM-PBSA: Per-residue energy decomposition

Per-residue decomposition analysis was performed to obtain a quantitative description of the energetic contribution of each amino acid with the ligands considered in the study. As evident from previous studies, binding free energy calculations and its per-residue decomposition

methods are crucial for determining the binding mechanisms of protein–protein or protein–ligand complexes [116, 173].

The molecular mechanics Poisson-Boltzmann surface area (MM-PBSA) method is a technique which evaluates the standard free energies or binding free energies of complexes by considering the initial and final sates of the system [174, 175]. Using data from a MD simulation where the protein complex has been simulated in a periodic water box with water and counterions, the structures are post-processed removing solvent and counterion molecules. The free energy G is calculated according to:

$$\Delta G = \Delta E_{MM} + \Delta G_{solvation} - TS_{MM}$$ (2.17)

where ΔG is the calculated average free energy, and ΔE_{MM} is the average molecular mechanical energy, $\Delta G_{solvation}$ is the molecular solvation free energy, and TS_{MM} is the solute entropy. The average molecular mechanical energy is defined as:

$$E_{MM} = E_{bond} + E_{angle} + E_{tors} + E_{vdw} + E_{elec}$$ (2.18)

where E_{bond}, E_{angle}, E_{tors}, E_{vdw}, and E_{elec} represent the bond, angle, torsion, van der Waals, and electrostatic terms in the molecular mechanical force field, respectively. While MM-PBSA is computationally efficient to perform, limitations of MM-PBSA must be considered. These include the lack of consideration for specific water interactions, and its sensitivity to the trajectory and induced fit effects [174]. Despite these uncertainties, MM-PBSA can provide useful estimates of binding free energy and has been applied to evaluate the relative stabilities of different biomolecular structures [176-179].

In Chapter 6, the g_mmpbsa tool implemented with GROMACS was used to calculate the contribution of individual residues in COX to ligand binding affinity [180]. This tool utilises Adaptive Poisson-Boltzmann Solver (APBS) packages to calculate enthalpic components of the interaction [181].

The free energy of solvation refers to the energy required for a solute to be transferred from vacuum into solvent. The MM-PBSA model calculates this using an implicit solvent model [181], expressed as:

$$G_{solvation} = G_{polar} + G_{nonpolar} \tag{2.19}$$

where G_{polar} and $G_{nonpolar}$ refer to electrostatic and non-electrostatic contributions, respectively. G_{polar} is estimated by solving the Poisson-Boltzmann equation [181]

$$\nabla \cdot [\varepsilon(r) \nabla \cdot \varphi(r)] - \varepsilon(r) \kappa(r)^2 \sinh[\varphi(r)] + \frac{4\pi \rho^f(r)}{kT} = 0 \tag{2.20}$$

where $\varphi(r)$ denotes the electrostatic potential, $\varepsilon(r)$ the dielectric constant, and $\rho^f(r)$ is the fixed charge density. The term κ^2 refers to the reciprocal of the Debye length, which is dependent on the ionic strength of the solution. The non-electrostatic term of solvation free energy $G_{nonpolar}$ incorporates attractive and repulsive forces between the solvent and solute generated by cavity formation, as well as van der Waals interactions [182]

$$G_{nonpolar} = G_{cavity} + G_{vdw} \tag{2.21}$$

where G_{cavity} is the work done by the solute to create a cavity in the solvent, and G_{vdw} is the attractive van der Waals energy between solvent and solute. A range of models can be used to estimate these terms, such as solvent accessible surface area (SASA) only, solvent accessible volume (SAV) only, SASA-SAV model, and combining these models with the Weeks-Chandler-Andersen (WCA) theory in the SASA-SAV-WCA non-polar model [180]. The g_mmpbsa tool enables decomposition of the binding energy on a per residue basis. The E_{MM}, G_{polar}, and $G_{nonpolar}$ terms are initially calculated for individual atoms in the bound and unbound forms, and their subsequent contribution to the binding energy ΔR_x^{BE} of residue x is calculated with:

$$\Delta R_x^{BE} = \sum_{i=0}^{n} \left(A_i^{bound} - A_i^{free} \right) \tag{2.22}$$

where A_i^{bound} and A_i^{free} represent the energy of atom i from x residue in bound and unbound forms, respectively, and n describes the total number of atoms in the residue. The sum of energy contributions over all residues is equivalent to the total binding energy of the complex.

2.4.4 Umbrella sampling

Predicting the rate of permeation of substances across cell membranes is important to drug design. To estimate the permeability of olive compounds to a lipid bilayer, MDS was used to estimate the free energy profile of the system around local minima, or the potential of mean force (PMF). However, for biological processes such as protein binding or membrane permeation the various minima can be separated by large energy barriers. This results in a failure of regular MD to adequately describe the energy surface. One method of obtaining sufficient sampling is umbrella sampling [183]. This introduces a biasing potential $W(x^N)$ to ensure membrane permeation is sampled along the reaction coordinate. Using an example of two interacting particles, the biasing potential would force prescribed separation distances to be sampled.

$$V'(x^N) = V(x^N) + W(x^N) \qquad (2.23)$$

Several simulations are carried out along the reaction coordinate by applying the bias $W(x^N)$. Each new position is called a 'window'. The particle at the reaction coordinate location of each specified window is often restrained with a harmonic potential. From the probability distributions calculated from MD, the effects of the restraints are analytically removed, then combined into a single PMF describing the complete interval of the coordinate using post-treatment analysis [107]. This can be done using the Weighted Histogram Analysis Method (WHAM) [184].

3. Olive Phenolic library

Manuscript published in association with this chapter:

Bonvino NP, Liang J, McCord ED, Zafiris E, Benetti N, Ray NB, Hung A, Boskou D, Karagiannis TC, *OliveNet™: A Comprehensive Library of Compounds from Olea europaea*. Database, 2018. 2018: p.bay106-bay106.

3.1 Introduction

Phenolic compounds derived from *Olea europaea* have demonstrated potent antioxidant and anti-inflammatory properties which may aid in the prevention of several chronic disease states. The characterisation and biological validation of compounds within the olive are being increasingly researched. However, the diversity and complexity of structures means their identification and quantification is challenging. Consequently, only a subset of phenolic compounds has been explored for biological activity and potential health effects. Although information describing the identification of olive-derived compounds is available, these are not easily searchable.

Therefore, as part of a wider project we developed a database of all compounds found in the olive: OliveNet. Data from published reports concerned with the identification and biological effects of compounds in *O. europaea* was sourced and critically assessed to construct the database. OliveNet consists of compounds divided into 13 main classes, including olive phenolics. For this study, compounds identified within this class were used to inform initial *in silico* analyses. Compounds found within the olive fruit, leaf, and pressed oil were obtained, as well as those found in wastewater and pomace produced during olive oil production.

The OliveNet library is freely available at www.mccordresearch.com.au [185].

3.2 Methods

Compounds were identified from a comprehensive review of scientific publications, including journal articles and books. The literature search was performed in PubMed and SciDirect from January—July 2016, using the search terms ("olea europeaea", "olive", "phenol", "polyphenol").

Original publications were selected based on the analytical methods used to identify and quantify the compounds present in the natural olive matrices. These involved a range of extraction processes, analytical separation, and quantification techniques. Generally, HPLC coupled with mass spectrometry/gas chromatography (GC-MS) was employed to separate and then quantify the unsaponificable compounds, including the phenolics [186-188]. High resolution multinuclear (1H, 13C, 31P) NMR was also used for elucidation of isolated compounds [189-191]. These techniques represent a higher sensitivity compared to other spectrophotometric techniques that have several limitations associated with their application [192]. If the methodology was not sufficiently documented or was considered inadequate, the paper was not included as a seminal reference. Information regarding known biological and pharmacological activity of the compounds was included.

Compound structures were obtained from PubChem where available, or if unavailable, drawn in ChemBioDraw Ultra 14.0 (PerkinElmer, Massachusetts, USA) as provided in the seminal reference [193]. All structures were saved as sdf files.

Resulting compounds were compiled and classified into classes and subclasses according to their chemical structure. A graph depicting phenolic compounds within their subclasses was drawn as a network graph using Gephi 0.9.1 [194]. Gephi is an open source software package used to explore and manipulate networks for graph and network analysis [194]. A graph dataset was required for the generation of a network graph. This was constructed by defining subclasses and individual compounds as 'nodes', and compounds within their subclasses as 'edges'. Node and edge matrices were constructed and imported into Gephi to generate the network graph.

3.3 Results and discussion

As a result of the comprehensive review, a total of 222 phenolic compounds were identified. Structurally, phenolic compounds are characterised by an aromatic ring with one or more hydroxyl groups. These were divided into 13 subclasses: simple phenols, methoxyphenols, hydroxybenzoic acids, hydroxyphenylacetic acids, hydroxycinnamic acids, secoiridoids, glucosides, flavonoids, hydroxyisochromans, coumarins, irridoids, lignans, and phenolic fatty acid esters as shown in Figure 3.1. The full phenolic database is shown in Appendix 9.1.
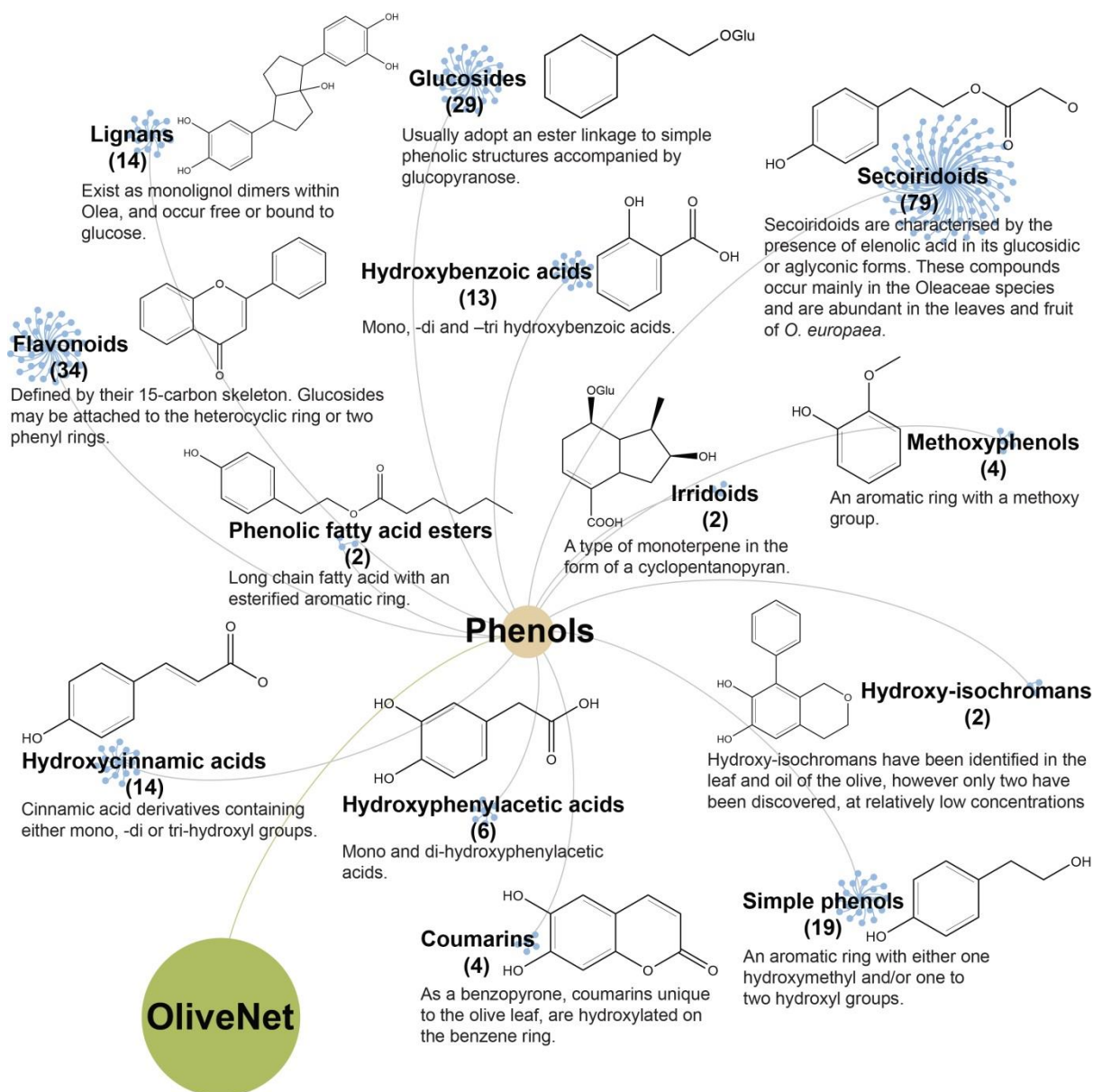
Figure 3.1: Subclasses of phenolic compounds identified in the OliveNet database. 222 phenolic compounds found through literature searches are categorised into 13 subclasses. Each ball is representative of a single compound. Representative chemical structures, key information, and the number of compounds found within each class, are shown.

Phenolic compounds contribute the stability of olive oil through resistance of oxidative rancidity [195]. Shown to be strong antioxidants and radical scavengers, phenolic compounds also possess important biological effects [196]. Oleuropein, hydroxytyrosol, and tyorosol are amongst the most common compounds investigated for their biological activities [196]. Numerous studies have examined their antioxidant properties, which have been mainly attributed to their othodiphenolic structure [196]. These compounds have also been

researched for their microbial properties, with hydroxytyrosol and oleuropein possessing antimicrobial activity against American Type Culture Collection (ATCC) and clinical bacterial strains [196].

Common disease states such as cancer, cardiovascular disease, arthritis, and neurodegenerative disease are known to be associated with chronic inflammation [197-201]. *In vivo* and *in vitro* studied have suggested that dietary intake of EVOO, which contains substantial concentrations of phenolic compounds, may reduce inflammatory responses in the body, thus reducing the risk of developing inflammatory diseases [197, 202-204]. Oleocanthal has previously been found to inhibit both COX-1 and COX-2 in a dose dependent manner similar to ibuprofen [70]. Oleocanthal can also attenuate inflammatory mediators such as inducible nitric oxide synthase (iNOS) which contribute to the pathogenesis of joint degenerative disease [205]. Furthermore, oleocanthal has been found to reduce markers of inflammation in Alzheimer's disease, and possesses anti-proliferative effects in human breast and prostate cancer lines [206, 207].

As a consequence of the variability and complexity of phenolic content among cultivars and limitations associated with analytical methodologies, research has only been focussed on a few compounds. Of the phenolic compounds identified in OliveNet, 45% are currently not commercially available (Table A1.1, Appendix 9.1). The presentation of uncharacterised compounds through the OliveNet library forms the basis for the validation of these compounds in their potential action against various disease states, such as inflammatory diseases.

3.4 Conclusion

In this chapter, a comprehensive literature search was performed to compile a database of olive phenolic compounds, which were classified according to their chemical structures. The OliveNet database forms a useful resource for those conducting research on *O. europaea*, and could form the basis of investigation for the biological effects of compounds or the synthesis of novel compounds. For the purpose of this study, olive phenolic compounds identified here form the basis of *in silico* analysis regarding potential COX inhibitory activity. Molecular docking of olive phenolic compounds with COX-1 and COX-2 proteins will be discussed in the next chapter.

4. Molecular docking

4.1 <u>Introduction</u>

An important tool in structure based drug design is molecular docking. This allows the interactions between a ligand and protein to be modelled at the atomic level, aiding in the characterisation of small molecule behaviour and examination of biochemical processes. To examine the COX inhibitory potential of all olive phenolic ligands identified in the library, molecular docking was employed to determine the binding affinity and orientation of ligands in the COX-1 and COX-2 active site. Using molecular docking, the olive phenolics were screened and ranked to produce lead compounds.

A reliable protein structure is essential for determining atomistic drug interactions. This chapter describes the initial generation of a model of human COX-1, followed by virtual screening of the olive library using a range of docking methods.

4.2 <u>Methods</u>

4.2.1 Homology modelling

Since there is currently no human x-ray crystal structure available for COX-1 in the Protein Data Bank (PDB), the structure was generated using homology modelling techniques. The amino acid sequence of COX-1 was retrieved from UniProt (ID: P23219). The template structure was identified using the blastp (protein-protein BLAST) algorithm, selecting the structure with the highest sequence identity [124]. The selected template structure (1CQE) was of ovine COX-1, had a resolution of 3.1 Å and 93% sequence identity. The homology model of COX-1 was built with Modeller 9.16, using the partial sequence (PRO 32—PRO 583), with ten models generated [122]. The models were evaluated based on the lowest Modeller zDOPE score and RMSD values [122]. The stereochemical quality of the model was then validated using PROCHECK and ProSA [133, 137]. The recent deposit of human COX2 was obtained from PDB, with accession code 5F1A (2.38 Å).

Prior to docking, the two structures were optimized to adopt energetically stable conformations using Schrodinger's Protein Preparation Wizard [208]. It involves the addition

and optimization of hydrogen bonds, termini capping, creation of disulphide bonds, followed by restrained minimization using the OPLS-AA (2005) force field to obtain the optimized geometry of the protein [99].

4.2.2 Ligand preparation

The database of phenolic compounds identified within *Olea europaea* was prepared for docking studies using the LigPrep utility of Schrodinger's package [209]. It generates all possible tautomeric, stereochemical and ionization variants of the input molecules, followed by energy minimization to obtain structures with optimized geometry. Ligands with a molar mass of < 500 g/mol were employed for docking.

We also compiled published inhibition data for putative COX inhibitors where enzyme immunoassay kits were used, and where reported IC_{50} values were comparable among assays. These compounds were intended to aid our docking setting selections following generation of correlation plots between published IC_{50} values and calculated binding energy. These compounds were obtained from PubChem and underwent preparation within the LigPrep utility.

4.2.3 Molecular docking

Ligand-protein docking utilized the quantum mechanics-polarized ligand docking (QMPLD) protocol from the Schrodinger Suite [94]. Receptor grid generation of the enzyme active site was conducted within the docking suite of Maestro [93]. Native ligands were present in the active site of both crystallised structures of COX-1 and COX-2. The native COX-1 ligand flurbiprofen (FLP) present in the template structure was centroid to the docking box. The active site of COX-2 was centred on active site residues Tyr385, Ser530, Arg120, and Tyr355, since the native bound salicylate (SAL) was too small to adequately dock larger ligands. An extensive search on the centroid box was performed within 2 x 2 x 2 nm^3 of these coordinates. Flexibility of the hydroxyl groups of residues at the binding site was allowed, while the rest of the protein was rigid. Docking was carried out using the extra precision (XP) docking protocol of Glide. After initial docking in XP mode, the atomic charges of the ligand were calculated in the binding pocket with the QMPLD workflow. The new charges were generated using semi-empirical methods based on the Mulliken charge model [210]. A maximum of 5 poses per ligand was selected for redocking in XP mode, and final pose selection was based on van der Waals and electrostatic interactions

The R^2 values reported in Figure 4.2 were determined by linear regression using GraphPad Prism 6.0.

The root mean square deviation (RMSD) of the native and docked ligands was measured in Maestro using the superposition feature, whilst the ligand interaction tool was used to analyse ligand-residue contacts. Figures were generated using VMD 1.9.1 [211]. All computation works were performed on a Windows 7 workstation equipped with an Intel Core i5 (3.00GHz) and 8GB of RAM.

4.3 <u>Results and discussion</u>

4.3.1   *Homology model generation and quality*

While the ovine x-ray crystal structure of COX-1 has been available since 1994, the crystal structure of human COX-1 (hCOX-1) is not yet available [212]. PDB structure 1CQE (ovine ortholog) was selected as the template for the generation of hCOX-1, since it had the highest sequence identity (93%) following the BLAST-protein search. Most residues were conserved between the template and target sequence following alignment. The stereochemical quality of the model was examined using Procheck (Table 4.1), which demonstrated 92.6% of residues were in the most favoured regions, while 7.4% were within the allowed regions of the Ramachandran plot. This analysis suggested that the quality of the model was comparable to refined structures, where more than 90% of residues in the most favoured regions deemed a reliable model [213].

Table 4.1: Statistical values of non-glycine and non proline residues in Ramachandran plot obtained from Procheck

Residues in most favoured regions	92.6%
Residues in allowed regions	7.4%
Residues in generously allowed regions	0.00%
Resides in disallowed regions	0.00%
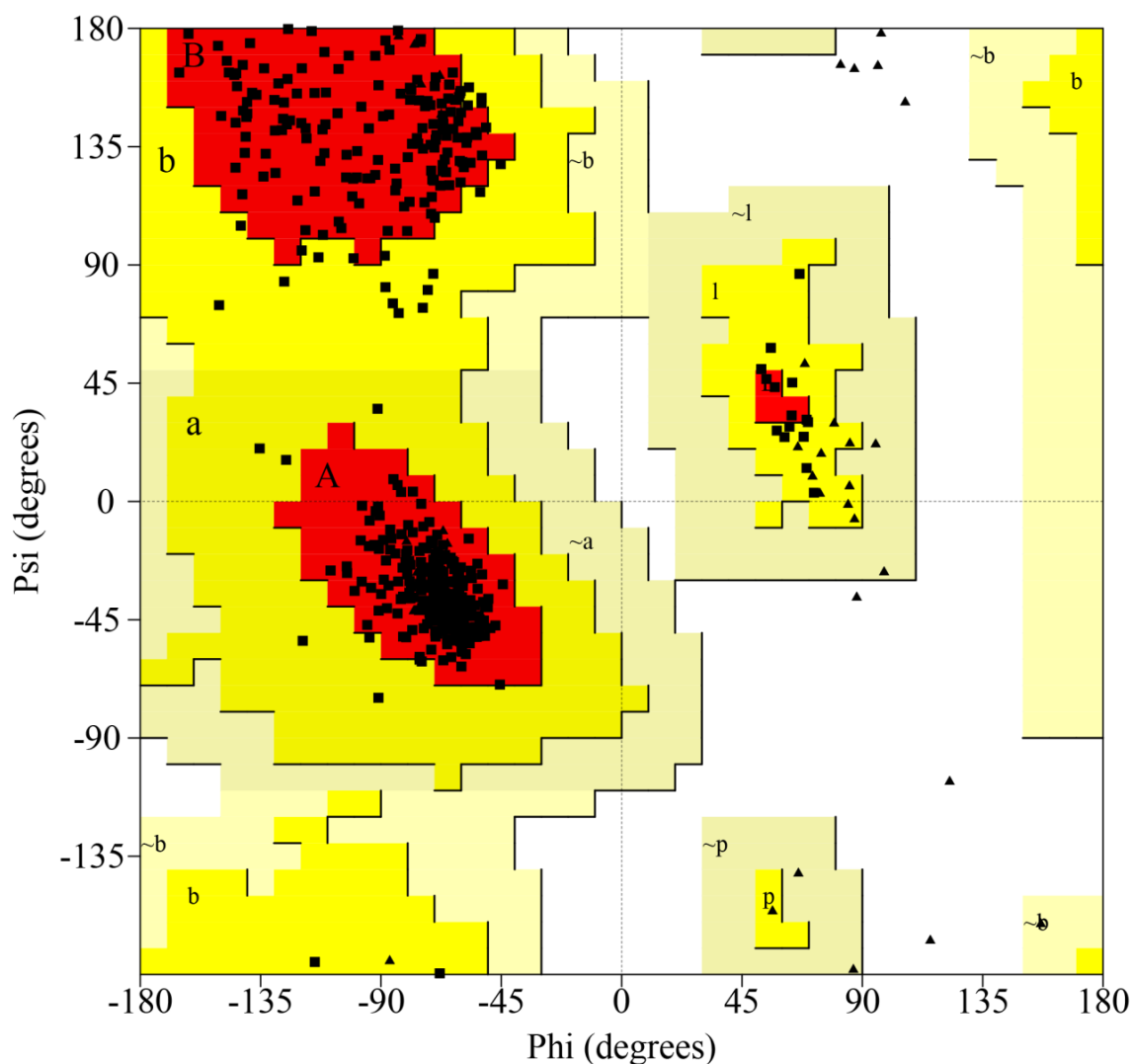
Figure 4.1: Validation of homology model structure using Procheck. Ramachandran plot shows structural validation of COX-1 homology model, with the most favoured regions, allowed regions and generously allowed regions depicted in red, yellow and light brown, respectively. Glycine residues are shown as triangles.

4.3.2   Optimization of docking protocol

In order to optimise the docking protocol, a reference set of published inhibition data for COX-1 and COX-2 inhibitors was assembled (Appendix 9.2.1). Data using enzyme immunoassay kits was selected to limit variability in data. Correlation plots comparing published IC_{50} values against predicted energy of interaction with the docking program were generated. Several docking methods were tested, and the quantum mechanics-polarised ligand (QMPLD) protocol provided the best correlation between these two sets of data [94]. R^2 values of 0.6048 and 0.6852 for COX-1 and COX-2 respectively were produced (Figure 4.2). This method has been used in previous studies [214].

Additionally, the QMPLD docking protocol reliably predicted the experimentally observed native ligands. RMSD values of less than 2 Å are generally considered acceptable [96]. The RMSD values between the native and docked FLP ligand was 0.51 Å for COX-1, whilst salicylate had an RMSD of 0.48 Å for COX-2 (Figure 4.3).
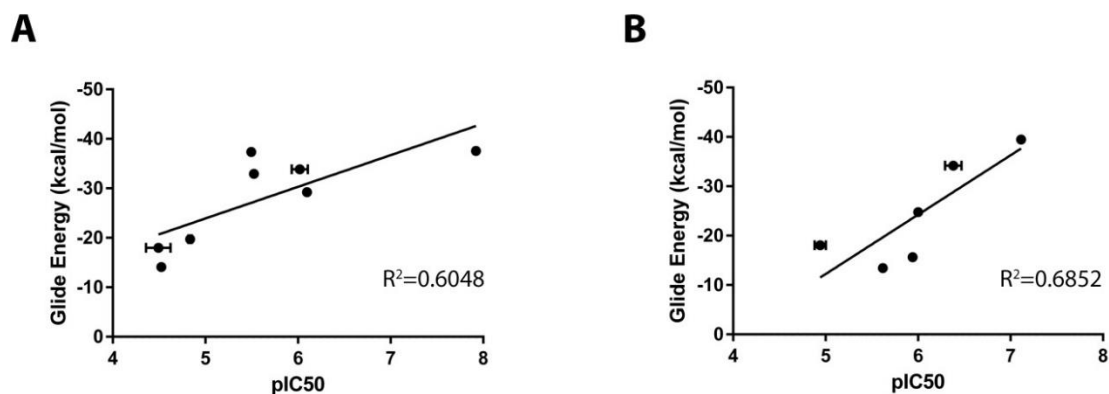
Figure 4.2: Standard curve comparing experimental pIC_{50} values to Glide energies predicted by QMPLD for COX-1 (A) and COX-2 (B). Error bars indicate the standard deviation of pIC_{50} values for experimental binding affinities.
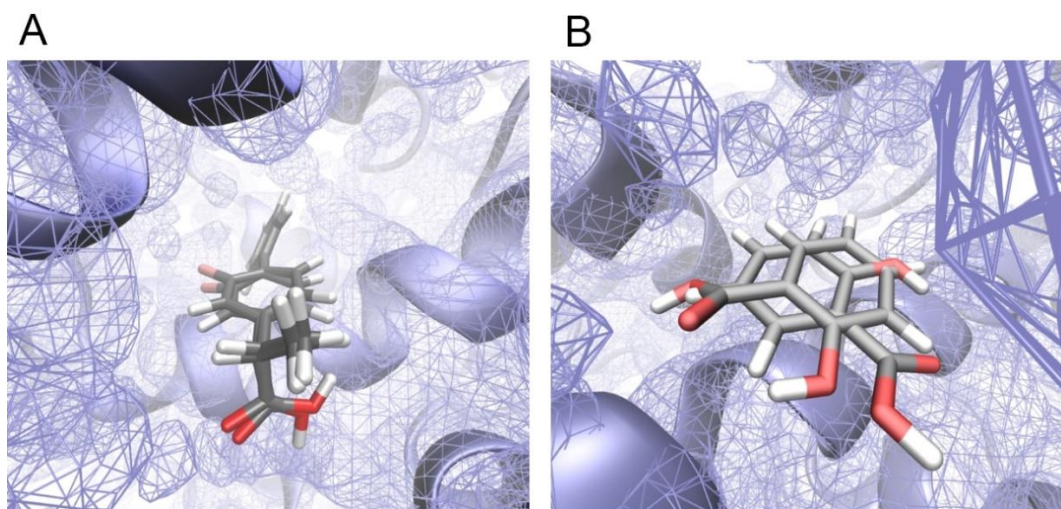
Figure 4.3: Native and docked complex of COX-1 (A) with FLP and COX-2 (B) with salicylate. The molecular surface of active site residues is displayed with a mesh representation.

4.3.3   Predicted top binding compounds

Docking was performed using the QMPLD method. 159 phenolic compounds with a molecular mass of less than 500 g/mol identified in *olea* were docked to COX-1 and COX-2. This was due to the likelihood that larger compounds, such as oleuropein, are largely metabolised following ingestion [215]. Of these 159 compounds, 155 docked to COX-1 and 132 to COX-2. Of the 13 chemical classes in which the compounds are classified, six classes contained ligands that are amongst the top ten in predicted binding energy. Some compounds

produced comparable docking affinities to some putative COX inhibitors (Table 4.2 and 4.3). These six classes were: phenolic fatty acid esters, secoiridoids, hydroxycinnamic acids, flavonoids, glucosides, and lignans (Figure 4.4).
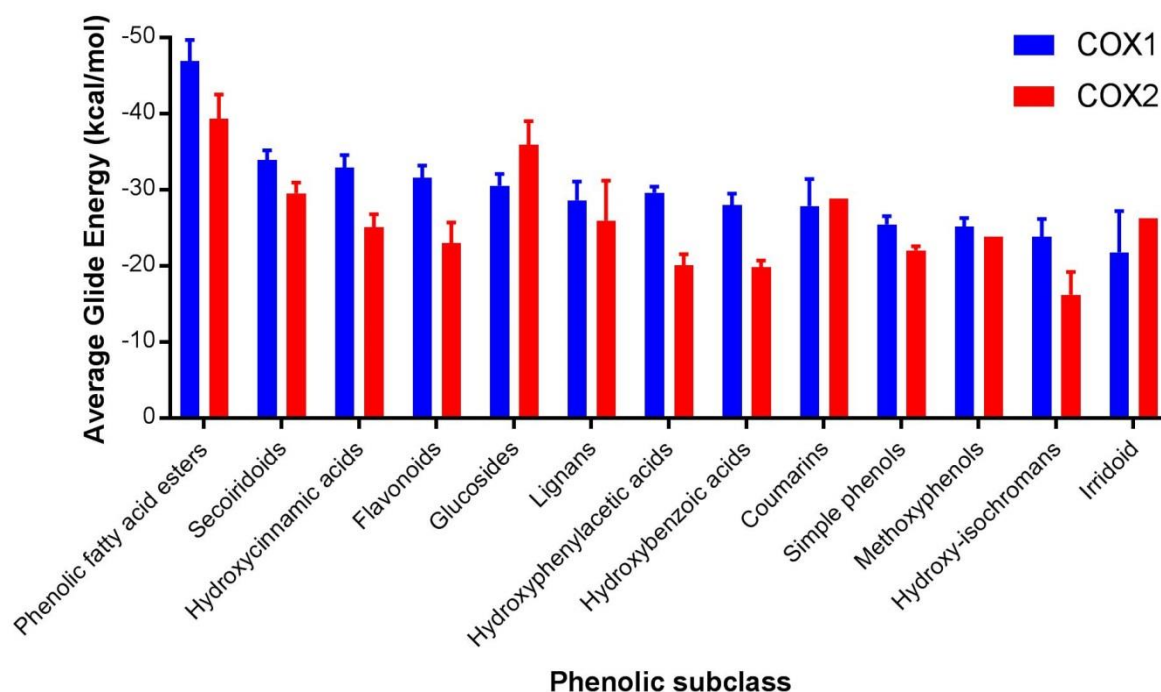
Figure 4.4: Average Glide energy of olive compounds in phenolic subclasses docked to COX-1 and COX-2. Error bars indicate standard error.
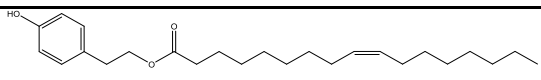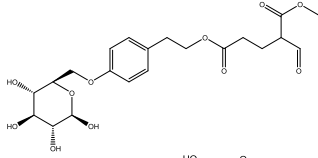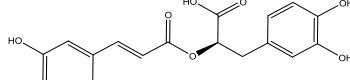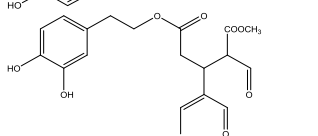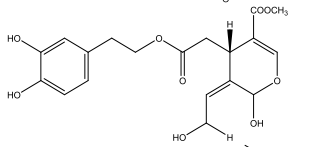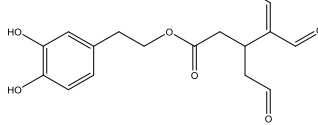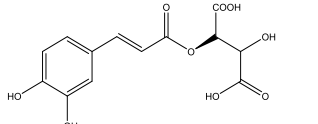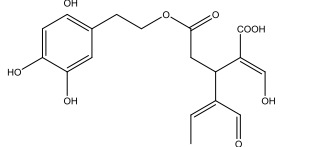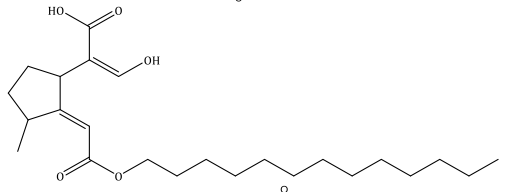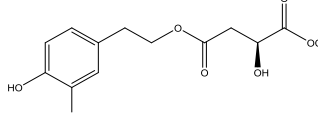
Tables 4.2 and 4.3 list the top 10 binding compounds with the strongest Glide energy to COX-1 and COX-2 respectively. The full table of docking results is shown in Appendix 9.2.2. Figure 4.5 depicts the 2D ligand interaction diagrams of the two strongest binding ligands to COX-1 and COX-2 (A-D), as well as oleocanthal (E-F) based on its known COX inhibitory behaviour [70].

For COX-1, phenolic fatty acid ester is the top binding compound. This may be owing to its long unsaturated tail, bearing resemblance to the native fatty acid substrates of COX. Secoiridoids are vastly over-represented, accounting for six out of the top ten ligands. Hydroxycinnamic acids comprise the remaining top binding ligands, which are structurally similar to the secoiridoids with the exception of having a carbonyl (C=O) group instead of an ester.

For COX-2, there is a greater diversity of chemical classes amongst the top ten. While secoirioids were still prominent and 1-oleyltyrosol again was one of the top binding compounds, remaining classes composing the top ligands were glucosides, lignans, and flavonoids. This suggests that COX-2 is more liable to interact with a larger range of olive compounds, while COX-1 is largely dominated by the secoiridoids.

When the top twenty ligands are taken into consideration for binding to COX, secoiridoids were again vastly represented. In COX-1, fourteen ligands were in this class, compared to twelve for COX-2. All top six classes of ligands were represented in the twenty strongest binding ligands to COX-2, while only four in COX-1. When the strongest binding compounds were extended to include the top thirty, it was apparent that secoiridoids still dominated the majority of strong binding compounds with twenty to COX-1 and nineteen to COX-2. While there more classes of olive phenolics represented in the top thirty binding compounds to COX-1, the number of top binding compounds was more concentrated amongst the top six classes in COX-2. Glucosides and lignans were better represented in top binding to COX-2. Through this it is further demonstrated that there is a trend for COX-2 to bind to a larger range of olive compounds.

Table 4.2: Top 10 greatest binding scores to COX-1

Ligand	Glide energy (kcal/mol)	MW	Class	Compound structure
1-oleyltyrosol	-51	389	Phenolic fatty acid ester	
Ligstroside derivative 2	-48	453	Secoiridoids	
Rosmarinic acid	-47	361	Hydroxycinnamic acids	
Oleuropeindial (keto form)	-45	378	Secoiridoids	
10-Hydroxy oleuropein aglycone	-45	394	Secoiridoids	
Oleacein	-44	321	Secoiridoids	
Caftaric acid	-44	312	Hydroxycinnamic acids	
Demethyloleuropein aglycone (enol form)	-43	364	Secoiridoids	
Deoxyloganic acid lauryl ester	-43	408	Phenolic fatty acid ester	
Methyl malate-hydroxytyrosol ester	-43	284	Secoiridoids	

Note: For comparison, Glide energies for the following ligands are provided as follows: native ligand (FLP = -42 kcal/mol); non-selective NSAIDs (ibuprofen = -33 kcal/mol) (naproxen = -37 kcal/mol); COX-2 selective NSAIDs (rofecoxib = -18 kcal/mol) (celecoxib = -20 kcal/mol).

Table 4.3: Top 10 greatest binding scores to COX-2

Ligand	Glide energy (kcal/mol)	MW	Class	Compound structure
Ligstroside derivative 2	-50	453	Secoiridoids	
1-oleyltyrosol	-44	389	Phenolic fatty acid ester	
Hydroxytyrosol diglucoside	-44	478	Glucosides	
Luteolin-4'-O-glucoside	-42	448	Flavonoids	
10-Hydroxy-10-methyl oleuropein aglycone	-41	408	Secoiridoids	
Oleuropeindial-Lactone	-40	378	Secoiridoids	
Berchemol	-40	376	Lignans	
Hydroxytyrosol-3-β-glucoside	-40	316	Glucosides	
(+)-1-Acetoxypinoresinol-4''-O-methyl ether	-39	431	Lignans	
Oleuropeindial (keto form)	-39	378	Secoiridoids	

Note: For comparison, the Glide energies for the following ligands are provided as follows: native ligand (SAL = -27 kcal/mol); non-selective NSAIDs (ibuprofen = -16 kcal/mol) (naproxen = -18 kcal/mol); COX-2 selective NSAIDs (rofecoxib = -34 kcal/mol) (celecoxib = -39 kcal/mol).

Figure 4.5: 2D Ligand interaction diagram for COX-1 and COX-2 with 1-oleyltyrosol (A, B), ligstroside derivative 2 (C, D), and oleocanthal (E, F) showing protein residues located within 0.4 nm of the ligand. Hydrogen bonds are represented by purple arrows, pi-cation by red lines, and pi-pi stacking by green lines. Residues shown in green are hydrophobic, blue are polar, purple are positively charged, and red are negatively charged. Glycine residues are shown in white. Residue numbering for COX-1 is equivalent to COX-2 when 31 is added.

4.3.4   Overall trends in predicted ligand-COX interactions

The docking demonstrates that overall, the majority of olive phenolics bind with a marginally higher affinity to COX-1 than COX-2 (Figure 4.4). The main residues implicated in binding with COX-1 and COX-2 shown in this data are ARG 120, TYR 355, TYR 385, and SER 530. The location of these residues is depicted in Figure 4.6. Some of these active site residues of the cyclooxygenase proteins have been studied in some detail.
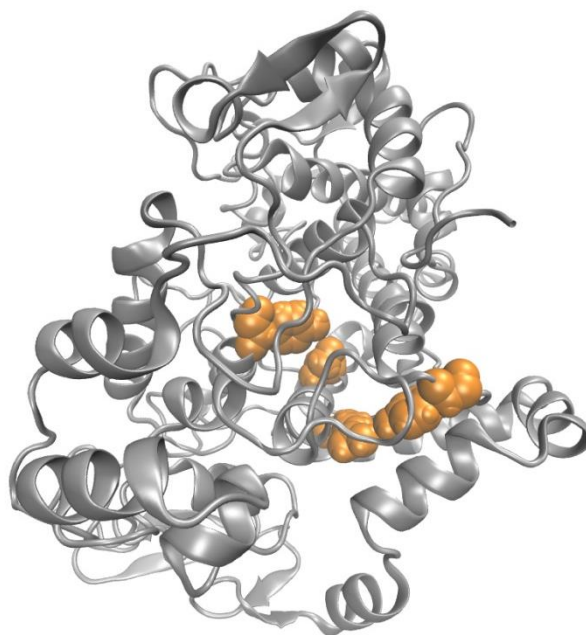
Figure 4.6: 3-dimensional structure COX-2, with active site residues ARG 120, TYR 355, TYR 385, and SER 530 shown in orange.

With its native ligand, the majority of interactions between arachidonic acid and COX-1 involve hydrophobic residues. TYR 385 abstracts the pro*S* hydrogen from C-13 of arachidonate; GLY 533 and TYR 348 orient C-13 of arachidonate for hydrogen abstraction; ARG 120 facilitates the electrostatic interaction with the carboxylate anion of arachidonate; and VAL 349, TRP 387, and LEU 534 dictate the orientation of arachidonate such that an optimal arrangement is achieved to yield PGG$_2$ [216].

In COX-1, ARG 120 is situated near the entrance of the main binding channel within ionic bond distance to its natural substrate arachidonic acid and has a role in catalytic efficiency [217, 218]. In COX-2, ARG 120 has been suggested to have less influence on the chemical kinetic activity of the protein [218]. Its role involves the formation of an ion pair with the carboxylate group of NSAIDs. The mutation of ARG 120 decreases the inhibitory effect of these NSAIDs but increases the potency of inhibitors of COX-1 lacking the carboxylic acid moiety [218, 219].

Located on the opposite side of the channel from ARG 120, TYR 355 dictates the stereospecificity of the enzymes towards NSAIDs [9, 220, 221]. It has also been suggested to contribute to the negative allosteric effect of arachidonate in COX-1 [220, 222]. Acetylation by aspirin occurs at SER 530 [223]. It was previously found that TYR 385 and SER 530 are able to cooperate in the chelation of negative charges, determining the selectivity of aspirin for covalent modification of SER 530 [224]. The mutation of TYR 385 to PHE reduces aspirin acetylation of SER 530 by more than 90% [224, 225]. It was proposed that hydrogen bonding by TYR 385 stabilises the negative charge of the tetrahedral intermediate that is developed during SER 530 acetylation [224, 225]. The crystal structure of diclofenac bound to COX-2 shows that the ligand binds in an inverted conformation compared to arachidonic acid, with its carboxylate group hydrogen bonded to TYR 385 and SER 530, rather than ARG 120 [224].

Phenolic ligands docked to COX-1 demonstrate a pattern of binding to some key binding site residues. The ligands ligstroside derivative 2, rosmarinic acid, oleuropeindial, and 10-hydroxyoleuropein aglycone are ranked 2-5 of the phenolic ligands, forming hydrogen bonds with ARG 120 and MET 522. With the exception of oleuropeindial, these highly ranking ligands also demonstrated hydrogen bond formation with GLY 526. Further, pi-alkyl bonds were formed with VAL 349.

As there was a much greater diversity of ligand classes found in the top ten ranking ligands for COX-2, trends in residue interactions were not as apparent in COX-2 compared to COX-1. The top 5 ranking ligands (ligstroside derivative 2, 1-olyltyrosol, hydroxytyrosol diglucoside, luteolin-4'-O-glucoside, and 10-hydroxy-10-methyl oleuropein aglycone) formed H-bonds with SER 119 and LYS 83 in COX-1. In COX-2, an additional side pocket is present that is not accessible to ligands binding to COX-1. The residues located in this pocket include VAL 523, ARG 513, and VAL 434 [9].

An absence of interactions with these residues among strong binding ligands suggests that olive phenolics may bind to COX proteins non-selectively. In other studies, structure-activity relationships (SAR) have previously found that the length of the S-alkyl chain is a determinant of potency for COX-2 selectivity, with heptyl chains being optimal. Potency and selectivity for COX-2 was enhanced when a triple bond was inserted into the heptyl chain, and also in sulfides compared to corresponding sulfoxides or sulfones [216, 226]. As the top

binding ligands do not tend to display this structural pattern, it is further suggested that olive phenolic ligands are binding non-selectively to COX-1 and COX-2. Nevertheless, 1-oleyltyrosol (with a long hydrophobic tail, similar to those described in the literature mentioned here) is predicted to bind strongly to both COX-1 and COX-2. This suggests that this ligand may be in the vicinity to potentially access the unique COX-2 side pocket. Molecular dynamics simulations may be able to shed light on the motions of this ligand subsequent to binding and assess the ligand's capability to access the COX-2 side pocket.

The following sections will describe docking results of selected ligands from the top docking subclasses of phenols.

4.3.5   Interactions with specific ligand classes

4.3.5.1   Hydroxycinnamic acids

Table 4.4: Ligand interactions with selected hydroxycinnamic acids

Ligand		Glide energy (kcal/mol)	H-bond	Pi-pi stacking	Pi-cation	Salt bridge
Rosmarinic acid	COX-1	-47	ARG 83, THR 89, ARG 120 x2, TYR 355			ARG 120
	COX-2	-34			LYS 83	ARG 120
Caftaric acid	COX-1	-44	ARG 120, SER 353			ARG 120
	COX-2	-27	LYS 83, ARG 120 x2			LYS 83 x2
Chlorogenic acid	COX-1	-34	GLU 524 x2			
	COX-2	-39	GLU 524			ARG 120

Hydroxycinnamic acids demonstrated a preference for binding to COX-1. ARG 120 was shown to form hydrogen bonds and salt bridges with the ligands. Rosmarinic acid demonstrated a strong binding affinity of -47 kcal/mol to COX-1 and -34 kcal/mol to COX-2, forming salt bridges with ARG 120 in both proteins. Caftaric acid had a preference for binding to COX-1 compared to COX-2, with a Glide energy of -44 kcal/mol with COX-1 and -27 with COX-2. While caftaric acid had a higher number of hydrogen bonds and salt bridge formation with COX-2, COX-1 demonstrated a higher binding affinity. This may be due to

differing orientations of the ligand, with its position in COX-1 resulting in hydrophobic interactions with more residues compared to COX-2. In contrast, chlorogenic acid demonstrated a higher affinity for binding with COX-2 compared to COX-1; -39 kcal/mol and -34 kcal/mol respectively. A combination of salt bridge and hydrogen bond formation in COX-2 may have contributed to the stronger binding energy.

4.3.5.2 Flavonoids

Table 4.5: Ligand interactions with selected flavonoids

Ligand		Glide energy (kcal/mol)	H-bond	Pi-pi stacking	Pi-cation	Salt bridge
Methoxyluteolin	COX-1	-38	ARG 120, MET 522			
	COX-2	-16	LYS 83	TYR 355		
Luteolin-4'-O-glucoside	COX-1	-33	ARG 83	TYR 355		
	COX-2	-42	LYS 83, PHE 470	TYR 115		
Delphinidin	COX-1	-31	ARG 120, MET 522			
	COX-2	-32	LYS 83, GLU 524	TYR324 x2	LYS 83, ARG 120	

Like the hydroxycinnamic acids, flavonoids demonstrated a stronger preference for binding to COX-1 compared to COX-2. Methoxyluteolin exhibited a large difference in its preference for COX-1 binding; -38 kcal/mol to COX-1 compared to -16 kcal/mol to COX-2. Pi-pi stacking observed with TYR 355 in COX-2 is absent in COX-1, as well as a hydrogen bond formed with LYS 83. With COX-1, methoxyluteolin forms hydrogen bonds with ARG 120 and MET 522. The interactions with COX-2 are located near the entrance of the active site, with the hydroxyl group forming a H-bond with LYS 83 being exposed to solvent. Methoxyluteolin is positioned deeper in the active site of COX-1, having more hydrophobic interactions with residues within the protein, contributing to its higher binding affinity. Luteolin-4'-O-glucoside exhibited higher binding affinity to COX-2 compared to COX-1. While pi-pi stacking was shown in both interactions, an additional hydrogen bond was observed with COX-2. Delphinidin had a similar binding affinity to both COX-1 and COX-2. While there were more pi-pi stacking and pi-cation interactions with COX-2, a greater amount of hydrophobic interactions were observed with COX-1.

4.3.5.3 Lignans

Table 4.6: Ligand interactions with selected lignans

Ligand		Glide energy	H-bond	Pi-pi stacking	Pi-cation	Salt bridge
D-(+)-Erythro-1-(4-hydroxy-3-methoxy)- 214 -phenyl-1,2,3-propantriol	COX-1	-33	MET 522, SER 530 x2			
	COX-2	-28	SER 530			
Berchemol	COX-1	-31	ARG 120, TYR 355, SER 530		ARG 120	
	COX-2	-40	SER 119	TYR 115	LYS 83	
Pinoresinol	COX-1	-26	ARG 120, SER 530		ARG 120	
	COX-2	-3	ARG 120	TYR 355	LYS 83	

Of the ten lignans, all bound to COX-1 while six bound to COX-2 with varying preference for each isoenzyme. D-(+)-Erythro-1-(4-hydroxy-3-methoxy)-214-phenyl-1,2,3-propantriol bound to both COX-1 and COX-2, with an affinity of -33 and -28 kcal/mol respectively. SER 530 formed hydrogen bonds in interactions with both proteins, with additional hydrogen bonds with SER 530 and MET 522 potentially contributing to a slightly higher binding affinity with COX-1. Berchemol demonstrated a stronger binding affinity to COX-2, having a Glide energy of -40 kcal/mol compared to -31 kcal/mol in COX-1. ARG 120 formed a hydrogen bond and pi-cation interactions in COX-1, while SER 119, TYR 115, and LYS 83 in COX-2 formed a hydrogen bond and pi-pi stacking, as well as pi-cation interactions. Despite a range of molecular interactions of pinoresinol with COX-2, the binding affinity exhibited was much lower compared to COX-1. Pinoresinol bound to COX-1 with an affinity of -26 kcal/mol, compared to -3 kcal/mol in COX-2. ARG 120 was again implicated in these interactions.

4.3.5.4 Glucosides

Table 4.7: Ligand interactions with selected glucosides

Ligand		Glide energy	H-bond	Pi-pi stacking	Pi-cation	Salt bridge
Hydroxytyrosol diglucoside	COX-1	-38	ARG 83, GLU 524 x2			

Ligand		Glide energy	H-bond	Pi-pi stacking	Pi-cation	Salt bridge
	COX-2	-44	LYS 83 x2, PHE 470, GLU 524 x2	TRP 100		
Salidroside	COX-1	-31	ARG 120			
	COX-2	-32	LYS 83, TYR 115, GLU 524		LYS 83	
Hydroxytyrosol-3-β-glucoside	COX-1	-25	TYR 385, MET 522			
	COX-2	-40	LYS 83			ARG 120 x2

All eight glucosides bound to COX-1, while four bound to COX-2 with similar or greater affinity compared to COX-1. Hydroxytyrosol diglucoside bound strongly to both COX-1 and COX-2 with a Glide energy of -38 and -44 kcal/mol respectively. GLU 524 formed two hydrogen bonds in interactions with both proteins. Residue 83, being arginine in COX-1 and lysine in COX-2 formed hydrogen bonds with hydroxytyrosol diglucoside. Interactions with COX-2 exhibited additional hydrogen bonds and pi-pi stacking with TRP 100, contributing to its slightly stronger binding energy. Salidroside interacted with both COX isoenzymes with similar Glide energy; -31 kcal/mol with COX-1 and -32 kcal/mol with COX-2. Differing interactions were observed, with a hydrogen bond with ARG 120 coupled with more hydrophobic residue interactions in COX-1, and a higher number of hydrogen bonds and pi-cation interactions in COX-2. Hydroxytyrosol-3-β-glucoside demonstrated a greater affinity for COX-2; -40 kcal/mol compared to -25 kcal/mol in COX-1. While hydrogen bonds were involved in binding with both enzymes, ARG 120 formed two salt bridge interactions with the ligand only in COX-2.

4.3.5.5 Secoiridoids

Table 4.8: Ligand interactions with selected secoiridoids

Ligand		Glide energy	H-bond	Pi-pi stacking	Pi-cation	Salt bridge
Ligstroside derivative 2	COX-1	-48	ARG 120 x2, MET 522		ARG 120	
	COX-2	-50	LYS 83, ARG 120, TYR 385, MET 522	TYR 355		

Oleocanthal	COX-1	-42		
	COX-2	-30	LYS 83, ARG 120	
Hydroxytyrosol acyclodihydroelenolate	COX-1	-42	ARG 120, GLU 524	
	COX-2	-23	SER 119, ARG 120	
10-Hydroxy-10-methyl oleuropein aglycone	COX-1	-39	GLU 493	
	COX-2	-41	LYS 83	LYS 83

Having the greatest overall binding, a stronger affinity for binding to COX-1 was also observed in the secoiridoids. Ligstroside derivative 2 bound very strongly to both COX-1 and COX-2, producing a binding affinity of -48 kcal/mol and -50 kcal/mol respectively. ARG 120 and MET 522 formed hydrogen bonds with ligstroside derivative 2 in both isoforms of COX. In COX-1, the phenol ring formed pi-cation interactions with ARG 120, while TYR 355 in COX-2 engaged in pi-pi stacking. In both instances, ligstroside derivative 2 was able to penetrate deep into the binding pocket of the COX isoenzymes. Oleocanthal bound with a stronger affinity for COX-1; -42 kcal/mol compared to -30 kcal/mol in COX-2. While there were hydrogen bonds present in binding with COX-2 that were absent in COX-1, oleocanthal was positioned deeper within the binding site where a greater number of hydrophobic interactions with residues occurred. Hydroxytyrosol acyclodihydroelenolate demonstrated a preference for binding to COX-1, producing a binding affinity of -42 kcal/mol compared to -23 kcal/mol in COX-2. ARG 120 formed hydrogen bonds in both enzymes, as well as GLU 524 in COX-1 and SER 119 in COX-2. 10-hydroxy-10-methyl oleuropein aglycone bound to COX-2 with a slightly stronger affinity compared to COX-1, potentially due to the pi-cation interaction formed with LYS 83 in COX-2 which is absent in COX-1.

Secoiridoids produced a majority of the overall greatest binding affinity scores, with ligstroside derivative 2 being the strongest binding ligand to COX-2 and second strongest to COX-1. Ligands were found to bind on the internal surface of the constriction at the base of the funnel-shaped entrance to the COX active site. The constriction is composed of the residues ARG 120, TYR 355, and GLU 524 [9]. In interactions with secoiridoids, the phenol ring engaged in pi-pi stacking and pi-cation interactions with ARG 120 and TYR 355 with both COX isomers (Figure 4.5). These interactions are not consistently present in binding of

the same ligands to COX-2; instead, ARG 120 forms H-bonds with the alkoxy group of the backbone bond structure in COX-2, instead of the carbonyl group in COX-1.

4.3.5.6 *Phenolic fatty acid esters*

Table 4.9: Ligand interactions with phenolic fatty acid esters

Ligand		Glide energy	H-bond	Pi-pi stacking	Pi-cation	Salt bridge
1-oleyltyrosol	COX-1	-51	GLU 524		ARG 120	
	COX-2	-44	SER 119		LYS 83	
Deoxyloganic acid lauryl ester	COX-1	-43	THR 89, ARG 120			
	COX-2	-35	ARG 120			LYS 83

Both phenolic fatty acid esters bound with a high affinity to COX-1 and COX-2. 1-oleyltyrosol bound with the highest affinity to COX-1, producing a Glide energy of -51 kcal/mol in COX-1 and -44 kcal/mol in COX-2. The phenol ring engaged in pi-cation interactions with both enzymes. Deoxyloganic acid lauryl ester bound to COX-1 and COX-2 with a binding affinity of -43 and -35 kcal/mol respectively. ARG 120 formed hydrogen bonds in both interactions, while THR 89 formed a hydrogen bond in COX-1 and LYS 83 formed a salt bridge in COX-2.

4.4 Conclusion

Overall, the olive phenolic compounds suggested non-specific inhibition of COX isoenzymes, with the predicted binding scores suggesting a marginally greater affinity for COX-1. The secoiridoids were the class of phenolic compounds producing the greatest binding affinity scores, with COX-2 suggested to bind to a greater ranger of ligand classes compared to COX-1.

The strongest binding secoiridoid was ligstroside derivative 2, which demonstrated a high binding affinity to both COX-1 and COX-2 (Table 4.2 and 4.3). 1-oleyltyrosol is a phenolic fatty acid, the structure of which bears resemblance to the native fatty acid substrates of COX. Its structure enables binding in a manner similar to other fatty acids such as arachidonic acid, α-linolenic acid, and palmitic acid [227]. Possessing a long chain, 1-oleyltyrosol binding strongly to both COX-1 and COX-2 may suggest that the additional side

pocket in the COX-2 active site could be exploited by this and other similar olive phenolic compounds. While the present docking results did not predict this, other conformations should be explored in future work. For both compounds identified as the strongest binders in this docking study, there is little known about their biological activity. 1-oleyltyrosol was identified to be present in the seed of the olive fruit and leaves [228, 229], and displays weak antibacterial activity [230]. Ligstroside derivative 2 was identified in olive pomace and wastewater [231]. Both these compounds are present in the olive in low concentrations.

Close examination of specific interactions between the top ranked ligands and COX-1 and COX-2 indicated instances where interactions are formed with COX-2 residues that are absent in COX-1, and vice versa. These specific residues and interactions could be further investigated to determine whether they could be exploited to enhance specificity. The docking performed here provides an initial screening step in understanding the inhibitory potential of olive phenolic ligands in COX-1 and COX-2 enzymes. It is possible that physiological interactions between COX enzymes and potential inhibitors may involve kinetic parameters that are not directly apparent from structures, as docking methods utilise a static protein structure [232].

In the landmark paper by Beauchamp et al. found that oleocanthal was able to inhibit COX-1 and COX-2 in a manner similar to ibuprofen [70]. Responsible for the stinging sensation following ingestion of olive oil, this compound has become of interest in naturally occurring compounds with pharmacological properties over the years [233, 234]. Here, oleocanthal was ranked as the 12th strongest binder to COX-1 and 44th strongest to COX-2 (Table A2.3h, Appendix 9.2.2).

From the 159 compounds docked, 1-oleyltyrosol and ligstroside derivative 2, as well as oleocanthal for its prevalence in literature were ultimately chosen for further analysis. The next chapter will detail further screening of the olive phenolic library based on bioavailable characteristics, as well as simulations of compounds of interest to examine membrane permeability. Mechanisms of inhibition of these olive compounds using molecular dynamic simulations will be discussed in Chapter 6.

5. ADMET & preliminary study in membrane permeability

5.1 Introduction

The successful development of a drug relies critically on understanding its pharmacokinetics and potential toxicity. This chapter describes the analysis of absorption, distribution, metabolism, excretion and toxicity (ADMET) properties for compounds in the phenolic library, followed by the selection of the most suitable candidate compounds based on both docking and ADMET scores.

To reach their intended target, most drugs need to pass through at least one cellular membrane. As an adjunct to the membrane permeability of selected compounds was investigated by simulating the process of membrane diffusion at an atomic level using steered molecular dynamics (SMD) simulations. This was followed by umbrella sampling to estimate approximate permeation potentials of mean force (PMF) for the main compounds of interest. These techniques provide insight into the bioavailability of olive phenolic compounds.

5.2 Methods

5.2.1 ADMET

The ADMET properties of each phenolic compound was measured to assess the pharmacokinetic attributes of the compound within the human body using the 'ADMET Descriptors' function within Discovery Studio 4.1 (BIOVIA Discovery Studio, San Diego, US). The module is used to quantitatively predict properties by a set of rules that specify ADMET characteristics. For obtained hits, some important ADMET descriptors—for example: human intestinal absorption, aqueous solubility, blood brain barrier (BBB), plasma protein binding, and hepatotoxicity were calculated. Human intestinal absorption and aqueous

solubility descriptors predict the absorption and solubility, respectively [235, 236]. BBB, plasma protein binding, and hepatotoxicity descriptors predict blood brain penetration of a molecule after oral administration, binding to carrier proteins in the blood, and potential human hepatotoxicity [237, 238]. The compounds that fulfilled the acceptable criteria for these descriptors were subsequently selected for molecular docking studies.

5.2.2 *Steered molecular dynamics and umbrella sampling*

All steered molecular dynamics (SMD) simulations were performed using GROMACS 4.6.5 software package [139, 239]. The membrane bilayer topology was obtained from Lipidbook [240, 241]. The 1,2-dioleyl-sn-glycero-3-phosphocholine (DOPC) membrane consisted of 72 lipids per leaflet and each system was hydrated with ~6100 water molecules. The system corresponded to cell dimensions of $5.2 \times 5.2 \times 11.5$ nm^3. Due to the periodic boundary conditions, the membrane spanned the *x-y*-plane in a continuous manner. Simulations were run using TIP3P water [242] and CHARMM27 lipid parameters [243-245]. Small-molecular parameters were generated using SwissParam [246]. The ligand was positioned above the membrane at a distance equivalent to half the thickness of the membrane (~1.5 nm). MD simulations were performed under the isothermal-isobaric (NPT) ensemble using the GROMACS 4.6.5 package [139, 239] and CHARMM27 force field [243, 245]. A steepest-descent minimisation scheme was initially applied to the system. The temperature was kept stable at 310K using the Nose-Hoover ensemble [157, 158, 247], and the pressure was controlled semi-isotropically by a Parrinello-Rahman barostat [159, 248, 249]. All bond lengths were constrained by LINCS algorithms [250]. Long-range electrostatic forces were evaluated using the particle-mesh Ewald scheme (grid spacing 0.12 nm) [154], while short-range non-bonded interactions were calculated using a cut-off ratio of 1.4 nm for both Coulomb and van der Waals potentials. The spring constant was set to 1000 kJ/mol/nm and a pull rate of 0.01 nm/ps was implemented. The time length for the simulations was 800 ps for olive phenolic ligands, and 1000 ps for H33342 and propidium. The mean force profile for each ligand was calculated by averaging the outcome of 10 independent runs.

The free energy of methyl malate-β-hydroxytyrosol ester and oleocanthal across the bilayer was computed from the potential of mean force (PMF) using umbrella sampling. SMD simulations pulling the molecule through the membrane beginning and ending in the aqueous phase was conducted as described above. 30 configurations were generated along the z-axis direction (reaction coordinate). The z coordinates of centre of mass (COM) distance between

the ligand and membrane in each configuration differed by about 0.1 nm. Each window was equilibrated for 100 ps, and a production run of 10 ns was continued for sampling. The PMF profile was calculated using the Weighted Histogram Analysis Method (WHAM) [184], implemented in GROMACS package as 'g_wham' [251].

5.3 Results and Discussion

5.3.1 Phenolic ADMET

The results presented in Figure 5.1 describe the ADMET properties of all olive phenolic compounds in terms of their ability to be absorbed into the gastrointestinal tract, and their blood brain barrier penetration. These parameters were calculated based on their logP value and polar surface area. LogP describes the logarithm of the octanol/water partition coefficient, providing a measure of the lipophilicity of a compound. Lipophilicity is one of the most critical parameters for passive membrane partitioning, with an increased logP enhancing permeability [252]. The logP value of a compound also has importance in the prediction of biological activities, and toxicological end points [253].
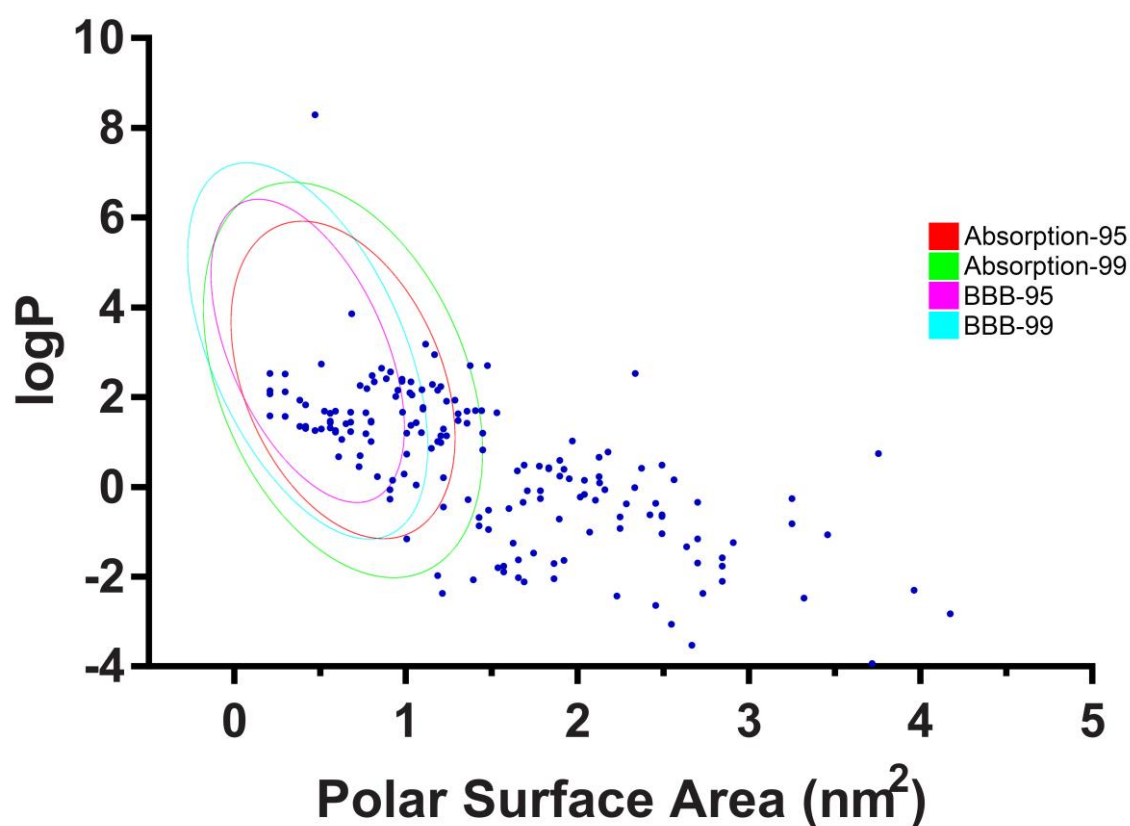
Figure 5.1: Polar Surface Area (PSA) vs logP for olive phenolics showing 95% and 99% confidence limits, denoted by ellipses corresponding to blood brain barrier (BBB) and intestinal absorption models.

Of the 222 phenolic compounds analysed, 107 were found to fall within in 99% confidence limit for intestinal absorption, and this fell to 100 for the 95% confidence limit. For BBB penetration, 88 were within the 99% confidence limit, and 68 within the 95% confidence limit. Other biological parameters relating to ADMET were also calculated with Discovery Studio. 214 were found to not be inhibitors of cytochrome P450 2D6, 154 to be non-hepatotoxic, and 192 to be low binders of plasma protein.

In the search for a novel compound with potential COX inhibitory activity, the results from docking and ADMET were combined. A list of strong binders which were ADMET approved was curated. Compounds were considered ADMET acceptable if they fell within both 99 and 95 confidence limits for intestinal absorption. The list was further refined based on commercial availability. Oleocanthal (OLEO) was chosen as a point of comparison based on its known ibuprofen-like activity and prevalence in literature [70]. Methyl malate-β-hydroxytyrosol ester (MMHTE) was found to be the only compound that bound with a strong affinity to both COX-1 and COX-2 compared to OLEO, as well as fulfilling ADMET criteria and being non-commercially available (Table A1.1, Appendix 9.1).

Table 5.1: Docking summary for MMHTE and OLEO
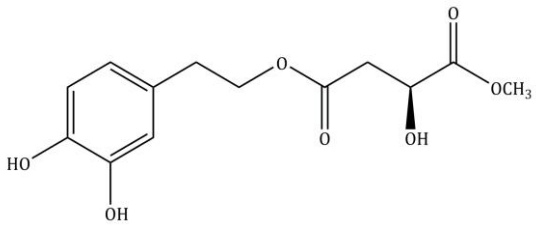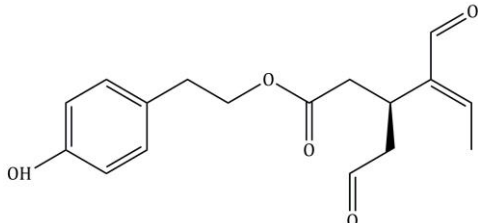
Phenolic ligand	MW	COX-1 Docking		COX-2 Docking	
		Rank	Energy (kcal/mol)	Rank	Energy (kcal/mol)
MMHTE	284	10	-43	34	-33
OLEO	305	12	-42	44	-30

Table 5.2: ADMET summary for MMHTE and OLEO

	MMHTE	OLEO
Human intestinal absorption	Yes	Yes
Blood brain barrier penetration	No	Yes
Hepatototoxicity	No	No
Cytochrome P450 2D6 inhibition	No	No
Plasma protein binding	No	No
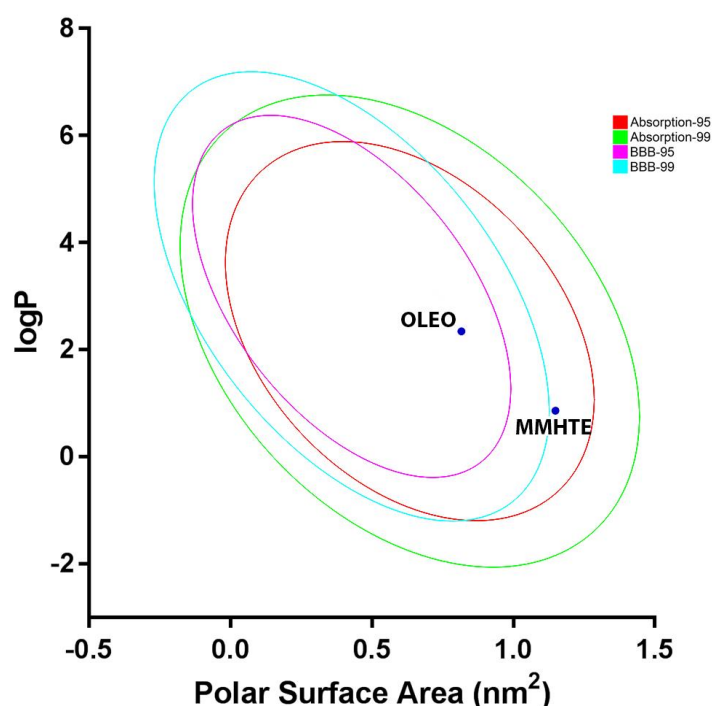
Figure 5.2: Polar Surface Area (PSA) vs logP for MMHTE and OLEO showing 95% and 99% confidence limits, denoted by ellipses corresponding to blood-brain barrier (BBB) and intestinal absorption models

Human intestinal absorption is one of the major factors affecting oral absorption, an important step in the early phase of lead discovery and optimisation [254, 255]. Through Discovery Studio, it was found that both MMHTE and OLEO were predicted to be absorbed in the intestinal tract (Figure 5.2). Blood brain barrier (BBB) penetration is an important property in the design of drugs for targeting or avoiding the brain [254]. MMHTE had a logP value of 0.86 and PSA of 1.15 nm^2, and these values for OLEO were 2.34 and 0.82 nm^2. OLEO was predicted to fall within the confidence limits for BBB penetration, while MMHTE was at the margin of, but not within these limits (Figure 5.2).

More than 90% of market withdrawals are due to drug toxicity, with hepatotoxicity being a major cause for the high attrition rate of drugs [254, 256]. MMHTE and OLEO were predicted to be non-hepatotoxic (Table 5.2). The human cytochromes P450 (CYPs) are a factor in the prediction of drug metabolism, as drugs are mainly broken down into harmless soluble metabolites excreted through urine or bile. More than 90% of approved therapeutic drugs are metabolised by five major isoforms: CYP1A2, 2C9, 2C19, 2D6, and 3A4 [254, 257]. It was predicted that both MMHTE and OLEO are classified as inhibitors of the CYP 2D6 enzyme (Table 5.2).

As there is a high concentration of plasma proteins found in the bloodstream, the degree to which candidate drugs bind to these proteins can affect the efficiency of their distribution. The plasma protein binding for MMTHE and OLEO were predicted to be false, suggesting that these compounds can efficiently reach their targets for the exertion of pharmacological effects (Table 5.2).

Oleocanthal has an absorption in the human body of 60—90% and has been studied for its anti-inflammatory and anti-cancer properties [70, 258, 259]. It has also been investigated in brain health, linked to a reduced risk of Alzheimer's disease (AD) [233, 260]. A 2013 study by Abuznait et al. enhanced clearance of β-amyloid, one of the characteristic proteins that accumulate in AD [260]. The various biological effects of oleocanthal indicate its ability to be absorbed through the gastrointestinal tract, as well as the potential penetration of the BBB. MMHTE is a novel compound, and its biological effects have not yet been studied. The ADMET results outline above indicate that there may be some similarities to oleocanthal.

Further work will involve analysing the ADMET properties of the olive phenolic library using other ADMET programs to obtain additional biological parameters such as Caco-2 and MDCK cell permeability, human Ether-a-go-go Related Gene (hERG) inhibition, and P-glycoprotein inhibition. This can be done using SwissADME and Schrodinger's QikProp [261, 262].

The potential bioavailability of these two compounds is further investigated SMD simulations, which will be presented in the next section.

5.3.2 Steered dynamics simulations

The potential bioavailability of MMHTE and OLEO was further investigated using SMD simulations. As well as these two compounds, positive and negative controls were also selected based on the ADMET results and known biological activity. ADMET data showed that hydroxytyrosol fell within the limits for intestinal absorption and BBB penetration, while elenolic acid diglucoside was well outside (Figure 5.1). These olive phenolic compounds were selected as positive and negative controls, respectively. Hoescht 33342 is biologically known to be membrane permeable, while propidium is known to be non-permeable to cells [263, 264]. Hence, these compounds were selected as additional positive and negative controls, respectively.
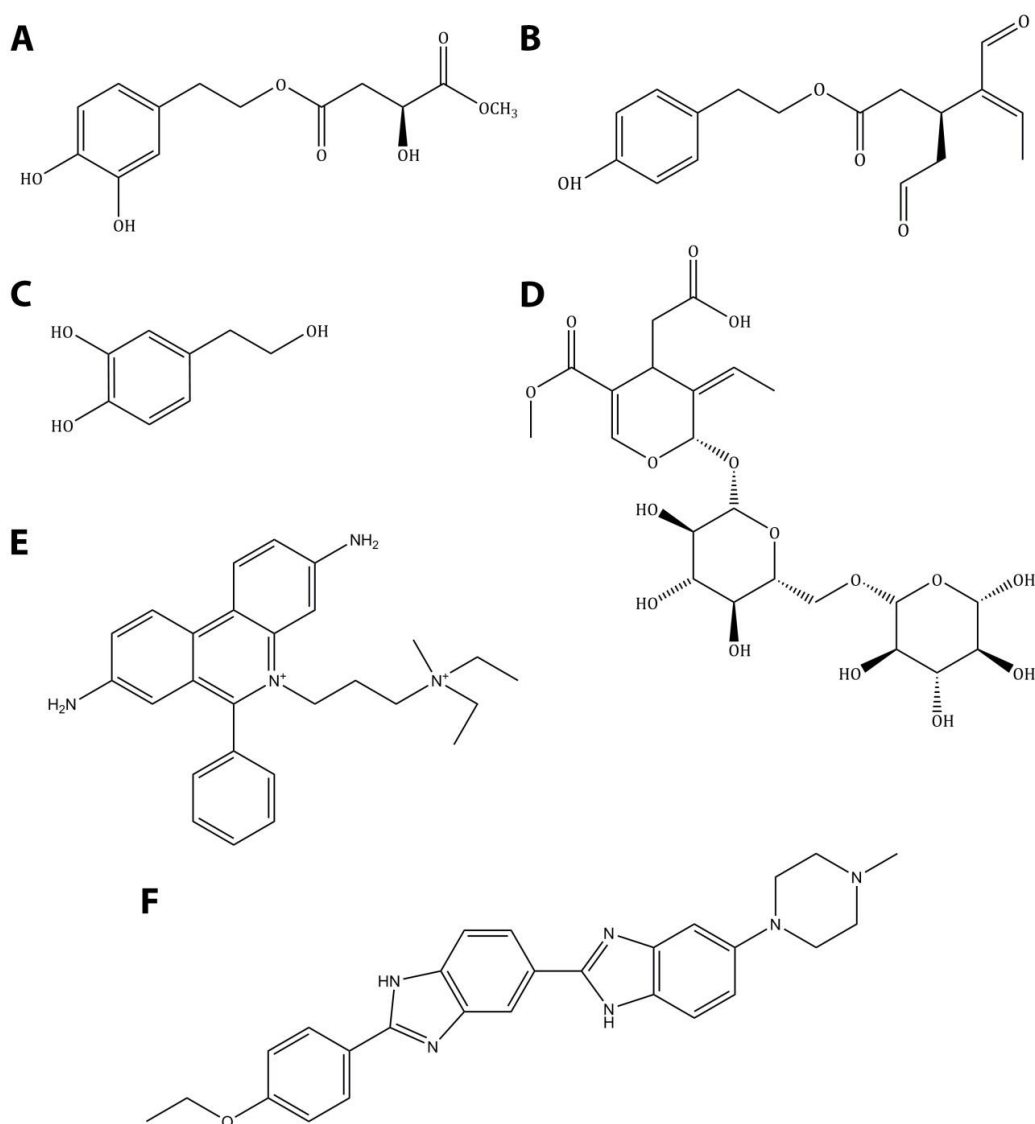
Figure 5.3: Compounds used for SMD simulations – methyl malate-β-hydroxytyrosol ester (A), oleocanthal (B), hydroxytyrosol (C), elenolic acid diglucoside (D), propidium (E), and Hoescht 33342 (F).

The profiles for the average pulling force over 10 runs of SMD simulations for each ligand passing through the DOPC membrane are presented in Figure 5.4. Increased force appears as the ligands begin to enter the membrane, implying energy barriers are encountered. As ligands travel through the lipid bilayers, pulling force either plateaus or further increases, before decreasing and returning to zero as ligands exit the membrane and reenter the water phase. An peak occurs as ligands approach the phosphocholine group of the lipid bilayer, suggesting that another energy barrier is encountered. As the force approaches zero, it is suggested that ligands have completeley passed through the lipid bilayer, hence returning to the water phase. HT, OLEO, and MMHTE demonstrate similar force profiles, while EADG, H33342, and propidium demonstrate higher force peaks.
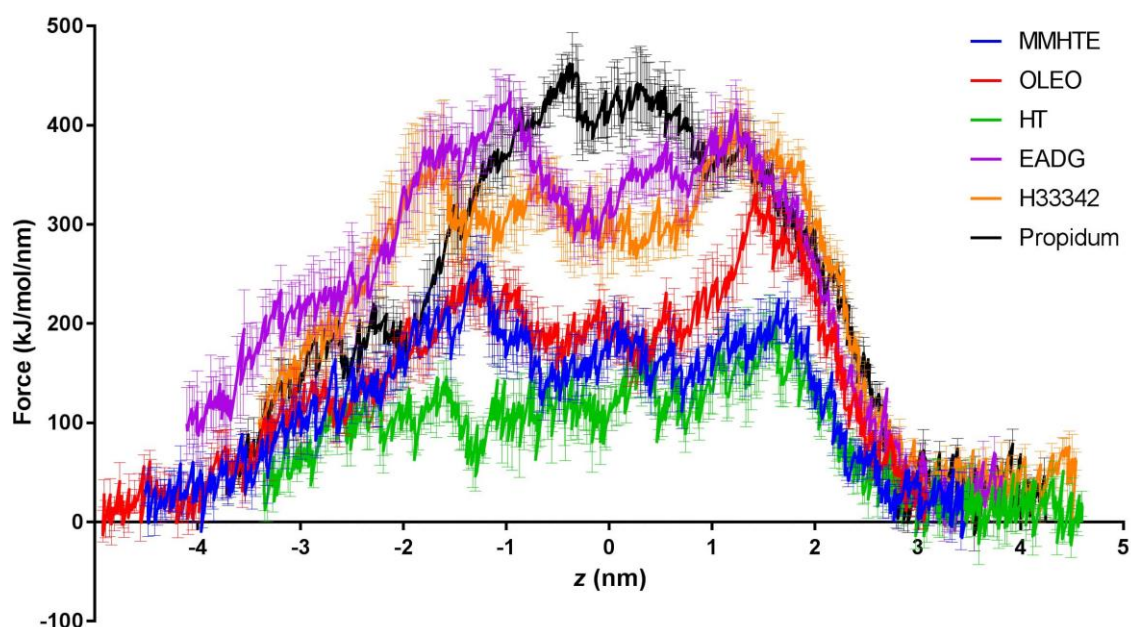
Figure 5.4: Force profile as ligands are pulled through DOPC membrane with respect to position of the ligand along the bilayer normal (*z*) for methyl malate-β-hydroxytyrosol ester (MMHTE), oleocanthal (OLEO), hydroxytyrosol (HT), elenoic acid diglucoside (EADG), Hoescht 33342 (H33342), and propidium. Forces are shown as an average of ten runs for each ligand, with error bars showing the standard error of every second data point.

The force profile of HT shown in Figure 5.4 demonstrates that it is able to easily pass through the lipid bilayer, with the least force required out of all the ligands. The maximum force of 202 kJ/mol/nm occurs at ~300 ps, as HT passes through phosphocholine head groups and approaches the fatty acid tails of the membrane. The force gradually decreases as HT exits the membrane.

MMHTE also demonstrates a small force barrier for the penetration of the DOPC membrane. Two peaks are observed as the ligand enters and exits the lipid membrane at ~200 and ~500 ps, reaching maximum forces of 224 kJ/mol/nm and 262 kJ/mol/nm respectively. As MMHTE travels through the membrane, the force plateaus at ~180 kJ/mol/nm, indicating some lipophilicity of MMHTE. Figure 5.5 presents the various orientations assumed by MMHTE as it passes through the DOPC membrane. As the ligand enters the membrane, the phenol head group is orientated downwards. Following entry of the membrane, MMHTE assumes a flat position. Gradually, MMHTE turns and orients itself such that the tail of the ligand exits the membrane first.
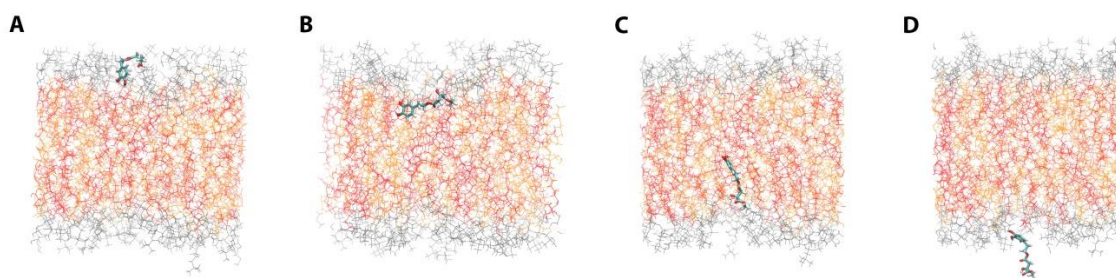
Figure 5.5: Permeation of MMHTE through DOPC membrane at 146 (A), 253 (B), 434 (C), and 608 (D) ps.

Similar to MMHTE, OLEO also demonstrated that a relatively small amount of force was required to penetrate the membrane. Initially, there is a slightly higher force barrier encountered as OLEO enters the membrane, reaching a peak of 330 kJ/mol/nm at ~200 ps. This force quickly decreases, plateauing at ~180 kJ/mol/nm from ~300—500 ps, before entering a slight blip at ~500 ps as the ligand exits the membrane. The orientations of OLEO as it passes through the membrane are shown in Figure 5.6 below. As OLEO enters the membrane, it assumes a folded position with neither head nor tail pointing downwards. Similar to MMHTE, OLEO assumes a flat positions as it moves through the fatty layer. This flat position is retained until the ligand exits the membrane, led by the tail-end of the compound.
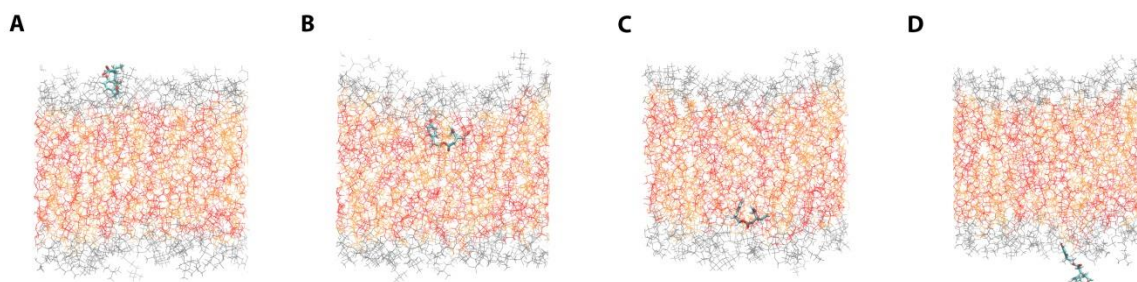
Figure 5.6: Permeation of OLEO through DOPC membrane at 110 (A), 232 (B), 462 (C), and 582 (D) ps.

Elenolic acid diglucoside (EADG) is an olive phenolic compound, classified as a secoiridoid and found in the olive fruit, leaf, and pomace [265-267]. With a molecular weight of 553 g/mol, EADG is a larger compound compared to oleocanthal and MMHTE (305 g/mol and 284 g/mol respectively). EADG has a high polar surface area of 2.55 nm^2 and low logP of 3.06, placing it outside the limits for intestinal absorption and BBB penetration in Figure 5.1. The low lipophilicity of this compound suggests that it may not pass through the cell membrane easily, hence SMD simulations were performed with EADG as a means of negative control. As expected, the force profile for EADG reaches a higher magnitude compared to the other olive phenolic compounds. As with the other ligands, two peaks occur at ~300 and ~500 ps, with the maximum force being 433 kJ/mol/nm. It should be noted that the DOPC membrane is severely disrupted as EADG is passing through. To allow the membrane to recover, SMD simulation of EADG was also conducted with a pull rate reduced by a magnitude of 10 (1.0 nm/ns) (Figure A3.2, Appendix 9.3.2). While the force value calculated here should not be taken as a direct quantitative measure, for the purposes of this section the significantly higher force is interpreted to mean that a greater force is required for this compound to pass through the membrane, compared to other olive phenolic compounds of interest.

Propidium iodide is a commonly used fluorescent stain for the detection of dead cells in a population, since it is not permeant to live cells [263]. Thus, propidium was used as an additional negative control in this study. Its force profile demonstrates that greater force is required to pass through the membrane. A peak force of 462 kJ/mol/nm occurs at the midpoint of the simulation. Unlike the other compounds, the shape of the force profile shows a single peak. Rather than encountering force barriers around the phosphocholine head

groups as observed with other compounds, the largest force barrier occurs at the midpoint of the lipid bilayer. This may demonstrate the lipophobic nature of the compound.

Hoescht stains are blue fluorescent dyes, with Hoescht 33342 (H33342) commonly used to stain DNA [268]. It works by binding to the minor groove of double stranded DNA and is known to be cell-permeable in both live and fixed cells [264]. The force profile for H33342 shown in Figure 5.3 shows that although the initial force barrier is larger (403 kJ/mol/nm at ~350 ps), this quickly decreases, and a plateau persists between ~400-700 ps at ~300 kJ/mol/min, before peaking again as the ligand exits the membrane at ~700 ps and rapidly decreasing as H33342 re-enters the water phase. Taking longer to exit, the persistent force plateau observed as the ligand traverses the lipid bilayer is suggestive of its lipophilic properties, aided in part by the ethyl group present in its structure (Figure 5.2F) [269].

In a study by Fox et al., the interaction of two NSAIDs, salicylate and ibuprofen, with vesicles of 1,2-dimyristoyl-*sn*-glycero-3-phosphocholine (DMPC) were examined using optical-trapping confocal Raman microscopy [270]. While the membrane was permeable to both drugs, it was found that ibuprofen preferentially accumulates in the membrane while salicylate does not. This suggested that ibuprofen localises in the hydrophobic acyl chain region of the membrane, whereas salicylate is weakly associated with the phospholipid headgroups [270]. These results are in qualitative agreement with the force profile presented in Figure 5.4. The plateaus observed for OLEO and MMHTE agree with the behaviour of ibuprofen. An MD simulation study of NSAIDs in lipid membranes noted that compared to aspirin, lipid interactions with ibuprofen were largely governed by the hydrophobic tail rather than the polar headgroup, even when ibuprofen was charged [271]. Displaying similar behaviour to the weaker association of salicylate with the membrane, this trend was not seen in the force profile of HT (Figure 5.4) [270]. To further elucidate the membrane penetrating properties of OLEO and MMHTE, the PMF was subsequently calculated.

5.3.3   Umbrella sampling

Umbrella sampling was performed for OLEO and MMHTE to calculate the potential of mean force (PMF).
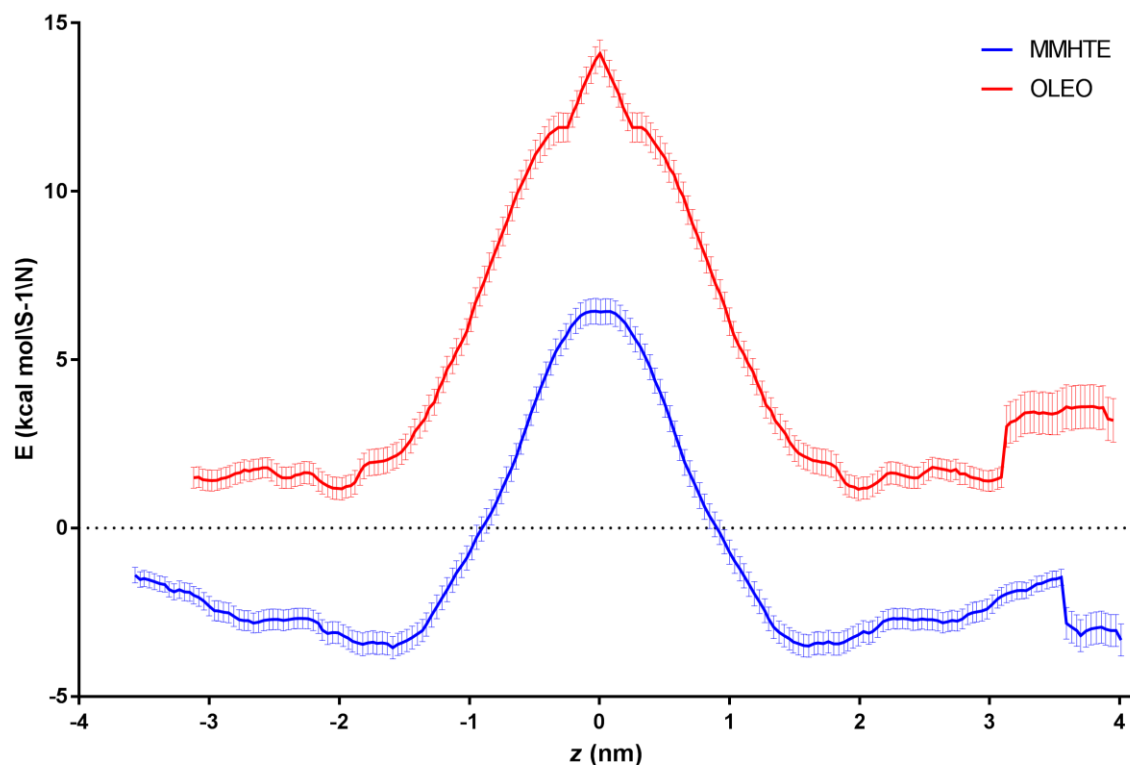
Figure 5.7: Potential mean force curve for MMHTE and OLEO pulled through DOPC membrane (symmetrised). Error bars are standard deviations calculated from 200 bootstrapped PMFs.

Figure 5.7 presents the symmetrised PMF curves for the penetration of MMHTE and OLEO through the DOPC lipid bilayer. It was found that OLEO produced a ΔG of 14.13 kcal/mol, and MMHTE yielded a ΔG of 9.40 kcal/mol. Consistent with the force curve in Figure 5.4, these two olive phenolic compounds have a similar amount of force required to penetrate the lipid bilayer, with OLEO requiring marginally greater force.

PMF was determined by pulling the ligand across the entire membrane, then symmetrising the resulting profile, such that the profiles were adjusted to be identical on both sides of the membrane centre starting and ending at 0. This method has previously been shown to produce a more accurate result compared to simulating over half the range for twice as long i.e. from one side to the membrane centre [107].
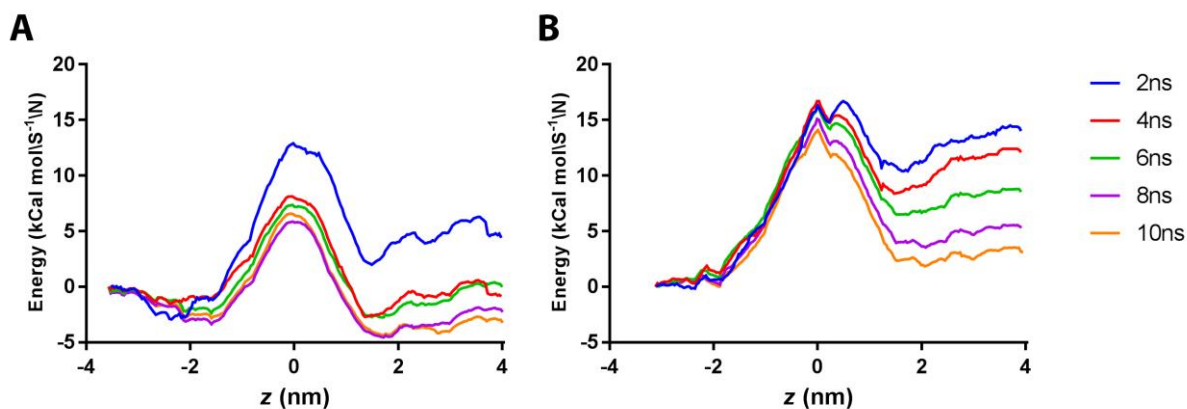
Figure 5.8: Potential mean force curve for MMHTE (A) and OLEO (B) pulled through DOPC membrane with each window simulated for varying lengths, prior to symmetrisation.

Prior to symmetrising, the degree of asymmetry suggested accumulated error (Figure 5.8). The asymmetry between the two end points is ~3 kcal/mol, potentially indicating slow converging orthogonal degrees of freedom [107]. This may be due to residual disturbance to the membrane structure from the starting states generated with SMD. To rectify this, a longer equilibration of the starting states may be required [107, 272]. This will be done in further work.

A study by Boggara and Krishnamoorti calculated PMFs of aspirin and ibuprofen in bilayers of dipalmitoylphosphatidycholine (DPPC) [271]. Using the GROMOS force field with the simple-point charge water model, the PMF was calculated for one monolayer and assumed to apply to the other monolayer. The trends in K values obtained from MD results of the study agreed with experimental trends [271]. In an anionic environment, the ΔG of ibuprofen was found to be -36.6 ± 3.0 kJ/mol (equivalent to 8.75 ± 0.72 kcal/mol) [271]. This value approximately agrees with the ΔG value for MMHTE of 9.40 kcal/mol (Figure 5.7). Although the calculated ΔG for OLEO slightly deviates from these values (14.13 kcal/mol), further equilibration of these simulations may provide more accurate estimates in further work.

It has been suggested that membrane fluidity may be a prerequisite of COX binding [273, 274]. The flexibility of COX proteins at the junction between the membrane bound domain and the catalytic domain substrates provide an access path to the cyclooxygenase active site from within the lipid bilayer [212, 275]. Additionally, the inhibition of facilitated sulfate transport may contribute to the toxicity of NSAIDs on gastromucosal cells [276-278]. One of

the suggested mechanisms for the gastrointestinal (GI) toxicity of traditional NSAIDs was the direct interaction of the drugs with zwitterionic phospholipids lining the GI mucosa [279, 280]. Therefore, understanding the interaction between potential COX inhibitors with membranes gains further importance for predicting the biological activity of these compounds.

5.4 Conclusion

In this chapter, the full library of olive phenolic compounds was analysed for their ADMET properties. Combining the results of the ADMET study and docking results obtained in Chapter 4 resulted in the selection of a novel compound which demonstrated favourable ADMET properties and strong binding to COX-1 and COX-2: methyl malate-β-hydroxytyrosol ester (MMHTE).

Following selection of the compounds of interest, a preliminary study of membrane permeability was performed. SMD simulations were performed to assess the force required to penetrate a DOPC membrane. Umbrella sampling was performed for MMHTE and oleocanthal to produce a potential mean force (PMF) curve for penetration of the DOPC membrane. While the PMF provided a more quantitative energy estimate, a longer equilibration time may be required for a more reliable estimate. Further work is required and will be conducted in future studies. For this study, it is shown that MMHTE and oleocanthal require similar amounts of force to pass through a lipid bilayer, which is as expected due to their similar molecular structure.

The membrane permeability of olive compounds has important implications for their bioavailability and toxicity, as well as their potential efficacy in accessing the COX active site for potential modulation of the protein. In the following chapter, the molecular mechanism for the COX inhibitory potential of the olive compounds selected in these chapters will be examined.

6. Molecular dynamics simulation

6.1 <u>Introduction</u>

Molecular dynamics (MD) simulation is a technique that can be used to effectively understand the structure and function of macromolecules. Using MD simulations, this chapter examines the dynamic properties of COX-1 and COX-2 bound to olive phenolic ligands selected from the previous chapters. This includes oleocanthal (OLEO), 1-oleyltyrosol (1OL), ligstroside derivative 2 (LG2), and methyl malate-β-hydroxytyrosol ester (MMHTE). The native ligands present within the original crystal structures of the protein were also included, being flurbiprofen (FLP) in COX-1 and salicylate (SAL) in COX-2. While the COX enzymes are monotopic membrane proteins, the membrane is excluded from simulations in this chapter for computational efficiency. The extraction of COX from the membrane is unlikely to induce a major conformational change altering the native membrane-bound structure, as the catalytic domain is structurally homologous to the soluble enzyme canine myeloperoxidase [8, 281].

The dynamical response of the protein to binding of different ligands may offer alternative quantitative or semi-quantitative measures which might correlate with the biological activity of the ligand. The COX protein channel undergoes structural fluctuations, such as through reduced disorder in helix D residues upon ligand binding to COX-2 [282]. The shift in orientation of this helix enables the opening and closure of the gate region in the active site of both COX-1 and COX-2 [74, 283]. Protein dynamics has been shown to play a role in protein-ligand interactions, determining the binding affinity for a ligand [284]. Additionally, it has been suggested that allosteric perturbations can change the relative distributions of states within an ensemble [285].

Thus, further predictive measures are proposed based on structural changes in order to complement more traditional measures of bioactivity based on ligand binding energy. This involves performing classical MDS, as well as employing additional analytical methods using

MD data such as essential dynamics (ED), network analysis, and Molecular Mechanics-Poisson Boltzmann Surface Area (MM-PBSA) calculations. This chapter will describe the use of these techniques to elucidate the mechanism for the binding of selected olive compounds to COX proteins.

6.2 Methods

6.2.1   MD simulation procedure

MD simulations were conducted using GROMACS 4.6.5 software package to further investigate the dynamics of the binding mode of ligands determined from molecular docking [139, 239]. The topology files of COX-1 and -2, as well as the ligands, were generated using pdb2gmx within Gromacs and SwissParam respectively [246]. MD simulations were run using the CHARMM27 force field [243, 245]. Both COX-1 and COX-2 complexes were solvated using TIP3P water [242] in a truncated octahedron period box, with a minimum of 1.2 nm distance between any protein complex atom to the closest box edge. Chloride counterions were added to the solvated system to neutralise the charge. Energy minimisation was performed on the system using a steepest-descent gradient method for a maximum of 50,000 steps.   Next, each protein-ligand complex was restrained using an isothermal-isochoric (NVT) ensemble for 100 ps. This was followed by pressure stabilisation using the isothermal-isobaric ensemble (NPT) for 100 ps.

MD simulations were performed for 100 ns with a time-step of 2 fs. The temperature was kept stable at 300 K using velocity rescaling with a stochastic term [286], and the pressure controlled isotropically at 1.0 bar using a Parrinello-Rahman barostat [159, 248, 249]. Trajectories were written every 10 ps. Each production run was submitted in triplicate, with velocities randomly generated according to a Maxwell distribution. All bond lengths were constrained by the LINCS algorithm [250]. Long-range electrostatic forces were evaluated using the particle-mesh Ewald scheme (PME) (grid spacing 0.16 nm) [154], while short-range nonbonded interactions were calculated using cutoff ratios of 1.4 nm for both Coulomb and van der Waals potentials.

Molecular dynamics parameters (mdp) files used for the energy minimization, NVT, NPT and production runs are provided in Appendix 9.4.1. Root mean square deviation (RMSD)

and root mean square fluctuation (RMSF) were calculated. Protein structures were visualised using visual molecular dynamics (VMD) 1.9.1 [211].

MDS applications were completed on computing platforms provided by Victorian Life Sciences Computation Initiative (VLSCI) and National Computational Infrastructure (NCI), specifically IBM Blue Gene/Q (Avoca) and IBM iDataplex x86 system (Barcoo) from VLSCI, and an Intel Xeon Sandy Bridge and Broadwell processor cluster (Raijin) at NCI.

6.2.2 Essential dynamics analysis

Utilising principal component analysis (PCA), ED was used to filter large-scale concerted motions from the MD trajectories of COX complexes [166, 287]. Three independent 10 ns segments from each triplicate trajectory with different initial atomic velocities were concatenated to obtain a single 30 ns trajectory, representing different sampling directions around the starting structure. A covariance matrix was generated from the atomic fluctuations in the trajectory, which was then diagonalized to obtain a set of eigenvectors and corresponding eigenvalues. For the COX complexes, the protein and hydrogen covariance matrices were built and diagonalized using the g_covar tool within the GROMACS package. Projections of trajectories onto the eigenvectors were performed using the g_anaeig tool within GROMACS. The concerted motions were visualised using a Tcl script combined with VMD to plot porcupine representations of the motions [211, 288].

6.2.3 Network analysis

Modular NETwork Analysis (MONETA) was used with the MD data of COX complexes to build and analyse intermolecular communications between the residues of COX-1 and -2 in their bound and unbound forms [168]. MONETA builds a modular network representation of a protein, composing clusters of residues representing independent dynamic segments (IDSs) and chains of residues representing communication pathways (CPs). IDSs were identified using Local Feature Analysis (LFA), a statistical method based on PCA that transforms global modes of the protein into 'local modes' to describe local dynamical behaviour independent from the rest of the protein motions [169, 170]. Calculations were performed on a 10 ns segment of trajectories of the apo forms of COX, as well as COX bound to each ligand.

Distance matrices comprised of the average smallest distance between each Cα residue pair were computed using the Cpptraj module of AMBER 12 [289, 290] and the g_mdmat tool

implemented in GROMACS 5.0.5 [139, 291]. Pairs of residues were considered neighbours if the average smallest distance between them was lower than a threshold of 0.37 nm. MONETA employs the concept of communication propensity to characterise CPs [171]. CPs are grown so that any two adjacent residues are connected by non-covalent interactions, and that every residue in the CP is connected to any other point by a short commute time (CT). Non-bonded interactions along the trajectory were computed using LIGPLOT [292]. Two residues were considered interacting when they formed at least one non-bonded interaction for at least 50% of the simulation time. The CT threshold value was 0.1 for all studied systems.

The protein structure, interactions, and communication paths were visualised with PyMOL and Gephi modules incorporated in MONETA [168, 194]. Protein visualisation was also performed using VMD 1.9.1 [211].

6.2.4 MM-PBSA: Per-residue energy decomposition

Per-residue decomposition analysis was performed to obtain a quantitative description of the energetic contribution of each amino acid with the ligands considered. As evident from previous studies, binding free energy calculations and their per-residue decomposition methods are crucial for determining the binding mechanisms of protein–protein or protein–ligand complexes [116, 173].

MM-PBSA calculations were performed on three independent 1 ns segments of the generated trajectory with different initial atomic velocities, and an average was taken. It has been found that this method generates more accurate calculations than analysing a single trajectory spanning a longer timeframe [293]. This is likely due to an amplification of force field errors along the MD simulation time frames, rather than errors in non-converged energy in the system [293, 294].

Calculations were performed using the g_mmpbsa tool [180, 181]. Contributions from the electrostatic energy, van der Waals energy, and polar solvation energy terms were calculated using adaptive Poisson-Boltzmann Solver [295]. For MM-PBSA calculation, the grid spacing was set to 0.05 nm. Values of 80 and 2 were taken for solvent dielectric constant and solute dielectric constant respectively. The non-polar energy contribution was approximated using

solvent-accessible surface area (SASA). Probe radius for SASA estimation was set to 0.14 nm. Entropic energy terms were not included in the calculation.

6.3 Results and discussion

6.3.1 *Comparison of ligand effects on overall protein structure and motions: RMSD and RMSF*

To assess the stability of the protein-ligand complexes, they were each independently simulated in triplicate in a water box for 100 ns. RMSD for the protein relative to backbone atoms are shown in Figure 6.1. Each curve was calculated as an average over the three trajectories at each time point. The graph shows that the structural drift for each protein-ligand complex is relatively consistent, producing an RMSD of approximately 0.30 nm in COX-1 and 0.25 nm in COX-2. The standard error for the RMSD of the protein complexes is shown in Figure A4.1 in Appendix 9.4.2.

The RMSD for COX-1 and COX-2 begins to stabilise at around 15 ns, an indication that the chemical and physical properties of the system have equilibrated. After 15 ns, the average RMSD of the COX-1 complexes are: 0.30 nm in its apo form, 0.31 nm for FLP, 0.30 nm for 1OL, 0.31 for LG2, 0.29 for OLEO, and 0.28 for MMHTE. For COX-2, the average RMSD values are: 0.24 in its apo form, 0.28 for SAL, 0.23 for 1OL, 0.23 for LG2, 0.23 for OLEO, and 0.27 for MMHTE. Thus, COX-1 exhibits more consistent RMSD values irrespective of ligand, while COX-2 exhibits a greater spread of average RMSD values, and is more sensitive to ligand type.
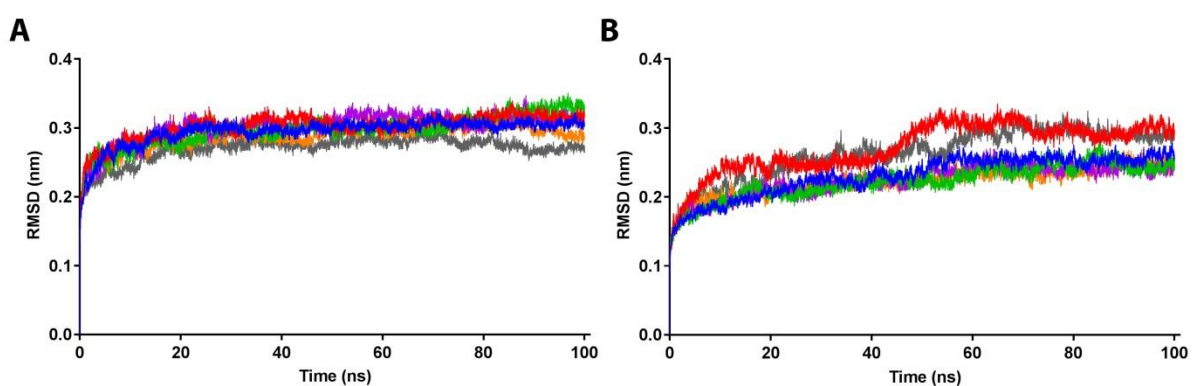
Figure 6.1: RMSD of backbone atoms of COX-1 (A) and COX-2 (B) bound to olive ligands for 100 ns: apo (blue), native ligand (red), 1-oleyltyrosol (green), ligstroside derivative 2 (purple), oleocanthal (orange), and methyl malate-β-hydroxytyrosol ester (grey).
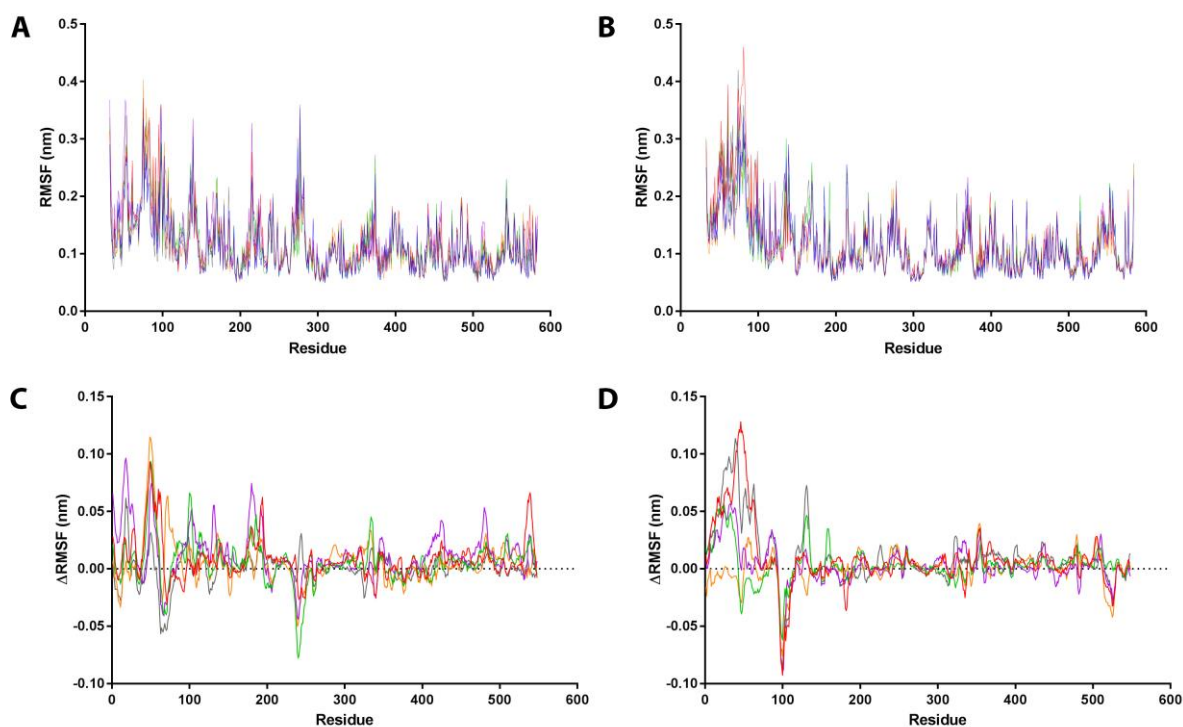
Figure 6.2: RMSF for backbone atoms of COX-1 (A) and COX-2 (B) for bound olive ligands for 100 ns: apo (blue), native ligand (red), 1-oleyltyrosol (green), ligstroside derivative 2 (purple), oleocanthal (orange), and methyl malate-β-hydroxytyrosol ester (grey). In (C) and (D) the average values of 5 residues are taken to produce a "moving average" of RMSF values, with the RMSF values for the apo form subtracted from ligand bound forms.
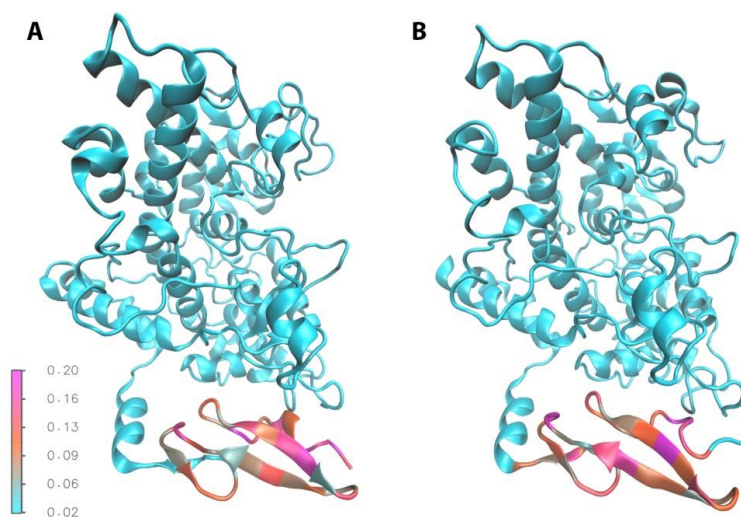
Figure 6.3: RMSF for apo form of COX-1 (A) and COX-2 (B). Colour scale bar indicates root mean square fluctuation in nm.

Root mean square fluctuation (RMSF) is a measure of flexibility of the backbone chain per residue, calculated by taking into account the fluctuation of the protein with respect to its average structure. This is shown in Figure 6.2, and the standard error of these plots shown in Figure A4.1 in Appendix 9.4.2. The RMSF graphs shown in Figure 6.2 demonstrate that the

greatest flexibility was in the first 100 residues of the protein. Other residues demonstrate a higher rigidity, and likely contribute to the stability of the protein. To illustrate the location of these residues within the protein, Figure 6.3 depicts ribbon structures of COX-1 and COX-2 colour coded according to RMSF. It can be observed that the first 100 residues have higher flexibility, with this region corresponding to the N-terminal epidermal growth factor (EGF)-like domain and the membrane binding domain (MBD). The larger C-terminal globular catalytic domain containing the active site was shown to be structurally stable with comparatively lower fluctuation for both COX-1 and COX-2 throughout the span of the simulation.

In Figure 6.2 C and D, a "moving average" of RMSF values was produced. This was done by taking the average of five sequential residues in a set, with each set beginning with the next residue. This was followed by a subtraction of these values for the apo form from ligand bound forms. Through this, noise is reduced to demonstrate differences in RMSF values that can be attributed to different ligands bound to COX-1 and COX-2 relative to their apo forms. The greatest peaks and troughs of RMSF differences tend to follow the fluctuations observed in their graphs for the original RMSF (Figure 6.2A and B). This indicates that structural changes occurring due to the type of bound ligand are occurring at residues that contain the greatest deviation from the average structure. While there tends to be little difference in RMSF values for most residues, it can be noted that the N-terminal residues of COX-2 bound to OLEO display a similar RMSF to that of the apo form (Figure 6.2D). This may suggest that binding to OLEO does not significantly alter the flexibility of the EGF-like domain in COX-2.

The bottom of the NSAID binding site of COX-2 can assume two possible conformations [275]. These involve helix D, the first half of which comprises residues 107-116 and lies in the same plane as other helices of the domain, being proposed to interact with the membrane [275]. Residues 117-121 comprise the second half of the helix, which extends out of the plane of the membrane to form part of the cyclooxygenase active site [275]. The structural changes involved in the open binding site are from the second half of helix D located in the binding site, rather than the membrane binding function of the protein [275]. The ability of the protein to transition from open to closed forms enables substrates and inhibitors to access the internal binding site from within the membrane [275]. From the RMSF graphs in Figures 6.2 and 6.3, residues 117-121 of the second half of helix D demonstrate greater flexibility in

COX-2 bound to LG2 and MMHTE. A similar trend is observed in COX-1, with a greater difference observed in COX-1 bound to LG2. This may suggest that the binding site assumes different conformations when bound to these two ligands compared to other olive phenolic ligands.

6.3.2   Essential dynamics analysis

Motions of the protein were analysed to extract the most essential movements which might be central to the function of the protein.

6.3.2.1   Conformational sampling

Conformations explored in MD simulations are only a subset of those possible that the protein can assume. To correlate MD data with characteristics of the protein, it must be ensured that there is sufficient sampling efficiency. To do this, principal components analysis (PCA) was used to investigate the conformational sampling. The advantage of utilising PCA in finding large scale motions in protein structures is that most of the fluctuations are able to be captured in the first few principal modes, enabling analyses in only a few dimensions for easier visual inspection of results [296]. In this analysis, the first 10 eigenvectors are shown to contribute to approximately 80% and 70% of total fluctuation in COX-1 and COX-2 respectively (Figure A4.2, Appendix 9.4.3).

Insufficient sampling often leads to protein motions along principal components resembling random diffusion, providing a less accurate description of protein behaviour. More specifically, the first few principal components of protein simulations resemble cosines, where this sampling bears more resemblance to random diffusion rather than relevant motions [296]. In this study, the cosine content of the first 10 principal components from PCA analysis of single replicas and concatenated trajectories were calculated (Table A4.1, Appendix 9.4.3). These values range from 0 and 1, with 0 indicating no cosine and 1 indicating a perfect cosine. It has previously been shown that a cosine content close to 1 may be representative of insufficient sampling where large scale motions along the eigenvector resemble random diffusion [297]. The cosine content of the first three principal components is generally lower in the concatenated trajectory compared to single replicas, suggesting that a higher number of MD simulations is more effective in obtaining reliable conformational sampling.

Another technique for investigating the sufficiency of conformation sampling is to examine principal modes of conformational space explored after dividing the simulation into two or more parts [296]. This is examined with the subspace overlap of the first few principal modes, where a wide region of the conformational space should be sampled with a partial overlap between different trajectories [296]. This is demonstrated by 2D plots of PCA-1 against PCA-2 and PCA-3 in Figure 6.4A and B, as well as in Figure A4.3 in Appendix 9.4.3 for all other ligands. These plots show that COX-1 and COX-2 occupy a distinct region re-sampled by replicas. These plots also show that there is generally more variation along the first principal component, especially in COX-2 where the plots tend to occupy a more flattened elliptical area compared to COX-1.

In addition, the projection and its corresponding distribution along each principal component was measured. Figure 6.4 shows the first, second, and twentieth projection for COX-2 apo along with their corresponding distributions (C-H). The first three projections and distributions for all apo and bound forms of COX-1 and COX-2 are shown in Figure A4.3 in Appendix 9.4.3. These results indicate that the distribution of motion along the first few principal components is anharmonic, producing two or more distinct peaks. In comparison, the twentieth principal component displays a narrow Gaussian shape—indicative of more random motion. Hence, it was determined that the first three principal components of the concatenated trajectory can be used as a reference conformational space for subsequent analysis of protein dynamics.
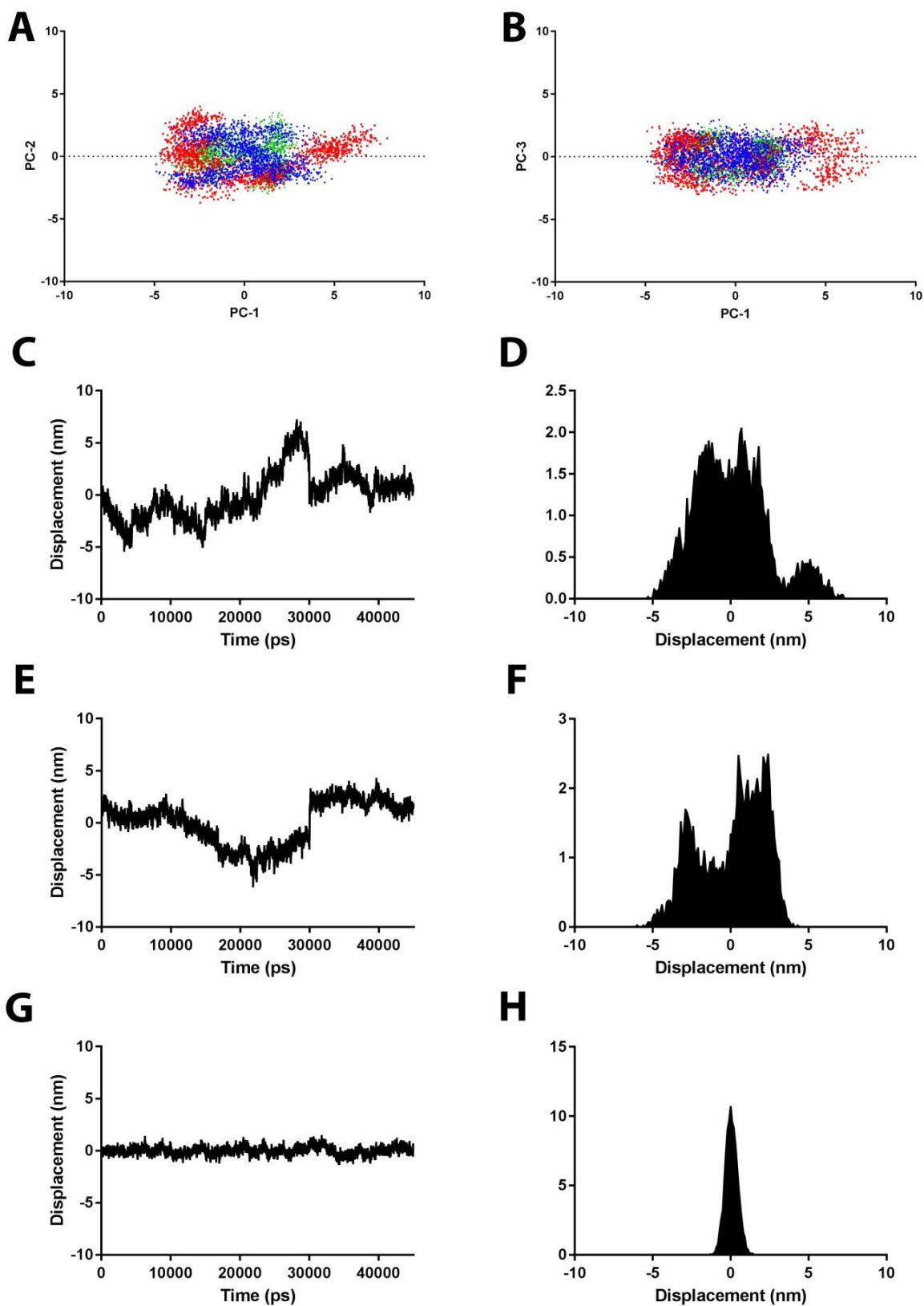
Figure 6.4: 2D projection for COX-2 apo along PC-1 against PC-2 (A), and PC-1 against PC-3 (B). Projection of motion for PC-1 (C), PC-2 (E), and PC-20 (G), along with their respective distribution (D, F, H) is shown.

6.3.2.2 Concerted motions along the first eigenvector

Essential dynamics analyses identified a number of segments which tend to move in concerted "blocks" in both COX-1 and COX-2. In this section we focus on several such blocks, especially for those for which there is some known functional significance. These include residues 32-85 (N-terminal epidermal growth factor-like domain, which is presently named "EGF block"), and 86-115 (membrane binding domain; "MBD block"). Further blocks of residues identified by essential dynamics and show significant motion, but which are not presently known for particular functions, are named simply as follows: 126-171 (C-terminal globular catalytic domain; "GCD1 block"), 209-229 ("GCD2"), and 235-290 ("GCD3"). Smaller blocks which move in concert within the above-defined larger ones have letters appended to their names alphabetically.

The overall concerted motions of each eigenvector may be described in terms of the relative rotations and translations of these blocks. Changes in the nature of the concerted motions of these blocks upon binding could suggest change in functional properties associated with those clusters of residues. The movements of residues in these corresponding blocks for the first eigenvector of each complex are described in the following section.

The movements of amino acid residues of the protein are described with respect to a perspective in which the first principal axis is aligned parallel to the page. Motion of blocks of residues within all three eigenvectors is shown in Table A4.4 of Appendix 9.4.3. The magnitude of motion is described relative to the magnitude of displacement observed within each eigenvector of the protein, and may be depicted using porcupine plots, where the motion vectors are centred at each amino acid. These porcupine representations are shown in Figure 6.5 and Appendix 9.4.3, Figure A4.5. A table describing the values of colours pertaining to the magnitude of movement for each representation is listed in Table A4.2 of Appendix 9.4.3.
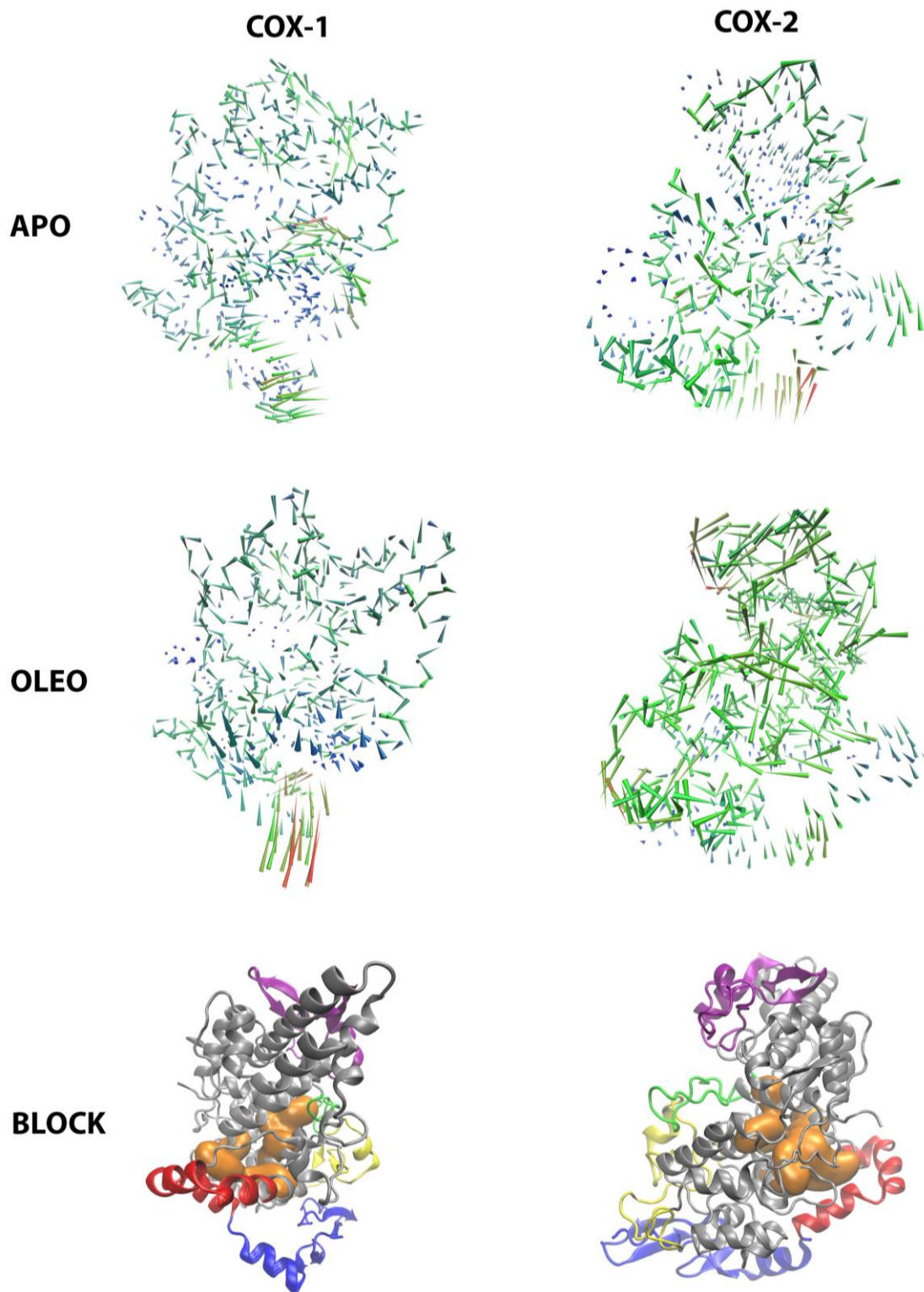
Figure 6.5: Porcupine plot of movement in the first eigenvector of COX-1 and COX-2 in its apo form, and bound to OLEO. Blocks of residues showing concerted movement are also depicted. Each coloured segment represents a block, with the surface of active site residues coloured in orange. Each block represented by a different colour: EGF block (blue), MBD block (red), GCD1 (yellow), GCD2 (green), and GCB3 (purple).

Corresponding to the EGF-like domain of the proteins, EGF block residues were shown to be the most dynamic of the whole protein. These residues demonstrated the most movement in all bound and unbound forms of COX-1 and COX-2. Concerted movement occurred in segments within this group of residues. Thus, the EGF block was further divided into 3 smaller sections: EGFa spanning residues 32-47, EGFb spanning 48-69, and EGFc spanning 70-85.

In the apo form of COX-1, segment EGFc moved in an anticlockwise direction. For COX-1 bound to FLP, EGF block residues of the first eigenvector demonstrated upward anticlockwise motion with a large magnitude. For COX-1 bound to 1OL, movement of EGF block residues in was concerted in the first eigenvector at segments EGFb and EGFc, moving in a clockwise direction upwards and downwards respectively. For COX-1 bound to OLEO, EGFa and EGFb moved as a unit in an anticlockwise direction with a large magnitude, and EGFc was moved upwards. COX-1 bound to MMHTE displayed movement in EGFc residues, moving upwards with a large magnitude in the first eigenvector.

In COX-2, residues in the EGF block moved as a whole unit rather than segments as observed in COX-1. EGF block residues of the apo form of COX-2 moved with a small magnitude in a downward clockwise direction in the first eigenvector. For COX-2 bound to SAL, clockwise displacement was observed. COX-2 bound to 1OL demonstrated a smaller magnitude of displacement, with EGF block residues rotating in an anticlockwise direction in the first eigenvector. In COX-2 bound to LG2, there was clockwise movement of a moderate magnitude. For COX-2 bound to OLEO, residues of the EGF block moved in an anticlockwise direction. In COX-2 bound to MMHTE, residues displayed movement of a larger magnitude, moving in a clockwise direction.

MBD block

The MBD of the COX proteins comprises four short, consecutive, amphipathic α-helices. Three of these helices lie in the same plane, while the fourth helix protrudes upwards into the catalytic domain [9].

Movement of MBD block residues was not observed in the first eigenvector for the apo form of COX-1 and COX-1 bound to 1OL, however movement did appear in other eigenvectors. There was moderate movement was observed in the first eigenvector of COX-1 bound to

FLP. MBD block residues moved with an upward motion of a large magnitude in the first eigenvector in COX-1 bound to OLEO. COX-1 bound to LG2 and MMHTE did not display concerted motion of MBD block residues.

MBD block residues displayed movement across all forms of COX-2, with the exception of COX-2 bound to 1OL. Upward motion of a moderate magnitude was observed in the first eigenvector of the apo form. Upward motion was also observed in the first eigenvector of COX-2 bound to SAL, as well as the in COX-2 bound to MMHTE. Moderate anticlockwise motion of these residues was observed in the first eigenvector of both COX-2 bound to LG2 and OLEO.

GCD1 block

As with the EGF block, GCD1 block residues also displayed concerted motion in segments. The residues of GCD1 block were divided into segments GCD1a spanning residues 126-171, GCD1b spanning 134-148, and GCD1c spanning residues 149-171.

For block 3 residues in COX-1, segment GCD1b displayed the most movement. In the apo form, GCD1b residues rotated in a clockwise direction in the first eigenvector. In COX-1 bound to LG2, GCD1b residues displayed a moderate magnitude of motion downwards in the first eigenvector. A moderate magnitude of motion was also observed for residues of segment GCD1b in COX-1 bound to OLEO, with upward translation in the first eigenvector. Segments GCD1a and GCD1b moved in unison in the first eigenvector of COX-1 bound to 1OL, where these residues moved downward with a moderate magnitude. For COX-1 bound to FLP, segment GCD1b moved in an anticlockwise direction. Residues of GCD1 block did not display concerted motion in COX-1 bound to MMHTE.

GCD2 block

Residues of GCD2 block in COX-1 displayed movement in the apo form and COX-1 bound to FLP, moving in an anticlockwise direction. These residues moved upwards with a moderate magnitude in COX-1 bound to 1OL. GCD2 block residues did not display concerted movement in forms of COX-1 bound to LG2, OLEO, and MMHTE.

For COX-2, residues in GCD2 block only displayed movement in the first eigenvector when bound to LG2 and OLEO, moving anticlockwise with a small magnitude in both cases.

GCD3 block

The movement of residues in GCD3 block was also divided into segments, with GCD3a spanning residues 233-264, and GCD3b spanning 265-290.

In COX-1, GCD3 block residues moved as a single unit. Downward motion was observed in the apo form and COX-1 bound to FLP. For COX-1 bound to 1OL, GCD3 block residues moved upwards with a small magnitude. Concerted motion of GCD3 block residues was not observed in COX-1 bound to LG2, OLEO, and MMHTE.

For COX-2 bound to 1OL, GCD3 block residues also moved as a single unit, translating in a downward motion with a small magnitude. Segment GCD3a displayed concerted motion in the apo form of the protein, as well as in COX-2 bound to MMHTE where upward motion of a small magnitude was observed. GCD3 block residues in COX-2 bound to LG2 all moved in a downward anticlockwise motion. It was observed that residues of segment GCD3a moved with a small magnitude of motion, while residues of GCD3b moved with a moderate magnitude of motion. Movement of GCD3 block residues in the first eigenvector was not observed in COX-2 bound to SAL and OLEO.

6.3.2.3   Overall characteristics of ligand-dependent concerted motions

The analysis of porcupine representations of protein motion in Figures 6.6 and Figure A4.5 in Appendix 9.4.3 show that movements of COX-1 and COX-2 do not appear to differ between the various ligand bound forms, nor between the unbound forms. Movement around the active site tended to be more conserved and restricted. The residues surrounding the outside of the protein appeared to display larger fluctuations, moving with a greater displacement compared to the rest of the protein. Residues moved with concerted motion in blocks of residues, with two of these blocks corresponding to the EGF domain and MBD of the COX proteins. The large fluctuation of these residues is consistent with RMSF data (Figures 6.2 and 6.3).

Along with the catalytic domain, the EGF domain forms the dimer interface, placing the two MBDs in a homodimer approximately 25 Å apart [9]. While it is unclear whether EGF domains have functional significance, they are commonly found in several families of membrane proteins and secreted proteins [9, 298]. Typically occurring in the primary sequence N-terminal to a membrane anchor, EGF domains are always located on the

extracytoplasmic face of the membrane [9]. It has been suggested that EGF domains of COX proteins play a role in the integration into the lipid bilayer [248, 299].

The MBD of COX contains four α-helices, one of which is merges upwards into the catalytic domain. Aromatic and hydrophobic residues of these helices form a large hydrophobic surface on the exterior of the protein, which interacts with the hydrophobic interior of the bilayer [300]. The depth of the COX membrane binding surface is only adequate for interaction with one leaflet of the lipid bilayer, representing a monotopic mechanism for integrating into the membrane [300].

The essential dynamics analysis performed in this section identified that residues of the EGF domain and MBD in blocks 1 and 2 of COX display concerted movements. These movements may have an effect on dimerization and membrane insertion of COX-1 and COX-2. While residues of GCD1-3 blocks are not known to have any functional significance, these blocks may form "dynamical domains" that contain functional modules. Residues that function together may be linked using this analysis, which may not be obvious from examination of protein structure alone. Further work will be required to elucidate the implications of these dynamical domains.

6.3.3   Network analysis using MONETA

The intra-protein communication pathways of all bound and unbound forms of COX were characterised using MONETA. A residue communication network was constructed based on the geometry of the protein, exploring long-range interactions and dynamical correlations to generate clusters of interacting residues calculated from MD trajectories.

6.3.3.1 Independent Dynamic Segments

Identification of Independent Dynamic Segments (IDSs) is based on Local Feature Analysis (LFA), a statistical technique that was originally developed for image analysis [169]. LFA extracts outputs of reduced dimensionality from PCA and was adapted to study essential dynamics in proteins, and implemented in MONETA to identify locally coupled residues, or IDSs [168].
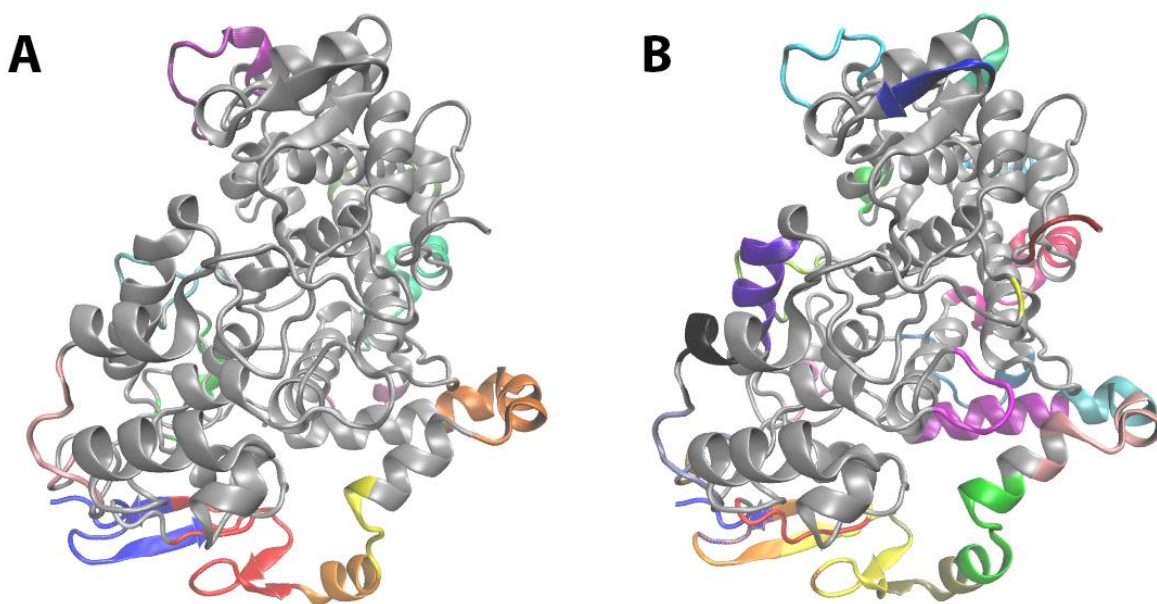
Fig 6.6: IDS of COX2 apo (A) and COX2 bound to oleocanthal (B). Individual IDSs are shown by coloured segments along the protein.

Similar to essential dynamics described in the previous section, LFA also utilises PCA to examine the dynamical behaviour of protein. In the case of LFA, regions fluctuating as independent units away from the rest of the protein were identified, with these IDSs describing 80% of the total atomic fluctuation. These patterns are shown in Figure 6.6 for the apo form of COX-2 and COX-2 bound to OLEO. It can be seen that the various IDSs are located around the outside of the protein away from the active site, with regions that overlap with the blocks identified in essential dynamics. In particular, overlapping residues occur in blocks 1 and 2 corresponding to the EGF and MBD, supporting the dynamic nature of these domains within COX. Residues of GCD1-3 blocks, which are currently not known to have any significant functions within COX, were also found to be included in the IDSs identified by MONETA, further suggesting that there may be implications for dynamical domains within proteins. It is believed that IDSs may play a crucial role in binding or allosteric propagation by shifting their energy content [301-303].

6.3.3.2 Protein communication network

Figure 6.7 represents the intra-protein network graph in the apo form of COX-2 and COX-2 bound to OLEO. Each residue is represented by a node linked by an edge indicating a connection—or communication path (CP)—between residues, such as peptide bonds or non-covalent interactions. Remaining network graphs for all forms of COX are shown in Figure A4.6 in Appendix 9.4.4. Network metrics describing these graphs are listed in Table A4.5 in Appendix 9.4.4.
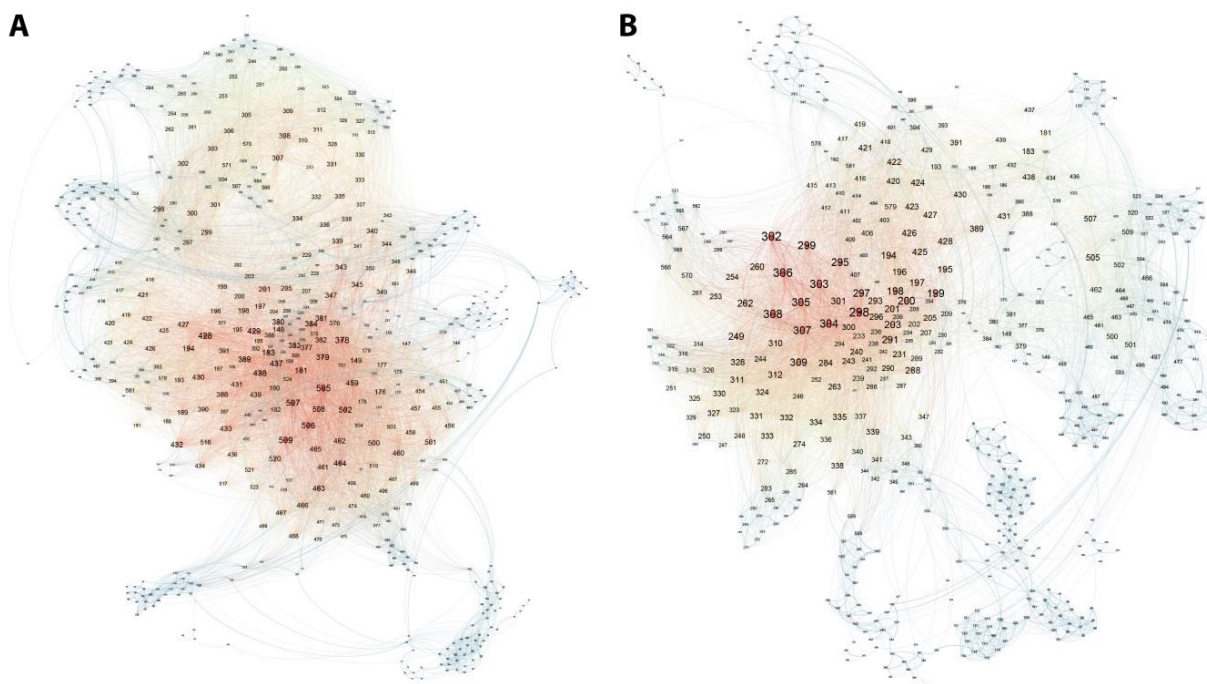
Figure 6.7: Communication pathway network for COX-2 in its apo form (A) and bound to OLEO (B). Nodes are coloured according to degree with red indicating that a residue has a higher number of connections. Nodes are labelled with their corresponding residue number, with a larger font indicating higher degrees.

From examination of the network graphs, changes in the communication network are more apparent in COX-2 compared to COX-1. In general, the apo form of COX-2 forms larger clusters of well-connected residues, the communications of which are disrupted when ligands are bound to the protein. The number of CPs diminishes as ligands are bound. The apo form of COX-2 contains 12,805 CPs, which is reduced when COX-2 is bound to SAL (9,937), 1OL (9,728), LG2 (9,940), OLEO (9,463), and MMHTE (8,340). These characteristics indicate that ligand binding causes a global perturbation in the communication network of COX-2. The most well-connected residues of the network are located away from the active site. This suggests that even if a well-connected residue were to be removed or altered, there may not be a significant impact on the ligand binding ability of COX isoenzymes.

Characteristics of communications between residues of the protein network were described by metrics measuring the influence of each node within the graph. The number of connections of a single node is described by degrees, as shown in Figure 6.7 for COX-2 in its apo form and bound to OLEO. Betweenness centrality measures how often a node appears on the shortest paths between nodes in the network. Residues with a high betweenness centrality may act as bottlenecks in a network that when disrupted, may render

101

communication within the protein network less efficient. Eigencentrality, or eigenvector centrality, is a measure of the node's importance based on how well connected a particular node is to other well-connected nodes. Closeness centrality refers to the average distance from a single node to all other nodes in the network. Another measure of importance is PageRank, which is based on the Google Algorithm to determine the probability of being at a page after a number of clicks [304].

Network metrics were examined for residues with known functions: the active site, peroxidase active site, EGF, MBD, and glycosylated residues. These values were expressed as a percentage change with respect to the apo form of COX-1 and COX-2, shown in Figure 6.8. Network metrics for these residues are listed in Table A4.6 in Appendix 9.4.4. From Figure 6.8 it can be observed that in general the degrees (A and B), betweenness centrality (C and D), and eigencentrality (E and F) tend to have fewer and sharper peaks in COX-2 compared to COX-1. This suggests that ligand binding has a greater influence on fewer residues in COX-2, showing more localised effects on these network properties of functional residues. For closeness centrality, COX-1 and COX-2 display a similar number of peaks with similar sharpness, with COX-2 generally displaying decreased closeness centrality as ligands are bound to the protein. There was no apparent trend in PageRank values (G and H).
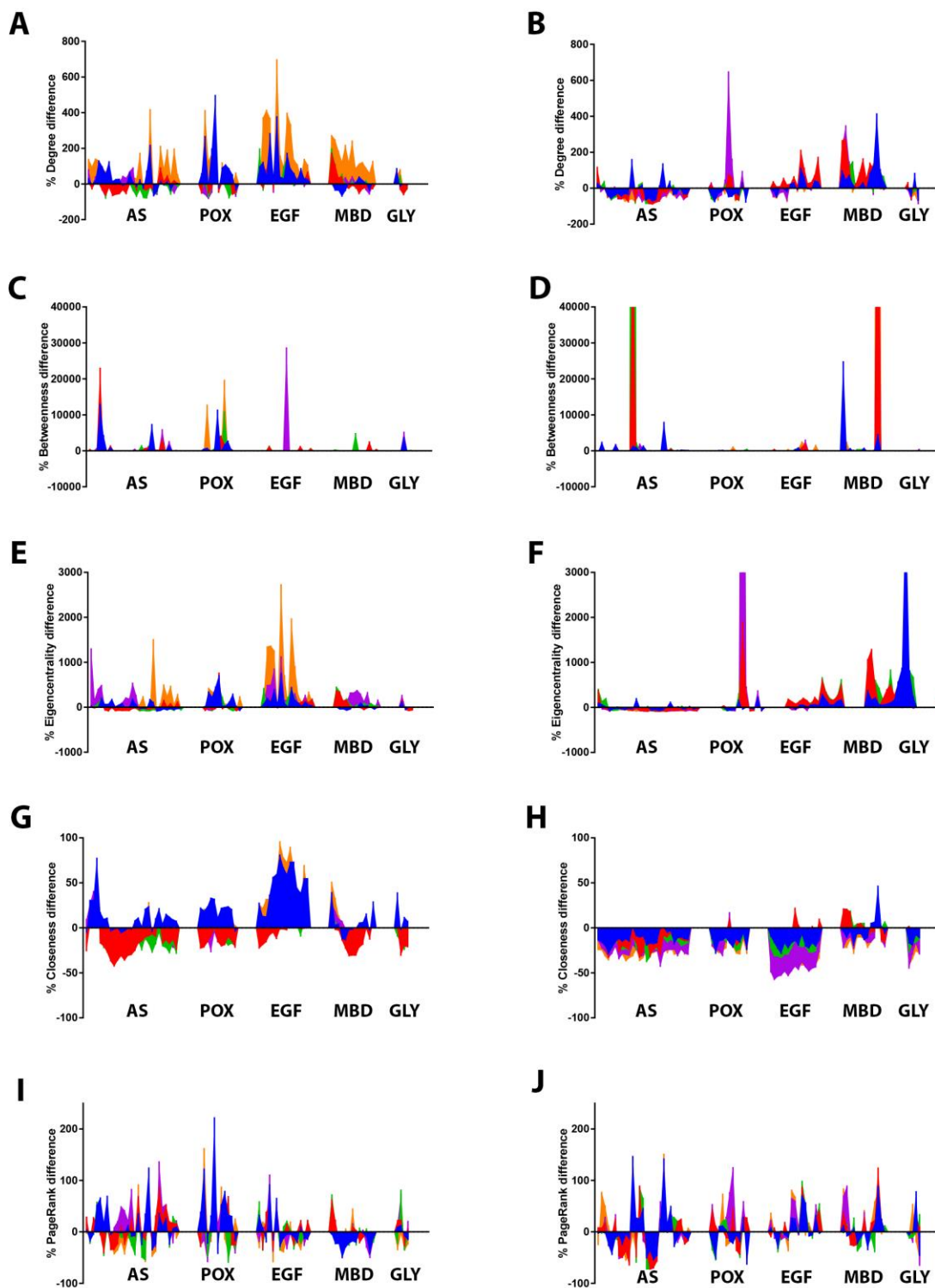
Figure 6.8: Percentage change in network metrics for functional residues of COX-1 (A, C, E, G, I) and COX-2 (B, D, F, H, J). Each ligand bound form of COX is represented by colours: native ligand (blue), 1OL (red), LG2 (green), OLEO (purple), and MMHTE (orange). Five clusters on each graph depict metrics for functional residues of COX: the active site (AS), peroxidase active site (POX), EGF domain, MBD, and glycosylated residues (GLY). Network metrics shown are shown as a percentage change with respect to apo: degrees (A-B), betweenness centrality (C-D), eigencentrality (E-F), closeness centrality (G-H), and PageRank (I-J).

Residues of the active site generally displayed a decrease in degrees after ligand binding for both COX-1 and COX-2, indicating that the communication of residues within the active site may be disrupted. An increase in the degrees of active site residues for COX-1 bound to FLP and MMHTE is observed, with HIS 513 displaying larger changes that are not present in the other ligands. While this is not observed in COX-2, binding to SAL causes a larger increase in degrees for the residues LEU 352 and PHE 518 of the active site. In general, the betweenness centrality of active site residues fluctuated moderately, with more localised peaks occurring. In COX-1, binding to FLP, 1OL, and LG2 displayed an increased betweenness centrality for ARG 120, while this value for LEU 352 in COX-2 was drastically increased with the protein bound to 1OL and LG2 compared to apo. Eigencentrality was mostly increased in ligand bound forms of COX-1, with greater peaks observed for LEU 93 of COX-1 bound to OLEO, and HIS 513 bound to MMHTE. There was modest fluctuation in the eigencentrality of active site residues in COX-2, displaying a general decrease with respect to apo. LEU 93 of COX-2 also had an increased eigencentrality when bound to 1OL and LG2. For closeness centrality, active site residues consistently had a decreased value when ligands were bound to COX-2. In COX-1, these values for the protein bound to FLP. MMHTE, and OLEO were increased, while closeness centrality decreased for COX-1 bound to 1OL and LG2. Across the ligand bound forms of COX-1, residue ARG 120 exhibited an increased value.

Compared to the apo form, residues of the peroxidase active site in ligand bound COX-1 generally exhibited increased network metrics. In particular, FLP and MMHTE bound COX-1 displayed higher degrees and closeness centrality compared to other ligands. On the other hand, peroxidase active site residues of ligand bound COX-2 displayed more moderate changes. ILE 274 in COX-2 bound to OLEO was observed to exhibit greatly increased values compared to other ligands for degrees, eigencentrality, and closeness centrality. Wang et al. found that bioflavonoids were able to bind to the peroxidase active site, functioning as reducing co-substrates for the COX enzymes [305]. This occurs through direct interaction with haematin, facilitating the electron transfer from bioflavonoids to haematin. When PGG_2 was used as a substrate, bioflavonoids were able to simulate the catalytic activity of COX-1 and -2 [305]. This may suggest that phenolic compounds may modulate the activity of COX through binding to sites distance from the main catalytic site, and changes associated with the peroxidase active site residues may have an effect on its function.

The EGF domain and MBD domain of the COX proteins displayed more moderate fluctuations in network metrics compared to the active site and peroxidase active site. Like other regions of the protein, FLP and MMHTE bound forms were observed to have increased values for degrees, eigencentrality, and closeness centrality in COX-1. In COX-1 bound to OLEO, ASP 52 was observed to have higher degrees compared to other ligands and residues. In the MBD of COX-2, VAL 103 was observed to have increased degrees and closeness centrality, as well as in FLP and OLEO bound protein for betweenness centrality and eigencentrality. Glycosylated residues did not display obvious trends in the fluctuation of network metrics with respect to apo, with the exception of a decreased closeness centrality in COX-2 which was also observed with other regions of the protein.

The identification of concerted motion and communication networks in proteins provides a technique for the characterisation of protein behaviour, allowing examination of more subtle changes that may not be immediately observable with conventional MD analysis methods. MONETA is a tool that can be used for identifying allosteric communication pathways, and further experiments are required to probe their existence in COX based on the network connections identified in this present study. MONETA has previously been useful for identifying changes in proteins with large conformational changes, as well as communications between spatially distant residues [306]. In the modular network analysis shown in this section, changes in network properties of residue communication network were identified. Further research may be necessary to examine the implications of these subtle changes in protein networks.

6.3.4   *Residue contributions to binding free energy: MM-PBSA*

Using trajectory data, the key residues contributing to ligand binding were elucidated using the MM-PBSA tool in GROMACS.
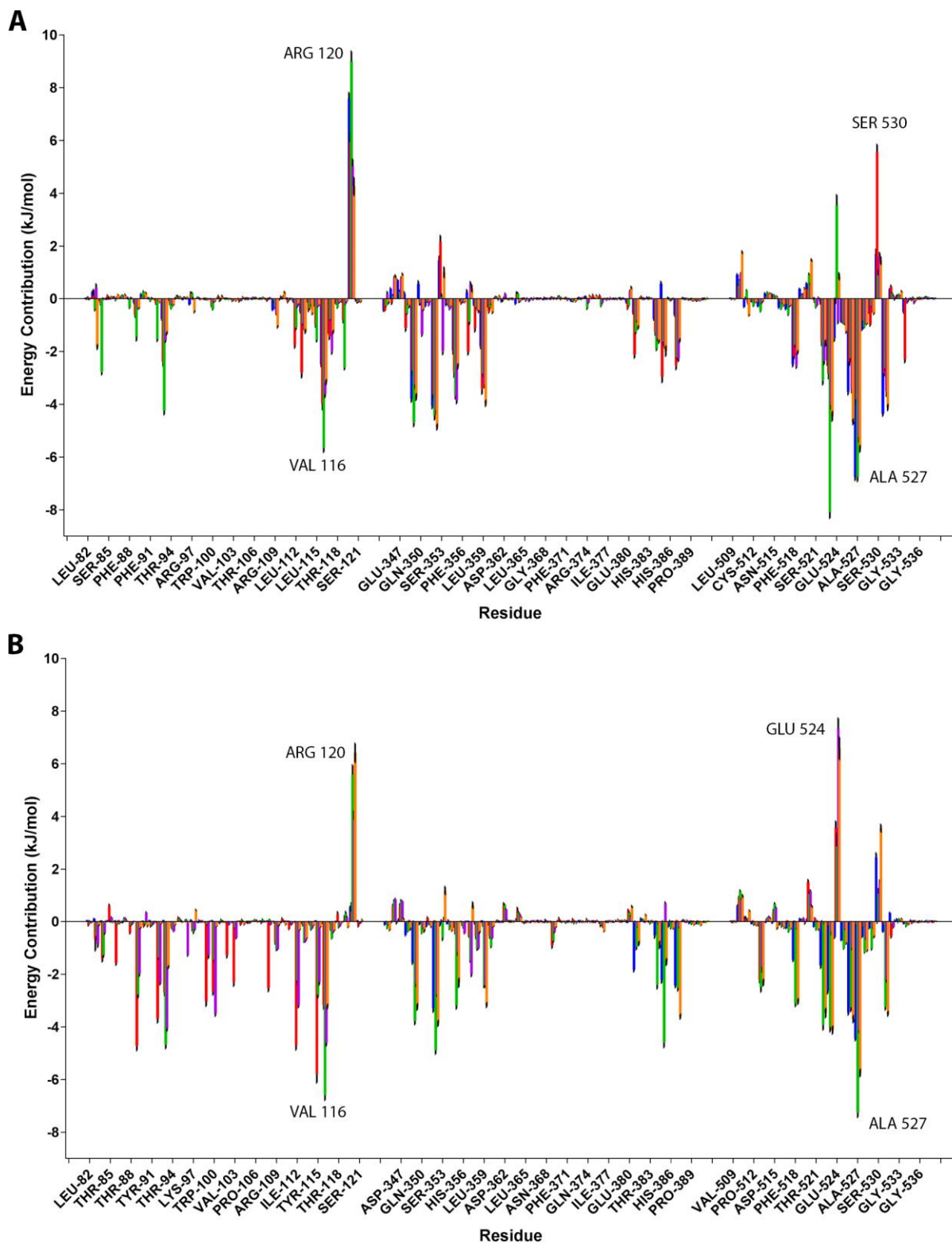
Figure 6.9: Residue contributions to binding of olive ligands to selected areas of COX-1 (A) and COX-2 (B) proteins. Each bound ligand is represented by colours: native ligand (blue), 1OL (red), LG2 (green), OLEO (purple), and MMHTE (orange). Error bars are pooled standard deviation from triplicate trajectories.

Figure 6.9 describes the per-residue contributions to the protein-ligand interactions to provide further insight into key interactions between the selected olive phenolic compounds and COX-1 and COX-2 enzymes. It is observed that there are three main regions along the amino acid sequence where more favourable and non-favourable interactions are occurring. Figure 6.9 shows these selected regions, while the full graph is shown in Figure A4.7 in Appendix 9.4.5. The main areas of residues contributing to ligand binding to COX-1 and COX-2 are residue numbers 82-121, 346-391, and 509-538. A more negative energy contribution corresponds to a more favourable interaction, while a more positive energy indicates contribution to a less favourable interaction with a specified protein residue.
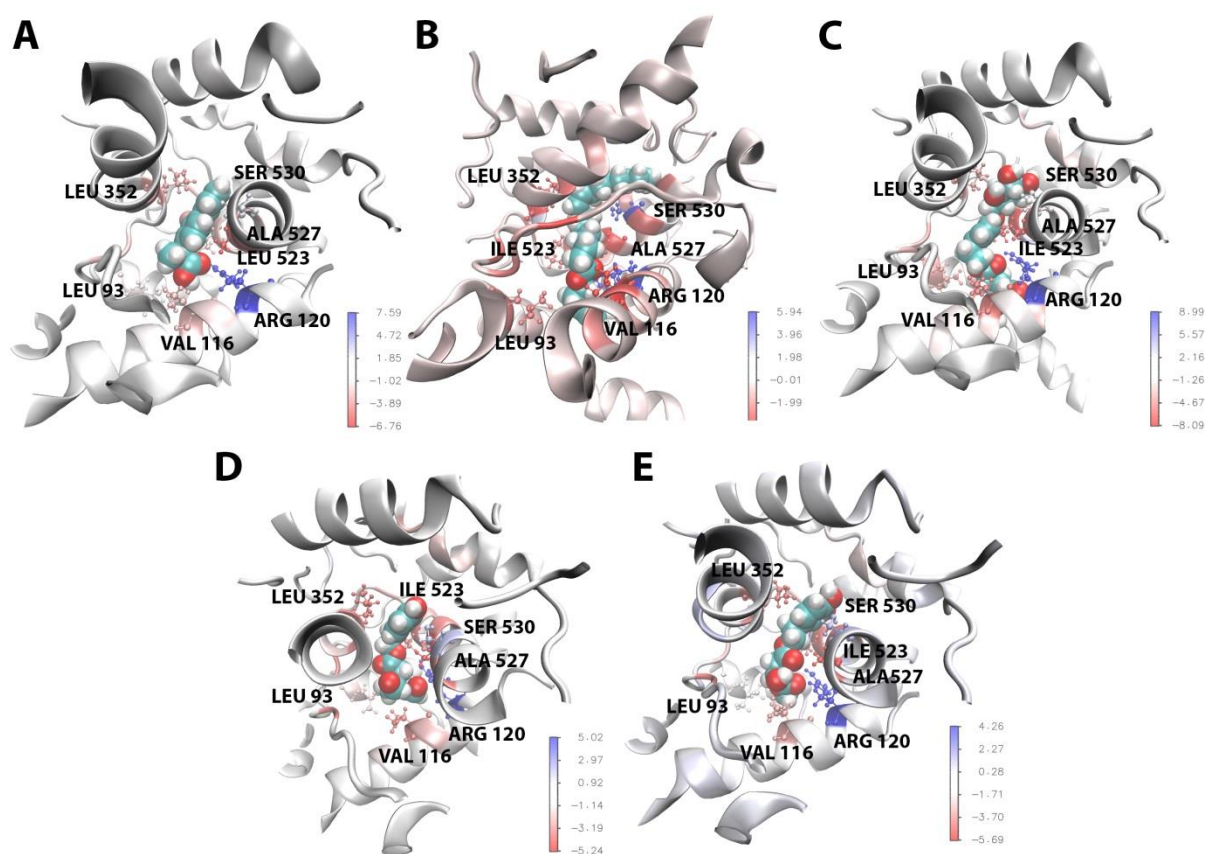
Figure 6.10: Residue contribution in kJ/mol to binding of olive ligands with COX-1: FLP (A), 1OL (B), LG2 (C), OLEO (C), and MMHTE (E). Residues are coloured according to their energy contribution in kJ/mol, coloured from blue to red with red indicating a more favourable contribution.
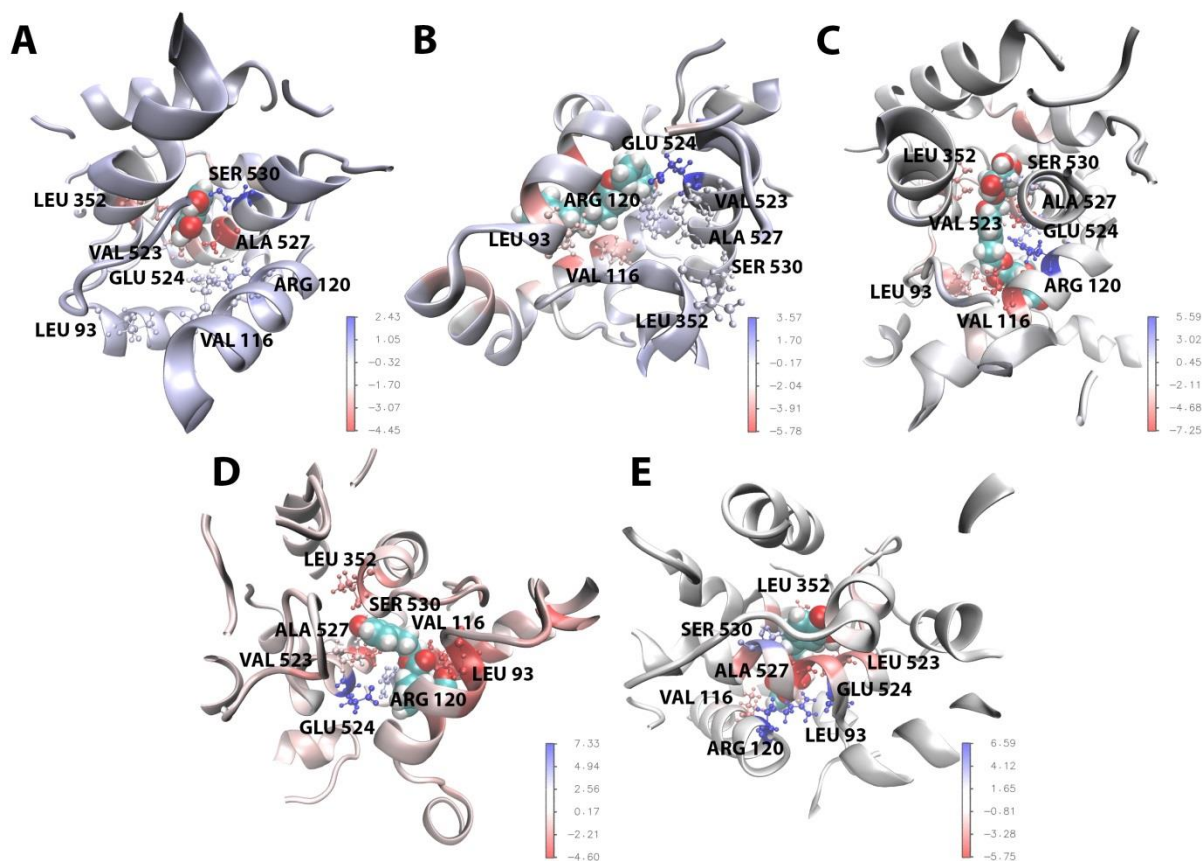
Figure 6.11: Residue contribution to binding of olive ligands with COX-2: SAL (A), 1OL (B), LG2 (C), OLEO (C), and MMHTE (E). Residues are coloured according to their energy contribution in kJ/mol, coloured from blue to red with red indicating a more favourable contribution.

Residues contributing favourably to binding of ligands to both COX-1 and COX-2 included LEU 93, VAL 116, LEU 352, and ALA 527. These interactions are depicted in Figures 6.10 and 6.11. 2D ligand interaction diagrams showing hydrogen bonds in these interactions are shown in Figures A4.8 and A4.9 in Appendix 9.4.5. ALA 527 contributed particularly favourable for ligand interaction, particularly in COX-1 for FLP (-6.76 kJ/mol) and LG2 (-6.75kJ/mol). ALA 527 forms part of the small hydrophobic pocket present in COX-2, along with VAL 349, SER 530, and LEU 531. In COX-2, ALA 527 had an energy contribution of -7.25 kJ/mol to binding with LG2. VAL 116 and LEU 352 also contributed favourably to ligand binding in both COX-1 and COX-2.

Other residues also contributing favourably to ligand interactions in both COX-1 and COX-2 include VAL 349, TYR 355, LEU 359, TYR 385, TRP 387. VAL 349 is located in the side pocket of COX-2. Mutagenesis of this residue resulted in alteration in the side pocket of COX-2 altering the kinetics of inhibition by indomethacin [307]. VAL 349 is also implicated in the ability of naproxen to bind to COX-2, with V349I and V349L mutants causing more sensitive inhibition [308]. Naproxen has also been shown to interact with LEU 359 and TRP 387, with TRP 387 interactions seemingly unique among carboxylate-containing compounds [308]. TYR 355 is situated on the opposite side from ARG 120, located at the entrance of the channel. Modification of TYR 355 in COX-1 has been suggested to alter the stereochemical specificity towards inhibitors containing 2-phenylpropionic acid groups, such as ibuprofen and flurbiprofen [300]. TYR 385 was shown to play a critical role in the acetylation of COX-2 by aspirin due to its hydrogen bonds possessing the ability to stabilise the negative charge of the tetrahedral intermediate formed during acetylation of SER 530 [225].

Favourable interactions were also observed with MET 522, GLY 526, and LEU 531. GLY 526 is situated on a helix between VAL 523 and SER 530, and has an important role in the positioning of reacting fatty acid intermediates in COX-2 [309]. MET 522 and GLY 526 are also implicated in COX-2 inhibition of diclofenac [224]. The rotation of LEU 531 is implicated in the binding of oxicams, a class of NSAIDs, by opening a novel pocket not utilised in the binding of other NSAIDs [310, 311]. LEU 531 also plays a minor role in aligning arachidonic acid optimally below TYR 385 for hydrogen abstraction [311].

6.3.4.2 Unfavourable energy contributions

Interestingly, ARG 120 was found to contribute unfavourably to protein interaction with all ligands in both isoenzymes, having high positive energy contributions ranging from 4.26 to 8.99 kJ/mol in COX-1, and 3.88 to 6.41 in COX-2 for LG2, oleocanthal, and MMHTE in COX-2. Similarly, GLU 524 is found to contribute unfavourably to protein interaction to ligands, particularly in COX-2 with oleocanthal (7.33 kJ/mol) and MMHTE (6.59 kJ/mol). SER 530 interactions are also shown to be unfavourable, especially with 1OL in COX-1 (5.57 kJ/mol), and MMHTE in COX-2 (3.55 kJ/mol).

ARG 120 is known to be important in the function of the cyclooxygenase enzyme, with appropriate positioning of the carboxylate of arachidonate to interact with this residue necessary for conversion to PGG$_2$ [220]. ARG 120 forms the most critical residue for binding of arachidonate binding in COX-1, forming an ionic linkage with the carboxylate group of arachidonate [220, 312]. It has been suggested that interactions of the carboxylate group of inhibitors with this residue is essential for time-dependent inhibition [312]. The K_m for arachidonate in a R120Q mutant of oPGHS-1 is 500-1000 times higher than that of native oPGHS-1 [312]. On the other hand, a R120Q mutant of hPGHS-2 has a similar K_m to that of native hPGHS-2, suggesting that hydrophobic residues of the active site may have a more significant effect in substrate binding [220, 313]. This may explain the slightly elevated interaction free energy between ARG 120 and olive ligands in COX-1 compared to COX-2.

Residues ARG 120 and GLU 524 are located in the binding site of COX-1 and COX-2, forming part of a restriction at the entrance along with TYR 355 [224]. SER 530 forms the site of acetylation in the mechanism of aspirin inhibition. SER 530 is situated below TYR 385, positioned such that its acetylation results in a blockage of access to the upper part of the binding channel [300]. There two residues have been found to cooperate in the chelation of negative charges, having a role in determining the selectivity of aspirin for covalent modification of SER 530 [224]. It was found that a mutation of ARG 120 to alanine had no effect on diclofenac inhibition of COX-2, however the S530A mutation produced a dramatic attenuation [224].

While these residues are shown here to contribute unfavourably to the ligand interaction, it may be that they instead have a role in maintaining the shape of the binding pocket and facilitate the orientation of ligands for an interaction more essential for catalysis. A crystal structure of ovine PGHS-1/*S*-flurbiprofen inhibitor complex demonstrates the formation of a

salt bridge between the carboxylate group of GLU 524 and ARG 120 [300]. In a mutagenesis study, it was concluded that GLU 524 does not play a role in either catalysis or substrate binding [312]. In the same study, the role of ARG 120 was determined to be for the affinity of the enzyme for arachidonate, rather than catalysis [312].

Additionally, crystallographic waters are present in the vicinity of residues with unfavourable energy contributions, including ARG 120. The energy contribution of water molecules are only indirectly taken into account in MM-PBSA methods, as solvent is represented as a continuous medium rather than an ensemble of explicit solvent molecules [314]. In MD simulations of COX-1 with carboxylated inhibitors, it was shown that a network of water molecules form a bridge between the guanidine head of ARG 120 and the hydroxyl group of SER 530 to enhance the polar interactions of the inhibitors [315]. The water bridge in the COX-1 binding pocket was shown to aid in stabilising the protein in an open conformation by preventing the LEU 531 side chain from rotating back into the closed conformation [315]. Since water molecules may play a role in mediating key interactions between the ligand and protein, further work may be required to take this into account.

Another known caveat of the MM-PBSA method is the lack of conformational entropy. Only the local stiffness of the utilised binding conformation is considered, while information on possible conformational changes of the ligand or protein is excluded from analysis [314]. Compared to other methods, MM-PBSA produces energies with poor precision, often better than docking but worse than free energy perturbation methods [314, 316]. Nevertheless, it is a useful tool for the understanding of affinities and trends observed in docking studies, as well as highlighting key residues contributing to ligand interactions.

6.3.4.3 Ligand interactions

There were some minor differences in residue interactions between COX-1/2 and various ligands. For the native ligands, interactions with FLP in COX-1 were not significantly different from other ligands. With COX-2, SAL displayed some less favourable interaction with residues LEU 93, VAL 116 and ARG 513. While there was a positive energy contribution to GLU 524 with olive ligands, this was not the case with SAL where there was -0.58 kJ/mol of energy contributed towards the interaction.

1OL had more favourable energy contributions from ILE112, PHE 381, and LEU 354 in COX-1 compared to other residues. Its interaction with SER 530 was observed to be less favourable, with an energy contribution of 5.57 kJ/mol in 1OL compared to approximately

1.50 kJ/mol with other ligands. In interactions with COX-2, there were more negative energy contributions with residues PRO 86, PHE 99, VAL 103, LEU 108, and ILE 112. Along with SAL, 1OL displayed less positive energy contributions with ARG 120 compared to other olive ligands, for example 0.38 kJ/mol in SAL compared to 6.42 kJ/mol in MMHTE. While other ligands had a favourable interaction with LEU 352 ranging from -4.88 to -2.76 kJ/mol, energy contribution of this residue towards 1OL was of a lesser magnitude (-0.01 kJ/mol).

In COX-1 interactions with LG2, there were more favourable interactions with PRO 84, LEU 115 and VAL 119, where energy contributions from those residues with other ligands were not apparent. Stronger energies were also observed with LEU 93 and VAL 116. Energy contribution of GLU 524 was negative with all ligands, except for LG2 where there was a positive energy contribution of 3.54 kJ/mol. While all ligands had strong interactions with ARG 120 and ILE 523, energy contributions with LG2 were more exaggerated, being the most positive and most negative respectively. Similarly, in COX-2 more negative energy contributions were observed in interactions with LG2 for residues VAL 349, LEU 352, TYR 355, and TYR 385. Interaction with ALA 527 was very favourable, producing an energy contribution of -7.25 kJ/mol.

For OLEO, there was a slightly more favourable energy interaction with LEU 117 in COX-1. Energy contributions with SER 353 were positive for all ligands in COX-1, except for OLEO where a negative energy contribution of -1.99 kJ/mol was observed. Interactions with COX-2 residues were largely similar to other ligands. Along with MMHTE, a more positive contribution was observed with GLU 524, producing energy of 7.33 kJ/mol.

Interactions with COX-1 and MMHTE were largely similar to other ligands. While ARG 83 had a slightly positive contribution with most ligands, a negative energy contribution of -1.83 kJ/mol was observed with MMHTE. Slightly more positive energies were observed for GLU 510 and GLU 520. In COX-2, a more positive energy contribution of 6.59 kJ/mol with ARG 120 was shown, as well as a slightly higher energy of 1.19 kJ/mol with SER 353 compared to other ligands.

6.3.4.4 Contrasting COX-1 and COX-2

One of the main differences between the two COX isoforms is the substitution of valine with isoleucine in COX-1. Being located at the entrance of the binding pocket, this residue plays a key role in determining the selectivity of a residue to either COX-1 or COX-2. This substitution causes the COX-2 active site to be roughly 20% larger than that of COX-1, as the

bulkier side chain of ILE 523 sterically inhibits access to the side pocket in the binding site [275, 317]. The volume of the inhibitor binding site and secondary pocket in COX-2 was calculated to be 394 Å^3 in COX-2, while the volume of the binding site of COX-1 was 316 Å^3 [232, 318]. The MM-PBSA data shown here indicates that ligands bind strongly with ILE 523 in COX-1, with ILE 523 contributing favourably, especially in the binding of LG2 with an energy contribution of -8.03 kJ/mol. Interaction energy is also favourable with other ligands (FLP -2.36; 1OL -2.85; oleocanthal -4.02; MMHTE -4.46 kJ/mol). In COX-2, VAL 523 also contributes favourably to the ligand interaction, albeit to a lesser degree (SAL -2.63; 1OL -0.15; LG2 -3.99; oleocanthal -0.37; MMHTE -4.12 kJ/mol).

In COX-2, ILE 434 in COX-1 in replaced with valine, which allows neighbouring PHE 518 to swing out of the way to further enable access to the side cavity [317]. The PBSA data shows that PHE 518 consistently contributes approximately -2 kJ/mol towards interactions with COX-1, while in COX-2 energy contribution ranges from -0.02 kJ/mol with 1OL to approximately -3 kJ/mol with LG2 and MMHTE. This variation between the isoforms may suggest that the movement of PHE 518 in COX-1 is more restricted than in COX-2, where the orientation of the side chain may be more flexible.

Another key difference between COX-1 and COX-2 is that of the replacement of arginine with histidine at residue 513 in COX-1. Rather than altering the shape of the binding pocket, the chemical environment is changed, allowing polar moieties to interact in the side pocket of COX-2 [317]. This is also shown in the PBSA data, where HIS 513 in COX-1 does not contribute strongly to the interaction (approximately -0.20 kJ/mol) compared to ARG 513 in COX-2. Energy contribution with this residue is stronger with 1OL (-2.20 kJ/mol), LG2 (-2.48 kJ/mol), oleocanthal (-1.71 kJ/mol), and MMHTE (-2.31 kJ/mol).

6.4 <u>Conclusion</u>

In this chapter, the mechanisms of selected olive compounds binding to COX were studied using MD simulations to examine their inhibitory potential. Additional metrics complementing binding free energy were calculated to determine the influence of ligand binding on the protein. The stability of protein-ligand complexes was studied using classical MD simulations, and functional insights into the structure of the protein were gained through examining protein dynamics with porcupine plot analysis and modular network analysis. The

contribution of individual residues to ligand binding was analysed using MM-PBSA methods.

Essential dynamics and network analysis identified distinct dynamical modules, which show some differences in motions depending on apo or ligand binding forms. Some of these dynamical modules might have functional consequences, especially in those of ligands affection motions in the EGF and MBD modules.

Network analysis showed that COX-2 inter-residue connectivity was generally more sensitive to ligand binding than COX-1. COX-2 demonstrated far greater disturbances in network properties compared to COX-1, although these changes were localised at a few distinct residues.

It was found through MM-PBSA that some residues of known importance for binding, such as ARG 120, paradoxically contributed unfavourably in terms of free energies. Rather than direct binding with ligands, the importance of ARG 120 may instead lie in maintaining structural integrity of the binding site. It may also be plausible that ARG 120 forms vital indirect interactions with ligands via water or ion molecules. This requires further exploration with more rigorous free energy methods.

7. Conclusions

In this thesis, computational techniques were employed to examine the potential of olive derived compounds in inhibiting COX-1 and COX-2 enzymes. Following the creation of an olive phenolic library, compounds were screened using docking and ADMET methods to select candidate compounds for further analysis using MD simulations. Molecular docking against the COX isoenzymes yielded two novel phenolic compounds: 1-oleyltyrosol and ligstroside derivative 2. Based on both ADMET and docking analysis, methyl malate-β-hydroxytyrosol ester was another novel compound that was selected for further analysis. Along with oleocanthal, an initial insight was gained in the membrane permeability of these compounds using steered MD simulations and umbrella sampling.

As well as the native ligands, olive phenolic compounds bound to COX-1 and COX-2 proteins were analysed using MD simulations to study the mechanisms of COX inhibition. It was found that in addition to being the most flexible, the MBD and EGF-like domain of the COX proteins displayed dynamic behaviour where differences in concerted motion were observed. Additional dynamical domains of unknown functional significance were also identified, which may be explored in further work. Residue communication networks of the different forms of COX were also examined, which demonstrated that residues of COX-2 exhibited greater disturbances in connectivity at distinct residues compared to COX-1. MM-PBSA elucidated the energy contribution of individual residues to binding of phenolic ligands. While key binding site residues were highlighted, it was revealed that some residues of known importance contributed unfavourably to ligand binding, such as ARG 120. Further examination using more rigorous free energy methods may be required.

In addition to further examination of the protein dynamics and network as mentioned above, future research would also focus on refining the membrane permeability study of the olive ligands. The work presented in Chapter 5 would be performed with adjustments to parameters, such as longer equilibration times, for the production of a more accurate estimate of free energy. MMHTE has been shown in this study to be a promising candidate as a potential therapeutic in COX inhibition. Since it is not yet commercially available, this

compound was selected to be synthesised. Future studies will be performed utilising this compound in *in vitro* studies.

Overall, this thesis has identified potential candidate compounds for the inhibition of COX enzymes, as well as the elucidation of mechanisms of action using *in silico* methodologies. With further work using *in vivo* and *in vitro* studies, novel olive phenolic compounds may provide an alternative in anti-inflammatory therapies. In addition, this study has provided further insight into the mechanism of more subtle, acute benefits that may be accumulative with consistent consumption of dietary phenolics.

8. References

1.   Crofford, L.J., *COX-1 and COX-2 tissue expression: implications and predictions.* J Rheumatol Suppl, 1997. **49**: p. 15-9.
2.   Whittle, B.J.R., et al., *Selective inhibition of prostaglandin production in inflammatory exudates and gastric mucosa.* Nature, 1980. **284**(5753): p. 271-273.
3.   Conaghan, P.G., *A turbulent decade for NSAIDs: update on current concepts of classification, epidemiology, comparative efficacy, and toxicity.* Rheumatology International, 2012. **32**(6): p. 1491-1502.
4.   Vane, J.R. and R.M. Botting, *Mechanism of action of nonsteroidal anti-inflammatory drugs.* Am J Med, 1998. **104**(3a): p. 2S-8S; discussion 21S-22S.
5.   Chandrasekharan, N.V., et al., *COX-3, a cyclooxygenase-1 variant inhibited by acetaminophen and other analgesic/antipyretic drugs: cloning, structure, and expression.* Proc Natl Acad Sci U S A, 2002. **99**(21): p. 13926-31.
6.   Zarghi, A. and S. Arfaei, *Selective COX-2 Inhibitors: A Review of Their Structure-Activity Relationships.* Iranian Journal of Pharmaceutical Research : IJPR, 2011. **10**(4): p. 655-683.
7.   Davies, N.M., et al., *Cyclooxygenase-3: axiom, dogma, anomaly, enigma or splice error?--Not as easy as 1, 2, 3.* J Pharm Pharm Sci, 2004. **7**(2): p. 217-26.
8.   Picot, D., P.J. Loll, and R.M. Garavito, *The X-ray crystal structure of the membrane protein prostaglandin H2 synthase-1.* Nature, 1994. **367**(6460): p. 243-249.
9.   Garavito, R.M., M.G. Malkowski, and D.L. DeWitt, *The structures of prostaglandin endoperoxide H synthases-1 and -2.* Prostaglandins Other Lipid Mediat, 2002. **68-69**: p. 129-52.
10.   Otto, J.C. and W.L. Smith, *Photolabeling of Prostaglandin Endoperoxide H Synthase-1 with 3-Trifluoro-3-(m-Iiodophenyl)diazirine as a Probe of Membrane Association and the Cyclooxygenase Active Site.* Journal of Biological Chemistry, 1996. **271**(17): p. 9906-9910.
11.   Otto, J.C., D.L. DeWitt, and W.L. Smith, *N-glycosylation of prostaglandin endoperoxide synthases-1 and -2 and their orientations in the endoplasmic reticulum.* Journal of Biological Chemistry, 1993. **268**(24): p. 18234-18242.
12.   Smith, W.L. and I. Song, *The enzymology of prostaglandin endoperoxide H synthases-1 and -2.* Prostaglandins & Other Lipid Mediators, 2002. **68–69**: p. 115-128.
13.   Hata, A.N. and R.M. Breyer, *Pharmacology and signaling of prostaglandin receptors: Multiple roles in inflammation and immune modulation.* Pharmacology & Therapeutics, 2004. **103**(2): p. 147-166.
14.   Samad, T.A., A. Sapirstein, and C.J. Woolf, *Prostanoids and pain: unraveling mechanisms and revealing therapeutic targets.* Trends Mol Med, 2002. **8**(8): p. 390-6.
15.   Tilley, S.L., T.M. Coffman, and B.H. Koller, *Mixed messages: modulation of inflammation and immune responses by prostaglandins and thromboxanes.* The Journal of Clinical Investigation, 2001. **108**(1): p. 15-23.
16.   Davis, R.J., et al., *EP4 prostanoid receptor-mediated vasodilatation of human middle cerebral arteries.* Br J Pharmacol, 2004. **141**(4): p. 580-5.

17.   Walch, L., et al., *Prostanoid EP(1)- and TP-receptors involved in the contraction of human pulmonary veins.* British Journal of Pharmacology, 2001. **134**(8): p. 1671-1678.

18.   Breyer, M.D. and R.M. Breyer, *G protein-coupled prostanoid receptors and the kidney.* Annu Rev Physiol, 2001. **63**: p. 579-605.

19.   Kennedy, C.R., et al., *Salt-sensitive hypertension and reduced fertility in mice lacking the prostaglandin EP2 receptor.* Nat Med, 1999. **5**(2): p. 217-20.

20.   Miner, J. and A. Hoffhines, *The Discovery of Aspirin's Antithrombotic Effects.* Texas Heart Institute Journal, 2007. **34**(2): p. 179-186.

21.   Vane, J.R. and R.M. Botting, *Anti-inflammatory drugs and their mechanism of action.* Inflamm Res, 1998. **47 Suppl 2**: p. S78-87.

22.   Stone, E., *An Account of the Success of the Bark of the Willow in the Cure of Agues. In a Letter to the Right Honourable George Earl of Macclesfield, President of R. S. from the Rev. Mr. Edmund Stone, of Chipping-Norton in Oxfordshire.* Philosophical Transactions, 1763. **53**: p. 195-200.

23.   Jack, D.B., *One hundred years of aspirin.* The Lancet, 1997. **350**(9075): p. 437-439.

24.   Maclagan, T., *The Treatment of Rheumatism by Salicin and Salicylic Acid.* British Medical Journal, 1876. **1**(803): p. 627-627.

25.   Vane, J.R., *Inhibition of prostaglandin synthesis as a mechanism of action for aspirin-like drugs.* Nat New Biol, 1971. **231**(25): p. 232-5.

26.   McGettigan, P. and D. Henry, *Use of Non-Steroidal Anti-Inflammatory Drugs That Elevate Cardiovascular Risk: An Examination of Sales and Essential Medicines Lists in Low-, Middle-, and High-Income Countries.* PLOS Medicine, 2013. **10**(2): p. e1001388.

27.   Singh, G., *Gastrointestinal complications of prescription and over-the-counter nonsteroidal anti-inflammatory drugs: a view from the ARAMIS database. Arthritis, Rheumatism, and Aging Medical Information System.* Am J Ther, 2000. **7**(2): p. 115-21.

28.   *Recommendations for the medical management of osteoarthritis of the hip and knee: 2000 update. American College of Rheumatology Subcommittee on Osteoarthritis Guidelines.* Arthritis Rheum, 2000. **43**(9): p. 1905-15.

29.   Zhang, W., et al., *EULAR evidence based recommendations for the management of hip osteoarthritis: report of a task force of the EULAR Standing Committee for International Clinical Studies Including Therapeutics (ESCISIT).* Annals of the Rheumatic Diseases, 2005. **64**(5): p. 669-681.

30.   Zhang, W., et al., *OARSI recommendations for the management of hip and knee osteoarthritis, Part II: OARSI evidence-based, expert consensus guidelines.* Osteoarthritis Cartilage, 2008. **16**(2): p. 137-62.

31.   Jones, R., *Nonsteroidal anti-inflammatory drug prescribing: past, present, and future.* The American Journal of Medicine, 2001. **110**(1, Supplement 1): p. S4-S7.

32.   Ward, R., *Identifying and assessing benefit-risk in primary care--a family physician's perspective.* Rheumatology (Oxford), 2010. **49 Suppl 2**: p. ii18-ii23.

33.   García Rodríguez, L.A. and L. Barreales Tolosa, *Risk of Upper Gastrointestinal Complications Among Users of Traditional NSAIDs and COXIBs in the General Population.* Gastroenterology, 2007. **132**(2): p. 498-506.

34.   Warner, T.D., et al., *Nonsteroid drug selectivities for cyclo-oxygenase-1 rather than cyclo-oxygenase-2 are associated with human gastrointestinal toxicity: A full in vitro analysis.* Proceedings of the National Academy of Sciences, 1999. **96**(13): p. 7563-7568.

35. Laine, L., *Approaches to nonsteroidal anti-inflammatory drug use in the high-risk patient.* Gastroenterology, 2001. **120**(3): p. 594-606.

36. Singh, G., *Recent considerations in nonsteroidal anti-inflammatory drug gastropathy.* Am J Med, 1998. **105**(1b): p. 31s-38s.

37. Lanas, A., et al., *A nationwide study of mortality associated with hospital admission due to severe gastrointestinal events and those associated with nonsteroidal antiinflammatory drug use.* Am J Gastroenterol, 2005. **100**(8): p. 1685-93.

38. Antman, E.M., et al., *Use of nonsteroidal antiinflammatory drugs: an update for clinicians: a scientific statement from the American Heart Association.* Circulation, 2007. **115**(12): p. 1634-42.

39. Fitzgerald, G.A., *Coxibs and cardiovascular disease.* N Engl J Med, 2004. **351**(17): p. 1709-11.

40. Kearney, P.M., et al., *Do selective cyclo-oxygenase-2 inhibitors and traditional non-steroidal anti-inflammatory drugs increase the risk of atherothrombosis? Meta-analysis of randomised trials.* Bmj, 2006. **332**(7553): p. 1302-8.

41. Bombardier, C., et al., *Comparison of upper gastrointestinal toxicity of rofecoxib and naproxen in patients with rheumatoid arthritis. VIGOR Study Group.* N Engl J Med, 2000. **343**(21): p. 1520-8, 2 p following 1528.

42. Bresalier, R.S., et al., *Cardiovascular events associated with rofecoxib in a colorectal adenoma chemoprevention trial.* N Engl J Med, 2005. **352**(11): p. 1092-102.

43. Sun, S.X., et al., *Withdrawal of COX-2 selective inhibitors rofecoxib and valdecoxib: impact on NSAID and gastroprotective drug prescribing and utilization.* Curr Med Res Opin, 2007. **23**(8): p. 1859-66.

44. Greenberg, J.D., et al., *The COX-2 inhibitor market withdrawals and prescribing patterns by rheumatologists in patients with gastrointestinal and cardiovascular risk.* Clin Exp Rheumatol, 2009. **27**(3): p. 395-401.

45. McGettigan, P. and D. Henry, *Use of non-steroidal anti-inflammatory drugs that elevate cardiovascular risk: an examination of sales and essential medicines lists in low-, middle-, and high-income countries.* PLoS Med, 2013. **10**(2): p. e1001388.

46. Warner, T.D. and J.A. Mitchell, *COX-2 selectivity alone does not define the cardiovascular risks associated with non-steroidal anti-inflammatory drugs.* Lancet, 2008. **371**(9608): p. 270-3.

47. McGettigan, P. and D. Henry, *Cardiovascular risk and inhibition of cyclooxygenase: a systematic review of the observational studies of selective and nonselective inhibitors of cyclooxygenase 2.* Jama, 2006. **296**(13): p. 1633-44.

48. Nissen, S.E., et al., *Cardiovascular Safety of Celecoxib, Naproxen, or Ibuprofen for Arthritis.* New England Journal of Medicine, 2016. **375**(26): p. 2519-2529.

49. Hennekens, C.H., *Update on aspirin in the treatment and prevention of cardiovascular disease.* American Journal of Managed Care, 2002. **8**(22; SUPP): p. S691-S700.

50. Harris, R.E., et al., *Aspirin, ibuprofen, and other non-steroidal anti-inflammatory drugs in cancer prevention: a critical review of non-selective COX-2 blockade (review).* Oncol Rep, 2005. **13**(4): p. 559-83.

51. Greenhough, A., et al., *The COX-2/PGE 2 pathway: key roles in the hallmarks of cancer and adaptation to the tumour microenvironment.* Carcinogenesis, 2009. **30**(3): p. 377-386.

52. Zhou, Y., et al., *Nonsteroidal anti-inflammatory drugs can lower amyloidogenic Abeta42 by inhibiting Rho.* Science, 2003. **302**(5648): p. 1215-7.

53. Keys, A., et al., *The diet and 15-year death rate in the seven countries study.* Am J Epidemiol, 1986. **124**(6): p. 903-15.

54. Caramia, G., et al., *Virgin olive oil in preventive medicine: From legend to epigenetics.* European Journal of Lipid Science and Technology, 2012. **114**(4): p. 375-388.

55. Gimeno, E., et al., *The effects of harvest and extraction methods on the antioxidant content (phenolics, alpha-tocopherol, and beta-carotene) in virgin olive oil.* Food Chemistry, 2002. **78**(2): p. 207-211.

56. Owen, R.W., et al., *Phenolic compounds and squalene in olive oils: the concentration and antioxidant potential of total phenols, simple phenols, secoiridoids, lignansand squalene.* Food Chem Toxicol, 2000. **38**(8): p. 647-59.

57. Covas, M.-I., et al., *Minor Components of Olive Oil: Evidence to Date of Health Benefits in Humans.* Nutrition Reviews, 2006. **64**: p. S20-S30.

58. Omar, S.H., *Oleuropein in olive and its pharmacological effects.* Sci Pharm, 2010. **78**(2): p. 133-54.

59. Bulotta, S., et al., *Beneficial effects of the olive oil phenolic components oleuropein and hydroxytyrosol: focus on protection against cardiovascular and metabolic diseases.* J Transl Med, 2014. **12**: p. 219.

60. Covas, M.I., R. de la Torre, and M. Fito, *Virgin olive oil: a key food for cardiovascular risk protection.* Br J Nutr, 2015. **113 Suppl 2**: p. S19-28.

61. Kimura, Y. and M. Sumiyoshi, *Olive leaf extract and its main component oleuropein prevent chronic ultraviolet B radiation-induced skin damage and carcinogenesis in hairless mice.* J Nutr, 2009. **139**(11): p. 2079-86.

62. Perugini, P., et al., *Efficacy of oleuropein against UVB irradiation: preliminary evaluation.* Int J Cosmet Sci, 2008. **30**(2): p. 113-20.

63. Petroni, A., et al., *Inhibition of platelet aggregation and eicosanoid production by phenolic components of olive oil.* Thromb Res, 1995. **78**(2): p. 151-60.

64. Rosignoli, P., et al., *Effect of olive oil phenols on the production of inflammatory mediators in freshly isolated human monocytes.* J Nutr Biochem, 2013. **24**(8): p. 1513-9.

65. Khanal, P., et al., *p-HPEA-EDA, a phenolic compound of virgin olive oil, activates AMP-activated protein kinase to inhibit carcinogenesis.* Carcinogenesis, 2011. **32**(4): p. 545-53.

66. Elnagar, A.Y., P.W. Sylvester, and K.A. El Sayed, *(-)-Oleocanthal as a c-Met inhibitor for the control of metastatic breast and prostate cancers.* Planta Med, 2011. **77**(10): p. 1013-9.

67. Cusimano, A., et al., *Oleocanthal exerts antitumor effects on human liver and colon cancer cells through ROS generation.* Int J Oncol, 2017. **51**(2): p. 533-544.

68. Angeloni, C., et al., *Bioactivity of Olive Oil Phenols in Neuroprotection.* Int J Mol Sci, 2017. **18**(11).

69. Li, W., et al., *Inhibition of tau fibrillization by oleocanthal via reaction with the amino groups of tau.* J Neurochem, 2009. **110**(4): p. 1339-51.

70. Beauchamp, G.K., et al., *Phytochemistry: ibuprofen-like activity in extra-virgin olive oil.* Nature, 2005. **437**(7055): p. 45-6.

71. Breslin, P.A.S., T.N. Gingrich, and B.G. Green, *Ibuprofen as a Chemesthetic Stimulus: Evidence of a Novel Mechanism of Throat Irritation.* Chemical Senses, 2001. **26**(1): p. 55-65.

72. Bisswanger, H., *Enzyme assays.* Perspectives in Science, 2014. **1**(1): p. 41-55.

73. Nina, M., S. Berneche, and B. Roux, *Anchoring of a monotopic membrane protein: the binding of prostaglandin H2 synthase-1 to the surface of a phospholipid bilayer.* Eur Biophys J, 2000. **29**(6): p. 439-54.

74.   Limongelli, V., et al., *Molecular basis of cyclooxygenase enzymes (COXs) selective inhibition.* Proceedings of the National Academy of Sciences of the United States of America, 2010. **107**(12): p. 5411-5416.

75.   Lei, J., et al., *Mechanistic Insights into a Classic Wonder Drug—Aspirin.* Journal of the American Chemical Society, 2015. **137**(1): p. 70-73.

76.   Chavatte, P. and A. Farce, *A Computational View of COX-2 Inhibition.* Anti-Cancer Agents in Medicinal Chemistry, 2006. **6**(3): p. 239-249.

77.   Philippe, C. and F. Amaury, *A Computational View of COX-2 Inhibition.* Anti-Cancer Agents in Medicinal Chemistry, 2006. **6**(3): p. 239-249.

78.   Palomer, A., et al., *Identification of Novel Cyclooxygenase-2 Selective Inhibitors Using Pharmacophore Models.* Journal of Medicinal Chemistry, 2002. **45**(7): p. 1402-1411.

79.   Palomer, A., et al., *Structure-based design of cyclooxygenase-2 selectivity into ketoprofen.* Bioorg Med Chem Lett, 2002. **12**(4): p. 533-7.

80.   Kim, H.J., et al., *Computational studies of COX-2 inhibitors: 3D-QSAR and docking.* Bioorg Med Chem, 2004. **12**(7): p. 1629-41.

81.   Moitessier, N., et al., *Towards the development of universal, fast and highly accurate docking/scoring methods: a long way to go.* Br J Pharmacol, 2008. **153 Suppl 1**: p. S7-26.

82.   Rarey, M., et al., *A fast flexible docking method using an incremental construction algorithm.* J Mol Biol, 1996. **261**(3): p. 470-89.

83.   Perola, E., W.P. Walters, and P.S. Charifson, *A detailed comparison of current docking and scoring methods on systems of pharmaceutical relevance.* Proteins, 2004. **56**(2): p. 235-49.

84.   Trott, O. and A.J. Olson, *AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading.* J Comput Chem, 2010. **31**(2): p. 455-61.

85.   Korb, O., et al., *Potential and Limitations of Ensemble Docking.* Journal of Chemical Information and Modeling, 2012. **52**(5): p. 1262-1274.

86.   Cross, J.B., et al., *Comparison of Several Molecular Docking Programs: Pose Prediction and Virtual Screening Accuracy.* Journal of Chemical Information and Modeling, 2009. **49**(6): p. 1455-1474.

87.   Kalyaanamoorthy, S. and Y.-P.P. Chen, *Modelling and enhanced molecular dynamics to steer structure-based drug discovery.* Progress in Biophysics and Molecular Biology, 2014. **114**(3): p. 123-136.

88.   De Vivo, M., et al., *Role of Molecular Dynamics and Related Methods in Drug Discovery.* Journal of Medicinal Chemistry, 2016. **59**(9): p. 4035-4061.

89.   Brooijmans, N. and I.D. Kuntz, *Molecular recognition and docking algorithms.* Annu Rev Biophys Biomol Struct, 2003. **32**: p. 335-73.

90.   Wodak, S.J. and J. Janin, *Computer analysis of protein-protein interaction.* Journal of Molecular Biology, 1978. **124**(2): p. 323-342.

91.   Salemme, F.R., *An hypothetical structure for an intermolecular electron transfer complex of cytochromes c and b5.* Journal of Molecular Biology, 1976. **102**(3): p. 563-568.

92.   Kuntz, I.D., et al., *A geometric approach to macromolecule-ligand interactions.* Journal of Molecular Biology, 1982. **161**(2): p. 269-288.

93.   Schrödinger, *Maestro.* 2016, LLC: New York

94.   Cho, A.E., et al., *Importance of accurate charges in molecular docking: quantum mechanical/molecular mechanical (QM/MM) approach.* J Comput Chem, 2005. **26**(9): p. 915-31.

95.   Wang, Z., et al., *Comprehensive evaluation of ten docking programs on a diverse set of protein-ligand complexes: the prediction accuracy of sampling power and scoring power.* Physical Chemistry Chemical Physics, 2016. **18**(18): p. 12964-12975.

96.   Pagadala, N.S., K. Syed, and J. Tuszynski, *Software for molecular docking: a review.* Biophysical Reviews, 2017. **9**(2): p. 91-102.

97.   Bissantz, C., G. Folkers, and D. Rognan, *Protein-Based Virtual Screening of Chemical Databases. 1. Evaluation of Different Docking/Scoring Combinations.* Journal of Medicinal Chemistry, 2000. **43**(25): p. 4759-4767.

98.   Perola, E., W.P. Walters, and P.S. Charifson, *A detailed comparison of current docking and scoring methods on systems of pharmaceutical relevance.* Proteins: Structure, Function, and Bioinformatics, 2004. **56**(2): p. 235-249.

99.   Kaminski, G.A., et al., *Evaluation and reparametrization of the OPLS-AA force field for proteins via comparison with accurate quantum chemical calculations on peptides.* Journal of Physical Chemistry B, 2001. **105**(28): p. 6474-6487.

100.   Friesner, R.A., et al., *Glide: A New Approach for Rapid, Accurate Docking and Scoring. 1. Method and Assessment of Docking Accuracy.* Journal of Medicinal Chemistry, 2004. **47**(7): p. 1739-1749.

101.   Eldridge, M.D., et al., *Empirical scoring functions: I. The development of a fast empirical scoring function to estimate the binding affinity of ligands in receptor complexes.* J Comput Aided Mol Des, 1997. **11**(5): p. 425-45.

102.   Kola, I. and J. Landis, *Can the pharmaceutical industry reduce attrition rates?* Nature Reviews Drug Discovery, 2004. **3**: p. 711.

103.   Kola, I. and J. Landis, *Can the pharmaceutical industry reduce attrition rates?* Nat Rev Drug Discov, 2004. **3**(8): p. 711-5.

104.   Tsaioun, K., M. Bottlaender, and A. Mabondzo, *ADDME--Avoiding Drug Development Mistakes Early: central nervous system drug discovery perspective.* BMC Neurol, 2009. **9 Suppl 1**: p. S1.

105.   Avdeef, A., *The rise of PAMPA.* Expert Opin Drug Metab Toxicol, 2005. **1**(2): p. 325-42.

106.   van Breemen, R.B. and Y. Li, *Caco-2 cell permeability assays to measure drug absorption.* Expert Opin Drug Metab Toxicol, 2005. **1**(2): p. 175-85.

107.   Lee, C.T., et al., *Simulation-Based Approaches for Determining Membrane Permeability of Small Compounds.* Journal of chemical information and modeling, 2016. **56**(4): p. 721-733.

108.   van de Waterbeemd, H. and E. Gifford, *ADMET in silico modelling: towards prediction paradise?* Nat Rev Drug Discov, 2003. **2**(3): p. 192-204.

109.   Stenberg, P., et al., *Experimental and computational screening models for the prediction of intestinal drug absorption.* J Med Chem, 2001. **44**(12): p. 1927-37.

110.   Scheife, R.T., *Protein binding: what does it mean?* Dicp, 1989. **23**(7-8 Suppl): p. S27-31.

111.   Schuster, D., C. Laggner, and T. Langer, *Why drugs fail--a study on side effects in new chemical entities.* Curr Pharm Des, 2005. **11**(27): p. 3545-59.

112.   Low, Y., et al., *Predicting drug-induced hepatotoxicity using QSAR and toxicogenomics approaches.* Chem Res Toxicol, 2011. **24**(8): p. 1251-62.

113.   Vyas, V.K., et al., *Homology Modeling a Fast Tool for Drug Discovery: Current Perspectives.* Indian Journal of Pharmaceutical Sciences, 2012. **74**(1): p. 1-17.

114.   Chandonia, J.M. and S.E. Brenner, *Implications of structural genomics target selection strategies: Pfam5000, whole genome, and random approaches.* Proteins, 2005. **58**(1): p. 166-79.

115. Lesk, A.M. and C.H. Chothia, *The Response of Protein Structures to Amino-Acid Sequence Changes.* Philosophical Transactions of the Royal Society of London. Series A, Mathematical and Physical Sciences, 1986. **317**(1540): p. 345-356.

116. Gohlke, H., C. Kiel, and D.A. Case, *Insights into Protein–Protein Binding by Binding Free Energy Calculation and Free Energy Decomposition for the Ras–Raf and Ras–RalGDS Complexes.* Journal of Molecular Biology, 2003. **330**(4): p. 891-913.

117. de Graaf, C., et al., *Catalytic Site Prediction and Virtual Screening of Cytochrome P450 2D6 Substrates by Consideration of Water and Rescoring in Automated Docking.* Journal of Medicinal Chemistry, 2006. **49**(8): p. 2417-2430.

118. Diller, D.J. and R. Li, *Kinases, Homology Models, and High Throughput Docking.* Journal of Medicinal Chemistry, 2003. **46**(22): p. 4638-4647.

119. Evers, A. and T. Klabunde, *Structure-based Drug Discovery Using GPCR Homology Modeling: Successful Virtual Screening for Antagonists of the Alpha1A Adrenergic Receptor.* Journal of Medicinal Chemistry, 2005. **48**(4): p. 1088-1097.

120. McGovern, S.L. and B.K. Shoichet, *Information decay in molecular docking screens against holo, apo, and modeled conformations of enzymes.* J Med Chem, 2003. **46**(14): p. 2895-907.

121. Hongmao, S., *Chapter 4 - Homology Modeling and Ligand-Based Molecule Design*, in *A Practical Guide to Rational Drug Design.* 2016, Woodhead Publishing. p. 109-160.

122. Eswar, N., et al., *Comparative protein structure modeling using Modeller.* Curr Protoc Bioinformatics, 2006. **Chapter 5**: p. Unit-5.6.

123. Berman, H.M., et al., *The Protein Data Bank.* Nucleic Acids Research, 2000. **28**(1): p. 235-242.

124. Altschul, S.F., et al., *Basic local alignment search tool.* Journal of Molecular Biology, 1990. **215**(3): p. 403-410.

125. Pearson, W.R. and D.J. Lipman, *Improved tools for biological sequence comparison.* Proc Natl Acad Sci U S A, 1988. **85**(8): p. 2444-8.

126. Pitman, M.R. and R.I. Menz, *2 - Methods for Protein Homology Modelling*, in *Applied Mycology and Biotechnology*, D.K. Arora, R.M. Berka, and G.B. Singh, Editors. 2006, Elsevier. p. 37-59.

127. Rost, B., *Twilight zone of protein sequence alignments.* Protein Eng, 1999. **12**(2): p. 85-94.

128. Gribskov, M., A.D. McLachlan, and D. Eisenberg, *Profile analysis: detection of distantly related proteins.* Proc Natl Acad Sci U S A, 1987. **84**(13): p. 4355-8.

129. Henikoff, S. and J.G. Henikoff, *Position-based sequence weights.* J Mol Biol, 1994. **243**(4): p. 574-8.

130. Altschul, S.F., et al., *Gapped BLAST and PSI-BLAST: a new generation of protein database search programs.* Nucleic Acids Res, 1997. **25**(17): p. 3389-402.

131. Fiser, A., *Template-Based Protein Structure Modeling.* Methods in molecular biology (Clifton, N.J.), 2010. **673**: p. 73-94.

132. Brooks, B.R., et al., *CHARMM: A program for macromolecular energy, minimization, and dynamics calculations.* Journal of Computational Chemistry, 1983. **4**(2): p. 187-217.

133. Laskowski, R.A., et al., *{PROCHECK}: a program to check the stereochemical quality of protein structures.* J. Appl. Cryst., 1993. **26**: p. 283-291.

134. Hooft, R.W., et al., *Errors in protein structures.* Nature, 1996. **381**(6580): p. 272.

135. Ramachandran, G.N., C. Ramakrishnan, and V. Sasisekharan, *Stereochemistry of polypeptide chain configurations.* J Mol Biol, 1963. **7**: p. 95-9.

136. Hovmoller, S., T. Zhou, and T. Ohlson, *Conformations of amino acids in proteins.* Acta Crystallogr D Biol Crystallogr, 2002. **58**(Pt 5): p. 768-76.

137. Wiederstein, M. and M.J. Sippl, *ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins.* Nucleic Acids Research, 2007. **35**(Web Server issue): p. W407-W410.

138. Eisenberg, D., R. Luthy, and J.U. Bowie, *VERIFY3D: assessment of protein models with three-dimensional profiles.* Methods Enzymol, 1997. **277**: p. 396-404.

139. Berendsen, H.J.C., D. van der Spoel, and R. van Drunen, *GROMACS: A message-passing parallel molecular dynamics implementation.* Computer Physics Communications, 1995. **91**(1): p. 43-56.

140. Hermans, J., et al., *A consistent empirical potential for water–protein interactions.* Biopolymers, 1984. **23**(8): p. 1513-1518.

141. MacKerell, A.D., et al., *All-atom empirical potential for molecular modeling and dynamics studies of proteins.* J Phys Chem B, 1998. **102**(18): p. 3586-616.

142. Hospital, A., et al., *Molecular dynamics simulations: advances and applications.* Advances and Applications in Bioinformatics and Chemistry : AABC, 2015. **8**: p. 37-47.

143. Rueda, M., et al., *A consensus view of protein dynamics.* Proceedings of the National Academy of Sciences of the United States of America, 2007. **104**(3): p. 796-801.

144. Weiner, S.J., et al., *A new force field for molecular mechanical simulation of nucleic acids and proteins.* Journal of the American Chemical Society, 1984. **106**(3): p. 765-784.

145. Cornell, W.D., et al., *A Second Generation Force Field for the Simulation of Proteins, Nucleic Acids, and Organic Molecules.* Journal of the American Chemical Society, 1995. **117**(19): p. 5179-5197.

146. Damm, W., et al., *OPLS all‑atom force field for carbohydrates.* Journal of Computational Chemistry, 1997. **18**(16): p. 1955-1970.

147. Ott, K.-H. and B. Meyer, *Parametrization of GROMOS force field for oligosaccharides and assessment of efficiency of molecular dynamics simulations.* Journal of Computational Chemistry, 1996. **17**(8): p. 1068-1084.

148. Mackerell, A.D., Jr., *Empirical force fields for biological macromolecules: overview and issues.* J Comput Chem, 2004. **25**(13): p. 1584-604.

149. Guvench, O. and A.D. MacKerell, Jr., *Comparison of protein force fields for molecular dynamics simulations.* Methods Mol Biol, 2008. **443**: p. 63-88.

150. Swope, W.C., et al., *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, 1982. **76**(1): p. 637-649.

151. Leach, A.R., *Molecular modelling: principles and applications.* 2001: Pearson education.

152. Metropolis, N., et al., *Equation of State Calculations by Fast Computing Machines.* The Journal of Chemical Physics, 1953. **21**(6): p. 1087-1092.

153. Ewald, P.P., *Die Berechnung optischer und elektrostatischer Gitterpotentiale.* Annalen der Physik, 1921. **369**(3): p. 253-287.

154. Darden, T., D. York, and L. Pedersen, *Particle mesh Ewald: An N·log(N) method for Ewald sums in large systems.* The Journal of Chemical Physics, 1993. **98**(12): p. 10089-10092.

155. Essmann, U., et al., *A smooth particle mesh Ewald method.* The Journal of Chemical Physics, 1995. **103**(19): p. 8577-8593.

156. Berendsen, H.J.C., et al., *Molecular dynamics with coupling to an external bath.* The Journal of Chemical Physics, 1984. **81**(8): p. 3684-3690.

157. Nosé, S., *A molecular dynamics method for simulations in the canonical ensemble.* Molecular Physics, 1984. **52**(2): p. 255-268.

158. Hoover, W.G., *Canonical dynamics: Equilibrium phase-space distributions.* Physical Review A, 1985. **31**(3): p. 1695-1697.

159. Parrinello, M. and A. Rahman, *Polymorphic transitions in single crystals: A new molecular dynamics method.* Journal of Applied Physics, 1981. **52**(12): p. 7182-7190.

160. Jakopitsch, C., et al., *Role of the main access channel of catalase-peroxidase in catalysis.* J Biol Chem, 2005. **280**(51): p. 42411-22.

161. Ahalawat, N. and R.K. Murarka, *Conformational changes and allosteric communications in human serum albumin due to ligand binding.* J Biomol Struct Dyn, 2015. **33**(10): p. 2192-204.

162. Pearson, K., *LIII. On lines and planes of closest fit to systems of points in space.* The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science, 1901. **2**(11): p. 559-572.

163. David, C.C. and D.J. Jacobs, *Principal Component Analysis: A Method for Determining the Essential Dynamics of Proteins.* Methods in molecular biology (Clifton, N.J.), 2014. **1084**: p. 193-226.

164. Hotelling, H., *Analysis of a complex of statistical variables into principal components.* Journal of Educational Psychology, 1933. **24**(6): p. 417-441.

165. Abdi, H. and L.J. Williams, *Principal component analysis.* Wiley Interdisciplinary Reviews: Computational Statistics, 2010. **2**(4): p. 433-459.

166. Balsera, M.A., et al., *Principal Component Analysis and Long Time Protein Dynamics.* The Journal of Physical Chemistry, 1996. **100**(7): p. 2567-2572.

167. Amadei, A., et al., *An efficient method for sampling the essential subspace of proteins.* J Biomol Struct Dyn, 1996. **13**(4): p. 615-25.

168. Allain, A., et al., *Allosteric pathway identification through network analysis: from molecular dynamics simulations to interactive 2D and 3D graphs.* Faraday Discussions, 2014. **169**(0): p. 303-321.

169. Penev, P.S. and J.J. Atick, *Local feature analysis: a general statistical theory for object representation.* Network: Computation in Neural Systems, 1996. **7**(3): p. 477-500.

170. Zhang, Z. and W. Wriggers, *Local feature analysis: a statistical theory for reproducible essential dynamics of large macromolecules.* Proteins, 2006. **64**(2): p. 391-403.

171. Chennubhotla, C. and I. Bahar, *Signal Propagation in Proteins and Relation to Equilibrium Fluctuations.* PLOS Computational Biology, 2007. **3**(9): p. e172.

172. Atilgan, A.R., P. Akan, and C. Baysal, *Small-world communication of residues and significance for protein dynamics.* Biophys J, 2004. **86**(1 Pt 1): p. 85-91.

173. Hou, T., et al., *Characterization of Domain–Peptide Interaction Interface: Prediction of SH3 Domain-Mediated Protein–Protein Interaction Network in Yeast by Generic Structure-Based Models.* Journal of Proteome Research, 2012. **11**(5): p. 2982-2995.

174. Adcock, S.A. and J.A. McCammon, *Molecular Dynamics: Survey of Methods for Simulating the Activity of Proteins.* Chemical Reviews, 2006. **106**(5): p. 1589-1615.

175. Kollman, P.A., et al., *Calculating Structures and Free Energies of Complex Molecules: Combining Molecular Mechanics and Continuum Models.* Accounts of Chemical Research, 2000. **33**(12): p. 889-897.

176. Kuhn, B. and P.A. Kollman, *Binding of a Diverse Set of Ligands to Avidin and Streptavidin: An Accurate Quantitative Prediction of Their Relative Affinities by a Combination of Molecular Mechanics and Continuum Solvent Models.* Journal of Medicinal Chemistry, 2000. **43**(20): p. 3786-3791.

177. Rastelli, G., et al., *Fast and accurate predictions of binding free energies using MM-PBSA and MM-GBSA.* J Comput Chem, 2010. **31**(4): p. 797-810.

178. Brice, A.R. and B.N. Dominy, *Analyzing the robustness of the MM/PBSA free energy calculation method: Application to DNA conformational transitions.* Journal of Computational Chemistry, 2011. **32**(7): p. 1431-1440.

179. Brown, S.P. and S.W. Muchmore, *Large-Scale Application of High-Throughput Molecular Mechanics with Poisson−Boltzmann Surface Area for Routine Physics-Based Scoring of Protein−Ligand Complexes.* Journal of Medicinal Chemistry, 2009. **52**(10): p. 3159-3165.

180. Kumari, R., R. Kumar, and A. Lynn, *g_mmpbsa—A GROMACS Tool for High-Throughput MM-PBSA Calculations.* Journal of Chemical Information and Modeling, 2014. **54**(7): p. 1951-1962.

181. Baker, N.A., et al., *Electrostatics of nanosystems: Application to microtubules and the ribosome.* Proceedings of the National Academy of Sciences, 2001. **98**(18): p. 10037-10041.

182. Still, W.C., et al., *Semianalytical treatment of solvation for molecular mechanics and dynamics.* Journal of the American Chemical Society, 1990. **112**(16): p. 6127-6129.

183. Torrie, G.M. and J.P. Valleau, *Nonphysical sampling distributions in Monte Carlo free-energy estimation: Umbrella sampling.* Journal of Computational Physics, 1977. **23**(2): p. 187-199.

184. Kumar, S., et al., *THE weighted histogram analysis method for free-energy calculations on biomolecules. I. The method.* Journal of Computational Chemistry, 1992. **13**(8): p. 1011-1021.

185. Bonvino, N.P., et al., *OliveNet™: a comprehensive library of compounds from Olea europaea.* Database, 2018. **2018**: p. bay016-bay016.

186. de la Torre-Carbot, K., et al., *Characterization and quantification of phenolic compounds in olive oils by solid-phase extraction, HPLC-DAD, and HPLC-MS/MS.* J Agric Food Chem, 2005. **53**(11): p. 4331-40.

187. Boskou, G., et al., *Antioxidant capacity and phenolic profile of table olives from the Greek market.* Food Chemistry, 2006. **94**(4): p. 558-564.

188. Quirantes-Pine, R., et al., *HPLC-ESI-QTOF-MS as a powerful analytical tool for characterising phenolic compounds in olive-leaf extracts.* Phytochem Anal, 2013. **24**(3): p. 213-23.

189. Christophoridou, S., et al., *Separation and identification of phenolic compounds in olive oil by coupling high-performance liquid chromatography with postcolumn solid-phase extraction to nuclear magnetic resonance spectroscopy (LC-SPE-NMR).* J Agric Food Chem, 2005. **53**(12): p. 4667-79.

190. Hatzakis, E., et al., *Determination of phospholipids in olive oil by 31P NMR spectroscopy.* J Agric Food Chem, 2008. **56**(15): p. 6232-40.

191. Karkoula, E., et al., *Direct measurement of oleocanthal and oleacein levels in olive oil by quantitative (1)H NMR. Establishment of a new index for the characterization of extra virgin olive oils.* J Agric Food Chem, 2012. **60**(47): p. 11696-703.

192. Gómez-Caravaca, A.M., et al., *9 - Bioactive Phenolic Compounds from Olea europaea: A Challenge for Analytical Chemistry A2 - Boskou, Dimitrios*, in *Olive and Olive Oil Bioactive Constituents.* 2015, AOCS Press. p. 261-298.

193. Kim, S., et al., *PubChem Substance and Compound databases.* Nucleic Acids Res, 2016. **44**(D1): p. D1202-13.

194. Bastian, M., S. Heymann, and M. Jacomy, *Gephi: An Open Source Software for Exploring and Manipulating Networks.* 2009. 2009.

195. Gutfinger, T., *Polyphenols in olive oils.* Journal of the American Oil Chemists' Society, 1981. **58**(11): p. 966-968.

196. Tuck, K.L. and P.J. Hayball, *Major phenolic compounds in olive oil: metabolism and health effects.* The Journal of Nutritional Biochemistry, 2002. **13**(11): p. 636-644.

197. Cicerale, S., L.J. Lucas, and R.S. Keast, *Antimicrobial, antioxidant and anti-inflammatory phenolic activities in extra virgin olive oil.* Curr Opin Biotechnol, 2012. **23**(2): p. 129-35.

198. McGeer, P.L., E.G. McGeer, and C. Schwab, *Inflammatory Processes Exacerbate Degenerative Neurological Disorders*, in *Current Hypotheses and Research Milestones in Alzheimer's Disease*, R.B. Maccioni and G. Perry, Editors. 2009, Springer US: Boston, MA. p. 117-124.

199. Karlson, E.W., et al., *Biomarkers of inflammation and development of rheumatoid arthritis in women from two prospective cohort studies.* Arthritis & Rheumatism, 2009. **60**(3): p. 641-652.

200. Mathieu, P., et al., *Visceral Obesity.* The Link Among Inflammation, Hypertension, and Cardiovascular Disease, 2009. **53**(4): p. 577-584.

201. Solinas, G., et al., *Tumor-associated macrophages (TAM) as major players of the cancer-related inflammation.* Journal of Leukocyte Biology, 2009. **86**(5): p. 1065-1073.

202. Corona, G., J.P. Spencer, and M.A. Dessi, *Extra virgin olive oil phenolics: absorption, metabolism, and biological activities in the GI tract.* Toxicol Ind Health, 2009. **25**(4-5): p. 285-93.

203. Konstantinidou, V., et al., *In vivo nutrigenomic effects of virgin olive oil polyphenols within the frame of the Mediterranean diet: a randomized controlled trial.* Faseb j, 2010. **24**(7): p. 2546-57.

204. Khymenets, O., et al., *Mononuclear cell transcriptome response after sustained virgin olive oil consumption in humans: an exploratory nutrigenomics study.* Omics, 2009. **13**(1): p. 7-19.

205. Iacono, A., et al., *Effect of oleocanthal and its derivatives on inflammatory response induced by lipopolysaccharide in a murine chondrocyte cell line.* Arthritis & Rheumatism, 2010. **62**(6): p. 1675-1682.

206. Pitt, J., et al., *Alzheimer's-associated Aβ oligomers show altered structure, immunoreactivity and synaptotoxicity with low doses of oleocanthal.* Toxicology and Applied Pharmacology, 2009. **240**(2): p. 189-197.

207. Elnagar, A.Y., P.W. Sylvester, and K.A. El Sayed, *(−)-Oleocanthal as a c-Met Inhibitor for the Control of Metastatic Breast and Prostate Cancers.* Planta Med, 2011. **77**(10): p. 1013-1019.

208. Sastry, G.M., et al., *Protein and ligand preparation: parameters, protocols, and influence on virtual screening enrichments.* J Comput Aided Mol Des, 2013. **27**(3): p. 221-34.

209. Fukunishi, Y., et al., *Prediction of synthetic accessibility based on commercially available compound databases.* Journal Of Chemical Information And Modeling, 2014. **54**(12): p. 3259-3267.

210. Mulliken, R.S., *Electronic Population Analysis on LCAO–MO Molecular Wave Functions. I.* The Journal of Chemical Physics, 1955. **23**(10): p. 1833-1840.

211. Humphrey, W., A. Dalke, and K. Schulten, *VMD: visual molecular dynamics.* J Mol Graph, 1996. **14**(1): p. 33-8, 27-8.

212. Picot, D., P.J. Loll, and R.M. Garavito, *The X-ray crystal structure of the membrane protein prostaglandin H2 synthase-1.* Nature, 1994. **367**(6460): p. 243-9.

213. Laskowski, R.A., et al., *PROCHECK: a program to check the stereochemical quality of protein structures.* Journal of Applied Crystallography, 1993. **26**(2): p. 283-291.

214. De Colibus, L., et al., *More-powerful virus inhibitors from structure-based analysis of HEV71 capsid-binding molecules.* Nat Struct Mol Biol, 2014. **21**(3): p. 282-288.

215. de Bock, M., et al., *Human absorption and metabolism of oleuropein and hydroxytyrosol ingested as olive (Olea europaeaL.) leaf extract.* Molecular Nutrition & Food Research, 2013. **57**(11): p. 2079-2085.

216. Fitzpatrick, F., *Cyclooxygenase Enzymes: Regulation and Function.* Current Pharmaceutical Design, 2004. **10**(6): p. 577-588.

217. Bhattacharyya, D.K., et al., *Involvement of Arginine 120, Glutamate 524, and Tyrosine 355 in the Binding of Arachidonate and 2-Phenylpropionic Acid Inhibitors to the Cyclooxygenase Active Site of Ovine Prostaglandin Endoperoxide H Synthase-1.* Journal of Biological Chemistry, 1996. **271**(4): p. 2179-2184.

218. Hood, W.F., et al., *Characterization of celecoxib and valdecoxib binding to cyclooxygenase.* Mol Pharmacol, 2003. **63**(4): p. 870-7.

219. Mancini, J.A., et al., *Arginine 120 of Prostaglandin G/H Synthase-1 Is Required for the Inhibition by Nonsteroidal Anti-inflammatory Drugs Containing a Carboxylic Acid Moiety.* Journal of Biological Chemistry, 1995. **270**(49): p. 29372-29377.

220. Smith, W.L., D.L. DeWitt, and R.M. Garavito, *Cyclooxygenases: structural, cellular, and molecular biology.* Annu Rev Biochem, 2000. **69**: p. 145-82.

221. Segal, R.A., et al., *Differential Utilization of Trk Autophosphorylation Sites.* Journal of Biological Chemistry, 1996. **271**(33): p. 20175-20181.

222. Dubois, R.N., et al., *Cyclooxygenase in biology and disease.* FASEB J, 1998. **12**(12): p. 1063-73.

223. Loll, P.J., D. Picot, and R.M. Garavito, *The structural basis of aspirin activity inferred from the crystal structure of inactivated prostaglandin H2 synthase.* Nat Struct Biol, 1995. **2**(8): p. 637-43.

224. Rowlinson, S.W., et al., *A novel mechanism of cyclooxygenase-2 inhibition involving interactions with Ser-530 and Tyr-385.* J Biol Chem, 2003. **278**(46): p. 45763-9.

225. Hochgesang, G.P., S.W. Rowlinson, and L.J. Marnett, *Tyrosine-385 Is Critical for Acetylation of Cyclooxygenase-2 by Aspirin.* Journal of the American Chemical Society, 2000. **122**(27): p. 6514-6515.

226. Kalgutkar, A.S., et al., *Covalent modification of cyclooxygenase-2 (COX-2) by 2-acetoxyphenyl alkyl sulfides, a new class of selective COX-2 inactivators.* J Med Chem, 1998. **41**(24): p. 4800-18.

227. Calder, P.C., *Polyunsaturated fatty acids and inflammation.* Prostaglandins Leukot Essent Fatty Acids, 2006. **75**(3): p. 197-202.

228. Duran, M., et al., *Bitter phenolic glucosides from seeds of olive (Olea europaea).* Grasas y Aceites, 1994.

229. Bianco, A., et al., *Phenolic components of Olea Europaea – isolation of tyrosol derivatives.* Natural Product Research, 2004. **18**(1): p. 29-32.

230. Ding, L.-J., et al., *Hypocrol A, a new tyrosol derivative from a sponge-derived strain of the fungus Hypocrea koningii.* Natural Product Research, 2016. **30**(14): p. 1633-1638.

231. Cardoso, S.M., et al., *Oleuropein/ligstroside isomers and their derivatives in Portuguese olive mill wastewaters.* Food Chemistry, 2011. **129**(2): p. 291-296.

232. Callan, O.H., O.Y. So, and D.C. Swinney, *The kinetic factors that determine the affinity and selectivity for slow binding inhibition of human prostaglandin H synthase 1 and 2 by indomethacin and flurbiprofen.* J Biol Chem, 1996. **271**(7): p. 3548-54.

233. Parkinson, L. and R. Keast, *Oleocanthal, a Phenolic Derived from Virgin Olive Oil: A Review of the Beneficial Effects on Inflammatory Disease.* International Journal of Molecular Sciences, 2014. **15**(7): p. 12323-12334.

234. Lucas, L., A. Russell, and R. Keast, *Molecular mechanisms of inflammation. Anti-inflammatory benefits of virgin olive oil and the phenolic compound oleocanthal.* Curr Pharm Des, 2011. **17**(8): p. 754-68.

235. Egan, W.J. and G. Lauri, *Prediction of intestinal permeability.* Advanced Drug Delivery Reviews, 2002. **54**(3): p. 273-289.

236. Egan, W.J., K.M. Merz Jr, and J.J. Baldwin, *Prediction of drug absorption using multivariate statistics.* Journal of Medicinal Chemistry, 2000. **43**(21): p. 3867-3877.

237. Cheng, A. and K.M. Merz Jr, *Prediction of aqueous solubility of a diverse set of compounds using quantitative structure-property relationships.* Journal of Medicinal Chemistry, 2003. **46**(17): p. 3572-3580.

238. Cheng, A. and S.L. Dixon, *In silico models for the prediction of dose-dependent human hepatotoxicity.* Journal of Computer-Aided Molecular Design, 2003. **17**(12): p. 811-823.

239. Hess, B., et al., *GROMACS 4: Algorithms for Highly Efficient, Load-Balanced, and Scalable Molecular Simulation.* J Chem Theory Comput, 2008. **4**(3): p. 435-47.

240. Siu, S.W., et al., *Biomolecular simulations of membranes: physical properties from different force fields.* J Chem Phys, 2008. **128**(12): p. 125103.

241. Domanski, J., et al., *Lipidbook: a public repository for force-field parameters used in membrane simulations.* J Membr Biol, 2010. **236**(3): p. 255-8.

242. Jorgensen, W.L., et al., *Comparison of simple potential functions for simulating liquid water.* The Journal of chemical physics, 1983. **79**(2): p. 926-935.

243. Bjelkmar, P., et al., *Implementation of the CHARMM Force Field in GROMACS: Analysis of Protein Stability Effects from Correction Maps, Virtual Interaction Sites, and Water Models.* Journal of Chemical Theory and Computation, 2010. **6**(2): p. 459-466.

244. Feller, S.E. and A.D. MacKerell, *An Improved Empirical Potential Energy Function for Molecular Simulations of Phospholipids.* The Journal of Physical Chemistry B, 2000. **104**(31): p. 7510-7515.

245. Vanommeslaeghe, K., et al., *CHARMM General Force Field (CGenFF): A force field for drug-like molecules compatible with the CHARMM all-atom additive biological force fields.* Journal of computational chemistry, 2010. **31**(4): p. 671-690.

246. Zoete, V., et al., *SwissParam: a fast force field generation tool for small organic molecules.* J Comput Chem, 2011. **32**(11): p. 2359-68.

247. Nosé, S., *A unified formulation of the constant temperature molecular dynamics methods.* The Journal of Chemical Physics, 1984. **81**(1): p. 511-519.

248. Parrinello, M. and A. Rahman, *Crystal Structure and Pair Potentials: A Molecular-Dynamics Study.* Physical Review Letters, 1980. **45**(14): p. 1196-1199.

249. Parrinello, M. and A. Rahman, *Strain fluctuations and elastic constants.* The Journal of Chemical Physics, 1982. **76**(5): p. 2662-2666.

250. Hess, B., et al., *LINCS: a linear constraint solver for molecular simulations.* Journal of computational chemistry, 1997. **18**(12): p. 1463-1472.

251. Hub, J.S., B.L. de Groot, and D. van der Spoel, *g_wham—A Free Weighted Histogram Analysis Implementation Including Robust Error and Autocorrelation Estimates.* Journal of Chemical Theory and Computation, 2010. **6**(12): p. 3713-3720.

252. Refsgaard, H.H.F., et al., *In Silico Prediction of Membrane Permeability from Calculated Molecular Parameters.* Journal of Medicinal Chemistry, 2005. **48**(3): p. 805-811.

253. Bennion, B.J., et al., *Predicting a Drug's Membrane Permeability: A Computational Model Validated With in Vitro Permeability Assay Data.* J Phys Chem B, 2017. **121**(20): p. 5228-5237.

254. Cheng, F., et al., *In Silico ADMET Prediction: Recent Advances, Current Challenges and Future Trends.* Current Topics in Medicinal Chemistry, 2013. **13**(11): p. 1273-1289.

255. Hou, T., et al., *ADME Evaluation in Drug Discovery. 7. Prediction of Oral Absorption by Correlation and Classification.* Journal of Chemical Information and Modeling, 2007. **47**(1): p. 208-218.

256. Schuster, D., C. Laggner, and T. Langer, *Why Drugs Fail - A Study on Side Effects in New Chemical Entities.* Current Pharmaceutical Design, 2005. **11**(27): p. 3545-3559.

257. Williams, J.A., et al., *Drug-drug interactions for UDP-glucuronosyltransferase substrates: a pharmacokinetic explanation for typically observed low exposure (AUCi/AUC) ratios.* Drug Metab Dispos, 2004. **32**(11): p. 1201-8.

258. Tuck, K.L., et al., *The in vivo fate of hydroxytyrosol and tyrosol, antioxidant phenolic constituents of olive oil, after intravenous and oral dosing of labeled compounds to rats.* J Nutr, 2001. **131**(7): p. 1993-6.

259. Miro-Casas, E., et al., *Hydroxytyrosol disposition in humans.* Clin Chem, 2003. **49**(6 Pt 1): p. 945-52.

260. Abuznait, A.H., et al., *Olive-oil-derived oleocanthal enhances beta-amyloid clearance as a potential neuroprotective mechanism against Alzheimer's disease: in vitro and in vivo studies.* ACS Chem Neurosci, 2013. **4**(6): p. 973-82.

261. Daina, A., O. Michielin, and V. Zoete, *SwissADME: a free web tool to evaluate pharmacokinetics, drug-likeness and medicinal chemistry friendliness of small molecules.* Scientific Reports, 2017. **7**: p. 42717.

262. Schrödinger Release 2017-4, *QikProp.* 2017, Schrödinger, LLC: New York, NY.

263. Lecoeur, H., *Nuclear apoptosis detection by flow cytometry: influence of endogenous endonucleases.* Exp Cell Res, 2002. **277**(1): p. 1-14.

264. Lalande, M.E., V. Ling, and R.G. Miller, *Hoechst 33342 dye uptake as a probe of membrane permeability changes in mammalian cells.* Proceedings of the National Academy of Sciences of the United States of America, 1981. **78**(1): p. 363-367.

265. De Marco, E., et al., *Characterization and fractionation of phenolic compounds extracted from olive oil mill wastewaters.* Food Chemistry, 2007. **104**(2): p. 858-867.

266. Ryan, D., et al., *Identification of phenolic compounds in tissues of the novel olive cultivar hardy's mammoth.* J Agric Food Chem, 2002. **50**(23): p. 6716-24.

267. Cardoso, S.M., et al., *Characterisation of phenolic extracts from olive pulp and olive pomace by electrospray mass spectrometry.* Journal of the Science of Food and Agriculture, 2005. **85**(1): p. 21-32.

268. Latt, S.A., et al., *Recent developments in the detection of deoxyribonucleic acid synthesis by 33258 Hoechst fluorescence.* J Histochem Cytochem, 1975. **23**(7): p. 493-505.

269. Vekshin, N., *Binding of Hoechst with nucleic acids using fluorescence spectroscopy.* Journal of Biophysical Chemistry, 2011. **2**(04): p. 443.

270. Fox, C.B., R.A. Horton, and J.M. Harris, *Detection of Drug−Membrane Interactions in Individual Phospholipid Vesicles by Confocal Raman Microscopy.* Analytical Chemistry, 2006. **78**(14): p. 4918-4924.

271. Boggara, M.B. and R. Krishnamoorti, *Partitioning of Nonsteroidal Antiinflammatory Drugs in Lipid Membranes: A Molecular Dynamics Simulation Study.* Biophysical Journal, 2010. **98**(4): p. 586-595.

272. Paloncýová, M., K. Berka, and M. Otyepka, *Convergence of Free Energy Profile of Coumarin in Lipid Bilayer.* Journal of Chemical Theory and Computation, 2012. **8**(4): p. 1200-1211.

273. Mitchell, J.A., et al., *Selectivity of nonsteroidal antiinflammatory drugs as inhibitors of constitutive and inducible cyclooxygenase.* Proc Natl Acad Sci U S A, 1993. **90**(24): p. 11693-7.

274. Barbato, F., M.I. La Rotonda, and F. Quaglia, *Interactions of nonsteroidal antiinflammatory drugs with phospholipids: comparison between octanol/buffer partition coefficients and chromatographic indexes on immobilized artificial membranes.* J Pharm Sci, 1997. **86**(2): p. 225-9.

275. Luong, C., et al., *Flexibility of the NSAID binding site in the structure of human cyclooxygenase-2.* Nat Struct Biol, 1996. **3**(11): p. 927-33.

276. Balasubramanian, S.V., R.M. Straubinger, and M.E. Morris, *Salicylic acid induces changes in the physical properties of model and native kidney membranes.* J Pharm Sci, 1997. **86**(2): p. 199-204.

277. Sagawa, K., et al., *Ibuprofen-induced changes in sulfate renal transport.* J Pharmacol Exp Ther, 1998. **287**(3): p. 1092-7.

278. Tomisato, W., et al., *Membrane permeabilization by non-steroidal anti-inflammatory drugs.* Biochem Biophys Res Commun, 2004. **323**(3): p. 1032-9.

279. Lichtenberger, L.M., *The hydrophobic barrier properties of gastrointestinal mucus.* Annu Rev Physiol, 1995. **57**: p. 565-83.

280. Lichtenberger, L.M., *Where is the evidence that cyclooxygenase inhibition is the primary cause of nonsteroidal anti-inflammatory drug (NSAID)-induced gastrointestinal injury? Topical injury revisited.* Biochem Pharmacol, 2001. **61**(6): p. 631-7.

281. Zeng, J. and R.E. Fenna, *X-ray crystal structure of canine myeloperoxidase at 3 A resolution.* J Mol Biol, 1992. **226**(1): p. 185-207.

282. Kurumbail, R.G., et al., *Structural basis for selective inhibition of cyclooxygenase-2 by anti-inflammatory agents.* Nature, 1996. **384**(6610): p. 644-8.

283. Furse, K.E., et al., *Molecular Dynamics Simulations of Arachidonic Acid Complexes with COX-1 and COX-2: Insights into Equilibrium Behavior.* Biochemistry, 2006. **45**(10): p. 3189-3205.

284. Seo, M.-H., et al., *Protein conformational dynamics dictate the binding affinity for a ligand.* Nature Communications, 2014. **5**: p. 3724.

285. Nussinov, R. and C.-J. Tsai, *Allostery in Disease and in Drug Discovery.* Cell, 2013. **153**(2): p. 293-305.

286. Bussi, G., D. Donadio, and M. Parrinello, *Canonical sampling through velocity rescaling.* J Chem Phys, 2007. **126**(1): p. 014101.

287. Amadei, A., A.B. Linssen, and H.J. Berendsen, *Essential dynamics of proteins.* Proteins, 1993. **17**(4): p. 412-25.

288. Schulz, R. and U. Kleinekathofer, *Transitions between closed and open conformations of TolC: the effects of ions in simulations.* Biophys J, 2009. **96**(8): p. 3116-25.

289. Case, D., et al., *AMBER 12.* 2012, University of California.

290. Pearlman, D.A., et al., *AMBER, a package of computer programs for applying molecular mechanics, normal mode analysis, molecular dynamics and free energy calculations to simulate the structural and energetic properties of molecules.* Computer Physics Communications, 1995. **91**(1): p. 1-41.

291.   Abraham, M.J., et al., *GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers.* SoftwareX, 2015. **1-2**: p. 19-25.

292.   Wallace, A.C., R.A. Laskowski, and J.M. Thornton, *LIGPLOT: a program to generate schematic diagrams of protein-ligand interactions.* Protein Engineering, Design and Selection, 1995. **8**(2): p. 127-134.

293.   Hou, T., et al., *Assessing the performance of the MM/PBSA and MM/GBSA methods: I. The accuracy of binding free energy calculations based on molecular dynamics simulations.* Journal of chemical information and modeling, 2011. **51**(1): p. 69-82.

294.   Su, P.-C., et al., *Comparison of Radii Sets, Entropy, QM Methods, and Sampling on MM-PBSA, MM-GBSA, and QM/MM-GBSA Ligand Binding Energies of F. tularensis Enoyl-ACP Reductase (FabI).* Journal of computational chemistry, 2015. **36**(25): p. 1859-1873.

295.   Robert, K., A.B. Nathan, and J.A. McCammon, *iAPBS: a programming interface to the adaptive Poisson–Boltzmann solver.* Computational Science & Discovery, 2012. **5**(1): p. 015005.

296.   Hess, B., *Similarities between principal components of protein dynamics and random diffusion.* Phys Rev E Stat Phys Plasmas Fluids Relat Interdiscip Topics, 2000. **62**(6 Pt B): p. 8438-48.

297.   Hess, B., *Convergence of sampling in protein simulations.* Phys Rev E Stat Nonlin Soft Matter Phys, 2002. **65**(3 Pt 1): p. 031910.

298.   Campbell, I.D. and P. Bork, *Epidermal growth factor-like modules.* Current Opinion in Structural Biology, 1993. **3**(3): p. 385-392.

299.   Garavito, R.M., D. Picot, and P.J. Loll, *Prostaglandin H synthase.* Current Opinion in Structural Biology, 1994. **4**(4): p. 529-535.

300.   Picot, D., P.J. Loll, and R.M. Garavito, *The X-ray crystal structure of the membrane protein prostaglandin H2 synthase-1.* Nature, 1994. **367**: p. 243.

301.   Piazza, F. and Y.H. Sanejouand, *Long-range energy transfer in proteins.* Phys Biol, 2009. **6**(4): p. 046014.

302.   Juanico, B., et al., *Discrete breathers in nonlinear network models of proteins.* Phys Rev Lett, 2007. **99**(23): p. 238104.

303.   Csermely, P., R. Palotai, and R. Nussinov, *Induced fit, conformational selection and independent dynamic segments: an extended view of binding events.* Trends Biochem Sci, 2010. **35**(10): p. 539-46.

304.   Page, L., et al., *The PageRank Citation Ranking: Bringing Order to the Web.* 1999, Stanford InfoLab.

305.   Wang, P., H.-W. Bai, and B.T. Zhu, *Structural Basis for Certain Naturally Occurring Bioflavonoids to Function as Reducing Co-Substrates of Cyclooxygenase I and II.* PLOS ONE, 2010. **5**(8): p. e12316.

306.   Chauvot de Beauchêne, I., et al., *Hotspot Mutations in KIT Receptor Differentially Modulate Its Allosterically Coupled Conformational Dynamics: Impact on Activation and Drug Sensitivity.* PLOS Computational Biology, 2014. **10**(7): p. e1003749.

307.   Prusakiewicz, J.J., et al., *Molecular basis of the time-dependent inhibition of cyclooxygenases by indomethacin.* Biochemistry, 2004. **43**(49): p. 15439-45.

308.   Duggan, K.C., et al., *Molecular Basis for Cyclooxygenase Inhibition by the Non-steroidal Anti-inflammatory Drug Naproxen.* The Journal of Biological Chemistry, 2010. **285**(45): p. 34950-34959.

309.   Schneider, C., W.E. Boeglin, and A.R. Brash, *Identification of two cyclooxygenase active site residues, Leucine 384 and Glycine 526, that control carbon ring cyclization in prostaglandin biosynthesis.* J Biol Chem, 2004. **279**(6): p. 4404-14.

310. Hermanson, D.J., et al., *Oxicams Bind in a Novel Mode to the Cyclooxygenase Active Site via a Two-water-mediated H-bonding Network.* The Journal of Biological Chemistry, 2014. **289**(10): p. 6799-6808.

311. Vecchio, A.J., D.M. Simmons, and M.G. Malkowski, *Structural Basis of Fatty Acid Substrate Binding to Cyclooxygenase-2.* The Journal of Biological Chemistry, 2010. **285**(29): p. 22152-22163.

312. Bhattacharyya, D.K., et al., *Involvement of arginine 120, glutamate 524, and tyrosine 355 in the binding of arachidonate and 2-phenylpropionic acid inhibitors to the cyclooxygenase active site of ovine prostaglandin endoperoxide H synthase-1.* J Biol Chem, 1996. **271**(4): p. 2179-84.

313. Rieke, C.J., et al., *The role of arginine 120 of human prostaglandin endoperoxide H synthase-2 in the interaction with fatty acid substrates and inhibitors.* J Biol Chem, 1999. **274**(24): p. 17109-14.

314. Genheden, S. and U. Ryde, *The MM/PBSA and MM/GBSA methods to estimate ligand-binding affinities.* Expert Opinion on Drug Discovery, 2015. **10**(5): p. 449-461.

315. Khan, Y.S., H. Gutiérrez-de-Terán, and J. Åqvist, *Probing the Time Dependency of Cyclooxygenase-1 Inhibitors by Computer Simulations.* Biochemistry, 2017. **56**(13): p. 1911-1920.

316. Changhao, W., et al., *Calculating protein–ligand binding affinities with MMPBSA: Method and error analysis.* Journal of Computational Chemistry, 2016. **37**(27): p. 2436-2446.

317. Zarghi, A. and S. Arfaei, *Selective COX-2 Inhibitors: A Review of Their Structure-Activity Relationships.* Iran J Pharm Res, 2011. **10**(4): p. 655-83.

318. Nicholls, A., K.A. Sharp, and B. Honig, *Protein folding and association: insights from the interfacial and thermodynamic properties of hydrocarbons.* Proteins, 1991. **11**(4): p. 281-96.

9. Appendix

9.1 Olive phenolic library

Table A1.1: Phenolic compounds found in the olive

Compound Class	MW	Olive matrix					Commercial availability (Y/N)
		Olive fruit	Leaf	Pomace	Wastewater	Extra-virgin oil	
Simple phenols							
Hydroxytyrosol	154	De Marco, 2007	Savarese, 2007	Lozano-Sanchez, 2011		Dierkes, 2012	Y
Homovanillyl alcohol	168		Eyles, 2007			Boskou, 2006	Y
Tyrosol	138	Romero, 2002		Lozano-Sanchez, 2011	Artajo, 2006	Suarez, 2008	Y
Syringaldehyde	182				Boskou, 2006		Y
3,4-Dihydroxyphenylglycol	171	Marsilio, 2005					Y
Catechol	110	Romero, 2002			Lozano-Sanchez, 2011	Brenes, 2004	Y
Phenol	94					Vichi, 2008	Y
p-cresol	108				Artajo, 2006	Vichi, 2008	Y
m-cresol	108					Vichi, 2008	Y
o-cresol	108					Vichi, 2008	Y
4-Ethylguaiacol	152				Lozano-Sanchez, 2011	Vichi, 2008	Y
4-Ethylphenol	122					Vichi, 2008	Y
4-Vinylguaiacol	150				Limiroli, 1996	Vichi, 2008	Y
4-Vinylphenol	120			Obeid, 2007		Vichi, 2008	Y
4-Methylcatechol	124			Obeid, 2007			Y
3,4,5-Trimethoxybenzoic acid	212			Obeid, 2007			Y
3,4-Dimethoxybenzoic acid	182			Obeid, 2007			Y
2,6-Dimethoxybenzoic acid	182			Obeid, 2007			Y
4-Hydroxybenzaldehyde	122						Y
Hydroxybenzoic acids							
Syringic acid	198	Ryan, 1999		Alu'datt, 2010		Cioffi, 2010	Y

	Mass						
Quinic acid	192		Quirantes-Pine, 2013	Lozano-Sanchez, 2011	Lozano-Sanchez, 2011		Y
2,3-dihydrocaffeic acid	182	Owen, 2003				Bendini, 2007	Y
Shikimic acid	174			Peralbo-Molina, 2012			Y
Gallic acid	170	McDonald, 2001		Peralbo-Molina, 2012		Cioffi, 2010	Y
Vanillic acid	168	Romero, 2002		Peralbo-Molina, 2012	De Marco, 2007	De la Torre-Carbot, 2005	Y
Phloretic acid	166	Owen, 2003					Y
Protocatechuic acid	154	Boskou, 2006		Alu'datt, 2010			Y
Gentisic acid	154					Bendini, 2007	Y
4-hydroxybenzoic acid	138	Boskou, 2006	Quirantes-Piné, 2013	Alu'datt, 2010		Caponio, 2005	Y
2,4 dihydroxybenzoic acid	154	McDonald, 2001				Carrasco-Pancorbo, 2005	Y
2,6-Dihydroxybenzoic acid	154	Bianco, 2003				Bianco, 2003	Y
4-O-methyl-D-glucuronic acid	208	Guinda, 2010					Y
Hydroxyphenylacetic acids							
p-Hydroxyphenylacetic acid	152	Boskou, 2006		Cardoso, 2005		Caponio, 2005	Y
3,4-Dihydroxyphenylacetic acid	168	Boskou, 2006				Bendini, 2007	Y
4-Hydroxy-3-methoxy-phenylacetic acid	182					Bendini, 2007	Y
Homoveratric acid	196	Bianco, 2003				Bianco, 2003	Y
Homovanillic acid	182	Ryan, 2002			Artajo, 2006		Y
2,5-Dihydroxyphenylacetic acid	168			Obeid, 2007			Y
Hydroxycinnamic acids							
Rosmarinic acid	361			Peralbo-Molina, 2012			Y
Chlorogenic acid	354	Ryan, 2003		Cardoso, 2005			Y
Sinapic acid	224	Ryan, 1999		Alu'datt, 2010			Y

Compound	MW						
Ferulic acid	194	Boskou, 2006	Li, 2003	Peralbo-Molina, 2012		Cioffi, 2010	Y
Caffeic acid	180	De Marco, 2007		Obeid, 2007	Mulinacci, 2001	Mateos, 2001	Y
p-**Coumaric acid**	164	Ryan, 2002	Quirantes-Pine, 2013	Peralbo-Molina, 2012	Suarez, 2010	De la Torre-Carbot, 2005	Y
o-**Coumaric acid**	164	McDonald, 2001		Peralbo-Molina, 2012		Mateos, 2001	Y
m-**Coumaric acid**	164	Bianco, 2000				Caponio, 1999	Y
Dihydro-*p*-coumaric acid	166			Obeid, 2007			Y
Hydroxycaffeic acid	196					Caponio, 1999	Y
Cinnamic acid	148	Boskou, 2006		Peralbo-Molina, 2012		Mateos, 2001	Y
Caffeoylglucose	342			Obeid, 2007			Y
Caftaric acid	312			Obeid, 2007			Y
β-Hydroxy verbascoside	641						N
Flavonoids							
Hesperidin	610	Kalua, 2006		Alu'datt, 2010			Y
Rutin	610	De Marco, 2007	Savarese, 2007	Cardoso, 2005	Mulinacci, 2001		Y
Luteolin-7,4-O-diglucoside	609		Quirantes-Pine, 2013				N
Vicenin-2	595	Bouaziz, 2005					Y
Cyanidin-3-O-rutinoside	595	Romero, 2002					Y
Scolymoside	594	Bouaziz, 2010	Quirantes-Pine, 2013	Cardoso, 2005			Y
Luteolin-4'-O-rutinoside	594			Obeid, 2007			N
Lucidumoside C	584		Quirantes-Pine, 2013				Y
Isorhoifolin	578	Romero, 2002	Quirantes-Pine, 2013	Obeid, 2007			Y
Quercetin-3-O-glucoside	464	Bouaziz, 2005					Y
Chrysoeriol-7-O-glucoside	462	Bouaziz, 2005	Altıok, 2008				Y
Cyanidin-3-O-glucoside	449	Romero, 2002					Y
Luteolin-7-O-glucoside	448	De Marco, 2007	Laguerre, 2009	Cardoso, 2005	Mulinacci, 2001	Yorulamz, 2011	Y

Luteolin-4'-O-glucoside	448	De Marco, 2007	Quirantes-Pine, 2013	Cardoso, 2005			Y
Luteolin-6-C-glucoside	448	Bouaziz, 2005					Y
Luteolin-7-O-rutinoside	595	Bouaziz, 2005					Y
Luteolin-3′,7-O-diglucoside	611			Obeid, 2007			Y
Luteolin-8-C-glucoside	448			Obeid, 2007			Y
Quercitrin	448	De Marco, 2007		Obeid, 2007			Y
Apigenin-7-O-glucoside	432	Obeid, 2007	Obeid, 2007	Peralbo-Molina, 2012	Suarez, 2010		Y
Taxifolin	304		Japón-Luján, 2007	Peralbo-Molina, 2012		Bendini, 2007	Y
Quercetin	302	Rigane, 2011	Obeid, 2007	Alu'datt, 2010			Y
Chrysoeriol	300	Bouaziz, 2005				Dierkes, 2012	Y
Diosmetin	300		Meirinhos, 2005				Y
Methoxyluteolin	316					De la Torre-Carbot, 2005	Y
Cyanidin (cation)	287	Ryan, 1999					Y
Eriodictyol	287		Hvattum, 2002				Y
Luteolin	286	Ryan, 2002	Obeid, 2007	Lozano-Sanchez, 2011	Lozano-Sanchez, 2011	Fu, 2009	Y
Quercetin 3-O-rutinoside	611	Vlahov, 1992					Y
Apigenin	270	Bouaziz, 2005	Obeid, 2007	Lozano-Sanchez, 2011	Lozano-Sanchez, 2011	Fu, 2009	Y
Delphinidin	303			Obeid, 2007			Y
Delphinidin-3-O-glucoside	465			Obeid, 2007			Y
Hesperitin	302			Obeid, 2007			Y
Apigenin-7-O-rutinoside	579			Obeid, 2007			Y
Lignans							
Syringaresinol	418		Christophoridou, 2005			Garcia-Villalba, 2010	Y
1-Acetoxypinoresinol	416	Lopez, 2008		Lozano-Sanchez, 2011	Suarez, 2010	Fu, 2009	Y
Pinoresinol	358	Bonoli, 2004		Suarez, 2010	Suarez, 2010	Fu, 2009	Y

Name	MW	Ref 1	Ref 2	Ref 3	Ref 4	Y/N
Hydroxypinoresinol	374				Obeid, 2007	Y
Berchemol	376	Christophoridou, 2005				Y
3-Acetyloxy berchemol	418			Obeid, 2007		N
(-)-Olivil	376		Tsukamoto, 1984			Y
(+)-Fraxiresinol-1-β-D-glucopyranoside	537		Bianco, 1993			N
D-(+)-Erythro-1-(4-hydroxy-3-methoxy)-214 - phenyl-1,2,3-propantriol	214	Obeid, 2007		Obeid, 2007	Obeid, 2007	N
(+)-1-Acetoxypinoresinol-4''-O-methyl ether	431		Tsukamoto, 1984			N
(+)-1-Hydroxypinoresinol-4''-O-methyl ether	404		Tsukamoto, 1984			N
(+)-1-Acetoxypinoresinol-4'-β-D-glucopyranoside-4''-O-methyl ether	594		Tsukamoto, 1984			N
(+)-1-Hydroxypinoresinol-4'-β-D-glucopyranoside	537		Tsukamoto, 1984			N
(+)-1-Acetoxypinoresinol-4'-β-D-glucopyranoside	578		Tsukamoto, 1984			N
Hydroxy-isochromans						
1-(3'-Methoxy-4'-hydroxy)- phenyl-6,7-dihydroxyisochroman	288		Bianco, 2002		Bianco, 2001	N
1-Phenyl-6,7-dihydroxyisochroman	242		Bianco, 2002		Bianco, 2001	N
Secoiridoids						
Nüzhenide oleoside	1059	Cardoso, 2006		Cardoso, 2006		N
Oleuropein pentamer	2692	Cardoso, 2006		Cardoso, 2006		N
Oleuropein tetramer	2154	Cardoso, 2006		Cardoso, 2006		N
Oleuropein trimer	1616	Cardoso, 2006		Cardoso, 2006		N
Oleuropein dimer	1076	Cardoso, 2006		Cardoso, 2006		N
Nüzhenide 11-Methyl oleoside	1072	Silva, 2006				N
Oleuropein diglucoside	702	Bouaziz, 2010	Molina-Alcaidea, 1996	Cardoso, 2005		N

Neo-nüzhenide	702	Di Donna, 2007				N
Nüzhenide	686	Bouaziz, 2010	Ryan, 2002	Obeid, 2007		N
10-Hydroxyoleuropein	556	Cardoso, 2005	Caruso, 2000	Peralbo-Molina, 2012		Y
Dihydro-oleuropein	544	Obeid, 2007		Peralbo-Molina, 2012		N
Oleuropein	540	De Marco, 2007	Di Nino, 1997	Cardoso, 2005	Suarez, 2010	Bianco, 1998	Y
Oleuroside	540	De Marco, 2007	Di Nino, 1997	Obeid, 2007		N
Oleuroside-10-carboxylic acid	585			Obeid, 2007	Obeid, 2007	N
Oleuropein-3'-O-β-D-glucopyranoside	540			Obeid, 2007		N
Ligstroside-3'-O-β-D-glucopyranoside	524			Obeid, 2007		N
Demethyloleuropein	526	De Marco, 2007	Savarase, 2007		Bianco, 1998	N
Oleoside dimethylester	418			Peralbo-Molina, 2012		N
Elenolic acid glucoside	404	De Marco, 2007	Ryan, 2002	Cardoso, 2005		N
Elenolic acid diglucoside	553			Obeid, 2007		N
Secologanic acid	375			Peralbo-Molina, 2012		N
Secologanol	390				Fu, 2009	Y
Secologanin	388			Peralbo-Molina, 2012		Y
Oleuropein aglycone (3,4-DHPEA-EA)	378	De Marco, 2007		Cardoso, 2005	Suarez, 2010	Fu, 2009	N
7-Deoxyloganic acid	360			Peralbo-Molina, 2012		N
Oleacein (Dialdehydic form of decarboxymethyl Oleuropein aglycon)	321				Karkoula, 2012	Y
3,4-DHPEA-EDA (Oleuropein-aglycone di-aldehyde)	320	De Marco, 2007		Peralbo-Molina, 2012	Suarez, 2010	N
Cornoside	316	Bianchi, 2003				Y

Compound						
Oleocanthal (Dialdehydic form of decarboxymethyl Ligstroside aglycon)	305				Smith, 2005	Y
p-HPEA-EDA	304	Obeid, 2007		Peralbo-Molina, 2012	Suarez, 2010	Perez-Trujillo, 2010	N
Elenolic acid	242	De Marco, 2007		Peralbo-Molina, 2012	Mulinacci, 2001	Dierkes, 2012	Y
Elenolic acid methylester	256		Gariboldi, 1986			Dierkes, 2012	N
Dialdehydic elenolic acid decarboxymethyl	184			Lozano-Sanchez, 2011	Lozano-Sanchez, 2011		N
Dialdehydic elenolic ester decarboxymethyl	198			Lozano-Sanchez, 2011			N
Methyl malate-hydroxytyrosol ester	284	Bianco, 2006					N
Oleuricine A	716		Wang, 2009				N
Oleuricine B	716		Wang, 2009				N
Hydroxytyrosil-elenolate	392		Gariboldi, 1986				N
10-Hydroxy oleuropein aglycone	394					Garcia-Villalba, 2010	N
10-Hydroxy oleuropein aglycone decarboxymethyl	336			Obeid, 2007			N
10-Hydroxy-10-methyl oleuropein aglycone	408			Obeid, 2007			N
Monoaldehydic form of Ligstroside aglycon	365	Kanakis, 2013				Kanakis, 2013	N
Monoaldehydic form of Oleuropein aglycon	380	Kanakis, 2013				Kanakis, 2013	N
Oleuropeindial (enol form)	378	Obeid, 2007					N
demethyloleuropein aglycone (enol form)	364	Obeid, 2007					N
Demethyloleuropein aglycone	364	Obeid, 2007					N
Demethyloleuropein	526	Obeid, 2007					N
Demethyloleuropein aglycone dialdehyde	364	Obeid, 2007					N
3,4-DHPEA-DEDA (Oleuropein aglycone decarboxymethyl dialdehyde form)	320	Obeid, 2007			Scalzo, 1993		N

3,4-DHPEA-DEDA (acetal)	366	Obeid, 2007			N
Oleuropeindial (keto form)	378	Obeid, 2007			N
Oleuropeindial (Cannizzaro-like product of oleuropeindial)	396	Obeid, 2007			N
Oleuropeindial - Lactone (Cannizzaro-like product of oleuropeindial)	378	Obeid, 2007			N
Elenolic acid dialdehyde	242	Obeid, 2007			N
DEDA (Decarboxymethyl elenolic acid dialdehyde)	184	Obeid, 2007			N
Hydroxytyrosol acetate	196	Morello, 2004	Quirantes-Pine, 2013	Brenes, 1999	Y
DEDA acetal	230	Obeid, 2007			N
Demethyl elenolic acid	228	Obeid, 2007			N
Tyrosol acetate	222			Mateos, 2001	Y
Ligstroside	524	De Marco, 2007	Savarase, 2007	De Marco, 2007	Y
Ligstroside aglycone methyl acetal	376			Obeid, 2007	N
Ligstroside aglycone	380		Savarase, 2007		N
Demethylligstroside	510	Sivakumar, 2005			N
Ligstroside derivative 1	523			Cardoso, 2011	N
Ligstroside derivative 2	453			Cardoso, 2011	N
Ligstroside derivative 3	685			Cardoso, 2011	N
Ligstroside derivative 4	847			Cardoso, 2011	N
Ligstroside derivative 5	909			Cardoso, 2011	N
Jaspolyoside	926		Perez-Bonilla, 2011		N
Jaspolyanoside	911		Perez-Bonilla, 2011		N
Isojaspolyoside A	926		Perez-Bonilla, 2011		N

(+)-Cycloolivil	376		Tsukamoto, 1984			Y
Hemiacetal of dialdehydic oleuropein aglycone decarboxymethyl	334	Christophoridou, 2005				N
Hemiacetal of dialdehydic ligstroside aglycone decarboxymethyl	318	Christophoridou, 2005				N
7''-S-Hydroxyoleuropein	557	Di Donna, 2007	Quirantes-Pine, 2013			N
Oleuropein-3''-Methyl ether	555			Obeid, 2007		N
3,4-DHPEA-DETA	350			Obeid, 2007		N
Decarboxymethyl ligstroside aglycone	304				Lozano-Sanchez, 2010	N
Hydroxytyrosol acyclodihydroelenolate	382	Obeid, 2007		Obeid, 2007		N
Hydroxytyrosil elenolate	364			Obeid, 2007		N
Coumarins						
Esculin	340		Tsukamoto, 1984			Y
Esculetin	178		Tsukamoto, 1984			Y
Scopoletin	192		Tsukamoto, 1984			Y
Scopolin	354		Tsukamoto, 1984			Y
Irridoids						
Loganic acid	375			Peralbo-Molina, 2012		Y
Loganin	390			Peralbo-Molina, 2012		Y
Glucosides						
Hydroxytyrosol rhamnoside	301			Peralbo-Molina, 2012		N
β-Hydroxy-acetoside	640	Cecchi, 2013				N
Verbascoside	624	Rigane, 2011	Laguerre, 2009	Cardoso, 2005	Mulinacci, 2001	Y
Oxidized verbascoside	623			Cardoso, 2005		N
Isoverbascoside	624	Obeid, 2007	Laguerre, 2009	Innocenti, 2006		Y

Oxidized isoverbascoside	623			Cardoso, 2005		N
Acetoside	623		Ryan, 1999	Rodriguez, 2009		Y
Isoacteoside	625			Obeid, 2007		Y
Suspensaside	641			Rodriguez, 2009		Y
Hellicoside	657			Rodriguez, 2009		Y
Orbanchoside	623			Rodriguez, 2009		N
Wedelosin	758			Rodriguez, 2009		N
Oleoside-11-Methylester	403			Peralbo-Molina, 2012		Y
6'-O-[(2E)-2,6-Dimethyl-8-hydroxy- 2-octenoyloxy]-secologanoside	559		Karioti, 2006			N
4'-O-β-D-Glucosyl-9-O-(6''-deoxysaccharosyl)olivil	851		Schumacher, 2002			N
Verucosin	345				Christophoridou, 2005	N
Quercetin-3-rhamnoside	449	Savarese, 2007				N
Quercetin-7-O-glucoside	464	Bouaziz, 2005				Y
Caffeoyl-6'-secologanoside	552	Obeid, 2007		Obeid, 2007	Obeid, 2007	N
Comselogoside	536	Obeid, 2007		Obeid, 2007		N
Oleoside	390	Bouaziz, 2010	Obeid, 2007	Cardoso, 2005	Fu, 2009	Y
Secologanoside	390		Obeid, 2007		Fu, 2009	N
Hydroxytyrosol-1′-β-glucoside	316	Cardoso, 2005	Savarese, 2007	Cardoso, 2005	Bianco, 1998	N
Hydroxytyrosol-3-β-glucoside	316	Rubio-Senet, 2013	Savarese, 2007	Rubio-Senet, 2013	Bianco, 1998	N
Hydroxytyrosol-4-β-glucoside	316	Rubio-Senet, 2013	Savarese, 2007	Rubio-Senet, 2013	Bianco, 1998	N
Hydroxytyrosol diglucoside	478			Peralbo-Molina, 2012		N
Salidroside	300	Romero, 2002		Peralbo-Molina, 2012		Y

6'-β-D-Glucopyranosyl oleoside	553	Savarese, 2007	Obeid, 2007		N
6'-Rhamnopyranosyl oleoside	537	Savarese, 2007	Obeid, 2007		N
Methoxyphenols					
Guaiacol	124			Reiners, 1998	Y
2-Methoxy-4-vinylphenol	150		Poerschmann, 2013		Y
Isoeugenol	165		Poerschmann, 2013		Y
Homovanillin	166		Poerschmann, 2013		Y
Phenolic fatty acid esters					
1-oleyltyrosol	389	Maestroduran, 1994			N
Deoxyloganic acid lauryl ester	408		Rigane, 2011		N

9.2 Molecular docking

9.2.1 Inhibitory data for common COX inhibitors

Table A2.1: Maximal inhibitory data for common COX-1 inhibitors

PubChem CID	Common name	pIC$_{50}$	PubChem AID
2662	Celecoxib	4.83	161333; 161494; 392041
3033	Diclofenac	7.92	312490; 348028; 392041
3177	Dup-697	6.09	161655; 161679; 161680
3672	Ibuprofen	5.52	161666; 161679
5090	Rofecoxib	4.49	161496
156391	Naproxen	5.49	161655; 161679; 161680
443373	SC-57666	4.52	332217; 370919
445154	Resveratrol	6.01	332217; 370919

Table A2.2: Maximal inhibitory data for common COX-2 inhibitors

PubChem CID	Common name	pIC$_{50}$	PubChem AID
2244	Aspirin	5.612	649270; 494635
2662	Celecoxib	7.115	649270; 1125534
3672	Ibuprofen	5.941	494635; 443725
5090	Rofecoxib	6.381	1125534; 724444
156391	Naproxen	4.942	724444; 587449
445154	Resveratrol	6.002	289279

9.2.2   Glide energy values for all olive phenolics to COX-1 and COX-2

Table A2.3a: Glide energy values for simple phenols

Ligand	Glide Energy (kcal/mol)	
	COX-1	COX-2
Hydroxytyrosol	-25.793	-22.835
Hydroxytyrosol acetate	-33.633	-30.276
Tyrosol acetate	-38.545	-26.245
Homovanillyl alcohol	-26.521	-22.246
Tyrosol	-21.946	-21.249
Syringaldehyde	-31.248	-23.576
3,4-Dihydroxyphenylglycol	-26.317	-21.336
Catechol	-21.198	-20.306
Phenol	-18.624	-19.287
p-cresol	-20.984	-21.39
m-cresol	-20.499	-20.771
o-cresol	-21.599	-19.031
4-Ethylguaiacol	-26.092	-22.513
4-Ethylphenol	-23.361	-20.197
4-Vinylguaiacol	-25.448	-22.776
4-Vinylphenol	-19.647	-22.616
4-Methylcatechol	-23.463	-23.131
3,4,5-Trimethoxybenzoic acid	-30.07	-21.691
3,4-Dimethoxybenzoic acid	-29.354	-20.856
2,6-Dimethoxybenzoic acid	-28.799	-18.148
4-Hydroxybenzaldehyde	-22.345	-23.03

Table A2.3b: Glide energy values for hydroxybenzoic acids

Ligand	Glide Energy (kcal/mol)	
	COX-1	COX-2
Syringic acid	-30.082	-20.603
Quinic acid	-28.679	-14.147
2,3-dihydrocaffeic acid	-32.855	-23.753
Shikimic acid	-26.019	-23.685
Gallic acid	-26.925	-18.698
Vanillic acid	-27.074	-19.318
Phloretic acid	-29.041	-21.047
Protocatechuic acid	-24.64	-19.297
Gentisic acid	-24.705	-17.7
4-hydroxybenzoic acid	-24.049	-20.941
2,4 dihydroxybenzoic acid	-25.367	-23.958
2,6-Dihydroxybenzoic acid	-24.972	-15.181
4-O-methyl-D-glucuronic acid	-16.17	N/A
3,4-dihydroxyphenylethyl 4-formyl -3-	-41.404	N/A

Ligand		
formylmethyl-4-hexenoate		
4-hydroxyphenylethyl 4-formyl -3-formylmethyl-4-hexenoate	-38.095	N/A

Table A2.3c: Glide energy values for hydroxyphenylacetic acids

Ligand	Glide Energy (kcal/mol)	
	COX-1	COX-2
p-Hydroxyphenylacetic acid	-26.789	-17.634
3,4-Dihydroxyphenylacetic acid	-28.266	-19.638
4-Hydroxy-3-methoxy-phenylacetic acid	-31.033	-19.389
Homoveratric acid	-32.921	-26.703
Homovanillic acid	-29.916	-22.037
2,5-Dihydroxyphenylacetic acid	-29.003	-15.314

Table A2.3d: Glide energy values for hydroxycinnamic acids

Ligand	Glide Energy (kcal/mol)	
	COX-1	COX-2
Rosmarinic acid	-47.359	-33.965
Chlorogenic acid	-33.93	-38.948
Sinapic acid	-35.45	-23.094
Ferulic acid	-32.578	-22.715
Caffeic acid	-30.285	-20.614
p-Coumaric acid	-27.78	-18.149
o-Coumaric acid	-29.063	-21.458
m-Coumaric acid	-28.412	-22.435
Dihydro-p-coumaric acid	-29.513	-21.047
Hydroxycaffeic acid	-29.938	-24.327
Cinnamic acid	-26.671	-18.82
Caffeoylglucose	-33.1	-34.194
Caftaric acid	-44.118	-26.471

Table A2.3e: Glide energy values for flavonoids

Ligand	Glide Energy (kcal/mol)	
	COX-1	COX-2
Quercetin-3-O-glucoside	-31.025	N/A
Chrysoeriol-7-O-glucoside	-36.912	-29.258
Cyanidin-3-O-glucoside	-17.559	-5.546
Luteolin-7-O-glucoside	-34.11	-38.19
Luteolin-4'-O-glucoside	-32.51	-42.307
Luteolin-6-C-glucoside	N/A	-8.989
Luteolin-8-C-glucoside	N/A	N/A

Ligand		
Quercetin-3-rhamnoside	-15.864	N/A
Quercetin-7-O-glucoside	-31.838	N/A
Quercitrin	N/A	N/A
Apigenin-7-O-glucoside	-42.693	N/A
Taxifolin	-36.837	-25.505
Quercetin	-37.581	-32.21
Chrysoeriol	-26.775	-15.527
Diosmetin	-37.694	-14.046
Methoxyluteolin	-37.829	-15.94
Cyanidin (cation)	-28.554	-31.798
Eriodictyol	-34.389	-26.867
Luteolin	-35.87	-31.249
Apigenin	-29.846	-13.708
Delphinidin	-31.314	-31.481
Delphinidin-3-O-glucoside	-17.408	-3.801
Hesperitin	-35.649	-25.266

Table A2.3f: Glide energy values for lignans

Ligand	Glide Energy (kcal/mol)	
	COX-1	COX-2
Syringaresinol	-28.202	-17.701
1-Acetoxypinoresinol	-8.022	N/A
Pinoresinol	-26.272	-3.106
Hydroxypinoresinol	-33.612	N/A
Berchemol	-31.434	-40.041
3-Acetyloxy berchemol	-25.851	-27.856
(-)-Olivil	-30.699	N/A
D-(+)-Erythro-1-(4-hydroxy-3-methoxy)-214 - phenyl-1,2,3-propantriol	-32.532	-27.779
(+)-1-Acetoxypinoresinol-4"-O-methyl ether	-30.251	-39.494
(+)-1-Hydroxypinoresinol-4"-O-methyl ether	-39.448	N/A

Table A2.3g: Glide energy values for hydroxy-isochromans

Ligand	Glide Energy (kcal/mol)	
	COX-1	COX-2
1-(3'-Methoxy-4'-hydroxy)- phenyl-6,7-dihydroxyisochroman	-27.18	-20.492
1-Phenyl-6,7-dihydroxyisochroman	-20.536	-11.946

Table A2.3h: Glide energy values for secoiridoids

Ligand	Glide Energy (kcal/mol)	
	COX-1	COX-2
Elenolic acid glucoside	-19.212	-19.172
Secologanic acid	-14.816	N/A
Secologanol	-27.035	-29.139
Secologanin	-25.574	N/A
Secologanoside	-28.068	-5.008
Methyl malate-hydroxytyrosol ester	-43.03	-32.925
Oleuropein aglycone (3,4-DHPEA-EA)	-38.088	-33.702
7-Deoxyloganic acid	-25.114	-11.685
Oleacein (Dialdehydic form of decarboxymethyl Oleuropein aglycon)	-44.865	-34.428
3,4-DHPEA-EDA (Oleuropein-aglycone di-aldehyde)	-42.234	-36.475
Cornoside	-34.365	-11.717
Oleocanthal (Dialdehydic form of decarboxymethyl Ligstroside aglycon)	-42.425	-30.408
p-HPEA-EDA	-42.425	-30.408
Elenolic acid	-18.761	-11.183
Dialdehydic elenolic acid decarboxymethyl	-22.577	-24.075
Hydroxytyrosil-elenolate	-35.757	-36.375
10-Hydroxy oleuropein aglycone	-45.202	-38.795
10-Hydroxy oleuropein aglycone decarboxymethyl	-40.335	-32.177
10-Hydroxy-10-methyl oleuropein aglycone	-38.715	-40.673
Monoaldehydic form of Ligstroside aglycon	-38.796	-26.739
Monoaldehydic form of Oleuropein aglycon	-27.634	-33.922
Oleuropeindial (enol form)	-41.052	-27.813
demethyloleuropein aglycone (enol form)	-43.261	-34.088
Demethyloleuropein aglycone	-40.345	-36.147
Demethyloleuropein aglycone dialdehyde	-35.221	-34.088
3,4-DHPEA-DEDA (Oleuropein aglycone decarboxymethyl dialdehyde form)	-39.665	-34.663
3,4-DHPEA-DEDA (acetal)	-40.584	-33.613
Oleuropeindial (keto form)	-45.252	-39.385
Oleuropeindial (Cannizzaro-like product of oleuropeindial)	-35.834	-35.639
Oleuropeindial - Lactone (Cannizzaro-like product of oleuropeindial)	-37.689	-40.125
Elenolic acid dialdehyde	-27.784	-18.019
DEDA (Decarboxymethyl elenolic acid	-24.116	-22.513

dialdehyde)		
DEDA acetal	-31.296	-22.547
Oleoside-11-Methylester	-22.977	-32.621
Demethyl elenolic acid	-25.833	-17.567
(+)-Cycloolivil	-21.807	N/A
Oleoside dimethylester	-25.639	-17.307
Oleanolic acid demethyl	N/A	N/A
Hemiacetal of dialdehydic oleuropein aglycone decarboxymethyl	-36.984	-36.487
Hemiacetal of dialdehydic ligstroside aglycone decarboxymethyl	-34.032	N/A
Dialdehydic elenolic ester decarboxymethyl	-22.98	-26.047
Elenolic acid methylester	-15.632	N/A
Ligstroside aglycone	-37.703	-34.122
Ligstroside derivative 2	-48.222	-49.682
Ligstroside aglycone methyl acetal	-40.105	-35.28
3,4-DHPEA-DETA	-40.784	-31.546
Decarboxymethyl ligstroside aglycone	-42.425	-33.763
Hydroxytyrosol acyclodihydroelenolate	-42.071	-23.099
Hydroxytyrosol elenolate	-35.757	-36.375

Table A2.3i: Glide energy values for coumarins

Ligand	Glide Energy (kcal/mol)	
	COX-1	COX-2
Esculin	-38.905	-30.335
Esculetin	-24.749	-25.235
Scopoletin	-28.289	-25.334
Scopolin	-19.65	-34.739

Table A2.3j: Glide energy values for irridoids

Ligand	Glide Energy (kcal/mol)	
	COX-1	COX-2
Loganic acid	-14.176	N/A
Loganin	-29.453	-26.266

Table A2.3k: Glide energy values for glucosides

Ligand	Glide Energy (kcal/mol)	
	COX-1	COX-2
Verucosin	-24.374	N/A
Oleoside	-35.417	N/A

Ligand		
Hydroxytyrosol rhamnoside	-31.813	-28.606
Hydroxytyrosol-1′-β-glucoside	-31.554	N/A
Hydroxytyrosol-3-β-glucoside	-25.303	-39.91
Hydroxytyrosol-4-β-glucoside	-26.997	N/A
Hydroxytyrosol diglucoside	-37.63	-43.855
Salidroside	-31.292	-31.509

Table A2.3l: Glide energy values for methoxyphenols

Ligand	Glide Energy (kcal/mol)	
	COX-1	COX-2
Guaiacol	-22.192	N/A
2-Methoxy-4-vinylphenol	-24.445	N/A
Isoeugenol	-25.7	N/A
Homovanillin	-28.446	-23.912

Table A2.3m: Glide energy values for phenolic fatty acid esters

Ligand	Glide Energy (kcal/mol)	
	COX-1	COX-2
1-oleyltyrosol	-50.865	-43.895
Deoxyloganic acid lauryl ester	-43.078	-34.807

9.3 Membrane permeability simulations

9.3.1   mdp files for SMD and umbrella sampling

9.3.1.1 SMD energy minimisation

```
; minim.mdp - used as input into grompp to generate em.tpr
; Parameters describing what to do, when to stop and what to save
integrator  = steep     ; Algorithm (steep = steepest descent
minimization)
emtol       = 1000.0    ; Stop minimization when the maximum force <
1000.0 kJ/mol/nm
emstep      = 0.01      ; Energy step size
nsteps      = 50000     ; Maximum number of (minimization) steps to
perform

; Parameters describing how to find the neighbors of each atom and how to
calculate the interactions
nstlist     = 1         ; Frequency to update the neighbor list and long
range forces
ns_type     = grid      ; Method to determine neighbor list (simple, grid)
rlist       = 1.4       ; Cut-off for making neighbor list (short range
forces)
coulombtype = PME       ; Treatment of long range electrostatic
interactions
rcoulomb    = 1.4       ; Short-range electrostatic cut-off
rvdw        = 1.4       ; Short-range Van der Waals cut-off
pbc         = xyz       ; Periodic Boundary Conditions
```

9.3.1.2 SMD equilibration

```
title         = NPT Equilibration
define        = -DPOSRES            ; position restrain the protein
; Run parameters
integrator  = md                    ; leap-frog integrator
nsteps      = 50000                  ; 2 * 50000 = 100 ps
dt          = 0.002                  ; 2 fs
; Output control
nstxout     = 1000                   ; save coordinates every 2 ps
nstvout     = 1000                   ; save velocities every 2 ps
nstenergy   = 1000                   ; save energies every 2 ps
nstlog      = 1000                   ; update log file every 2 ps
; Bond parameters
continuation        = yes          ; Initial simulation
constraint_algorithm = lincs       ; holonomic constraints
constraints         = all-bonds ; all bonds (even heavy atom-H bonds)
constrained
lincs_iter          = 1            ; accuracy of LINCS
lincs_order         = 4            ; also related to accuracy
; Neighborsearching
cutoff-scheme   = Verlet
ns_type     = grid                  ; search neighboring grid cels
nstlist     = 20                    ; 40 fs
rlist       = 1.4                   ; short-range neighborlist cutoff (in nm)
rcoulomb    = 1.4                   ; short-range electrostatic cutoff (in nm)
rvdw        = 1.4                   ; short-range van der Waals cutoff (in nm)
; Electrostatics
coulombtype     = PME               ; Particle Mesh Ewald for long-range
electrostatics
pme_order   = 4                     ; cubic interpolation
fourierspacing = 0.16               ; grid spacing for FFT
; Temperature coupling is on
tcoupl      = V-rescale             ; modified Berendsen thermostat
tc-grps     = DOPC_LIG   Water      ; two coupling groups - more accurate
tau_t       = 0.1       0.1         ; time constant, in ps
ref_t       = 310       310         ; reference temperature, one for each
group, in K
; Pressure coupling is on
pcoupl          = Berendsen     ; Pressure coupling on in NPT, also
weak coupling
pcoupltype          = semiisotropic ; not uniform scaling of x-y-z box
vectors
tau_p       = 2.0 2.0           ; time constant, in ps
ref_p       = 1.0 1.0           ; reference pressure (in bar)
compressibility     = 4.5e-5 4.5e-5   ; isothermal compressibility, bar^-1
refcoord_scaling    = com
; Periodic boundary conditions
pbc     = xyz                       ; 3-D PBC
; Dispersion correction
DispCorr    = EnerPres              ; account for cut-off vdw scheme
; Velocity generation
gen_vel     = no                    ; Velocity generation is off afer NVT
; COM motion removal
; These options remove COM motion of the system
nstcomm         = 10
comm-mode       = Linear
comm-grps       = System
```

9.3.1.3 SMD pulling simulation

```
title         = Umbrella pulling simulation
define        = -DPOSRES_B
; Run parameters
integrator  = md
dt          = 0.002
tinit       = 0
nsteps      = 400000     ; 2 * 400000 = 800 ps
nstcomm     = 10
```

```
; Output parameters
nstxout     = 5000      ; every 10 ps
nstvout     = 5000
nstfout     = 500
nstxtcout   = 500       ; every 1 ps
nstenergy   = 500
; Bond parameters
constraint_algorithm    = lincs
constraints             = all-bonds
continuation            = no        ; not continuing from NPT
; Single-range cutoff scheme
cutoff-scheme   = Verlet
nstlist     = 5         ; 10 fs
ns_type     = grid      ; search neighbouring grid cells
rlist       = 1.4       ; short-range neighborlist cutoff (in nm)
rcoulomb    = 1.4       ; short-range electrostatic cutoff (in nm)
rvdw        = 1.4       ; short-range van der Waals cutoff (in nm)
; PME electrostatics parameters
coulombtype     = PME   ; Particle Mesh Ewald for long-range electrostatics
fourierspacing  = 0.12
fourier_nx      = 0
fourier_ny      = 0
fourier_nz      = 0
pme_order       = 4     ; cubic interpolation
ewald_rtol      = 1e-5
optimize_fft    = yes
; Berendsen temperature coupling is on in two groups
Tcoupl      = Nose-Hoover
tc_grps     = DOPC_LIG   Water
tau_t       = 0.5        0.5
ref_t       = 310        310
; Pressure coupling is on
Pcoupl          = Parrinello-Rahman
pcoupltype      = semiisotropic ; not uniform scaling of x-y-z box vectors
tau_p       = 1.0    1.0
compressibility = 4.5e-5    4.5e-5
ref_p       = 1.0 1.0
refcoord_scaling = com
; Velocity generation
gen_vel     = yes       ; assign velocities from Maxwell distribution
gen_seed    = -1        ; generate random velocities with each run
gen_temp    = 310       ; so that first step will be randomised
; Periodic boundary conditions are on in all directions
pbc     = xyz
; Long-range dispersion correction
DispCorr    = EnerPres
; Pull code
pull_nstxout    = 500           ; every 1 ps
pull_nstfout    = 500           ; every 1 ps
pull            = umbrella
pull_geometry   = direction
pull_dim        = N N Y
pull_vec1       = 0 0 -1        ; pull down z-axis
pull_start      = yes       ; define initial COM distance > 0
pull_ngroups    = 1
pull_group0     = DOPC
pull_group1     = LIG
pull_rate1      = 0.01      ; 0.01 nm per ps = 10 nm per ns
pull_k1         = 1000      ; kJ mol^-1 nm^-2
```

9.3.1.4 Umbrella sampling equilibration

```
title       = Umbrella pulling simulation
define      = -DPOSRES_B
; Run parameters
integrator  = md
dt          = 0.002
tinit       = 0
nsteps      = 50000     ; 100 ps
```

```
nstcomm       = 10
; Output parameters
nstxout       = 5000        ; every 10 ps
nstvout       = 5000
nstfout       = 5000
nstxtcout     = 5000
nstenergy     = 5000
; Bond parameters
constraint_algorithm    = lincs
constraints             = all-bonds
continuation            = no
; Single-range cutoff scheme
cutoff-scheme   = Verlet
nstlist       = 5
ns_type       = grid
rlist         = 1.4
rcoulomb      = 1.4
rvdw          = 1.4
; PME electrostatics parameters
coulombtype       = PME
fourierspacing    = 0.12
fourier_nx        = 0
fourier_ny        = 0
fourier_nz        = 0
pme_order         = 4
ewald_rtol        = 1e-5
optimize_fft      = yes
; Berendsen temperature coupling is on in two groups
Tcoupl        = Nose-Hoover
tc_grps       = DOPC_LIG    Water
tau_t         = 0.5         0.5
ref_t         = 310         310
; Pressure coupling is on
Pcoupl            = Parrinello-Rahman
pcoupltype        = semiisotropic          ; not uniform scaling
tau_p             = 1.0    1.0
compressibility = 4.5e-5    4.5e-5
ref_p             = 1.0 1.0
refcoord_scaling = com
; Generate velocities is on
gen_vel       = yes
; Periodic boundary conditions are on in all directions
pbc       = xyz
; Long-range dispersion correction
DispCorr      = EnerPres
; Pull code
pull              = umbrella
pull_geometry     = distance
pull_dim          = N N Y
pull_start        = yes
pull_ngroups      = 1
pull_group0       = DOPC
pull_group1       = LIG
pull_init1        = 0
pull_rate1        = 0.0
pull_k1           = 1000        ; kJ mol^-1 nm^-2
pull_nstxout      = 1000        ; every 2 ps
pull_nstfout      = 1000        ; every 2 ps
```

9.3.1.5   Umbrella sampling simulation

```
title         = Umbrella pulling simulation
define        = -DPOSRES_B
; Run parameters
integrator  = md
dt          = 0.002
tinit       = 0
nsteps      = 5000000    ; 10 ns
nstcomm     = 10
```

```
; Output parameters
nstxout      = 50000      ; every 100 ps
nstvout      = 50000
nstfout      = 5000
nstxtcout    = 5000       ; every 10 ps
nstenergy    = 5000
; Bond parameters
constraint_algorithm    = lincs
constraints             = all-bonds
continuation            = yes
; Single-range cutoff scheme
cutoff-scheme   = Verlet
nstlist      = 5
ns_type      = grid
rlist        = 1.4
rcoulomb     = 1.4
rvdw         = 1.4
; PME electrostatics parameters
coulombtype     = PME
fourierspacing  = 0.12
fourier_nx      = 0
fourier_ny      = 0
fourier_nz      = 0
pme_order       = 4
ewald_rtol      = 1e-5
optimize_fft    = yes
; Berendsen temperature coupling is on in two groups
Tcoupl       = Nose-Hoover
tc_grps      = DOPC_LIG    Water
tau_t        = 0.5        0.5
ref_t        = 310        310
; Pressure coupling is on
Pcoupl          = Parrinello-Rahman
pcoupltype      = semiisotropic
tau_p           = 1.0  1.0
compressibility = 4.5e-5  4.5e-5
ref_p           = 1.0  1.0
refcoord_scaling = com
; Generate velocities is off
gen_vel      = no
; Periodic boundary conditions are on in all directions
pbc      = xyz
; Long-range dispersion correction
DispCorr     = EnerPres
; Pull code
pull            = umbrella
pull_geometry   = direction
pull_dim        = N N Y
pull_vec1       = 0 0 -1
pull_start      = yes
pull_ngroups    = 1
pull_group0     = DOPC
pull_group1     = LIG
pull_init1      = 0
pull_rate1      = 0.0
pull_k1         = 1000       ; kJ mol^-1 nm^-2
pull_nstxout    = 1000       ; every 2 ps
pull_nstfout    = 1000       ; every 2 ps
```
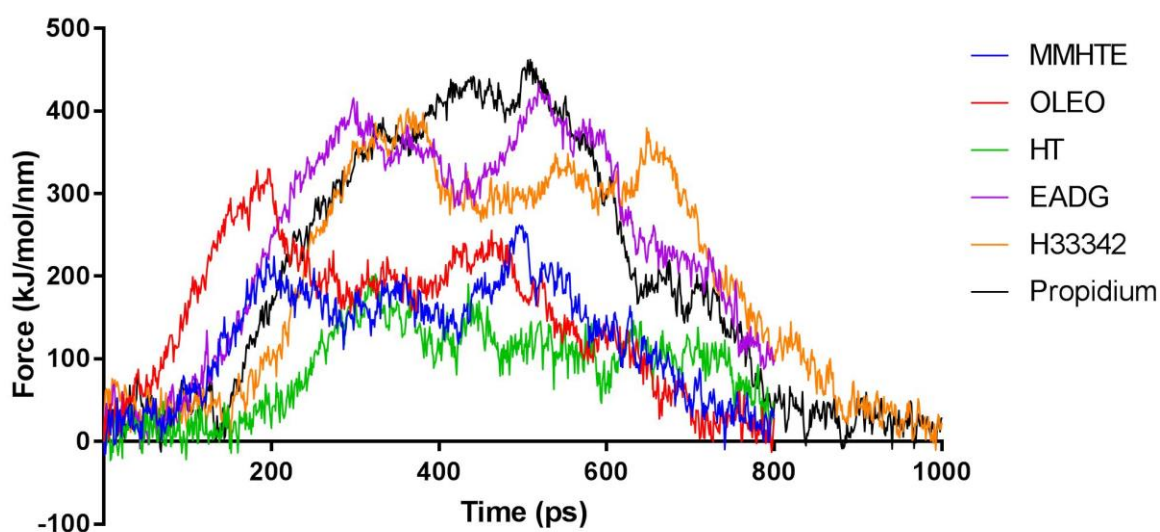
Figure A3.1: Force profile as ligands are pulled through DOPC membrane with respect to time for methyl malate-β-hydroxytyrosol ester (MMHTE), oleocanthal (OLEO), hydroxytyrosol (HT), elenoic acid diglucoside (EADG), Hoescht 33342 (H33342), and propidium. Forces are shown as an average of ten runs for each ligand.
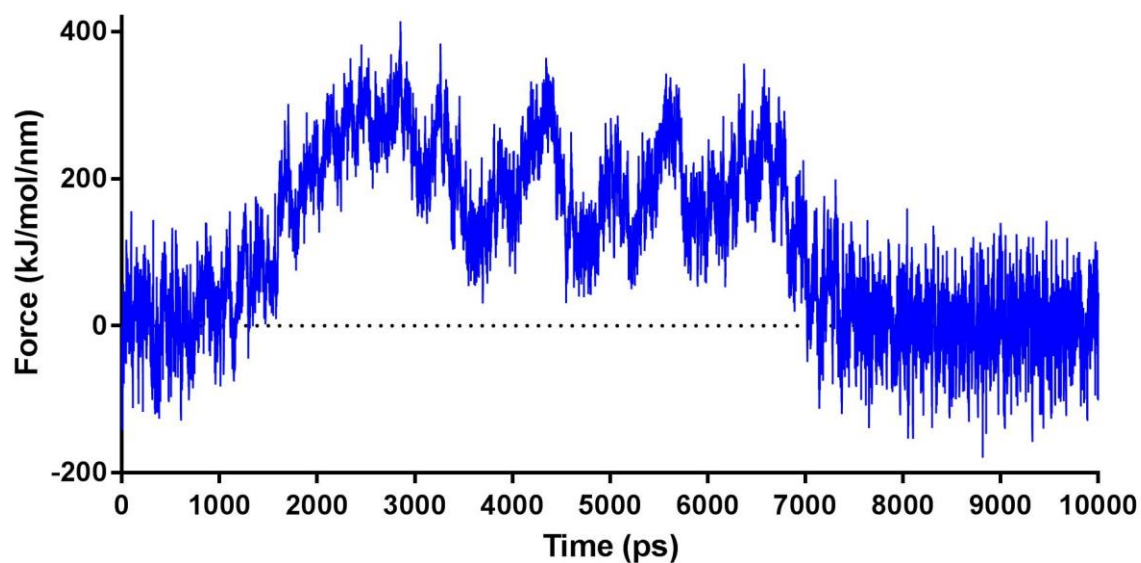
Figure A3.2: Force profile for elenolic acid diglucoside passing through DOPC membrane with a pull rate of 1.0 nm/ns with respect to time.

9.4 <u>Molecular dynamics simulation</u>

9.4.1   *mdp files for classical MD simulations*

9.4.1.1 Energy minimisation
```
title            = Minimization  ; Title of run

; Parameters describing what to do, when to stop and what to save
integrator       = steep         ; Algorithm (steep = steepest descent
minimization)
emtol            = 1000.0        ; Stop minimization when the maximum force
< 10.0 kJ/mol
emstep       = 0.01      ; Energy step size
nsteps           = 50000         ; Maximum number of (minimization) steps to
perform
energygrps       = Protein       ; Which energy group(s) to write to disk

; Parameters describing how to find the neighbors of each atom and how to
calculate the interactions
nstlist          = 1             ; Frequency to update the neighbor list
and long range forces
cutoff-scheme    = Verlet
ns_type          = grid          ; Method to determine neighbor list
(simple, grid)
rlist            = 1.0           ; Cut-off for making neighbor list
(short range forces)
coulombtype      = PME           ; Treatment of long range
electrostatic interactions
rcoulomb         = 1.0           ; long range electrostatic cut-off
rvdw             = 1.0           ; long range Van der Waals cut-off
pbc              = xyz           ; Periodic Boundary Conditions
```

9.4.1.2 NVT
```
title        = Protein-ligand complex NVT equilibration
define       = -DPOSRES  ; position restrain the protein and ligand
; Run parameters
integrator  = md         ; leap-frog integrator
nsteps       = 50000      ; 2 * 50000 = 100 ps
dt           = 0.002      ; 2 fs
; Output control
nstxout      = 500        ; save coordinates every 1.0 ps
nstvout      = 500        ; save velocities every 1.0 ps
nstenergy    = 500        ; save energies every 1.0 ps
nstlog       = 500        ; update log file every 1.0 ps
energygrps   = Protein
; Bond parameters
continuation     = no           ; first dynamics run
constraint_algorithm = lincs    ; holonomic constraints
constraints      = all-bonds    ; all bonds (even heavy atom-H bonds)
constrained
lincs_iter       = 1            ; accuracy of LINCS
lincs_order      = 4            ; also related to accuracy
; Neighborsearching
cutoff-scheme    = Verlet
ns_type          = grid        ; search neighboring grid cells
nstlist          = 10          ; 20 fs, largely irrelevant with Verlet
rcoulomb         = 1.4         ; short-range electrostatic cutoff (in nm)
rvdw             = 1.4         ; short-range van der Waals cutoff (in nm)
; Electrostatics
coulombtype      = PME         ; Particle Mesh Ewald for long-range
electrostatics
pme_order        = 4          ; cubic interpolation
fourierspacing  = 0.16        ; grid spacing for FFT
; Temperature coupling
tcoupl       = V-rescale                    ; modified Berendsen
thermostat
```

```
tc-grps      = Protein_LIG Water_and_ions   ; two coupling groups - more
accurate
tau_t        = 0.1   0.1                     ; time constant, in ps
ref_t        = 300   300                     ; reference temperature, one
for each group, in K
; Pressure coupling
pcoupl       = no            ; no pressure coupling in NVT
; Periodic boundary conditions
pbc          = xyz           ; 3-D PBC
; Dispersion correction
DispCorr     = EnerPres   ; account for cut-off vdW scheme
; Velocity generation
gen_vel      = yes           ; assign velocities from Maxwell distribution
gen_temp     = 300           ; temperature for Maxwell distribution
gen_seed     = -1            ; generate a random seed
```

9.4.1.3 NPT

```
title       = Protein-ligand complex NPT equilibration
define      = -DPOSRES  ; position restrain the protein and ligand
; Run parameters
integrator  = md         ; leap-frog integrator
nsteps      = 50000      ; 2 * 50000 = 100 ps
dt          = 0.002      ; 2 fs
; Output control
nstxout     = 500        ; save coordinates every 1.0 ps
nstvout     = 500        ; save velocities every 1.0 ps
nstenergy   = 500        ; save energies every 1.0 ps
nstlog      = 500        ; update log file every 1.0 ps
energygrps  = Protein LIG
; Bond parameters
continuation        = yes          ; first dynamics run
constraint_algorithm = lincs    ; holonomic constraints
constraints         = all-bonds    ; all bonds (even heavy atom-H bonds)
constrained
lincs_iter          = 1            ; accuracy of LINCS
lincs_order         = 4            ; also related to accuracy
; Neighborsearching
cutoff-scheme   = Verlet
ns_type         = grid     ; search neighboring grid cells
nstlist         = 10       ; 20 fs, largely irrelevant with Verlet
rcoulomb        = 1.4      ; short-range electrostatic cutoff (in nm)
rvdw            = 1.4      ; short-range van der Waals cutoff (in nm)
; Electrostatics
coulombtype     = PME      ; Particle Mesh Ewald for long-range
electrostatics
pme_order       = 4        ; cubic interpolation
fourierspacing  = 0.16     ; grid spacing for FFT
; Temperature coupling
tcoupl      = V-rescale                      ; modified Berendsen
thermostat
tc-grps      = Protein_LIG Water_and_ions   ; two coupling groups - more
accurate
tau_t        = 0.1   0.1                     ; time constant, in ps
ref_t        = 300   300                     ; reference temperature, one
for each group, in K
; Pressure coupling
pcoupl       = Parrinello-Rahman            ; pressure coupling is on for
NPT
pcoupltype   = isotropic                     ; uniform scaling of box
vectors
tau_p        = 2.0                           ; time constant, in ps
ref_p        = 1.0                           ; reference pressure, in bar
compressibility = 4.5e-5                     ; isothermal compressibility
of water, bar^-1
refcoord_scaling    = com
; Periodic boundary conditions
pbc          = xyz           ; 3-D PBC
; Dispersion correction
DispCorr     = EnerPres   ; account for cut-off vdW scheme
```

157

```
; Velocity generation
gen_vel      = no          ; velocity generation off after NVT
```

9.4.1.4 MD run

```
title        = Protein-ligand complex MD simulation
; Run parameters
integrator = md          ; leap-frog integrator
nsteps       = 50000000  ; 2 * 50000000 = 100000 ps (100 ns)
dt           = 0.002     ; 2 fs
; Output control
nstxout              = 0          ; suppress .trr output
nstvout              = 0          ; suppress .trr output
nstenergy            = 5000       ; save energies every 10.0 ps
nstlog               = 5000       ; update log file every 10.0 ps
nstxout-compressed   = 5000       ; write .xtc trajectory every 10.0 ps
compressed-x-grps    = System
energygrps           = Protein LIG
; Bond parameters
continuation    = no            ; first dynamics run
constraint_algorithm = lincs    ; holonomic constraints
constraints     = all-bonds     ; all bonds (even heavy atom-H bonds)
constrained
lincs_iter      = 1             ; accuracy of LINCS
lincs_order     = 4             ; also related to accuracy
; Neighborsearching
cutoff-scheme   = Verlet
ns_type         = grid         ; search neighboring grid cells
nstlist         = 10           ; 20 fs, largely irrelevant with Verlet
rcoulomb        = 1.4          ; short-range electrostatic cutoff (in nm)
rvdw            = 1.4          ; short-range van der Waals cutoff (in nm)
; Electrostatics
coulombtype     = PME          ; Particle Mesh Ewald for long-range
electrostatics
pme_order       = 4            ; cubic interpolation
fourierspacing  = 0.16         ; grid spacing for FFT
; Temperature coupling
tcoupl       = V-rescale                   ; modified Berendsen
thermostat
tc-grps      = Protein_LIG Water_and_ions  ; two coupling groups - more
accurate
tau_t        = 0.1   0.1                    ; time constant, in ps
ref_t        = 300   300                    ; reference temperature, one
for each group, in K
; Pressure coupling
pcoupl       = Parrinello-Rahman           ; pressure coupling is on for
NPT
pcoupltype   = isotropic                    ; uniform scaling of box
vectors
tau_p        = 2.0                          ; time constant, in ps
ref_p        = 1.0                          ; reference pressure, in bar
compressibility = 4.5e-5                    ; isothermal compressibility
of water, bar^-1
; Periodic boundary conditions
pbc          = xyz         ; 3-D PBC
; Dispersion correction
DispCorr     = EnerPres  ; account for cut-off vdW scheme
; Velocity generation
gen_vel      = yes         ; assign velocities from Maxwell distribution
gen_seed     = -1          ; generate random velocities with each run
gen_temp     = 300         ; so that first step will be randomised
```
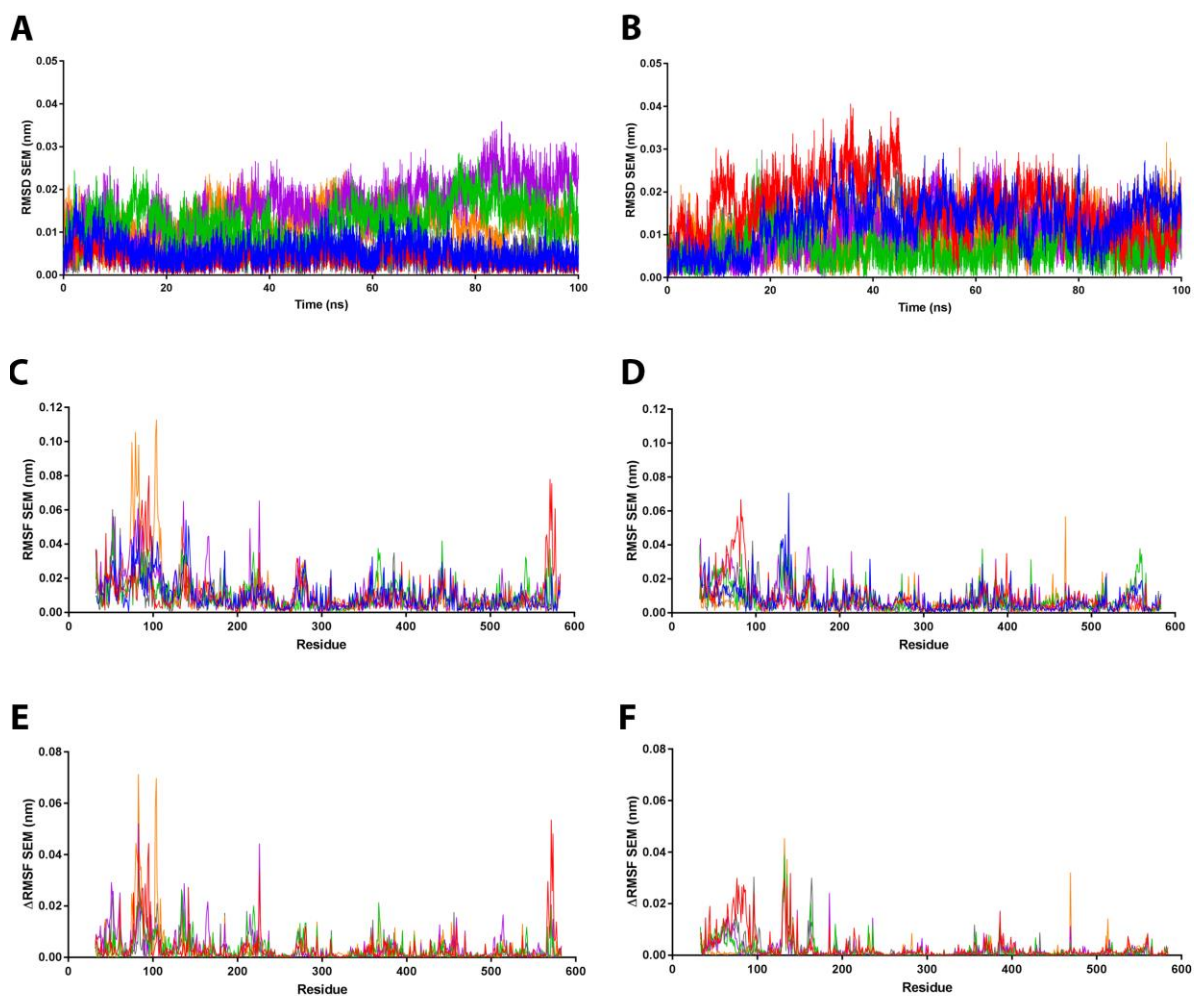
Figure A4.1: Standard error for 100 ns trajectories of COX-1 (A, C, E) and COX-2 (B, D, F) proteins bound to olive ligands: apo (blue), native ligand (red), 1-oleyltyrosol (green), ligstroside derivative 2 (purple), oleocanthal (orange), and methyl malate-β-hydroxytyrosol ester (grey). The standard error for the RMSD of backbone atoms is shown (A-B). RMSF for the protein fit to the backbone is shown with respect to time. Standard error in RMSF for protein fit to the protein backbone is shown with respect to residue (C-D). The standard error for the difference in RMSF when values for the apo form is substracted from ligand bound forms of the protein is also shown with respect to residue (E-F).

9.4.3 Essential Dynamics

9.4.3.1 Essential dynamics values

Table A4.1a: Cosine content for triplicate and concatenated trajectories for COX-1

Traj	1	2	3	4	5	6	7	8	9	10
Apo 1-01	0.825	0.374	0.198	0.000	0.015	0.080	0.115	0.028	0.026	0.000
Apo 1-02	0.880	0.387	0.056	0.002	0.087	0.075	0.122	0.002	0.126	0.004
Apo 1-03	0.941	0.151	0.060	0.034	0.039	0.000	0.068	0.001	0.074	0.003
Apo 1-cat	0.264	0.203	0.143	0.028	0.103	0.007	0.000	0.015	0.035	0.029
FLP 1-01	0.827	0.465	0.291	0.007	0.066	0.008	0.005	0.002	0.044	0.002
FLP 1-02	0.906	0.471	0.239	0.269	0.000	0.029	0.005	0.024	0.001	0.018
FLP 1-03	0.928	0.508	0.270	0.008	0.049	0.003	0.003	0.008	0.010	0.006
FLP 1-cat	0.518	0.415	0.270	0.047	0.014	0.002	0.126	0.015	0.107	0.001
1OL 1-01	0.916	0.337	0.249	0.281	0.308	0.066	0.003	0.036	0.023	0.152
1OL 1-02	0.880	0.687	0.286	0.019	0.031	0.123	0.017	0.044	0.014	0.072
1OL 1-03	0.865	0.618	0.113	0.192	0.003	0.068	0.052	0.012	0.000	0.024
1OL 1-cat	0.557	0.397	0.107	0.389	0.016	0.051	0.027	0.053	0.002	0.000
LG2 1-01	0.436	0.378	0.019	0.021	0.011	0.008	0.002	0.006	0.028	0.087
LG2 1-02	0.847	0.496	0.343	0.044	0.060	0.209	0.016	0.008	0.080	0.094
LG2 1-03	0.875	0.424	0.031	0.128	0.022	0.071	0.026	0.003	0.040	0.045
LG2 1-cat	0.257	0.342	0.045	0.135	0.003	0.037	0.001	0.007	0.002	0.045
OLEO 1-01	0.863	0.564	0.016	0.031	0.000	0.011	0.000	0.000	0.054	0.053
OLEO 1-02	0.775	0.497	0.247	0.163	0.128	0.018	0.050	0.028	0.004	0.000
OLEO 1-03	0.908	0.032	0.031	0.000	0.023	0.001	0.000	0.050	0.051	0.044
OLEO 1-cat	0.000	0.006	0.000	0.143	0.093	0.108	0.048	0.001	0.018	0.037
MMHTE 1-01	0.798	0.065	0.045	0.061	0.031	0.021	0.011	0.000	0.149	0.023
MMHTE 1-02	0.571	0.508	0.391	0.418	0.009	0.001	0.037	0.100	0.000	0.001
MMHTE 1-03	0.888	0.287	0.158	0.306	0.070	0.014	0.118	0.001	0.074	0.028
MMHTE 1-cat	0.442	0.284	0.134	0.280	0.034	0.001	0.105	0.016	0.004	0.019

Table A4.1b: Cosine content for triplicate and concatenated trajectories for COX-2

Traj	1	2	3	4	5	6	7	8	9	10
Apo 2-01	0.014	0.026	0.056	0.064	0.010	0.030	0.002	0.014	0.010	0.018
Apo 2-02	0.822	0.705	0.175	0.069	0.000	0.013	0.070	0.033	0.005	0.037
Apo 2-03	0.756	0.081	0.043	0.033	0.006	0.000	0.008	0.071	0.003	0.016
Apo 2-cat	0.447	0.582	0.107	0.000	0.082	0.039	0.001	0.016	0.029	0.001
SAL 2-01	0.721	0.691	0.016	0.004	0.019	0.002	0.170	0.097	0.026	0.013
SAL 2-02	0.130	0.227	0.000	0.001	0.000	0.048	0.082	0.015	0.003	0.031
SAL 2-03	0.876	0.044	0.227	0.023	0.001	0.061	0.013	0.002	0.101	0.014
SAL 2-cat	0.727	0.539	0.052	0.114	0.191	0.112	0.012	0.002	0.007	0.012
1OL 2-01	0.820	0.029	0.045	0.003	0.021	0.059	0.009	0.037	0.000	0.003
1OL 2-02	0.766	0.175	0.002	0.001	0.010	0.002	0.011	0.011	0.009	0.081
1OL 2-03	0.819	0.631	0.007	0.004	0.030	0.010	0.002	0.001	0.004	0.037
1OL 2-cat	0.274	0.183	0.001	0.342	0.022	0.142	0.018	0.001	0.002	0.081
LG2 2-01	0.875	0.112	0.068	0.033	0.017	0.002	0.000	0.000	0.010	0.003
LG2 2-02	0.512	0.172	0.000	0.120	0.000	0.000	0.038	0.002	0.012	0.009
LG2 2-03	0.684	0.370	0.142	0.075	0.061	0.000	0.012	0.000	0.011	0.001
LG2 2-cat	0.032	0.005	0.027	0.006	0.000	0.022	0.119	0.070	0.041	0.030
OLEO 2-01	0.863	0.564	0.016	0.031	0.000	0.011	0.000	0.000	0.054	0.053
OLEO 2-02	0.775	0.497	0.247	0.163	0.128	0.018	0.050	0.028	0.004	0.000
OLEO 2-03	0.908	0.032	0.031	0.000	0.023	0.001	0.000	0.050	0.051	0.044
OLEO 2-cat	0.000	0.006	0.000	0.143	0.093	0.108	0.048	0.001	0.018	0.037
MMHTE 2-01	0.287	0.452	0.000	0.053	0.000	0.002	0.015	0.020	0.000	0.005
MMHTE 2-02	0.808	0.005	0.086	0.052	0.015	0.025	0.000	0.008	0.002	0.000
MMHTE 2-03	0.425	0.085	0.008	0.005	0.049	0.005	0.204	0.000	0.034	0.035
MMHTE 2-cat	0.312	0.351	0.092	0.023	0.003	0.022	0.001	0.000	0.038	0.003
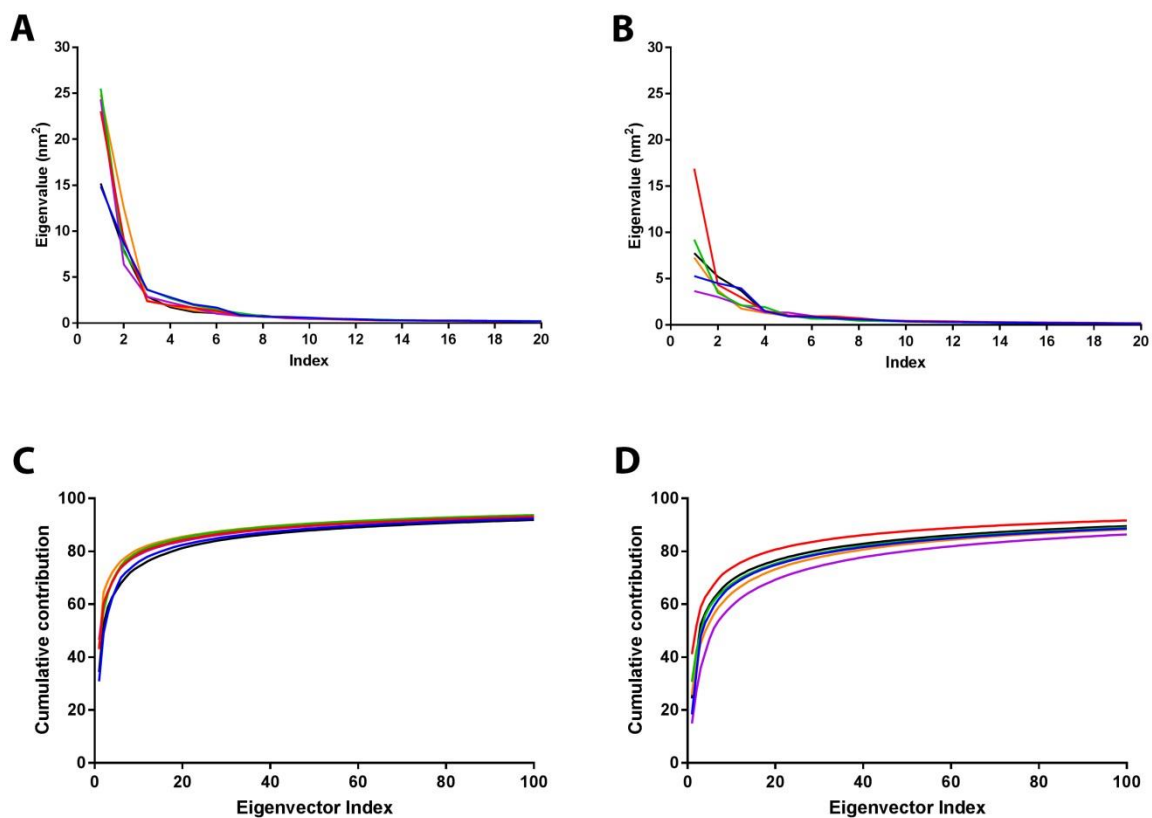
Figure A4.2: Eigenvalues (A, B) and cumulative contribution (C, D) for COX-1 and COX-2 in its apo form (blue), and bound to its native ligand (red), 1OL (green), LG2 (purple), OLEO (orange), and MMHTE (black).

Table A4.2: Colour scale values for porcupine plots

	Eigenvector	COX1	COX2
APO	1	0.25-13.55	0.15-9.15
	2	0.14-7.98	0.50-20.14
	3	1.85-16.28	1.03-21.32
FLP/SAL	1	0.19-12.20	0.39-15.79
	2	0.21-9.69	0.56-22.17
	3	0.08-8.86	1.19-9.39
1OL	1	0.11-14.59	0.07-14.80
	2	0.13-10.21	0.21-14.40
	3	0.26-6.43	0.13-5.58
LG2	1	0.66-13.49	0.05-9.01
	2	0.55-11.44	1.24-8.07
	3	0.12-9.49	1.16-9.42
OLEO	1	0.17-12.16	0.57-11.38
	2	0.14-7.33	0.14-7.21
	3	0.03-4.48	0.23-8.89
MMHTE	1	0.78-12.01	0.20-24.04
	2	0.68-22.42	0.16-8.08
	3	1.30-14.22	0.14-9.37
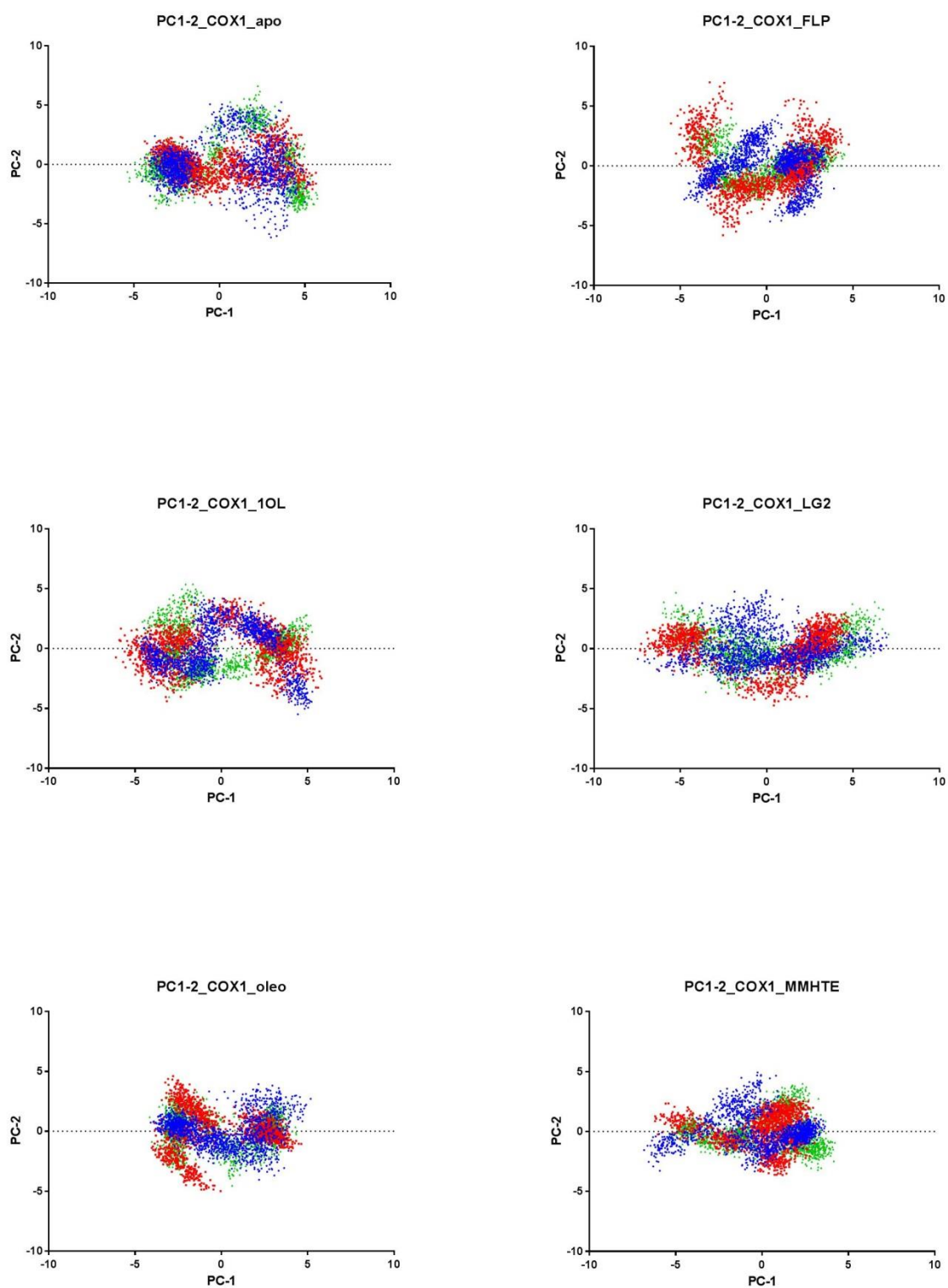
Figure A4.3a: 2D plot for PC-1 against PC-2 for COX-1 complexes
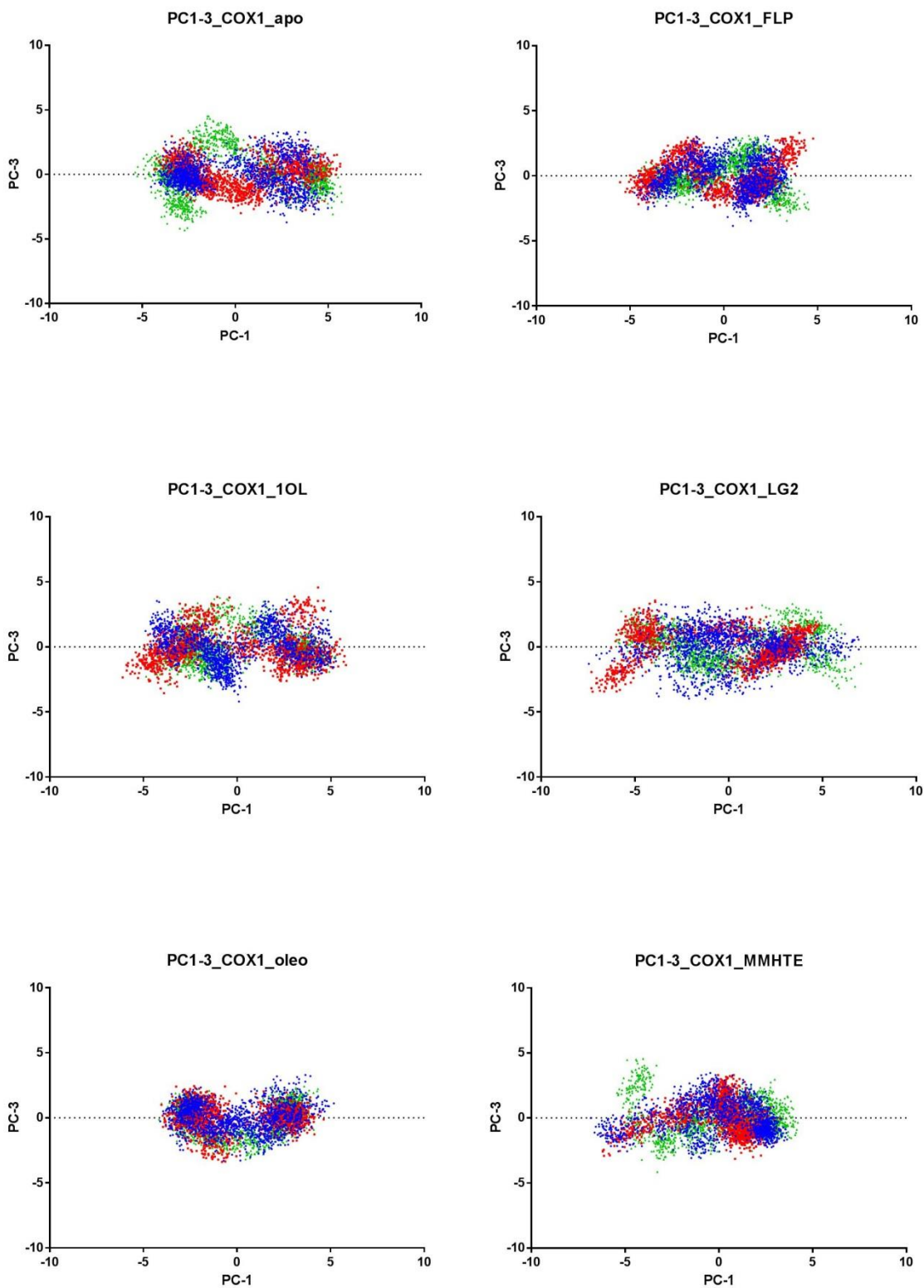
Figure A4.3b: 2D plot for PC-1 against PC-3 for COX-1 complexes
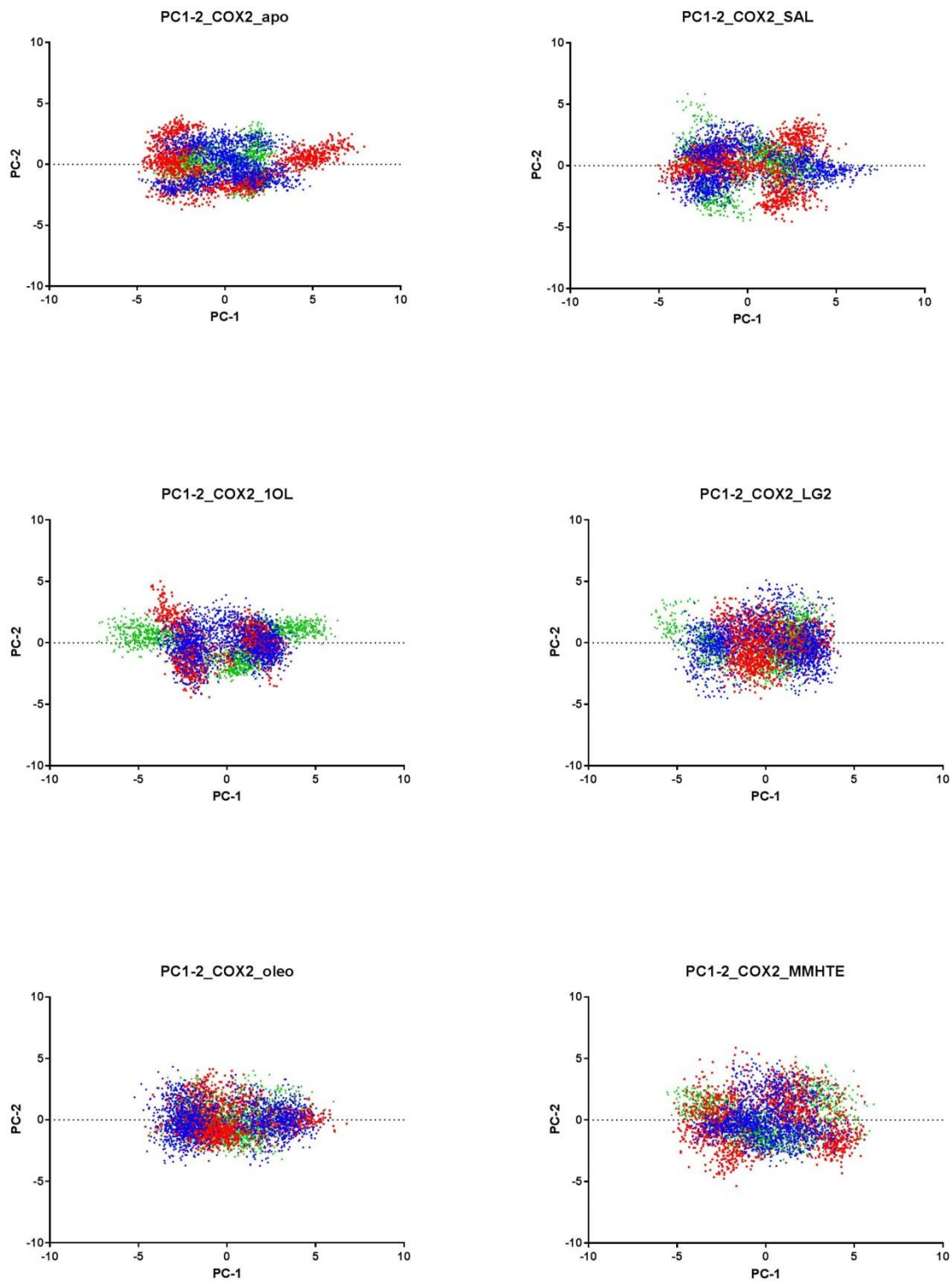
Figure A4.3c: 2D plot for PC-1 against PC-2 for COX-2 complexes

166
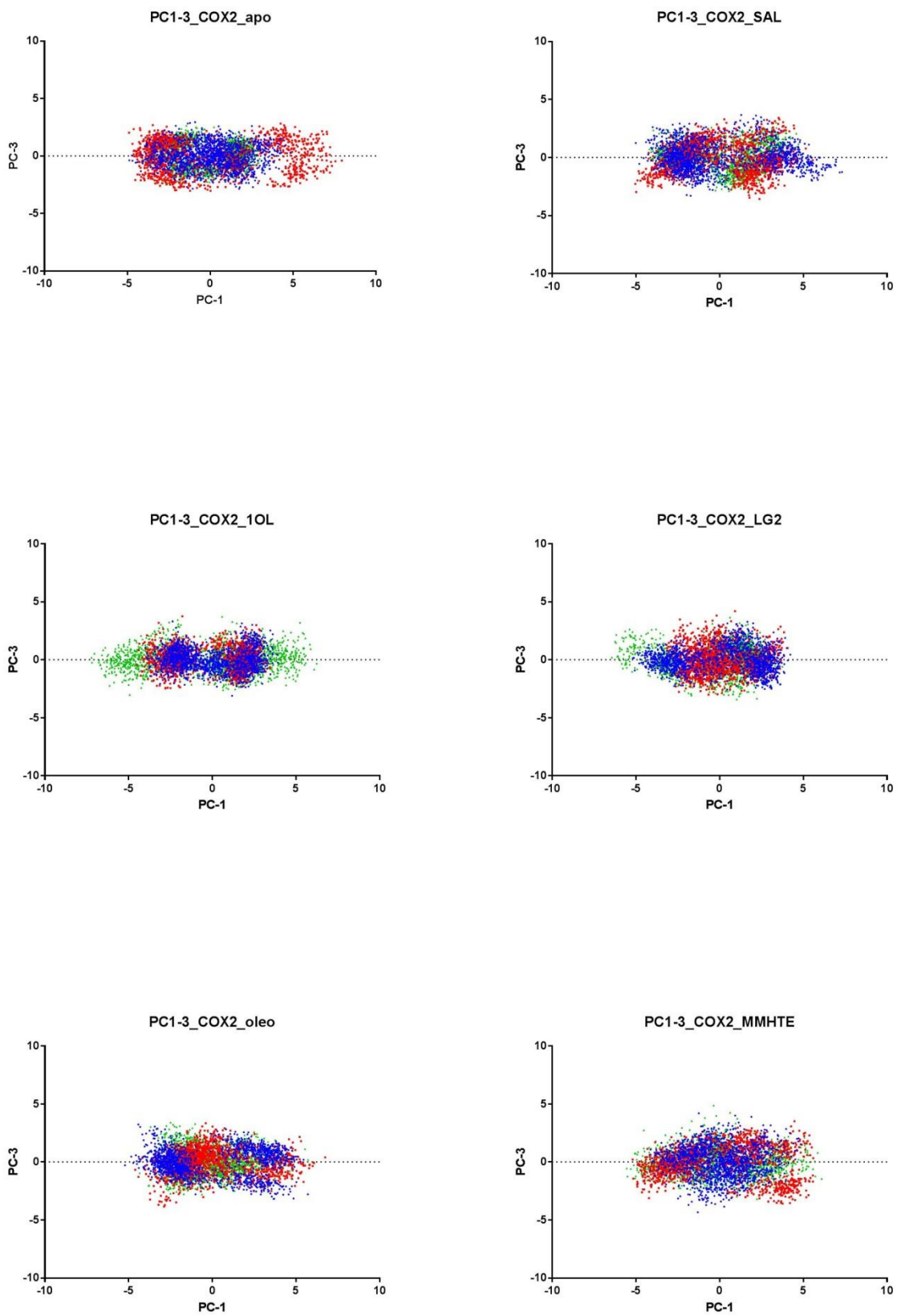
Figure A4.3d: 2D plot for PC-1 against PC-3 for COX-2 complexes
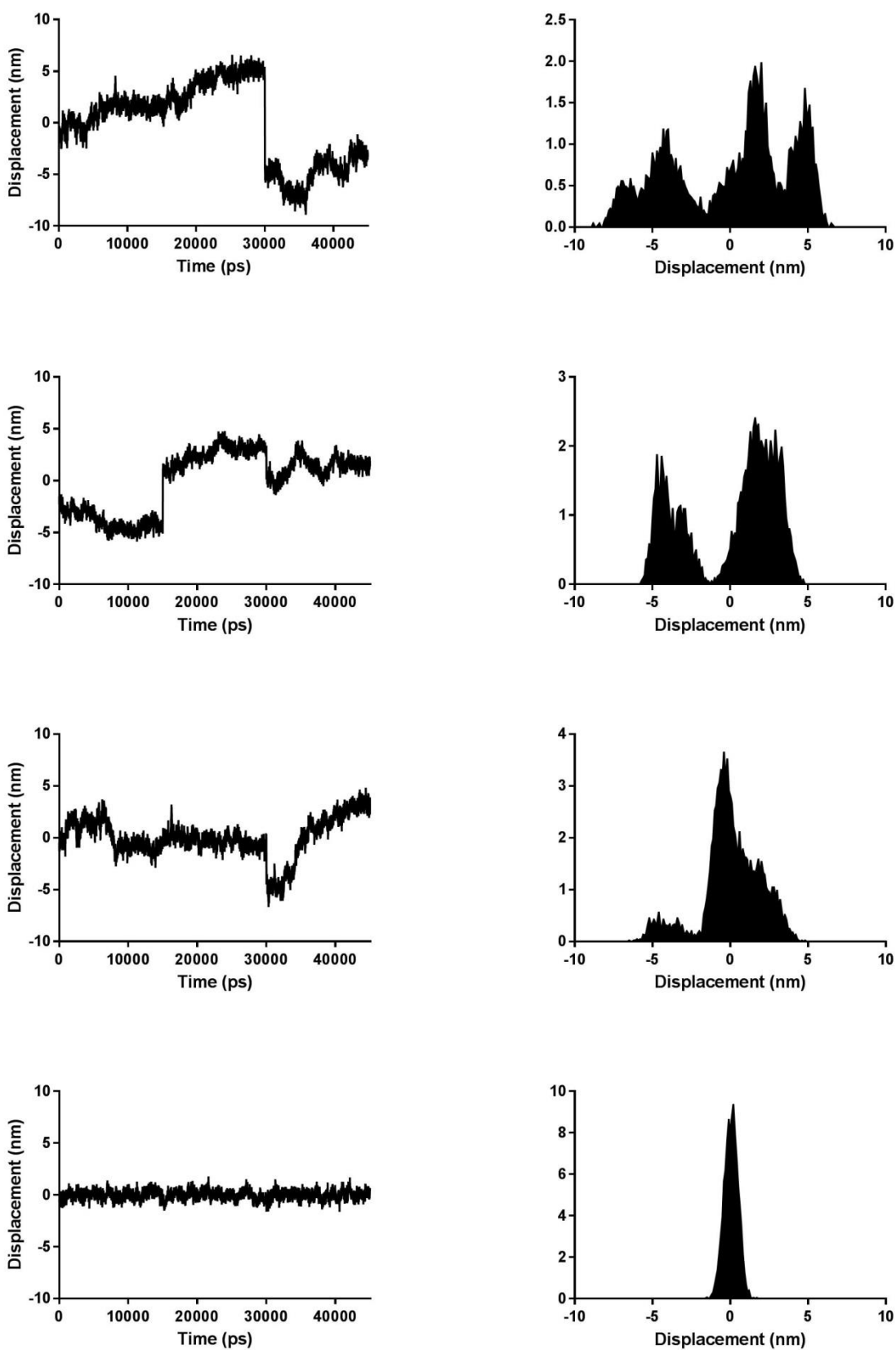
Figure A4.4a: Projection and distribution for COX-1 in its apo form of PCA-1, PCA-2, PCA-3, and PCA-20.
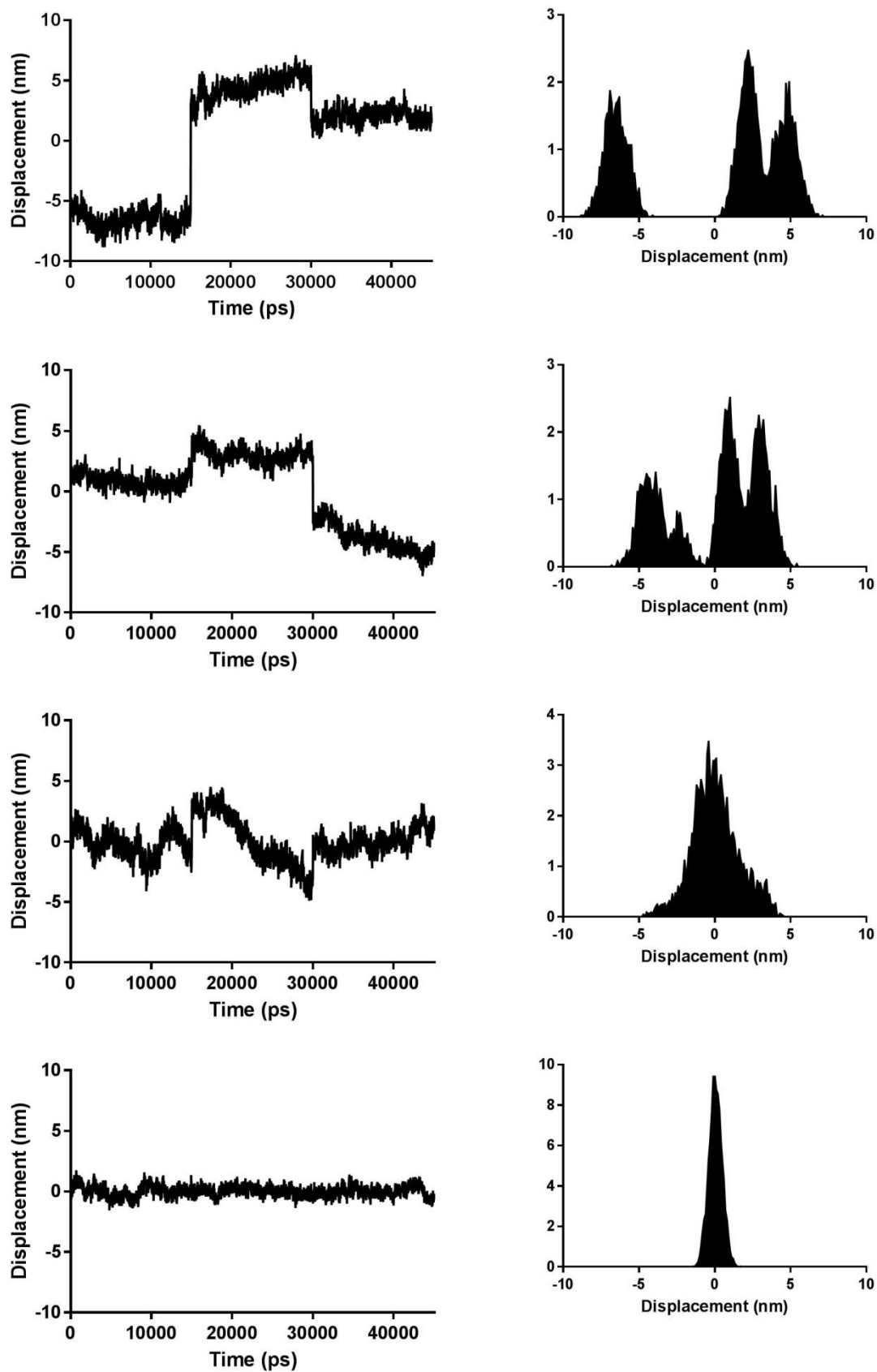
Figure A4.4b: Projection and distribution for COX-1 bound to FLP of PCA-1, PCA-2, PCA-3, and PCA-20.
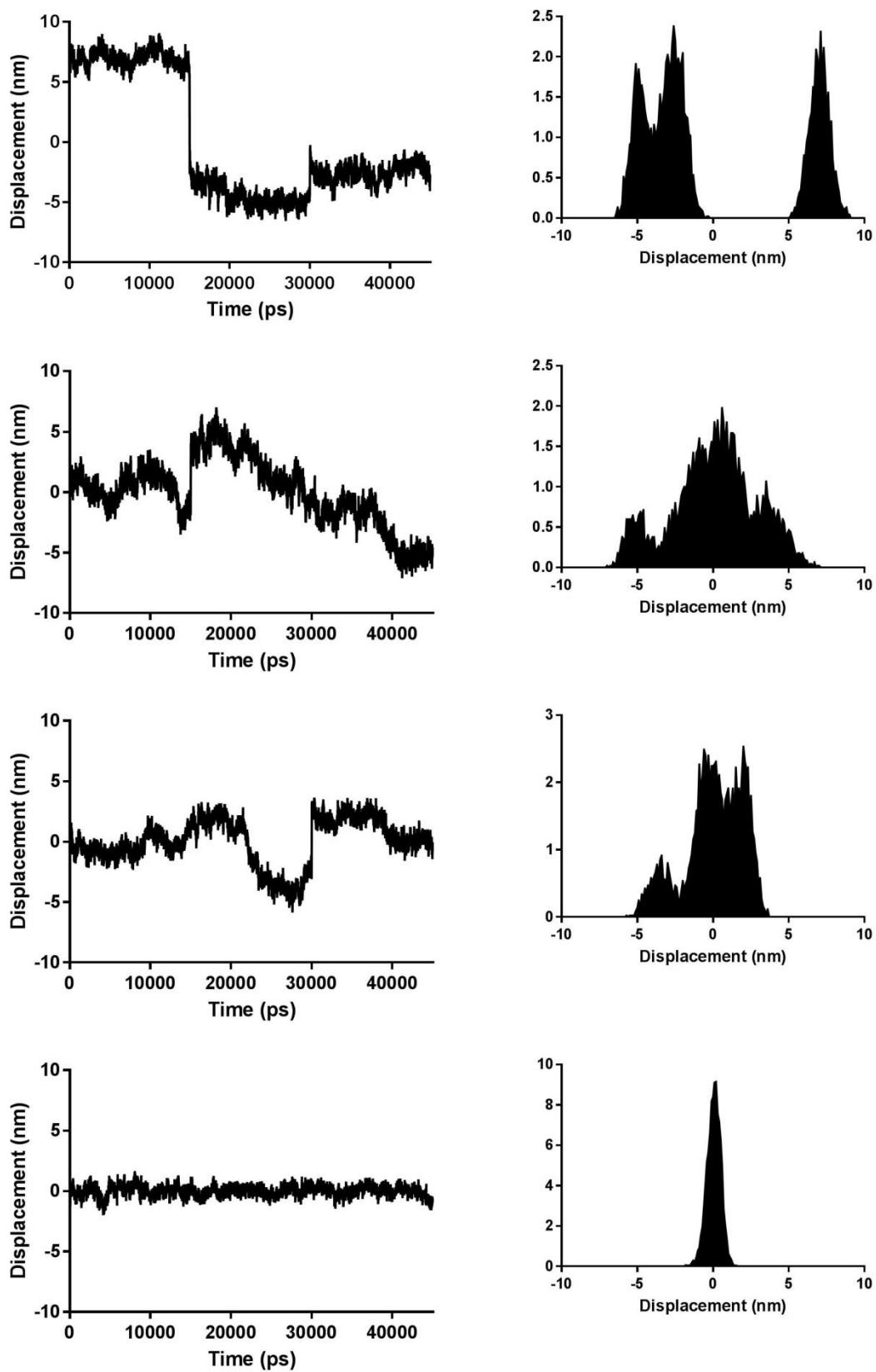
Figure A4.4c: Projection and distribution for COX-1 bound to 1OL of PCA-1, PCA-2, PCA-3, and PCA-20.
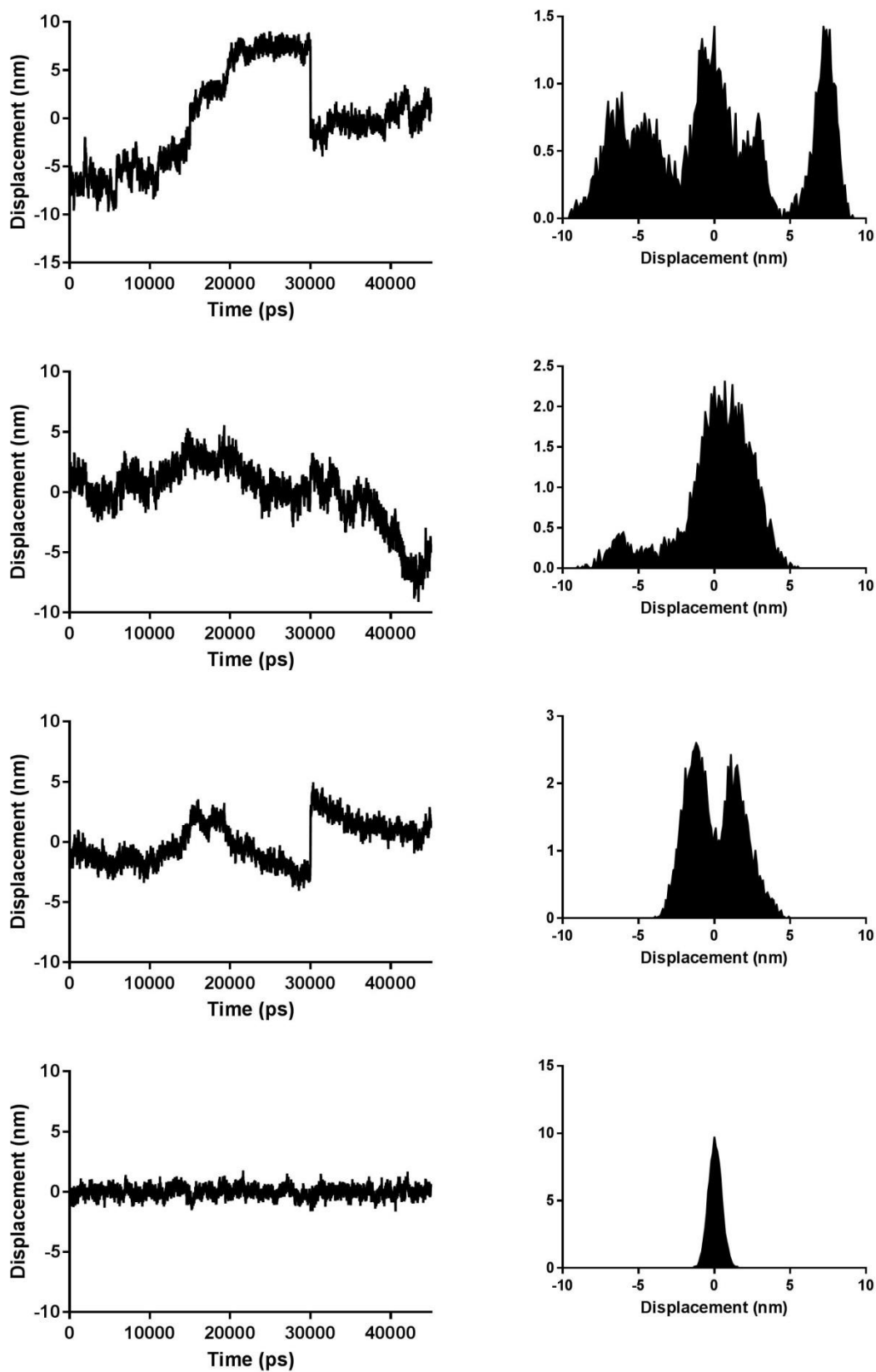
Figure A4.4d: Projection and distribution for COX-1 bound to LG2 of PCA-1, PCA-2, PCA-3, and PCA-20.
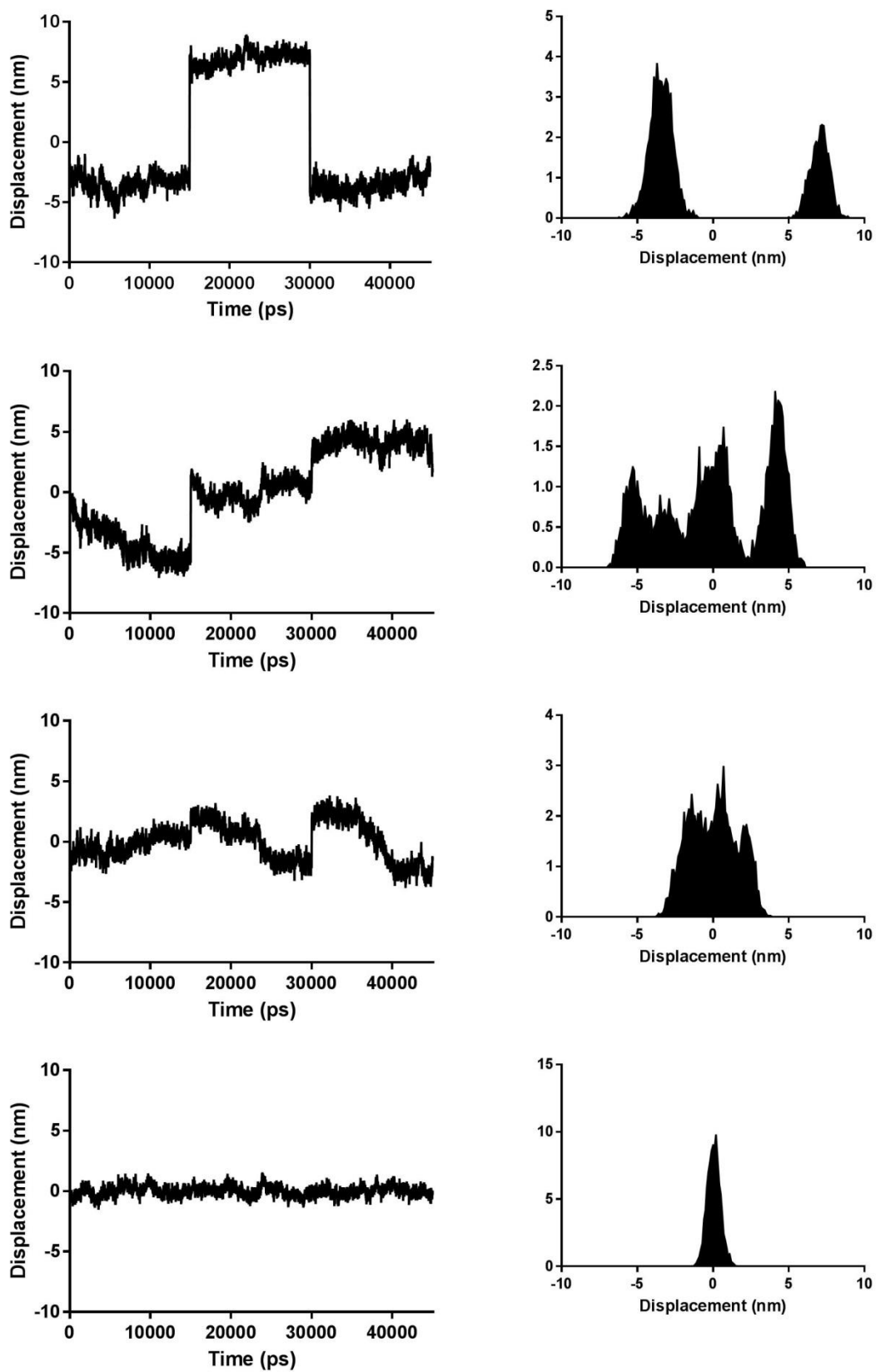
Figure A4.4e: Projection and distribution for COX-1 bound to OLEO of PCA-1, PCA-2, PCA-3, and PCA-20.
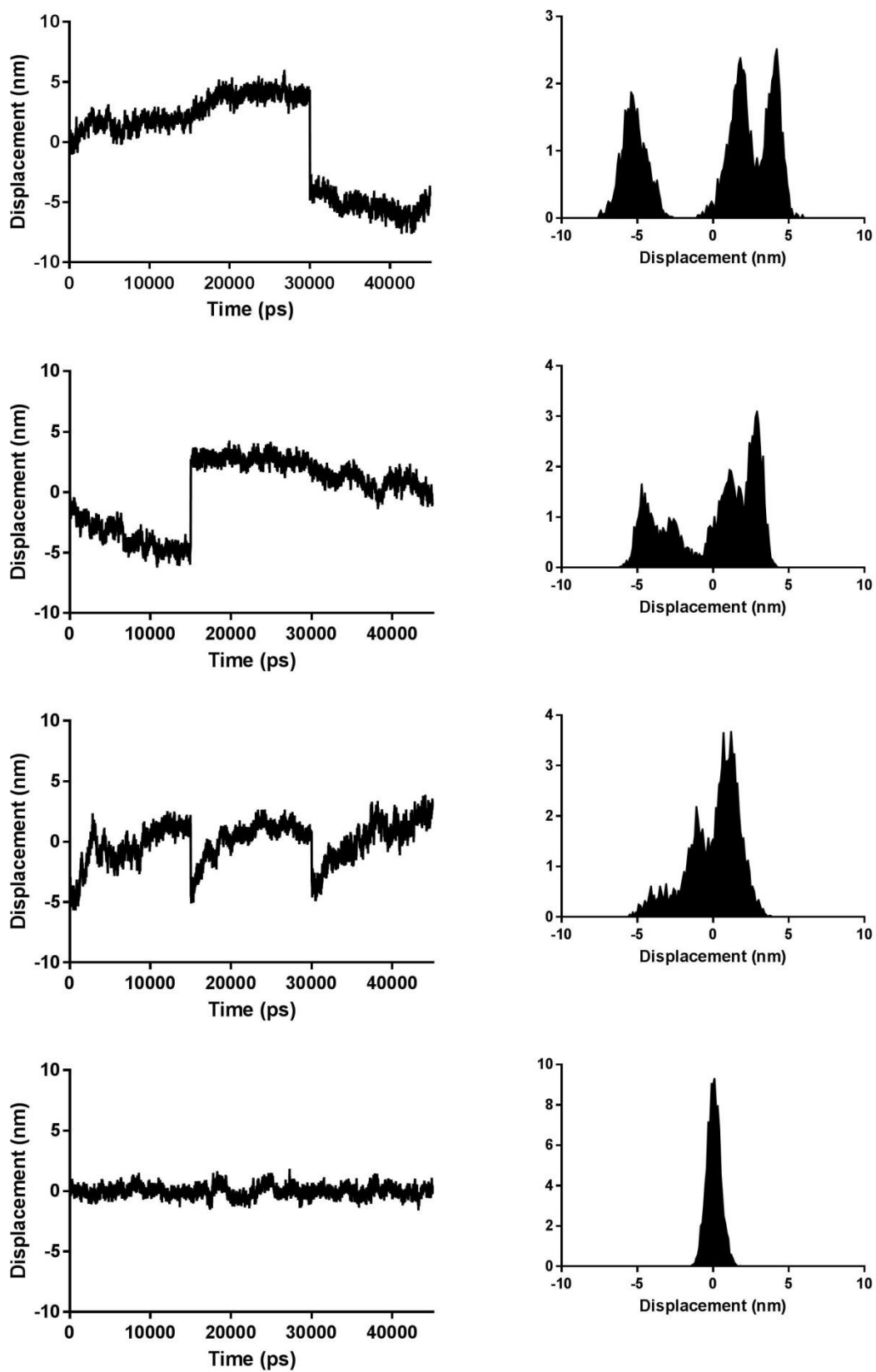
Figure A4.4f: Projection and distribution for COX-1 bound to MMHTE of PCA-1, PCA-2, PCA-3, and PCA-20.

173
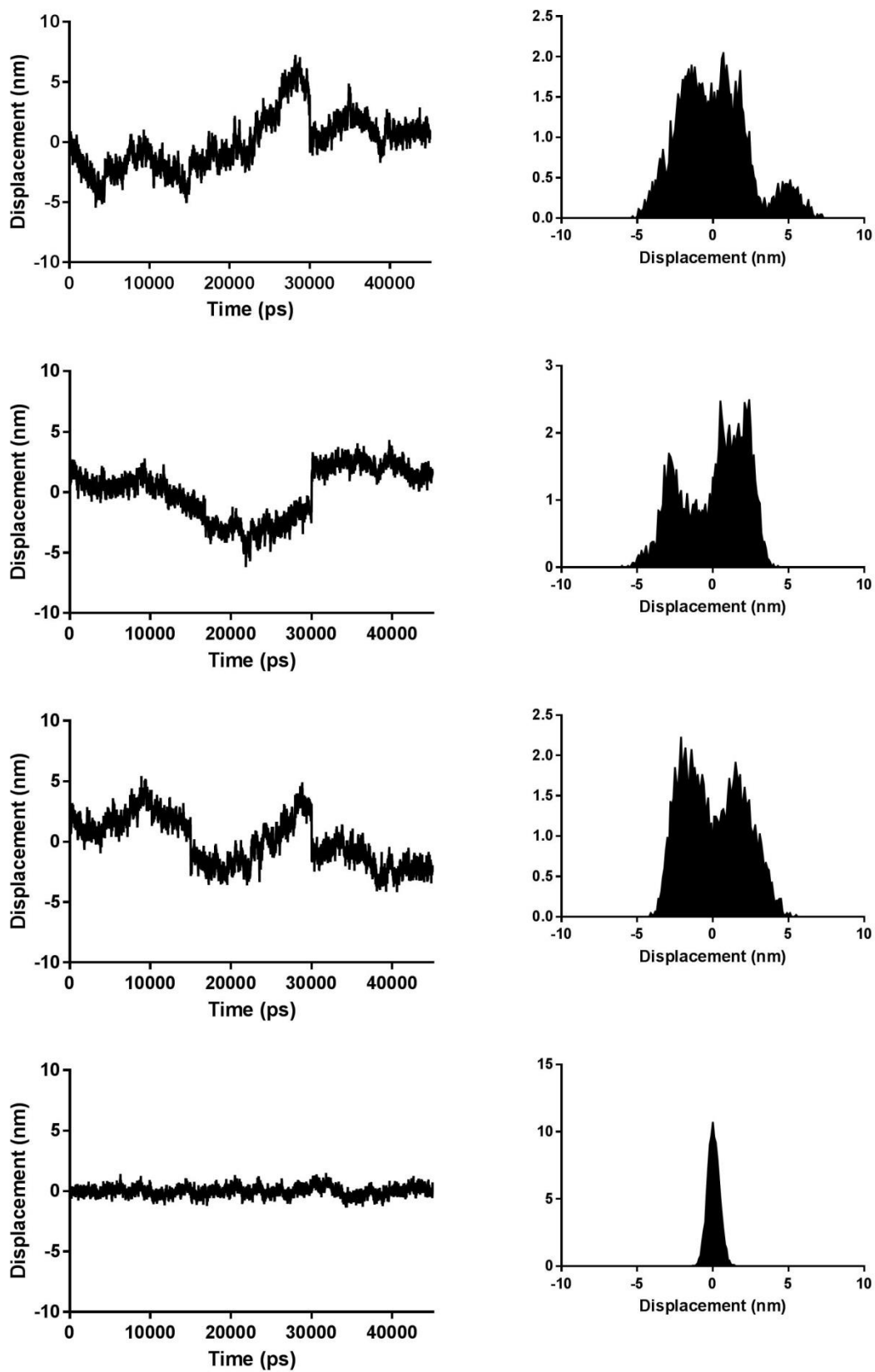
Figure A4.4g: Projection and distribution for COX-2 in its apo form of PCA-1, PCA-2, PCA-3, and PCA-20.

174
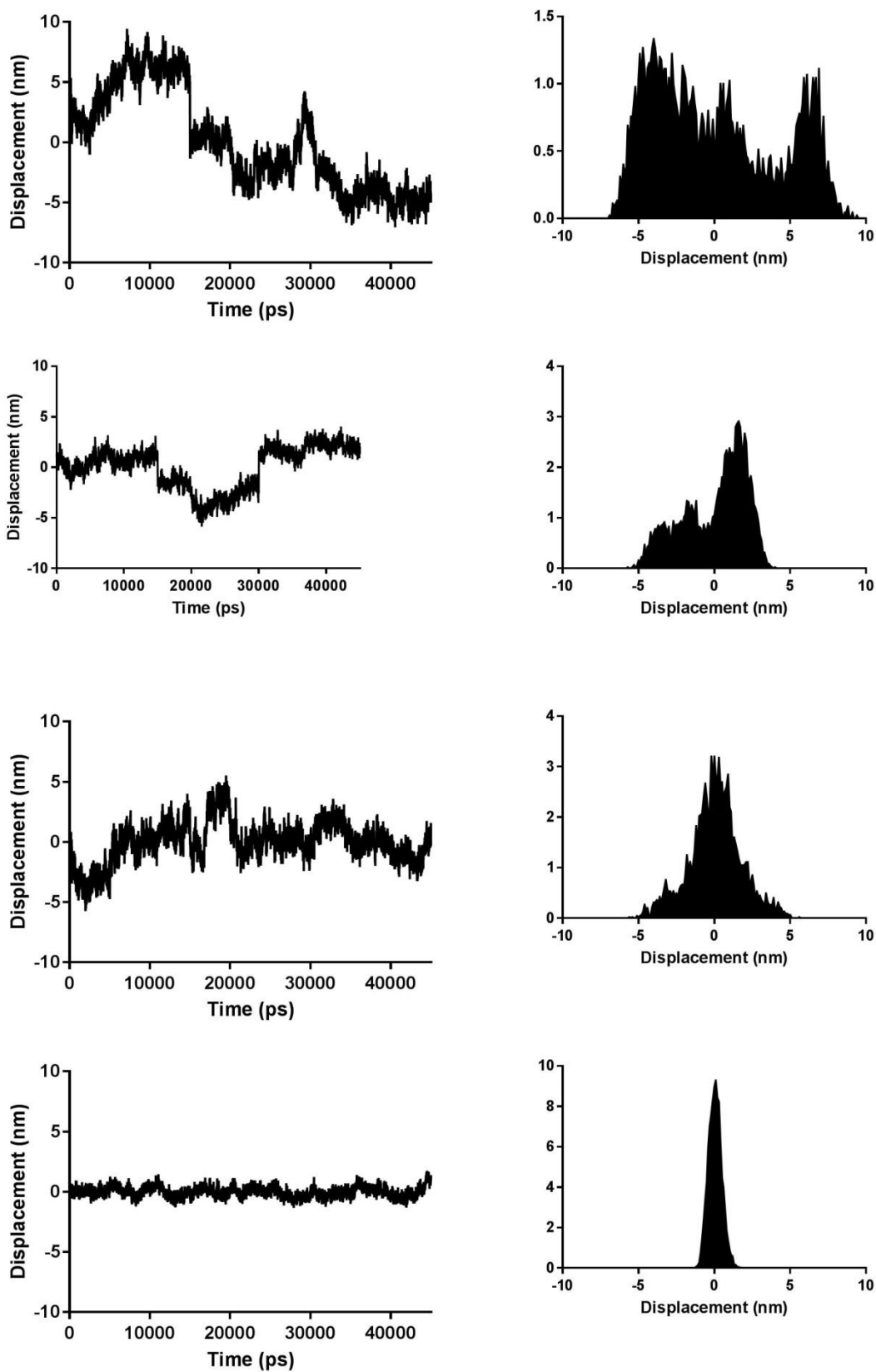
Figure A4.4h: Projection and distribution for COX-2 bound to SAL of PCA-1, PCA-2, PCA-3, and PCA-20.
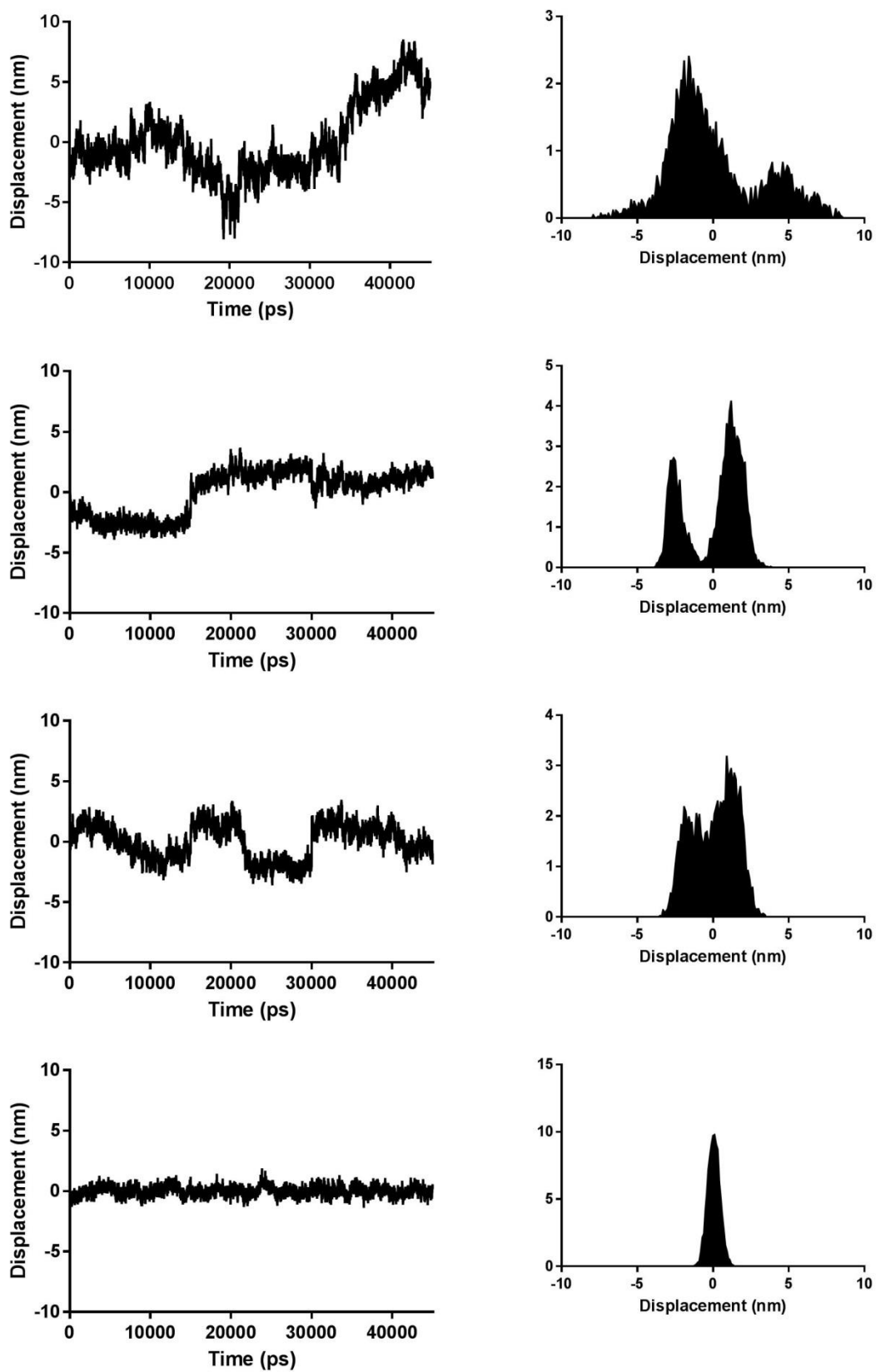
Figure A4.4i: Projection and distribution for COX-2 bound to 1OL of PCA-1, PCA-2, PCA-3, and PCA-20.
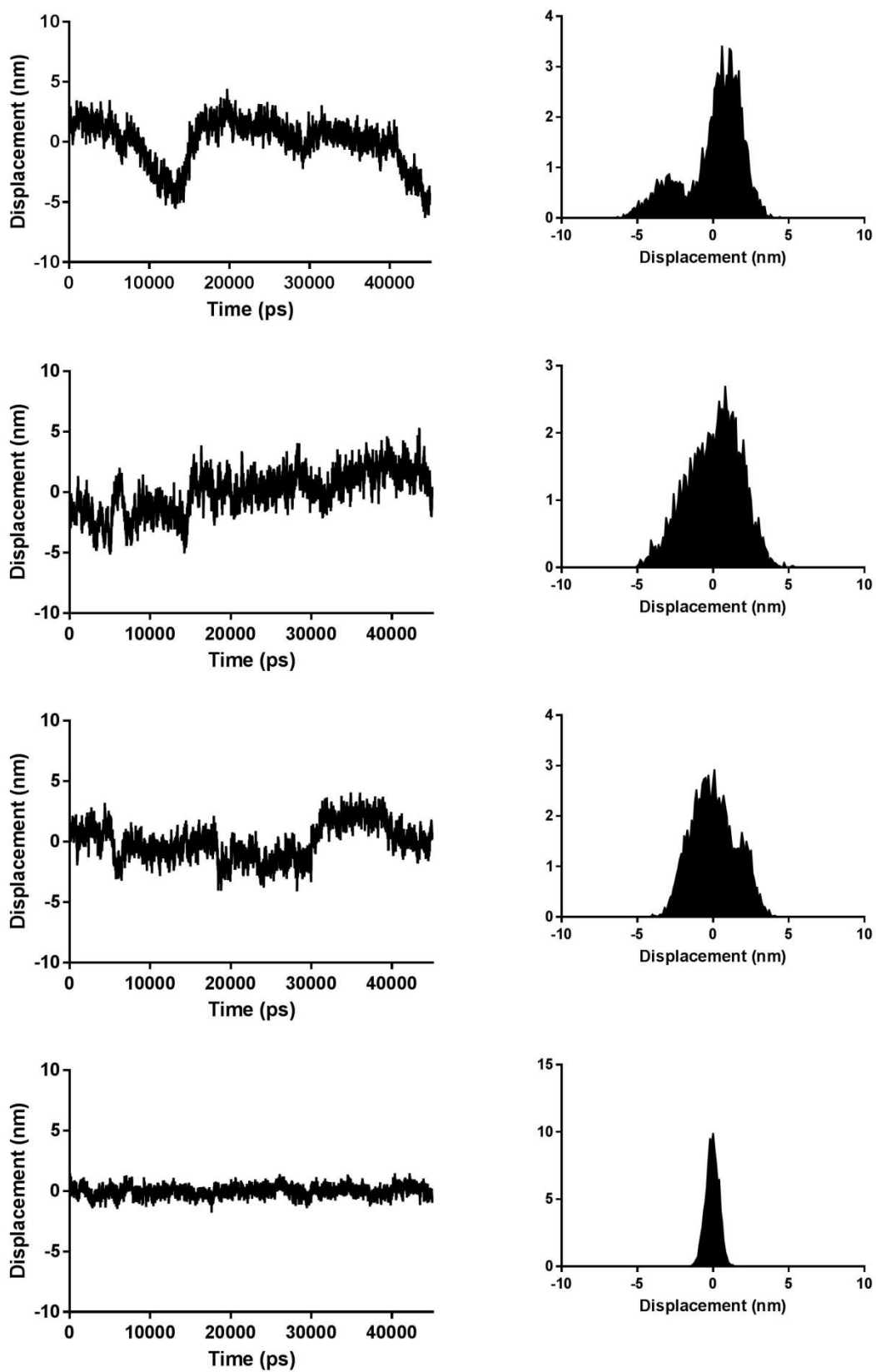
Figure A4.4j: Projection and distribution for COX-2 bound to LG2 of PCA-1, PCA-2, PCA-3, and PCA-20.
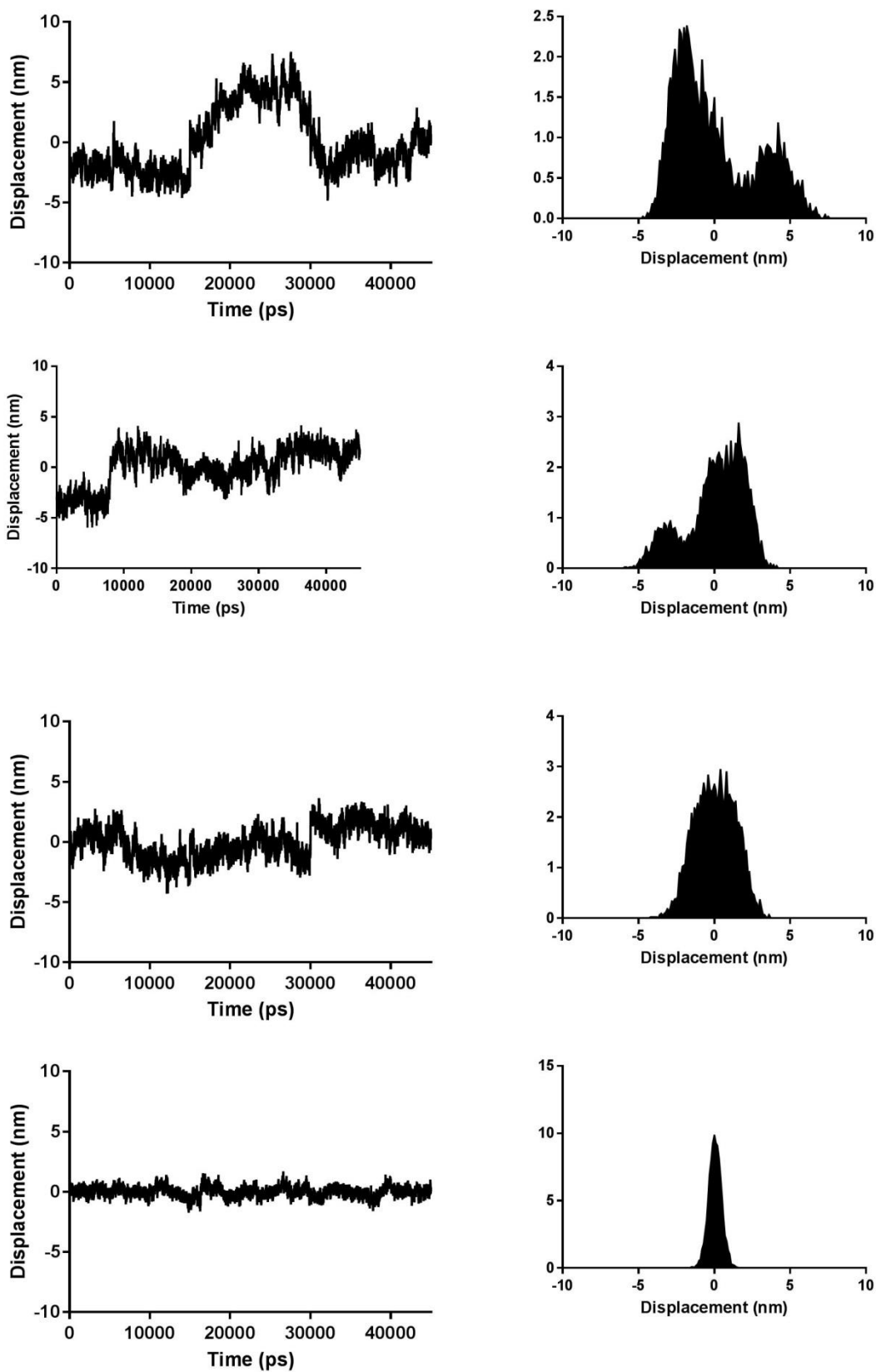
Figure A4.4k: Projection and distribution for COX-2 bound to OLEO of PCA-1, PCA-2, PCA-3, and PCA-20.

178
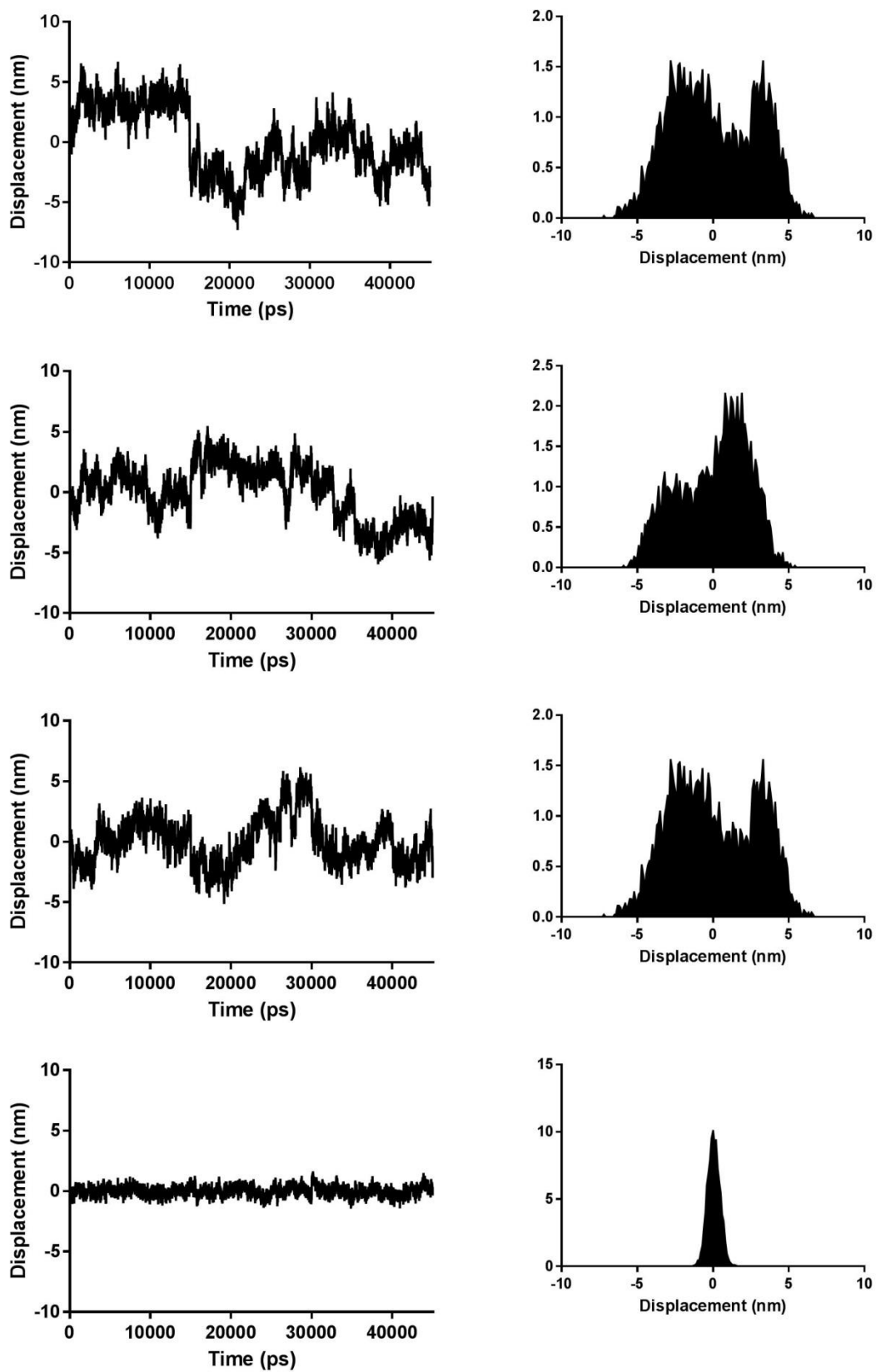
Figure A4.4l: Projection and distribution for COX-2 bound to MMHTE of PCA-1, PCA-2, PCA-3, and PCA-20.

179

COX-1   **COX-2**

E1

E2

E3
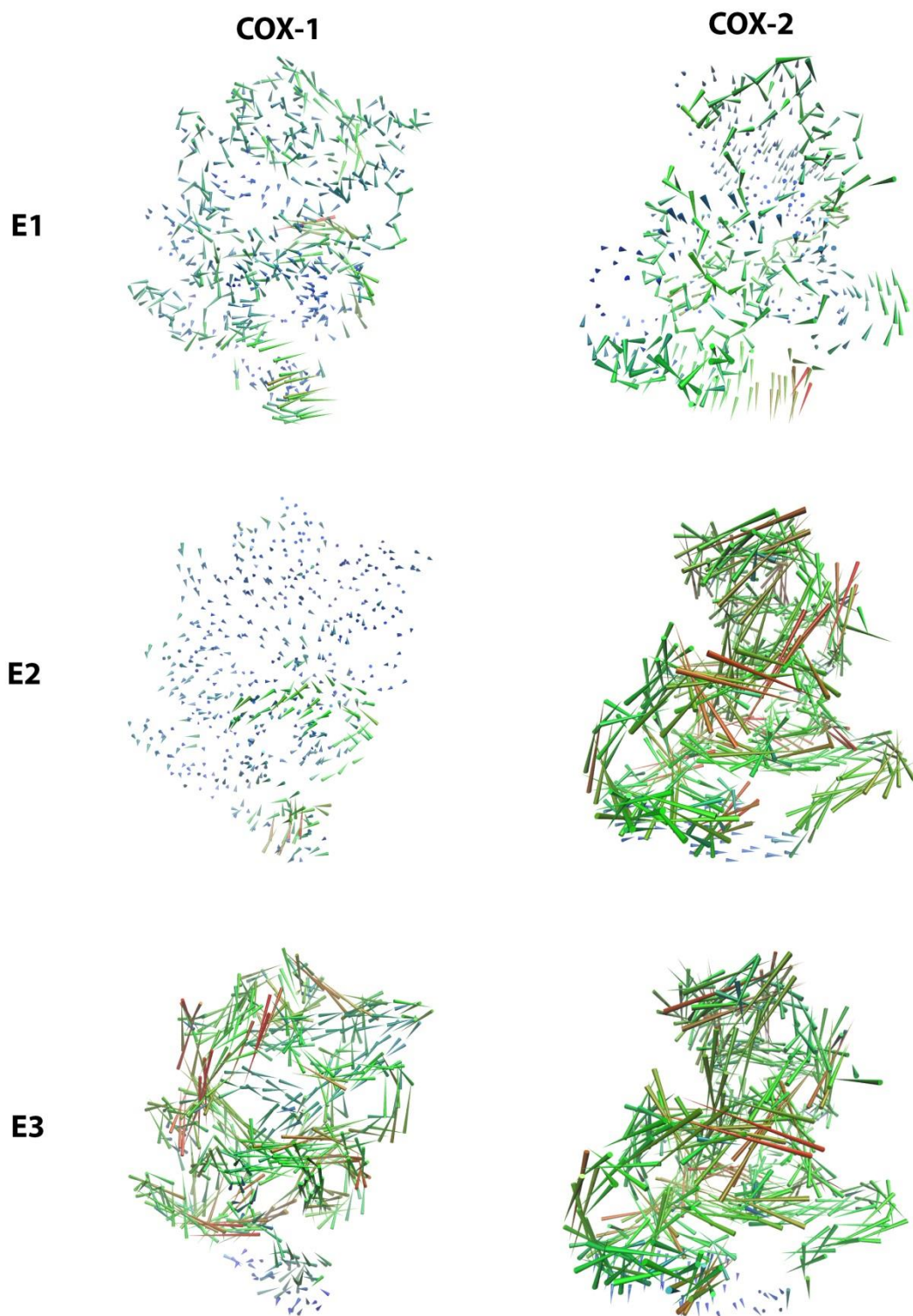
Figure A4.5a: Porcupine plots for the first three eigenvectors of the apo form of COX-1 and COX-2
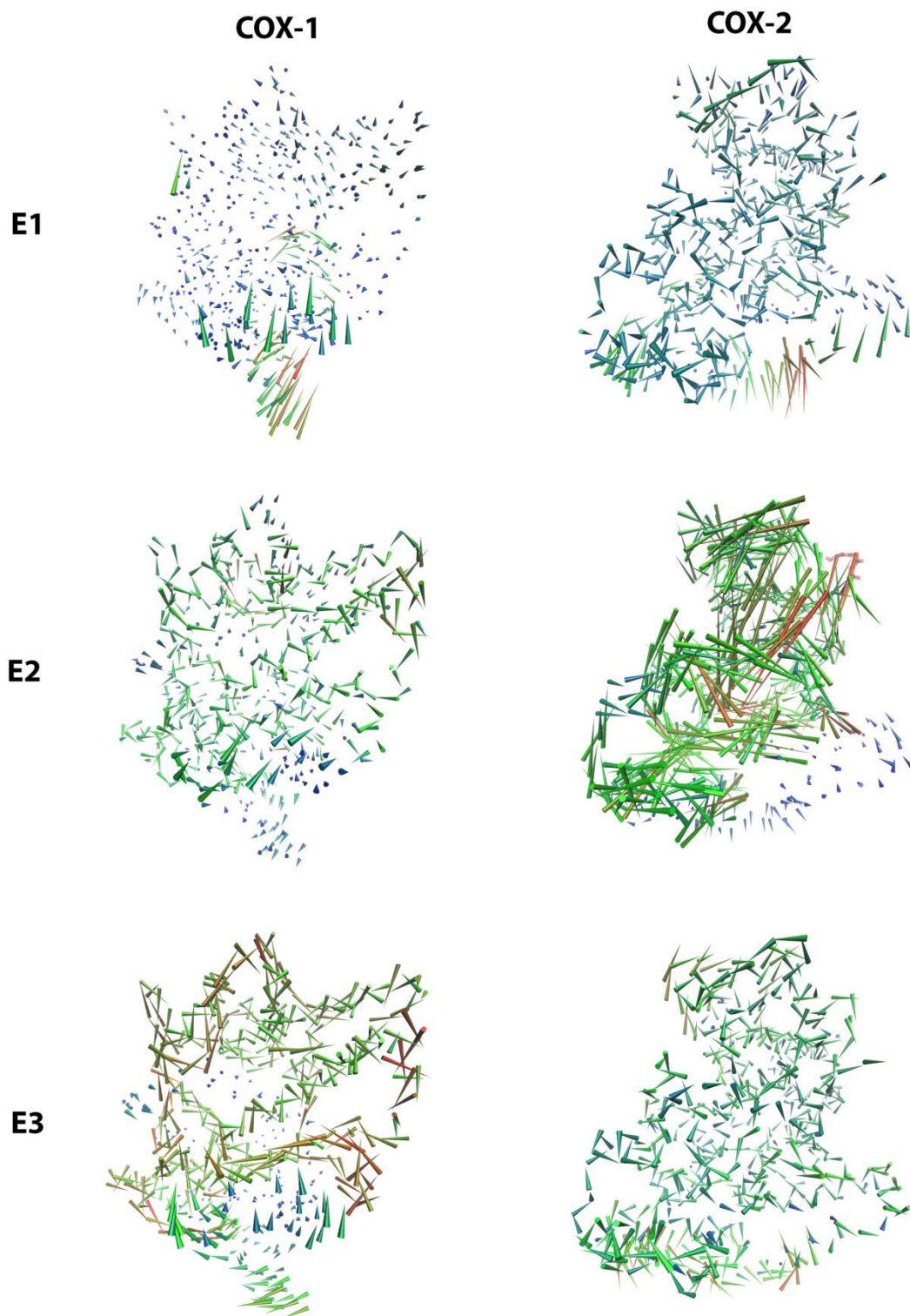
Figure A4.5b: Porcupine plots for the first three eigenvectors of COX-1 and COX-2 bound to their native ligands (FLP and SAL respectively)
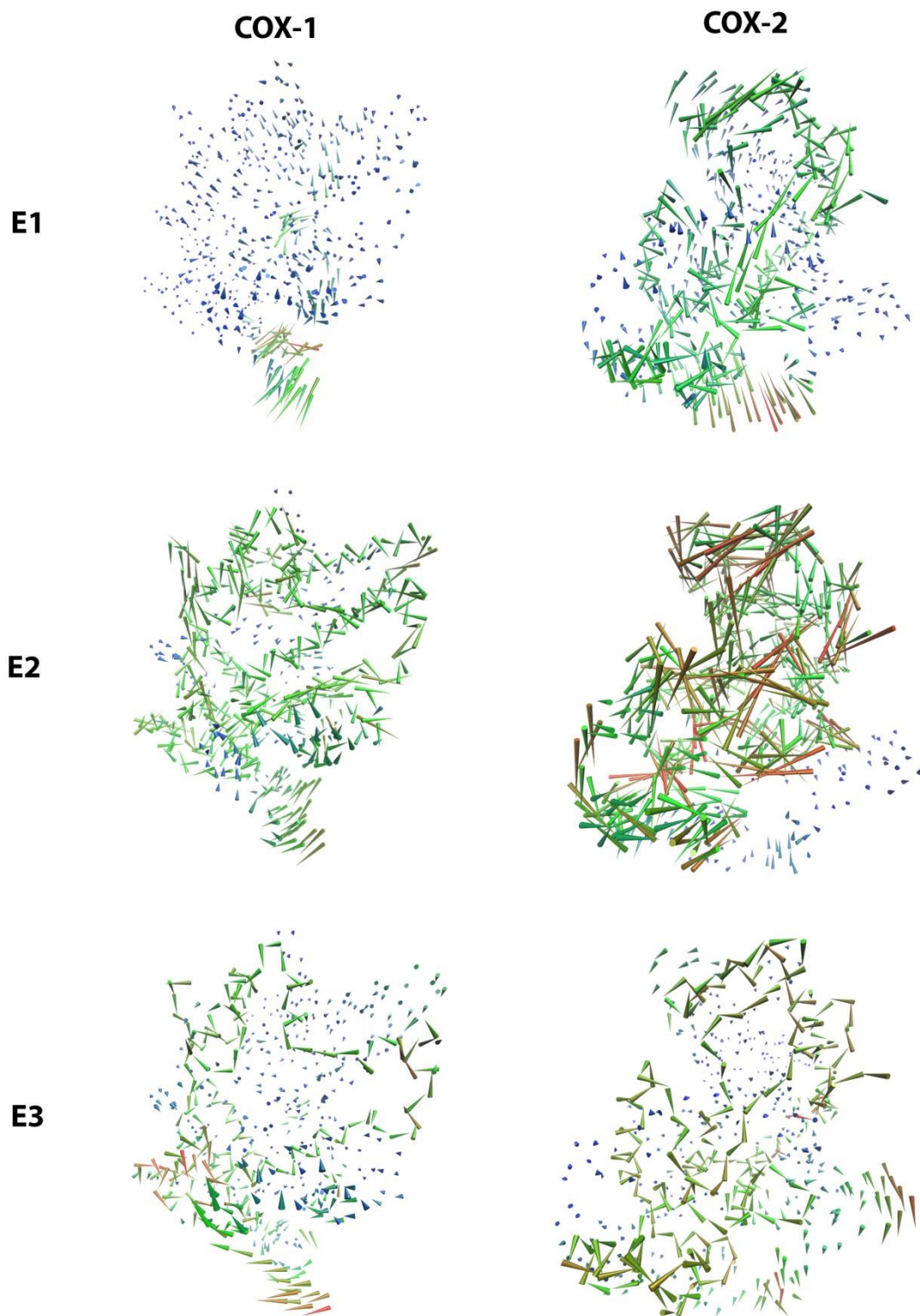
Figure A4.5c: Porcupine plots for the first three eigenvectors of COX-1 and COX-2 bound to 1OL

COX-1   **COX-2**

E1

E2

E3
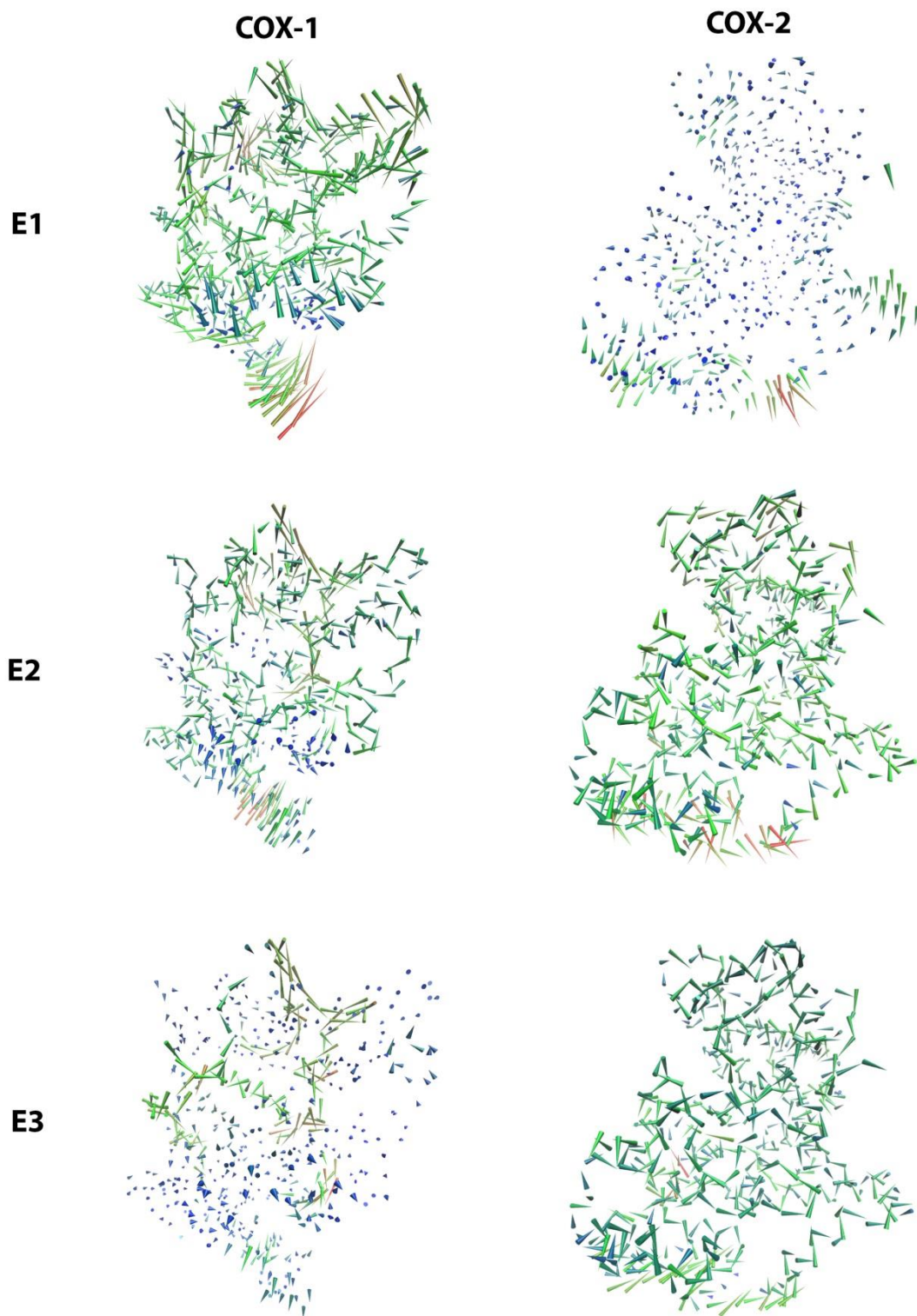
Figure A4.5d: Porcupine plots for the first three eigenvectors of COX-1 and COX-2 bound to LG2
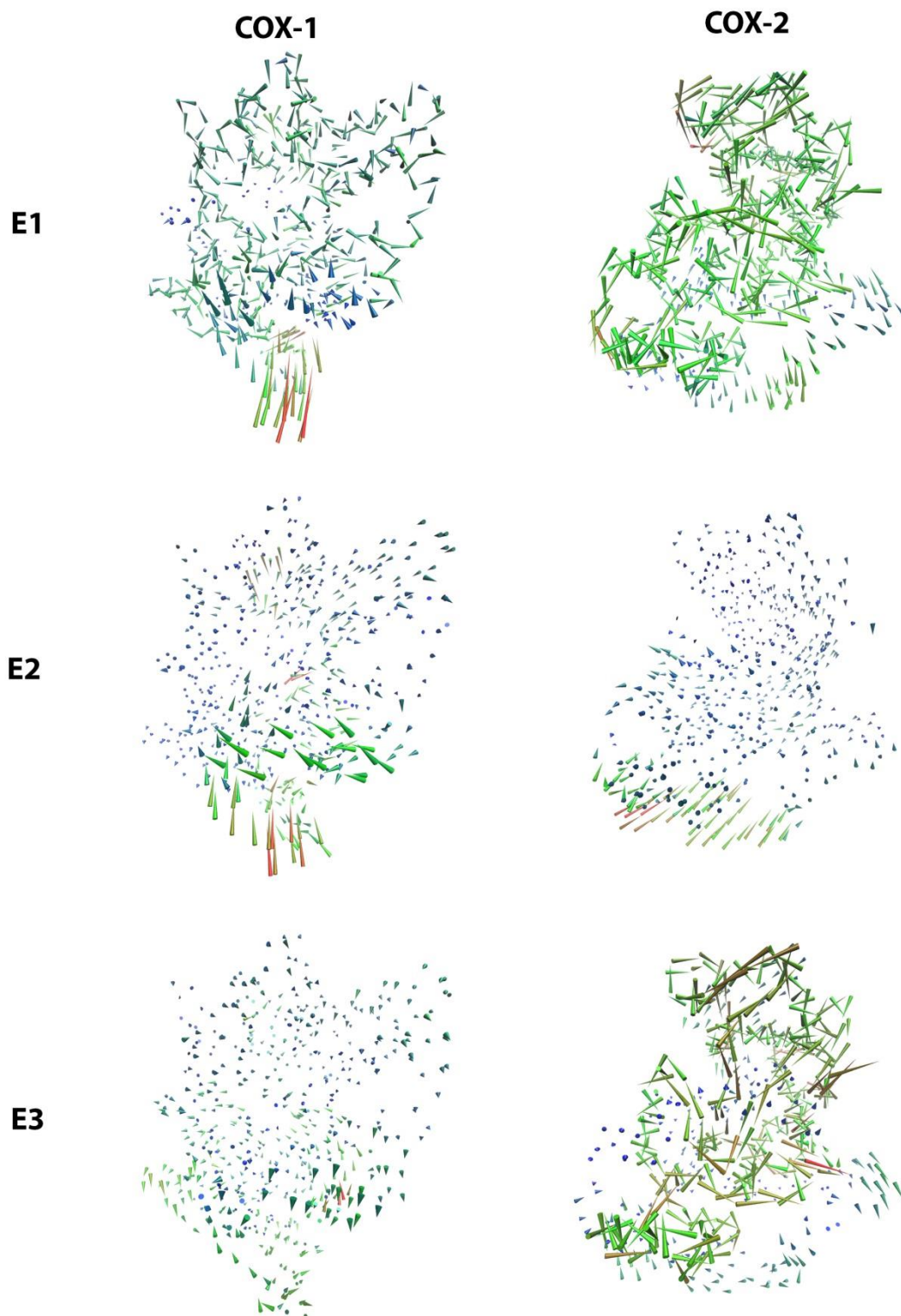
Figure A4.5e: Porcupine plots for the first three eigenvectors of COX-1 and COX-2 bound to OLEO
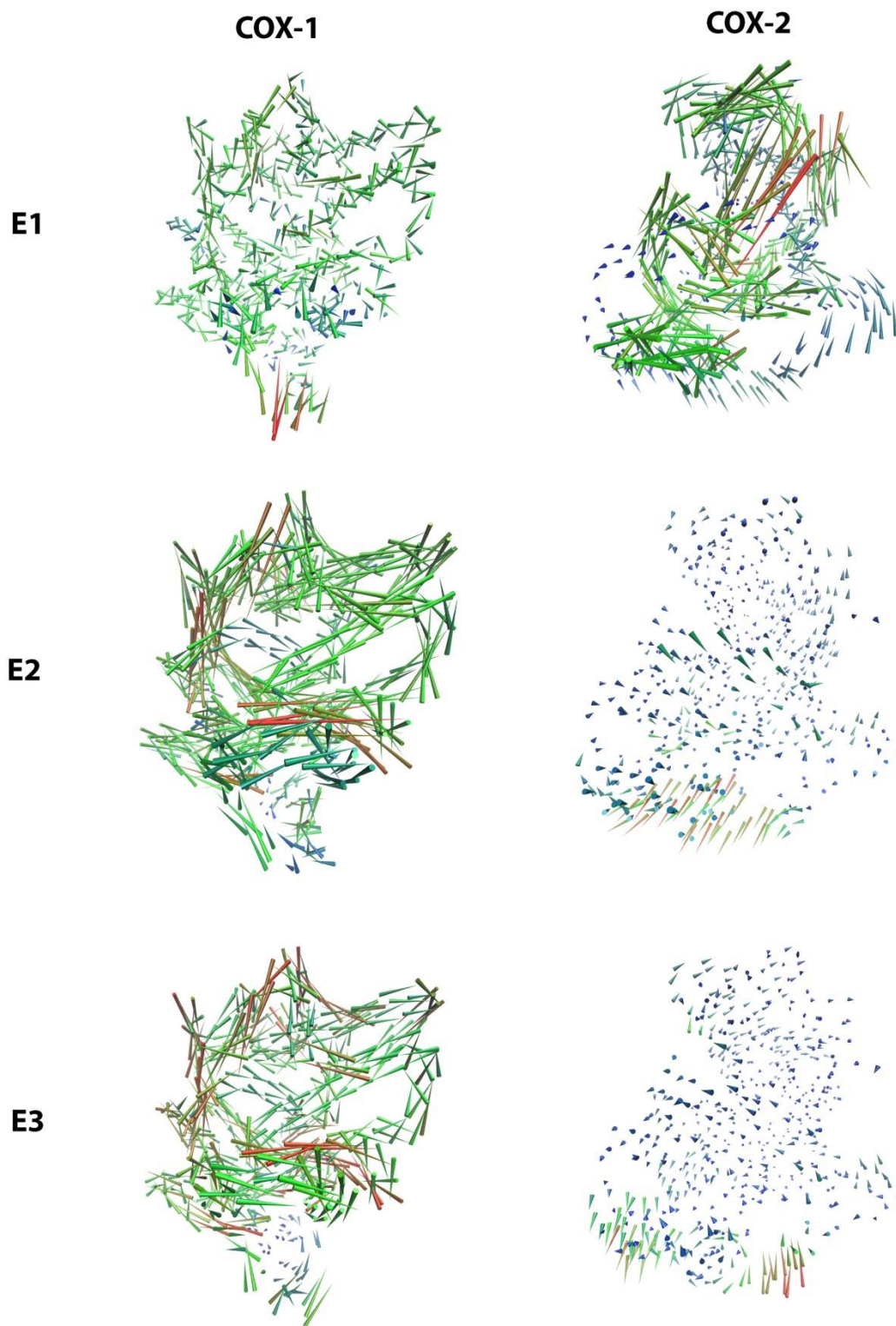
Figure A4.5f: Porcupine plots for the first three eigenvectors of COX-1 and COX-2 bound to MMHTE

Table A4.3: Blocks of residues and segments of concerted motion as identified with principal components analysis of essential dynamics of proteins

Residues		Segment		
		A	**B**	**C**
Block 1	**32-85**	32-47	48-69	70-85
Block 2	**86-115**			
Block 3	**126-171**	126-133	134-148	149-171
Block 4	**209-229**			
Block 5	**235-290**	233-264	265-290	

Table A4.4a: Concerted movements of blocks of residues in COX-1 bound to various ligands

COX-1		APO	FLP	1OL	LG2	OLEO	MMHTE
Block 1	E1	1C=ACW	L-UP-ACW	1B=U-CW; 1C=D-CW	1B+C=L-UP-ACW	1A+B=L-ACW; 1C=L-UP	1C=L-UP
	E2	1B=UP; 1C=M-ACW	1A=S; 1C=S	M-ACW	1A+B=L-UP-CW	1A+B=M-ACW; 1C=UP	1A=S
	E3	1A=S	M-CW	1C=L-CW	1B=M-UP-ACW	M-CW	1A+B=S-UP-ACW; 1C=M-UP
Block 2	E1		M			L-UP	
	E2	M-CW		S-UP		M-CW	
	E3		M-UP	M-UP		M-DN	
Block 3	E1	3B=CW	3B=ACW	3A+B=M-DN	3B=M-DN	3B=M-UP	
	E2	3B=M-UP	3C=S		3B=M-UP	3B=M-UP	
	E3				3B=M-DN-CW	3A+B=M-UP	
Block 4	E1	ACW	ACW	M-UP			
	E2					M-UP-CW	
	E3					M-ACW	
Block 5	E1	M-DN	DN	S-UP			
	E2		5A=S			5A=S-DN-ACW; 5B=M-DN	
	E3					S-UP	

Note: Magnitude of motion is denoted by S, M, and L for small, moderate, and large respectively. Upward and downward movement is denoted by UP and DN. CW denotes clockwise motion, and ACW denotes anticlockwise motion.

Table A4.4b: Concerted movements of blocks of residues in COX-2 bound to various ligands

COX-2		APO	SAL	1OL	LG2	OLEO	MMHTE
Block 1	**E1**	DN-CW	CW	ACW	M-CW	M-ACW	M-CW
	E2	S-CW	S-CW	S	CW	M-CW	L-ACW
	E3	S-DN	L-DN	S-ACW	M-CW	S-ACW	L-ACW
Block 2	**E1**	M-UP	S-UP		M-ACW	M-ACW	M-UP
	E2		S-UP	S			
	E3			M-DN		M-CW	S-UP
Block 3	**E1**	3B=M-UP		S	3A=S-UP; 3B=M-CW; 3C=M-UP	S-DN	S
	E2					3A+B=S-ACW; 3C=M-UP	M
	E3	3C=M-ACW				S-UP	S-DN
Block 4	**E1**				S-ACW	S-ACW	
	E2						S-UP
	E3					S-ACW	
Block 5	**E1**	5A=S		S-DN	5A=S-DN-ACW; 5B=M-DN-ACW		5A=S-UP
	E2					5A=S-DN	S
	E3	5B=L				M-DN	S-UP

Note: Magnitude of motion is denoted by S, M, and L for small, moderate, and large respectively. Upward and downward movement is denoted by UP and DN. CW denotes clockwise motion, and ACW denotes anticlockwise motion.

9.4.4   Modular network analysis

9.4.4.1   Metrics for overall graphs

Table A4.5a: Overall metric values for COX-1 network analysis

	APO	**FLP**	**1OL**	**LG2**	**OLEO**	**MMHTE**
Nodes	552	552	552	552	552	552
Edges	8419	9293	6874	7248	6420	7756
Average Degree	26.011	29.054	20.667	21.812	19.022	23.547
Avg. Weighted Degree	39.489	49.902	33.384	35.159	31.739	37.21
Network Diameter	13	11	15	13	13	10
Graph Density	0.047	0.053	0.038	0.04	0.035	0.043
Modularity	0.637	0.59	0.717	0.711	0.68	0.676
Connected Components	1	1	1	1	1	1
Avg. Clustering Coefficient	0.613	0.634	0.623	0.628	0.609	0.621
Avg. Path Length	4.366	3.882	5.082	4.924	4.647	4.082

Table A4.5b: Overall metric values for COX-2 network analysis

	APO	SAL	1OL	LG2	OLEO	MMHTE
Nodes	552	552	552	552	552	552
Edges	14165	11271	11062	10795	10531	9488
Average Degree	46.395	36.004	32.246	34.203	34.286	30.217
Avg. Weighted Degree	61.178	50.504	49.746	48.931	45.895	42.696
Network Diameter	11	12	9	12	17	14
Graph Density	0.084	0.65	0.064	0.062	0.062	0.055
Modularity	0.509	0.576	0.597	0.655	0.574	0.59
Connected Components	1	1	1	1	1	1
Avg. Clustering Coefficient	0.626	0.624	0.635	0.637	0.631	0.641
Avg. Path Length	3.174	3.726	3.542	3.79	4.429	4.422
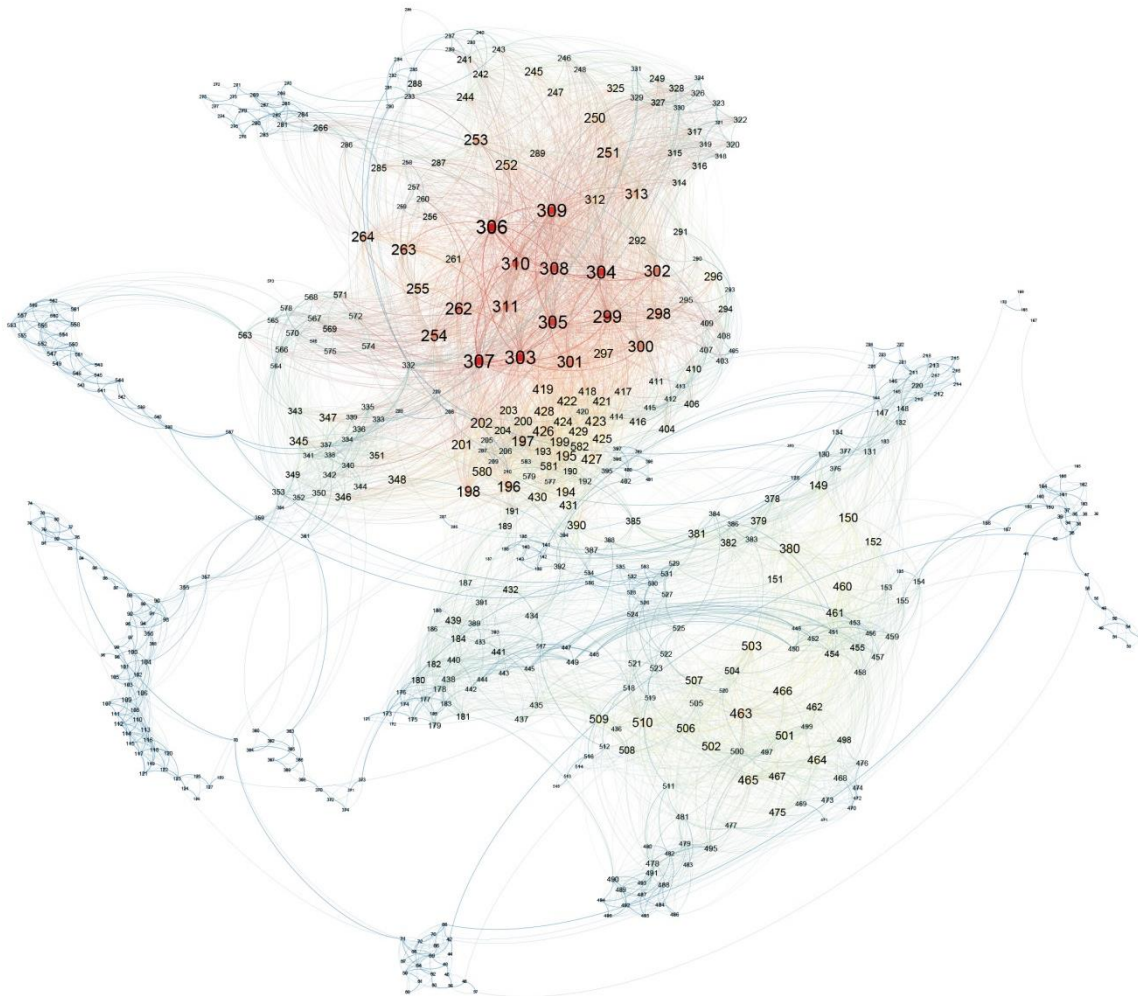
Figure A4.6a: Communication pathway network for COX-1 in its apo form
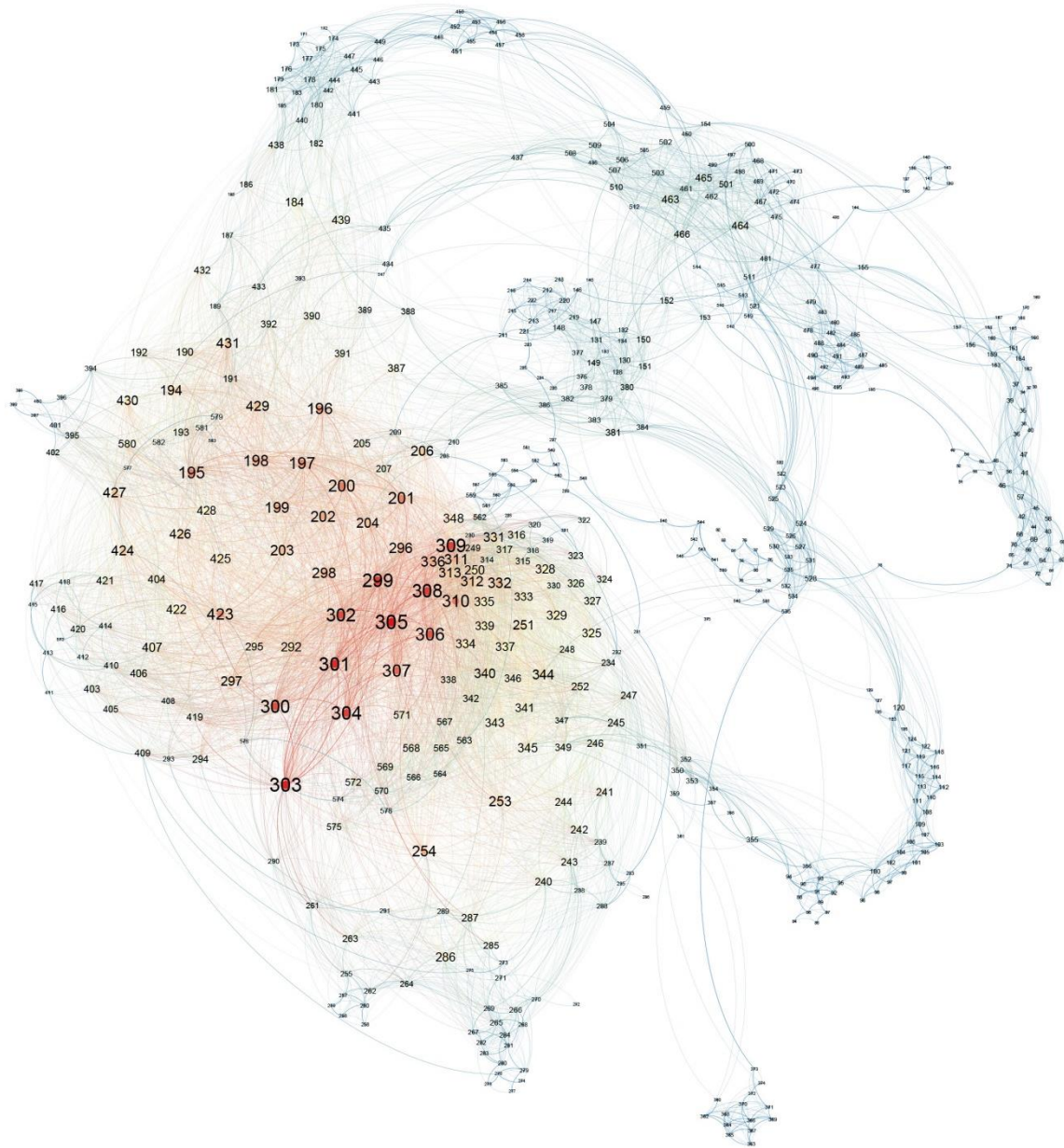
Figure A4.6b: Communication pathway network for COX-1 bound to FLP
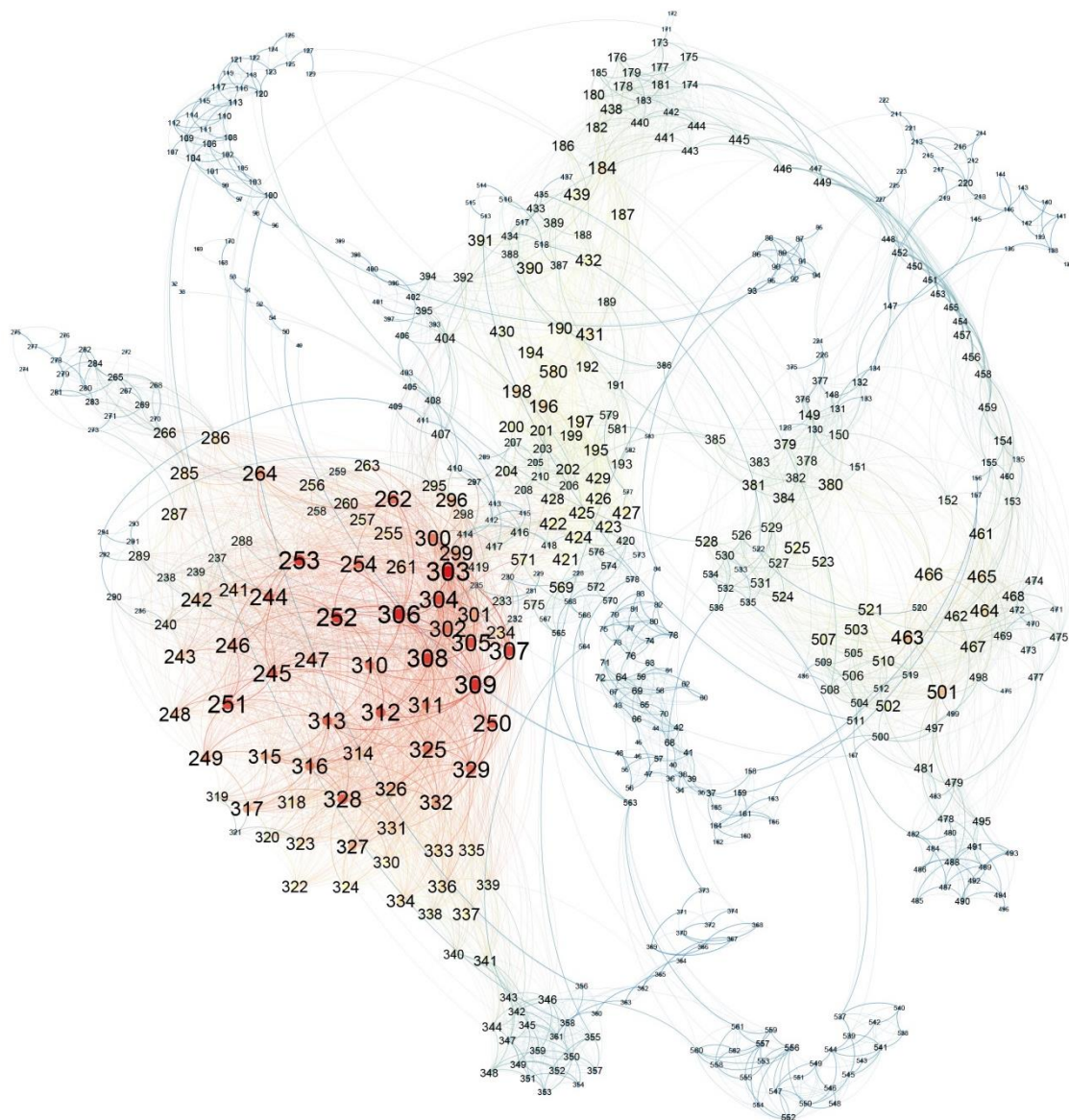
Figure A4.6c: Communication pathway network for COX-1 bound to 1OL
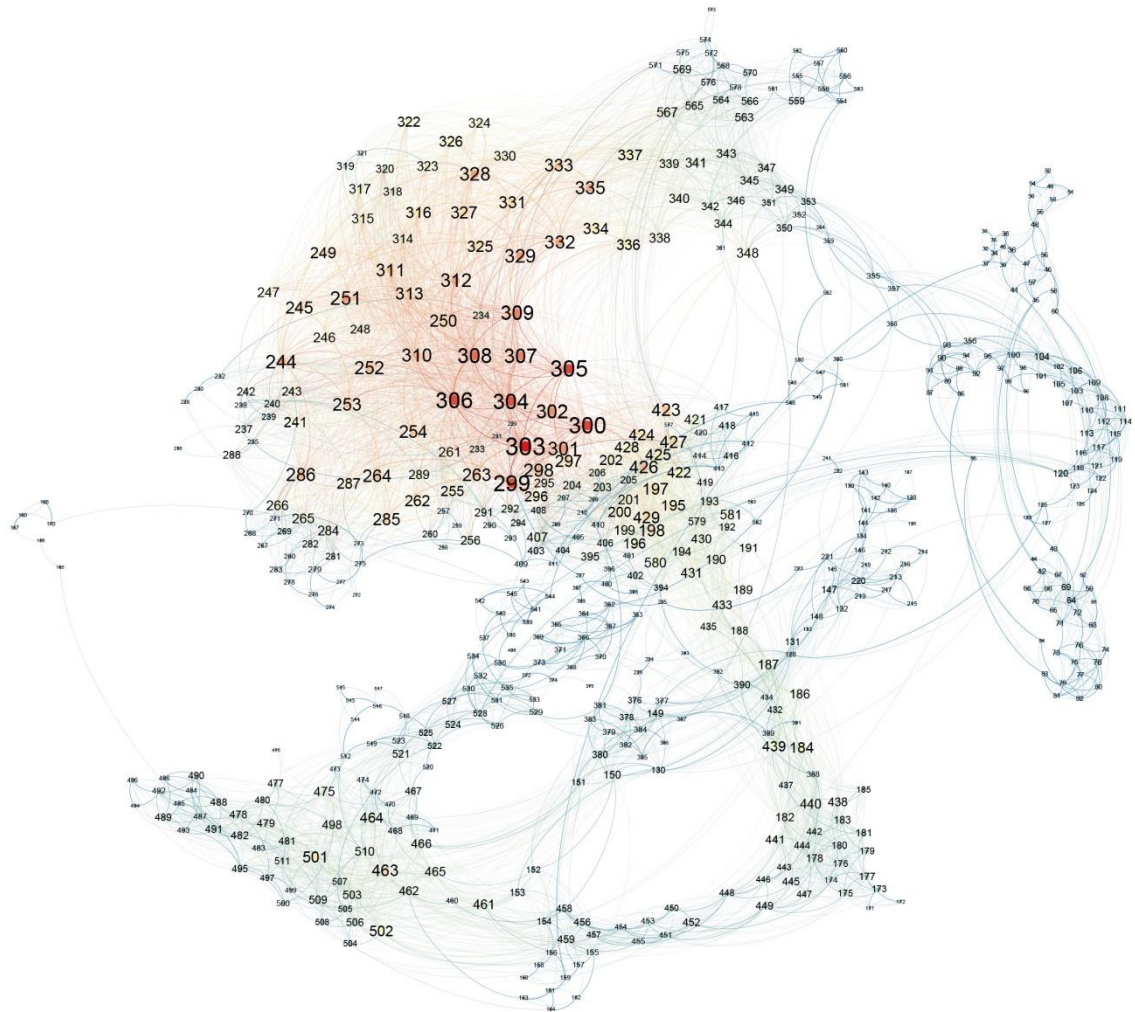
Figure A4.6d: Communication pathway network for COX-1 bound to LG2
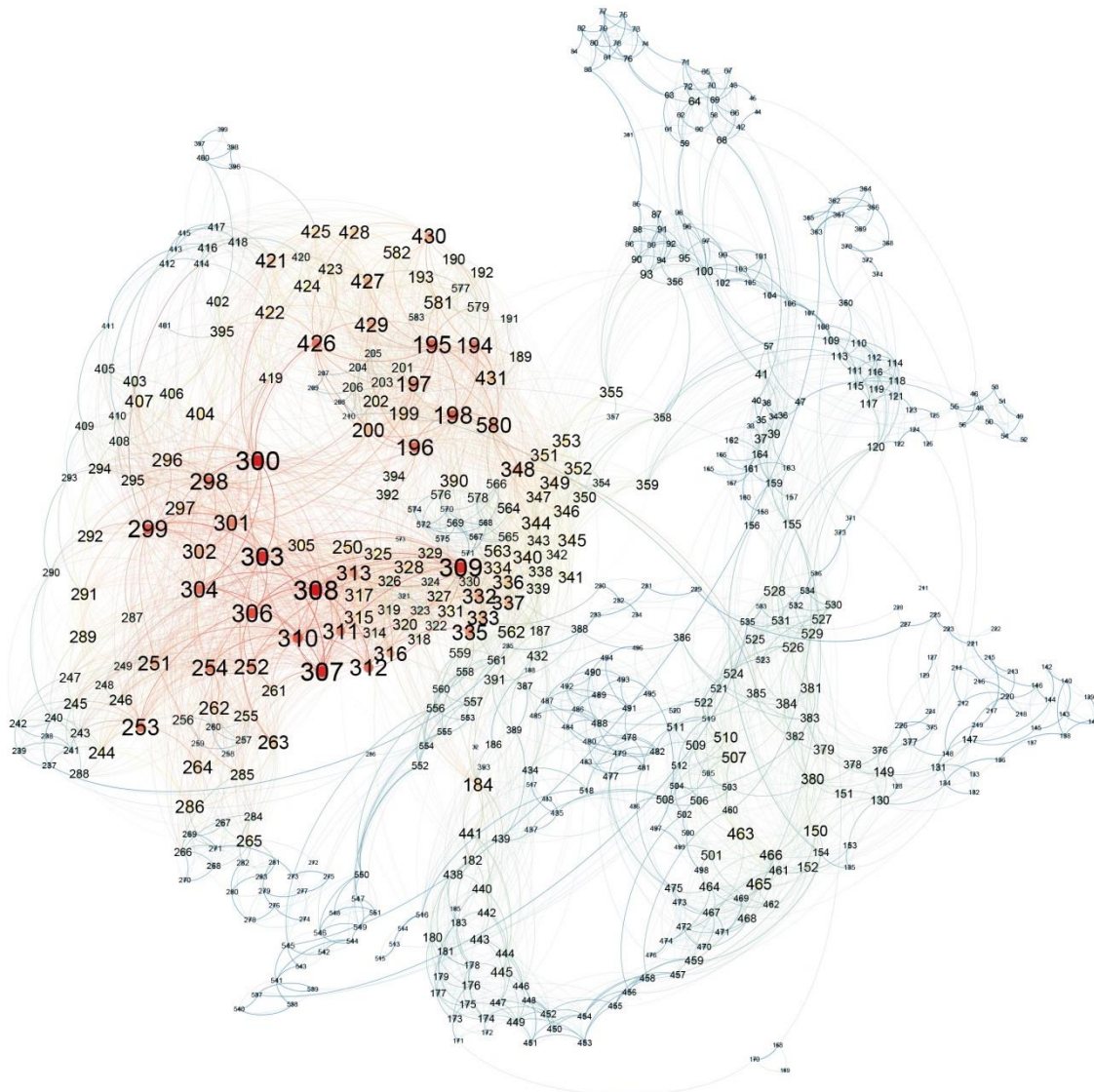
Figure A4.6e: Communication pathway network for COX-1 bound to OLEO
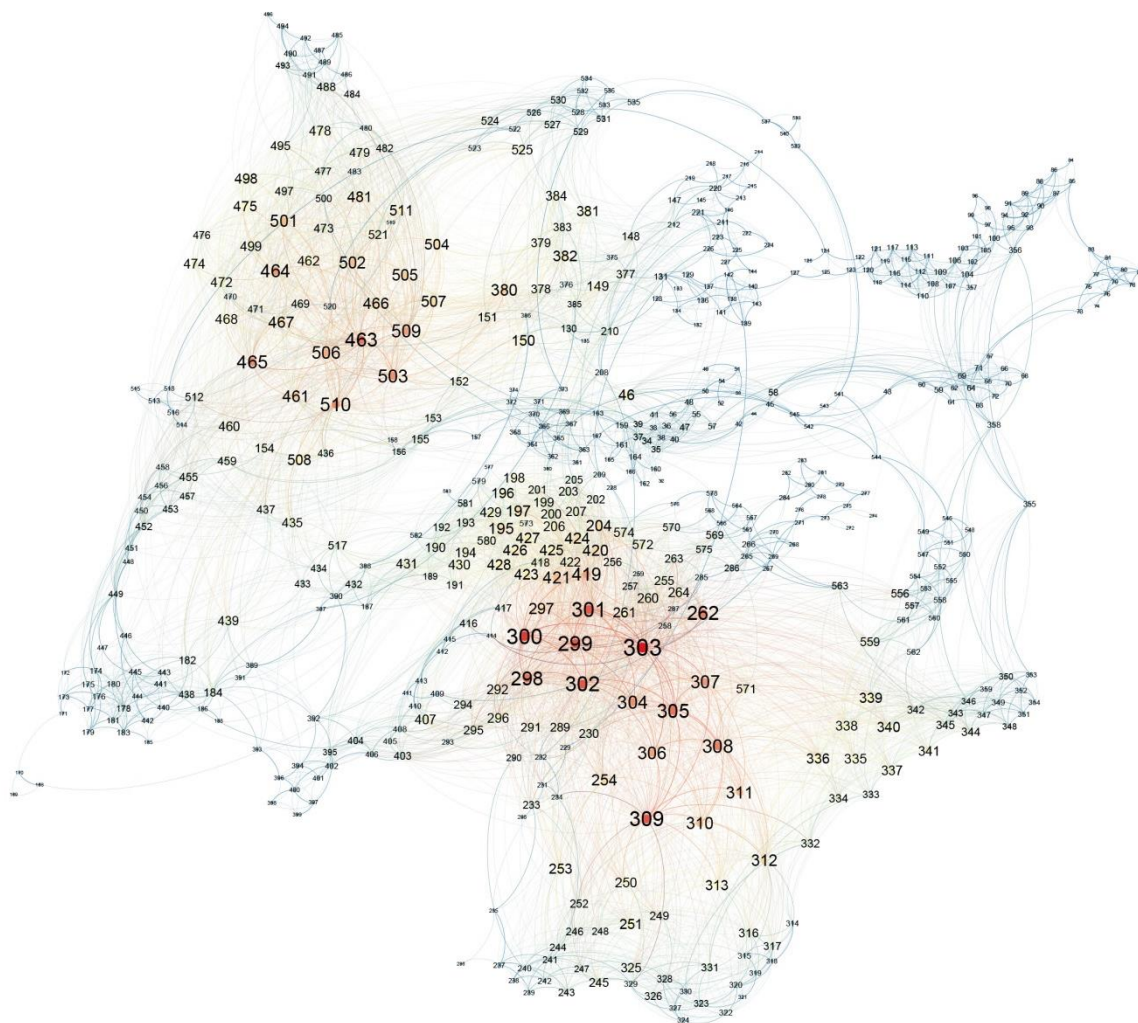
Figure A4.6f: Communication pathway network for COX-1 bound to MMHTE
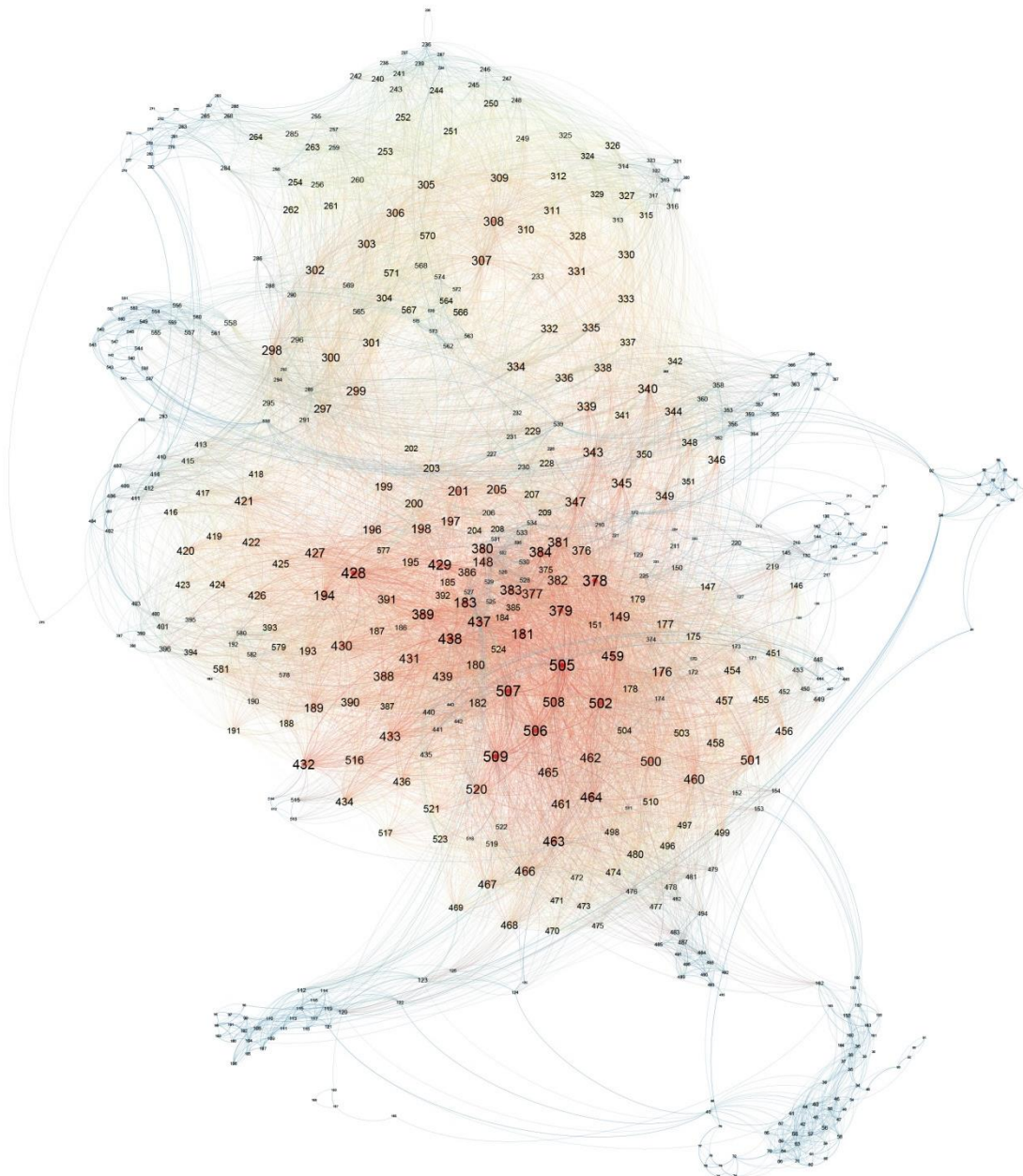
Figure A4.6g: Communication pathway network for COX-2 in its apo form
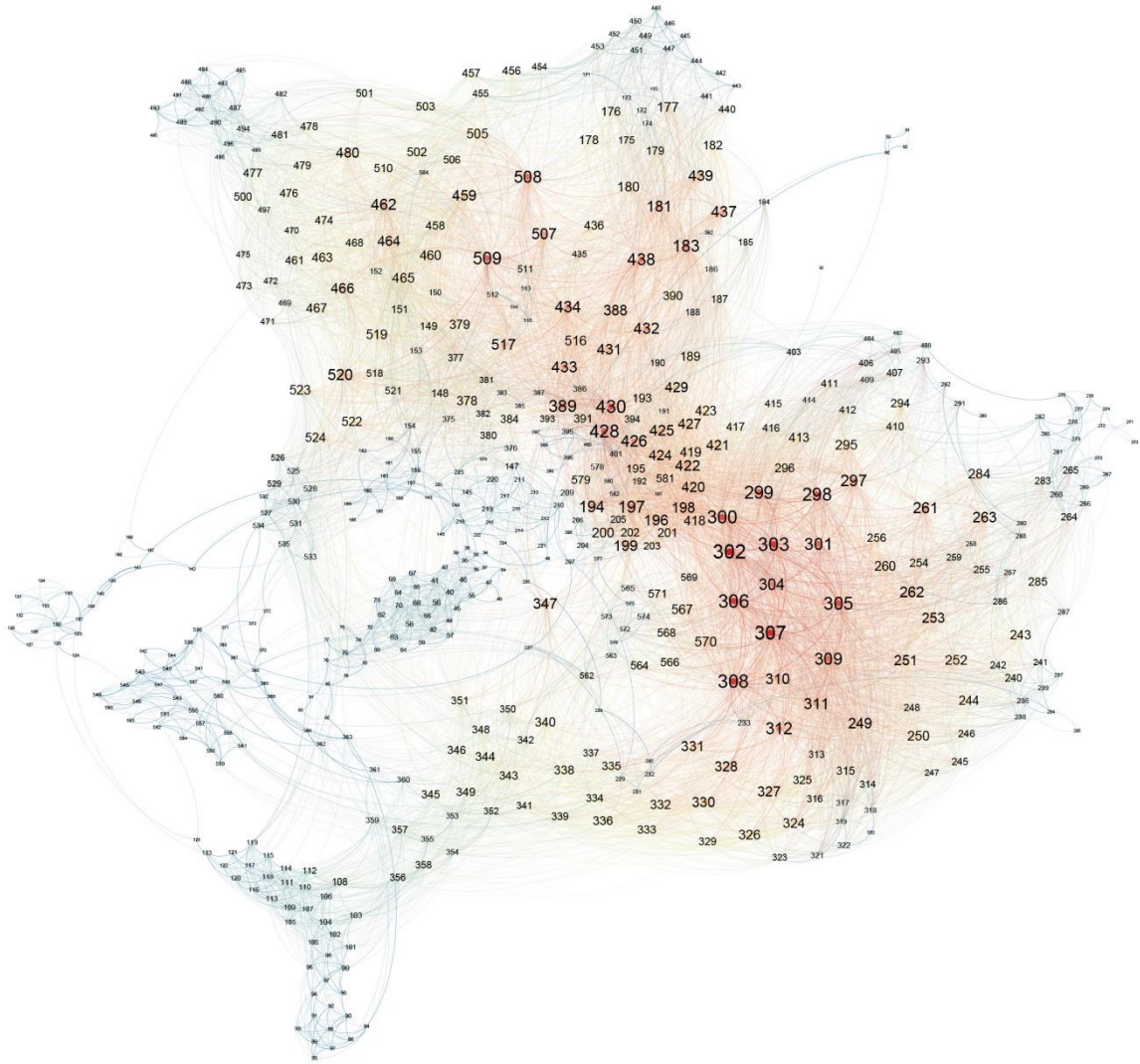
Figure A4.6h: Communication pathway network for COX-2 bound to SAL
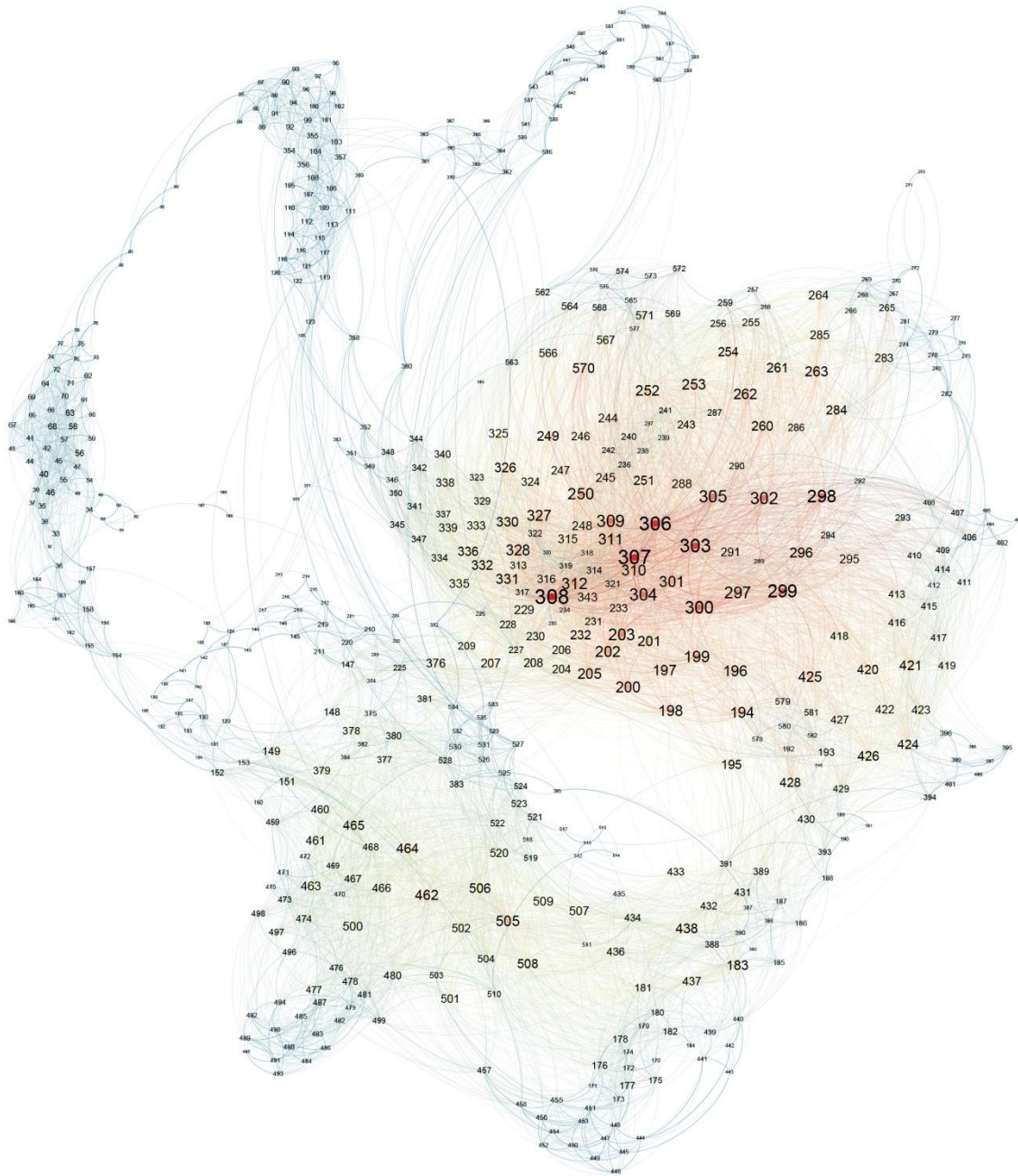
Figure A4.6i: Communication pathway network for COX-2 bound to 1OL
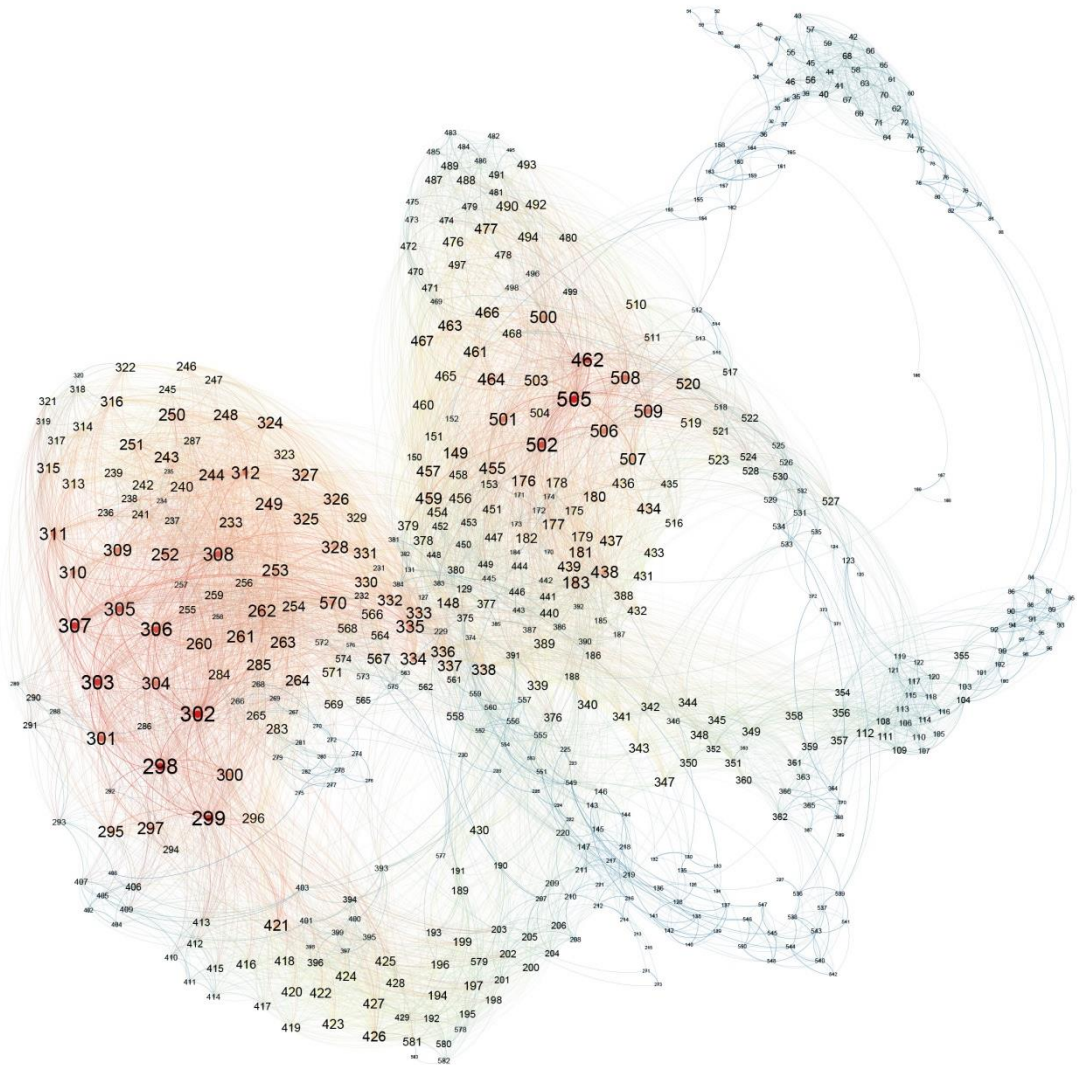
Figure A4.6j: Communication pathway network for COX-2 bound to LG2
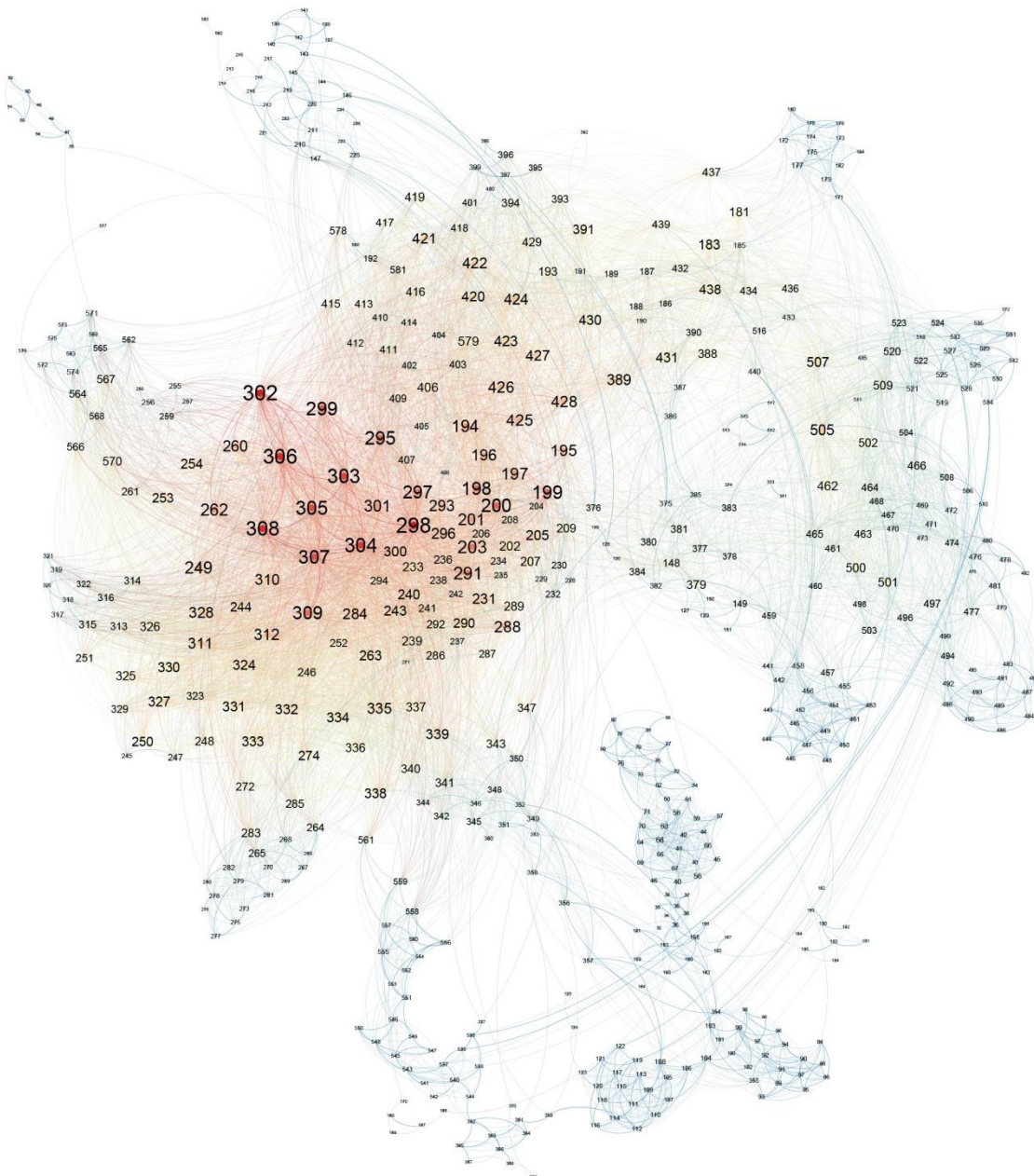
Figure A4.6k: Communication pathway network for COX-2 bound to OLEO

Figure A4.6l: Communication pathway network for COX-2 bound to MMHTE

Table A4.6a: Degrees for COX-1 residues

Residue	APO	FLP	1OL	LG2	OLEO	MMHTE
Active site						
93	12	16	13	14	22	29
116	16	16	11	13	14	32
117	14	17	15	17	20	34
120	12	28	12	23	20	28
205	27	54	17	26	17	48
209	18	29	6	3	4	31
344	35	80	27	31	44	55
345	59	69	19	33	41	54
348	56	72	22	43	54	38
349	40	52	17	33	44	37
352	28	25	18	18	41	28
353	31	30	13	22	42	23
355	18	30	17	18	31	29
359	15	17	17	9	29	21
381	55	46	34	15	28	73
384	24	30	31	16	26	66
385	42	31	29	9	23	37
387	33	58	21	6	15	17
513	5	16	4	4	4	26
518	26	8	15	9	12	21
523	31	17	32	15	12	30
526	14	17	27	15	27	44
527	23	21	27	21	24	44
530	16	19	24	14	17	47
531	28	17	29	11	21	41
533	9	9	11	9	4	27
534	16	17	17	13	13	25
Peroxidase site						
203	54	87	24	31	27	57
207	13	48	21	14	5	67
211	13	14	6	3	2	30
222	4	12	3	3	2	11
240	9	54	31	24	19	39
274	4	3	2	4	4	2
290	14	27	15	16	18	31
388	19	40	20	15	17	18
391	27	51	38	5	27	16
409	31	48	12	17	21	31
504	43	27	22	19	12	71
EGF						

32	1	2	1	3	1	1
35	8	18	6	5	14	38
38	6	9	11	8	11	31
41	7	27	15	14	23	33
44	6	9	4	7	3	6
47	5	24	10	10	16	40
50	5	7	5	5	7	13
53	3	4	3	6	4	6
56	4	11	8	10	4	20
59	8	15	12	13	11	35
62	8	13	8	7	7	19
65	9	11	12	11	9	19
68	10	19	16	10	15	15
71	12	16	15	15	10	29
72	10	13	17	17	11	21
MBD						
73	4	4	11	12	9	15
76	8	5	17	17	11	28
79	8	5	11	10	9	23
82	7	2	10	11	8	16
85	5	3	4	5	7	16
88	11	11	8	9	12	25
91	9	9	8	11	13	31
94	11	12	9	8	12	21
97	7	10	3	9	7	15
100	17	21	12	17	20	38
103	14	15	11	17	9	27
106	20	12	14	19	8	33
109	20	21	14	17	18	46
Glycosylation						
68	10	19	16	10	15	15
144	6	5	4	9	6	11
410	41	42	15	16	18	31
580	64	60	46	43	56	47

Table A4.6b: Betweenness centrality for COX-1 residues

Residue	APO	FLP	1OL	LG2	OLEO	MMHTE
Active site						
93	1691.14	527.37	10464.89	1828.82	4122.11	653.20
116	267.45	31.62	190.89	71.89	111.19	156.63
117	442.33	64.01	933.83	1315.79	4002.58	1148.96
120	41.59	5605.79	9884.64	5373.77	2083.04	559.68
205	11.42	558.26	3.10	66.80	44.81	430.01
209	139.85	78.44	0.00	21.57	1.98	289.17
344	474.96	5473.86	7892.39	3452.91	1978.99	5112.05
345	5338.44	1725.01	2091.80	1436.74	636.71	613.33
348	25449.16	5738.12	1038.46	17655.66	17907.25	79.78
349	1711.07	1139.31	184.29	2305.49	980.35	28.49
352	550.11	738.57	1364.42	385.79	764.30	62.97
353	2869.09	4653.96	94.54	2332.54	2916.64	201.21
355	16768.98	4668.21	5622.27	12522.16	5788.20	3408.19
359	975.25	873.30	360.51	406.16	6720.45	2632.10
381	8759.08	12358.13	2754.20	11639.14	2380.54	4350.87
384	384.75	1478.96	889.60	6495.91	1556.87	1539.13
385	2985.97	1201.43	24289.51	2364.54	27736.47	442.33
387	624.98	3931.77	7755.10	9565.03	258.43	47.18
513	4.00	377.70	136.48	11.34	152.88	56.84
518	305.48	182.83	1459.28	704.58	1286.00	84.00
523	134.69	357.44	563.67	335.44	114.70	12.38
526	45.42	234.09	1622.80	1499.63	2835.10	388.78
527	634.36	539.86	250.66	2376.69	1057.65	758.61
530	107.25	1803.01	210.54	368.53	3073.69	567.61
531	1006.77	848.08	301.86	2.53	433.18	505.80
533	0.93	0.00	0.19	0.21	1.25	6.72
534	3365.66	3610.19	1127.31	399.04	460.11	1241.52
Peroxidase site						
203	309.76	2161.55	30.92	454.02	341.20	842.27
207	121.08	1183.38	18.97	328.98	8.07	15836.27
211	1075.87	1896.50	2318.26	1117.50	1351.22	1437.79
222	25.89	103.25	82.88	203.17	0.00	0.50
240	1.12	243.78	75.43	8.05	31.44	62.06
274	4.39	1.95	236.56	2.77	57.47	0.00
290	1.76	50.29	56.17	304.58	47.38	546.47
388	155.89	4696.74	145.67	367.70	4225.33	511.73
391	216.39	1163.33	1393.72	0.45	1292.13	71.84
409	1021.04	1104.95	144.05	377.22	47.71	2288.85
504	224.75	448.72	50.23	168.04	35.76	453.54
EGF						
32	0.00	0.00	0.00	2.36	0.00	0.00
35	409.54	65.38	115.73	227.69	49.66	102.00

38	25.00	1.20	377.13	404.28	18.97	8.49
41	8896.76	987.33	5757.36	13377.56	7104.36	111.99
44	257.38	18.80	7.74	26.55	13.92	0.00
47	6254.52	724.35	462.48	94.43	2669.77	113.75
50	544.17	560.83	730.92	0.17	492.71	344.22
53	0.42	1.80	0.58	0.37	408.09	0.00
56	591.41	11.10	23.61	84.03	2.90	1.59
59	454.61	27.22	52.24	502.01	832.91	660.58
62	276.41	89.26	3.55	1.51	3.98	25.43
65	2.80	13.68	53.98	2.20	1.10	2.50
68	1919.10	218.01	1391.09	0.81	1584.01	1.22
71	2012.01	4872.92	206.46	201.30	251.40	3417.35
72	215.35	309.01	1729.70	1905.66	153.22	15.61
MBD						
73	124.23	312.68	98.59	638.58	431.18	328.50
76	620.37	1.50	1612.16	1870.18	1708.37	2049.94
79	620.37	7.47	701.44	38.03	91.84	23.57
82	3364.83	0.00	577.61	5215.62	825.28	176.95
85	717.89	17.70	244.59	1547.39	458.41	787.32
88	6387.68	318.43	0.00	213.40	197.97	527.77
91	38.68	0.00	0.00	2013.51	204.90	1478.52
94	212.59	10.31	1.27	6.60	13.78	43.74
97	82.67	5.35	11.99	211.07	8.03	8.91
100	1438.18	1902.17	2571.01	1391.00	1380.09	548.26
103	62.27	202.55	1742.05	249.79	14.15	345.14
106	5288.81	64.89	1643.38	2457.57	92.49	847.52
109	386.59	789.37	2423.43	1279.17	968.82	723.82
Glycosylation						
68	1919.10	218.01	1391.09	0.81	1584.01	1.22
144	24.47	990.35	325.63	900.00	1378.58	38.03
410	77.08	70.46	141.10	23.53	54.02	99.64
580	1118.76	339.94	921.10	4141.76	1851.93	87.96

Table A4.6c: Eigencentrality for COX-1 residues

Residue	APO	FLP	1OL	LG2	OLEO	MMHTE
Active site						
93	0.0074	0.0106	0.0064	0.0102	0.1047	0.0077
116	0.0070	0.0069	0.0051	0.0090	0.0204	0.0089
117	0.0061	0.0072	0.0069	0.0122	0.0309	0.0101
120	0.0050	0.0159	0.0073	0.0160	0.0303	0.0081
205	0.2344	0.4093	0.0597	0.1936	0.1823	0.3464
209	0.1381	0.2510	0.0158	0.0033	0.0088	0.1149
344	0.2105	0.6083	0.1588	0.2156	0.4455	0.3019
345	0.4181	0.5265	0.0467	0.1850	0.4236	0.3350
348	0.3437	0.5580	0.0792	0.2541	0.5453	0.1250
349	0.2036	0.3985	0.0302	0.1599	0.4345	0.1009
352	0.1300	0.1102	0.0248	0.0603	0.4179	0.0569
353	0.1356	0.1577	0.0174	0.0768	0.4047	0.0274
355	0.0322	0.0946	0.0196	0.0327	0.2089	0.0438
359	0.0442	0.0764	0.0336	0.0256	0.1681	0.0695
381	0.1771	0.1100	0.0581	0.0158	0.0583	0.2928
384	0.0708	0.0465	0.0571	0.0157	0.0557	0.2485
385	0.1403	0.0977	0.0618	0.0085	0.0635	0.1163
387	0.1160	0.3321	0.0520	0.0094	0.0669	0.0230
513	0.0047	0.0102	0.0035	0.0016	0.0025	0.0760
518	0.0625	0.0026	0.0263	0.0084	0.0212	0.0250
523	0.0788	0.0140	0.0710	0.0251	0.0266	0.1488
526	0.0199	0.0100	0.0540	0.0160	0.0493	0.1212
527	0.0477	0.0142	0.0587	0.0233	0.0441	0.1734
530	0.0253	0.0150	0.0498	0.0121	0.0307	0.1478
531	0.0609	0.0220	0.0599	0.0097	0.0399	0.1079
533	0.0090	0.0054	0.0210	0.0078	0.0051	0.0359
534	0.0398	0.0283	0.0300	0.0103	0.0199	0.0363
Peroxidase site						
203	0.5101	0.6708	0.0791	0.2842	0.3083	0.4231
207	0.0832	0.3792	0.0684	0.0838	0.0213	0.4455
211	0.0109	0.0370	0.0036	0.0014	0.0018	0.0497
222	0.0012	0.0070	0.0009	0.0012	0.0007	0.0053
240	0.0469	0.3825	0.4097	0.2680	0.1526	0.1133
274	0.0012	0.0016	0.0008	0.0039	0.0018	0.0006
290	0.1686	0.2357	0.1683	0.1650	0.2349	0.3103
388	0.0687	0.1603	0.0469	0.0319	0.0800	0.0582
391	0.0900	0.3657	0.0953	0.0099	0.1858	0.0453
409	0.2910	0.3992	0.0349	0.0783	0.2474	0.1597
504	0.1247	0.0239	0.0504	0.0392	0.0196	0.4359
EGF						
32	0.0002	0.0004	0.0002	0.0012	0.0005	0.0009
35	0.0024	0.0075	0.0029	0.0020	0.0144	0.0351

38	0.0018	0.0042	0.0054	0.0034	0.0112	0.0271
41	0.0023	0.0119	0.0084	0.0074	0.0224	0.0318
44	0.0020	0.0040	0.0021	0.0043	0.0015	0.0026
47	0.0013	0.0109	0.0041	0.0047	0.0156	0.0359
50	0.0012	0.0023	0.0014	0.0019	0.0031	0.0094
53	0.0007	0.0008	0.0008	0.0023	0.0018	0.0014
56	0.0009	0.0050	0.0034	0.0047	0.0024	0.0187
59	0.0024	0.0064	0.0067	0.0074	0.0099	0.0243
62	0.0025	0.0051	0.0046	0.0040	0.0049	0.0092
65	0.0032	0.0044	0.0074	0.0068	0.0067	0.0070
68	0.0034	0.0078	0.0092	0.0062	0.0127	0.0063
71	0.0039	0.0064	0.0084	0.0088	0.0067	0.0178
72	0.0034	0.0049	0.0100	0.0101	0.0077	0.0081
MBD						
73	0.0012	0.0011	0.0059	0.0069	0.0050	0.0031
76	0.0021	0.0010	0.0093	0.0098	0.0065	0.0045
79	0.0021	0.0011	0.0054	0.0052	0.0047	0.0034
82	0.0018	0.0004	0.0049	0.0054	0.0041	0.0030
85	0.0013	0.0006	0.0017	0.0022	0.0057	0.0032
88	0.0043	0.0050	0.0035	0.0048	0.0185	0.0057
91	0.0041	0.0046	0.0035	0.0067	0.0199	0.0082
94	0.0048	0.0100	0.0038	0.0047	0.0194	0.0057
97	0.0026	0.0034	0.0008	0.0055	0.0063	0.0029
100	0.0075	0.0108	0.0059	0.0112	0.0334	0.0085
103	0.0063	0.0054	0.0057	0.0122	0.0077	0.0093
106	0.0098	0.0046	0.0073	0.0142	0.0065	0.0144
109	0.0090	0.0083	0.0072	0.0127	0.0246	0.0138
Glycosylation						
68	0.0034	0.0078	0.0092	0.0062	0.0127	0.0063
144	0.0037	0.0013	0.0013	0.0044	0.0033	0.0028
410	0.3936	0.3520	0.0577	0.0815	0.2129	0.1225
580	0.5569	0.4832	0.1548	0.2576	0.5127	0.3539

Table A4.6d: Closeness centrality for COX-1 residues

Residue	APO	FLP	1OL	LG2	OLEO	MMHTE
Active site						
93	0.2457	0.2317	0.1817	0.2148	0.2701	0.2005
116	0.1873	0.2457	0.1912	0.1876	0.2385	0.1885
117	0.1877	0.2460	0.1940	0.2024	0.2645	0.1966
120	0.1709	0.3043	0.2260	0.2132	0.2446	0.1945
205	0.2807	0.3301	0.2315	0.2452	0.2448	0.3183
209	0.2796	0.3078	0.2218	0.2161	0.2291	0.2982
344	0.3034	0.3382	0.2195	0.2528	0.2886	0.2912
345	0.3230	0.3096	0.2025	0.2469	0.2845	0.2840
348	0.3550	0.3487	0.2024	0.2788	0.3215	0.2425
349	0.3071	0.3049	0.1946	0.2453	0.2855	0.2389
352	0.2972	0.2794	0.2036	0.2325	0.2849	0.2410
353	0.3013	0.2897	0.1910	0.2379	0.2817	0.2325
355	0.2850	0.2821	0.1964	0.2399	0.2769	0.2418
359	0.2793	0.2734	0.1948	0.2306	0.2704	0.2397
381	0.3461	0.3756	0.2569	0.2583	0.2839	0.3360
384	0.3043	0.3399	0.2509	0.2578	0.2803	0.3080
385	0.3321	0.3395	0.2814	0.2485	0.3131	0.2914
387	0.3034	0.3681	0.2710	0.2668	0.2680	0.2613
513	0.2099	0.2611	0.1903	0.1723	0.2053	0.2702
518	0.2633	0.2507	0.2445	0.1902	0.2580	0.2543
523	0.2644	0.2928	0.2375	0.2157	0.2407	0.2533
526	0.2409	0.2939	0.2609	0.2267	0.2686	0.2811
527	0.2877	0.3032	0.2360	0.2285	0.2505	0.2855
530	0.2730	0.3161	0.2357	0.2122	0.2534	0.2850
531	0.2888	0.3251	0.2366	0.2070	0.2407	0.2871
533	0.2523	0.2766	0.2188	0.2068	0.1962	0.2428
534	0.2985	0.3222	0.2258	0.2122	0.2299	0.2449
Peroxidase site						
203	0.3037	0.3725	0.2336	0.2610	0.2748	0.3217
207	0.2912	0.3474	0.2276	0.2391	0.2356	0.3537
211	0.2488	0.3082	0.2080	0.2151	0.2175	0.2773
222	0.2002	0.2676	0.1807	0.1899	0.1459	0.2299
240	0.2114	0.2800	0.2075	0.2049	0.1999	0.2459
274	0.1842	0.2068	0.1554	0.1673	0.1543	0.1548
290	0.2517	0.3082	0.1995	0.2243	0.2104	0.3031
388	0.2931	0.3611	0.2367	0.2464	0.2826	0.3075
391	0.2829	0.3510	0.2532	0.2267	0.2894	0.2725
409	0.2571	0.3115	0.2192	0.2118	0.2300	0.2763
504	0.2903	0.2776	0.2261	0.2319	0.2334	0.2995
EGF						
32	0.1477	0.1892	0.1136	0.1381	0.1511	0.1874
35	0.1919	0.2177	0.1549	0.1570	0.1866	0.2346

38	0.1749	0.1972	0.1469	0.1578	0.1792	0.2313
41	0.1769	0.2423	0.1621	0.1757	0.2123	0.2346
44	0.1345	0.2107	0.1269	0.1379	0.1507	0.1894
47	0.1512	0.2416	0.1320	0.1412	0.1893	0.2355
50	0.1167	0.2118	0.1144	0.1311	0.1507	0.2292
53	0.1046	0.1753	0.1027	0.1311	0.1505	0.1867
56	0.1326	0.2124	0.1290	0.1413	0.1542	0.2300
59	0.1228	0.2132	0.1372	0.1458	0.1808	0.2335
62	0.1225	0.2127	0.1320	0.1378	0.1556	0.2048
65	0.1359	0.1962	0.1441	0.1380	0.1559	0.1921
68	0.1526	0.2130	0.1503	0.1379	0.1811	0.1919
71	0.1386	0.2152	0.1372	0.1457	0.1619	0.2354
72	0.1360	0.2113	0.1513	0.1589	0.1620	0.1923
MBD						
73	0.1278	0.1787	0.1395	0.1572	0.1635	0.1935
76	0.1476	0.1529	0.1441	0.1575	0.1755	0.1970
79	0.1476	0.1549	0.1430	0.1482	0.1628	0.1683
82	0.1668	0.1460	0.1429	0.1638	0.1800	0.1752
85	0.1921	0.1657	0.1473	0.1820	0.1920	0.1773
88	0.2297	0.2239	0.1571	0.1881	0.2249	0.1994
91	0.2272	0.2224	0.1571	0.2056	0.2254	0.2007
94	0.2278	0.2282	0.1572	0.1884	0.2252	0.1976
97	0.1877	0.1991	0.1465	0.1893	0.1957	0.1691
100	0.2290	0.2441	0.1863	0.2176	0.2342	0.2002
103	0.1897	0.2202	0.1868	0.1933	0.1967	0.2001
106	0.2276	0.2200	0.1977	0.2137	0.2083	0.2044
109	0.1898	0.2455	0.2036	0.2073	0.2373	0.2024
Glycosylation						
68	0.1526	0.2130	0.1503	0.1379	0.1811	0.1919
144	0.2237	0.2188	0.1540	0.1859	0.1948	0.2140
410	0.2769	0.3108	0.2210	0.2086	0.2411	0.2787
580	0.3150	0.3397	0.2471	0.2664	0.2839	0.2934

Table A4.6e: PageRank for COX-1 residues

Residue	APO	FLP	1OL	LG2	OLEO	MMHTE
Active site						
93	0.0018	0.0019	0.0023	0.0019	0.0023	0.0017
116	0.0022	0.0018	0.0020	0.0017	0.0017	0.0021
117	0.0020	0.0019	0.0026	0.0021	0.0024	0.0023
120	0.0018	0.0028	0.0021	0.0029	0.0024	0.0020
205	0.0013	0.0022	0.0012	0.0016	0.0012	0.0019
209	0.0011	0.0013	0.0006	0.0006	0.0007	0.0012
344	0.0019	0.0032	0.0021	0.0020	0.0027	0.0024
345	0.0029	0.0027	0.0019	0.0021	0.0025	0.0022
348	0.0029	0.0030	0.0019	0.0027	0.0032	0.0016
349	0.0022	0.0022	0.0017	0.0022	0.0027	0.0014
352	0.0017	0.0016	0.0019	0.0015	0.0025	0.0013
353	0.0019	0.0019	0.0014	0.0017	0.0027	0.0011
355	0.0021	0.0023	0.0020	0.0019	0.0026	0.0018
359	0.0013	0.0012	0.0017	0.0010	0.0025	0.0013
381	0.0033	0.0031	0.0030	0.0018	0.0029	0.0033
384	0.0016	0.0022	0.0027	0.0020	0.0027	0.0031
385	0.0025	0.0019	0.0023	0.0013	0.0024	0.0016
387	0.0020	0.0027	0.0016	0.0009	0.0013	0.0008
513	0.0007	0.0016	0.0008	0.0011	0.0011	0.0011
518	0.0018	0.0011	0.0013	0.0014	0.0016	0.0009
523	0.0020	0.0015	0.0026	0.0016	0.0014	0.0012
526	0.0012	0.0017	0.0023	0.0018	0.0029	0.0017
527	0.0017	0.0020	0.0023	0.0023	0.0026	0.0020
530	0.0014	0.0017	0.0021	0.0018	0.0021	0.0019
531	0.0020	0.0015	0.0024	0.0014	0.0023	0.0018
533	0.0009	0.0010	0.0011	0.0012	0.0007	0.0011
534	0.0015	0.0016	0.0017	0.0017	0.0018	0.0013
Peroxidase site						
203	0.0024	0.0033	0.0017	0.0017	0.0017	0.0022
207	0.0009	0.0020	0.0015	0.0011	0.0007	0.0024
211	0.0016	0.0012	0.0015	0.0008	0.0007	0.0017
222	0.0010	0.0013	0.0010	0.0009	0.0009	0.0008
240	0.0007	0.0022	0.0016	0.0013	0.0014	0.0014
274	0.0010	0.0006	0.0008	0.0007	0.0012	0.0009
290	0.0008	0.0012	0.0010	0.0010	0.0012	0.0015
388	0.0012	0.0022	0.0016	0.0013	0.0015	0.0010
391	0.0016	0.0021	0.0027	0.0006	0.0020	0.0009
409	0.0016	0.0021	0.0012	0.0013	0.0014	0.0015
504	0.0025	0.0021	0.0019	0.0017	0.0015	0.0032
EGF						
32	0.0006	0.0008	0.0007	0.0009	0.0005	0.0004
35	0.0020	0.0022	0.0013	0.0013	0.0021	0.0024

38	0.0015	0.0012	0.0021	0.0017	0.0018	0.0017
41	0.0016	0.0031	0.0025	0.0025	0.0034	0.0022
44	0.0014	0.0012	0.0009	0.0012	0.0009	0.0006
47	0.0017	0.0028	0.0022	0.0019	0.0026	0.0026
50	0.0020	0.0014	0.0019	0.0013	0.0019	0.0013
53	0.0013	0.0012	0.0013	0.0015	0.0012	0.0008
56	0.0016	0.0015	0.0018	0.0019	0.0010	0.0013
59	0.0020	0.0019	0.0020	0.0021	0.0019	0.0024
62	0.0019	0.0017	0.0014	0.0012	0.0014	0.0014
65	0.0019	0.0015	0.0019	0.0017	0.0016	0.0015
68	0.0021	0.0023	0.0026	0.0016	0.0025	0.0013
71	0.0025	0.0022	0.0023	0.0023	0.0019	0.0023
72	0.0021	0.0018	0.0026	0.0026	0.0020	0.0018
MBD						
73	0.0011	0.0010	0.0018	0.0019	0.0019	0.0013
76	0.0021	0.0016	0.0026	0.0026	0.0022	0.0021
79	0.0021	0.0016	0.0019	0.0017	0.0019	0.0018
82	0.0019	0.0009	0.0017	0.0018	0.0017	0.0016
85	0.0013	0.0009	0.0010	0.0010	0.0012	0.0014
88	0.0020	0.0015	0.0016	0.0014	0.0016	0.0017
91	0.0014	0.0012	0.0016	0.0016	0.0017	0.0021
94	0.0017	0.0014	0.0018	0.0013	0.0015	0.0015
97	0.0013	0.0013	0.0010	0.0014	0.0011	0.0012
100	0.0024	0.0024	0.0022	0.0023	0.0026	0.0022
103	0.0020	0.0018	0.0019	0.0022	0.0014	0.0018
106	0.0026	0.0015	0.0023	0.0023	0.0013	0.0024
109	0.0027	0.0023	0.0023	0.0021	0.0022	0.0026
Glycosylation						
68	0.0021	0.0023	0.0026	0.0016	0.0025	0.0013
144	0.0010	0.0011	0.0012	0.0018	0.0018	0.0010
410	0.0018	0.0017	0.0013	0.0012	0.0012	0.0012
580	0.0028	0.0023	0.0030	0.0024	0.0034	0.0021

Table A4.6f: Degrees for COX-2 residues

Residue	APO	SAL	1OL	LG2	OLEO	MMHTE
Active site						
93	10	13	22	18	15	11
116	18	20	18	20	14	21
117	17	17	18	24	16	17
120	31	18	14	19	16	18
205	111	44	91	35	85	76
209	61	34	59	28	67	41
344	94	65	39	49	35	48
345	112	59	44	44	41	54
348	80	55	35	45	40	19
349	102	63	21	46	32	32
352	16	42	22	30	15	6
353	30	30	7	4	9	7
355	28	30	28	34	16	16
359	28	34	12	37	6	9
381	123	39	46	16	56	30
384	133	58	16	16	51	30
385	68	17	6	9	22	15
387	67	23	13	28	27	19
513	12	14	3	12	3	4
518	21	50	17	25	11	31
523	74	71	51	56	41	35
526	26	37	25	26	22	20
527	36	22	26	43	31	21
530	39	26	32	34	16	17
531	30	29	29	26	19	24
533	45	29	19	20	15	21
534	38	25	20	24	19	18
Peroxidase site						
203	87	51	108	34	116	88
207	77	17	71	17	68	42
211	33	21	23	22	16	10
222	5	4	4	4	3	5
240	53	52	47	63	79	35
274	10	5	18	14	75	11
290	29	15	47	30	77	45
388	117	83	41	50	64	32
391	102	60	34	29	78	42
409	27	36	37	25	53	41
504	72	14	58	53	39	28
EGF						
32	1	1	1	1	1	1
35	20	16	28	20	10	10

38	19	9	21	9	10	8
41	25	25	28	25	18	19
44	17	16	27	19	13	13
47	18	14	22	15	5	4
50	4	5	5	5	5	5
53	3	4	5	5	4	4
56	32	29	38	34	28	25
59	7	15	22	20	14	8
62	14	23	30	26	14	11
65	17	16	22	21	17	18
68	29	28	39	29	25	26
71	16	23	28	24	23	20
72	8	11	22	21	11	15
MBD						
73	3	6	11	6	10	11
76	4	6	15	11	18	16
79	4	7	9	9	10	9
82	4	5	4	10	7	3
85	10	10	13	10	10	11
88	9	10	17	10	10	11
91	9	12	24	20	18	22
94	15	15	24	16	16	20
97	7	13	17	9	10	14
100	8	18	17	6	5	19
103	6	31	29	28	19	18
106	14	29	22	25	16	13
109	19	26	21	32	17	18
Glycosylation						
68	29	28	39	29	25	26
144	19	10	7	11	6	9
410	29	54	39	24	45	18
580	37	11	32	34	4	6

Table A4.6g: Betweenness centrality for COX-2 residues

Residue	APO	SAL	1OL	LG2	OLEO	MMHTE
Active site						
93	319.97	960.37	160.46	268.44	464.842761	17.1937745
116	49.33	20.60	112.40	25.00	441.47564	1310.49145
117	160.40	35.87	139.60	412.77	181.460051	95.1777724
120	3744.92	9.33	1.68	545.35	374.220287	165.514922
205	4339.63	2124.70	5040.84	1693.64	1947.92895	6962.39163
209	218.65	1196.61	2893.71	4342.83	4497.17182	4099.52274
344	1758.31	3335.26	349.49	3247.79	38.0482632	997.476146
345	2400.37	3055.97	253.52	1611.17	828.633045	2454.40066
348	746.41	1122.85	392.16	1265.92	404.215328	23.4982063
349	2913.01	2855.11	365.53	1232.47	3579.52699	2170.2068
352	0.87	589.17	5990.97	2575.33	112.376599	12.711584
353	5.84	66.62	146.77	0.00	98.2925923	66.4542876
355	196.48	125.49	465.73	4134.26	757.134595	303.23311
359	69.90	635.97	158.87	512.46	1136.16607	1104.00377
381	1442.68	1193.72	4584.31	4.98	4038.61708	4088.52603
384	1862.92	6329.79	88.47	186.85	3172.78006	3239.60252
385	128.12	2.93	24.66	57.92	16.7511974	80.7832152
387	214.04	14.55	52.93	20.59	89.3856851	351.772425
513	53.36	24.11	0.00	83.02	0	0
518	4.20	151.45	15.64	36.88	117.436857	344.107761
523	571.18	1790.23	589.31	2872.36	960.762828	938.986669
526	44.31	324.53	413.47	163.33	129.64339	189.453944
527	2273.56	569.09	3529.32	5208.16	4211.8037	3015.73481
530	151.92	154.07	299.90	593.04	23.996191	125.076364
531	451.29	101.70	1896.54	708.43	433.446039	2102.91159
533	710.77	1340.72	1554.42	2199.76	2303.53787	2544.89986
534	1263.99	3075.71	969.12	3371.95	3894.01183	4570.61422
Peroxidase site						
203	2367.86	1011.46	7406.96	4516.19	3319.59756	4410.50518
207	1128.83	609.27	3575.28	858.63	1836.17131	1333.24487
211	330.13	767.87	605.74	839.08	930.471003	1232.52839
222	11.30	14.93	3.82	50.03	6.16355662	13.4682669
240	281.25	329.02	119.38	169.06	142.884343	19.0050715
274	565.27	26.90	722.07	568.65	548.18356	552.116141
290	22.71	107.32	13.96	16.88	300.644154	8.91084738
388	495.14	513.08	262.49	216.03	831.370869	1143.7204
391	1138.39	952.74	500.15	1044.65	1343.52732	549.734414
409	47.33	49.45	148.32	104.71	59.3958521	35.9557322
504	67.13	1.24	441.39	193.90	223.519696	167.781699
EGF						
32	0.00	0.00	0.00	0.00	0	0
35	1189.61	7352.63	739.22	1738.39	56.0702351	124.724231

38	732.12	219.27	236.23	68.08	510.538342	950.207687
41	1280.61	133.16	37.85	35.49	3.30315813	135.538418
44	47.25	102.25	20.85	2.55	3.04938704	1.90740648
47	205.45	93.57	58.04	22.76	1286.94377	264.319753
50	953.88	1020.89	21.82	194.41	314.523803	387.122434
53	142.52	105.02	966.82	269.86	290.898155	719.536877
56	747.66	1801.05	2752.30	5719.84	6137.8478	8067.38445
59	12.42	157.72	3.68	5.14	338.635438	2.37306845
62	40.75	883.61	81.50	1327.25	667.354813	150.259219
65	272.15	17.70	12.54	14.69	3.42867488	193.044746
68	1947.29	1121.92	1485.83	74.99	1200.01379	572.70386
71	129.30	344.17	40.75	77.12	2326.56283	487.874231
72	1025.86	9.85	1107.25	1063.81	819.483802	435.703568
MBD						
73	0.00	13.83	61.37	12.72	20.1038906	249.503978
76	314.49	10.05	96.27	1792.33	8285.33093	596.933267
79	428.11	778.84	186.26	816.51	273.673977	3732.68089
82	1630.70	1354.90	190.76	7054.24	6086.46426	358.113062
85	423.61	23.32	2481.49	0.00	0	6.64313333
88	14.33	23.32	79.95	0.00	0	6.64313333
91	176.17	469.16	1371.09	1109.51	1172.22796	1748.52166
94	1994.13	1093.28	320.72	107.02	973.614492	371.807125
97	286.73	653.46	4.67	2.11	44.6095439	0.97173937
100	77.16	233.87	6.15	2.48	24.6812792	342.472669
103	0.00	1058.22	389.29	1141.74	3702.89024	46.9532838
106	252.09	188.78	289.33	75.48	293.414658	71.6815347
109	667.90	156.66	56.66	660.54	76.3930393	139.713754
Glycosylation						
68	1947.29	1121.92	1485.83	74.99	1200.01379	572.70386
144	589.94	1351.22	222.40	130.08	23.4554895	715.285162
410	394.16	528.30	74.86	19.40	225.786722	106.962031
580	4.24	0.25	15.63	28.93	0	0.20238095

Table A4.6h: Eigencentrality for COX-2 residues

Residue	APO	SAL	1OL	LG2	OLEO	MMHTE
Active site						
93	0.0022	0.0046	0.0109	0.0115	0.004308	0.004013
116	0.0113	0.0193	0.0092	0.0279	0.005852	0.007717
117	0.0130	0.0159	0.0093	0.0408	0.005817	0.00653
120	0.1045	0.0165	0.0067	0.0269	0.006441	0.007054
205	0.6469	0.3825	0.6456	0.1344	0.592982	0.621946
209	0.3689	0.2370	0.3720	0.0997	0.465147	0.293964
344	0.5454	0.3697	0.2639	0.2768	0.217203	0.353658
345	0.6882	0.3086	0.2904	0.2400	0.221573	0.380259
348	0.5017	0.3374	0.1981	0.2379	0.213677	0.076329
349	0.6989	0.4121	0.1185	0.2155	0.171379	0.182549
352	0.0839	0.2593	0.0729	0.1598	0.047853	0.00583
353	0.1365	0.1200	0.0189	0.0124	0.018201	0.021006
355	0.1072	0.1052	0.0145	0.0656	0.00649	0.007012
359	0.1065	0.1093	0.0073	0.1467	0.003086	0.004181
381	0.8109	0.2055	0.2187	0.0795	0.213642	0.138454
384	0.9019	0.3701	0.0402	0.0634	0.206122	0.078742
385	0.5168	0.0940	0.0076	0.0356	0.071501	0.014375
387	0.4837	0.1573	0.0407	0.1437	0.090326	0.02829
513	0.0790	0.0861	0.0007	0.0640	0.000501	0.001604
518	0.1386	0.3191	0.0475	0.1355	0.013253	0.033886
523	0.5005	0.4456	0.1257	0.2979	0.08132	0.03826
526	0.1474	0.1883	0.0481	0.1234	0.02812	0.017134
527	0.1280	0.0742	0.0392	0.1809	0.03693	0.017291
530	0.2265	0.1236	0.0698	0.1707	0.020017	0.01338
531	0.1123	0.1516	0.0599	0.1138	0.030785	0.04022
533	0.2285	0.1315	0.0489	0.0587	0.038636	0.059124
534	0.1474	0.1063	0.0273	0.0815	0.014067	0.011071
Peroxidase site						
203	0.5511	0.4803	0.8022	0.2021	0.805421	0.728018
207	0.4773	0.0890	0.4923	0.0686	0.467105	0.335159
211	0.1609	0.0466	0.0874	0.0530	0.043022	0.019438
222	0.0057	0.0020	0.0022	0.0024	0.000653	0.001306
240	0.1485	0.4067	0.3562	0.6284	0.605734	0.260037
274	0.0034	0.0016	0.0676	0.0606	0.53651	0.035006
290	0.1029	0.0951	0.4118	0.3376	0.582157	0.398023
388	0.8367	0.6185	0.1150	0.2824	0.276696	0.051535
391	0.7347	0.4177	0.1435	0.1513	0.438419	0.198524
409	0.0877	0.3001	0.2870	0.1969	0.422637	0.320671
504	0.5142	0.0706	0.1514	0.3295	0.07344	0.03677
EGF						
32	0.0002	0.0003	0.0006	0.0003	0.000185	0.000185
35	0.0073	0.0071	0.0175	0.0126	0.002777	0.002756

38	0.0068	0.0034	0.0128	0.0051	0.002718	0.002237
41	0.0074	0.0125	0.0178	0.0189	0.006344	0.007491
44	0.0053	0.0086	0.0169	0.0149	0.00467	0.005281
47	0.0054	0.0067	0.0142	0.0103	0.000993	0.000932
50	0.0006	0.0010	0.0010	0.0013	0.000789	0.000821
53	0.0005	0.0008	0.0015	0.0013	0.000629	0.000668
56	0.0092	0.0138	0.0224	0.0226	0.008822	0.008717
59	0.0020	0.0080	0.0145	0.0154	0.004976	0.003039
62	0.0041	0.0113	0.0182	0.0186	0.004469	0.00424
65	0.0052	0.0087	0.0143	0.0163	0.006132	0.006942
68	0.0082	0.0138	0.0227	0.0212	0.008287	0.009599
71	0.0047	0.0117	0.0171	0.0181	0.007586	0.007707
72	0.0019	0.0054	0.0120	0.0139	0.002774	0.005528
MBD						
73	0.0005	0.0024	0.0058	0.0039	0.003049	0.003748
76	0.0006	0.0016	0.0082	0.0048	0.005014	0.005594
79	0.0005	0.0017	0.0032	0.0038	0.002301	0.00248
82	0.0006	0.0011	0.0009	0.0034	0.001549	0.000511
85	0.0018	0.0031	0.0056	0.0053	0.002854	0.004024
88	0.0017	0.0031	0.0085	0.0053	0.002854	0.004024
91	0.0021	0.0044	0.0128	0.0195	0.005187	0.008505
94	0.0031	0.0085	0.0119	0.0085	0.004329	0.007427
97	0.0014	0.0096	0.0088	0.0046	0.002763	0.005552
100	0.0017	0.0169	0.0088	0.0047	0.001423	0.007873
103	0.0011	0.0536	0.0159	0.0449	0.011469	0.007386
106	0.0049	0.0456	0.0123	0.0460	0.006083	0.006984
109	0.0078	0.0275	0.0112	0.0609	0.005672	0.008565
Glycosylation						
68	0.0082	0.0138	0.0227	0.0212	0.008287	0.009599
144	0.0738	0.0074	0.0032	0.0225	0.002853	0.002394
410	0.1332	0.4623	0.3043	0.2166	0.325964	0.105364
580	0.2441	0.0867	0.2232	0.1452	0.01976	0.025234

Table A4.6i: Closeness centrality for COX-2 residues

Residue	APO	SAL	1OL	LG2	OLEO	MMHTE
			Active site			
93	0.2521	0.2175	0.2274	0.2413	0.214564	0.189542
116	0.3048	0.2549	0.2577	0.2728	0.235672	0.228631
117	0.3251	0.2593	0.2583	0.2936	0.239253	0.221641
120	0.3546	0.2513	0.2379	0.2882	0.236278	0.225727
205	0.4401	0.3420	0.4022	0.3370	0.313068	0.335566
209	0.4043	0.3399	0.3911	0.3485	0.318681	0.316122
344	0.4356	0.3811	0.3174	0.3439	0.283145	0.290153
345	0.4415	0.3736	0.3512	0.3325	0.313246	0.308857
348	0.4162	0.3635	0.3226	0.3178	0.291689	0.261014
349	0.4370	0.3683	0.3272	0.3258	0.3006	0.289847
352	0.3397	0.3410	0.3044	0.3189	0.273449	0.232098
353	0.3431	0.3117	0.2862	0.2570	0.263889	0.259538
355	0.3505	0.3124	0.2344	0.2961	0.236481	0.221196
359	0.3476	0.3141	0.2531	0.3021	0.239149	0.230737
381	0.4451	0.3573	0.3953	0.2751	0.338244	0.324308
384	0.4528	0.3738	0.3071	0.3036	0.330534	0.322978
385	0.3848	0.3147	0.2906	0.2942	0.280693	0.277024
387	0.3800	0.3152	0.2980	0.2876	0.28213	0.280407
513	0.3115	0.3070	0.2163	0.2681	0.176716	0.209426
518	0.3337	0.3325	0.3029	0.2940	0.24316	0.274676
523	0.3856	0.3642	0.3439	0.3220	0.290306	0.276606
526	0.3530	0.3405	0.3317	0.3066	0.265927	0.275225
527	0.3861	0.3136	0.3232	0.3251	0.292153	0.270363
530	0.3779	0.3345	0.3362	0.3217	0.258929	0.269174
531	0.3751	0.3348	0.3264	0.2991	0.283436	0.295126
533	0.3916	0.3418	0.3339	0.3178	0.27998	0.27495
534	0.3710	0.3207	0.3158	0.3082	0.259906	0.272772
			Peroxidase site			
203	0.4152	0.3620	0.4028	0.3422	0.330534	0.328172
207	0.4229	0.3008	0.3975	0.3339	0.317031	0.311828
211	0.3673	0.3125	0.3583	0.3172	0.272233	0.263636
222	0.2787	0.2398	0.2528	0.2486	0.190657	0.208003
240	0.3315	0.2939	0.3198	0.2773	0.29138	0.255684
274	0.2399	0.1988	0.2644	0.2265	0.281986	0.231902
290	0.3290	0.2566	0.2975	0.2652	0.292153	0.257236
388	0.4238	0.3681	0.3401	0.3125	0.311828	0.302581
391	0.4165	0.3485	0.3481	0.3158	0.324499	0.293085
409	0.3132	0.2918	0.3149	0.2769	0.281986	0.269174
504	0.3738	0.2818	0.3571	0.3183	0.286383	0.265286
			EGF			
32	0.2049	0.1845	0.2004	0.1568	0.1033	0.114173
35	0.2750	0.2285	0.2451	0.1883	0.115514	0.130786

38	0.2748	0.2108	0.2453	0.1876	0.125627	0.146309
41	0.2536	0.1797	0.2097	0.1681	0.118546	0.133253
44	0.2309	0.1905	0.2068	0.1624	0.118495	0.131409
47	0.2310	0.1781	0.2064	0.1623	0.106659	0.11696
50	0.1812	0.1505	0.1710	0.1397	0.093263	0.104002
53	0.1642	0.1402	0.2012	0.1397	0.093248	0.103708
56	0.2415	0.1929	0.2506	0.1921	0.128259	0.148598
59	0.2071	0.1905	0.2025	0.1625	0.126871	0.120017
62	0.2260	0.1795	0.2100	0.1792	0.126871	0.132963
65	0.2314	0.1779	0.2055	0.1679	0.118648	0.133221
68	0.2330	0.1927	0.2384	0.1685	0.127399	0.133608
71	0.2273	0.1795	0.2061	0.1681	0.12734	0.133414
72	0.1909	0.1634	0.2106	0.1784	0.139247	0.125742
MBD						
73	0.1613	0.1536	0.1961	0.1653	0.124576	0.125627
76	0.1620	0.1439	0.1964	0.1837	0.1396	0.12577
79	0.1535	0.1473	0.1812	0.1834	0.137338	0.132229
82	0.2030	0.1620	0.1848	0.2079	0.15565	0.144203
85	0.2080	0.1927	0.2174	0.2186	0.199205	0.185148
88	0.2071	0.1927	0.2188	0.2186	0.199205	0.185148
91	0.2519	0.2174	0.2558	0.2663	0.217443	0.220576
94	0.2584	0.2289	0.2287	0.2303	0.214647	0.200364
97	0.2326	0.2460	0.2220	0.2203	0.194837	0.194356
100	0.2305	0.2465	0.2206	0.2283	0.203246	0.221108
103	0.1951	0.2868	0.2562	0.2790	0.254034	0.215487
106	0.2804	0.2859	0.2436	0.2845	0.240716	0.23547
109	0.2914	0.2721	0.2407	0.3127	0.228536	0.236684
Glycosylation						
68	0.2330	0.1927	0.2384	0.1685	0.127399	0.133608
144	0.3284	0.2634	0.2591	0.2725	0.231027	0.213566
410	0.3412	0.3124	0.3170	0.2846	0.280264	0.252174
580	0.3401	0.2998	0.3241	0.2970	0.239774	0.241985

Table A4.6j: PageRank for COX-2 residues

Residue	APO	SAL	1OL	LG2	OLEO	MMHTE
Active site						
93	0.0018	0.0017	0.0019	0.0019	0.001911	0.001312
116	0.0015	0.0014	0.0016	0.0014	0.001489	0.002606
117	0.0013	0.0013	0.0016	0.0016	0.001716	0.002014
120	0.0018	0.0013	0.0013	0.0014	0.001704	0.002062
205	0.0033	0.0018	0.0033	0.0019	0.00308	0.002939
209	0.0019	0.0017	0.0025	0.0017	0.002615	0.002239
344	0.0029	0.0027	0.0015	0.0024	0.001383	0.001947
345	0.0034	0.0026	0.0017	0.0022	0.00172	0.002278
348	0.0024	0.0023	0.0015	0.0023	0.001737	0.001072
349	0.0030	0.0026	0.0010	0.0023	0.001529	0.001623
352	0.0007	0.0018	0.0013	0.0017	0.000926	0.000666
353	0.0012	0.0015	0.0006	0.0005	0.000751	0.000673
355	0.0012	0.0016	0.0024	0.0023	0.001786	0.001677
359	0.0012	0.0018	0.0012	0.0020	0.000993	0.001442
381	0.0036	0.0017	0.0020	0.0010	0.002634	0.001782
384	0.0038	0.0024	0.0010	0.0010	0.002295	0.002035
385	0.0020	0.0009	0.0006	0.0007	0.001137	0.001203
387	0.0020	0.0011	0.0008	0.0014	0.001301	0.00133
513	0.0007	0.0008	0.0008	0.0009	0.001063	0.000727
518	0.0008	0.0020	0.0010	0.0013	0.000956	0.002069
523	0.0023	0.0028	0.0025	0.0026	0.002265	0.002331
526	0.0011	0.0016	0.0015	0.0014	0.001498	0.00154
527	0.0017	0.0012	0.0017	0.0022	0.002162	0.001646
530	0.0014	0.0013	0.0018	0.0018	0.001135	0.001377
531	0.0014	0.0014	0.0017	0.0015	0.001278	0.001664
533	0.0016	0.0015	0.0012	0.0014	0.001247	0.001567
534	0.0018	0.0015	0.0014	0.0016	0.001792	0.001906
Peroxidase site						
203	0.0025	0.0019	0.0037	0.0016	0.003931	0.003101
207	0.0024	0.0011	0.0027	0.0011	0.002646	0.001752
211	0.0014	0.0016	0.0015	0.0016	0.001455	0.001467
222	0.0007	0.0008	0.0007	0.0008	0.000761	0.001214
240	0.0020	0.0021	0.0017	0.0024	0.002532	0.001399
274	0.0014	0.0012	0.0012	0.0013	0.002539	0.001112
290	0.0011	0.0009	0.0016	0.0013	0.002535	0.001621
388	0.0032	0.0030	0.0020	0.0022	0.00262	0.002039
391	0.0029	0.0023	0.0016	0.0014	0.002922	0.001941
409	0.0011	0.0015	0.0015	0.0013	0.001798	0.001567
504	0.0021	0.0008	0.0027	0.0023	0.002151	0.001831
EGF						
32	0.0004	0.0004	0.0004	0.0004	0.000487	0.000543
35	0.0021	0.0020	0.0023	0.0023	0.001826	0.002145

|						
38	0.0020	0.0014	0.0019	0.0012	0.001876	0.001773
41	0.0026	0.0027	0.0023	0.0024	0.002182	0.002351
44	0.0019	0.0018	0.0022	0.0019	0.001667	0.001713
47	0.0020	0.0018	0.0019	0.0017	0.001682	0.001259
50	0.0012	0.0016	0.0012	0.0013	0.00204	0.002239
53	0.0011	0.0013	0.0011	0.0013	0.001703	0.001789
56	0.0033	0.0031	0.0031	0.0034	0.003755	0.003557
59	0.0010	0.0017	0.0018	0.0019	0.001758	0.001148
62	0.0016	0.0026	0.0025	0.0026	0.001889	0.001503
65	0.0019	0.0017	0.0018	0.0020	0.002086	0.002246
68	0.0032	0.0029	0.0032	0.0027	0.003003	0.003179
71	0.0018	0.0025	0.0023	0.0023	0.002837	0.002502
72	0.0014	0.0014	0.0020	0.0022	0.001713	0.001993
MBD						
73	0.0009	0.0010	0.0012	0.0008	0.001456	0.001573
76	0.0013	0.0012	0.0015	0.0015	0.002512	0.002128
79	0.0016	0.0015	0.0012	0.0013	0.001645	0.001613
82	0.0012	0.0013	0.0009	0.0015	0.001306	0.001039
85	0.0019	0.0015	0.0014	0.0012	0.001368	0.001313
88	0.0017	0.0015	0.0016	0.0012	0.001368	0.001313
91	0.0016	0.0016	0.0021	0.0019	0.002215	0.002279
94	0.0025	0.0016	0.0021	0.0018	0.002127	0.002123
97	0.0013	0.0013	0.0015	0.0011	0.001456	0.001522
100	0.0012	0.0014	0.0015	0.0007	0.000824	0.001968
103	0.0010	0.0020	0.0024	0.0021	0.001893	0.001857
106	0.0015	0.0018	0.0018	0.0016	0.001686	0.001394
109	0.0018	0.0018	0.0018	0.0020	0.001813	0.00192
Glycosylation						
68	0.0032	0.0029	0.0032	0.0027	0.003003	0.003179
144	0.0012	0.0012	0.0012	0.0011	0.000949	0.001865
410	0.0011	0.0021	0.0015	0.0011	0.001608	0.000913
580	0.0012	0.0006	0.0013	0.0016	0.000402	0.000512

9.4.5 MM-PBSA
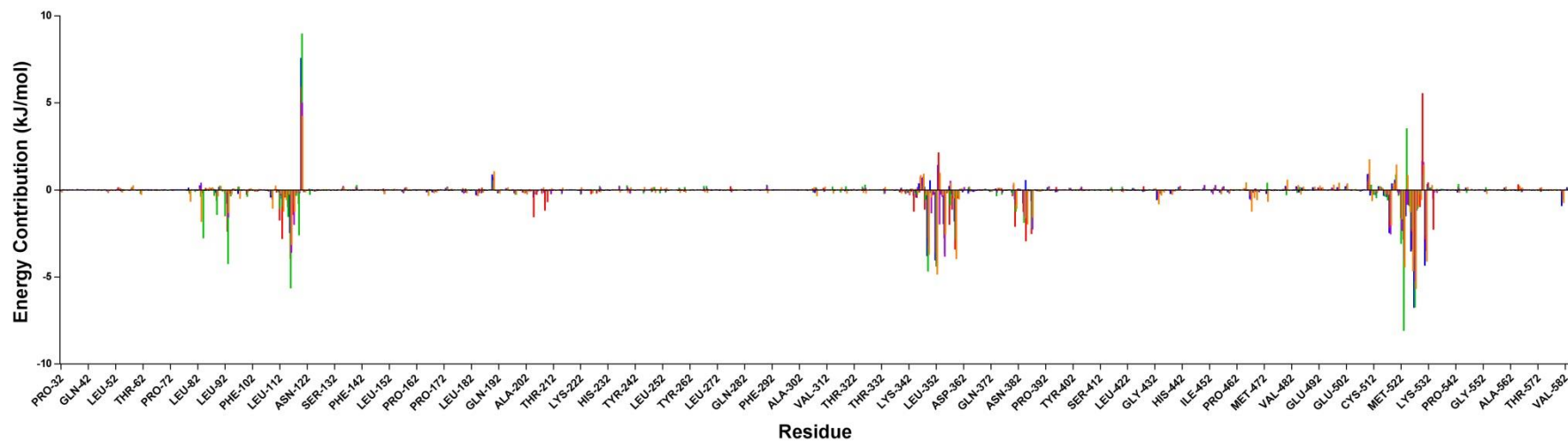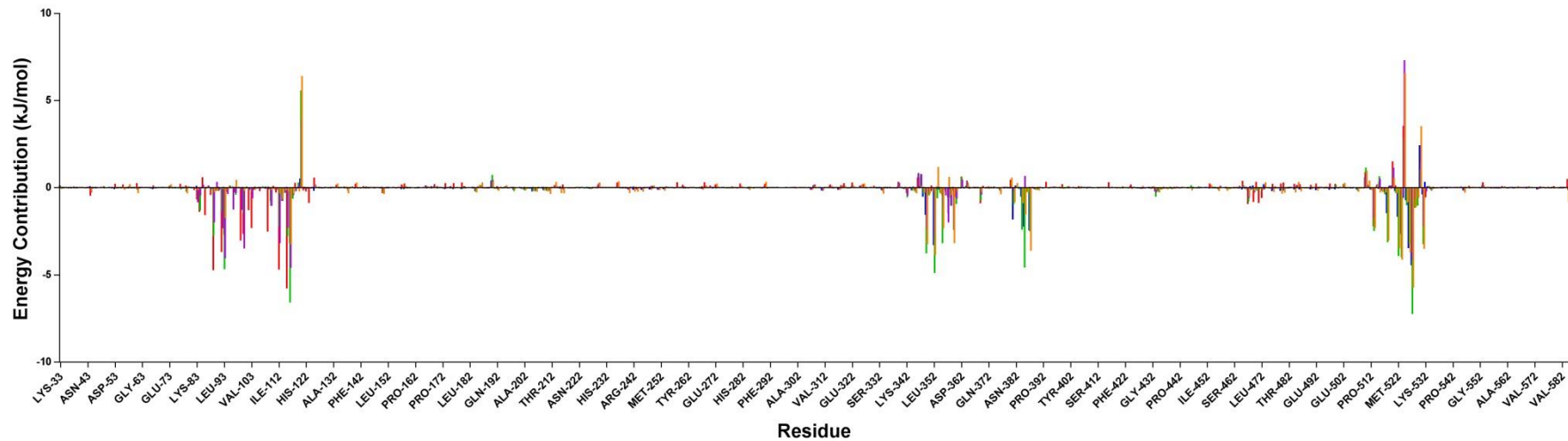
Figure A4.7: Residue contribution for binding of olive ligands to COX-1 (A) and COX-2 (B) proteins with the native ligand (blue), 1OL (red), LG 2 (green), OLEO (purple), and MMHTE (orange).
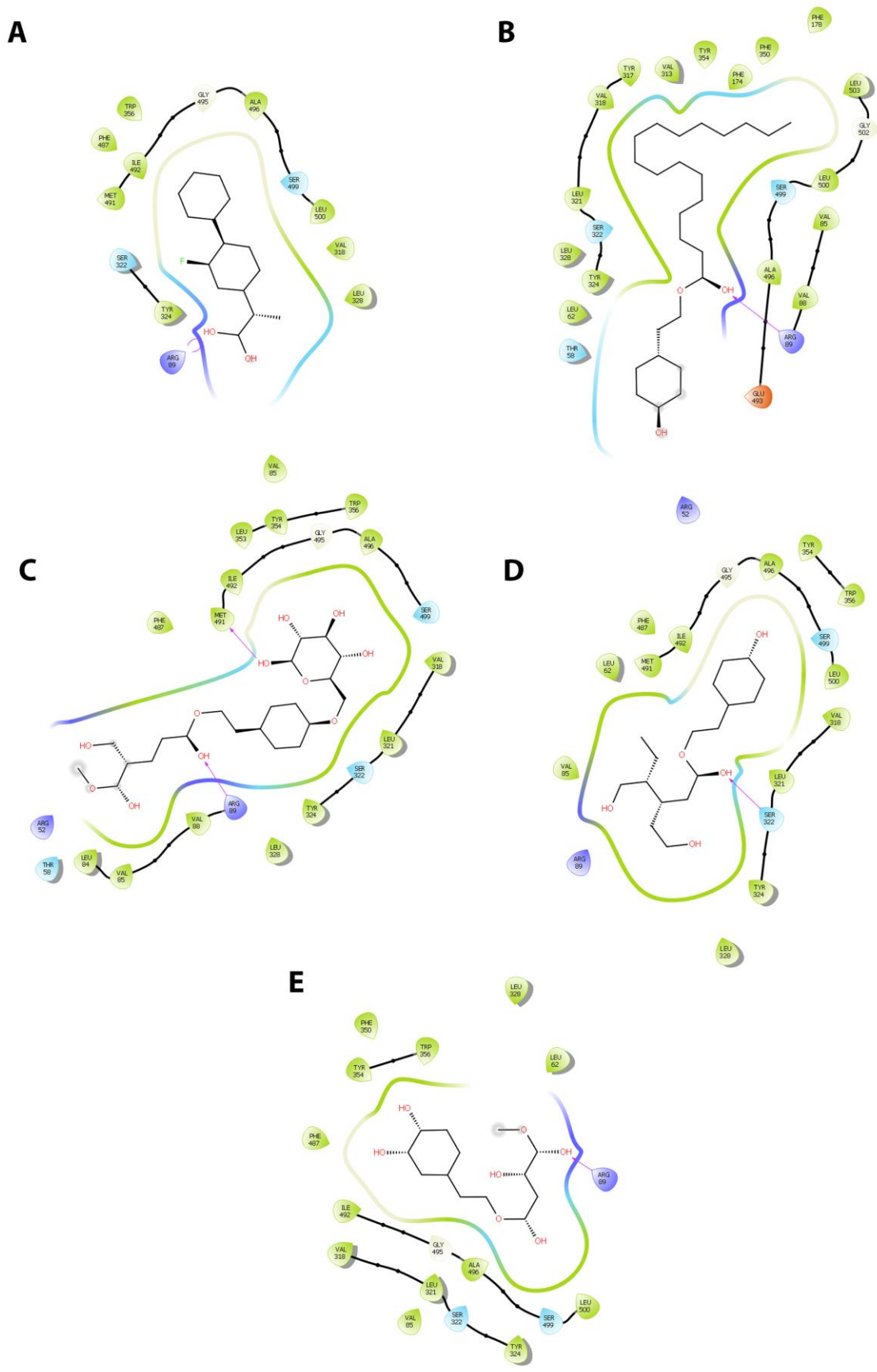
Figure A4.8: 2D interaction diagram for dynamic COX-1 complexes with FLP (A), 1OL (B), LG2 (C), OLEO (D), and MMHTE (E). Hydrogen bonds are represented by purple arrows.
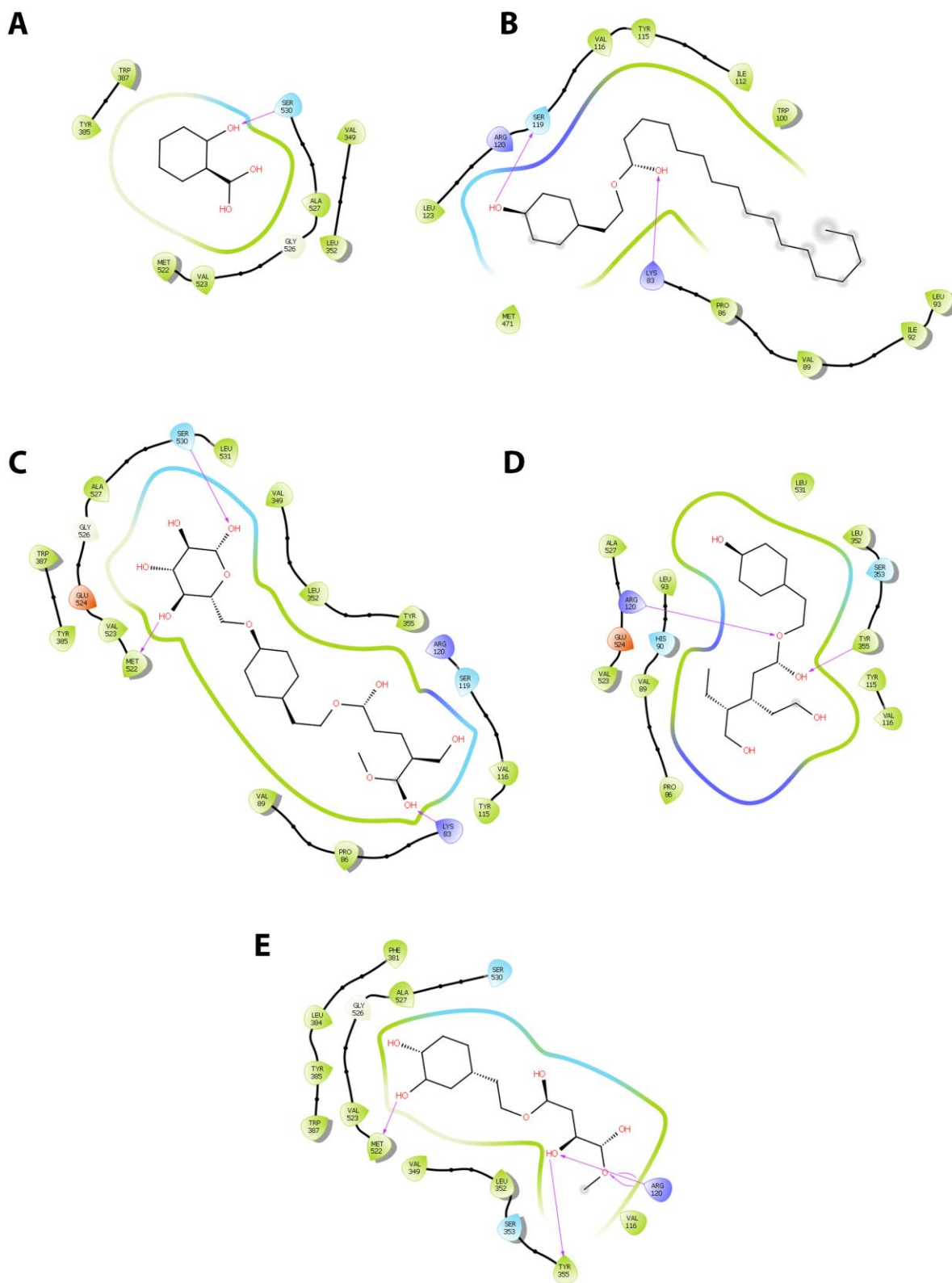
Figure A4.9: 2D interaction diagram for dynamic COX-2 complexes with SAL (A), 1OL (B), LG2 (C), OLEO (D), and MMHTE (E). Hydrogen bonds are represented by purple arrows.