 doi:10.20944/preprints201902.0124.v4

Ligand-Protein Interactions: A Hybrid *ab initio*/Molecular Mechanics Computational Study

Yornei R. Perez[a], ***Dinais Alvarez***[a], and ***Aldo F. Combariza***[*,a]

[a]*in silico Molecular Modeling and Computational Simulation Research Group, Department of Biology and Chemistry, Faculty of Education and Sciences, University of Sucre, Colombia*

[*]aldo.combariza@unisucre.edu.co

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 29 May 2019 doi:10.20944/preprints201902.0124.v4

Abstract

Cyclooxygenases (COX), or prostaglandin endoperoxide synthases (PTGS), are key enzimes in the synthesis of prostaglandins, which are chemical species critical in mediating inflammatory processes. There are two highly homologous COX isoforms: COX-1 and COX-2. COX-1 is involved in the production of prostaglandins, chemical compuounds that take part in physiological processes such as: protection of the gastric epithelium, maintenance of renal flow, platelet aggregation, neutrophil migration and, also, are expressed in the vascular endothelium. Meanwhile, COX-2 is induced by proinflammatory stimuli. It is very frequent the use of nonsteroidal antiinflammatory drugs (NSAIDs) to counteract the symptoms of inflammatory processes. These drugs, in addition to its benefits, can cause side effects on peopleâĂŹs health, such as cardiovascular and respiratory problems, among others. In the past years, it has been recognized the potential of plants secondary metabolites as pharmacological agents, prompting the need for investigations that shed light into its mechanism of action. In this work we have applied computational techniques, based on quantum chemistry and mechanical statistics, to study the protein-ligand interaction involving COX's and secondary metabolites from natural sources. Our aim is to determine the structure activity interplay in processes involving the participation of secondary plant metabolites such as luteolin, galangin, kaempferol, apigenin, morine and quercetin on the inactivation of COX's. From molecular docking analysis, we have extracted the energetics of the COX-(1,2)/metabolite coupling. By defining energy based factors, we have determined a procedure that predicts the chemical species with highest stability and selectivity towards inactivation of COX-2 over COX-1. The results are discussed with regard to conformational features of the selected ligands and its intermolecular strong/weak interactions inside the active-sites of the COX's hosts.

Keywords enzymes Cyclooxygenase, ligand, metabolites, antiinflammatory, Molecular Docking.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 29 May 2019 doi:10.20944/preprints201902.0124.v4

Introduction

The awe-inspiring development of computational performance in the last decades has open the door to a new way of doing science, which is usually referred as computational experimentation or *in silico* experimentation.[1–3] In general, an *in silico* experiment begins by modeling a real system using solid theoretical physical principles, followed by the simulation of the system through computer programmed mathematical algorithms.[3–5]

In this work, we use a hybrid *ab initio*/mechanical-statistical technique to study the features of biological macromolecules, proteins, and their interactions with low molecular weight chemical species, called ligands. These protein-ligand systems, are a paradigmatic example of the structure-activity principle.[5,6] We use here the Molecular Docking (MDock) technique, which allows us to predict the most appropriate conformation of a molecule (ligand) when it binds to another (protein) to form a stable complex.[6] In fact, the use of MDock its seen as a advantageous technique in the study of protein-ligand interactions within the field of drug discovery and development.[7–9] Our goal is to understand the structure-activity relationship of the ligand-protein interactions from an atomistic point of view and, therefore, to attain relevant information to disclose the effect of a biological active molecule on a specific proteic target.[10]

One relevant proteic target, very useful in the clinical research field, is the Prostaglandin endoperoxide synthase (PTGS) or Cyclooxygenase (COX) enzyme.[11,12] The first isoform is constitutive, and executes indispensable functions in the organism, while the second isoform intervenes in inflammatory processes, becoming the target of study for the search of novel anti-inflammatory compounds.[13] Both, COX-1 and COX-2, have a high structural similarity, only differing by the relative position of three amino acids. This structural similarity brings about non-selectivity between COX-1 and COX-2, *e.g.*, non steroidal anti-inflammatory drugs (NSAIDs) are not selective for any form of COX, which causes serious consequences in the organism.[13,14]

The lack of NSAIDs-COX selectivity, has fueled the search for new bio-active molecules. Sometimes these species are obtained from natural sources, *e.g.*, secondary metabolites from plants or animals, some of which are recognized to have a higher COX selectivity, helping to avoid or minimize the side effects produced by NSAIDs in the body.[15,16] Within the secondary metabolites obtained from plants, we have a very wide group of chemicals known as flavonoids, which include: flavones,

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 29 May 2019 doi:10.20944/preprints201902.0124.v4

flavonols, antocyanidins and isoflavonols.[17] These metabolites are recognized as anti-allergenic, anti-thrombotic, anti-oxidant and anti-inflammatory species.[18] Some examples of vegetal species which are used to extract these compunds are: *Dioclea grandiflora, Sophora flavescens, Swinglea glutinosa* and *Muntingia calabura*.[15] The aim of this work is to use tools of computational biochemistry to describe the structure-activity relationship between COX-ligand systems, which is of paramount importance in inflammatory inhibition, involving the action of secondary metabolites.

Methodology

COX structural characterization

COX-1,2 enzymes belong to the oxidoreductase protein family, and forms complexes with selective inhibitors such as SC-558.[19] We have used structural information from the PDB-database, as well as software such as PyMOL[20] to clean up the reported structures, for example, eliminating hemo groups, water molecules and ligands present in the structural data.

COX-1

From the Protein Data Bank (PDB)[21], we have extracted 217 structures directly related with COX-1. These structures were discriminated according to the species from which they have been obtained, and are distributed as follows: *Homo sapiens* (51), *Bos taurus* (29), *Mus musculus* (25), *Ovis aries* (22), *Escherichia coli* (12), *Rattus norvegicus* (12), *Zea mays* (10), others species (53).[21] Currently, there is a record of 171 ligands that form complexes with COX-1, among which are flurbiprofen, arachidonic acid, ibuprofen, sodium ion and nitric acid.[21,22]

COX-2

For COX-2, we have found 220 structures from the PDB, which were extracted from species like: *Homo sapiens* (51), *Mus musculus* (43), *Bos taurus* (28), *Ovis aries* (21), *Escherichia coli* (12), *Rattus norvegicus* (12), *Saccharomyces cerevisiae* (8), others (42).[21] For this COX isoform, there are 178 ligands recognized, among which are the following: acetylsalicylic acid, flurbiprofen, urea, arachidonic acid and ibuprofen.[22]

For COX-1 we have used the model corresponding to *Ovis aries* (PDB ID: code 3N8Z) and for

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 29 May 2019 doi:10.20944/preprints201902.0124.v4

COX-2, the one for *Mus musculus* (PDB ID: code 1CX2). The experimental method used to dilucidate the tertiary structures of COXâĂŹs was reported to be Xray's crystallography.[19,23] These structures were selected because, according to Kurumbail *et al*[19], human COX-2 has a similarity of 87% with the murine enzyme, and the amino acid sequence that forms the active site is conserved between both species, therefore, we expect that human COX-2, and mainly the active site, will behave similarly to murine.

Ligands characterization

The ChemSpider[24] database was used to retrieve the chemical species of interest.[25] These species are secondary metabolites from plants, belonging to the flavonoid family. We have selected the following metabolites as subjects of study:

- Apigenin (5,7-Dihydroxy-2-(4-hydroxyphenyl)-4H-chromen-4-one)[26]
- Luteolin (2-(3,4-Dihydroxyphenyl)-5,7-dihydroxy-4H-chromen-4-one)[27]
- Quercetin (2-(3,4-Dihydroxyphenyl)-3,5,7-trihydroxychromen-4-one)[28]
- Morin (2-(2,4-Dihydroxyphenyl)-3,5,7-trihydroxy-4H-chromen-4-one)[29]
- Galangin (3,5,7-Trihydroxy-2-phenyl-4H-chromen-4-one)[30]
- Kaempferol (3,5,7-Trihydroxy-2-(4-hydroxyphenyl)-4H-chromen-4-one)[31]

Ligand Geometry optimization

Geometry optimization is a molecular modeling technique that allows us to obtain the structural conformations of lowest energy of a molecule by determining the critical points of the potential energy function and the rotations of chemical bonds that form the molecule.[32,33] With this information we predict the tridimensional disposition of the atoms in a molecule creating a potential energy surface.[32–36]

MOPAC[37] was used to optimize the ligands chemical structures. MOPAC is a program based on the theoretical principles of quantum mechanics (QM), that is, the use of semi-empirical methods to predict properties of chemical structures and model chemical reactions.[38] MNDO, MINDO/3, AM1 and PM3 were the Hamiltonians used to achieve molecular orbitals, heat of formation and

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 29 May 2019 doi:10.20944/preprints201902.0124.v4

its derivatives with respect to molecular geometry.[38–41] Followed by the calculation of properties like: vibration spectra, thermodynamic quantities and constant forces that corresponds to molecules, radicals, ions and polymers.[39,40] After the ligand optimization, we define parameters like the torsion center and the rotable bonds of the ligands.

Molecular Docking

Before the MDock process, we add the hydrogen bonds missing to COX structures. This step is carried out because the hydrogen atoms have a small nuclei and are very mobile, so they are not detected by the X rays in crystallographic analysis. AutoDock Vina (AUV)[42] was used to guide the MDock Process. AUV uses a Lamarckian Genetic Algorithm (LGA) and a free-energy Empirical Punctuation Function (EPF) that allows a faster search method and provides reproducible results in larger systems.[43–45] This software uses a semi-empiric force-field to evaluate the free-energy of conformations along the coupling simulations. At the same time, the force-field is parameterized using experimental information recover from numerous protein-inhibitor systems.[43] We consider ten conformations for each protein-ligand pair.

Selectivity and Stability Indexes

We have developed an energetically based index of selectivity and stability for each COX-ligand pair. The stability index is calculated by normalizing every Protein-Ligand interaction energy using the lowest calculated interaction energy. Likewise, the selectivity index was determined by dividing the binding energy of each ligand/COX-2 interaction between the binding energy of each ligand-COX-1 interaction, taking into account the correspondence between the ligands. Finally, the values obtained were normalized.

Results and Discussion

Molecular characterization of COX

The tertiary structures of COX are very similar, thus, they share the same folding units. Because of this, we use the COX-1 structure to characterize both enzymes. We can divide the tertiary structure of COX-1 in three folding units, showed in Figure 1.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 29 May 2019 doi:10.20944/preprints201902.0124.v4

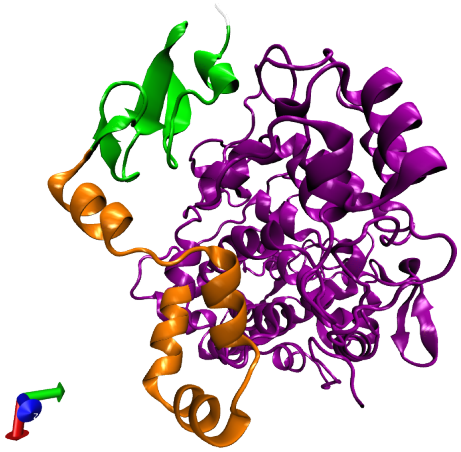

Figure 1: Folding units of COX-1. First unit is show in green, the second is show in orange and the third is show in violet.

The first folding unit is formed by residues 34-72 in positions amino-terminal and form a small compact domain binded by three disulfide bonds, as Figure 2A shows. The global conformation of this domain, including the disposition of the disulfide bonds, is very similar to the conformation of the epidermal growth factor and binds covalently to the principal chain of the enzyme by another disulfide bridge, that is, Cys-37-Cys-159. The second unit is formed by residues 73-116 and form a structure of alpha helix A, B, C and D that are located at one side of the first folding unit, as we illustrate in Figure 2B. These alpha helix are highly amphipathic and represents a structural motif for the insertion of enzyme to the lipid bilayer. And the last folding unit consist of a large catalytic globular domain formed by residues 117-587. This catalytic domain is a globular structure that contains the cyclooxygenase and peroxidase active sites. Both sites, peroxidase and cyclooxygenase, comprise two differents lobes in the polypeptidic chain although interlaced with each other. The largest lobe is made by a structure formed by seven alpha helices, whereas the smallest lobe of the catalytic domain is formed by six alpha helices. These alpha helices form a set, more or less parallel, with its axes (see Figure 2C).

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 29 May 2019 doi:10.20944/preprints201902.0124.v4

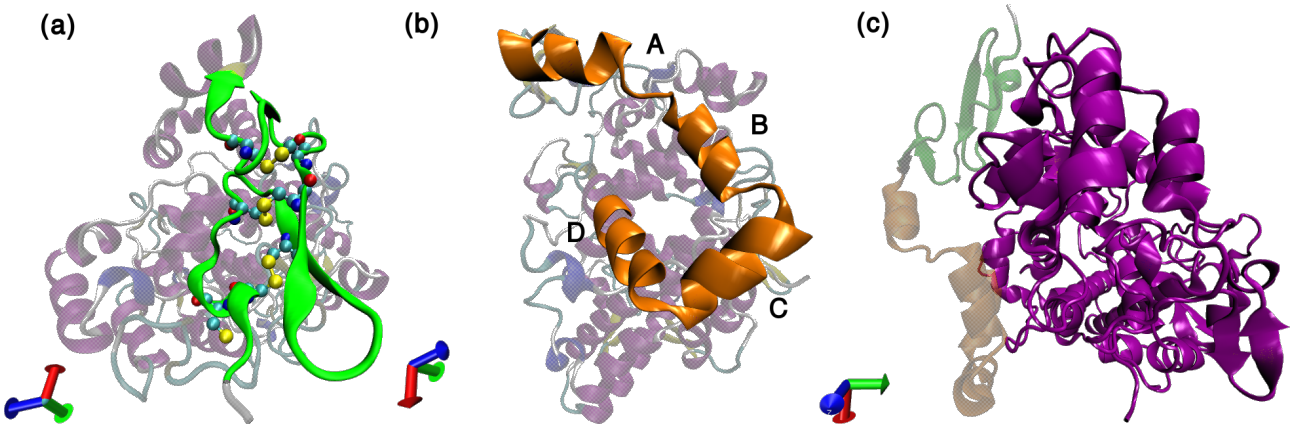

Figure 2: Folding units of COX-1. (a) first folding unit in green and the three disulfide bonds are shown in yellow; (b) in orange is show the second folding unit and the four alpha helices (A, B, C and D); (c) in violet, the third folding unit is represent by alpha helices and beta sheets.

Characterization of the secondary structures of COX

Table 1: COX-1 secondary structure characterization

Coil		3-10 Helix	Alpha helix	Beta sheet
32-34	324	35-38	74-81	46-50
39-41	354-362	281-283	86-93	54-58
63	370-377	363-366	97-105	64-65
70	391-394	388-390	107-121	71-72
73	403	498-500	139-143	130-131
82-85	411	538-540	174-182	149-150
94-96	418	547-550	196-207	245-248
106	428-429		238-244	251-252
129	431		296-319	255-257
135-138	433-439		325-353	260-262
148	459-462		379-385	395-397
154-155	471-477		404-407	400-402
167	483-484		412-417	
172-173	496-497		419-427	
188	501-502		445-458	
211	510-511		463-470	
236-237	517-518		478-482	
253-254	537		485-495	
263-264	546		503-509	
271-275	551-552		520-536	
280	562-563		553-561	
284	571		564-570	
286	576-580			
	582-584			

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 29 May 2019 doi:10.20944/preprints201902.0124.v4

Table 2: COX-2 secondary structure characterization

Coil		3-10 HÃl'lices	HÃl'lices alfa	LÃąminas Beta
33	263-264	363-366	74-81	46-49
40	271-275	388-390	86-94	55-58
63	324	498-500	97-104	64-65
70	354-362	538-540	106-121	71-72
73	370-377		139-143	245-248
82-85	391		174-182	251-252
95-96	396		196-209	255-257
105	401		238-243	260-262
129-130	403		296-317	
138	418		325-353	
148	428-431		379-384	
154-155	437-439		404-407	
167	459-461		411-417	
172-173	471-477		419-427	
188	483-484		444-458	
195	501		463-470	
210-212	510-511		478-482	
236-237	561-563		485-495	
244	576-580		503-509	
253-254			520-533	
			553-560	
			564-571	

Topological and molecular characterization of the active sites of COX

The active sites of COX enzymes have a high similarity. Both active sites consist of a hydrophobic channel with an approximate lenght of 25 Angstroms. This channel originates in the Membrane Binding Domain (MBD) and extends to the nucleus of the globular domain.[19,23,46] The amino acids surrounding the hydrophobic channel of COX active sites are: Leu117, Arg120, Val434 (COX-2), Ile434 (COX-1), Phe205, Phe209, Val344, Ile345, Tyr348, Val349, Leu352, Ser353, Tyr355, Leu359, Phe381, Leu384, Tyr385, Trp387, Phe518, Ile523 (COX-1), Val523 (COX-2), Gly526, Ala527, Ser530, His513 (COX-1), Arg513 (COX-2), Leu531, Gly533, Leu534, shown in Figure 3. Only three of the channel residues are polar Arg120, Ser353 y Ser530.

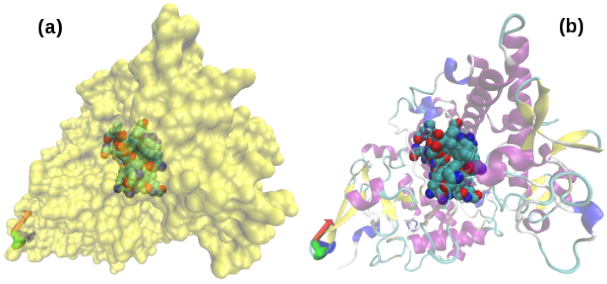

Figure 3: Amino acids of COX active site. (a) shown at the center of the figure are the amino acids that form the active site of the enzyme, highlighting the third folding unit.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 29 May 2019 doi:10.20944/preprints201902.0124.v4

The residue Ser530 is acetylated by aspirin, while the residue Arg120 binds to the carboxylate groups of the fatty acids and many NSAIDs.[47–50]

The size of the active site of COX-2 is approximately about 20% more large that the COX-1 enzyme. This difference in the active sites of COX is very important due to the variation of three residues at positions 434, 513 and 523 of both enzymes, as we show in Figure 4. The variation in the size of both active sites has been an important feature for the development of specific NSAIDs for COX-2.[51] Figure 5 shows, an superposition of the active sites of both enzymes with the residues that differentiate them, at the same point.

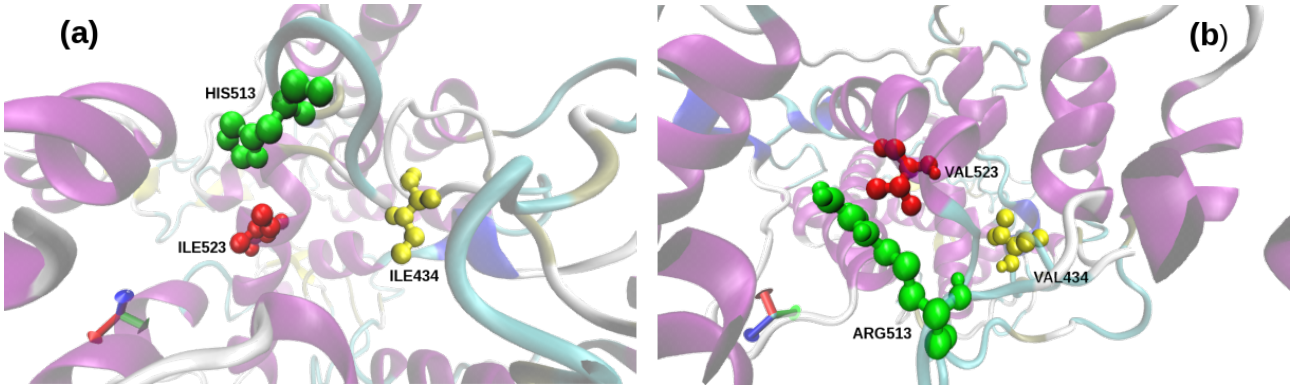

Figure 4: Amino acids that differentiate the active sites of COX. A. Amino acid Ile434 is highlighting in yellow, His513 in green and Ile523 in red to COX-1. B. Amino acids of COX-2 are highlighting as: Val434 in yellow, Arg513 in green and Val523.

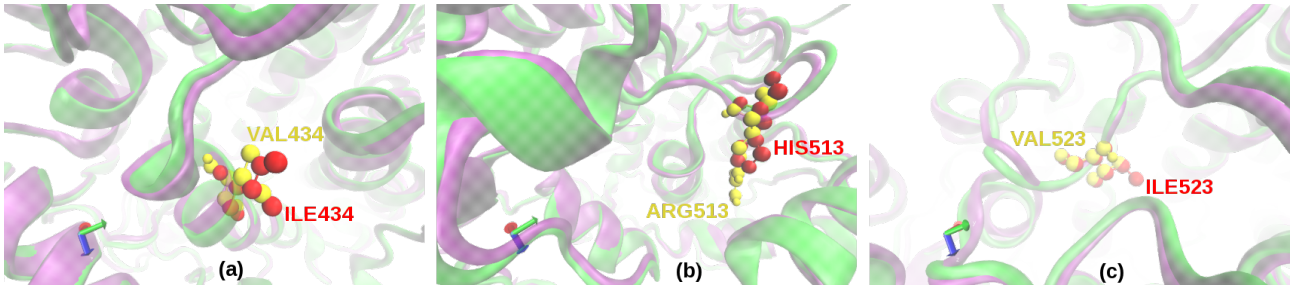

Figure 5: Superposition of COX active sites. COX-1: green, COX-2: violet. (a) isoleucin for COX-1 and valine for COX-2; (b) histidin for COX-1 and arginin for COX-2; (c) isoleucin for COX-1 and valin for COX-2.

Geometry optimization

Table 3 shows the values obtained of minimization energies calculations and the evaluation of molecular HOMO and LUMO orbitals for each ligand. These values are ordered according to the HOMO energy from the best nucleophile to the best electrophile. The HOMO orbital is the last orbital that is doubly occupied and depicts the location of the pair of electrons most susceptible to electrophilic

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 29 May 2019 doi:10.20944/preprints201902.0124.v4

attack. Taking into account the values shown in Table 1, we can see that, of all the ligands, galangin is the best nucleophile, while apigenin is the worst nucleophile.

Table 3: Minimization energy values and molecular orbitals HOMO and LUMO [kJ/mol]

Ligand	Galangin	Kaempferol	Quercetin	Morin	Luteolin	Apigenin
Total Energy	-3452.9	-3748.7	-4044.0	-4043.9	-3749.0	-3453.7
HOMO	-8.974	-9.075	-9.099	-9.147	-9.351	-9.458
LUMO	-0.755	-0.997	-1.084	-0.8	-1.18	-1.071

It should be noted that the calculations of the energies of the molecular orbitals HOMO and LUMO was carried out to know the distribution of the electrical charges of the molecules. Therefore, interactions with electrophilic groups are controlled by the HOMO orbital, while the nucleophilic attacks are controlled by the LUMO orbitals. Thus, Figure 6 allow us to observe the charge distribution and regions of the molecule with high electron density, that is, regions susceptible to electrophilic and nucleophilic attacks. These areas also show the electrostatic potential of the molecules, the red color shows the negative part and the blue color shows the positive part. The green and yellow regions show the poorest electron-rich zones.

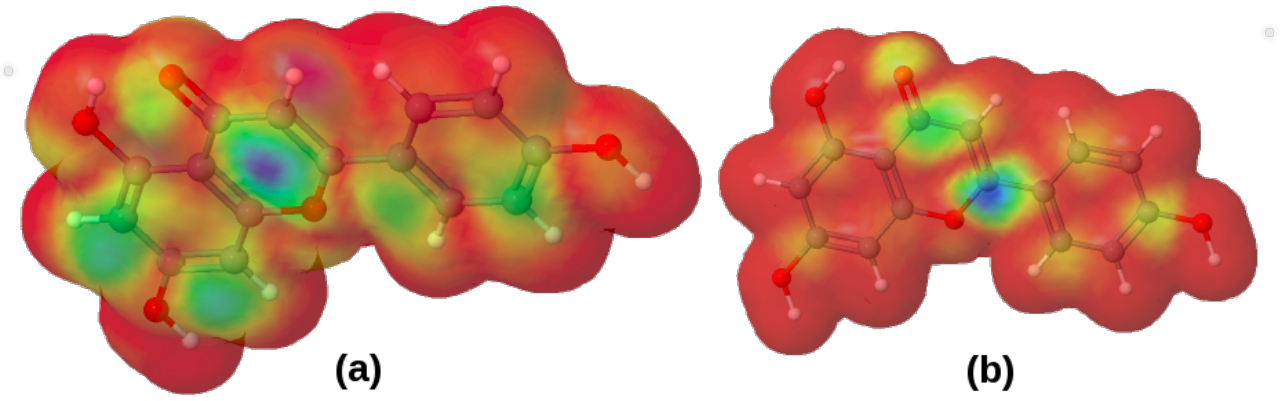

Figure 6: Apigenin electronic density maps. (a) HOMO orbital: in red, the zone most susceptible to electrophilic attacks; (b) LUMO orbital: the region most susceptible to nucleophilic attacks.

On the other hand, Table 4 shows the geometry optimization energy values for each ligand in gaseous state, that is, before the coupling process occurs, as well as the optimization energies after the coupling process and the energy differences for COX-1/ligand interaction. Table 5 shows the values corresponding to COX-2/ligand interactions.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 29 May 2019 doi:10.20944/preprints201902.0124.v4

Table 4: Optimization and pos-coupling energies of COX-1/ligand complex.

COX-1						
Ligand	Luteolin	Galangin	Apigenin	Kampferol	Morin	Quercetin
Optimization Energy	-3749.02	-3452.95	-3453.72	-3748.74	-4043.90	-4044.03
MOPAC Energy	-3746.46	-3452.89	-3307.05	-3745.36	-4042.01	-4043.43
Energy Difference	-2.56	-0.05	-146.67	-3.38	-1.89	-0.60

Table 5: Optimization and pos-coupling energy values of COX-2/ligand complex

COX-2						
Ligand	Luteolin	Galangin	Apigenin	Kaempferol	Morin	Quercetin
Optimization Energy	-3749.02	-3452.95	-3453.72	-3748.74	-4043.90	-4044.03
MOPAC Energy	-3746.82	-3452.80	-3449.55	-3745.36	-4041.95	-4042.41
Energy Difference	-2.20	-0.15	-4.18	-3.38	-1.95	-1.63

Molecular Docking

For each protein-ligand complex, we have selected the top ten conformations and for each one we evaluated the binding energy of the complex formed. Table 6 shows the interaction energy for each conformation for the Galangin/COX-1 system, as well as the amino acids and hydrogen bonds involved. For every conformation there is, at least, the formation of one hydrogen bond, except conformation 9 where it is shown the formation of 3 hydrogen bonds. The conformation number nine have the lowest energy-binding value, which means that in this conformation the ligand have more stability in the active site of enzyme compared with other conformations.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 29 May 2019 doi:10.20944/preprints201902.0124.v4

Table 6: COX-1/Galangin complex. 1-10: obtained conformations and bindng energy values X: amino acids within the interaction orbit of the ligand, for each conformation. Line upwards: all aminoacids present in all of the conformations. Highlighted in green: amino acids interacting with the ligand and repeated in the indicated conformations.

Amino acids	Conformations									
	1	2	3	4	5	6	7	8	9	10
TYR 355	X	X	X	X	X	X	X	X	X	X
SER 530	X	X	X	X	X	X	X	X	X	X
ILE 523	X	X	X	X	X	X	X	X	X	X
MET 522	X	X	X	X	X	X	X	X	X	X
LEU 531	X	X	X	X	X	X	X	X	X	X
VAL 349	X	X	X	X	X	X	X	X	X	X
ALA 527	X	X	X	X	X	X	X	X	X	X
TRP 387	X		X	X	X	X	X	X	X	X
SER 353	X	X	X	X	X	X	X	X		X
LEU 352	X	X	X	X	X	X	X	X		X
TYR 385	X		X	X	X	X	X	X		X
PHE 518		X		X	X	X	X	X	X	X
GLY 526		X							X	
TYR 348	X		X							
PHE 381		X								
ARG 120									X	
VAL 116		X								
LEU 359		X								
LEU 384									X	
Bind energy	-6.96	-6.96	-6.97	-7.2	-7.18	-7.19	-7.18	-7.18	-7.26	-7.18

Most stable configuration

The selection criteria of the most stable configuration for each protein-ligand coupling, was based on energetic values, that is, the lowest interaction energy indicate the highest stability of the complex.

For each of the 6 protein-ligands pairs, the best conformation based on the aforementioned energy terms was selected. The binding energy values between all the ligands were compared in order to determine which of them best matches the active site of the protein. Table 7 shows the best conformation

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 29 May 2019 doi:10.20944/preprints201902.0124.v4

for each COX-1/ligand pair and the amino acids that interact with the ligand in each conformation. The interaction is remarkable by hydrogen bonds, attraction forces, electrostatic interactions and Van der Waals dispersion forces. Table 8 shows this information for COX-2/ligand system.

Table 7: Best conformation of COX-1/ligand complex

COX-1						
Ligand	Apigenin	Kaempferol	Luteolin	Morin	Galangin	Quercetin
Conformations	1-10	1	5	6	9	3
Binding energy	-8.88	-8.24	-7.79	-7.71	-7.26	-6.65
Amino acids	TRP 387	TRP 387	TRP 387	TRP 387	TRP 387	TRP 387
	VAL 349	HIS 388	VAL 349	LEU 390	VAL 349	VAL 349
	SER 530	HIS 386	SER 530	PHE 210	SER 530	SER 530
	ILE 523	LEU 390	ILE 523	ALA 202	ILE 523	ILE 523
	GLY 526	MET 391	LEU 384	GLN 203	GLY 526	TYR 355
	ALA 527	TYR 385	TYR 385	HIS 207	MET 522	SER 353
	LEU 384	ASN 382	MET 522	ASN 382	ALA 527	TYR 385
	LEU 531	ALA 202	ALA 527	TYR 385	LEU 531	LEU 352
	MET 522	THR 206	ARG 120	THR 206	PHE 518	VAL 116
	LEU 352	HIS 207	VAL 116	HIS 388	ARG 120	TYR 348
	ARG 120	PHE 210	PHE 381	HIS 386	LEU 384	ARG 120
	TYR 355		PHE 518		TYR 355	MET 522
	VAL 116		LEU352			ALA527

Apigenin shows the lowest coupling energy in complex with COX-1, which means that this ligand have an advantage in energetic terms that others. Apigenin interact with the same amino acids in all conformations of the COX-1/Apigenin system and have the same energetic values. The ligands that follows Apigenin in terms of high energetic stability are: Kaempferol, Luteolin, Morin, Galangin and Quercetin, repectively.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 29 May 2019 doi:10.20944/preprints201902.0124.v4

Table 8: Best conformation of COX-2/ligand complex

COX-2						
Ligand	Apigenine	Kaempferol	Luteolin	Morin	Galangin	Quercetin
Conformations	7	3	1	1	3 and 8	1 and 9
Binding energy	-8.93	-7.51	-7.58	-7.91	-7.55	-7.76
Amino acids	SER 530	SER 353	VAL 523	HIS 90	TYR 355	TYR 355
	ALA 378	VAL 523	TYR 355	SER 353	VAL 523	VAL 5239
	ILE 377	HIS 90	SER 530	VAL 349	HIS 90	SER 530
	PHE 381	LEU 352	LEU 352	LEU 352	SER 353	TYR 385
	PHE 209	TYR 355	TYR 385	SER 530	LEU 352	LEU 352
	PHE 205	PHE 518	ALA 527	VAL 523	ALA 527	SER 353
	VAL 228	ILE 517	HIS 90	TYR 355	PHE 518	GLN 192
	LEU 534	GLN 192	ARG 513	MET 522	ALA 516	HIS 90
	GLY 533	ALA 516	TYR 348	LEU 384	ILE 517	PHE 518
	ASN 375	GLY 526	SER 353	TYR 385	GLN 192	ALA 527
	PHE 529	PHE 381		TRP387		
	LYS 532			PHE 518		
	ILE 124					

Table 8 shows that conformation 7 (Apignenin) have the lowest value of coupling energy, namely, is the most stable conformation of the ligand in the active site of enzyme in the coupling process. Galangin and Quercetin shows 2 conformations with the same amino acids and the same values of coupling energy. In this case, the ligands that follows Apigenin in high stability are: Morin, Quercetin, Luteolin, Galangin and Kaempferol, respectively. On the other hand, electrostatic terms has been the more relevants energetic terms to choose the conformations and the more stable ligands.

Figure 7 shows, highlighted in red, five zones of high electronic density of Apigenin, in which the hydroxyl groups of the molecule are found. Due to the high electronic density, the formation of electrostatic interactions is possible. The amino acids that interact with Apigenin are: SER 530, ALA 378, ILE 377, PHE 381, PHE 209, PHE 205, VAL 228, LEU 534, GLY 533, ASN 375, PHE 529,

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 29 May 2019 doi:10.20944/preprints201902.0124.v4

LYS 532 and ILE 124. There is formation of a hydrogen bond between the amino group of ALA 378 and a hydroxyl group (OH-) of Apigenin, which helps to stabilize said conformation.

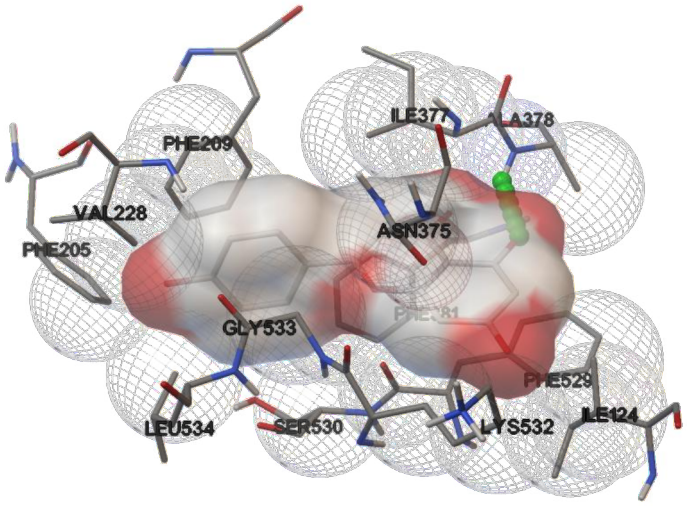

Figure 7: COX-2/Apigenin complex. The ligand is surrounded by amino acids that participate in the interaction with the active site of the enzyme. A hydrogen bond is shown in green.

Table 8 shows a comparison between the COX-1/ligand and COX-2/ligand interactions. Following the energetic terms, Apigenin have the most high stability in the active site of COX-1 with a value of -8.88 and a value of -8.93 in COX-2. While, the comparison of the coupling energy of Apigenin in COX-1 and COX-2, it shows that the stability is higher in the COX-1/Apigenin system.

Table 9: Comparison of ligand/COX-1 system and ligand/COX-2 system

Ligand	Apigenin		Galangin		Kaempferol		Luteolin		Morin		Quercetin	
Enzymes	COX-1	COX-2	COX-1	COX-2	COX-1	COX-2	COX-1	COX-2	COX-1	COX-2	COX-1	COX-2
Bind energy	-8.88	-8.93	-7.26	-7.55	-8.24	-7.51	-7.79	-7.58	-7.71	-7.91	-6.65	-7.76
Amino acids	TRP 387	SER 530	TRP 387	TYR 355	TRP 387	SER 353	TRP 387	VAL 523	TRP 387	HIS 90	TRP 387	TYR 355
	VAL 349	ALA 378	VAL 349	VAL 523	HIS 388	VAL 523	VAL 349	TYR 355	LEU 390	SER 353	VAL 349	VAL 523
	SER 530	ILE 377	SER 530	HIS 90	HIS 386	HIS 90	SER 530	SER 530	PHE 210	VAL 349	SER 530	SER 530
	ILE 523	PHE 381	ILE 523	SER 353	LEU 390	LEU 352	ILE 523	LEU 352	ALA 202	LEU 352	ILE 523	TYR 385
	GLY 526	PHE 209	GLY 526	LEU 352	MET 391	TYR 355	LEU 384	TYR 385	GLN 203	SER 530	TYR 355	LEU 352
	ALA 527	PHE 205	MET 522	ALA 527	TYR 385	PHE 518	TYR 385	ALA 527	HIS 207	VAL 523	SER 353	SER 353
	LEU 384	VAL 228	ALA 527	PHE 518	ASN 382	ILE 517	MET 522	HIS 90	ASN 382	TYR 355	TYR 385	GLN 192
	LEU 531	LEU 534	LEU 531	ALA 516	ALA 202	GLN 192	ALA 527	ARG 513	TYR 385	MET 522	LEU 352	HIS 90
	MET 522	GLY 533	PHE 518	ILE 517	THR 206	ALA 516	ARG 120	TYR 348	THR 206	LEU 384	VAL 116	PHE 518
	LEU 352	ASN 375	ARG 120	GLN 192	HIS 207	GLY 526	VAL 116	SER 353	HIS 388	TYR 385	TYR 348	ALA 527
	ARG 120	PHE 529	LEU 384		PHE 210	PHE 381	PHE 381		HIS 386	TRP 387	ARG 120	
	TYR 355	LYS 532	TYR 355				PHE 518			PHE 518	MET 522	
	VAL 116	ILE 124					LEU 352				ALA 527	

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 29 May 2019 doi:10.20944/preprints201902.0124.v4

Stability Index SI

We have been established a stability index according to the energy values (bind energies) of each protein-ligand interaction. We quantified the force exerted by the interaction of the ligand with the active site of enzyme. The interaction index is defined as follows:

$$IE = \frac{E_i}{E_{minor}} \tag{1}$$

Where, E_i represent the bind energy of each protein-ligand interaction and E_{minor} is the minor bind energy of all protein-ligand interaction. The stability index is then normalized as follows:

$$NSI = \frac{(E_j - E_{minor})}{(E_{mayor} - E_{minor}} \tag{2}$$

Where, NSI is the normalized stability index, E_i represent the bind energy of each protein-ligand interaction, E_{minor} is the minor bind energy of all protein-ligand interaction and E_{higher} is the higher bind energy of all protein-ligand interactions.

Figure 8(a) represent the greater stability of Apigenin in the active site of COX-1, followed by Kaempferol, Luteolin, Morin, Galangin and Quercetin featuring the lowest stability. On the other hand, Figure 8(b) shows Apigenin as the more stable ligand in the active site of COX-2, followed by Morin, Quercetin, Luteolin, Galangin and Kaempferol, respectively. In the region of the ligands with the lowest stability, that is, 0.51, we find Luteolin, Morin, Galangin and Quercetin.

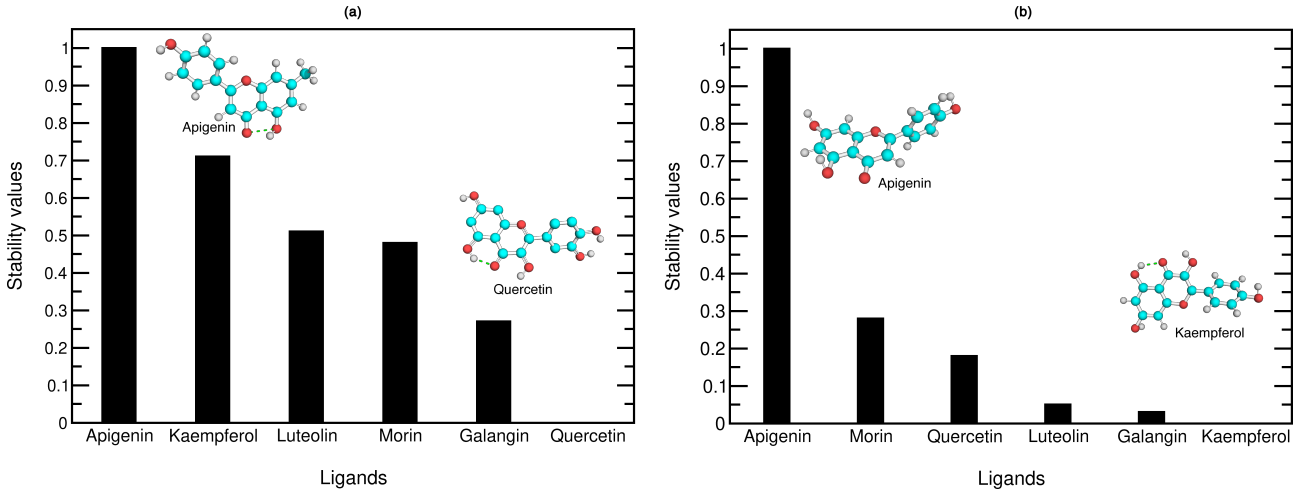

Figure 8: Normalized sStability index vs Ligand: (a) NSI for COX-1; (b) NSI for COX-2.

Quercetin have the lowest relative stability index of all ligands in complex with COX-1, while

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 29 May 2019 doi:10.20944/preprints201902.0124.v4

the same ligand has a higher stability index in complex with COX-2 (IE 0.18), which means that this ligand can be the structure with more energetic differences between both active sites.

On the other hand, Kaempferol have the higher stability index in COX-1 (IE 0.71) while this index is close to zero in COX-2. This suggest that Kaempferol is in a low energy state in the active site of COX-1. Table 10 shows a comparison of the energetic parameters and the formation of interactions between the ligands with IE values more representative.

Table 10: Energetic parameters and hydrogen bonds formation of Apigenin, Quercetin and Kaempferol.

Ligand	Apigenin		Quercetin		Kaempferol	
Enzymes	COX-1	COX-2	COX-1	COX-2	COX-1	COX-2
Heat of Formation kJ mol	-362.63	-235.67	-917.04	-813.97	-467.72	-469.34
Total Energy eV	-3307.05	-3449.55	-4043.43	-4042.41	-3745.36	-3745.36
Electronic Energy eV	-21346.54	-21448.13	-25951.83	-25836.36	-23672.59	-23641.45
Core-Core Repulsion eV	18039.48	17998.58	21908.39	21793.95	19927.22	19896.09
Ionization Potential eV	9.21	9.52	9.17	9.09	9.06	9.17
HOMO Energy eV	-9.21	-9.52	-9.17	-9.09	-9.06	-9.17
LUMO Energy eV	-1.36	-0.99	-1.21	-1.37	-1.53	-1.44

Figure 9 shows the interaction of Apigenin, Quercetin and Kamepferol with COX's, highlighting the relevant interacting AA's of active pockets and hydrogen bridges formed.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 29 May 2019 doi:10.20944/preprints201902.0124.v4

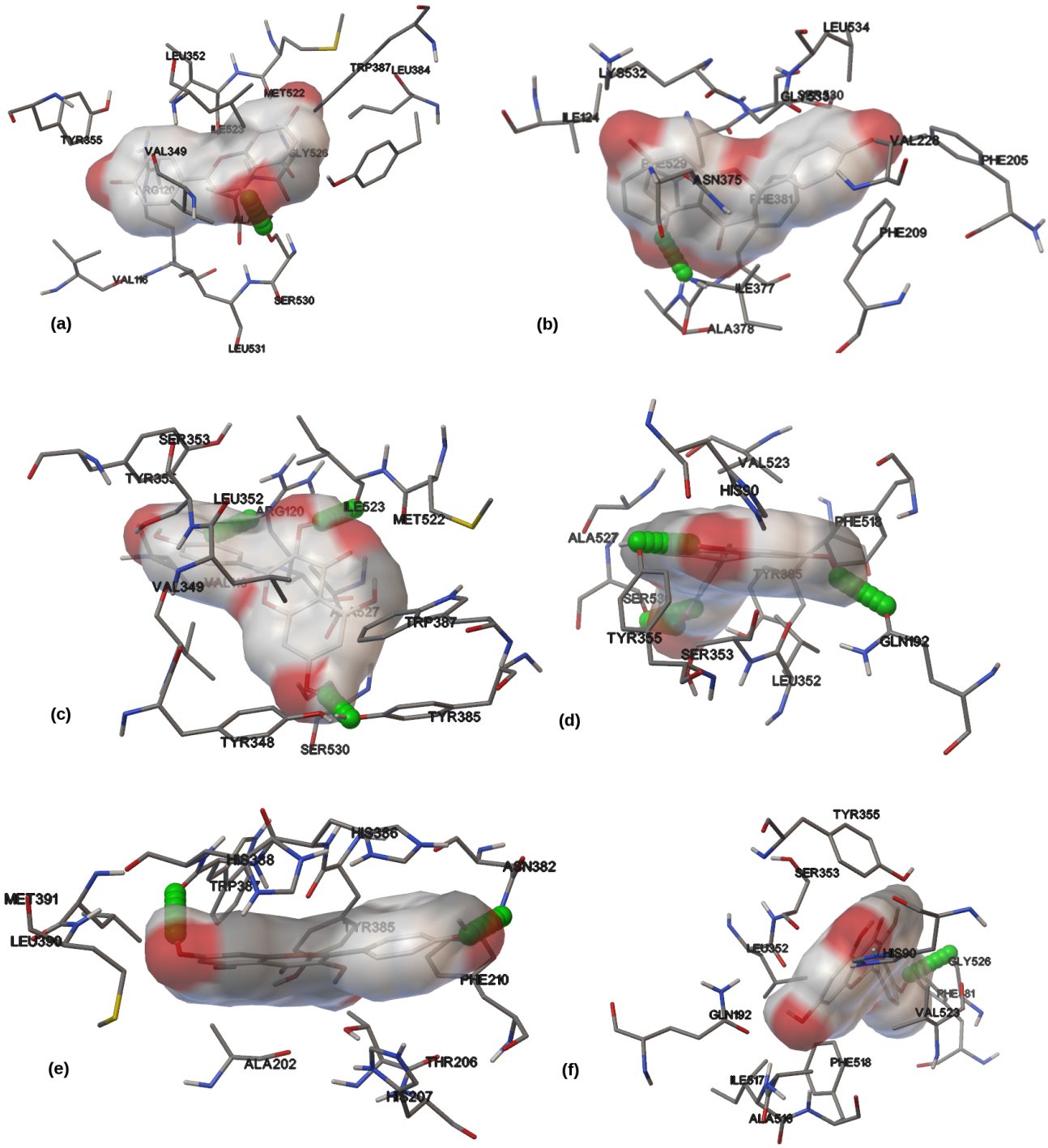

Figure 9: COX-Ligands complex. (a) COX-1/Apigenin system; (b) COX-2/Apigenin system; (c) COX-1/Quercetin system; (d) COX-2/Quercetin system; (e) COX-1/Kaempferol system; (f) COX-2/Kaempferol system. In all complexes, the ligand is surrounded by amino acids of the enzyme active site. In green the hydrogen bonds formed.

Ligand-Protein Selectivity Index of PLSI

The PLSI is a very important parameter to quantify the potential that some ligands have to inactivate a enzyme. This index shows the affinity for one or another isomeric form of COX in numerical and relative form.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 29 May 2019 doi:10.20944/preprints201902.0124.v4

The PLSI is calculated by dividing the binding energy of each COX-2/ligand interaction between the binding energy of each COX-1/ligand interaction, taking into account the correspondence between the ligands.

The PLSI is defined as follows:

$$IS = \frac{\frac{E_i,COX2 - E_{min},COX2}{E_{hig},COX2 - E_{min},COX2}}{\frac{E_i,COX1 - E_{min},COX1}{E_{hig},COX1 - E_{min},COX1}} \tag{3}$$

Quercetin was the ligand with a higher PLSI in the active sites of both enzymes, followed by Galangin, Morin, Apigenin, Luteolin and Kaempferol.

Figure 10 shows the PLSI. First, Quercetin have the higher value of selectivity with 1.17, while Kaempferol shows the lowest value with 0.91. Finally, we can notice that the difference in the PLSI values is not very marked among the ligands compared with the PLSI values, for which there are more marked differences.

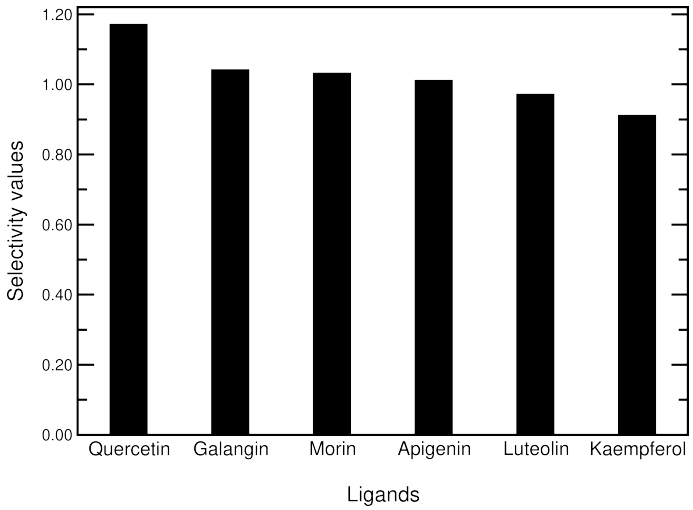

Figure 10: Selectivity index of ligands.

Conclusions

We have developed a procedure to comparatively quantify the degree of selectivity and stability of chemical species in contact with a biologically active proteic agent. We have established a normalized stability index (NSI) based on docking binding protein-ligand energies, thus quantifying their degree of electrostatic (strong) and van der Waals (weak) intermolecular interaction. The NSI for every flavonoid in our set is defined as the normalized ratio between the calculated binding energy of each possible ligand-protein pair and the difference between the limit values. Also, from the ob-

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 29 May 2019 doi:10.20944/preprints201902.0124.v4

tained NSI's, we were able to define a protein-ligand stability factor (PLSF). The PLSF measures the affinity of a given set of bio-active chemical species towards a competing set of proteic targets. In our specific case, we have a set of COX-2/COX-1 inhibition competing flavonoids. Our chosen set of flavonoid species has been quantum-chemically characterized, highlighting its more elctrostatically active zones, as seen in the electronic density surface maps calculated from semiempirical hamiltonians.

Our results suggest that inside the active site of the COX's there is a complex interplay between the hydroxil groups of the ligand and the side chains of the sorrounding AA's. These hydroxil groups, are a critical variable, considering the set of flavonoids studied. However, the data shows that the formation of hydrogen bridges was not the most relevant factor for ligand stability/selectivity inside the active site. Thus, adscribing the potential stability or selectivity of a chemical species in terms of the number of hydrogen bridges formed with lateral chains of AA's, which are within the sphere of strong electrostatic or weak van der waals influence, is not the correct assesment. Although hydrogen bridges are important to characterize the set of acting ligand-protein forces, they do not determine the coupling stability. Based on the stability/selectivity criteria devised in this work, we could infer the independency between both factors. That means, a highly stable chemical species inside an protein active site does not guarantee an equally high degree of selectivity, wich could be unambiguously extracted from the apigenin data, which shows a high degreee of stability which is not equaled for the COX-2/COX-1 PLSI. From our flavonoid set of bioactive chemicals, we could mark Quercetin as the best candidate for the spcific selective inactivation of the highly homologous isomers of COX, having an intermediate level of stability and a large selectivity potential towards COX-2.

Finally, we are extending the present hybrid *ab initio* / mechanical statistics methodology to study wider sets of chemically bio-active relevant species related to activation/inactivation of protein targets, in order to make a more sensible screening of candidates to be subjected to in-vitro or in-vivo experimentation.

 doi:10.20944/preprints201902.0124.v4

Supplementary Information SI

Acknowledgments

We gratefully acknowledged the Biomedical Research Group and the Research Group in Natural Products at the University of Sucre for its colaboration in the developemnt of this work.

References

1. Aldo F. Combariza, German Sastre, and Avelino Corma. *J. Phys. Chem. C.*, 115:875–884, 2011.
2. Aldo F Combariza, Diego A Gomez, and German Sastre. *Chem. Soc. Rev.*, 42:114–27, 2013.
3. Sergio Marti, Vicente Moliner, Juan Andres, Maite Roca, Violeta Lopez-canut, Estanislao Silla, Inaki Tunon, and Juan Bertran. *An. Quim.*, 107:144–153, 2011.
4. A F Combariza and G Sastre. *J. Phys. Chem. C.*, 115:13751, 2011.
5. Shaun M Lippow and Bruce Tidor. *Curr. Opin. Biotechnol.*, 18:305–311, 2007.
6. Eric Paquet and Herna L. Viktor. *Biomed Res. Int.*, 2015:1–18, 2015.
7. Leonardo G. Ferreira, Ricardo N. Dos Santos, Glaucius Oliva, and Adriano D. Andricopulo. *Molecules.*, 20:13384–13421, 2015.
8. I Levine, A Rodriguez, A Pascual, and J Roman. *Quantum Chemistry*. Pearson Education, New York, USA, 6th ed edition, 2009.
9. Johnny D Perez. *J. Trop. Eng.*, 24:117–127, 2014.
10. Douglas B. Kitchen, Helene Decornez, John R. Furr, and Jurgen Bajorath. *Nat. Rev. Drug Discov.*, 3:935–949, 2004.
11. F A Fitzpatrick. *Curr. Pharm. Des.*, 10:577–88, 2004.
12. Carol A Rouzer and Lawrence J Marnett. *J. Lipid Res.*, 50 Suppl:S29–34, 2009.

 doi:10.20944/preprints201902.0124.v4

13. Vittorio Limongelli, Massimiliano Bonomi, Luciana Marinelli, Francesco Luigi Gervasio, Andrea Cavalli, Ettore Novellino, and Michele Parrinello. *Proc. Natl. Acad. Sci. U. S. A.*, 107: 5411–6, 2010.

14. R M Garavito and D L DeWitt. *Biochim. Biophys. Acta.*, 1441:278–287, 1999.

15. Mona S Mohammed, Wadah J A Osman, Elrashied A E Garelnabi, Zuheir Osman, Bashier Osman, Hassan S Khalid, Magdi A Mohamed, and J A Osman. *J. Phytopharm.*, 3:275–285, 2014.

16. Miran Aswad, Mahmoud Rayan, Saleh Abu-Lafi, Mizied Falah, Jamal Raiyn, Ziyad Abdallah, and Anwar Rayan. *Inflamm. Res.*, 67:67–75, 2018.

17. Shashank Kumar, Abhay K Pandey, Shashank Kumar, and Abhay K Pandey. *Sci. World J.*, 2013: 1–16, 2013.

18. Hong-Jhang Chen, Cheng-Pei Chung, Wenchang Chiang, and Yun-Lian Lin. *Food Chem.*, 126: 1741–1748, 2011.

19. Ravi G Kurumbail, Anna M Stevens, James K Gierse, Joseph J McDonald, Roderick A Stegeman, Jina Y Pak, Daniel Gildehaus, Julie M Iyashiro, Thomas D Penning, Karen Seibert, Peter C Isakson, and William C Stallings. *Nature.*, 384:644–648, 1996.

20. Shuguang Yuan, H.C. Stephen Chan, and Zhenquan Hu. *Wiley Interdiscip. Rev. Comput. Mol. Sci.*, 7:e1298, 2017.

21. Rcsb pdb: Homepage. URL `https://www.rcsb.org/`.

22. David S Wishart, Yannick D Feunang, An C Guo, Elvis J Lo, Ana Marcu, Jason R Grant, Tanvir Sajed, Daniel Johnson, Carin Li, Zinat Sayeeda, Nazanin Assempour, Ithayavani Iynkkaran, Yifeng Liu, Adam Maciejewski, Nicola Gale, Alex Wilson, Lucy Chin, Ryan Cummings, Diana Le, Allison Pon, Craig Knox, and Michael Wilson. *Nucleic Acids Research*, 46(D1):D1074–D1082, 2018. ISSN 0305-1048.

23. Ranjinder S Sidhu, Jullia Y Lee, Chong Yuan, and William L Smith. *Biochemistry.*, 49:7069–7079, 2010.

24. ChemSpider: Search and share chemistry. URL `http://www.chemspider.com/`.

 doi:10.20944/preprints201902.0124.v4

25. Harry E. Pence and Antony Williams. *J. Chem. Educ.*, 87:1123–1124, 2010.

26. URL `http://www.chemspider.com/Chemical-Structure.4444100.html`.

27. URL `http://www.chemspider.com/Chemical-Structure.4444102.html`.

28. Brian Y. Feng, Brandon H. Toyama, Holger Wille, David W. Colby, Sean R. Collins, Barnaby C.H. May, Stanley B. Prusiner, Jonathan Weissman, and Brian K. Shoichet. *Nat. Chem. Biol.*, 4:197–199, 2008.

29. URL `http://www.chemspider.com/Chemical-Structure.4444989.html`.

30. URL `http://www.chemspider.com/Chemical-Structure.4444935.html`.

31. URL `http://www.chemspider.com/Chemical-Structure.4444395.html`.

32. Gabriel. Cuevas and Fernando Cortes. *Introduccion a la quimica computacional*. Fondo de Cultura Economica, Mexico, 1ra ed. edition, 2003.

33. James B. Foresman and AEleen Frisch. *Exploring Chemistry With Electronic Structure Methods*. Gaussian, Inc., Wallingford, CT, 3th ed. edition, 2015.

34. H. Bernhard Schlegel. *Wiley Interdiscip. Rev. Comput. Mol. Sci.*, 1:790–809, 2011.

35. Jingjing Zheng and Michael J. Frisch. *J. Chem. Theory Comput.*, 13:6424–6432, 2017.

36. H. Bernhard Schlegel. *J. Comput. Chem.*, 24:1514–1527, 2003.

37. Stewart Computational Chemistry: MOPAC Home Page. URL `http://openmopac.net/`.

38. James J. P. Stewart. *J. Comput. Aided. Mol. Des.*, 4:1–103, 1990.

39. Donald B. Boyd, David W. Smith, James J. P. Stewart, and Erich Wimmer. *J. Comput. Chem.*, 9: 387–398, 1988.

40. James J. P. Stewart. *J. Mol. Model.*, 10:155–164, 2004.

41. Michael J. S. Dewar and Walter Thiel. *J. Am. Chem. Soc.*, 99:4899–4907, 1977.

42. AutoDock Vina: molecular docking and virtual screening program. URL `http://vina.scripps.edu/`.

 doi:10.20944/preprints201902.0124.v4

43. Oleg Trott and Arthur J. Olson. *J. Comput. Chem.*, 31:NA–NA, 2009.

44. Garrett M. Morris, David S. Goodsell, Robert S. Halliday, Ruth Huey, William E. Hart, Richard K. Belew, and Arthur J. Olson. *J. Comput. Chem.*, 19:1639–1662, 1998.

45. Jan Fuhrmann, Alexander Rurainski, Hans-Peter Lenhof, and Dirk Neumann. *J. Comput. Chem.*, 31:NA–NA, 2010.

46. D Picot, P J Loll, and R M Garavito. *Nature.*, 367:243–249, 1994.

47. D L DeWitt, E A El-Harith, S A Kraemer, M J Andrews, E F Yao, R L Armstrong, and W L Smith. *J. Biol. Chem.*, 265:5192–8, 1990.

48. Marc Lecomte, Odette Laneuville, Chuan Ji, David L DeWitt, and William L Smith. *J. Biol. Chem.*, 269:13207–13215, 1994.

49. J A Mancini, D Riendeau, J P Falgueyret, P J Vickers, and G P O-Neill. *J. Biol. Chem.*, 270: 29372–29377, 1995.

50. C J Rieke, A M Mulichak, R M Garavito, and W L Smith. *J. Biol. Chem.*, 274:17109–17114, 1999.

51. Afshin Zarghi and Sara Arfaei. *Iran. J. Pharm. Res.*, 10:655–683, 2011.