Screening and docking molecular studies of natural products targeting overexpressed receptors HER-2 in breast cancer

Abstract

The first cancer to strike a community is breast cancer. Because of its extremely high mitotic activity, breast cancer that tests positive for HER 2 is thought to have a bad prognosis. Due to the effects caused by chemical drugs, patients are increasingly turning to natural medicine, such as phytotherapy and nutritherapy. The main objective of this study is to search, using a bioinformatics approach (molecular docking), for new non-toxic anti-cancer inhibitors by carrying out a screening of 102 ligands from natural and dietary compounds, likely to interact with the HER-2.The results of the virtual screening permit to choose 23 best compounds which can be proposed as the best inhibitors of HER-2. Lycopene would be a very promising ligand which presents a DeltaG of -9.82 kcal/mol, followed by Beta-carotene (DeltaG of -8.58), P-cumaric acid kcal/mol (DeltaG of -8.57) and Curcumin (DeltaG of -8.46). Another compounds; luteolin, anacardium (Anacardic acid) and alpha-Tocopherol were found to have the strongest inhibitory effects, with DeltaG values of -7.92 kcal/mol, -7.89 kcal/mol and-7.85 kcal/mol, respectively, and act directly on residues keys found in the hydrophobic pocket II (ATP binding site) and the hydrophobic region (the αC-β4 loop) of the EGFR domain. Pinoresino, Kaempferol and Caffeic acid with DeltaGs of -7.48 Kcal/mol, -6.88 Kcal/mol and -6.34 kcal/mol, and are three ligands specific to the conserved regions of the HER-2 receptor and interact with the tail respectively; C-terminal, the C-lobe activation loop and the N-lobe P loop of the tyrosine kinase domain. The comparison of Lapatinib (chemical compound) and **quercetin** (natural compound) have respectively DeltaG of -7.58 kcal/mol and -7.28 kcal/mol, form a hydrogen bond with the same residue of the hydrophobic region. All the natural molecules seem very promising and, after *in vitro/in vivo* tests, could constitute good substitutes for the chemotherapies currently used to treat breast cancers as well as other cancers.

Keywords: Brest cancer, HER-2, Molecular docking, Natural compounds

1. Introduction

After lung cancer, breast cancer is the most common malignancy in women and the second leading cause of cancer-related deaths (**1**). The World Health Organization (WHO) announced in early 2020 that the incidence of breast cancer is rising in developing nations as a result of rising life expectancies, increased urbanization, and adoption of western lifestyles. It is estimated that 627,000 women died of breast cancer in 2020, accounting for 15% of all female cancer deaths.

Estrogen and progesterone hormone receptor dysfunction is typically associated with breast tumors (**2, 3**) Furthermore, a great deal of research has been done on the overexpression of the human epidermal growth factor receptor 2 (HER2), EGFR1 (overexpression of the epidermal growth factor receptor), and PI3Ka (dysregulation of the ER+ and ER−) signaling pathways in breast cancer (**3**,**4)** Therefore, it is essential to continue discovering novel techniques and compounds that target these proteins. Roughly 20% to 25% of all breast cancers were caused by the transmembrane protein receptor known as human epidermal growth factor receptor 2 (HER2), which is encoded by the HER2 gene located on chromosome 17 long arm. The EGFR family, which consists of the four HER receptors HER4, HER3, HER2, and HER1, includes HER2 (**5**). Specific tyrosine kinase residues are phosphorylated and signaling proteins are activated upon HER2 receptor activation, which leads to the start of downstream signaling processes. Apoptosis, angiogenesis, cell proliferation, and survival are all regulated by the critical pathways induced by the HER2 receptor, which include the mitogen-activated protein kinase (MAPK) and phosphatidylinositol triphosphate kinase (PI3K) signaling mechanisms (**6**). In HER2+ breast cancers, overexpression of the HER2 receptor is known to be a HER2 activation mechanism. HER2-positive breast cancer remains a case study to this day. It is considered a cancer with a poor prognosis due to its high mitotic activity and ability to metastasize more easily. However, improved molecular genetic techniques have made it possible to study resistance to the administration of trastuzumab (HERCEPTIN) and generate new anti-HER2 targeted therapies. The monoclonal antibody pertuzumab and the tyrosine kinase inhibitor lapatinib specifically target HER2 receptors. The adverse effects of these two chemical and synthetic drugs include alopecia, nausea, vomiting, fatigue, fever, infection, diarrhea, muscle pain, paresthesia, cognitive disorders, cardiotoxicity, leukemia, gastrointestinal and dermatological reactions. Another several market drugs such as tamoxifen,

raloxifene, toremifene, and fulvestrant for the treatment of breast cancer are available, but each has its limitations, which cause irreversible side effects (**7**). The recent search is on for other, less toxic and natural molecules. Patients are therefore increasingly turning to natural medicine, such as phytotherapy and nutritherapy. There are various benefits to using natural products in the food and medicine development industries, such as their superior chemical diversity, biological potency, and structural complexity and optimize the regulation of natural product biosynthesis. Therefore, this study aimed to discover, *in silico*, a more selective natural compound targeting breast cancer for using as a therapeutic agent.

2. Material and Methods

2.1. Preparation of the protein

The crystal structure of the kinase domain of human HER2 was obtained from Protein Data Bank (PDB) (https://www.rcsb.org) (**8**), with PDB ID: 3PP0 (**9**). The structure was downloaded in .pdb format and was further prepared for the docking process.

2.2. Preparation of ligands

After an extensive literature search, 102 molecules that could have a positive interaction with the ErbB2 receptor tyrosine kinase domain in HER2+ breast cancer and that are derived from plants, microorganisms or food sources were selected. The ligand codes were obtained from PubChem (https://pubchem.ncbi.nlm.nih.gov/)(**10**), and Zinc Database (https://zinc.docking.org/) (**11**).

2.3. Pharmaco-toxicity study of ligands

In order to test the toxicity of the plant-derived molecules, we used the PKCSM Database online server (http://structure.bioc.cam.ac.uk/pkcsm) (**12**). The ligand codes obtained from PubChem and Zinc Database were copied to pkCSM Database to eliminate toxic ligands based on the following criteria : AMES toxicity, hERG K^+ channel inhibitors toxicity and Hepatotoxicity.

2.4. Molecular docking

Molecular docking was performed by the SwissDock server (http://www.swissdock.ch/) (**13**), which allows importing the target molecule "tyrosine kinase domain of the HER2 receptor" and the ligands in purpose of testing their interactions, with the aim of studying the interactions of

these two molecules. Investigating the outcomes enables the identification of the binding energy, the hydrogen bonds formed as well as the amino acids involved in these interactions.

2.5. Docking results visualisation

The visualisation of the molecular docking results from the SwissDock server is done using the UCSF Chimera software (https://www.cgl.ucsf.edu/chimera/) (**14**) .

3. Results and discussion

3.1. Ligand toxicity analysis

Initially, 102 natural compounds were obtained from the databases. These molecules can be found in food sources ; in plants or microorganisms. Some compounds were screened for their toxicity based on the predicted results of the mutagenicity screening (AMES Toxicity), HERG K+ channel inhibitor toxicity and hepatotoxicity. The results revealed that eight ligands are potentially toxic (**Table 1**). This toxicity could be skept by decreasing the administrated dose, respectively 0.558 log mg/kg/day, 0.36 log mg/kg/day, 0.654 log mg/kg/day, 0.144 log mg /kg/day, 0.82 log mg/kg/day for the ligands Genipin, Sauchinone , Denbinobin, Xenognosin and Kaempferol.

Table 1: Toxicity parameters of some compounds.

Molecules	Mutagenicity	K+ hERG 1	Na + hERG 2	Hepatotoxicity
Genipin	Yes	No	No	No
Sauchinone	Yes	No	No	No
Denbinobin	Yes	No	No	No
Furanodiene	No	No	No	No
Chalcones	No	No	No	No
Isoliquiritoside	No	No	No	No
Xenognosin	Yes	No	No	No
Kaempferol	Yes	No	No	No
Luteolin	Yes	No	No	No

Silibinin	Yes	No	Yes	No
Daidzeine	Yes	No	Yes	No

3.2. *Interaction of ligands with HER2*

In this study, the interaction of the 3PP0 protein against 84 ligands was investigated. The results are shown in Tab.2.

Table 2: docking results of total ligands with the tyrosine kinase domain obtained by Swissdock.

Ligands	Reesidue (s) target(s)	Length of hydrogen bong (A°)	Interraction energy (kcal/mol)
Crocetin	ALA 706 ALA 706	3.011 3.324	-9.46
Secoisolariciresinol diglucoside	GLU 757	1.841	-8.41
Lycopene	----------	No hydrogen bond	-9.82
P-coumaric acid	----------	No hydrogen bond	-8,57
Curcumin	----------	No hydrogen bond	-8.46
Pomiferin	----------	No hydrogen bond	-8.08
Formononetin	----------	No hydrogen bond	-8.07
Rosmarinc acid	----------	No hydrogen bond	-7.83
3,4-Dihydroxybenzoic	GLY 737	1.915	-6.22
4 P-hydroxybenzoique 1 acid	SER 779	2.213	-6.88
4 P-hydroxybenzoique 2 acid	----------	No hydrogen bond	-6.21
Gallic acid	VAL 777	2.517	-6.25
Gentisic acid	----------	No hydrogen bond	-6.93
Syringic acid	CYS 805	3.096	-6.51
Vanillic acid	VAL 777	2.191	-6.20
Catechines	VAL 777	3.187	-6.82

Epicatechines	GLN 943	1.819	-6.74
Biochanine A	LEU 726	3.347	-6.67
Ergostane	----------	No hydrogen bond	-6.18
Glycitein	----------	No hydrogen bond	-6.37
Daidzeine	VAL 777	2.114	-6.61
Genistein	CYS 947	1.995	-7.54
Malvidin	VAL 777	3.031	-7.32
Delphinidine	----------	No hydrogen bond	-6.70
Cyanidin	GLN 709 GLN 709	2.003 1.929	-7.00
Acetoxypinoresinol	----------	No hydrogen bond	-7.29
Pinoresinol	SER 728	2.523	-7.80
Hydroxytyrosol	ARG 849	2.215	-6.91
Tyrosol	ARG 849	2.129	-6.36
Secoisolariciresinol	----------	No hydrogen bond	-7.49
Enterodiol	PRO 945	2.169	-6.63
Enterolactone	----------	No hydrogen bond	-7.51
Capsaicin	----------	No hydrogen bond	-7.62

CAPE Caffeic-acid-phenethyl ester	----------	No hydrogen bond	-6.46
α-Linolenic acid	SER 728	3.160	-7,69
ChlorogeniqueHeriguard acid	GLN 943	2.140	-6.96
Ferulic acid	----------	No hydrogen bond	-7,37
Gingerol	----------	No hydrogen bond	-7,27
Petunidin	----------	No hydrogen bond	-7,03
Pelagronidine	GLY 778	2,052	-6,38
Homocastasterone	ASP838	2.716	-7.44
Cafeic	LEU1000	2.044	-6.71
Sinapic	CYS 805	2.052	-7.40
3-Hydroxybenzoic	---------	No hydrogen bond	-6.68
O-coumaric acid	----------	No hydrogen bond	-6.05
Diindolylmethane	----------	No hydrogen bond	-7,31
Naringenine	GLN943	1.841	-6.65
Indol 3-carbinol	----------	No hydrogen bond	-6.13
Kaempferol	ASP 863	3.572	-6.88

Dihydroresveratrol	---------	No hydrogen bond	-7.15
Resveratrol	----------	No hydrogen bond	-6.51
Sulforaphane	MET 801	3.708	-7.10
Myricetin	---------	No hydrogen bond	-7.03
Quercetin	ALA706 VAL 777	2.130 3.491	-7.28
Apigenin	CYS 947	2.064	-7.55
Luteolin	ALA 706 VAL 777	2.132 2.247	-7.92
Fisetin	VAL777	3.336	-6.55
Sauchinone	----------	No hydrogen bond	-6.83
Denbinobin	----------	No hydrogen bond	-7.01
Furanodiene	----------	No hydrogen bond	-6.82
Chalcone	----------	No hydrogen bond	-7.19
Lupane	----------	No hydrogen bond	-6.77
Genipin	VAL 777 VAL 777	2.696 3.142	-6.50
Opium	VAL777	2.054	-6.70
Pyocyanin	MET801	2.162	-7.65
Ginsenol	ALA706	2.131	-6.34
Menthol	VAL777	2.329	-6.10

Urocanic acid	MET801	2.182	-6.49
Anthranilic acid	ALA706 ALA706	2.410 2.177	-7.54
Anacardic acid	CYS805	2.049	-7.89
Diosmetin	VAL777 GLN709	2.073 2.348	-6.60
Khahalalide D	THR759	2.385	-7.39
Alpha-tocopherol	MET801	2.433	-7.85
Beta-carotene	----------	No hydrogen bond	-8.58
Choline	CYS802	2.442	-6.30
Sesamol	MET801	2.612	-6.09
Silibinin	VAL777 LEU711	2.197 2.546	-7.16
Xanthoxylin	VAL777	2.509	-6.26
Isofraxidin	MET801 VAL777	2.084 2.117	-7.77
Phloretic acid	MET801 VALL777	1.939 2.034	-7.06
Indole-3-carboxylic acid	GLN990 PHE731 GLN990	2.118 2.003 2.077	-7.81
Garlic	----------	No hydrogen bond	-6.69
Xenognosin	MET801	2.156	-6.86

3.3. Ligands interacting with conserved residues of the tyrosine kinase domain of EGFR family receptors.

It note that there are 47 complexes formed between the ligands and the tyrosine kinase domain of 3PP0 that have the lowest score energies compared to the other ligands by forming hydrogen bonds with essential residues conserved in the EGFR family (see Table 2).

Therefore, the best ligand according to the interaction energy is represented by Lycopene **Tab.2.** As for Lycopene from tomato, it presents an interaction energy of -9.82 kcal/mol with no hydrogen bonds were predicted by SwissDock suggesting the existence of other types of bonds (**Fig.1).**

(A) (B)

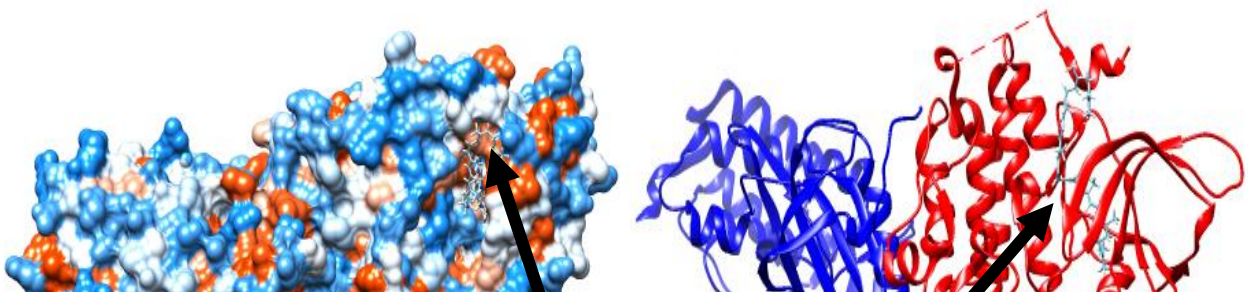

Figure.1 : Three-dimensional illustration of the 3PP0-Lycopene complex using the molecular surface (A) and the ribbon model (B)

Taking into account what has been cited in the literature including Met801 which is located in the Adenine region of the ATP binding site and Cys805 of hydrophobic pocket II, it has noted that some ligands form interactions with the ATP binding site (**Tab.2) (15).**

Compunds establishes a hydrogen bond with the residue Met 801 in the adenine region are as follow:

-Alpha-Tocopherol from sunflower oil has an interaction energy of –7.85 kcal/mol.

-Isofraxidinde of the species *Eleutherococcus senticosus* known as Siberian ginseng has an interaction energy of –7.77 kcal/mol**.**

-**Pyocyanin:** a blue green phenazine molecule, produced specifically by the bacterium *Pseudomonas aeruginosa*, has an interaction energy of -7.65 kcal/mol.

-**Sulforaphane**, mainly found in broccoli and cabbage., has an interaction energy of -7.10 kcal/mol.

-**Phloretic acid** belongs to the class of organic compounds present in peanuts and avocados has an interaction energy of –7.06 kcal/mol.

-**Xenognosin** present in common peas *Pisum sativumet* legumes, belongs to the class of organic compounds, presents an interaction energy of –6.86 kcal/mol.

-**Urocanic acid** essentially found in the fungus *Hippospongia communis* has an interaction energy of –6.49 kcal/mol.

-Sesamol of the sesame seed has an interaction energy of –6.09 kcal/mol.

The other compunds forming hydrogen bonds with cysteine residue 805 in the hydrophobic pocket II are: **Syringic acid** from olive oil and **Anacardic acid**, component of cashew nuts, present interaction energies (DeltaG) of -6.52 and -7.89 kcal/mol respectively.

These ligands interacting and forming hydrogen bonds with residues Met801 and Cys805 in the ATP binding site in the afore mentioned regions may have the potential to be competition inhibitors by blocking the access of ATP to its specific site on the tyrosine kinase domain **(16, 17).** In this study the ligand with the best interaction energy value with the ATP binding site is **Anacardic acid** with DG of -7.89 kcal/mo **(Fig.2).**

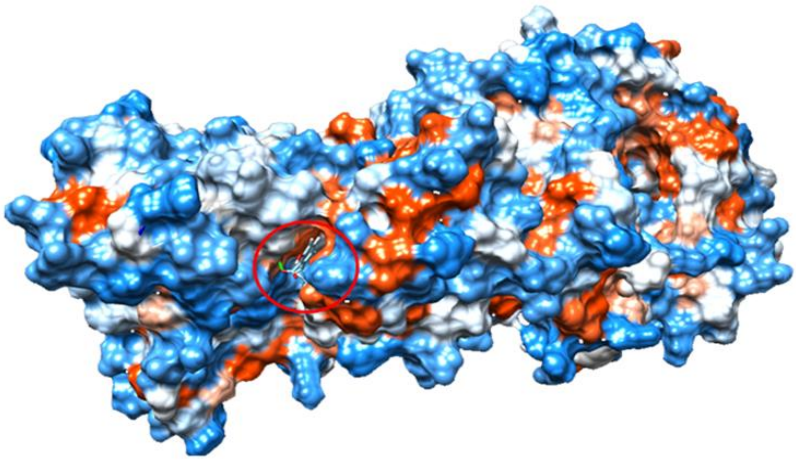

Figure.2: Three-dimensional illustration of the 3PP0 -Anacardic acide complex using the molecular surface.

Two olive oil ligands among the 21 ligands act on the N-lobe of the tyrosine kinase domain were selected according to their interactions with residues in the (Cα) helix span residues (729-744)~~.~~ **3,4-Dihydroxybenzoic acid** exhibits an interaction energy of -6.22 kcal/mol and establishes a hydrogen bond with the residue Gly737**. Sinapic acid** has an interaction energy of -7.40 kcal/mol and establishes a hydrogen bond with the residue Gly732.

Both ligands form hydrogen bonds, **3,4-dihydroxybenzoic acid** with residue Gly737and **Sinapic acid** with residue Gly732 in the (Cα) helix extent. This could destabilize the active open conformation of the activation loop (C-helix-in-DFG-in), no binding of the ATP substrate on its specific site, blockage of trans-autophosphorylation of the activation loop which maintains its inactive conformation, the ATP binding groove is no longer accessible, the tyrosines of the C-terminal tail will not be phosphorylated by the catalytic loop. This induces blockage of the downstream signaling cascade **(18).**

In the present study, the target with the best interaction energy value acting on the N-lobe (Cα-helix) is **Sinapic acid** with a DG of -7.40 kcal/mol (**Fig.3).**

(A) (B)

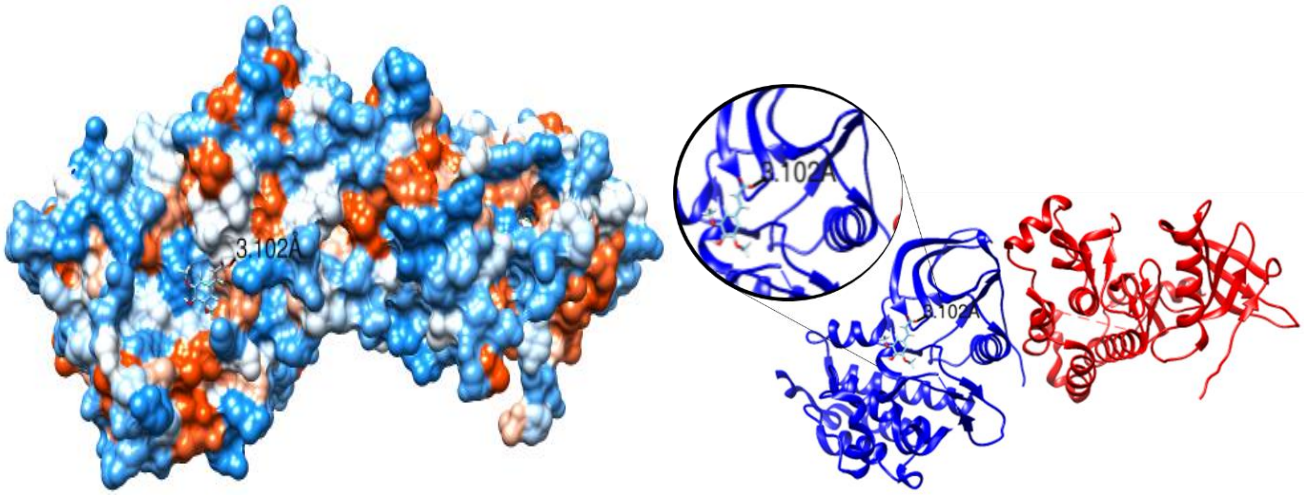

Figure.3: Three-dimensional illustration of the 3PP0 -Sinapic acide complex using the molecular surface (A) and the ribbon model (B)

Three of these 47 ligands act on the A-loop of the C-lobe of the tyrosine kinase domain. According to the literature, the 20-30 residue sequence of the activation loop (Asp831-Val852 in the EGFR family), which contains the conserved base DFG motif (Asp831-Phe-Gly833 in the EGFR family) and extends to an APE (Ala-Pro-Glu) motif, also contains the Tyr845 residue which is one of the target tyrosines for autophosphorylation by the catalytic loop (**16, 17, 18).**

-**Tyrosol** and **Hydroxytyrosol** of olive oil have interaction energies of -6.36 and -6.91kcal/mol respectively and each establishes a hydrogen bond with an estimated length of 2.129Å for tyrosol and 2.215Å for hydroxytyrosol with the same residue Arg849 of the chain A for tyrosol and B for hydroxytyrosol.

-**Homocasterone** from beans has an interaction energy of -7.44 kcal/mol establishing a strong hydrogen bond with an estimated length of 2.716Å with Asp838 of the A chain.

Thus the ligand **Homocasterone** forms a hydrogen bond with the residue Asp838, the ligands Tyrosol and Hydroxytyrosol, each form a hydrogen bond with the residue Arg849 of the activation loop. These ligands can therefore block its activation, its passage between the conformations, inactive (C-helix-out-DFGout) to the partially inactive conformation (C-helix-in-DFGout) and finally to the active conformation (C-helix-in-DFGin). This effect is possible by sequestering the trans autophosphorylation of these tyrosine residues to phosphotyrosines by the catalytic loop after dimerization. Thus the ATP binding groove is not accessible, no trans phosphorylation of the C-terminal tail and therefore no recruitment of adaptor proteins **(16, 17, 18, 19, 20).**

This stdudy shows that the ligand with the best interaction energy value that acts on loop A and lobe-C is **Homocasterone** with a DG of -7.44 kcal/mol (**Fig.4).**

(A) (B)

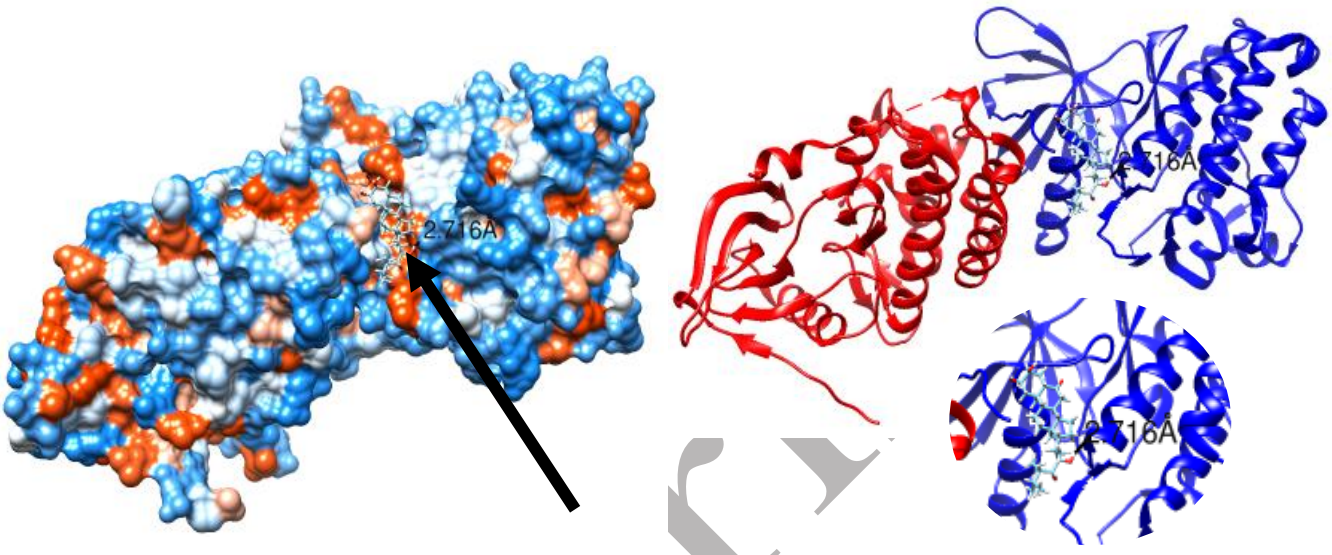

Figure.4: Three-dimensional illustration of the 3PP0 – Homocasterone complex using the molecular surface (A) and the ribbon model (B)

Other ligands were selected according to their interactions with residues in the hydrophobic region of ErbB2 which includes the following residues: Val773, Met774, Gly776, Val777, Gly778, and Val782 in the αC-β4 loop **(21)**. For exemple**, Pelagronidin** from grapes establishes a hydrogen bond with Gly778.

*Sixteen of these ligands establish a hydrogen bond with the same Val777 residue (**Tab.2) :** ***Vanillic acid*** from olive oil , ***Catechin*** from blackcurrant, ***Malvidin*** from grapes, ***Fisetin*** from strawberries and apples,***Genipin*** of the *Gardenia Jasminoides Ellis* species ,***Opuim***,.***Gallic acid***, ***Menthol*** from peppermint and tea, ***Diosmetin*** from sage,thyme, **Quercetin** from red onions or buckwheat, ***Luteolin*** from green peppers, olive oil and carrots, ***Xanthoxylin*** from fats and oils, herbs and spices, **Isofraxidin Phloretic acid**, ***Silibinin*** extracted from the milk thistle flower and ***Daidzein*** present in flax seeds,. Although these two ligands have a better interaction energy and strong bonds but according to the Pharmaco-toxicity test of the ligands showed that these 2 molecules have a mutagenic power therefore are toxic molecules.

Indeed these ligands forming hydrogen bonds with the residues (valine777) which belong to the hydrophobic region of ErbB2 (αC-β4 loop) and which interact with the activation loop, can destabilize these conformational changes, of the inactive conformation (C- helix-out-DFGout), to the partially inactive conformation (C-helix-in-DFGout) and finally to the active conformation (C-helix-in DFGin). This by sequestering its trans autophosphorylation by the catalytic loop after dimerization. Thus the ATP binding groove remains covered, no trans autophosphorylation of the C-terminal tail and therefore no recruitment of adapter proteins **(21).** This results indicated that the ligand which has the best interaction energy value interacts with the residues of the hydrophobic region in the loop (αC-β4) is **Luteolin** with a DG of -7.92 kcal/mol (**Fig.5**).

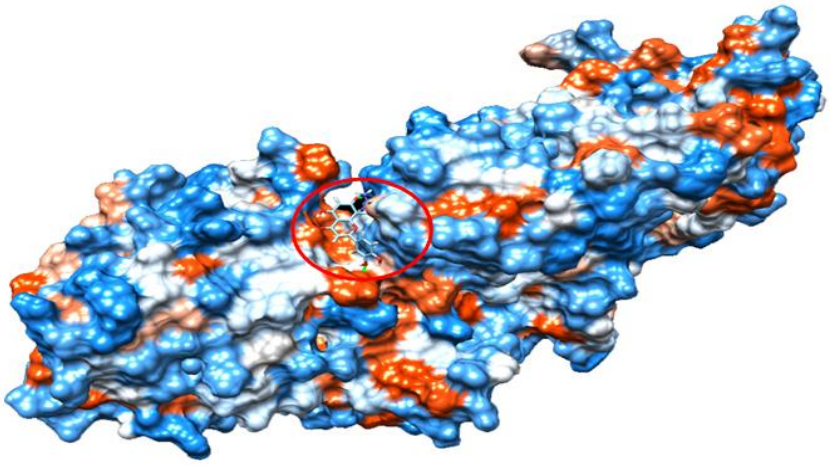

Figure.5. Three-dimensional illustration of the 3PP0 -Luteolin complex using the molecular surface.

According to the literature the residues of the hydrophobic region of loop A are Iso861, Thr862, Phe864, Leu866 and Leu869. Only the shared interaction between the two active and inactive conformations that takes place between Ser783 with in the hydrophobic region of ErbB2 (Cα-β4 loop) and residue Iso861 of loop A allows the interaction between the latter and the αC-β4 loop **(21)**.

Thus the ligands **Vanillic acid, Catechins, Malvidin, Fisetin, Genipin, Pelagronidin and 4 p-hydroxybenzoic acid**, forming hydrogen bonds with residues belonging to the hydrophobic region of ErbB2 (Cα-β4 loop) and interacting with the activation loop, can destabilize these conformational changes from the inactive conformation (C-helix-out-DFGout), to the partially inactive conformation (C-helix-in-DFGout) and finally to the active conformation (C-helix-in-DFGin), this by sequestering its trans autophosphorylation by the catalytic loop after

dimerization. Thus the ATP-binding groove remains covered with no trans autophosphorylation of the C-terminal tail and so no recruitment of adaptor proteins **(21)**.

In this study the ligand that has the best interaction energy value interacts with residues in the hydrophobic region (in the Cα-β4 loop) is **malvidin** with a DG of -7.32 kcal/mol (**Fig.6**).

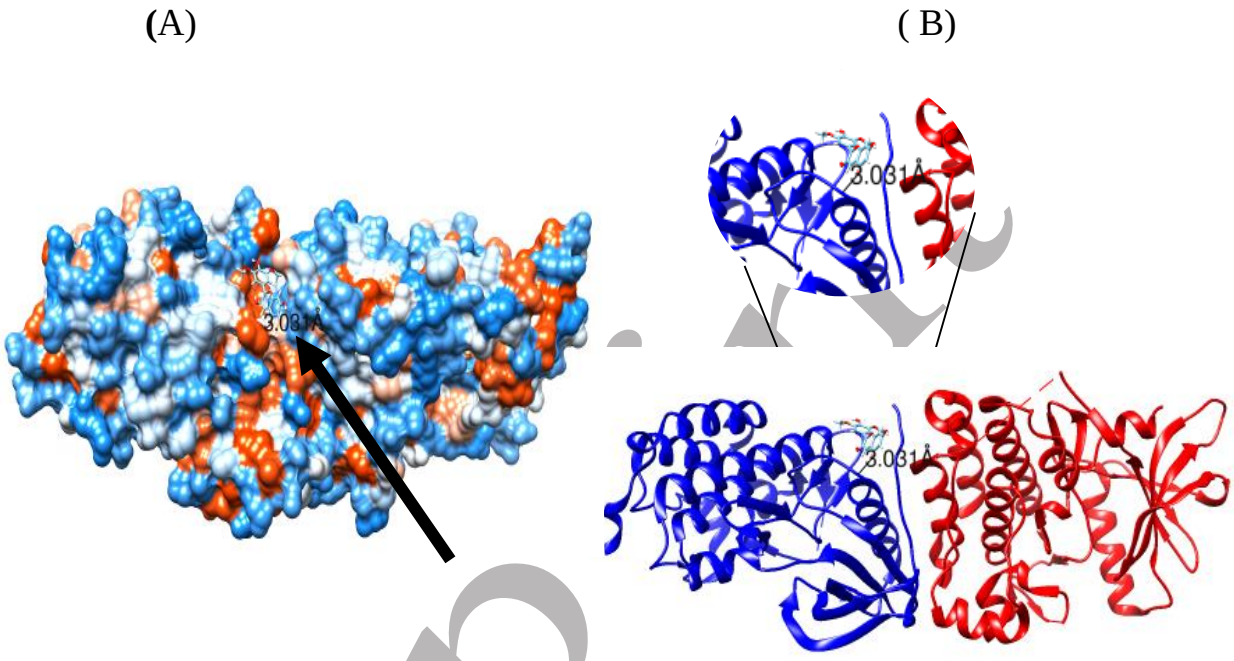

Figure 6: Three-dimensional illustration of the 3PP0 -Malvidin complex using the molecular surface (A) and the ribbon model (B)

The remaining five ligands were selected according to their interactions with residues that lie between the tyrosines of the C-terminal tail (Tyr874, Tyr992, Tyr1048, Tyr1068, Tyr1086, Tyr1101 and Tyr1173) (**Tab.2) (17).** The previously mentioned tyrosines correspond to residues trans-autophosphorylated by the catalytic loop so ligands that form hydrogen bonds, **Epicatechin** and **Naringenin** with residue Gln943, **Apigenin** with residue Cys947, **Genistein** with residue Cys947, that lie between these tyrosines may be susceptible to sequesterat the interaction of the C-terminal tail with the catalytic loop. Therefore the tyrosines will not be trans-autophosphorylated and thus no recruitment of the adaptor proteins **(17).**

The ligand that has the best interaction energy value with residues that lie between the tyrosines of the C-terminal tail is **Apigenin** with a DG of -7.55 kcal/mol (**Fig.7).**

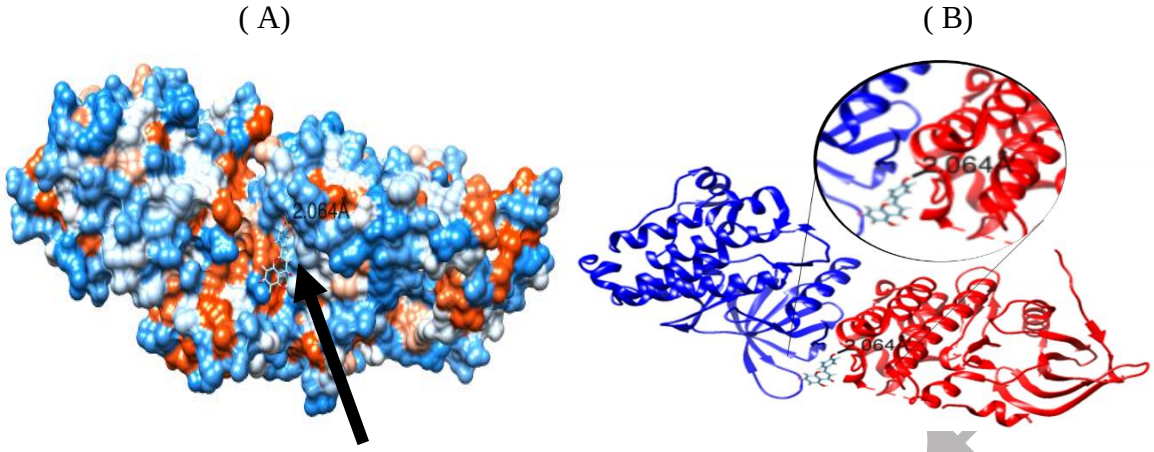

Figure 7 : Three-dimensional illustration of the 3PP0- Apigenin complex using the molecular surface(A) and the ribbon model (B)

3.4. Ligands interacting with specific residues of the HER-2 receptor tyrosine kinase domain.

One ligand acts on the P-loop of the N-lobe of the tyrosine kinase domain non-conserved in the EGFR family and specific for HER2 were selected in relation to their interaction with the residues in the extent (from residue Leu726 to residue Val734) **(9). Pinoresinol** from olive oil has an interaction energy of -7.48 kcal/mol and establishes a strong hydrogen bond whose length is estimated at 2.613 Å with the residue Ser728 (**Fig.8**). This ligand establishes a hydrogen bond with the P-loop, which can destabilize the open active conformation of the activation loop (C-helix-in-DFGin), not binding the ATP substrate to its specific site. This induces the blocking of the downstream signaling cascade **(16)**.

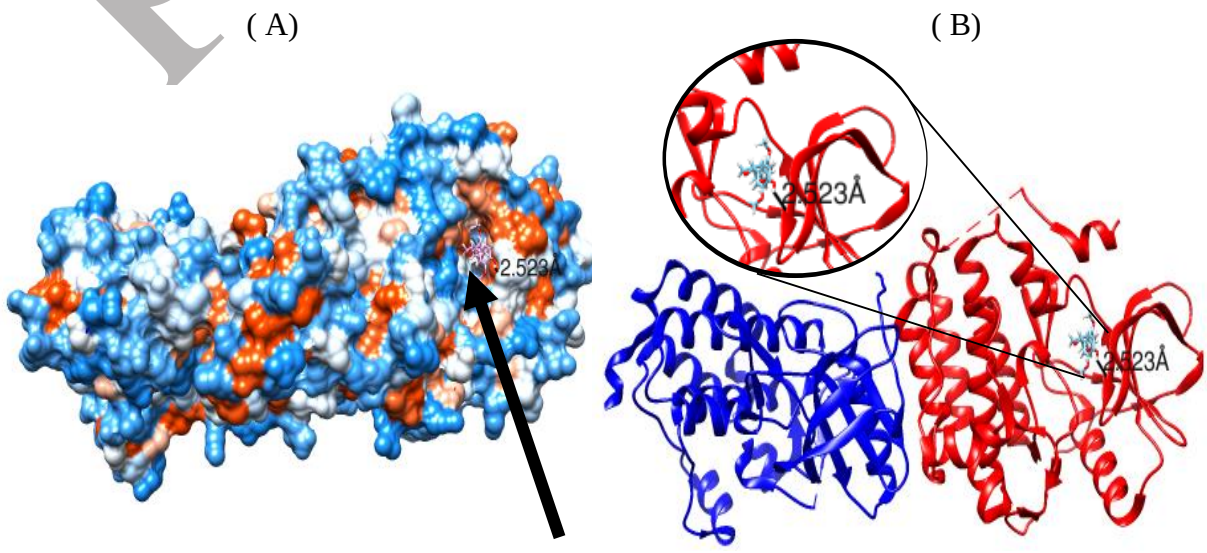

Figure 8. Three-dimensional illustration of the 3PP0-Pinoresinol complex using the molecular surface (A) and the ribbon model (B)

One ligand acts on the C-terminal tail of the non-conserved tyrosine kinase domain in the EGFR family and specific for HER2 was selected based on its interaction with the residues in the extended (from residue Pro999 to residue Leu1009) : **The caffeic acid** of olive oil has an interaction energy of -6.71 kcal/mol established a hydrogen bond whose length is estimated at 2.044Å with the residue **Leu1000 (9).**

There were two ligands, acting on the P-loop of the N-lobe of the tyrosine kinase domain selected according to their interaction with residues in the sequence extending from residue Leu726 to residue Val734 **(9). Biochanin A** from Soybean has an interaction energy of -6.67kcal/mol with Leu 726 **and α-Linolenic** acid from soybean, presents interaction energie of -7.69 kcal/mol and a strong hydrogen bond with an estimated length of 3.160Å with the residue Ser728.

The Cα-helix and P-loop are in close proximity and interact with ATP required for trans-autophosphorylation in the ATP binding site **(9, 17).** The ligands **Biochanin A**, **Pinoresinol** and **α-Linolenic** acid form hydrogen bonds with the P-loop, can destabilize the active open conformation of the activation loop (C-helix-in-DFGin) thus no ATP substrate binding at its specific site which induces blockage of the downstream signaling cascade **(18).**

The ligand that has the best interaction energy value with the N-lobe P-loop residues specific for HER2 is **α-Linolenic** with a DG of -7.69 kcal/mol (**Fig. 9**).

(A) (B)

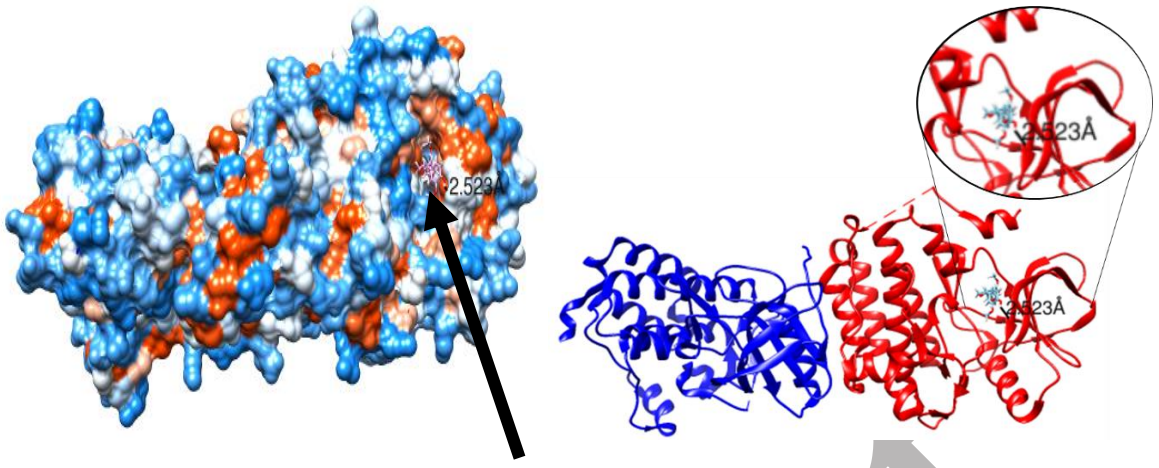

Figure 9. Three-dimensional illustration of the 3PP0 - α-Linolenic complex using the molecular surface (A) and the ribbon model (B)

As for the remaining ligand acting on the helix (Cα) of the N-lobe of the tyrosine kinase domain was selected according to its interaction with the residues in the stretch of the sequence extending from residue Pro761 to residue Ala775.

One ligand acts on the C-lobe activation loop of the HER2 receptor-specific domain tyrosine kinase domain was selected based on their interaction with residues spanning Asp863 to Val884 with the DFG motif (from residue Asp863 to residue Gly865) : **Kaempferol** Present an interaction energy of **-6.88** Kcal/mol establishes a hydrogen bond whose length is estimated at 3.572 Å with the residue ASP863 (**Fig.10**).

(A) (B)

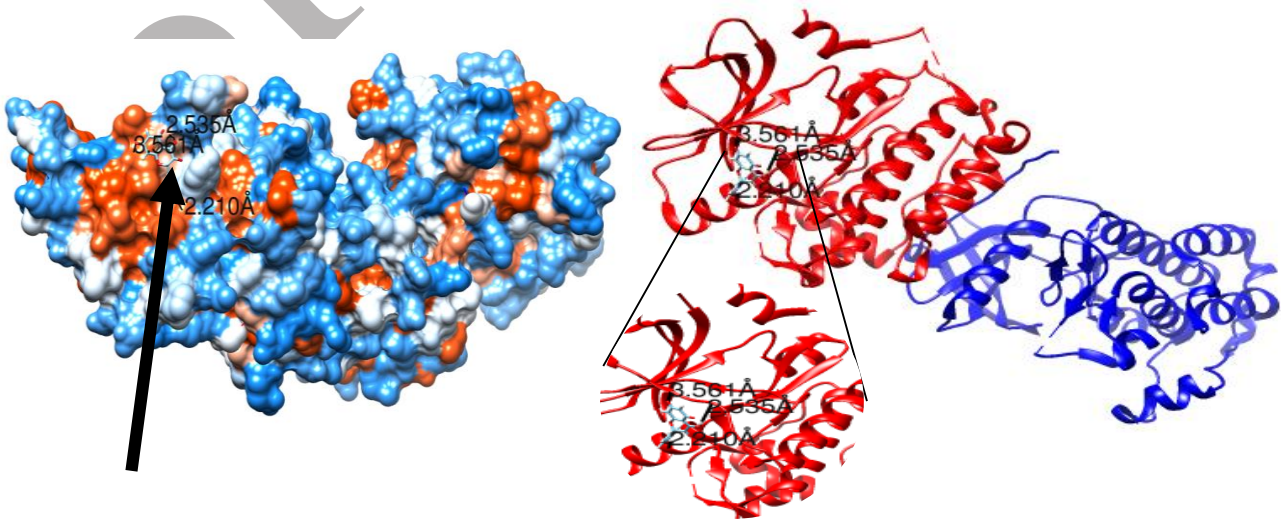

Figure 10. Three-dimensional illustration of the 3PP0 - Kaempferol complex using the molecular surface (A) and the ribbon model (B)

3.5. Visualization of Lapatinib Docking Results with 3PP0

Lapatinib is a ligand used as a chemical treatment in HER2+ breast cancer that is specific for inhibiting protein tyrosine kinase signaling pathways **(22).** Our results indicate that it has an interaction energy of -7.58 kcal/mol and establishes a hydrogen bond whose length is estimated at 2.303Å with the Val777 residue of the A chain. Lapatinib therefore interacts with the residues of the hydrophobic region of ErbB2 in the αC-β4 loop (**Fig.11 A**).

3.6. Comparison of the two ligands Lapatinib and Quercetin

The **quercetin** of red onions or buckwheat, has an interaction energy of -7.28 Kcal/mol establishes a hydrogen bond whose length is estimated at 2.130 Å with the residue Val777 (**Fig.11 A**).

Based on the visualization of the results of the two ligands, Lapatinib and **quercetin**, of which the first constitutes a modified chemical ligand and the second is a natural ligand coming mainly from red onions, it appears that : each of the two establishes a hydrogen bond of different lengths, of 2.303Å for Lapatinib and of 2.130Å for a **quercetin** , with the same residue valine777 of the hydrophobic region (in the αC-β4 loop), with different energies of interaction, -7.58kcal/mol for Lapatinib and -7.28kcal/mol for **quercetin** . Certainly, Lapatinib has a better interaction energy but remains as molecules of chemical origin which has its side effects, as for **quercetin**, in addition to its possible inhibition of 3PP0 has other beneficial effects on health.

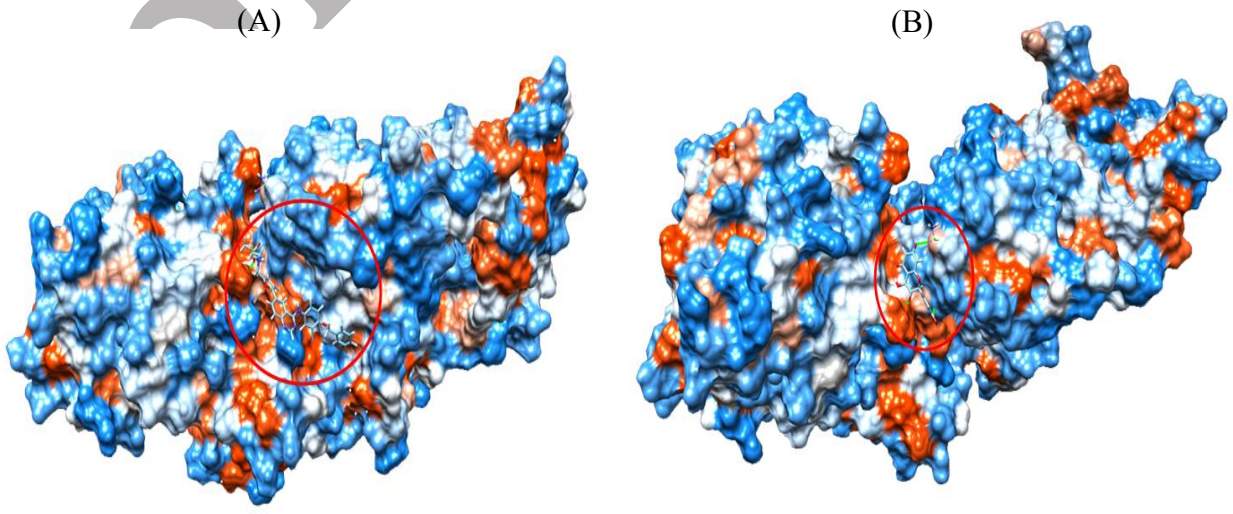

Figure 11. Three-dimensional illustration of the 3PP0 - Lapatinib (A) and 3PP0 - meletin (B) using the molecular surface

Indeed, quercetin or vitamin 'P' is a food-derived compound, a bioflavonoid, found in the pigments of colored fruits and vegetables, such as red onions, spinach, turmeric, apples, red grapes, carrots, berries, broccoli, green tea, lovage, but also chocolate or red wine. It is a natural antioxidant, helps fight against oxidative stress by capturing and blocking the activity of free radicals, but also by inhibiting the oxidation of lipids. It is also involved in the regulation of signaling pathways, cell cycle proliferation and the immune response.

In summary, investigating *in silico* before proceeding to the experimental stage can save a great deal of time and money. A number of ADMET factors, toxicological effects, and the likely active medication can all be predicted with the use of *in silico* technologies. The oral bioavailability of drugs was predicted using a number of prediction methodologies in this study, which may open the door to the creation of safer, innovative pharmaceuticals. Upon analyzing the screening and molecular docking studies, we reported that a large number of natural product could be employed as potential HER2 antagonists for the treatment of brest cancers. Additional wet-lab research is necessary to further evaluate these selected compounds.

Authors' Contribution

Conceived and designed experiments: Nesrine Lenchi
Conducted experiments: Nesrine Lenchi
Analyzed data: Nesrine Lenchi, Naima Maouche, Souad Khemili-Talbi.
Wrote the paper: Nesrine Lenchi

Data Availability

The data that support the findings of this study are available on request from the corresponding author.

Ethics

The authors have observed all ethical points including non-plagiarism, double publication, data distortion and data manipulation in this article.

Conflict of Interest

The author declares no known competing interest.

References

1. Domeyer PJ, Sergentanis TN (2020). New Insights into the Screening, Prompt Diagnosis, Management, and Prognosis of Breast Cancer. *J Oncol*.:8597892. doi: 10.1155/2020/8597892. PMID: 32308682; PMCID: PMC7142357.
2. Dai X, Xiang L, Li T, Bai Z (2016). Cancer hallmarks, biomarkers and breast cancer molecular subtypes. *Cancer J*.;7:1281-1294.
3. Hsu JL, Hung M-C (2016). The role of HER2, EGFR, and other receptor tyrosine kinases in breast cancer. *Cancer Metastasis Rev*.;35:575-588.
4. Maennling AE, Tur MK, Niebert M, et al. (2019). Molecular targeting therapy against EGFR family in breast cancer: progress and future potentials. *Cancers*.;11:1826.
5. Schlam I, Swain SM (2021). HER2-positive breast cancer and tyrosine kinase inhibitors: the time is now. *NPJ Breast Cancer*.;7:56-12.
6. Furrer D, Paquet C, Jacob S, Diorio C (2018). The human epidermal growth factor receptor 2 (HER2) as a prognostic and predictive biomarker: molecular insights into HER2 activation and diagnostic implications. *Cancer Progn*. doi:10.5772/intechopen.78271.
7. Sharma D, Kumar S, Narasimhan B (2018). Estrogen alpha receptor antagonists for the treatment of breast cancer: a review. *Chemistry Central Journal*.; 12(1):107. https://doi.org/10.1186/s13065-018- 0472-8 PMID: 30361894
8. Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN, Bourne PE (2000). The Protein Data Bank *Nucleic Acids Research* 28: 235-242 https://doi.org/10.1093/nar/28.1.235.
9. Aertgeerts K, Skene R, Yano J, Sang BC, Zou, H, Snell G, Jennings A, Iwamoto K, Habuka N, Hirokawa A, et al. (2011) Structural analysis of the mechanism of inhibition and allosteric activation of the kinase domain of HER2 protein. *J. Biol. Chem.*, *286*, 18756–18765.
10. Kim S, Chen J, Cheng T, Gindulyte A, He J, He S, Li Q, Shoemaker BA, Thiessen PA, Yu B, Zaslavsky L, Zhang J, Bolton EE. PubChem 2023 update. *Nucleic Acids Res*. 2023 Jan 6;51(D1):D1373-D1380. doi:10.1093/nar/gkac956.
11. Irwin JJ, Sterling T, Mysinger MM, Bolstad ES, and Coleman RG (2012). ZINC: A Free Tool to Discover Chemistry for Biology. *Journal of Chemical Information and Modeling 52* (7), 1757-1768 DOI: 10.1021/ci3001277.
12. Pires DE, Blundell TL, Ascher DB (2015). pkCSM: Predicting Small-Molecule Pharmacokinetic and Toxicity Properties Using Graph-Based Signatures. *J Med Chem*. ;58(9):4066-72. doi: 10.1021/acs.jmedchem.5b00104. Epub 2015 Apr 22. PMID: 25860834; PMCID: PMC4434528.
13. Grosdidier A, Zoete V, Michielin O (2011). SwissDock, a protein-small molecule docking web service based on EADock DSS, *Nucleic Acids Research*, Volume 39, Issue suppl_2, 1 July, Pages W270–W277, https://doi.org/10.1093/nar/gkr366.
14. Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, Ferrin TE. (2004) UCSF Chimera--a visualization system for exploratory research and analysis.. *J Comput Chem.* (13):1605-12.
15. Yim-im, W, Sawatdichaikul O, Semsri S, Horata N, Mokmak W, Tongsima S, Suksamrarn A , & Choowongkomon K (2014). Computational analyses of curcuminoid analogs against kinase domain of HER2. *BMC Bioinformatics*, 127063, 1–13.
16. Aller P. (2004). Etude Du domaine transmembranaire de recepteur tyrosine kinase dans un environnement membranaire. Aspects structuraux et mécanistiques explores par dynamique moleculaire *: HAL Id : tel-00009393*.
17. Martin-Fernandez, ML, Clarke DT, Roberts SK, Zanetti-Domingues LC, & Gervasio F L (2019). Structure and Dynamics of the EGF Receptor as Revealed by Experiments and

Simulations and Its Relevance to Non-Small Cell Lung Cancer. *Cells*, *8*(4), 316.https://doi.org/10.3390/cells8040316.

18. Modi V & Dunbrack RL (2019). Defining a new nomenclature for the structures of active and inactive kinases. *Proceedings of the National Academy of Sciences of the United States of America*, 116 (14), 6818–6827. https://doi.org/10.1073/pnas.1814279116.
19. Vijayan RSK, He P, Modi V, Duong-Ly KC, Ma H, Peterson JR, Dunbrack RL & Levy RM (2015). Conformational analysis of the DFG-out kinase motif and biochemical profiling of structurally validated type II inhibitors. *Journal of MedicinalChemistry*, 58(1), 466–479. https://doi.org/10.1021/jm501603h.
20. Klug LR, Kent JD & Heinrich MC (2018). Structural and clinical consequences of activation loop mutations in class III receptor tyrosine kinases. *Pharmacology and Therapeutics*, *191*: 123–134. https://doi.org/10.1016/j.pharmthera.2018.06.016
21. Shih AJ, Telesco SE, & Radhakrishnan R (2011). Analysis of somatic mutations in cancer: Molecular mechanisms of activation in the ErbB family of receptor tyrosine kinases. *Cancers*, *3*(1), 1195–1231. https://doi.org/10.3390/cancers3011195
22. Ulrich L, Okines AF (2021). Treating Advanced Unresectable or Metastatic HER2-Positive Breast Cancer: A Spotlight on Tucatinib. *Breast Cancer (Dove Med Press)*.;13:361-381 https://doi.org/10.2147/BCTT.S268451.