ORIGINAL ARTICLE

ISSN: 0331 - 670X

https://doi.org/10.51412/psnnjp.2024.28

The Nigerian Journal of PHARMACY
AS MEN OF HONOUR WE JOIN HANDS
THE OFFICIAL ORGAN OF THE PHARMACEUTICAL SOCIETY OF NIGERIA FOUNDED 1927

Identification of Some Bioactive Compounds from *Camellia sinensis* as Possible Inhibitors of human epidermal growth factor receptor 2 (HER2): A Structure-Based Drug Design for Breast Cancer Treatment.

Lilian N. Ebenyi[1]; Chidinma B. Godswill Egwuom[2]; Emmanuel M. Halilu[3]; Akachukwu P. Obialor[4,6;] Titilayo O. Johnson[5,6;] Abayomi E. Adegboyega[5,6] Victor U. Chigozie[7]and Basil U. Nwali[8].

[1]Department of Biotechnology, Ebonyi State University, Ebonyi, Nigeria.
[2]Department of Medicine, Friedrich-Alexander Universitat Erlangen-Nurnberg, Germany
[3]Faculty of Pharmacy, Cyprus International University, Haspolat, Nicosia, North Cyprus, via Mersin 10, Turkey.
[4] Department of Zoology, Faculty of Natural Sciences, University of Jos, Jos, Nigeria.
[5] Department of Biochemistry, Faculty of Basic Medical Science, College of Health Sciences, University of Jos, Jos, Nigeria.
[6] Bioinformatics unit, Jaris Computational Biology Centre, Jos, Nigeria.
[7]Department of Pharmaceutical Microbiology and Biotechnology, David Umahi Federal University of Health Sciences, Uburu, Ebonyi, Nigeria.
[8]Department of Biochemistry, Ebonyi State University, Ebonyi, Nigeria.

ARTICLE INFO

Article history:

Received 10th May 2024
Revised 4th September 2024
Accepted 16th September 2024
Online Published

Keywords:

HER2;
breast cancer;
Camellia sinensis;
molecular docking,
molecular dynamic simulation.

*Corresponding Author:
Lilian Nwanneka Ebenyi
Email: lilian.ebenyi@ebsu.edu.ng
Tel: +2348035851421

ABSTRACT

Background: Overexpression of HER2 has been related to a variety of malignancies, including breast cancer, and its inhibition has been established as an effective strategy for treating HER2-positive breast cancer. Because of its capacity to block carcinogenesis and reduce the proliferation of breast cancer cells, *Camellia sinensis* has been proven to be a source of anticancer agents.
Methods: In this study, the phytochemical library of *Camellia sinensis* was screened for inhibitory potentials against HER2 using molecular docking, pharmacophore modelling, ADMET studies, and molecular dynamic (MD) simulation.
Results: Gallocatechin, tricetinidin, SCHEMBL1950917, camellianin B, myricetin 3-glucoside, myricetin, camelliaside A, tricetin, faralateroside, and quercetin are the top-scoring compounds, with docking scores ranging from -9.327 kcal/mol to -8.147 kcal/mol. The selected compounds occupied the defined binding site and interacted with the same amino acid residues as the reference compound (03Q). The identified phytochemicals produced hydrophobic contacts with target amino acid residues of the HER2 ATP binding region in addition to one or more hydrogen bond interactions. Gallocatechin, possess favorable ADME properties and appeared to be the safest of all the chemicals, with an LD_{50} of 10,000 mg/kg, toxicity class 6, and no inclination toward any of the toxicity checkpoints. In the MD simulation, the gallocatechin-HER2 complex showed good stability, with GLN 799 and THR 862 retaining hydrogen bonds for 99% and 97% of the simulation, respectively.
Conclusion: The HER2-inhibiting potentials and favorable ADMET properties demonstrated by these compounds, especially gallocatechin, make them suitable for further experimental studies and development into drugs against HER2-positive breast cancer.

Introduction

Human epidermal growth factor receptor-2 (HER2) belongs to the epidermal growth factor family of tyrosine kinase receptors alongside HER1, HER3, and HER4[1]. The expression of HER receptors is crucial for cell division, differentiation, and survival[1,2]. Some adenocarcinomas, such as those of the breast, ovary, cervix, endometrial, lung, stomach, gastroesophageal junction, esophagus, and bladder, are linked to the aberrant overexpression of the HER2 protein. About 20 to 30% of human breast cancers are reported to have HER2 amplification or overexpression (1). HER2 amplification as a pivotal event in human breast carcinogenesis, occurs in about half of all in situ carcinomas[3]. Up to 25–50 copies of the HER2 gene and 40–100-fold increases in HER2 protein are found in HER2-positive breast cancers, leading to 2 million HER2 receptors expressed on the tumor cell surface[4]. The clinical aggressiveness of breast cancer and its biological characteristics are made worse by this aberration[5]. Some of the biological and clinical features of HER2-positive breast cancers are higher rates of proliferation; higher histologic and nuclear grade; more aneuploidy; lower estrogen receptors (ER) and progesterone receptors (PR) levels; lower sensitivity to endocrine therapy; a propensity to metastasize to central nervous system (CNS) viscera; and increased sensitivity to doxorubicin. The HER2-positive breast cancers develop and spread more rapidly than HER2-negative breast cancers, and they are more likely to respond to medications that inhibit the activity of the HER2 protein[6].

Inhibiting the expression and activity of HER2 has been found to be an effective strategy for the treatment of HER2-positive breast cancer and the prevention of the spread of malignant cells[7]. Increased understanding of the oncogenic activation pathways of the HER2 protein has led to the development of HER2-targeted treatments, such as trastuzumab, lapatinib, pertuzumab, and ado-trastuzumab emtansine (T-DM1), which are now frequently used in HER2-positive metastatic breast cancer[8]. First-line therapy for advanced HER2-positive breast cancer comprises a combination of taxane-based chemotherapy and dual HER2-inhibition with pertuzumab and trastuzumab[9], as recommended by international guidelines. The T-DM1 is used as a second-line treatment or for individuals who have progressed while on trastuzumab-based adjuvant therapy or within 6 months of stopping treatment[9]. Before the recent establishment of the third-line treatment, patients have been offered combinations based on capecitabine with either lapatinib or tratuzumab, or trastuzumab with chemotherapy, or trastuzumab with lapanitib. Although these treatments have greatly increased the chances of survival, some patients have been shown to end up suffering a relapse or disease progression. Furthermore, trastuzumab-related drug resistance issues have been documented[10], and while Lapatinib/Capecitabine combination therapy has shown synergistic advantages in the fight against breast cancer, it also comes with some unpleasant side effects[11]. Thus, advanced HER2-positive breast cancer remains incurable, and resistance to standard anti-HER2 medicines is essentially inevitable. The current biomedical and pharmacological landscape is driven by the imperative to face the challenge of developing ever-improved HER2-targeted therapy[8].

As scientists seek out better ways to treat breast cancer, natural and dietary chemicals have been the subject of a great deal of investigation. It has been shown that chemicals derived from natural sources have a considerable degree of anticancer action. Therefore, adopting a complementary therapy strategy can be a major help in this regard. Compounds found in nature have been shown to be effective in combating the aggressiveness of breast cancer, stopping the growth of cancer cells, and regulating the activity of pathways involved in cancer[12]. A very important dietary source of anti-cancer compounds that is increasingly gaining attention is *Camellia sinensis* (tea). Formerly limited to southern Asia, *Camellia sinensis* has now spread to the whole of Asia, Africa, and the Middle East. *Camellia sinensis* is used to make the many distinct types of tea, including white, yellow, green, oolong, dark, and black teas, all of which are oxidized to differing degrees depending on their processing. Green tea, which is the least processed, has significantly less caffeine than its highly processed counterparts, oolong and black tea. Several epidemiological studies have found that drinking green tea has health benefits and that tea intake is related to a lower incidence of several chronic illnesses, including cancer, though the evidence is still being validated [13,14]. Consumption of tea has been linked to anti-cancer action in a variety of studies, including epidemiological, clinical, and experimental studies (13). Based on a study by Mbuthia et al.[15], green tea infusions have potential ameliorative effects on breast cancer cells. An abundance of phytochemicals, some of which have anti-mutagenic, anti-tumor, anti-oxidant, anti-coagulant, anti-viral, anti-hypertensive, and cholesterol-lowering effects, have been discovered in *Camellia sinensis*[16].

The polyphenolic phytochemicals theaflavin-3,3′-digallate, predominantly present in black tea, and (−)-

epigallocatechin-3-gallate, particularly abundant in green tea, are thought to be the two most powerful anti-cancer agents found in tea, according to research[17]. These compounds have been the focus of several studies, all of which confirm their beneficial effects [16, 17]. Several other bioactive compounds including catechin, catechin gallate, gallocatechin, epicatechin digallates, epigallocatechin digallates, 3-O-methyl EC and EGC, and gallocatechin gallate are found in lower quantities in tea. A trace amount of 3′-O-methyl-EGCG was isolated from Oolong tea. Tea also contains flavonols including quercetin, myricitin, kaempferol, and their glycosides. Hence, in order to benefit from the rich therapeutic reservoir of tea, it is necessary to explore all the compounds known to be present in the plant for possible activity against important drug targets like HER2, which is the focus of this research. Consequently, the current study utilized in silico approaches to investigate the HER2-inhibiting potential of four hundred and five (405) *Camellia sinensis* compounds in an effort to identify alternative therapeutic options for the treatment of HER2-positive breast cancers.

Materials and methods

Protein preparation

The crystal structure of the kinase domain of human HER2 (PDB ID:3PP0) was retrieved from Protein Data Bank (PDB) repository. The protein was prepared using the protein preparation wizard panel of Glide (Schrödinger Suite 2021-2) where bond orders were assigned, hydrogen added and disulfide bonds created. Water molecules beyond 3.0 Å of the heteroatoms were removed and the structure was minimized using OPLS2005 and optimized using PROPKA. Subsequently, the receptor grid file was generated to define the binding pocket for the ligands.

Ligand preparation

Four hundred and five (405) compounds from *Camellia sinensis* were downloaded from the PubChem database and prepared for molecular docking. For protein preparation, the Glide Wizard panel (Schrodinger Suite 2021-2) was used to determine the bond order, hydrogens were added, disulfide bonds were created, and missing side chains and loops were filled using primes. OPLS2005 and PROPKA were used to reduce and optimize the structure, respectively, removing water molecules with heteroatom content greater than 3.0 and creating a receptor mesh to create binding sites for the ligand.

Receptor grid generation

Receptor grid generation allows defining the position and size of the protein's active site for ligand docking. The scoring grid was defined based on the co-crystalized ligand 0 3 Q (2 - { 2 - [4 - ({ 5 - c h l o r o - 6 - [3 - (trifluoromethyl)phenoxy]pyridin-3-yl}amino)-5H-pyrrolo[3,2-d]pyrimidin-5-yl]ethoxy}ethanol) using the receptor grid generation tool of Schrödinger Maestro 12.5. The van der Waals (vdW) radius scaling factor of nonpolar receptor atoms were scaled at 1.0, with a partial charge cut off of 0.25.

Protein-Ligand Docking

Glide tool of Schrödinger Maestro 12.8 was used to perform the molecular docking studies using the generated receptor grid file. The prepared ligands were docked using standard precision (SP), setting ligand sampling to flexible, with the ligand sampling set to none (refine only). The vdW radius scaling factor was scaled at 0.80 with a partial charge cut-off of 0.15 for ligand atoms.

Receptor-ligand complex pharmacophore modelling

The first three compounds ranked with highest binding affinity against the target protein was used to develop a receptor-ligand complex pharmacophore model using PHASE. Auto (E-pharmacophore) method was used, hypothesis was set with maximum number of features to be generated at 7, minimum feature-feature distance at 2.00, minimum feature-feature distance for feature of the same type at 4.00 and donors as vectors.

Pharmacology parameters

The absorption, distribution, metabolism, excretion and toxicity (ADMET) properties of the test compounds were determined using in silico integrative model predictions at the SwissADME (Daina et al., 2017) and PROTOX-II (Banerjee et al., 2018) server respectively.

Molecular Dynamics Simulation

The MD simulation was carried out using the Schrödinger LLC Desmond simulation software (18). The simulation system was configured to employ an orthorhombic box-shaped SPC solvent model. The NPT ensemble was applied with the conditions of 300 kelvin temperature and 1.01325 bar pressure. The OPLS 2005 force field settings were used, and the simulation lasted 10 ns. The particle mesh Ewald method was used to determine the long-range electrostatic interactions. The cutoff radius for Coulomb interactions was 9.0. The water molecules were described in detail using

the point-charge model. Temperature was controlled via the Nosé-Hoover chain coupling system, and pressure was controlled by the Martyna-Tuckerman-Klein chain coupling system with a coupling constant of 2.0 ps. The non-bonded forces were calculated using an r-RESPA integrator, with the short-range forces updated every step and the long-range forces updated every three steps. The trajectories were stored for examination at intervals of 10 ps. The ligand-protein interaction and its dynamics were examined using the Desmond MD Simulation Interaction Diagram. The stability of the complex was tracked using the RMSD and RMSF of the ligand and protein.

Results

Table 1 shows the docking scores of selected *Camellia sinensis* compounds against HER2 as well as that of the reference compound (03Q) with their docking scores ranging from -9.327 ((-)-Gallocatechin) to -8.147 (quercetin). The 2D and 3D representations of the interactions between the compounds with the highest scores and the target are shown in Fig. 1 to 7. The binding of 03Q with HER2 was mediated by a halogen bond with LYS 753 and Leu 796, a hydrogen bond with SER 728, and hydrophobic interactions with LEU 726, VAL 734, ALA 751, ILE 752, ALA 771, MET 774, LEU 785, LEU 796, VAL 797, LEU 800, MET 801, CYS 805, LEU 852 and PHE 864. The selected *Camellia sinensis* compounds also demonstrated hydrophobic interactions with these active site amino acid residues of HER2. In addition to the hydrophobic interactions, (-)-Gallocatechin, SCHEMBL1950917 and camellianin B formed hydrogen bonds with SER 728, ASP 808, CYS 805, MET 801 and GLN 799; SCHEMBL1950917 and camellianin B exhibited an additional hydrogen bonding with THR 862. Tricetinidin formed hydrogen bonds with ASP 808, ARG 813 and ARG 849 while Myricetin 3-Glucoside formed hydrogen bonds with ASP 863, ASP 808, GLU 812, SER 728, LEU 726, MET 801. Also, myricetin formed one hydrogen bond with MET 801 and a Pi-Pi stacking interaction with PHE 864 while Camelliaside A formed hydrogen bonds with ASP 808, ARG 849, SER 728 and MET 801.

Receptor-ligand complex pharmacophore modelling

The pharmacophore models of the standard ligand and seven selected compounds are shown in figure 8. The models show interactions such as A: hydrogen acceptor, D: hydrogen donor, H: hydrophobic interaction, N: and R: aromatic ring. Q30 interacted with the target through two hydrogen acceptors, three aromatic rings and hydrophobic interaction. (-)-Gallocatechin and 2-[2-[3-[[methyl(pyridine-3-carbonyl)amino]methyl]phenyl]-4-(trifluoromethyl)phenoxy]acetic acid has three hydrogen donors, two aromatic rings and two hydrogen acceptors respectively. Tricetinidin, Camelliani and Camelliaside A have two hydrogen acceptors, aromatic rings and hydrogen donors. Myricetin 3-Glucoside has five hydrogen donors and two aromatic rings. Myricetin has three aromatic rings, two hydrogen acceptors and one N.

ADMET properties of ten selected hit compounds

Table 2 to 6 shows the ADMET properties of ten selected hit compounds, demonstrating their lipophilicity, water solubility, druglikeness bioavailability, toxicity as well as other parameters. The consensus Log P values of the compounds range from 3.81 (SCHEMBL1950917) to -2.72 (Camelliaside A), all the values being less than 5 (table 2). SCHEMBL1950917 is moderately soluble while the other compounds are soluble in water (table 3). Tricetinidin, SCHEMBL1950917, Tricetin and quercetin violates none of the lipinski rules while gallocatechin and myricetin violate only one and these compounds possess a bioavailability score of 0.55 or 0.56. The remining compounds violate 2 or 3 lipinski rules and possess a bioavailability score of 0.17 (table 4).

Camellianin B, myricetin 3-Glucoside, myricetin, camelliaside A and faralateroside have low GI absorption, however, none of the compounds are BBB permeant. Tricetinidin, camellianin B and faralateroside are Pgp substrates. Tricetinidin, myricetin, tricetin and quercetin are CYP1A2 inhibitors while SCHEMBL1950917 is an inhibitor of all except CYP1A2. Quercetin and tricetin are inhibitors of CYP2D6 and CYP3A4; myricetin is also an inhibitor of CYP3A4 (table 5). None of the compounds are hepatotoxic; tricetinidin, myricetin, tricetin and quercetin are carcinogenic while SCHEMBL1950917, camellianin B, myricetin 3-Glucoside, camelliaside A and faralateroside are immunotoxic. Myricetin, tricetin and quercetin are mutagenic and none of them are cytotoxic (table 6).

Molecular dynamic simulation of (-)-Gallocatechin-HER2 complex

The RMSD value for the backbone atom of HER2 with respect to its initial position increased to 1.109 Å for the first 1 ns and to 1.700 Å at 4.12 ns, decreased to 1.369 Å at 7 ns, and then maintained more constant values ranging between

1.685 and 1.661 Å from 7.46 ns to 10 ns respectively. The RMSD value for the ligand increased to 0.854 Å at 1 ns and maintained a stable fluctuation throughout the simulation (figure 9). The protein RMSF maintained a constant range with fluctuations at various positions, the major ones being around positions 37 to 40 (figure 10). Ligand fluctuations, with respect to the protein, were found at positions 5-7, 16, 17 and 20-21(figure 10).

Figure 12 shows the interacting residues of HER2 (3PP0) and the types of interactions and contacts they made with (-)-Gallocatechin throughout the MD simulation. As shown in figure 12 A, ASP 808 interacted for the highest faction of the simulation time and the interactions were facilitated by hydrogen bonds, ionic and water bridges. Next is ASP 863 with hydrogen bonds and water bridges, followed by GLN 799 with only hydrogen bonds. The number of distinct interactions between the protein and the ligand during the course of the trajectory is displayed in the top panel of figure 12 B. The complex sustained a stable number of distinct interactions throughout the simulation. The bottom panel of figure 12 B displays, for each trajectory frame, which residues are involved in ligand interaction. Some residues, most notably, ASP 808 made two to four distinct types of contacts with the ligand, illustrated by a deeper orange color on the scale to the right of the scatter plot. GLN 799 maintained a single and constant interaction throughout the entire simulation time, as shown by the consistent light orange color. This is also illustrated in the ligand interaction diagram (figure 12 C). The hydrogen bond interaction with GLN 799 was stable for 99% of the simulation, while that of THR 862 was stable for 97% of the simulation.

Table 1: The binding affinity (kcal/mol) of the top ten ranked bioactive compounds of *Camellia sinensis* against HER2 protein target

Title	PubChem CID	Docking score (kcal/mol)
03Q (Reference compound)	16736274	-10.797
(-)-Gallocatechin	9882981	-9.327
Tricetinidin	11199650	-8.998
SCHEMBL1950917	44194465	-8.922
Camellianin B	5487342	-8.796
Myricetin 3-Glucoside	44259426	-8.634
Myricetin	5281672	-8.576
Camelliaside A	14890565	-8.413
Tricetin	5281701	-8.355
Faralateroside	20055968	-8.277
Quercetin	5280343	-8.147

03Q: 2-{2-[4-({5-chloro-6-[3-(trifluoromethyl)phenoxy]pyridin-3-yl}amino)-5H-pyrrolo[3,2-d]pyrimidin-5-yl]ethoxy}ethanol

SCHEMBL1950917:2-[2-[3-[[methyl(pyridine-3-carbonyl)amino]methyl]phenyl]-4-(trifluoromethyl) phenoxy]acetic acid.

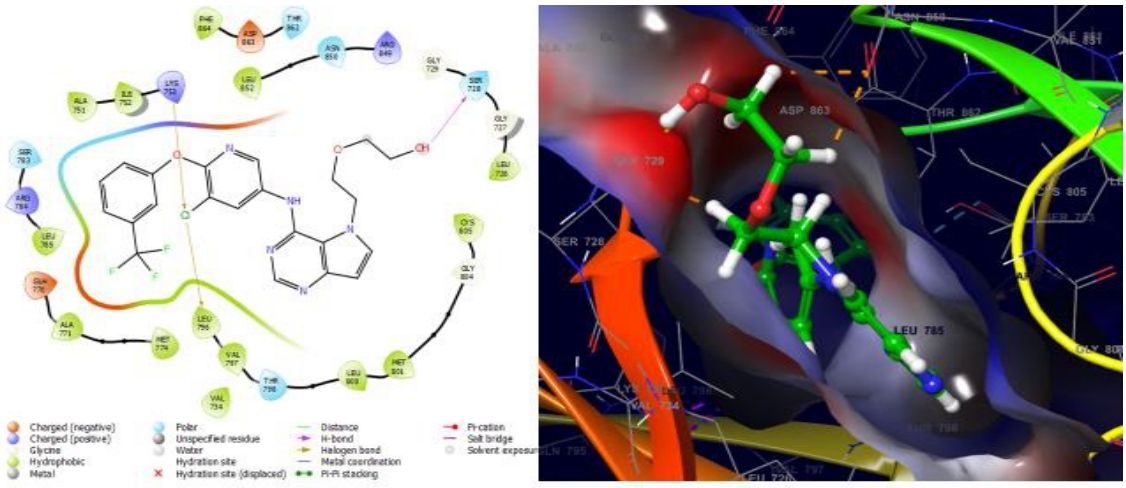

Figure 1: Molecular interactions of amino-acid residues of HER2 with 03Q showing the 2D (left) and 3D (right) views.

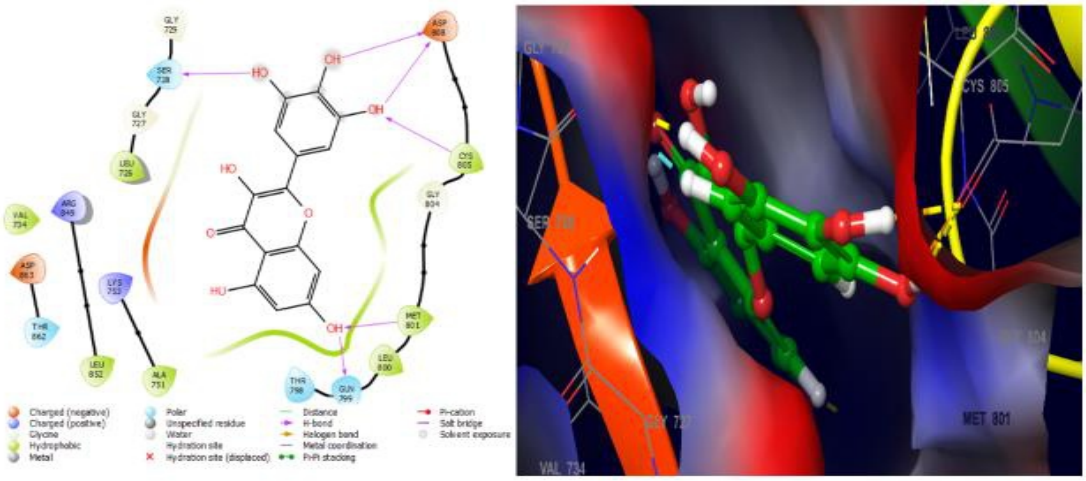

Figure 2: Molecular interactions of amino-acid residues of HER2 with (-)- Gallocatechin showing the 2D (left) and 3D (right) views.

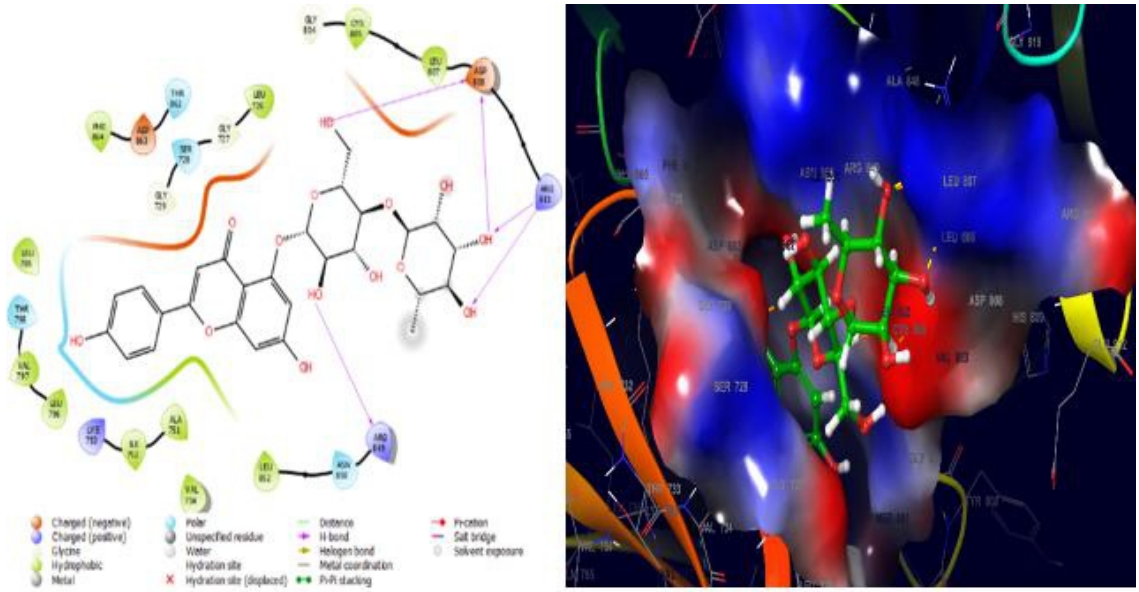

Figure 3: Molecular interactions of amino-acid residues of HER2 with tricetinidin showing the 2D (left) and 3D (right) views.

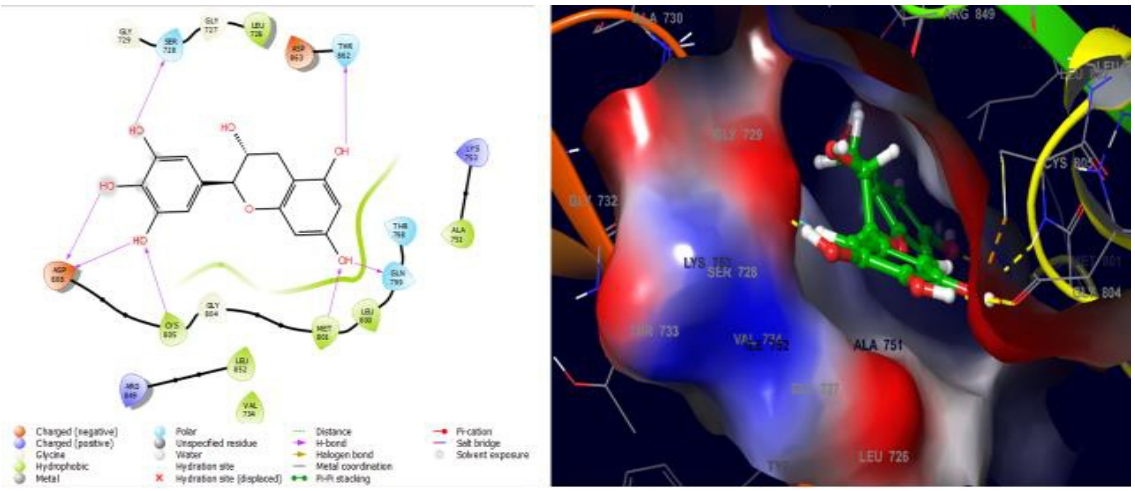

Figure 4: Molecular interactions of amino-acid residues of HER2 with SCHEMBL1950917 showing the 2D (left) and 3D (right) views

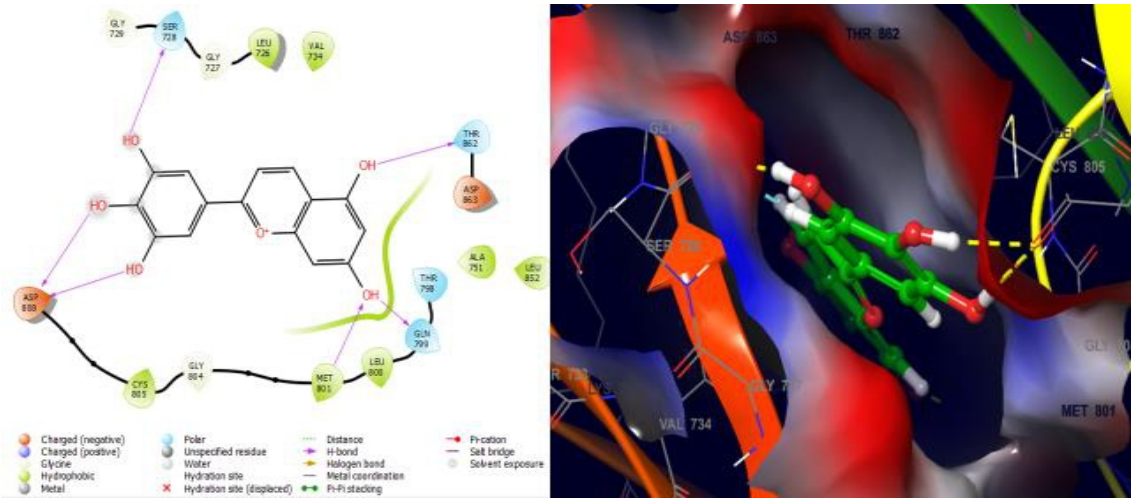

Figure 5: Molecular interactions of amino-acid residues of HER2 with Camellianin B showing the 2D (left) and 3D (right) views

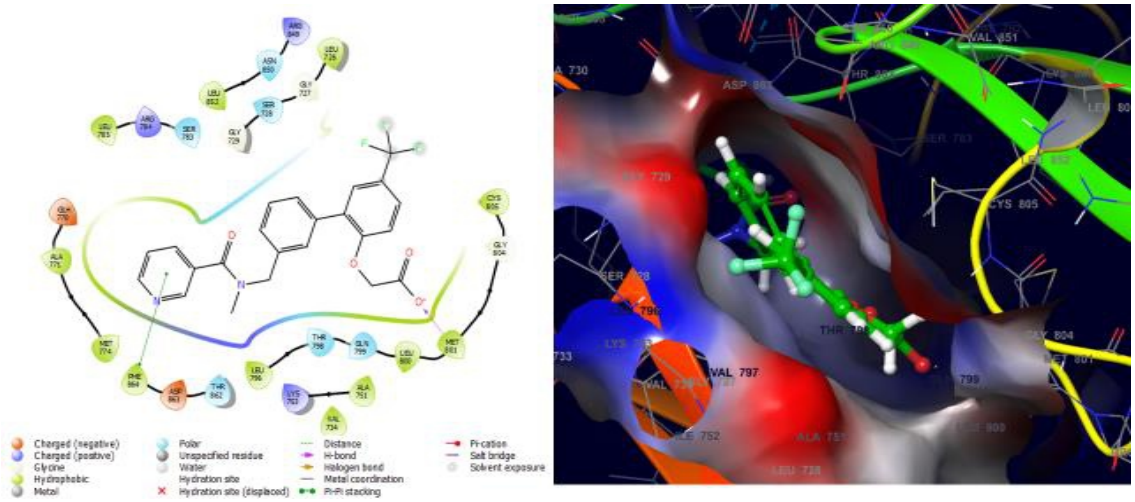

Figure 6: Molecular interactions of amino-acid residues of HER2 with Myricetin 3-Glucoside showing the 2D (left) and 3D (right) views

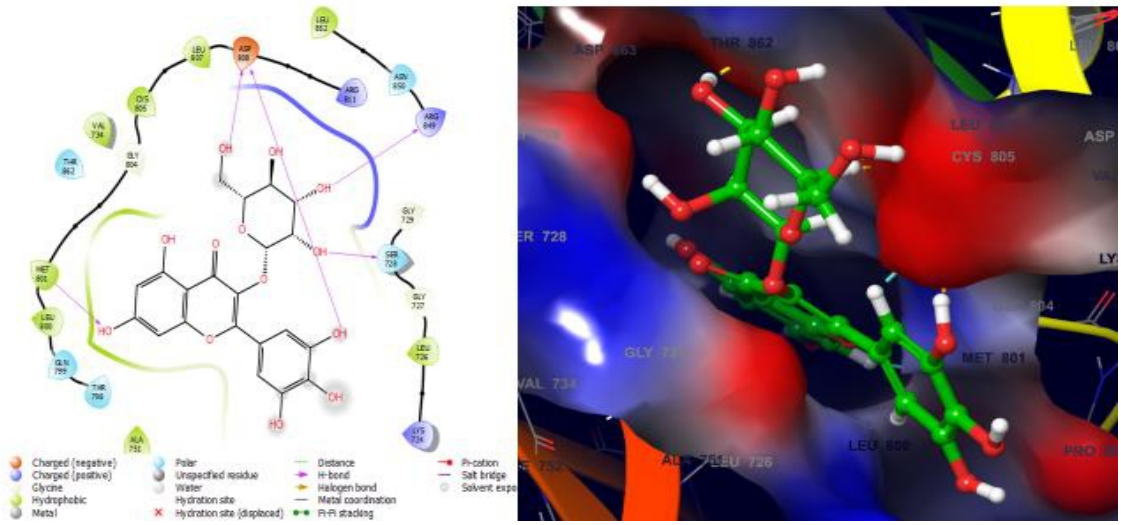

Figure 7: Molecular interactions of amino-acid residues of HER2 with Myricetin showing the 2D (left) and 3D (right) views

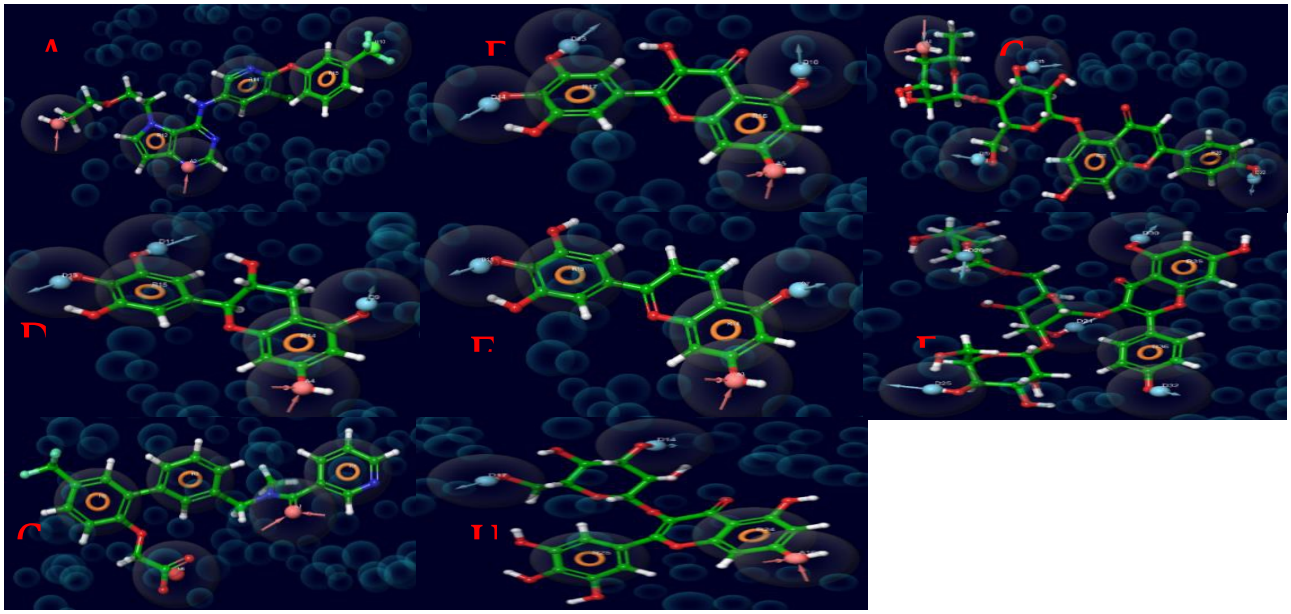

Figure 8: The receptor-ligand complex pharmacophore model of A:Q30; B: (-)-Gallocatechin; C:Tricetinidin;D:2-[2-[3-[[methyl(pyridine-3-carbonyl)amino]methyl]phenyl]-4-(trifluoromethyl)phenoxy]acetic acid; E: Camellianin B; F: Myricetin 3-Glucoside; G: Myricetin; H: Camelliaside A.

Table 2: The lipophilicity profile of the top 10 ranked phytochemical constituents of *Camellia sinensis.*

Molecule	MW	iLOGP	XLOGP3	WLOGP	MLOGP	Silicos-IT Log P	Consensus Log P
(-)-Gallocatechin;	306.27	0.98	0	0.93	-0.29	0.49	0.42
Tricetinidin	287.24	-2.42	1.05	2.91	0.32	0.26	0.42
SCHEMBL1950917	444.4	2.51	3.88	5.5	2.71	4.42	3.81
Camellianin B;	578.52	1.7	-0.38	-1.1	-2.96	-1.17	-0.78
Myricetin 3-Glucoside	480.38	0.18	0.01	-0.83	-3.07	-1.06	-0.96
Myricetin	318.24	1.08	1.18	1.69	-1.08	1.06	0.79
Camelliaside A	756.66	1.74	-2.52	-3.57	-5.49	-3.74	-2.72
Tricetin	302.24	1.33	2.17	1.99	-0.56	1.54	1.3
Faralateroside	756.66	2.1	-2.52	-3.57	-5.49	-3.74	-2.64
Quercetin	302.24	1.63	1.54	1.99	-0.56	1.54	1.23

Table 3: The water solubility profile of the top 10 ranked phytochemical constituents of *Camellia sinensis*.

Molecule	ESOL Log S	ESOL Solubility (mg/ml)	ESOL Class
(-)-Gallocatechin;	-2.08	2.57E+00	Soluble
Tricetinidin	-2.78	4.76E-01	Soluble
SCHEMBL1950917	-4.86	6.11E-03	Moderately soluble
Camellianin B;	-3.08	4.81E-01	Soluble
Myricetin 3-Glucoside	-2.91	5.93E-01	Soluble
Myricetin	-3.01	3.14E-01	Soluble
Camelliaside A	-2.57	2.02E+00	Soluble
Tricetin	-3.55	8.46E-02	Soluble
Faralateroside	-2.57	2.02E+00	Soluble
Quercetin	-3.16	2.11E-01	Soluble

Table 4: The drug-likeness properties of the top 10 ranked phytochemical constituents of *camellia sinensis*.

Molecule	Lipinski #violations	Bioavailability Score
(-)-Gallocatechin;	1	0.55
Tricetinidin	0	0.55
SCHEMBL1950917	0	0.56
Camellianin B;	3	0.17
Myricetin 3-Glucoside	2	0.17
Myricetin	1	0.55
Camelliaside A	3	0.17
Tricetin	0	0.55
Faralateroside	3	0.17
Quercetin	0	0.55

Table 5: The pharmacokinetics profile of the top 10 ranked phytochemical constituents of *Camellia sinensis*.

Molecule	GI absorption	BBB permeant	Pgp substrate	CYP1A2 inhibitor	CYP2C19 inhibitor	CYP2C9 inhibitor	CYP2D6 inhibitor	CYP3A4 inhibitor
(-)-Gallocatechin;	High	No	No	No	No	No	No	No
Tricetinidin	High	No	Yes	Yes	No	No	No	No
SCHEMBL1950917	High	No	No	No	Yes	Yes	Yes	Yes
Camellianin B;	Low	No	Yes	No	No	No	No	No
Myricetin 3-Glucoside	Low	No	No	No	No	No	No	No
Myricetin	Low	No	No	Yes	No	No	No	Yes
Camelliaside A	Low	No	No	No	No	No	No	No
Tricetin	High	No	No	Yes	No	No	Yes	Yes
Faralateroside	Low	No	Yes	No	No	No	No	No
Quercetin	High	No	No	Yes	No	No	Yes	Yes

Table 6: The toxicity profile of the top 10 ranked phytochemical constituents of *Camellia sinensis*.

Target	A	B	C	D	E	F	G	H	I	J
Hepatotoxicity	-	-	-	-	-	-	-	-	-	-
Carcinogenicity	-	+	-	-	-	+	-	+	-	+
Immunotoxicity	-	-	+	+	+	-	+	-	+	-
Mutagenicity	-	-	-	-	-	+	-	+	-	+
Cytotoxicity	-	-	-	-	-	-	-	-	-	-
LD50 (mg/kg)	10000	4000	1000	5000	5000	159	5000	3919	5000	159
Toxicity Class	6	5	4	5	5	3	5	5	5	3

A: (-)-Gallocatechin; B: Tricetinidin; C: SCHEMBL1950917; D: Camellianin B; E: Myricetin 3-Glucoside; F: Myricetin; G: Camelliaside A; H: Tricetin; I: Faralateroside; J: Quercetin

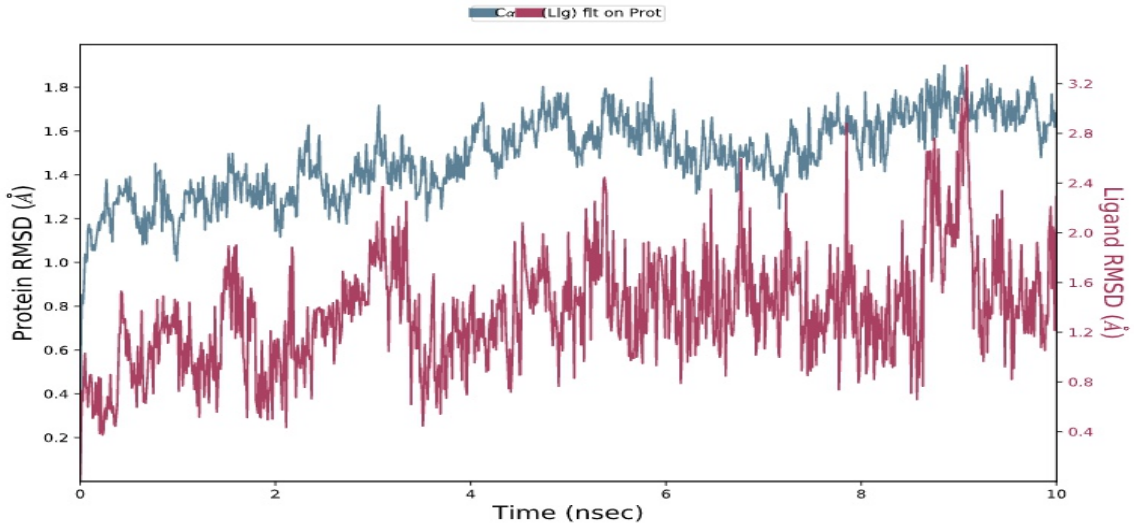

Figure 9: The protein-ligand Root Mean Square Deviation (RMSD) with respect to the reference frame (at time T = 0) during 10ns MD simulation. Left Y-axis: RMSD evolution of HER2 (3PP0); right Y-axis: RMSD for (-)-Gallocatechin.s

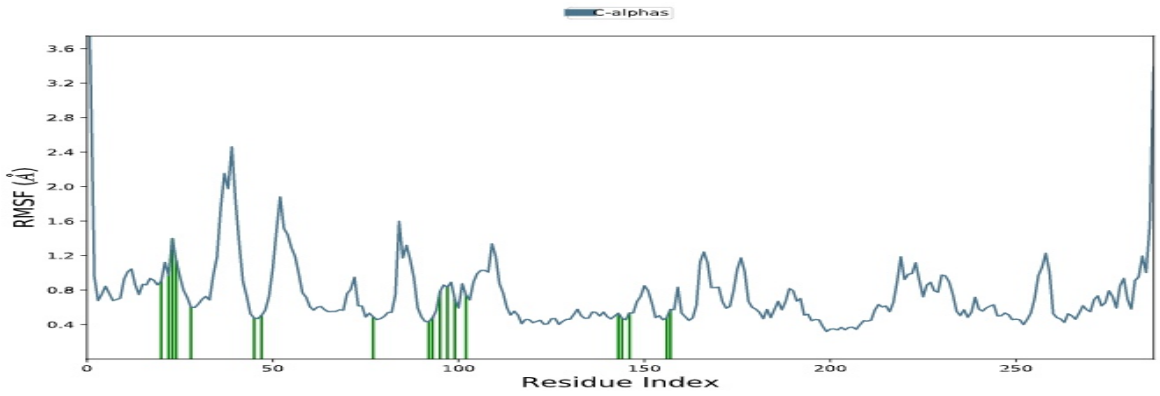

Figure 10: The Root Mean Square Fluctuation (RMSF) of HER2 (3PP0) during 10 ns MD simulation, representing local changes along the protein chain.

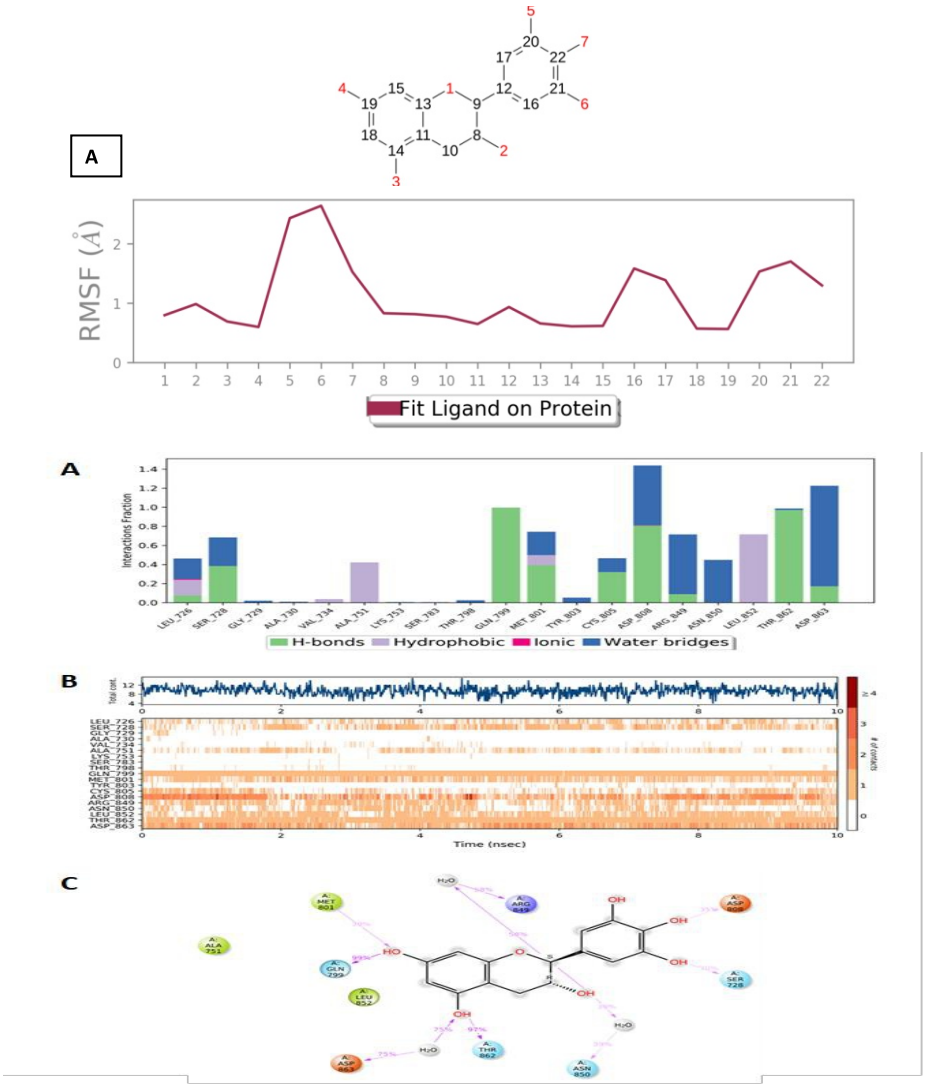

Figure 10: The Root Mean Square Fluctuation (RMSF) of HER2 (3PP0) during 10 ns MD simulation, representing local changes along the protein chain.Figure 10: Interactions and contacts (H-bonds, Hydrophobic, Ionic, and Water bridges) of HER2 (3PP0) with (-)-Gallocatechin throughout the MD simulation. A: Fractions of the simulation time a specific interaction (with respect to protein residues) is maintained. B: A timeline representation of the interactions and contacts. C: A schematic of detailed ligand atom interactions with the protein residues. Interactions that occur more than 30.0% of the simulation time in the selected trajectory (0 through 10 ns), are shown.

Discussion

HER2 is a tyrosine kinase receptor of type I that is involved in cell proliferation and differentiation[2] and its overexpression has been linked to different kinds of cancers, including breast cancer[19]. Identifying useful chemicals from plants with potential ability to suppress its overexpression could provide therapeutic benefits towards breast cancer control. *Camellia sinensis* is a well-known tea plant that is widely used as a healthful beverage and has been shown to have a variety of therapeutic properties. It's potential to limit the growth of breast cancer cells and stop carcinogenesis has been shown to be incredibly valuable in the development of innovative anticancer drugs[20]. In this study, the phytochemical library of *Camellia sinensis* was screened in the designated binding region of the HER2 crystal structure, and their binding energies were determined, in order to identify potential treatment candidates with high binding affinity against HER2.

The ten top-scoring compounds from the molecular docking analysis possess docking scores ranging from -9.327 kcal/mol to -8.147 kcal/mol (Table 1). The compounds in the order of binding affinity are gallocatechin, tricetinidin, SCHEMBL1950917, camellianin B, myricetin 3-glucoside, myricetin, camelliaside A, tricetin, faralateroside and quercetin. The 2D and 3D views of the protein-ligand interactions (Figures 1 to 7) show that some *Camellia sinensis* compounds occupied the designated binding site and interacted with the same amino acid residues the reference compound (03Q) interacted with. In addition to one or more hydrogen bond interactions, the compounds formed hydrophobic contacts with several amino acid residues of the ATP binding site of HER2. Molecular interaction with these amino acid residues is reported to be the target of most HER2 inhibitors. Beta-carotene-15,15′-epoxide, an inhibitor of HER2 was reported to act by forming hydrophobic interactions with LEU 726, LEU 800, ALA 751, VAL 734, LEU 796, LEU 755, ILE 767, THR 759, ALA 763, GLU 766, PHE 731, PHE 864, and LEU 852 inside the ATP binding domain of the protein. Lapatinib, despite having less hydrophobic contacts within the ATP binding domain, was revealed to have a very significant inhibitory power driven by the formation of hydrogen bonds with SER 728 and MET 801, and one pi stacking bond with PHE 864[21]. Similar interactions were observed in the molecular docking analysis of *Camellia sinensis* compounds against HER2 in this study and these could be responsible for the reported anticancer activity of green tea. The pharmacophore models of *Camellia sinensis* compounds on HER2 shows that hydrogen bond donors and acceptors as well as aromatic rings are the structural features of the compounds responsible for the observed molecular interactions (Figure 8). The presence of hydrogen bonds is known to improve the binding strength of a ligand to a protein target[22,23]. The generation of hydrogen bonds is a key molecular interaction that can increase the stability of *Camellia sinensis* compounds, which is a clear indication of their HER2-inhibitory capability. The existence of aromatic rings, according to the pharmacophore models, is also an essential structural property of the molecules. Around 20% of amino acids are aromatic in nature and interactions involving aromatic rings, such as protein-ligand interactions, are particularly crucial to diverse biological activities[23, 24]. Hence, the selected *Camellia sinensis* compounds may be considered as potential inhibitors of HER2 activity and possible anticancer agents since they possess the required structural characteristics, binding affinities, and molecular interactions. The results of the molecular dynamic simulation of the top-scoring compound further support this finding.

In addition to their inhibitory activity, most of the compounds showed acceptable ADMET properties. But, considering some of the compounds' ADMET outputs, lead optimization may be required to improve these drug characteristics for optimal outcomes. In silico ADMET analysis is a fast and low-cost method for evaluating the pharmacokinetics, drug-like properties, and toxicity of test compounds in drug development[24]. Water solubility and lipophilicity are two crucial physicochemical properties required for effective absorption and distribution of drugs[24]. As observed in this study, all of the selected *Camellia sinensis* compounds (apart from the moderately soluble SCHEMBL1950917) were predicted to be water soluble, meaning they are hydrophilic enough to pass through the aqueous blood. Besides, the low log P values of the compounds, which ranged between -2.72 and 3.81, indicate that they are not only water soluble, but also slightly lipophilic, suggesting that they may be able to pass through the intestinal lining and penetrate the target cell membrane to some extent, which is an important property of an orally bioavailable drug[24]. From the results of the drug-likeness prediction, tricetinidin, SCHEMBL1950917, tricetin, quercetin, gallocatechin and myricetin, with 1 or less lipinski violation and bioavailability score of 0.55 or 0.56, are likely to be good as orally administered drugs. An orally active drug should have no more than 5 hydrogen bond donors, no more than 10 hydrogen bond acceptors, a molecular weight of less than 500g/mol, and a log P of less

than 5, according to the Lipinski's Rule. If two or more of the rules are broken, a molecule is not orally active[25]).The bioavailability score, which incorporates total charge, TPSA, and the Lipinski filter, is a semi-quantitative assessment of the compounds' likelihood of being effective oral medicines[26]. The bioavailability score of 0.55 or 0.56 indicates that these drugs have a 55 or 56 percent chance of at least 10% oral bioavailability in rats or detectable human colon carcinoma (Caco-2) permeability, indicating that they are likely to be orally bioavailable[24]. This is also reflected in the GI absorption ability of these compounds. The GI absorption potential of tricetinidin, SCHEMBL1950917, tricetin, quercetin, and gallocatechin was predicted to be high, whereas the GI absorption potential of the remaining compounds was predicted to be low. None of the substances, however, have the ability to cross the blood-brain barrier (BBB).

As a substrate of the permeability glycoproteins (Pgp), a group of multidrug resistance proteins that actively flush foreign chemicals out of target organs through biological membranes for protective reasons[24], tricetinidin, camellianin B and faralateroside are not likely to successfully reach their target site of action. The CYP inhibitory potentials of Tricetinidin, SCHEMBL1950917, Myricetin, tricetin, and quercetin also suggest that these compounds may cause drug-drug interactions. This is because those CYP isoforms metabolize around fifty to ninety percent of medicines, and when they are blocked, pharmacokinetics-related drug-drug interactions do occur[27, 28].

According to toxicity predictions, (-)-Gallocatechin, with an LD_{50} of 10,000 mg/kg, of the toxicity class 6, and no inclination toward any of the toxicity check points, appears to be the safest of all the substances. Apart from SCHEMBL1950917 (LD_{50} 1000 mg/kg, class 4), myricetin (LD_{50} 159 mg/kg, class 3) and quercetin (LD_{50} 159 mg/kg, class 3), the remaining compounds in class 5 with LD_{50} 3919 to 5000 mg/kg are likewise relatively safe. Moreover, structural modification or optimization of the compounds may be required to eliminate some of the harmful properties while keeping their HER2 inhibiting ability.

As the top-scoring compound in the molecular docking study and because it gave a favorable ADME and toxicity profile, (-)-gallocatechin was chosen for the molecular dynamic simulation study. This is to further validate its potential as a HER2 inhibitor that could be developed for the treatment of HER2-positive breast cancers. The molecular dynamics simulation is a method and collection of algorithms for calculating and predicting the stability of molecules. It is one of the most important computational tools for understanding the strength of protein-ligand complexes and intermolecular interactions. Through deviation and fluctuation analysis, MD simulation helps to understand the ligand's dynamic behavior as well as its stability against the protein[29,30]. The MD simulation study of HER2 in complex with (-)-gallocatechin revealed the conformational stability, intermolecular interaction profile, and binding site occupancy of the complex, which is an important parameter for understanding the mechanism of inhibition predicted by the molecular docking approach. The RMSD number was used to calculate deviation in the protein's backbone during the simulation time[31]. The protein RMSD initially rose from 0 to 1.109 (at 1 ns), then to 1.700 by 4.12 ns, and finally stabilized between 1.685 and 1.661 until the simulation ended. The same applies to the ligand RMSD which shows an initial increase from 0 to 0.854 Å at 1 ns and maintained a stable fluctuation throughout the simulation. The initial fluctuation in the protein and ligand RMSD values is as a result of increase in the temperature of the system and the complex was able to stabilize after equilibration[29]. The ability to maintain stable RMSD values till the end of the simulation after an initial rise indication of the stability of the complex formed by HER2 in complex with (-)-gallocatechin.

The protein RMSF is useful for identifying local fluctuations throughout the protein chain, while the ligand RMSF (L-RMSF) is useful for assessing variations in ligand atom positions [29, 31]. The RMSF values of HER2 maintained a steady range, but there were fluctuations at some positions, the most significant of which were near positions 37 to 40. Ligand fluctuations in relation to the protein were discovered at positions 5-7, 16, 17, and 20-21. These positions represent the flexible loop regions in the structure of the protein and the ligand[31], and they are responsible for facilitating the induced-fit interaction of the ligand with the protein.

The types of interactions and contacts they made with HER2 throughout the MD simulation is very important in validating the inhibitory potential of (-)-Gallocatechin as predicted by the molecular docking analysis. Protein-ligand interactions (or 'contacts') are classified as Hydrogen Bonds, Hydrophobic, Ionic, and Water Bridges. Each interaction type has subcategories that may be examined using the 'Simulation Interactions Diagram' panel. As seen in Figure 13, ASP 808 interacted for the majority of the simulation duration, with interactions aided by hydrogen bonds, ionic and water bridges. Following that is ASP 863, which has both hydrogen bonds and water bridges, and

GLN 799, which exclusively has hydrogen bonds. It is important to note that the hydrogen bond interaction with GLN 799 and THR 862 was stable for 99% and 97% of the simulation respectively. The stable hydrogen bond interaction is a good indication of the stability of the protein-ligand complex, which is crucial for the HER2 inhibitory activity of (-)-Gallocatechin.

Conclusion

In order to uncover viable therapy options for HER2-positive breast cancer, the phytochemical library of *Camellia sinensis* was tested for inhibitory potentials against HER2 utilizing molecular docking, pharmacophore modelling, ADMET analyses and molecular dynamics simulation The top ten compounds in order of binding affinity are gallocatechin, tricetinidin, SCHEMBL1950917, camellianin B, myricetin 3-glucoside, myricetin, camelliaside A, tricetin, faralateroside, and quercetin. The selected compounds interacted majorly through hydrophobic contacts and hydrogen bonds with target amino acid residues inside the ATP binding domain of HER 2. Oral bioavailability is adequate for gallocatechin, tricetinidin, SCHEMBL1950917, tricetin, quercetin, and myricetin. With an LD_{50} of 10,000 mg/kg, toxicity class 6, and no trend toward any of the toxicity check points, gallocatechin looks to be the safest of all the compounds. Molecular dynamics simulation showed the formation of a stable complex between HER2 and gallocatechin, with GLN 799 and THR 862 maintaining hydrogen bond interaction for with 99% and 97% of the simulation time. However, longer MD simulation could be carried out to further appreciate the stability and assess any conformational changes. The compound could also be tested further, through *in vitro* and/or *in vivo* study, to validate its HER2-inhibitory potential for possible treatment of HER2-positive breast cancer.

References

1. Albagoush SA and Limaiem F (2022). HER 2. In StatPearls: StatPearls Publishing Treasure Island, FL. USA. Pp. 1-5.
2. Iqbal N, Iqbal N (2014) Human Epidermal Growth Factor Receptor 2 (Her2) in Cancers: Overexpression and Therapeutic Implications. Molecular Biology International 852-748. https://doi.org/10.1155/2014/852748
3. Mamani-Cancino AD, Veloz-Martinez MG, Casasola-Busteros L, Moctezuma-Meza C and Garcia-Cebada JM (2014) Frequency factor Her-2/neu overexpression in patients with breast cancer. Ginecologia y Obstetricia de Mexico, 82(6): 369-376. http://www.ncbi.nlm.nih.gov/pubmed/25016895
4. Chirstgen M, Bartels S, Luft A, Persing S, Henkel D, Lehmann U and Kreipe H (2018). Activating human epidermal growth factor receptor 2 (HER 2) gene mutation in bone metastases from breast cancer. Virchows Archive, 473(5), 577-582. https://doi.org/10.1007/s00428-01802414-1
5. King CR, Kraus MH and Aaronson SA (1985). Amplification of a Novel v-erb B-Related Gene in a Human Mammary Carcinoma. Science, 229 (4717), 974-976. https://doi.org/10.1126/science.2992089.
6. American Cancer Society. Breast Cancer Facts & Figures 2022-2024 report . https://www.cancer.org
7. Yang CS, Wang H, Li GX, Yang Z, Guan F and Jin H (2011). Cancer prevention by tea: Evidence from laboratory studies. Pharmacological Research, 64(2), 113-122. https://doi.org/10.1016/j.phrs.2011.03.001.
8. Mezni E, Vicier C, Guerin M, Sabatier R, Bertucci F and Goncalves A (2020). New Therapeutics in HER2-positive Advanced Breast Caancer: Towards a change in clinical practices? Cancers, 12(6), 1573. https://doi.org/10.3390/cancers/2061573.
9. Biganozoli L, Boers-Doets CB, Cardoso MJ, Carey LA, Cortes J, Curigliano G, Dieras V, El Sahir NS, Eniu A and Winer EP (2018). 4th ESO-ESMO International Consensus Guidelines for Advanced Breast Cancer (ABC 4). Annals of Oncology, 29(8), 1634-1657. https://doi.org/10.1093/amonc/mdy192.
10. Theillet C (2010). What do we learn from HER2-positive breast cancer genomic profiles? Breast Cancer Research, 12(3), 107. https://doi.org/10.1186/bcr2571.
11. Kroep JR, Linn SC, Boren E, Bloemendal HJ, Bass J, Mandjes IAM, Bosch J, Smit WM, de Graaf H, Schroder CP, Vermeulen GJ, Hop WCJ and Nortier JWR (2010) Lapatinib: clinical benefit in patients with HER2-positive advanced breast cancer. The Netherlands Journal of Medicine, 68(9), 371-376. https://www.ncbi.nlm.nih.gov/pubmed/20876920

.

12. Mitra S and Dash R (2018). Natural Products for the Management and prevention of Breast Cancer. Evidence-Based Complementary and Alternative Medicine, 1-23. https://doi.org/10.1155/2018/8324696.
13. Shirakami Y and Shimizu M (2018). Possible Mechanism of Green Tea and its constituents against cancer. Molecules, 23(9), 2284. https://doi.org/10.3390/molecules23092284.
14. Yuan JM, Sun C, and Butler LM (2011). Tea and cancer prevention: Epidemiological studies. Pharmacological Research, 64 (2), 123-135. https://doi.org/10.1016/j.phrs.2011.03.002
15. Sorokina M and Steinbeck C (2020). Review on natural products databases:where to find data in 2020. Journal of Cheminformatics, 12(1), 20. https://doi.org/101186/s1332-020-00424-9.
16. Schram L (2013). Going Green: The Role of the Green Tea component EGCG in Chemoprevention. Journal of Carcinogenesis and Mutagenesis, 04(02). https://doi.org/10.4172/2157-2518.
17. Khan N and Murkhtar H (2007). Tea polyphenols for health promotion. Life Science, 81 (7), 519-533. https://doi.org/10.1016/j.ifs.2007.06.011.
18. Schrodinger Release (2018) Desmond molecular dynamics system. D. E. Shaw Research, New York.
19. Burstein HJ (2005). The Distinctive Nature of HER2 positive breast cancer. New England Journal of medicine, 353(16), 1652-1654. https://doi.org/10.1056/nejmp058197.
20. Rafieian-kopaei M and Movahedi M (2017). Breast cancer chemopreventive and chemotherapeutic effects of Camellia sinensis (green tea): an updated review. Electronic physician 9(2), 3838-3844. https://doi.org/10.19082/3838.
21. Metibemu DS, Akinloye OA, Omotuyi IO, Okoye JO, Popoola MA and Akamo AJ (2012). Carotenoid-Enriched fractions from Spondias mombin demonstrated HER2 ATP kinase domain inhibition: Computational and In vivo animal model of Breast carcinoma studies. Frontiera in Oncology, 11, 1-11. https://doi.org/10.3389/fonc.2021.687190.
22. Chen D, Oezguen N, Urvil P, Ferguson C, Dann SM and Savidge TC (2016). Regulation of protein-ligand binding affinity by hydrogen bond pairing. Science Advances, 2(3). https://doi.org/10.1126/sciadv.1501240.
23. Samuel BB, Oluyemi WM, Johnson TO, Adegboyega AE (2021) Highthroughput virtual screening with molecular docking, phamacophore modelling and adme prediction to discover potential inhibitors of plasmodium falciparum lactate dehydrogenase (Pfldh) from compounds of combretaceae family. Trop J Nat Prod Res 5(9):1665–1672. https://doi.org/10.26538/tjnpr/v5.9.22.
24. Johnson TO, Adegboyega AE, Iwaloye O, Eseola OA, Plass W, Afolabi B, Rotimi D, Ahmed EI, Albrakati A, Batiha GE and Adeyemi OS (2021) Computational study of the therapeutic potentials of a new series of imidazole derivatives against SARS-CoV-2. J Pharmacol Sci. https://doi.org/10.1016/j.jphs.
25. Lipinski CA, Lombardo F, Dominy BW, Feeney PJ (2012) Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. Advanced Drug Delivery Rev 64:4–17
26. Testa B and Kramer SD (2009). The biochemistry of drug metabolism – An introduction part 5. Metabolism and bioactivity. Chemistry and Biodiversity 6(5), 591-684. https://doi.org/10.1002/cbdv.200900022.
27. Daina A, Michielin O, Zoete V (2017) SwissADME: a free web tool to evaluate pharmacokinetics, drug-likeness and medicinal chemistry friendliness of small molecules. Sci Rep 7:1–13. https://doi.org/10.1038/srep42717.
28. Huang SM, Strong JM, Zhang L, Reynolds KS, Nallani S, Temple R, Abraham S, AlHabet S, Baweja RK, Burckart GJ, Chung S, Colangelo P, Frucht D, Green MD, Hepp P, Karnaukhova E, Ko HS, Lee JI, Marroum PJ and Lesoko LJ (2008) Drug interations/review: New era in drug interaction evaluation: US Food and Drug Administration Update on CYP enzymes, transporters and the guidiance process. Journal of Clinical Pharmacology 48(6), 662-670.

https://doi.org/10.1177/0091270007312153.

29. Alturki NA, Mashraqi MM, Alzamami A, Alghamdi YS, Alharthi AA, Asiri SA, Ahmad S and Alshamrani S (2022). In-Silico Screening and Molecular Dyanmics Stimulation of Drug Bank Experimental Compounds against SARS-CoV-2. Molecules, 27(14), 4391. https://doi.org/10.3390/molecules27144391.

30. Balogun TA, Igbal MN, Saibu OA, Akintubosun MO, Lateef OM, Nneka UC, Abdullateef OT and Omoboyowa DA (2021). Discovery of potential HER2 inhibitors from Mangifera indica for the treatment of HER2 positive breast cancer: an integrated computational approach. Journal of Biomolecular Structure and Dynamics, 1-13. https://doi.org/10.1080/07391102.2021.1975570 .

31. Johnson TO, Adegboyega AE, Ojo OA, Yusuf AJ, Iwaloye O, Ugwah-Oguejiofor CJ, Asomadu RO, Chukwuma IF, Ejembi SA, Ugwuja EI, Alobaibi SS, Albogami SM, Batiha GES, Rajab BS and Conte-junior CA (2022) A Computational approach to elucidate the interactions of chemicals from *Artemisia annua* targeted toward SARS-CoV-2 main protease inhibition for COVID-19 treatment. Frontiers of Medicine 9:907583. https://doi.org/10.3389/fmed.2022.907583.