RBMC
Revista Brasileira Militar de Ciências

ISSN 2447-9071
doi https://doi.org/10.36414/rbmc.v7i19.105

In silico prospection of antineoplastic molecules from the *Artemisia annua* species

Prospecção in silico de moléculas antineoplásicas a partir da espécie Artemisia annua

Isabela Sacienti Lavezo[1], Juracy Cirino de Souza Neto[1], Túlio Nunes Pinto[1], Leonardo Luiz Borges[1,2]

[1] *Pontifícia Universidade Católica de Goiás – PUC Goiás*
[2] *Universidade Estadual de Goiás – UEG*

Abstract

Lung cancer kills the most men and the second that kills the most women (behind only breast cancer). The in silico study makes it possible to search for new drugs at low cost, with a greater possibility of rapid manufacturing and a lower future cost for their manufacture. The objective of this study was to analyze an antineoplastic activity of the compounds of Artemisia annua to obtain an active substance that can reach the molecular target of the cancer cells. Compounds with antineoplastic effects were selected using Scielo, PubMed, and ScienceDirect platforms. Afterward, the first screening of compound compounds was performed with a high ability to predict biological and pharmacological activity through the PASS Prediction, Pubchem, and Swiss ADME platforms. After the current screening, we determined the toxicological and molecular target prediction by the Portox II and Swiss Target Prediction platforms. As a final part, molecular docking and redocking were performed for a compound using the PDB server and the GOLD Suite 5.7.0 program. For another, we completed the pharmacophoric mapping using the Binding DB and PharmaGist database. The compounds scopoletin and caffeic acid were the most promising structures in silico models capable of interacting with EGFR (epidermal growth factor) and MM-9 (metalloproteinase type 9), respectively. The results obtained that these structures are promising to be tested in in vitro and in vivo tests about the antineoplastic activity. In addition, in silico analyses help to understand the biological effects of A. annua extracts regarding antineoplastic evidence.

Keywords: *Medicinal Plants; Scopoletin; Antineoplastic.*

Resumo

O câncer de pulmão é o com maior taxa de mortalidade entre homens e o segundo que mata mais mulheres (atrás apenas do câncer de mama). Os estudos in silico possibilitam a busca de novos candidatos a fármaco com menor custo e com mais rapidez. O objetivo deste estudo foi analisar a potencial atividade antineoplásica dos compostos de Artemisia annua para a obtenção de moléculas mais promissoras que possam atingir o alvo molecular das células cancerosas. Os compostos presents na espécie A. annua foram selecionados nas plataformas Scielo, PubMed e ScienceDirect. Posteriormente, a primeira triagem de compostos foi realizada para direcionar as potenciais atividades biológicas e farmacológicas por meio das plataformas Pubchem, PASS Prediction e SwissADME. Após esta triagem, foi realizada a predição toxicológica e na busca de alvos pelas plataformas Portox II e Swiss Target Prediction, respectivamente. Posteriormente, o docking foi realizado empregando o programa GOLD Suite 5.7.0. Também foi realizado o mapeamento farmacofórico usando o servidor PharmaGist. Os compostos escopoletina e ácido caféico foram as estruturas mais promissoras em modelos in silico capazes de interagir com EGFR (fator de crescimento epidérmico) e MM-9 (metaloproteinase tipo 9), respectivamente. Os resultados obtidos sugerem que essas estruturas são promissoras para serem testadas in vitro e in vivo sobre a atividade antineoplásica. Além disso, as análises in silico ajudam a compreender os efeitos biológicos dos compostos presents na espécie A. annua em relação às evidências antineoplásicas.

Palavras-chave: *Plantas Medicinas; Escopoletina; Antineoplásico.*

Contato para correspondência:
Leonardo Luiz Borges

E-mail:
leonardo.cbb@pucgoias.edu.br

Conflito de interesse: Não

Financiamento: Recursos próprios

Recebido: 18/10/2021
Aprovado: 04/11/2021

Introduction

In Brazil, lung cancer kills the most men and the second that kills the most women (behind only breast cancer). In 2019, according to SIM (Mortality Information System), there were 29,354 cancer deaths, 16,733 men, and 12,621 women. (2019 – Cancer Mortality Atlas – SIM). According to INCA (National Cancer Institute), the Estimates of new cases are 30,200, with 17,760 men and 12,440 women[1].

In Brazil, lung cancer kills the most men and the second that kills the most women (behind only breast cancer). In 2017, according to the SIM (Mortality Information System) of the Ministry of Health, more than 27 thousand people died from this cause[1]. According to the National Cancer Institute, in Brazil in 2016, 17,330 and 10,890 new tracheal, bronchial, or lung cancer were detected in men and women, respectively, with an incidence of 17.49 cases/ 100,000 men and 10.54 cases/100,000 women[1,2].

Although smoking is the leading cause of lung cancer, accounting for more than two-thirds of deaths globally, other carcinogens have also been identified, such as exposure to diesel engine exhaust and air pollution, among others[1].

Lung cancer starts in the cells that line the bronchi and parts of the lung, such as the bronchioles or alveoli. There are two main types of lung cancer, non-small cell lung cancer (80-85% of cases) and small cell lung cancer (10-15% of cases). Non-small cell lung cancer can be subdivided into adenocarcinoma, squamous cell carcinoma, large cell carcinoma (undifferentiated), adenosquamous carcinoma, and sarcomatoid carcinoma, the last two being less common. In addition, there are other types of lung cancer such as adenoid cystic carcinomas, lymphomas, sarcomas, carcinoid lung tumors. However, these have different treatments and are rarer. In Brazil, the incidence of adenocarcinoma has increased about squamous cell carcinoma[3].

Medicinal plants are an essential source for discovering new candidate bioactive structures, especially in the area of antineoplastic substances. For example, about 60% of the drugs currently used in cancers originate from medicinal plants, such as paclitaxel, vincristine, vinblastine, etoposide[2].

Artemisia annua is a plant species that was first mentioned as an alternative to treat fever and chills, as it has antimalarial mechanisms. As such, it is used throughout Asia and Africa as tea and juice to treat malaria. In addition, its main compound, artemisinin (ARS), is used alone in several parts of the world. The plant is restricted to its antimalarial activity and has antineoplastic activity *in vivo* and *in vitro*[3].

In silico studies can provide the prospect of new drug candidates with more incredible speed and lower cost. These studies also have the advantage of minimizing the use of laboratory animals in preclinical investigations. Furthermore, the *in silico* tools are based on the principle of chemical similarity. Similar substances bind to similar targets, where the comparison between compounds is made using similarity research based on ligands[4,5].

Thus, this study aimed to evaluate, using in silico tools candidate molecules for antineoplastic activity from the *Artemisia annua* species.

Methods

A virtual search was carried out for plants with possible antineoplastic activities against cancer on the platforms PubMed (https://pubmed.ncbi.nlm.nih.gov/), ScienceDirect (https://www.sciencedirect.com/) and Scielo (https://scielo.org/), where articles about the *Artemisia annua* species were selected.

Molecular structures were obtained from the PubChem website (https://pubchem.ncbi.nlm.nih.gov/). In the first screening of molecules, the PASS Prediction server (http://www.pharmaexpert.ru/passonline/)[6]. was used to structurally select the molecules with a greater probability of having antineoplastic activity. The pharmacokinetic prediction analysis was performed using the SwissADME server (http://www.swissadme.ch/)[7] and a toxicological prediction performed in Protox II (https://www.ncbi. nlm.nih.gov/pmc/articles/PMC6031011/)[8]. Finally, a projection to verify the most likely targets of the selected ligands was performed using the Swiss Target Prediction server (http://www.swisstargetprediction.ch/)[7].

Molecular docking was performed after defining the molecular targets of each compound. First, the Protein Data Bank website (PDB; https://www.rcsb.org/)[8] was used to obtain the crystallized molecular structure of metalloproteinase 9 (MMP-9) and then through the GOLD program Suite 5.7.0[9] models of interaction with selected ligands of the species *A. annua* were constructed. To validate the model used in the molecular docking analysis, a redocking analysis with the co-crystallized ligand was performed to verify the model's robustness.

For the pharmacophoric mapping, the Binding DB database (https://www.bindingdb.org/bind/index.jsp) was used to identify the most potent compounds (lowest IC_{50} values) with antagonist activity of the Epidermal growth factor receptor (EGFR). A pharmacophoric model was generated from a set through the ParmaGist server (https://bioinfo3d.cs.tau.ac.il/PharmaGist/)[9] of known molecules for verification of spatial features potentially responsible for the biological activity investigated.

Results and Discussion

After the bibliographical survey, 37 compounds were found in the species *A. annua*, 24 of which were selected after the first screening using the PASS prediction server for presenting properties of interest (properties associated with antineoplastic effect). In this first screening, the activities related to the antineoplastic impacts were selected (antineoplastic, antimutagenic, and TP53 expression enhancer). The inclusion criteria for these compounds was the use of a value of Pa> 0.7 (predicted probability of the structure to present a specific activity suggested in the software) and values of Pi<0.05 (predicted probability of the molecule not showing a particular activity indicated in the software). Such predictions are based on the premise that compounds with similar structures have similar biological activities, thus potentially interacting with similar active sites. In this way, the server compares the molecule understudy with a bank of molecules and establishes the level of structural similarity to predict possible biological activities[10].

Through the SwissADME server, structures classified as druglike, according to Lipinski's four criteria or "rule of five" were selected. According to these rules, a candidate molecule is more likely to be orally bioavailable if: the molecular weight is less than 500 Da (a), the calculated octanol/water partition coefficient (log P) is less than 5 (b), the compound has no more than 5 hydrogen bond donors (c) and the compound has no more than 10 hydrogen bond acceptors (d). In addition, gastrointestinal absorption, blood-brain barrier permeability, and interaction with cytochromes at the hepatic level were also predicted. These parameters assess the similarity of compounds to known drugs and represent a complex balance of several structural features. The present study selected substances with potential gastrointestinal absorption, without permeability to the blood-brain barrier, and lower interactions with cytochrome P450. Two molecules were finally selected within these criteria: caffeic acid and scopoletin (Table 1)[7].

After classifying the molecules, the prediction of toxicity was performed using Protox II. This step allows you to define safe dosage parameters for each compound, analyzing through a scale and acute toxicity, hepatotoxicity, carcinogenesis, cytotoxicity, mutagenicity, among others. Scopoletin had a median lethal dose (LD_{50}) of 3800mg/kg, caffeic acid had an LD_{50} of 2980 mg/kg. Both showed hepatotoxicity, immunotoxicity, and carcinogenicity. However, they can be considered safe because they need high doses to generate toxicities. The toxicity scale ranges from 1 to 6, with 1 being the most toxic and 6 the least toxic; we obtained class 5 in our compounds[11].

SwissTargetPrediction searches for the similarity between the structures of target molecules and ligands known from humans, evaluating the structure of a compound, which it needs to interact with a biologically active target. Thus, it allows bioinformatics-based analysis of resources and contributes to developing alternative research methods that rely on computer simulations, minimizing the use of animals, cost, and time in scientific research. Both compounds, caffeic acid, and scopoletin were evaluated for interaction with human receptors, with caffeic acid interacting with metalloproteinase type 9 (MMP-9) and scopoletin interacting with epidermal growth factor receptor (EGFR). MMP-9 is necessary for bone remodeling, wound healing, angiogenic revascularization of ischemic tissues, and remyelination; it is also involved in human pathological processes, one of its associations being cancer cell invasion, metastasis, and tumor progression. Its relationship with carcinogenesis is mainly due to the release of epidermal growth factor that leads to angiogenic change during carcinogenesis [12-14].

EGFR is part of the ErbB family of transmembrane receptor tyrosine kinases involved in signal transduction pathways that regulate proliferation and apoptosis. EGFR mutations are more frequent in tumors with adenocarcinoma histology, never smokers or light smokers, women with non-small cell lung cancer, and in patients with East Asian ethnicities[12-14].

Table 1. Biological, pharmacokinetic, and toxicological properties were obtained using in silico tools with compounds obtained from the Artemisia annua species.

Compounds	Biological activity	PA	PI	Lipinski	GI absorption	BB permeant	Interaction with cytochrome	Toxicity
Artemisinin	Antineoplasic (melanona)	0.889	0.003	Yes, 0 violation	High	Yes	CYP1A2 inhibitor	LD_{50}:4228 mg/kg
	Antineoplasic	0.853	0.007					Expected toxicity class: 5
	Antileukemic	0.806	0.004					
Caffeic acid	Membrane integrity agonist	0.955	0.003	Yes, 0 violation	High	No	None	LD_{50}: 2980 mg/kg
	Mucomembranous protector	0.945	0.003					Expected toxicity class: 5
	Antimutagenic	0.845	0.003					
	TP53 expression enhancer	0.776	0.014					
	MMP9 expression inhibitor	0.831	0.003					
	Cytoprotectant	0.702	0.005					
Scopoletin	Antimutagenic	0.898	0.002	Yes; 0 violation	High	Yes	CYP1A2 inhibitor	LD_{50}:3800 mg/kg
	Antineoplastic	0.723	0.022					Expected toxicity class: 5

The redocking of the co-crystallized ligand with the target MMP-9 (PDB: 2OW1) was successfully performed, and its results are shown in Fig. 1. The *in silico* experiments showed that caffeic acid works well at the binding site of MMP- 9, having 5 hydrogen bonds and binding distance between caffeic acid and MMP-9 proteins ranging from 1.9-6.5 (only one being 6.5). The result of the redocking analysis pointed to 7 poses with RMSD (Root Mean Standard Deviation) values lower than 1A.

In squamous cell carcinoma (SCC), the processes of invasion and metastasis are regulated by a complex system dependent on the interaction between neoplastic cells and host tumor cells, which occur as a consequence of degradation of the basement membrane and extracellular matrix by the metalloproteinase of the matrix (MMP). MM-9 degrades collagen type IV, which is essential for the basement membrane, participating in the invasion of the stroma and blood vessels, necessary for carcinogenesis[13,14].

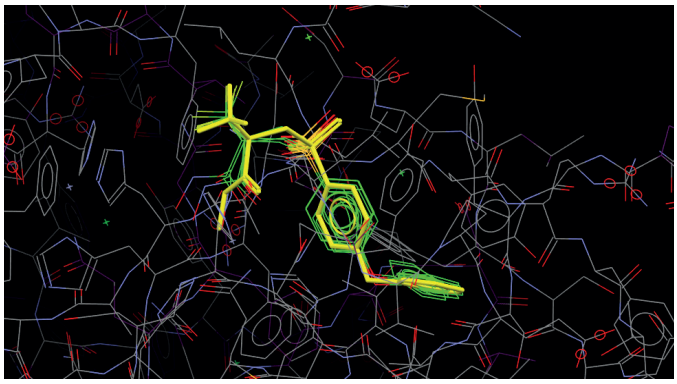
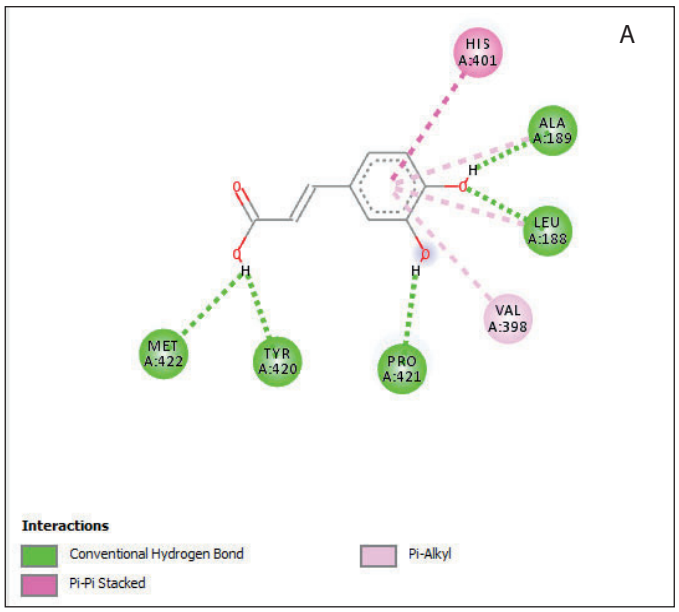


Figure 1. Result of redocking analysis with metalloproteinase-9 target (PDB: 2OW1) with its co-crystallized ligand ((2R)-2-amino-3,3,3-trifluoro-n-hydroxy-2-{[(4-phenoxyphenyl) sulfonyl]methyl}propanamide). This analysis presented 7 poses with RMSD (Root-mean-square deviation) values less than 1A.

The parameters used in the redocking analysis for the co-crystallized ligand of the MM-9 target were used for the molecular docking analysis with the caffeic acid ligand on this same target.

Molecular docking showed potential interaction between caffeic acid and ALA 189, LEU 188, PRO 421, TYR 420, MET 422 through hydrogen bonding, stronger bonds, and present ap-

propriate distances between them. Therefore, these bonds could more strongly support the structure at the active site of MM-9. Furthermore, it also interacts with VAL 398 through a π-alkyl bond which is a bond with lower strength, and with HIS 401 through a π-π bond which is the bond between electrons of two aromatic rings as shown in Figures 2 and 3.



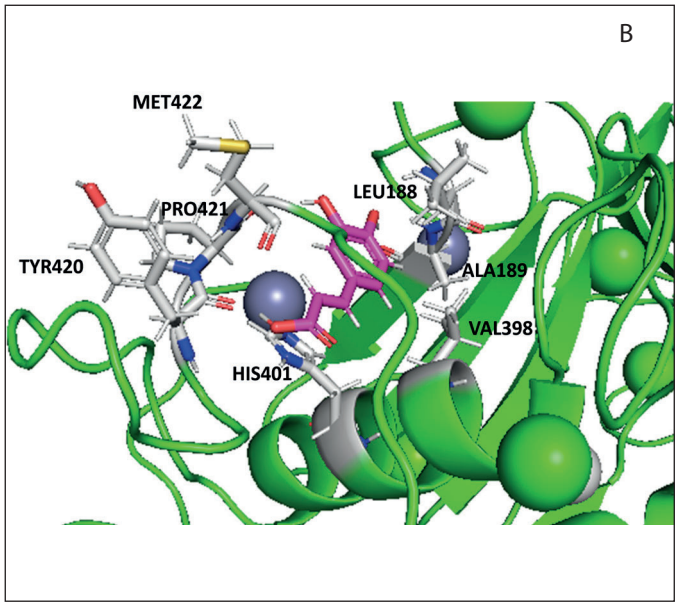


Figure 2. A. Two-dimensional representation of the potential interaction between caffeic acid and the active site of the MM-9 receptor. Five hydrogen bonds to caffeic acid are shown in green; a π-alkyl bond is shown in light pink, and a π-π bond is shown in dark pink. **B.** 3D representation of the interaction between caffeic acid and the active binding site of MM-9.

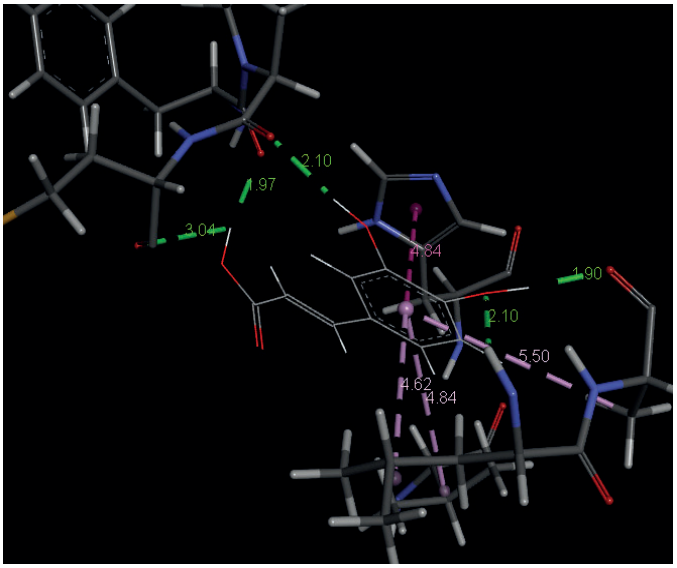


Figure 3. Representation of the interaction of caffeic acid with the active site of MM-9 with the respective distances of intermolecular interactions.

Pharmacophoric mapping allows interpreting the level of spatial similarity of the candidate structure with other binding molecules with known activities obtained from the Binding DB database. The analysis consists of four main steps, ligand representation (I), pair alignment (II), multiple alignments (III), and grouping and solution exit. This mapping revealed that the scopoletin molecule had alignment with the five most potent ligands that interact with the EGFR, as shown in Figure 4. Furthermore, scopoletin, a molecule also found in the plant species *A. annua*, shares two hydrogen bond donor groups and two aromatic groups with the 5 structures with the lowest IC50 values capable of interacting with the epidermal growth factor receptor.

Scopoletin interacts with the EGFR receptor through cell surface receptors, proteins located in the plasma membrane. These receptors are activated by stimuli from the external environment, generating intracellular signals leading to multiple molecular cascades. The successive phosphorylation of substrates activates the transcription of genes involved in proliferation, differentiation, invasion, angiogenesis, metastasis, and resistance to cell apoptosis. The role of EGFR in carcinogenesis is related to the mechanisms that lead to increased proliferative activity, invasiveness, angiogenesis, and resistance to chemotherapy and radiotherapy, which are paracrine and autocrine stimulation in the tumor microenvironment through increased production of ligands, overexpression of EGFR molecules in the membrane of tumor cells and activating mutations of the EGFR gene [9,14,15].

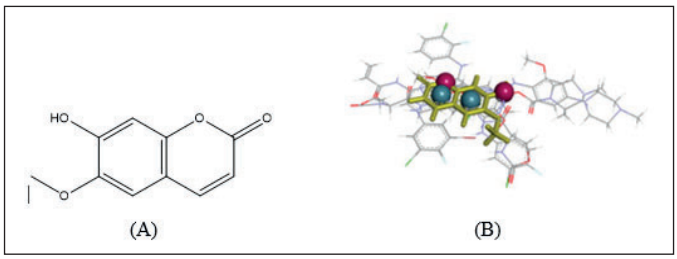


Figure. 4. A. Structure of scopoletin and adjustment of scopoletin to the EGFR antagonist pharmacophore. **B.** Pharmacophore features are color-coded for hydrogen bond donors (red) and aromatic rings (blue).

Conclusion

The compounds caffeic acid and scopoletin present in the *Artemisia annua* species were shown to be compounds with potential candidates for future tests in vitro and in vivo concerning antineoplastic activity. Furthermore, the results obtained here also point to two structures responsible for effects already identified in the *A. annua* extract. Thus, this study opens perspectives for targeted biological tests that explore the activities identified by the *in silico* tools.

References

1. Ribeiro AG, Baquero OS, Freitas CU d., Chiaravalotti Neto F, Cardoso MRA, Latorre M do RDO, et al. Incidence and mortality risk for respiratory tract cancer in the city of São Paulo, Brazil: Bayesian analysis of the association with traffic density. Cancer Epidemiol [Internet]. 2018;56(June):53–9. Available from: https://doi.org/10.1016/j.canep.2018.07.005
2. Costa-Lotufo L V., Montenegro RC, Alves APNN, Madeira SVF, Pessoa C, Moraes MEA de, et al. The Contribution of Natural Products as Source of New Anticancer Drugs: Studies Carried Out at the National Experimental Oncology Laboratory from the Federal University of Ceará. Rev Virtual Química. 2010;2(1):47–58.
3. Efferth T. From ancient herb to modern drug: Artemisia annua and artemisinin for cancer therapy. Semin Cancer Biol [Internet]. 2017;46:65–83. Available from: https://doi.org/10.1016/j.semcancer.2017.02.009
4. Trosset JY, Cavé C. In Silico Drug–Target Profiling. Methods Mol Biol. 2019;1953:89–103.
5. Tan BH, Pan Y, Dong AN, Ong CE. In vitro and in silico Approaches to Study Cytochrome P450-Mediated Interactions. J Pharm Pharm Sci. 2017;20(1):319–28.
6. Filimonov DA, Lagunin AA, Gloriozova TA, Rudik A V., Druzhilovskii DS, Pogodin P V., et al. Prediction of the biological activity spectra of organic compounds using the pass online web resource. Chem Heterocycl Compd. 2014;50(3):444–57.
7. Daina A, Michielin O, Zoete V. SwissADME: A free web tool to evaluate pharmacokinetics, drug-likeness and medicinal chemistry friendliness of small molecules. Sci Rep [Internet]. 2017;7(March):1–13. Available from: http://dx.doi.org/10.1038/srep42717
8. Banerjee P, Dehnbostel FO, Preissner R. Prediction is a balancing act: Importance of sampling methods to balance sensitivity and specificity of predictive models based on imbalanced chemical data sets. Front Chem. 2018;6(AUG):1–11.
9. Schneidman-Duhovny D, Dror O, Inbar Y, Nussinov R, Wolfson HJ. PharmaGist: a webserver for ligand-based pharmacophore detection. Nucleic Acids Res. 2008;36(Web Server issue):223–8.
10. Poroikov V V. Computer-Aided Drug Design: from Discovery of Novel Pharmaceutical Agents to Systems Pharmacology. Biochem Suppl Ser B Biomed Chem. 2020;14(3):216–27.
11. Banerjee P, Eckert AO, Schrey AK, Preissner R. ProTox-II: A webserver for the prediction of toxicity of chemicals. Nucleic Acids Res. 2018;46(W1):W257–63.
12. Daina A, Michielin O, Zoete V. SwissTargetPrediction: updated data and new features for efficient prediction of protein targets of small molecules. Nucleic Acids Res. 2019;47(W1):W357–3664.
13. Jorge SEDC, Kobayashi SS, Costa DB. Epidermal growth factor receptor (EGFR) mutations in lung cancer: Preclinical and clinical data. Brazilian J Med Biol Res. 2014;47(11):929–39.
14. Tochowicz A, Maskos K, Huber R, Oltenfreiter R, Dive V, Yiotakis A, et al. Crystal Structures of MMP-9 Complexes with Five Inhibitors: Contribution of the Flexible Arg424 Side-chain to Selectivity. J Mol Biol. 2007;371(4):989–1006.
15. Lopes GL, Vattimo EFQ, Junior GC. Identifying activating mutations in the EGFR gene: Prognostic and therapeutic implications in non-small cell lung cancer | Identificação de mutações ativadoras no gene EGFR: Implicações no prognóstico e no tratamento do carcinoma pulmonar de células não p. J Bras Pneumol. 2015;41(4):365–75.