STUDY OF COMPOUNDS ISOLATED FROM PLANT SPECIES WITH POTENTIAL ANTI-INFLAMMATORY AND ANTI-NEOPLASTIC ACTIVITY: A BIOCHEMOINFORMATICS APPROACH

Igor V. F. dos Santos[1], Njogu M. Kimanig[2,3], Martins D. de Cerqueira[4], Cleydson B. R. dos Santos[1], Aderaldo V. da Silva[1], Artemis S. N. Rodrigues[1], Luciana S. Lima[1], Elcimar S. Barros[1], Norberto S. Costa[1], Vitor H. S. Sanches[1]

[1] *Laboratory of Modeling and Computational Chemistry, Department of Biological and Health Sciences, Federal University of Amapá, Macapá 68902-280, AP, Brazil;*
[2] *Department of Physical Sciences, University of Embu, P.O. Box 6-60100, Embu,Kenya*
[3] *Natural Product Chemistry and Computational Drug Discovery Laboratory, P.O. Box 6-60100, Embu, Kenya*
[4] *GESNAT (Organic Natural Substances Study Group), Department of Organic Chemistry, Institute of Chemistry, Federal University of Bahia. Rua Barão de Jeremoabo, 147, Ondina University Campus, CEP: 40.170-115, Salvador - BA – Brazil.*

Inserir aqui por linha os numerais em sobrescrito e os endereços, em Times New Roman em itálico, 10.

Kewords: *Chemo-bioinformatics, anti-inflammatory and antineoplastic.*

1 INTRODUCTION

Chapada Diamantina National Park, located in the state of Bahia, Brazil, was originally conceived with the role of protecting a majestic landscape of waterfalls and clean rivers that flow through ancient sandstone mountains (Funch & Harley, 2007). Due to its exceptional biodiversity, CDNP is a prime location for discovering promising new compounds with potentially unreported bioactivities.

Chemo-bioinformatics uses computer simulations or theoretical analysis, rather than the traditional *in vivo* or *in vitro* tests, to accelerate drug discovery and development at a lower cost (Yusuf & History, 2023). This approach proves particularly valuable in exploring the potential anti-cancer (antineoplastic) and anti-inflammatory activities of various compounds (J. V. Cruz et al., 2021).

Lung cancer is a significant global health concern with an estimated 2 million diagnoses in 2018 (Bray et al., 2018). Smoking remains the primary cause of lung cancer accounting for 80% to 90% of reported cases (Ferlay et al., 2020). Skin cancer is another prevalent form of neoplasm (Nalli et al., 2018). It emerged as the most common cancer in Brazil and worldwide over the past decade (Silveira & Mauad, 2019) burdening healthcare systems and creating a significant public health concern (Duarte et al., 2018).

Interestingly, inflammation plays a complex role in cancer development. While it acts as a natural defense mechanism against internal and external threats, chronic inflammation can contribute to various diseases, including cardiovascular disease, rheumatoid arthritis, and even cancer (B Vendramini-Costa & E carvalho, 2012). Currently, epidemiological data indicate that over 25% of all cancers are related to chronic infections and other types of unresolved inflammation (B Vendramini-Costa & E carvalho, 2012).

This study aimed to investigate the potential antineoplastic, and anti-inflammatory activities of compounds isolated from CDNP plant species utilizing chemo-bioinformatics methods.

2 MATERIALS AND METHODS

The compounds were isolated from plant species of the Cerrado biome, in CDNP. They are part of the database of secondary metabolites described in several works of GESNAT (Group of Studies of Organic Natural Substances) from the Institute of Chemistry at the Federal University of Bahia. Chemical structures were drawn and their SMILES strings used for further analysis generated using ChemSketch (Freeware) 2020.1.2 software.

The prediction of biological activities was performed using the PASS software (Prediction of Activity Spectra for Substances) (D. A. Filimonov et al., 2014). This powerful tool boasts the capability to predict more than 3,500 types of biological activity (D. A. Filimonov et al., 2014) allowing researchers to virtually generate bioactivity profiles of small molecules. Notably, PASS software has garnered over 1,200 citations as of January 2021(https://www.way2drug.com/passonline/reference.php; consulted on: 10/24/2023).

The results are presented as probabilities: Pa (probability of being active) and Pi (probability of being inactive). The results with Pa > 0.3 were plotted for the antineoplastic and anti-inflammatory activities of the compound bank used.

The prediction of cytotoxic activity in cancer cell lines (CCL) was performed using the CLC-Pred 2.0 - http://www.way2drug.com/Cell-line/ (A. A. Lagunin et al., 2018). This webserver is used for *in silico* prediction of the cytotoxic effect of chemical compounds on non-transformed and cancerous cell lines based on the structural formula.

The prediction is based on PASS technology and the training set was created based on cytotoxicity data retrieved from ChEMBLdb (version 23) (https://www.ebi.ac.uk/chembldb/). Results with Pa > 0.3 for CCL were plotted to base the choice of molecular targets for subsequent analyses.

The SwissADME (http://www.swissadme.ch/index.php) was used to predict pharmacokinetic properties (Daina et al., 2014, 2017). It allows predicting physicochemical descriptors, ADME parameters, pharmacokinetic properties, and medicinal chemical compatibility.

The pharmacokinetic descriptors analyzed were: gastrointestinal absorption (GIA), violation of the rule of five (VROF), blood-brain barrier penetration (BBBP), skin permeation coefficient (Log Kp), lipophilicity (LogP), and solubility (LogS).

The software used for docking was Dockthor a receptor-ligand docking program (https://dockthor.lncc.br/v2/) (K. B. Santos et al., 2020). The target protein chain was kept in a neutral state and hydrogen was added to the ligands (according to the enablement available in the software). The grid size was kept standard (X=20, Y=20 and Z=20), as well as the discretization (0.25) and the precision of the algorithm. The redocking simulations were carried out to validate the analyses.

3 RESULTS AND DISCUSSION

PASS software was used to predict biological activity. **TABLE 1** highlights promising results, with compounds exhibiting significant anti-inflammatory and anti-neoplastic activity (Pa > 0.3). These findings guided the selection of molecular targets for further investigation.

Table 1: Biological activity prediction data via PASS software

Compound	Biological Activity			
	Antiinflammatory		Antineoplastic	
	Pa	Pi	Pa	Pi
Flavonoid 1	0.70	0.02	0.74	0.02
Flavonoid 2	0.69	0.02	0.71	0.03
Flavonoid 5	0.78	0.01	0.83	0.01
Polyisoprenylated benzophenone 2	0.33	0.14	0.79	0.01
Polyisoprenylated benzophenone 4	0.47	0.07	0.83	0.01
Myrciaine	0.33	0.14	0.59	0.05
2,2-dimethyl-5-hydroxi-7-phenyl	0.40	0.10	0.68	0.03

The compounds that showed potential antineoplastic activity (see **TABLE 2**) were subjected to prediction of cytotoxicity in CCL via CLC-Pred.

Table 2: Prediction of cytotoxicity in cancer cell lines via CLC-Pred.

Compound	Cancer cell Lineage Tissue	
	Lung	Skin
Flavonoid 1		
Flavonoid 2		
Flavonoid 5		
Polyisoprenylated benzophenone 4		
Myrciaine		
2,2-dimethyl-5-hydroxi-7-phenyl		

The green fields indicate a positive prediction (Pa > 0.3),
While the red field indicates a negative prediction (Pa < 0.3).

Most of the compounds obtained positive prediction (Pa > 0.3) for cancer cell lines originating from the lung and skin. These data served as a starting point for choosing molecular targets involved in carcinogenesis in these tissues.

The compounds selected for this study, in addition to the native ligands, were subjected to pharmacokinetic analysis via SwissADME. The objective was to select compounds with desirable pharmacokinetic profiles, taking the native ligands as reference control compounds in this case.

Table 3: prediction of pharmacokinetic and toxicological properties of compounds via SwissADME and PROTOX-II.

Compound*	GIA	ROFV	BBBP	Log *Kp* (cm/s)	LD_{50} (mg/Kg)	Toxicity Class	Hepatotoxicity	Carcinogenicity	Mutagenicity
1	▲	0	No	-5.77	4000	5	-	-	-
2	▲	0	No	-5.56	2570	5	-	-	-
3	▼	2	No	-8.64	5000	5	-	-	-
4	▲	0	No	-5.10	1250	4	-	-	-
5	▲	1	No	-4.84	1000	4	-	-	-
6	▲	0	No	-6.14	10000	6	-	-	-
7	▲	0	Yes	-4.91	500	4	-	-	-
8	▲	0	Yes	-6.35	125	3	+	-	+
9	▼	0	No	-5.92	2000	4	+	-	-
10	▲	0	Yes	-4.98	53	3	+	-	-
11	▲	0	Yes	-4.58	100	3	+	-	-

*: 1 (Flavonoid 1), 2 (Flavonoid 2), 3 (Flavonoid 5), 4 (Polyisoprenylated benzophenone 2), 5 (Polyisoprenylated benzophenone 4), 6 (Myrciaine), 7 (2,2-dimethyl-5-hydroxi-7-phenyl), 8 (Erlotinib), 9 (Vermurafenib), 10 (Diclofenac), 11 (Meclofenamic acid). GIA (gastrointestinal absorption: ▲ (high), ▼ (low). BBBP: Blood brain barrier penetration. ROFV: Rule of five violations. Hepatotoxicity, Carcinogenicity e Mutagenicity: + (active), - (inactive). Toxicity classes Class I: fatal if swallowed (LD50 ≤ 5), Class II: fatal if swallowed (5 < LD50 ≤ 50), Class III: toxic if swallowed (50 < LD50 ≤ 300), Class IV: harmful if swallowed (300 < LD50 ≤ 2000), Class V: may be harmful if swallowed (2000 < LD50 ≤ 5000), Class VI: non-toxic (LD50 > 5000).

The prediction of lipophilicity (LogP) via SwissADME is expressed by five methods: iLOGP, xLOGP3, WLOGP, MLOGP and Silicos-IT. Furthermore, the consensus LogP is the average between all the methods mentioned (Daina et al., 2014). This property affects, for example, the tendency of a compound to decompose in nonpolar versus aqueous environments. Therefore, increasing lipophilicity can generally lead to increased permeability, protein binding and volume of distribution (I. V. F. dos Santos et al., 2022; Kerns & Di, 2003).

Controls obtained LogP predictions in the range of 1.95 - 4.70 (see **TABLE 5**). In contrast, the tested compounds ranged from 0.03 to 5.69

Aqueous solubility is an important parameter for any drug candidate for oral or parenteral administration. Since a large amount of pharmaceutical ingredients must be administered in a small aqueous volume (Sepay et al., 2020). The consensus LogS prediction can be interpreted as follows: values between −4.0 to −6.0 indicate moderate solubility; −2.0 to −4.0 indicate good solubility and greater than -6.0 indicate poor solubility (Sepay et al., 2020). Three commercial controls presented consensus LogS compatible with moderate solubility and 1 presented low solubility.

Table *4*: Prediction of LogP values using different methods via SwissADME.

Compound (*)	LogP prediction					
	iLOGP	XLOGP3	WLOGP	MLOGP	Silicos-IT Log P	Consensus Log P
1	3.38	3.69	3.51	0.94	4.10	3.12
2	2.89	3.72	3.50	1.25	4.05	3.08
3	1.77	0.43	0.10	-2.06	-0.10	0.03
4	4.16	6.60	5.79	3.38	7.08	5.40
5	4.69	6.80	6.10	3.30	7.55	5.69
6	2.63	3.59	3.17	2.22	2.81	2.88
7	2.82	4.12	4.13	3.23	3.94	3.65
8	3.67	3.31	3.48	1.62	4.06	3.23
9	2.59	4.74	7.27	3.14	5.78	4.70
10	1.98	4.40	4.36	3.84	3.74	3.66
11	-6.76	5.16	4.74	2.73	3.87	1.95

*: 1 (Flavonoid 1), 2 (Flavonoid 2), 3 (Flavonoid 5), 4 (Polyisoprenylated benzophenone 2), 5 (Polyisoprenylated benzophenone 4), 6 (Myrciaine), 7 (2,2-dimethyl-5-hydroxi-7-phenyl), 8 (Erlotinib), 9 (Vermurafenib), 10 (Diclofenac), 11 (Meclofenamic acid).

Table *5*: Prediction of LogS values using different methods via SwissADME.

Compound (*)	LogS prediction			
	ESOL Log S	Ali Log S	Silicos-IT LogSw	Consensus LogS
1	-4.56	-5.25	-5.97	-5.26
2	-4.50	-5.09	-5.86	-5.15
3	-2.99	-3.99	-1.94	-2.97
4	-7.00	-8.26	-8.01	-7.76
5	-6.99	-8.47	-7.92	-7.79
6	-4.16	-5.2	-4.25	-4.54
7	-4.40	-4.45	-5.21	-4.69
8	-4.11	-4.56	-7.26	-5.31
9	-5.87	-6.58	-10.07	-7.51
10	-4.65	-5.15	-5.97	-5.26
11	-5.32	-5.94	-5.95	-5.74

*: 1 (Flavonoid 1), 2 (Flavonoid 2), 3 (Flavonoid 5), 4 (Polyisoprenylated benzophenone 2), 5 (Polyisoprenylated benzophenone 4), 6 (Myrciaine), 7 (2,2-dimethyl-5-hydroxi-7-phenyl), 8 (Erlotinib), 9 (Vermurafenib), 10 (Diclofenac), 11 (Meclofenamic acid).

The epidermal growth factor receptor tyrosine kinase domain (EGFTKD) (PDBID: 1M17) was used to investigate its anti-lung cancer activity, as described in previous studies Goda et al. (2022), Patel & Narechania, (2018) e Sarkar & Ganguly (2022).

Epidermal growth factor receptors are transmembrane receptors that play an important role in controlling cell growth and apoptosis. They have an extracellular binding fraction, a transmembrane component, and an intracellular tyrosine kinase unit. Mutations in them can lead to continued or abnormal activation of receptors, causing diseases such as non-small cell lung cancer (Patel & Narechania, 2018).

FIGURE 1 shows the BA (ΔG) of erlotinib (native ligand), gefitinib (commercial control) and compounds tested at the defined binding site of EGFTKD (PDB ID: 1M17). Flavonoid compounds 1 and 2 obtained the highest BA values (ΔG).

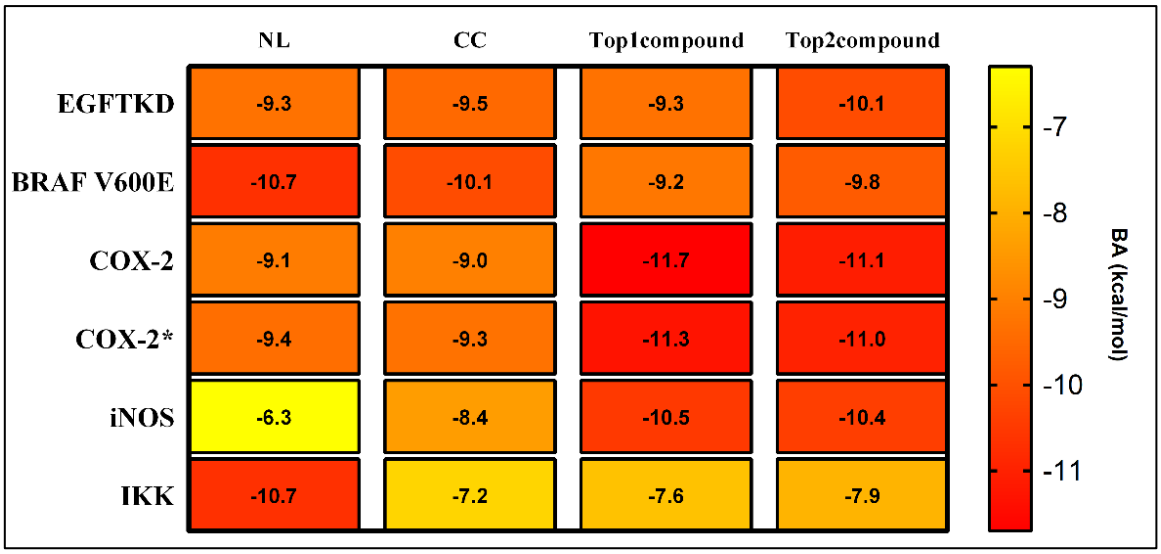

FIGURE 1: Heatmap graph showing BA (ΔG) values obtained. The Top1compound and Top2compound are shown respectively in parentheses: EGFTKD and BRAF V600E (Flavonoid 1, Flavonoid 2), COX-2 *Mus musculus* and iNOS (PBZP2, PBZP4), COX-2* *Homo sapiens* (PBZP2, Flavonoid 5), IKK (PBZP4, PBZP2). NL: native ligand; CC: commercial control.

CONCLUSIONS

Brazilian plant biodiversity, especially that native to the CDNP, can be a source of several bioactive compounds. In our work, we investigated potential bioactivities - via Chemo-bioinformatics tests - for 19 compounds isolated from CDNP plant species. Initially, PASS software was used, which indicated that the main bioactivities of the compounds were anti-inflammatory and antineoplastic.

The pharmacokinetic analysis considered several parameters, and the compounds studied demonstrated results similar to commercial controls. Regarding lipophilicity, the positive values obtained indicate that oral administration is possible. The aqueous solubility of most compounds was good or moderate. However, for 2 compounds suitable pharmaceutical preparations may be required, due to their low aqueous solubility. Predictions of toxicological properties indicated that the isolated compounds present lower toxicity than the commercial controls used. Furthermore, they do not present a positive prediction for hepatotoxicity, carcinogenicity, and mutagenicity.

The results of molecular docking analysis revealed that several compounds studied have a potential inhibitory action on targets related to lung and skin cancer. In particular, flavonoid 1 and flavonoid 2 compounds obtained BA values (ΔG) similar to commercial controls. In relation to anti-inflammatory activity, the compounds - PBZP2, PBZP4, flavonoid 5 - also obtained BA values (ΔG) like or higher than the values of commercial controls.

Finally, this study demonstrated how an appropriate Chemo-bioinformatics approach, and with free tools, can be used to investigate potential bioactivities of compounds isolated from plant biodiversity. It is important to highlight that to unequivocally confirm the bioactivities of the mentioned compounds, additional biological tests are required.

BIBLIOGRAPHIC REFERENCE

B Vendramini-Costa, D., & E carvalho, J. (2012). Molecular link mechanisms between inflammation and cancer. *Current Pharmaceutical Design*, *18*(26), 3831–3852.

Bray, F., Ferlay, J., Soerjomataram, I., Siegel, R. L., Torre, L. A., & Jemal, A. (2018). Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: A Cancer Journal for Clinicians*, *68*(6), 394–424.

Cruz, J. V., Giuliatti, S., Alves, L. B., Silva, R. C., Ferreira, E. F. B., Kimani, N. M., Silva, C. H. T. P., Souza, J. S. N. de, Espejo-Román, J. M., & Santos, C. B. R. (2021). Identification of novel potential cyclooxygenase-2 inhibitors using ligand- and structure-based virtual screening approaches. *Journal of Biomolecular Structure and Dynamics*, *0*(0), 1–23. https://doi.org/10.1080/07391102.2020.1871413

Daina, A., Michielin, O., & Zoete, V. (2014). iLOGP: a simple, robust, and efficient description of n-octanol/water partition coefficient for drug design using the GB/SA approach. *Journal of Chemical Information and Modeling*, *54*(12), 3284–3301.

Daina, A., Michielin, O., & Zoete, V. (2017). SwissADME: A free web tool to evaluate pharmacokinetics, drug-likeness and medicinal chemistry friendliness of small molecules. *Scientific Reports*, *7*(October 2016), 1–13. https://doi.org/10.1038/srep42717

dos Santos, I. V. F., Borges, R. S., Silva, G. M., de Lima, L. R., Bastos, R. S., Ramos, R. S., Silva, L. B., da Silva, C. H. T. P., & dos Santos, C. B. R. (2022). Hierarchical Virtual Screening Based on Rocaglamide Derivatives to Discover New Potential Anti-Skin Cancer Agents. *Frontiers in Molecular Biosciences*, *9*. https://doi.org/10.3389/fmolb.2022.836572

Duarte, A. F., Sousa-Pinto, B., Freitas, A., Delgado, L., Costa-Pereira, A., & Correia, O. (2018). Skin cancer healthcare impact: A nation-wide assessment of an administrative database. *Cancer Epidemiology*, *56*(September), 154–160. https://doi.org/10.1016/j.canep.2018.08.004

Filimonov, D. A., Lagunin, A. A., Gloriozova, T. A., Rudik, A. V., Druzhilovskii, D. S., Pogodin, P. V., & Poroikov, V. V. (2014). Prediction of the biological activity spectra of organic compounds using the pass online web resource. *Chemistry of Heterocyclic Compounds*, *50*(3), 444–457. https://doi.org/10.1007/s10593-014-1496-1

Funch, R. R., & Harley, R. M. (2007). Reconfiguring the boundaries of the Chapada Diamantina National Park (Brazil) using ecological criteria in the context of a human-dominated landscape. *Landscape and Urban Planning*, *83*(4), 355–362. https://doi.org/https://doi.org/10.1016/j.landurbplan.2007.06.003

Goda, M. S., Nafie, M. S., Awad, B. M., Abdel-Kader, M. S., Ibrahim, A. K., Badr, J. M., & Eltamany, E. E. (2022). In vitro and in vivo studies of anti-lung cancer activity of artemesia judaica L. Crude extract combined with LC-MS/MS metabolic profiling, docking simulation and HPLC-DAD quantification. *Antioxidants*, *11*(1). https://doi.org/10.3390/antiox11010017

Kerns, E. H., & Di, L. (2003). Pharmaceutical profiling in drug discovery. *Drug Discovery Today*, *8*(7), 316–323.

Lagunin, A. A., Dubovskaja, V. I., Rudik, A. V., Pogodin, P. V., Druzhilovskiy, D. S., Gloriozova, T. A., Filimonov, D. A., Sastry, N. G., & Poroikov, V. V. (2018). CLC-Pred: A freely available web-service for in silico prediction of human cell line cytotoxicity for drug-like compounds. *PLoS ONE*, *13*(1), 1–13. https://doi.org/10.1371/journal.pone.0191838

Nalli, A. D., Brown, L. E., Thomas, C. L., Sayers, T. J., Porco, J. A., & Henrich, C. J. (2018). Sensitization of renal carcinoma cells to TRAIL-induced apoptosis by rocaglamide and analogs. *Scientific Reports*, *8*(1), 1–11. https://doi.org/10.1038/s41598-018-35908-0

Patel, C. N., & Narechania, M. B. (2018). Targeting epidermal growth factor receptors inhibition in non-small-cell lung cancer: a computational approach. *Molecular Simulation*, *44*(17), 1478–1488. https://doi.org/10.1080/08927022.2018.1515484

Santos, K. B., Guedes, I. A., Karl, A. L. M., & Dardenne, L. E. (2020). Highly flexible ligand docking: benchmarking of the DockThor program on the LEADS-PEP protein–peptide data set. *Journal of Chemical Information and Modeling*, *60*(2), 667–683.

Sarkar, D., & Ganguly, A. (2022). Molecular Docking Studies with Garlic Phytochemical Constituents to Inhibit the Human EGFR Protein for Lung Cancer Therapy. *International Journal of Pharma and Bio Sciences*, *13*(2). https://doi.org/10.22376/ijpbs.2022.13.2.b1-14

Sepay, N., Sepay, N., Al Hoque, A., Mondal, R., Halder, U. C., & Muddassir, M. (2020). In silico fight against novel coronavirus by finding chromone derivatives as inhibitor of coronavirus main proteases enzyme. *Structural Chemistry*, *31*(5), 1831–1840. https://doi.org/10.1007/s11224-020-01537-5

Silveira, C. E. G., & Mauad, E. C. (2019). Analysis of a decade of skin cancer prevention using a mobile unit in Brazil. *Rural and Remote Health*, *19*(2), 82–87.

Yusuf, M., & History, A. (2023). REVIEW OPEN ACCESS Insights into the in-silico research: Current scenario, advantages, limits, and future perspectives ARTICLE INFO ABSTRACT. In *Life in Silico* (Vol. 1, Issue 1). https://life-insilico.com