RESEARCH ARTICLE

Assessment of *in vitro* and *in silico* Antiproliferative Effects of *p*-Cymene and Its Naturally Occurring 7-Oxygenated Derivatives on Human Cancer Cell Lines

Natália Plinio Souza[1], Ana Paula Gregório Alves Fontão[2], Gustavo Werneck de Souza e Silva[2], André Luiz Franco Sampaio[2], José Luiz Mazzei[1,3]*

Abstract

Objective: To evaluate the in vitro and silico antiproliferation of p-cymene, cumin aldehyde, cuminic acid, and cuminol against human cancer cell lines (HCCLs). The sodium salt of the organic acid was included after synthesis. **Methods:** The quality of the compounds was verified using analytical methods. A primary screening of the compounds at 100 µM was conducted (n=6) on nine HCCLs (SK-MEL-28, K562, Lucena, Jurkat, Caco-2, MDA-MB-231, THP-1, U87-MG, and Calu-3) and HEK-293 through the MTT method after 48h-incubation. The viability curve and apoptotic and necrotic cell populations of cumin aldehyde-treated Calu-3 cells were determined. The statistical significance relative to the vehicle was evaluated using ANOVA and Dunnett. The possibility of being active (Pa) on HCCLs and biological targets was assessed in silico using CLC-Pred. Moreover, the ADMET properties were predicted using three servers. **Results:** Only cumin aldehyde induced a low and significant (31±5%, p<0.001) in vitro antiproliferation, and even then, only on the Calu-3 line (IC50 650 µM). Only necrosis was significant (p<0.01) with 300 µM after 24h. The absence (p>0.05) of in vitro activity on the other HCCLs corroborated the low in silico probability of being active (Pa≤0.33), except for cumin aldehyde on the MDA-MB-231 line (Pa=0.47). P-cymene was proven to be the most toxic compound to human health. **Conclusion:** Excepting cumin aldehyde, the lack of the antiproliferation potential of these p-cymene derivatives was extensively demonstrated for the first time. Cuminaldehyde induced toxicity in a lung adenocarcinoma line, corroborating the literature. Six selective and three specific cell lines were proposed to evaluate the anticancer activity of the compounds in addressed studies, mainly involving those with Pa≥0.50. The inhibition of five targets seems to play a role in inducing HCCL antiproliferation. The ADMET estimated that cumin aldehyde, cuminol, and sodium cuminate are the safest compounds for human use.

Keywords: *p*-menthane- cumin seeds- cancer cells- antiproliferation- ADMET

Introduction

Several studies on *p*-cymene (4-isopropyltoluene, Compound 1 in Figure 1), a monoterpene with a *p*-menthane skeleton hydrocarbon structure, have reported its significant *in vitro* activity against specific cancer cells [1-3]. However, *p*-cymene has not been shown to induce antiproliferation in most human cancer cell lines (HCCLs) [1], even when its concentration is greater than 100 µM [2-4]. In addition, there have been studies on *in vitro* activity against cancer cells by natural 7-oxygenated derivatives of *p*-cymene, which are also *para-* (*p*-) oriented. Cuminaldehyde (2) is the most interesting of these derivatives because of its widely reported antiproliferative activity [5-9], whereas only one or no antiproliferation studies have been conducted for cuminic acid (3) [10] and cuminol (4), respectively.

A good advantage of the medicinal use of *p*-cymene and some of its derivatives is that they are widely
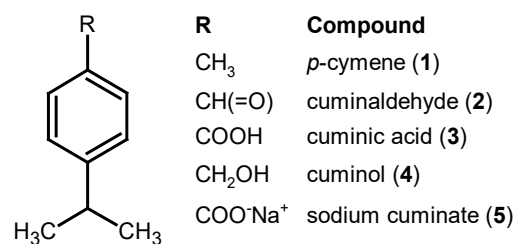

Figure 1. Chemical Structures of the Tested Compounds

*[1]Post-Graduate Program of Translational Research on Drugs and Medicines, Institute of Drug Technology, Oswaldo Cruz Foundation, Brazil. [2]Laboratory of Molecular Pharmacology, Department of Pharmacology, Institute of Drug Technology, Oswaldo Cruz Foundation, Brazil. [3]Laboratory of Technology for Biodiversity in Health, Department of Natural Products, Institute of Drug Technology, Oswaldo Cruz Foundation, Brazil. *For Correspondence: jose.mazzei@fiocruz.br*

available from natural sources [1-2, 5-6, 10-12]. For example, a richer source of *p*-cymene and its oxygenated derivatives is the essential oil of cumin seeds (*Cuminum cyminum*, family Apiaceae), whose contents can reach 3.5-12.0% *p*-cymene, 19.2-60.6% cuminaldehyde, 0.4-22.6% cuminol, and 1.4% cuminic acid [10-11, 13-16]. The species is native to Egypt, the Mediterranean region and South Asia [17], has a pantropical distribution and has been cultivated worldwide for commercial purposes, mainly in India [13-14, 18], achieving a worldwide production of 300,000 tons [13]. The anticancer activity of *C. cyminum* has been revised [18]. As an example of its use as a dietary protective alternative against cancer, cumin prevents the production of 1,2-dimethylhydrazine-induced colon cancer cells and the growth of breast cancer in rats [19]. Moreover, a nanoemulsion of cumin oil induced cytotoxicity against the SAS cell (tongue carcinoma) line (half maximal inhibitory concentration, IC_{50}, 1.5 µL/mL), whereas this effect was not detected in the non-malignant human line HEK-293 (embryonic kidney) [20].

The affordability, safety, and usefulness of this herb indicate its essential use as a natural substance for human medicine. Thus, studies on these cumin constituents should be considered to establish their potential anticancer action. The present study aimed to assess the *in vitro* and *in silico* antiproliferative activity of these compounds. In view of our expertise in terpene derivation (patent number WO2016/026014), we included the sodium salt of cuminic acid (5) in this investigation.

Materials and Methods

The menthane derivatives

The *p*-isomers of cymene, cuminol, cuminaldehyde, and cuminic acid were supplied with certification (GC purity) by Sigma-Aldrich (St. Louis, USA). Sodium *p*-cuminate was synthesized based on our patent of invention (number WO2016/026014), in which another *p*-menthane derivative was obtained by dropping 30% sodium methoxide on acid (100 mg/ml in methanol) in a stoichiometric quantity in an ice bath. The mixture was centrifuged at 4000 rpm at 15°C for 10 min, after which the supernatant was discarded. The precipitate was washed repeatedly with ice-cold methanol, followed by centrifugation until the crystals were no longer alkaline. The residual solvent was removed under a vacuum and a freeze dryer.

Chemical and thermal analysis

Gas chromatography (GC) was performed on an Agilent Technologies Model 7890 instrument with a 7683 injector at 280°C in split/splitless injection mode (25:10 split), and flame ionization detection at 300°C using a J&W Scientific HP-Innowax column (30 m×250 µm×0.250 µm) and helium as the carrier gas (1 ml/min). Programming: 50°C to 250°C at 20°C/min, held for 5 min. Fourier transform infrared spectrophotometry (FTIR, 4000-400 cm-1) was conducted at 1%w/w in spectroscopic grade KBr [21] in an IRAffinity–1S spectrometer (Shimadzu) at room temperature. Thermogravimetric analysis (TGA) was performed under N_2 from room temperature to 350°C

at 10°C/min on a Discovery TGA 550 analyzer.

Cell cultures

The HCCLs SK-MEL-28 (malignant melanoma), K562 (chronic myeloid leukemia), Lucena (derived from K562 cells that express P-glycoprotein (Pgp) and have a multidrug resistance phenotype), Jurkat (acute T cell leukemia), Caco-2 (colorectal adenocarcinoma), MDA-MB-231 (breast adenocarcinoma), THP-1 (acute monocytic leukemia), U87-MG (likely glioblastoma), Calu-3 (lung adenocarcinoma), and the non-malignant human line HEK-293 were used. The cell lines were seeded and subcultured in RPMI-1640 medium or Dulbecco's modified Eagle medium for suspension and adherent cell cultures, respectively, under a humidified atmosphere at 37°C and 5% CO_2. The media were supplemented with 10% fetal bovine serum, streptomycin, penicillin and trypsin.

MTT assay

Cells were seeded (100 µl, 2-5×10^4 cells/well) in 96-well plates containing medium and treated with compounds at 100 µM. The compounds were diluted previously in either water (sodium cuminate) or a 0.4% dimethyl sulfoxide (DMSO) in a phosphate-buffer (pH 7.2) solution. The total well volume was 200 µl. Staurosporine (1 µM in 0.4% DMSO) was applied as a positive control. The plates were kept at 37°C and 5% CO_2 for 48h.

The cell viability was determined using the 3-(4,5-dimethyltiazol-2-yl)-2,5-diphenyltetrazolium bromide (MTT) method [22]. MTT solution (20 µL, 2.5 mg/ml) was added to each well, followed by incubation at 37°C for 4h, centrifugation (1500 rpm, 10 min), and removal of the supernatant. After dilution of the formed formazan crystals with 100 µl DMSO, the absorbance at 540nm was read in a Victor X5 spectrophotometer (Perkin Elmer, USA). Cell viability (%) was calculated as the ratio between the absorbance of the sample well and that of the negative control and was expressed as the mean ± standard error of the mean (SEM) of two independent experiments performed in triplicate. The data were compared statistically to the negative control using analysis of variance (ANOVA) and, if pANOVA < 0.05, followed by Dunnett´s test ($p < 0.05$).

Determination of apoptosis/necrosis by flow cytometry analysis (Annexin V/propidium iodide)

For apoptosis determination, Calu-3 cells were seeded (0.5 mL at 2×10^5 cells/mL) in a 24-well plate and incubated (24h) in 5% CO_2 atmosphere at 37°C. Cells were treated in triplicate with cuminaldehyde 100 µM, for another 24 and 48h. Controls consisted of cells treated with vehicle (negative) or staurosporine (positive, 0.5 µM). Staining was performed as recommended by the manufacturer (Dead cell apoptosis kit – Invitrogen, Oregon, USA). False-positive interpretation due to propidium iodide (PI) over-labeling was avoided by adding it (Sigma-Aldrich) to samples at the acquisition time (within 5 min). Data was acquired (10^4 gated events) with a FACScalibur flow cytometer and analyzed using CellQuestTM software (Becton Dickinson, Bioscience, San Jose, USA). Early

apoptosis, late apoptosis, and necrosis were classified by the cell staining profile: Annexin V+/PI-, Annexin V+/PI+, and Annexin V-/PI+, respectively.

In silico evaluation of cytotoxicity in human cancer cell lines

The cytotoxicity of the compounds on human cancerous and non-malignant cell lines and possible mechanisms of action were evaluated *in silico* through the CLC-Pred 2.0 (Cell Line Cytotoxicity Predictor, way2drug.com/cell-line) webserver [23]. For that, the canonical SMILES (simplified molecular-input line-entry system) of those compounds was obtained from the PubChem database (pubchem.ncbi.nlm.nih.gov). Cytotoxicity is predicted in terms of Pa (probable activity) and Pi (probable inactivity), which range from 0 to 1. A threshold is considered when the value Pa is equal to Pi. When Pa > Pi, the compound is considered cytotoxic or active. The compound is considered highly cytotoxic and highly active when $Pa \geq 0.5$ [23]. Correlations (rho coefficient) between Pa of mechanisms of action and Pa of activity against HCCLs were examined through Spearman's Rank Order using GraphPad Prism 8 program.

Pharmacological and toxicity prediction

The drug-likeness, human gastrointestinal (GI) absorption, blood-brain barrier (BBB) permeability, and inhibition or induction of Pgp and human cytochrome P450 (CYP) enzymes by the compounds were estimated *in silico* through the CLC-Pred 2.0 (only on CYPs), SwissADME (swissadme.ch) [24, 25] and ADMETlab 2.0 (admetmesh.scbdd.com) [26] webservers. The toxicological profile of the compounds was investigated by applying the latter and the Protox-II server (tox. charite.de) [27]. The canonical SMILES was applied as the input data.

Results

Evaluation of the literature

The outcome results of the *in vitro* studies on the antiproliferative activity of the compounds on cancer cell lines are shown in Table 1. *p*-Cymene induced antiproliferation on B16-F10 (murine melanoma, IC_{50} of 20.06 μg/ml) [1] and MCF-7 (human breast adenocarcinoma, IC_{50} of 5.11 mM) lines [2]; the suppression of 12-*O*-tetradecanoyl-phorbol-13-acetate (TPA)-induced MMP-9 expression via the inhibition of the ERK1/2 and p38 MAPK signaling pathways in HT-1080 (epithelial fibrosarcoma) lines; and the inhibition (87% at 600 μM) of the invasive activity of these cells at the basal level [3]. However, *p*-cymene has not been shown to induce antiproliferation in several HCCLs, such as HT-1080 (≤ 800 μM) [3], A549 (lung adenocarcinoma), DLD-1 (colon adenocarcinoma) (≤ 100 μM), Hep3B (≤ 10 μM) [2, 4], and HepG2 (hepatic carcinomas) (≤ 25 μg/ml) [1].

Cuminaldehyde has induced cancer cell death *in vitro* and *in vivo* in COLO 205 (colorectal adenocarcinoma) (IC_{50} 16.31 μM, and a 48.9 - 69.4% reduction in tumors) [5], NCI-H520 (lung squamous carcinoma, with an IC_{50}

33.29 μM, and ~45%) [6], and A549 (IC_{50} 18.33 μM, ~50%), which act on topoisomerases 1 and 2 [7] and suppress, in a dose-dependent manner (≤ 0.25 mM), the formation of melanin in B16-F10 cells without affecting cell growth [9]. This compound was not toxic to A375 (malignant melanoma) [9] or to pheochromocytoma cells (a tumor of the adrenal glands) in PC12 mice [8]. Only one report evaluated cuminic acid activity, and the induction of cytotoxicity was observed in HepG2 (IC_{50} 6 μg/ml) cells without further details on the determination [10].

The *in silico* analysis performed by applying the CLC-Pred server on the cell lines that were studied *in vitro* until now is also shown in Table 1. Interestingly, Pa < 0.01 values were found, except for the activity of *p*-cymene in K562, which presented a Pa of 0.24; however, an expected threshold (Pi (0.13) was close to Pa) corroborated the absence of activity [1]. In addition, this absence of activity predicted through CLC-Pred is not consistent with what was observed in some studies in the literature with cuminaldehyde and cuminic acid, which indicated *in vitro* antiproliferative activity (Table 1).

Chemical characterization of p-cymene and its derivatives

Content by GC. The commercial compounds showed > 97% purity (Supplemental Table S1), according to the certification from the supplier. It was possible to detect other constituents that did not exceed 2.5% under GC conditions. *p*-Cymene contained only one other compound, whereas cuminol contained the most different constituents.

Synthesis of the sodium cuminate. The controlled addition of sodium methoxide over cuminic acid promoted, by ionic exchange, the formation of salt insoluble in alcohols, which precipitated, displacing the reaction. The chemical composition of the salt was confirmed by FTIR spectroscopy and no traces of cuminic acid or residual solvent were detected in the TGA profile, which indicated that the salt contained 3.2% water (Supplemental Figure S1 and S2, respectively).

In vitro antiproliferation in cancer cell lines

Using the MTT assay, *in vitro* antiproliferative effects of the compounds were investigated at a concentration of 100 μM in the non-malignant human cell line HEK-293 and in nine HCCLs: SK-MEL-28, K562, Lucena, Jurkat, Caco-2, MDA-MB-231, THP-1, U87-MG, and Calu-3 (Figure 2). The HCCLs employed herein would be sensitive to the compounds studied since an *in silico* evaluation (Figure 3a) indicated a Pa of 0.12-0.47, except for two of them. In addition, the moderate and selective *in silico* activity of cuminaldehyde on the cytotoxicity of the MDA-MB-231 cell line, which was predicted using the CLC-Pred server (Pa 0.47, Figure 3a), could be better investigated *in vitro*.

The performance of the *in vitro* assay was successful for all cell lines as seen by the cytotoxic responses to the positive control (p < 0.001, ANOVA followed by Dunnett´s test) (Figure 2). For Caco-2 (pANOVA < 0.05), HEK-293, MDA-MB-231, and Calu-3 (pANOVA < 0.001) lines, the null hypothesis in the differences in the cell viability means among the treatments with the monoterpenes and

Table 1. Reported Studies on *in vitro* Antiproliferation by *p*-Cymene and 7-Oxygenated Derivatives on Cancer Cell Lines and the *in silico* Activity Predicted.

Compound	Ref.	Cell line	IC$_{50}$ (µM)	Incubation time	Assay, statistical test (sample size)	CLC-Pred (Pa*)
p-Cymene	[1]	HepG2	nc (max. 186)	24h	Methyl-[3H]-thymidine,	< 0.01
		K562	nc (max. 186)		ANOVA/Newman-Keuls	0.24 (Pi 0.13)
		PBMC	nc (max. 186)		(2 × 2-3)	-
		B16-F10 (murine)	149			-
	[2]	MCF-7	5110 ± 500	72h	Sulforhodamine B assay,	< 0.01
		A549, HepG2, Hep3B	nc (max. 10,000)		regression analysis (3)	< 0.01
	[3]	HT-1080	nc (max. 800)	24h	CellTiter-Blue reagent, nd (6)	< 0.01
	[4]	A549, DLD-1	nc (max. 100)	48h	Resazurin reduction, nd (3)	< 0.01
Cuminaldehyde	[5]	COLO 205	16.31	48h	XTT, ANOVA/Bonferroni (3)	< 0.01
	[6]	NCI-H520	33.29	48h	XTT, ANOVA/Bonferroni (3)	< 0.01
	[7]	A549	18.33	48h	XTT, ANOVA/Tukey (3)	< 0.01
	[9]	A375	nc (max. 0.5)	72h	MTT, F-/t-Student-Welch tests (2)	< 0.01
		B16-F10 (murine)	700			-
	[10]	MCF-7	53	48h	MTT, ANOVA/Tukey (3)	< 0.01
Cuminic acid	[11]	HepG2	6	nd	MTT, nd	< 0.01

Ref, See references. *Probability to be active (Pa) or inactive (Pi, indicated when close to Pa); nc, Not cytotoxic concentration was reached (maximum (max.) concentration tested); nd, Not declared.

the negative control was rejected in one-way ANOVA. However, no compound decreased cell viability ($\geq$ 97%) significantly (p $\geq$ 0.99, post-hoc Dunnett′s test) for HEK-293 and Caco-2. Only cuminaldehyde induced a decrease in the proliferation of MDA-MB-231 (15 ± 7%) and Calu-3 (31 ± 5%) cells, but significantly only for the latter (p = 0.068 and p < 0.001, respectively, by post-hoc Dunnett′s test), corroborating the activity previously reported for another lung adenocarcinoma line in the literature [7]. Notably, cuminaldehyde did not induce the antiproliferation of the non-malignant cell line. In view of both the *in vitro* slight induction of cuminaldehyde against

MDA-MB-231 and the *in silico* prediction (Figure 3a), the *in vitro* assay was also performed at higher doses (100, 200, and 400 µM, six determinations each) but no antiproliferation was observed: cell viability of 102 ± 4%, 110 ± 2%, and 111 ± 3%, respectively (p = 0.29, one-way ANOVA). The antiproliferation effect of cuminaldehyde on Calu-3 cells was investigated with a viability curve (1-300 µM, 7 points), and a dose-dependent response was obtained, which reached significance (p < 0.001, ANOVA followed by Dunnett's test) at concentrations $\geq$ 100 µM. The IC$_{50}$ of 650 µM was estimated. There was neither enhanced apoptosis nor necrosis of Calu-3 cells

Table 2. Biological Targets (and Correlations) on Human Cancerous Cell Lines with High Interaction (Probability $\geq$ 0.7) of *p*-Cymene and Its 7-Oxygenated Derivatives.

Mechanism of action (human protein code)	*p*-Cymene	Cuminaldehyde	Cuminic acid	Cuminol	Spearman Cell lines (Pa $\geq$ 0.5)	*p*	*rho*
DNA-3-methyladenine glycosylase (AAG)	0.86	0.34	0.92	0.83	all	> 0.33	-
Regulator of G-protein signaling 17 (RGS17)	0.81	0.58	0.84	0.75	all	> 0.33	-
Solute carrier family 22 member 2 (SCL22A2)	0.76	0.31	0.61	0.55	CAKI-2, SK-LU-1,		
Cytochrome P450 2A13 (CYP2A13)	0.76	0.43	0.60	0.57	and RCC4	0.08	1.00
Deoxyhypusine hydroxylase (DOHH)	0.70	0.33	0.56	0.50			
Emopamil-binding protein-like (EBPL)	0.76	0.45	0.53	0.65	SK-MEL-1	0.08	1.00
Histamine *N*-methyltransferase (HNMT)	0.75	0.45	0.57	0.65			
DNA polymerase k (POLK)	0.74	0	0.71	0.68	all	> 0.33	-
RecQ-like DNA helicase (BLM)	0.52	0.34	0.70	0.56	all	> 0.33	-
Solute carrier organic anion transporter family member 2B1 (SLCO2B1)	0.44	0	0.70	0.64	all	> 0.33	-
Endoplasmic reticulum chaperone BiP (HSPA5)	0.37	0	0.42	0.73	all	> 0.33	-
Aldehyde dehydrogenase 1A1 (ALDH1A1)	0.19	0.77	0	0	-*	-	-

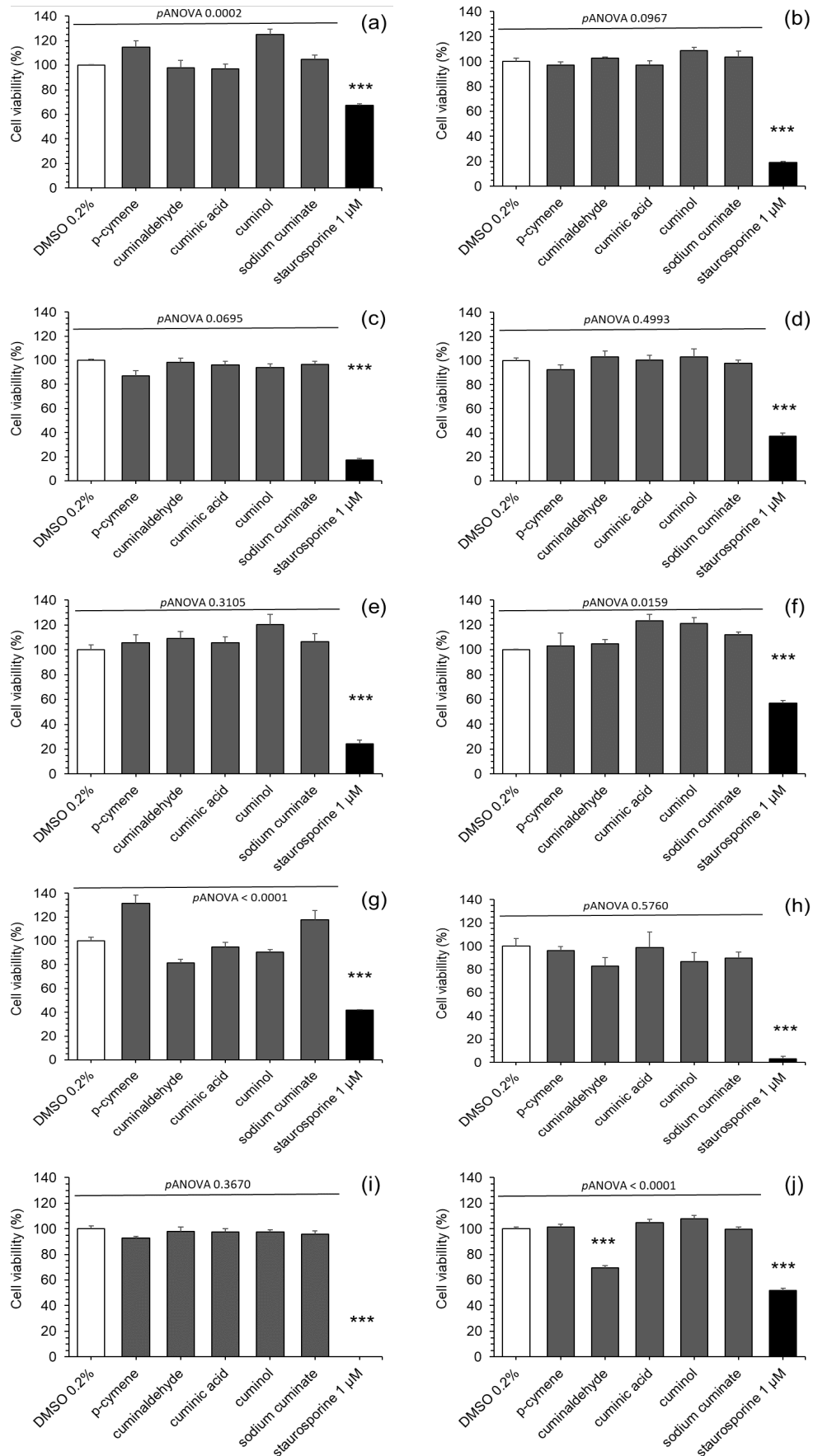* A few data of ALDH1A1 to calculate correlation.

Figure 2. Cell Viability (mean ± SEM, n = 6) through MTT assay after 48h-Incubation without (DMSO 0.2%) and with the Compounds at 100 μM. Positive control: staurosporine 1 μM. Significant cytotoxicity are indicated by ANOVA (p-values above the means) followed by Dunnett´s test (***p < 0.001) to the negative control (DMSO 0.2%). Cell lines: (a) HEK-293, (b) SK-MEL-28, (c) K562, (d) Lucena, (e) Jurkat, (f) Caco-2, (g) MDA-MB-231, (h) THP-1, (i) U87-MG, and (j) Calu-3.
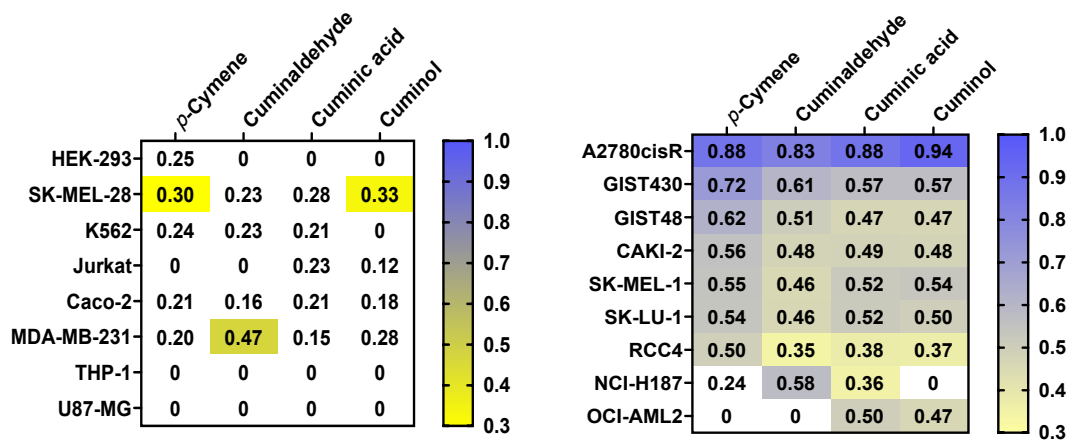
Figure 3. Heatmap on the Toxicity *in silico* of Human Cancer Cell Lines, Predicted Applying CLC-Pred, by the Compounds of the Present Study: (a) on cell lines assayed *in vitro* in this study; (b) of probability of Pa ≥ 0.5 when at least one of the compounds is active. Lucena and Calu-3 do not belong in the CLC-Pred 2.0 library. For SK-MEL-28, Jurkat, and MDA-MB-231, the highest probability of cytotoxicity is presented including the NCI-60 tumor cell-line panel.

after treatment with cuminaldehyde at 100 µM (p > 0.05) for 24h and 48h, as determined by flow cytometry analysis (Annexin V/propidium iodide) (Figure 4). However, when the assay was performed with cuminaldehyde at 300 µM (half of the IC_{50}), only a significant (p < 0.01, ANOVA followed by Dunnett´s test) but very low Calu-3 necrosis

(10.2 ± 2.6%, Annexin V-/PI+ cells) was detected.

Prediction of cytotoxicity in cancer cell lines and biological targets

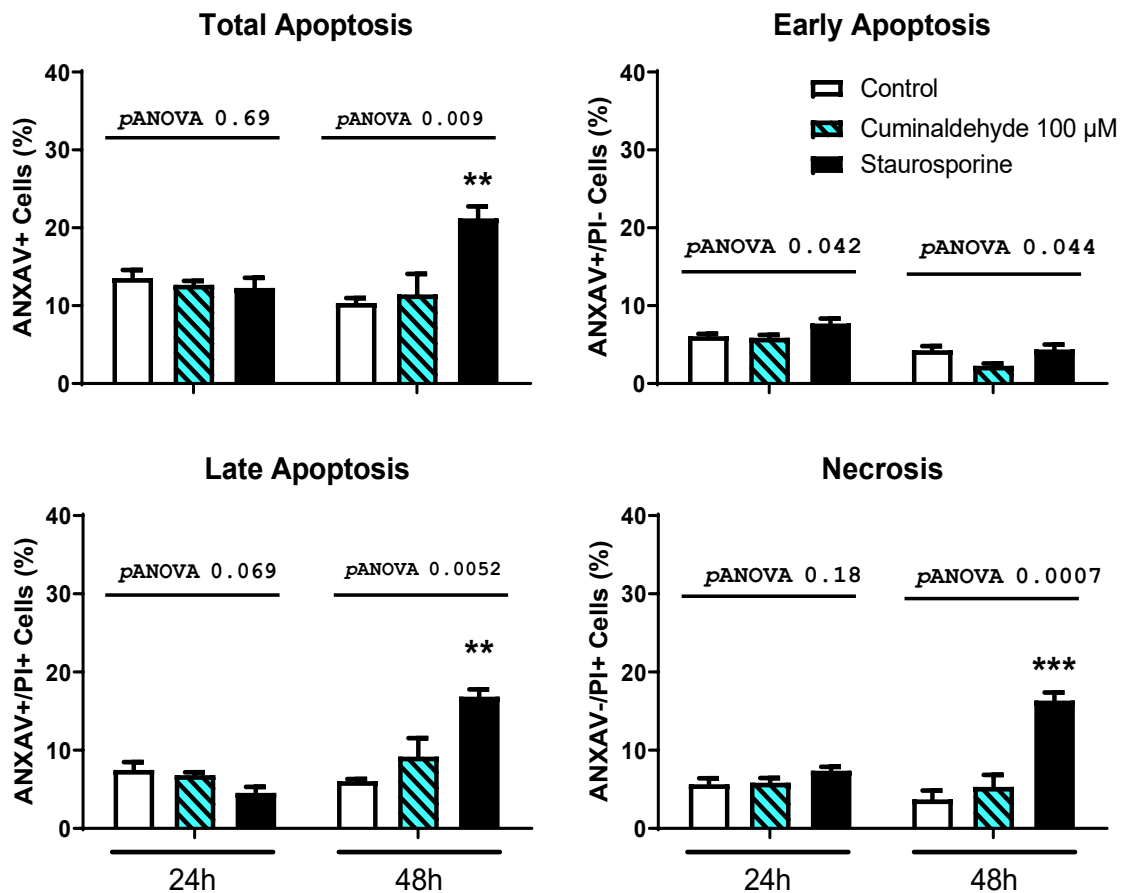The potential cytotoxicity of the compounds was determined on CLC-Pred in terms of their Pa and Pi values

Figure 4. Cell Staining Profile Apoptosis/Necrosis (10^4 Gated Events, Annexin V/Propidium Iodide, Dead Cell Apoptosis kit, mean ± SEM, n = 3) by flow cytometry analysis of Calu-3 line (2×10^5 cells/mL) after 24h- and 48h-periods of incubation at 5% CO_2 atmosphere and 37 °C. Negative and positive controls consisted in cells incubated with vehicle (blank bars) or staurosporine 0.5 µM (black bars), respectively. Significant events indicated by ANOVA (p values above the means) followed by Dunnett´s test (**p < 0.01 or ***p < 0.001) to the negative control.

against cancer cell lines. According to these data, among cancer cell lines with Pa $\geq$ 0.50 for at least one of the substances (Figure 3b), the very low Pi ($\leq$ 0.11) implied a high Pa-Pi (0.30-0.60), leading us to consider them to have low-to-high cytotoxic activity. Given that the server does not analyze ionic molecules [23], the cytotoxicity of sodium cuminate could not be predicted. Each compound had similar cytotoxic effects on six cell lines (cells, Pa): A2780cisR (cisplatin-resistant ovarian carcinoma, $\geq$ 0.83), GIST430 (gastrointestinal stromal tumor, $\geq$ 0.57), GIST48 (gastrointestinal stromal tumor, $\geq$ 0.47), CAKI-2 (kidney carcinoma, $\geq$ 0.48), SK-MEL-1 (metastatic melanoma, $\geq$ 0.46), and SK-LU-1 (lung adenocarcinoma, $\geq$ 0.46). Specific cell lines were identified as susceptible to each compound (cells, Pa): *p*-cymene against RCC4 (clear cell renal cell carcinoma, 0.50), cuminaldehyde against NCI-H187 (small cell lung carcinoma, 0.58), and cuminic acid and cuminol (Pa $\geq$ 0.47) against OCI-AML2 (adult acute myeloid leukemia) cells.

Considering Pa $\geq$ 0.7 of interactions for at least one of the substances, the CLC-Pred indicated only twelve biological targets in HCCLs (Table 2). Nevertheless, no families were attributed predominantly as one preferential path. Some correlations (p = 0.08, Spearman rho = 1.00) between Pa of these mechanisms and four of the potential HCCLs (Figure 3b) suggest the susceptible targets to cuminic acid (SCL22A2, CYP2A13, and DOHH) or cuminol (EBPL and HNMT) together with the highest affinity by *p*-cymene. Other significant correlations were not achieved (p > 0.33), but HSPA5 and the electrochemical transporter SLCO2B1 were pointed out as selective targets, respectively, to cuminol and concomitantly this compound and cuminic acid.

Pharmacological and toxicity estimation

The measurements of the *in silico* ADMET properties are provided in the Supplementary Material. Each compound exhibited molecular and structural characteristics similar to those of known drugs, according to the Lipinski, Verber, Egan, and GSK criteria (Table S2), suggesting that they have the potential to be effective medicines. The drug-likeness indicated by Ghose, Pfizer, and Golden Triangle criteria were controversial. The compounds showed high GI absorption, except for *p*-cymene according to the SwissADME server, suggesting that it is difficult for this compound to be absorbed across the intestine after oral intake. Cuminic acid and sodium cuminate showed a low ability to cross the BBB according to the ADMETlab server (Table S3). However, in the BOILED-Egg (Brain Or Intestinal EstimateD) model, all the compounds had high GI absorption and BBB permeability values (Figure S3). Sodium cuminate was the only one predicted to be excreted effectively from the central nervous system (CNS) by Pgp.

The servers indicated that none of the terpenes studied may be substrates or inhibitors (probability < 70%) of 8 among 13 CYP enzymes investigated. The servers showed inhibitory effects (> 50% probability) of *p*-cymene (divergent results) on CYP1A2, CYP2C19 (also cuminaldehyde and cuminol), CYP2D6 (also cuminol), and CYP2A13 (also cuminic acid and cuminol)

(Table S4). *p*-Cymene showed acute toxicity (class I) and carcinogenicity. Cuminic acid and sodium cuminate may induce hepatotoxicity (Table S5). None of the substances showed a high probability (> 50%) of activity for the Tox21 assays.

Discussion

Given the highly divergent results obtained with our *in silico* analysis versus the *in vitro* data found in the literature on the level of antiproliferative activity of those compounds in the same HCCLs, an overall evaluation of the studies performed thus far is needed. However, only the activity of cuminaldehyde has been well investigated previously, and appropriate statistical tests were applied to determine the significance of the differences between experimental groups via ANOVA followed by a post-hoc test [5-7, 28]. Interestingly, apart from only two studies, which performed the test with two experiments in duplicate or triplicate [1] or in sextuplicate [28], most of the studies were designed with few replicates, triplicate [2, 4-7] or duplicate [9], or did not declare the number of replicates [10]. Given that the pooled standard deviation of 9% antiproliferation was achieved in MTT assay, we also performed it at n = 6 to increase the power of post-hoc statistical tests of 78% to 97% (determined by G*Power) in detecting significatively (p < 0.05) at least 20% of the activity, and so contributing to the sensitivity.

Reviews on the *in vitro* cytotoxicity of *p*-cymene are found in the literature but include the ruthenium complexes of this compound [29]. Table 1 shows studies on the cancer antiproliferation of *p*-cymene without complexation, and few studies have investigated its activity. Antiproliferation of *p*-cymene and cuminic acid has been demonstrated only in single HCCLs, respectively, at high concentrations (IC$_{50}$ 5.11 $\pm$ 0.5 mM against MCF-7) [2] or reaching IC$_{50}$ 6 μM against HepG2 without detail [11] (Table 1).

In the present study, the HCCLs cytotoxicity of cuminol and sodium cuminate was investigated via *in vitro* and *in silico* approaches for the first time. In addition, the potential of cuminic acid was investigated extensively. The compounds were tested *in vitro* following a primary screening similar to the NCI-60 standard protocol for HCCL anticancer drug screening [30]. The NCI-60 screening is initially performed at a single concentration of 10 μM and, when toxicity is confirmed, the compound is tested in 5 of 10-fold dilutions up to 100 μM. Considering the reported absence of the *p*-cymene-induced toxicity (Table 1) in HCCLs, the single dose of 100 μM was primarily selected for all compounds in the present work and, considering the low toxic potential observed, the dilutions were extended to higher doses, to establish the IC$_{50}$ values (successfully for Calu-3 but not for MDA-MB-231). Both very low (Calu-3) or absence (other cell lines) toxicity at 100 μM and the high IC$_{50}$ value (Calu-3) indeed demonstrated very low toxicity of the compounds against the HCCLs assayed. The antiproliferative effect of cuminaldehyde on Calu-3 cells was significant and has not been reported until now. The absence of *in vitro* activity in the other cell lines, including cuminol, cuminic acid, and sodium cuminate, corroborates the *in silico*

prediction that indicated that the Pa values were ≤ 0.33 (Figure 3a). The *in silico* prediction showed Pa of 0.47 of cuminaldehyde activity for MDA-MB-231, but it was not significantly demonstrated *in vitro* ($p > 0.05$) by the post-hoc test. CLC-Pred estimates Pa by the qualitative prediction based on 128,545 structures [24]. However, the prediction against an NCI-60 tumor cell-line panel based on the Developmental Therapeutics Program's NCI-60 data demonstrated a lower Pa of cuminaldehyde against MDA-MB-231 (0.20) at both 1 and 10 nM, which is reasonable with our findings (no antiproliferation *in vitro*). Thus, a careful analysis must be considered to select HCCLs in future investigations, while considering Pa > 0.50 as an example of criterion. Our results also indicated the absence of antiproliferation by *p*-cymene at 100 µM, corroborating most of the data in the literature for eight other HCCLs (Table 1).

In view of these results, it is essential to determine which cell lines can be more sensitive to these abundant natural compounds in a prospective study. In fact, according to the predictions obtained with the CLC-Pred server, each compound in this study exhibited high cytotoxicity against A2780cisR cells (Pa ≥ 0.83). The antiproliferation of acyclic monoterpenes against the same cell line was recently reported [31], indicating that the activity is possibly associated with the structure of the monoterpene chain.

It has been reported that cuminaldehyde significantly changes the expressions of both pro-apoptotic (Bax and Bak) and anti-apoptotic (Bcl-2 and Bcl-XL) proteins in treated NCI-H520 cells [6]. CLC-Pred displayed fewer promising potential targets of cuminaldehyde than those for other compounds, probably due to less information about aldehyde drugs involved in the bases applied for the prediction. The server indicated only the oxidoreductase ALDH1A1 as a selective target for cuminaldehyde. The *in vitro* results corroborate the low probability of being active against proliferation for inducing only necrosis on one of the nine HCCLs. Correlation analysis indicated three and two targets more susceptible to cuminic acid and cuminol, respectively, and it was possible to identify the HCCLs for future investigations on this selectivity. HCCLs with the three highest Pa of toxicity (A2780cisR, GIST430, and GIST48, Figure 3) were not associated selectively with any interactions of Pa ≥ 0.7, possibly due to the contribution of interactions with other targets. It is interesting to note the higher affinity of all compounds at the AAG, which may make the tumor cell more sensitive to alkylating anticancer agents [32], and at the RGS17, which positive regulation is associated with an aggressive cancer phenotype [33].

Because of the promising cancer antiproliferation of the compounds in this study, an *in silico* evaluation was necessary to determine their drug-likeness and ADMET properties. Notably, in CYPs, *p*-cymene and cuminic acid are the most and the least active compounds, respectively. Interestingly, lung cancer can be associated with the polymorphisms of CYP2A13 [34], in which selective high inhibition (except for cuminic acid) was emphasized among the preferential mechanisms of action (Table 2). Cuminaldehyde and cuminol presented the best ADMET

parameters, especially with respect to drug-likeness properties and gastrointestinal absorption, mainly because they did not present high toxicity in the models evaluated, suggesting that these compounds among the compounds assayed are the most promising for oral drug development. Suitable formulations that increase the GI residence time of these compounds, as in modulation of mucoadhesion, can be developed especially against gastrointestinal stromal tumors. Otherwise, regional drug targeting in the gastrointestinal tract [35] may be considered a reliable pharmaceutical approach in other cancer therapies, given the moderate-to-high probability of metabolisms by cytochromes P450. The activity of the cuminaldehyde on HCCL NCI-H520 [6], as well as on the promisingly proposed NCI-H187 and its role as a CYP2A13 inhibitor, emphasizes the importance of studying it in delivery strategies against lung cancer, as by inhalation.

Furthermore, sodium cuminate has demonstrated to be a preferable drug candidate for controlled volume intermittent infusions as delivery strategy because of being less toxic, of low BBB permeability, susceptible to Pgp efflux from the CNS, and its acid form showing similar activity to *p*-cymene against six of nine promising HCCLs.

Overall, these findings can contribute to the value of designing forthcoming studies on the antiproliferation of cumin constituents derived from *p*-cymene, mainly cuminaldehyde and sodium cuminate, as promising compounds to fight against cancer.

Author Contribution Statement

Conceptualization and design: ALFS (*in vitro* experiments) and NPS (*in silico* determinations). Formal biological assays, methodology, and investigation: NPS, GWSS, APGAF, and ALFS. Design conception, interpretation of chemical assays, and writing: NPS and JLM. Supervision, validation, funding acquisition, and resources: ALFS and JLM. Project administration: JLM. All authors have read and agreed to the published version of the manuscript and accept responsibility for all aspects of the work.

Acknowledgements

1. Ferraz RPC, Bomfim DS, Carvalho NC, Soares MBP, Silva TB, Machado WJ, et al. Cytotoxic effect of leaf essential oil of *Lippia gracilis* Schauer (Verbenaceae). Phytomed. 2013;20(7):615-21. https://doi.org/10.1016/j.phymed.2013.01.015.

2. Fitsiou E, Anestopoulos I, Chlichlia K, Galanis A, Kourkoutas I, Panayiotidis MI, et al. Antioxidant and antiproliferative properties of the essential oils of *Satureja thymbra* and *Satureja parnassica* and their major constituents. Anticancer Res. 2016;36(11):5757-63. https://doi.org/10.21873/anticanres.11159.

3. Li J, Liu C, Sato T. Novel antitumor invasive actions of *p*-cymene by decreasing MMP-9/TIMP-1 expression ratio in human fibrosarcoma HT-1080 cells. Biol Pharm Bull. 2016;39(8):1247-53. https://doi.org/10.1248/bpb.b15-00827.

4. Bourgou S, Pichette A, Marzouk B, Legault J. Bioactivities of black cumin essential oil and its main terpenes from Tunisia. S Afr J Botany. 2010;76(2):210-6. https://doi.org/10.1016/j.sajb.2009.10.009.

5. Tsai KD, Liu YH, Chen TW, Yang SM, Wong HY, Cherng J, et al. Cuminaldehyde from *Cinnamomum verum* induces cell death through targeting topoisomerase 1 and 2 in human colorectal adenocarcinoma COLO 205 cells. Nutr. 2016;8(6):318. https://doi.org/10.3390/nu8060318.

6. Yang SM, Tsai KD, Wong HY, Liu YH, Chen TW, Cherng J, et al. Molecular mechanism of *Cinnamomum verum* component cuminaldehyde inhibits cell growth and induces cell death in human lung squamous cell carcinoma NCI-H520 cells *in vitro* and *in vivo*. J Cancer. 2016;7(3):251-61. https://doi.org/10.7150/jca.13689.

7. Chen TW, Tsai KD, Yang S, Wong WY, Liu YH, Cherng J, et al. Discovery of a novel anti-cancer agent targeting both topoisomerase I & II as well as telomerase activities in human lung adenocarcinoma A549 cells *in vitro* and *in vivo*: *Cinnamomum verum* component cuminaldehyde. Curr Cancer Drug Targets. 2016;16(9):796-806. https://doi.org/10.2174/1568009616666160426125526.

8. Morshedi D, Aliakbari F, Tayaranian-Marvian A, Fassihi A, Pan-Montojo F, Pérez-Sánchez H, et al. Cuminaldehyde as the major component of *Cuminum cyminum*, a natural aldehyde with inhibitory effect on alpha-synuclein fibrillation and cytotoxicity. J Food Sci. 2015;80(10):H2336-45. https://doi.org/10.1111/1750-3841.13016.

9. Nitoda T, Fan MD, Kubo I. Effects of cuminaldehyde on melanoma cells. Phytother Res. 2008;22(6):809-13. https://doi.org/10.1002/ptr.2374.

11. Abdel-Aal MM, El-Shafae AM, Abou-Hashem MM, Ateya AEM, Fekry M. Bioguided screening for cytotoxic active constituents of *Cuminum cyminum* volatile oil. Zagazig J Pharm Sci. 2018;27(2):1-7. https://doi.org/10.21608/zjps.2018.48350.

10. Ghosh S, Kundu M, Dutta S, Mahalanobish S, Ghosh N, Das J, et al. Enhancement of anti-neoplastic effects of cuminaldehyde against breast cancer via mesoporous silica nanoparticle based targeted drug delivery system. Life Sci. 2022;298:120525. https://doi.org/10.1016/j.lfs.2022.120525.

12. Valizadeh S, Mahmodi R, Fakheri T, Katiraie F, Rahmani V. Investigating the phytochemical, antibacterial and antifungal effects of *Thymus vulgaris* and *Cuminum cyminum* essential oils. Med Lab J. 2016;10(1):36-43. http://doi.org/10.18869/acadpub.mlj.10.1.36.

13. Srinivasan K. Cumin (*Cuminum cyminum*) and black cumin (Nigella sativa) seeds: Traditional uses, chemical constituents, and nutraceutical effects. Food Qual Saf. 2018;2(1):1-16. https://doi.org/10.1093/fqsafe/fyx031.

14. Agarwal U, Pathak DP, Kapoor G, Bhutani R, Roper R, Gupta V, et al. Review on *Cuminum cyminum* – Nature's magical seeds. J Chem Pharm Res. 2017;9(9):180-7.

15. Belal AA, Ahmed FBM. Effect of *Cuminum cyminum* L. oil on the chemical properties of sunflower reused oil. J Food Nutr Sci. 2017;5(4):162-6. https://doi.org/10.11648/J.JFNS.20170504.14.

16. Hartini AD, Jiwanakusuma P, Margono A, Suprastiwi E. Antifungal effect of *Cuminum cyminum* extract on *Candida albicans* ATCC 10231 (experimental laboratory). J Int Dental Med Res. 2019;12(1):101-4.

17. Rana VS. Chemical composition of the essential oil of *Cuminum cyminum* L. seeds from Western India. J Med Plants By-Prod. 2014;2:207-10. https://doi.org/10.22092/jmpb.2014.108735.

18. Tabarsa M, You S, Yelithao K, Palanisamy S, Prabhu NM, Nan M. Isolation, structural elucidation and immuno-stimulatory properties of polysaccharides from *Cuminum cyminum*. Carbohydr Polym. 2020;230:115636. https://doi.org/10.1016/j.carbpol.2019.115636.

19. Al-Snafi AE. The pharmacological activities of *Cuminum cyminum* - A review. IOSR J Pharm. 2016;6:46–65.

20. Nalini N, Manju V, Menon VP. Effect of spices on lipid metabolism in 1,2-dimethylhydrazine-induced rat colon carcinogenesis. J Med Food. 2006;9(2):237-45. https://doi.org/10.1089/jmf.2006.9.237.

21. Nirmala MJ, Durai L, Rao KA, Nagarajan R. Ultrasonic nanoemulsification of *Cuminum cyminum* essential oil and its applications in medicine. Int J Nanomedicine. 2020;15:795-807. https://doi.org/10.2147/ijn.s230893.

22. Silverstein RM, Bassler GC, Morrill TC. Spectrometric identification of organic compounds. 4th ed. Singapore: John Wiley & Sons; 1981. cap. 3.

23. Mosmann T. Rapid colorimetric assay for cellular growth and survival: Application to proliferation and cytotoxicity assays. J Immunol Methods. 1983;65(1-2):55-63. https://doi.org/10.1016/0022-1759(83)90303-4.

24. Lagunin AA, Rudik AV, Pogodin PV, Savosina PI, Tarasova OA, Dmitriev AV, et al. CLC-Pred 2.0: A freely available web application for *in silico* prediction of human cell line cytotoxicity and molecular mechanisms of action for druglike compounds. Int J Mol Sci. 2023;24(2):1689. https://doi.org/10.3390/ijms24021689.

25. 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. 2017;7:42717. https://doi.org/10.1038/srep42717.

26. Daina A, Zoete V. A BOILED-Egg to predict gastrointestinal absorption and brain penetration of small molecules. ChemMedChem. 2016;11(11):1117-21. https://doi.org/10.1002/cmdc.201600182.

27. Xiong G, Wu Z, Yi J, Fu L, Yang Z, Hsieh C, et al. ADMETlab 2.0: an integrated online platform for accurate

and comprehensive predictions of ADMET properties. Nucleic Acids Res. 2021;49(W1):W5-14. https://doi. org/10.1093/nar/gkab255.

28. 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. https://doi. org/10.1093/nar/gky318.

29. Balahbib A, Omari NE, Hachlafi NE, Lakhdar F, Menyiy NE, Salhi N, et al. Health beneficial and pharmacological properties of *p*-cymene. Food Chem Toxicol. 2021;153:112259. https:// doi.org/10.1016/j.fct.2021.112259

30. National Cancer Institute. Standard operating procedures for sample preparation (NCI-60 HTS384) [Internet]. NCI. Available from: https://dtp.cancer.gov/discovery_development/nci-60/handling_HTS384.htm.

31. Dascalu D, Isvoran A, Ianovici N. Predictions of the biological effects of several acyclic monoterpenes as chemical constituents of essential oils extracted from plants. Molecules. 2023;28(12):4640. https://doi.org/10.3390/molecules28124640.

32. Roth RB, Samson LD. 3-Methyladenine DNA glycosylase-deficient AAG null mice display unexpected bone marrow alkylation resistance. Cancer Res. 2002;62(3):656-60.

33. Hayes MP, Roman DL. Regulator of G protein signaling 17 as a negative modulator of GPCR Signaling In Multiple Human Cancers. AAPS J. 2016;18(3):550-9. Hhttps://doi. org/10.1208/S12248-016-9894-1.

34. Chiang HC, Lee H, Chao HR, Chiou YH, Tsou TC. Pulmonary CYP2A13 levels are associated with early occurrence of lung cancer-its implication in mutagenesis of non-small cell lung carcinoma. Cancer Epidemiol. 2013;37(5):653-9. https://doi. org/10.1016/j.canep.2013.04.010.

35. Hua S. Advances in oral drug delivery for regional targeting in the gastrointestinal tract - influence of physiological, pathophysiological and pharmaceutical factors. Front Pharmacol. 2020;11:524. https://doi. org/10.3389/fphar.2020.00524.