

Novel water-soluble lignin derivative BP-Cx-1: identification of components and screening of potential targets *in silico* and *in vitro*

Elena I. Fedoros^{1,5}, Alexey A. Orlov², Alexander Zhrebker^{2,8}, Ekaterina A. Gubareva¹, Mikhail A. Maydin¹, Andrey I. Konstantinov², Konstantin A. Krasnov³, Ruben N. Karapetian⁴, Ekaterina I. Izotova⁵, Sergey E. Pigarev⁵, Andrey V. Panchenko¹, Margarita L. Tyndyk¹, Dmitry I. Osolodkin^{6,7}, Evgeny N. Nikolaev^{8,9,10}, Irina V. Perminova² and Vladimir N. Anisimov¹

¹N.N. Petrov National Medical Research Center of Oncology, Saint-Petersburg 197758, Russia

²Department of Chemistry, Lomonosov Moscow State University, Moscow 119991, Russia

³Institute of Toxicology, Federal Medical-Biological Agency, Saint-Petersburg 192019, Russia

⁴CHEMDIV LTD, Moscow District, Khimki 141400, Russia

⁵Nobel LTD, Saint-Petersburg 192012, Russia

⁶Institute of Poliomyelitis and Viral Encephalites, Chumakov FSC R&D IBP RAS, Moscow 108819, Russia

⁷Sechenov First Moscow State Medical University, Moscow 119991, Russia

⁸Skolkovo Institute of Science and Technology, Skolkovo 143025, Russia

⁹Institute for Energy Problems of Chemical Physics, Russian Academy of Sciences, Moscow 119334, Russia

¹⁰Orekhovich Institute of Biomedical Chemistry, Russian Academy of Medical Sciences, Moscow 119121, Russia

Correspondence to: Vladimir N. Anisimov, **email:** aging@mail.ru
Irina V. Perminova, **email:** iperm@org.chem.msu.ru

Keywords: cancer; polyphenols; molecular targets; BP-Cx-1; ChEMBL

Received: December 15, 2017

Accepted: December 16, 2017

Published: April 06, 2018

Copyright: Fedoros et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License 3.0 (CC BY 3.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

ABSTRACT

Identification of molecular targets and mechanism of action is always a challenge, in particular – for natural compounds due to inherent chemical complexity. BP-Cx-1 is a water-soluble modification of hydrolyzed lignin used as the platform for a portfolio of innovative pharmacological products aimed for therapy and supportive care of oncological patients. The present study describes a new approach, which combines *in vitro* screening of potential molecular targets for BP-Cx-1 using Diversity Profile - P9 panel by Eurofins Cerep (France) with a search of possible active components *in silico* in ChEMBL - manually curated chemical database of bioactive molecules with drug-like properties. The results of diversity assay demonstrate that BP-Cx-1 has multiple biological effects on neurotransmitters receptors, ligand-gated ion channels and transporters. Of particular importance is that the major part of identified molecular targets are involved in modulation of inflammation and immune response and might be related to tumorigenesis. Characterization of molecular composition of BP-Cx-1 with Fourier Transform Ion Cyclotron Resonance Mass Spectrometry and subsequent identification of possible active components by searching for molecular matches *in silico* in ChEMBL indicated polyphenolic components, nominally, flavonoids, saponin, phenanthrenes, as the major carriers of biological activity of BP-Cx-1. *In vitro* and *in silico* target screening yielded overlapping lists of proteins: adenosine receptors, dopamine receptor DRD4, glucocorticoid receptor, serotonin receptor

5-HT₁, prostaglandin receptors, muscarinic cholinergic receptor, GABA_A receptor. The pleiotropic molecular activities of polyphenolic components are beneficial in treatment of multifactorial disorders such as diseases associated with chronic inflammation and cancer.

INTRODUCTION

Lignin and its derivatives are considered as good candidates for development of novel medicinal products thanks to their low toxicity and broad range of biological activity. Lignin derivatives are known to exhibit antiviral [1, 2], anti-inflammatory [3], antidiabetic [4, 5] and antitumor [6, 7] effects. In addition, hydrolyzed lignin is used as enterosorbent [8], as well as for production of biodegradable and safe drug-delivery nanoparticles, e.g., antibacterial silver-compositions [9, 10]. Antioxidant activity, which is inherent to polyphenolic compounds, is believed to be the leading mode of action of lignin-like compounds [11, 12, 13].

Due to poor solubility in water, lignin and its derivatives demonstrate low bioavailability in animals. Hence, preparation of water-soluble modifications of lignin, which are most suitable for development of pharmacologically active substances, is of particular interest. Liquid-phase hydrolysis is the method of choice for preparing water-soluble lignin derivatives. In our previous studies, we used a lignin-based enterosorbent - Polyphepan® as the starting material and treated it in the alkaline medium at the elevated pressure and temperature with continuous supply of oxygen, to form the target water-soluble modification of lignin – BP-Cx-1 [14]. In line with the protocol of its production, BP-Cx-1 is a multicomponent composition of polyphenolic compounds. It was further used as a ligand for preparation of novel pharmaceutical products BP-C1, BP-C2, and BP-C3.

Platinum containing antineoplastic investigational medicinal product (IMP) BP-C1 is at the moment on the stage of clinical trials in patients with metastatic breast cancer [15]. Investigational medicinal product BP-C2 is a composition of BP-Cx-1 with ammonium molybdate. It exposes radioprotective effects through protection and reconstruction of hematopoiesis in irradiated rats [16], inhibits radiation-induced skin damage in mice [17], and improves quality of life of cancer patients undergoing chemotherapy [18]. BP-C3 is a composition of BP-Cx-1 with iron complex, selenium, ascorbic acid, and retinol. It exhibits geroprotective activity [19] and reduces 5-fluorouracil toxicity in mice by accelerating recovery of leukopoiesis and protecting small intestine epithelium [20].

Despite multiple studies published on biological effects of lignin derivatives, molecular mechanism of their action remains unknown. This is because application of existing characterization methods to multicomponent compositions of lignin and its derivatives is very challenging. Fourier transform ion cyclotron resonance mass-spectrometry (FTICR MS) with mild ionization

techniques is the only method, which is capable to resolve this issue. Due to its high resolution and mass accuracy, this method detects thousands of molecular peaks in complex mixtures without preliminary fractionation [21]. The power of this method enables exact assignment of thousands formulae, which could be used to evaluate the pathways of product formation as well as to suggest individual components responsible for bioactivity. Recently, FTICR MS was applied to explore homologues series in lignin derivatives [22]. Despite clear advantage of FTICR MS over other analytical techniques, the limitation of this method is a tolerance to isomers. The possible way to overcome this limitation is the application of selective chemical modification with enumeration of reactive centers [23]. Searching for well-known molecules with matching molecular masses and the described biological activity profiles is another feasible approach.

The identification of the molecular targets for bioactive compounds is an essential step in every drug discovery project [24]. In order to complement *in vitro* experiments numerous *in silico* approaches based on either information about the structures of biotargets or ligands were developed [25, 26]. These techniques were broadly applied in natural compound drug research [27, 28]. However, in case of compound mixtures, when structural information is limited and the individual compounds are hard to elucidate, a range of strategies, united under a term ‘dereplication’, were suggested [29]. Dereplication, in a broad sense, can be defined as a process of quickly identifying known chemotypes [29]. Different variants of dereplication strategies were successfully applied for chemical profiling of natural products by means of combining NMR, MS data and search in compound and bioactivity databases [29].

The present study describes a dereplication strategy, combining FTICR MS data, *in vitro* screening of potential molecular targets for water-soluble lignin-based ligand BP-Cx-1 using Diversity Profile - P9 panel by Eurofins Cerep (France), with identification of possible active components of BP-Cx-1 by means of searching for their molecular formulas *in silico* in ChEMBL, which is large scale bioactive compounds database.

RESULTS AND DISCUSSION

BP-Cx-1 characterization

The BP-Cx-1 was characterized using a number of methods. Elemental composition was determined using automatic elemental analysis. The content of carbon was 66.7%, hydrogen–4.82%, and oxygen–28.48%, which

corresponded to H/C ratio of 0.87 and O/C ratio of 0.32. The obtained data demonstrate low oxidation degree and high contribution of aromatic fragments into molecular structures of the water-soluble modification of lignin – BP-Cx-1.

The structural group composition was determined using the most powerful method of structural analysis - NMR spectroscopy. ^{13}C and ^1H NMR spectra are shown in Supplementary Figures 1 and 2. The peaks at 55–58 ppm in ^{13}C NMR spectrum and at 3.3–4.3 ppm in ^1H NMR spectrum are related to methoxy groups which are constitutive parts of lignin (Supplementary Figure 3). High values of spectral density in the range from 108 to 165 ppm of ^{13}C NMR and in the range from 6.2 to 7.9 ppm of ^1H NMR are consistent with predominantly aromatic nature of the product used in this study. Quantitative data of ^{13}C and ^1H NMR spectroscopy are given in Table 1. They indicate that aromatic compounds represent main part of BP-Cx-1. It also contains carboxyl and carbonyl groups, which are formed during oxidation of the starting lignin material.

In general, the structural group composition of the product under study agrees well with the data for the lignin starting material - a cross-linked polymer composed of phenylpropanoid units [30].

High-resolution mass-spectrometry

FTICR mass-spectrum is presented in Figure 1. The most intensive peaks in the mass-spectrum correspond to singly charged ions within 200-900 m/z window. In addition, low-intensity peaks of doubly charged ions could be detected (shown in the inset in Figure 1), which extend mass window up to 1600 Da. High mass accuracy of the FTICR MS enabled identification of exact CHO molecular formulas for 2146 singly- and 1602 doubly-charged ions.

The FTICR MS data were used to calculate number-averaged elemental composition of the sample used in this study, which accounted for 65.7% (C), 5.52 (H), and 28.77 (O). The obtained content of elements yielded H/C ratio of 1.01 and O/C ratio of 0.33. The MS-derived values of H content and the H/C ratio are somewhat higher as compared to those measured by the elemental analysis, which might be indicative of preferential ionization of hydrogen enriched aliphatic acids as compared to phenolic compartments of the BP-Cx-1 sample [31]. Still, the data on elemental composition of the BP-Cx1 sample obtained by the two methods – ESI FTICR MS and elemental analysis – are rather consistent and indicative of polyphenolic nature of the sample under study.

The molecular formulae assigned to FTICR MS data were plotted in Van Krevelen diagram, which is widely used for visualization of molecular composition of the complex matrices [32]. It represents a 2D plot of H/C versus O/C ratios constructed from all the determined molecular formulae (Figure 2A). The advantage of this diagram is that it demonstrates the full elemental

composition space occupied by the sample rather than the average bulk values. The obtained Van Krevelen diagram shows that the singly-charged ions (highlighted in blue color) contribute the most into molecular heterogeneity of the BP-Cx-1 sample—the corresponding dots are scattered all over the diagram field, whereas the molecular compositions of doubly-charged ions are much more concentrated (highlighted in red color) and located mostly within a narrow region of H/C values from 0.5 to 1 and O/C values from 0.3 to 0.5.

The convenient approach to describing MS-derived data in Van Krevelen diagram is an assignment of identified compounds to different molecular classes: lignin, tannins, lipids, etc, as it is reported in [33]. This enabled us for assigning the majority of doubly-charged ions to lignin-like components because they are located in the region, which is typical for phenylpropanoic structures (Figure 2A), whereas the molecular formulae assigned to singly-charged ions occupy the regions, which are typical for lipids, lignins and condensed aromatics. The significant differences between elemental compositions of singly and doubly-charged ions are also revealed by closer consideration of molecular mass distribution of the identified ions, which is shown in Figure 2B.

So, H/C versus molecular mass plot shows that all medium-molecular weight compounds, which exceed 600 Da, correspond to lignin-like molecules. At the same time, saturated (aliphatic) and condensed species are characterized by lower molecular weight. This might be indicative of depolymerization of the starting lignin material, which gives rise to formation of heterogenic products. Aromatic ring cleavage and condensation of phenols are among the most probable reactions: the former might result in formation of aliphatic molecules with $\text{H/C} > 1.2$, whereas the latter might lead to formation of condensed products with $\text{H/C} < 0.5$ (Figure 2B). In the case of BP-Cx-1, both reactions might be induced by an attack of hydroxyl radicals present in alkaline conditions onto the polyphenolic structures of the starting lignin material, as it was described, e.g., in [34]. We also cannot exclude the presence of extractives (terpenoids) in the starting lignin material, which could contribute to the saturated low molecular weight components.

The conducted studies allowed us to conclude that the BP-Cx-1, lignin derivative used in this study, is characterized by dominating contribution of polyphenolic compounds with molecular weights ranging from 300 to 1200 Da along with the presence of low molecular weight compartments of aliphatic character (e.g., terpenoids, fatty acids), and condensed aromatic compounds (e.g., flavonoids).

In vitro studies

Given the results of structural characterization of the lignin derivative under study, which revealed its molecular

Table 1: Structural group composition of BP-Cx-1 as measured by ¹³C and ¹H NMR spectroscopy

Carbon content in the structural fragments,%						
O = CRR'	COO	C _{Ar-O(H,R)}	C _{Ar}	C _{Alk-O}	OCH ₃	C _{Alk}
4	6	8	45	15	7	15

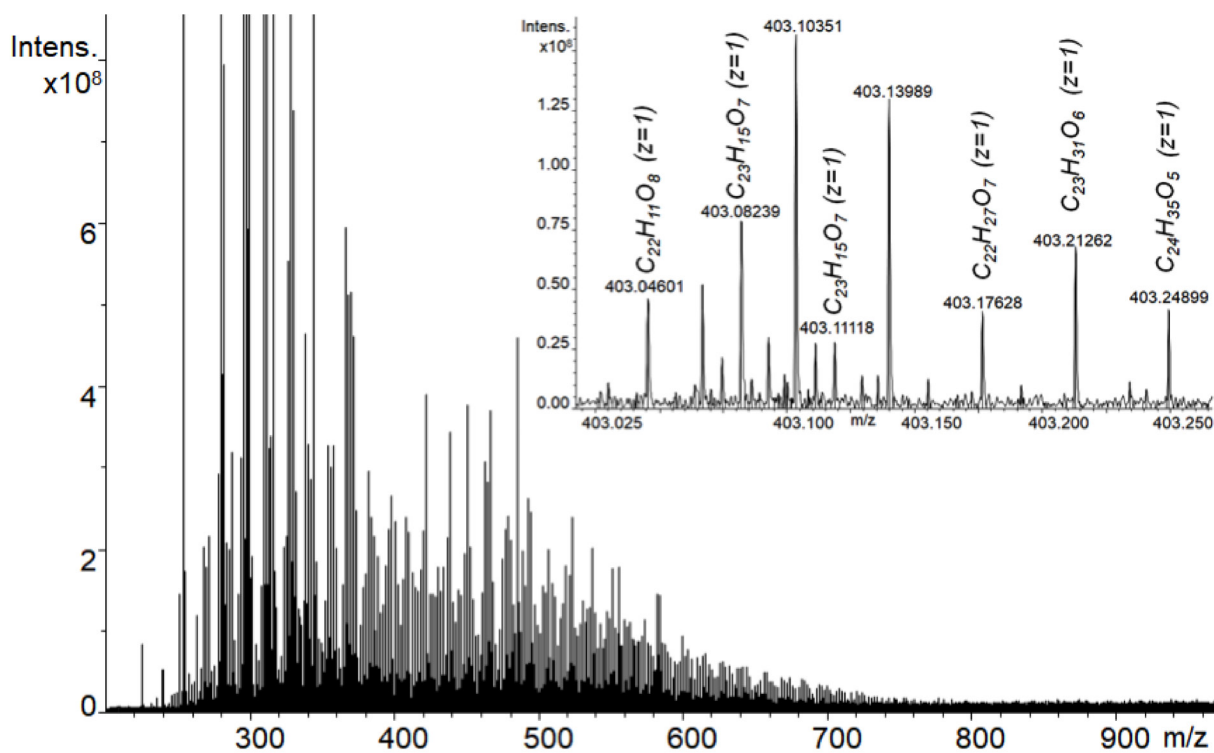
Hydrogen content in the structural fragments,%			
COOH+ArOH	ArH	CH ₃ O	AlkH
6.8	36	34.5	19.7

diversity, we have started exploration of its potential action targets, using the Diversity profile panel of Eurofins, which is composed of 71 binding and 27 enzyme assays (see Supplementary File 1). The binding assay panel has roughly an equal number of selective, central and peripheral therapeutically relevant targets. The enzyme assay panel is particularly focused on phosphatases and specific enzymes involved in cell cycle regulation. This profile is designed for the assessment of the biological diversity and potential adverse activity of Lead discovery candidates. Only the results of the binding assays were used in the present work.

The corresponding assay of BP-Cx-1 biological activity was conducted at the highest achievable concentration of 0.0042% (V/V). The thirteen cases of nonspecific interference of BP-Cx-1 with the assay and the pronounced BP-Cx-1 binding (over 50%) to 22 targets (Table 2) were identified.

The results show that interaction of BP-Cx-1 with multiple targets was detected. The majority of the identified targets belong to neurotransmitters receptors, ligand-gated ion channels and transporters, which are expressed in central and peripheral nervous systems. Given that drugs, which contain BP-Cx-1 ligand, are developed for antitumor therapy or as supportive care for cancer patients [35], of particular interest was to identify molecular targets, which are involved in modulation of inflammation and immune response.

Adenosine, adrenergic, bradykinin receptors, glucocorticoid receptor, histamine receptors, acetylcholine muscarinic receptors, purinergic receptors and prostanoid receptors are known to be involved in regulation of inflammation response [36]. Inhibition of bradykinin receptor B1 [37, 38], muscarinic receptors M1 and M2 [39], histamine receptor H1 [40] and prostanoid receptors EP2 [41], as well as activation

**Figure 1: ESI FTICR mass-spectrum of the lignin derivative used in this study - BP-Cx-1. Insert shows the presence of singly and doubly charged ions within 0.25 Da of mass-spectrum fragment.**

of glucocorticoid receptor GR [42] and adenosine receptors A(1), A(2A) [43, 44], is beneficial for reducing inflammation and pain.

There is reported evidence suggesting that chronic inflammations are tightly associated with cancer [45, 36]. This might indicate specific role of inflammation promoting receptors in tumorigenesis. The tumors recruit multiple paracrine, autocrine and systemic signaling factors for maintenance of high proliferative and metabolic level and for avoiding host immune response. As a result, simultaneous targeting of several proteins may be beneficial for treatment of the cancer patients.

Inhibitors and agonists of many receptors are currently considered as antitumor drug candidates. Thus, glutamate bradykinin and AMPA receptors have been implicated in cancer migration, invasion and metastasis [46, 47, 48]. Some data on prostaglandin receptor EP2 indicate that it is highly expressed in a variety of cancer cells and exacerbates disease pathology [49, 50, 41, 51]. Adenosine A(1) and A(2A) receptors agonists exhibit immunosuppressive effect in tumor stroma [52, 53]; this effect can be amplified by co-administration of selective inhibitors of purinergic receptors [54]. Dopamine D4 receptor inhibitors, which affect autophagy in glioblastoma stem cells and induce apoptosis were also demonstrated to have antitumor activity [55]. This leads to suggestions that β -adrenoceptors might be valuable drug targets to control the deleterious effects of β -adrenergic system in tumor development along with psychological stress in cancer patients [56, 57].

The direct experimental evidence with regard to possible effect of BP-Cx-1 onto various receptors, taking part in inflammatory reactions and tumor development, as well as with regard to its pharmacologically active concentrations is missing. However, the data from clinical trials of BP-Cx-1 containing products, as well as laboratory animal testing allows us a suggestion that its numerous documented cumulative anti-inflammatory

and antitumor effect in cancer patients may in part be attributed to the direct ligand-receptor interaction.

ChEMBL data mining

For suggesting structures, which might be related to the BP-Cx-1 bioactivity profile, we used the FTICRMS data on molecular formulae, which were identified in this sample (3748 formulas). We used these formulas for manual search in ChEMBL database. The related structures were extracted along with their bioactivity data. The text search was conducted in the target_dictionary.pref_name and assays.description fields, which contain information about protein target and biological assay conditions. This enabled extraction of 1474 relevant assays corresponding to 1316 ChEMBL compounds with interpretable activity values. Due to diversity of ChEMBL assays and activity types, the activity values were hashed to boolean flags, where '1' and '0' denoted the presence or absence of the compound activity against the target. Individual flags were summarized in boolean fingerprints (Supplementary File 2, Figure 3), which represent compound bioactivity profiles against all the targets screened. The obtained bioactivity profiles were compared with the experimental data by constructing the so-called "fingerprints". We proposed a parameter of the fitness of a ChEMBL compound to the BP-Cx-1 bioactivity profile: values of '1' and '-1' corresponded to the match or mismatch of activity of the ChEMBL compound and BP-Cx-1 against the target, respectively, and '0' was assigned in the absence of data (Supplementary File 3, Figure 3). All values were summed up giving fit_score - the final value, which characterized the fitness of a compound to the BP-Cx-1 bioactivity profile. Thus, 56 would represent the fit_score for an ideally fitted compound, that was tested and demonstrated matching activity in all 56 receptors. The maximum fit_score found for the

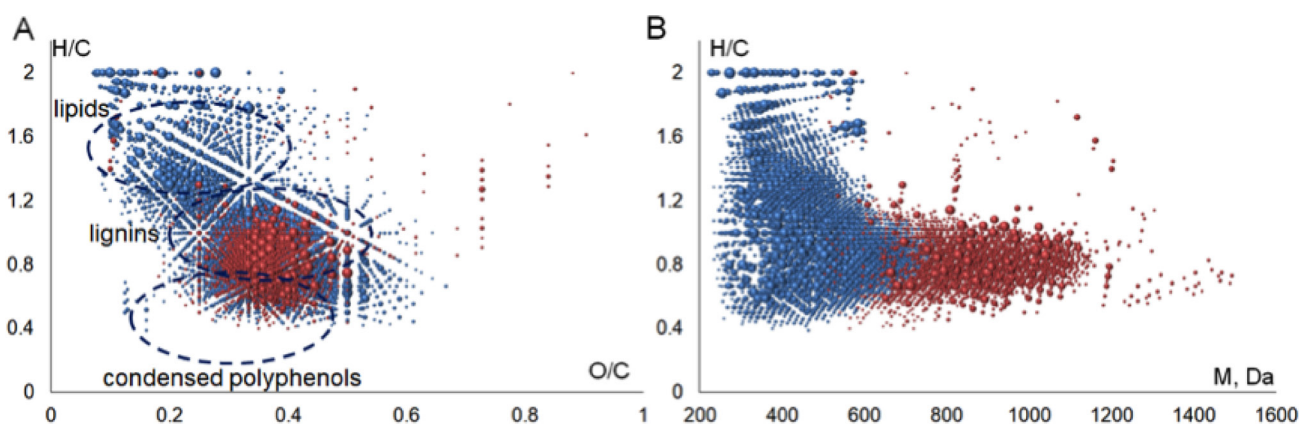


Figure 2: Visualization of FTICR MS data obtained for the BP-Cx-1 lignin derivative: (A) Van Krevelen diagram, (B) H/C versus molecular weight diagram. The molecular compositions of singly and doubly charged ions are highlighted in blue and red colors, respectively.

Table 2: Binding targets of BP-Cx-1 identified by Diversity Profile–P9 panel

Receptor family	Target	Protein/protein family symbol	Inhibition%
Adenosine receptors	A1 (h)	ADORA1	62.1%
	A2A (h)	ADORA2A	86.6%
	$\alpha 2$	ADORA2A-C	52.2%
Adrenergic receptors	$\beta 1$ (h)	ADRB1	51.5%
	$\beta 2$ (h)	ADRB2	71.4%
GABAA (γ -Aminobutyric acid)	BZD	GABRA1	53.8%
Glutamate	AMPA	GRIA 1-4	90.9%
	Kainate	GRIK 1-5	84.3%
Bradykinin receptors	B1 (h)	BDKRB1	76.7%
Dopamine receptors	D4.4 (h)	DRD4	62.7%
Histamine receptors	H1 (h)	HRH1	57.6%
Cholinergic receptors	$\alpha 4\beta 2$ (h)	CHRNA4, CHRNB2	73.5%
	M	CHRM1-5	96.3%
Vasopressin receptor	V2 (h)	AVPR2	69.8%
	V1a (h)	AVPR1A	56.4%
Ion channels	KATP	KCNJ	50.3%
Transporters	GABA	SLC6A1	71.5%
Thyroid Stimulating Hormone receptor	TRH1 (h)	TSHR	68.8%
Glucocorticoid receptor	GR (h)	NR3C1	63.2%
Serotonin receptor	5-HT1	HTR1 A,B,D,E,F	81%
	P2X	P2RX1-7	64%
Purinergic receptors	P2Y	P2RY1,2,4,6,11–14	52.6%
	EP2 (h)	PTGER2	52.4%
Prostanoid receptors	IP (PGI2)	PTGIR	68.7%

ChEMBL compounds was 4. The fit_score ≥ 1 (well-fitted compounds) was obtained for 354 ChEMBL compounds.

Of the 1316 ChEMBL compounds, 1005 had the activity against various targets matching that of BP-Cx-1, and 1968 did not match activity of BP-Cx-1. For the 354 well-fitted ChEMBL compounds these values were 396 and 6, correspondingly (Supplementary File 3). Data on activity of the ChEMBL compounds and BP-Cx-1 against the Diversity Profile targets (receptors and transporters) are graphically presented as a STRING protein association network in Figure 4. Network approach was used to analyze functional clusters of BP-Cx1 targets established *in vitro* and additional probable targets identified *in silico*.

As intracellular signaling resembles network rather than straightforward pathways, targeting several connected proteins would influence the pathway more efficiently than interaction with a single target. Network analysis in STRING revealed that the targets form a network which is enriched in interactions, i.e. it has more interactions among proteins than a random set of proteins (enrichment

p -value: $< 1.0e-16$). Interaction enrichment implies that proteins are at least partly connected and regulate the same function. In our network (Figure 4), some biological processes according to GO (gene ontology database) were enriched, including G-protein coupled receptor signaling, blood circulation, cellular calcium ion homeostasis, regulation of cell proliferation, inflammatory response.

As can be seen from Figure 4, established activity of ChEMBL compounds was matching that of BP-Cx-1 for 14 out of the 22 targets listed in Table 2, at that, well-fitted ChEMBL compounds had the activity against 9 targets (Supplementary File 3). The highest number of ChEMBL compounds with activities matching those of BP-Cx-1 were found for the following BP-Cx-1 targets: adenosine receptors A1 (ADORA1) – 71 records of matching activity among 91 tested compounds and 20 out of 20 records for the well-fitted ChEMBL compounds, A2 (ADORA2A) – 52 active out of the 63 tested (7 out of 7 for the well-fitted ChEMBL compounds), dopamine receptor D4.4 (DRD4) – 65 out of 73 for all ChEMBL compounds and 15 out of 15 for the well-

fitted compounds. For GR (NR3C1) glucocorticoid receptor matching activities were identified for 81 out of 103 tested compounds (29 out of 32 for the well-fitted), for 5-HT1 (HTR1 A,B,D,E,F in Table 2, HTR1A in Figure 4) serotonin receptor matching activities were identified for 59 out of 64 tested compounds (6 out of 6 for the well-fitted compounds) and for EP2 (PTGER2) receptor, matching activities were identified for 24 out of 27 compounds and for 10 out of 10 of the well-fitted compounds. A smaller number of matches were identified for 12 (PTRIG) ligands - 8 out of 25 tested compounds with $fit_score \geq 1$ were active (8 out of 8 for the well-fitted compounds) and for the muscarinic (CHRM1-5 in Table 2, CHRM1 in Figure 4) cholinergic receptor – 50 active compounds were identified in ChEMBL (52 records were tested), however, only 1 compound had the $fit_score \geq 1$. Compounds active against BZD site of GABAA receptor (GABRA1) were also identified: 8 compounds were identified, of which 5 were active (2 out of 2 for the well-fitted compounds).

Around 50 activity records against each of adrenergic ($\alpha 2$, ADRA2A; $\beta 1$, ADRB1; $\beta 1$, ADRB1) as well as , histamine (H1, HRH1) and vasopressin (V1a, AVPR1A) receptors were identified in ChEMBL database; activity profiles of all corresponding ChEMBL compounds were characterized by the $fit_score < 1$ (Supplementary File 3, Figure 4).

There was a number of receptors in the binding assays of the Diversity Profile–P9 panel, against which activity of BP-Cx-1 was smaller than 50% (Supplementary File 1). ChEMBL compounds were found active against 16 of these receptors (Figure 4). Thus, while 212 out of 283 tested ChEMBL compounds were active against cannabiod (CCKAR) and 145 out of 236 against opioid (OPRK) receptors, BP-Cx-1 had a low affinity to these (9% and 23%, respectively). Likewise, 191 out of 208 and 106 out of 145 ChEMBL compounds, were active against endothelial ETA (ENDRA) and ETB (ENDRB) receptors, respectively.

Structure	Activity Fingerprints					Fitness fingerprints					
	A1	A2A	GR	...	ETA	A1	A2A	GR	...	ETA	<u>fit_score</u>
	1	1	ND	...	ND	1	1	0	...	0	4
	ND	ND	1	...	ND	0	0	1	...	0	1
	ND	ND	1	...	ND	0	0	1	...	0	1
	ND	ND	ND	...	1	0	0	0	...	1	2

Figure 3: Examples of Boolean bioactivity fingerprints.

BP-Cx-1 also demonstrated low affinity to nuclear receptors: 30% for ER (ESR), 13% for PR (PGR) and 36% for AR. 275 out of 353 tested ChEMBL compounds were active against estrogen receptors; 98 out of 108 and 109 out of 136 were active against progesterone and androgen receptors, respectively. Interestingly, single instances of activity against these receptors were identified for the well-fitted ChEMBL compounds. Thus, activity records exist in ChEMBL database for: 4',5,7-Trihydroxyflavone against estrogen receptor, resorcylic acid lactone with molecular formula $C_{21}H_{28}O_6$ against progesterone receptor and tetracyclic triterpene ($C_{24}H_{30}O_6$) against androgen receptor.

To compare chemical composition of ChEMBL compounds versus BP-Cx-1 compounds, the molecules found in ChEMBL database were plotted in Van Krevelen diagram and in the H/C vs MW diagram (Figure 5). It can be deduced that the major portion out of 1316 identified formulas corresponded to saturated compounds (Figure 5A) with molecular masses below 500 Da (Figure 5B). By the structural type we assume them to be terpenoids and lignans. For example, the role of lignans in biological activity was reported in our previous work [58]. It should

be noted that the database query returned BP-Cx-1 components with the highest relative intensity in mass-spectrum. Given that 354 ChEMBL well-fitted compounds with fit-score ≥ 1 are low-molecular weight entities with high abundance (Figure 5), we hypothesize that saturated compounds play a significant role in interaction with the identified targets.

Additionally, we calculated Murcko scaffolds for all ChEMBL compounds (Table 3). It can be deduced from Table 3 that the scaffolds of well-fitted compounds match those of furanditerpenes derivatives (Table 3, 1A, 2B), tetracyclic triterpenes (Table 3, sapogenins, 2A, 1B) and their precursors—phenanthrene structures (Table 3, 4A), as well as flavonoids (Table 3, flavanes, 3A) and simple aromatic structures (Table 3, fragments of noncondensed lignin, 5A, 3B). Compounds of these types are part of various natural products [59, 60, 61], and they may also form in the course of successive acidic and alkaline hydrolysis of wood [62].

Polyphenolic compounds of flavonoid nature are the most investigated in the medical field (Table 3, 3A). Antiinflammatory [63], antitumor [64, 65], antiarthritic [66] activities were demonstrated for quercetin. The spectrum of biological activity of genistein isoflavone resembles that

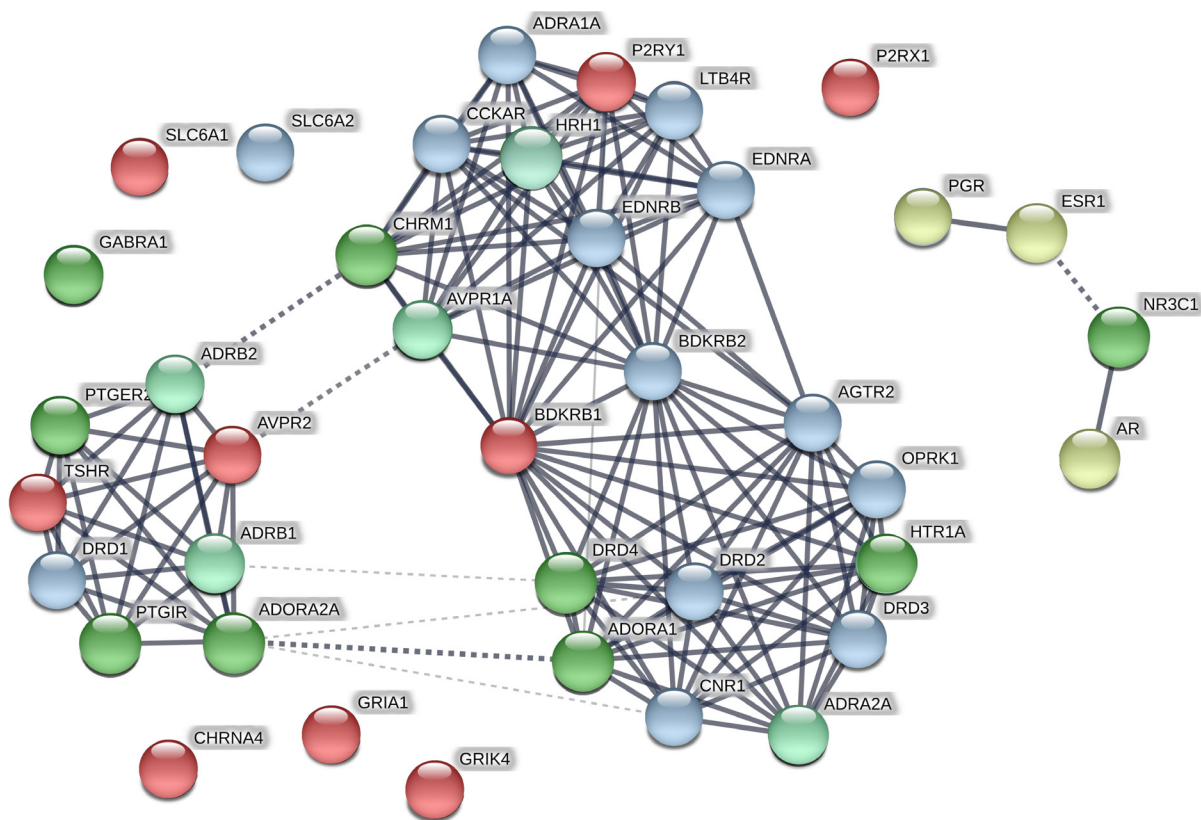


Figure 4: STRING protein association network for BP-Cx-1 and ChEMBL compounds targets. Bright-green—targets: activity of the well-fitted ChEMBL compound matched activity of BP-Cx-1; Light-green—targets: activity of other ChEMBL compounds matched activity of BP-Cx-1; Yellow—targets: activity of the well-fitted ChEMBL compound did not match activity of BP-Cx-1 (binding by BP-Cx-1 below 50%); Blue—targets: activity of other ChEMBL compounds did not match activity of BP-Cx-1 (binding by BP-Cx-1 below 50%); Red—BP-Cx-1 targets: no matching activity records were identified in the ChEMBL database.

of other flavonoids [67, 68]; besides, it is conformationally similar to estradiol and it might act as the latter [69]. It is the estrogen-like activity that is referred to as being responsible for the radioprotective effectiveness of this compound in ARS CD2F1 murine model (DRF 1.16 in case of s.c. prophylactic administration) where it acts via protection of bone marrow progenitor cell populations [70].

It is worthwhile to note that the present study demonstrated moderate activity of the BP-Cx-1 sample towards ER receptors (Supplementary File 1), while in our previous ARS study, efficacy of BP-C2, composition composed of BP-Cx-1 and molybdenum, was similar to that of genistein (DRF = 1.1). The results of that study suggested that the mechanism of action is associated with protection and acceleration of recovery of hematopoietic system of irradiated animals [16]. Apparently, induction of antiinflammatory signaling pathways, though not ER-induced, results in a similar *in vivo* activity of BP-C2.

Given that some flavonoids are reported to traverse the blood-brain barrier, when administered orally and/or parenterally [71], it might be expected that other compounds of this group could act on CNS receptors. The studies dedicated to these effects are very scarce, but in a study of hispidulin, for instance, ability of this compound to act as a ligand in recombinant GABAA/BZD receptors expressed by *Xenopus laevis* oocytes was demonstrated as well as a pronounced anticonvulsive effect was observed in seizure-prone Mongolian gerbils (*Meriones unguiculatus*) in a model of epilepsy of oral formulation of this compound at 10 mg/kg body weight [72].

Activity of furane diterpene compounds (Table 3, 1A, 2B) towards CNS receptors such as salvinorin A, a selective high efficacy kappa-opioid receptor (KOPr) agonist, is reported; some intensive studies are performed to evaluate furane diterpene products with a reduced burden of undesirable effects associated with salvinorin A [73].

Earlier reports demonstrated that BP-Cx-1, when administered intravenously, was able to traverse the blood

brain barrier [35]. Activity of BP-Cx-1 towards GABAA/BZD receptors and a lack of effect towards opioid receptors (Supplementary File 1) are demonstrated by the results of the present study.

Ginseng, *Pulsatilla* sp., *Aesculus pavia* and other plants are the natural sources of triterpene compounds (Table 3, 2A, 1B). Like flavonoids, natural saponines (glycosides) and sapogenins (aglycosides) are characterized by multiple biological activities: antitumor [74–77], antiinflammatory [78], pain reducing [79].

Phenanthrene structures (Table 3, 4A) are mainly obtained from Orchidaceae family plants. Phenanthrenes exhibit various biological activities, among which are anti-inflammatory, antiallergic [80] and cytotoxic [81] effects.

Thus, polyphenols' "molecular promiscuity", in a sense that a single molecule has multiple targets, is a characteristic feature for this class of compounds. It is important to note that the targets include diverse protein families such as GPCRs, enzymes, ion channels, transcription factors. Different polyphenolic molecules have been shown to influence signaling pathways associated with inflammation, oxidative stress, cancer, metabolic regulation, etc. [82–84].

CONCLUSIONS

The reported study, for the first time, employed interdisciplinary approach towards identification of molecular targets and biologically active structures of the novel water-soluble lignin derivative – BP-Cx-1, used as a platform for a portfolio of medicinal products for the main therapy and supportive care of oncological patients.

It was demonstrated that BP-Cx-1 exerts multiple pharmacological effects towards neurotransmitters receptors, ligand-gated ion channels and transporters; most of molecular targets are involved in modulation

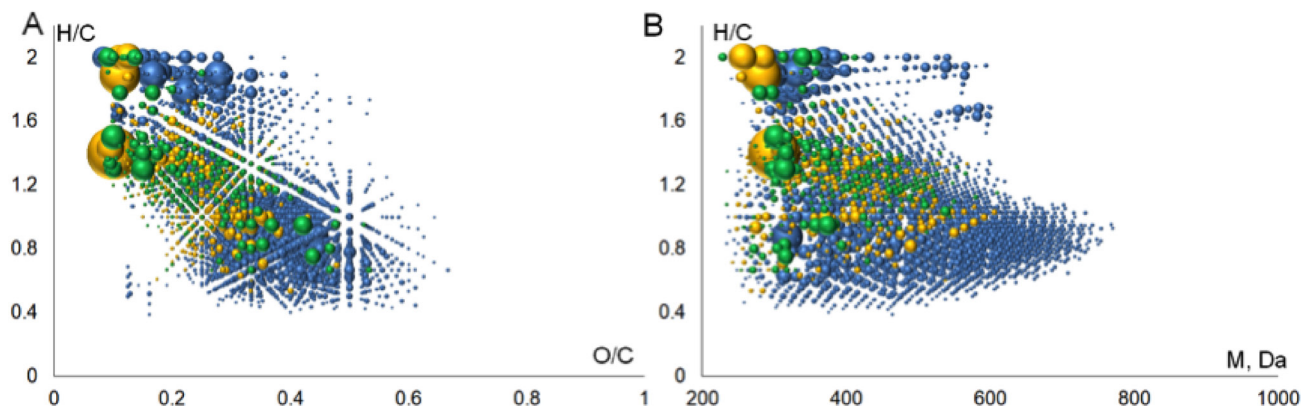
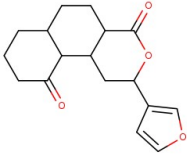
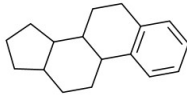
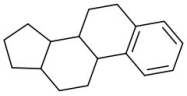
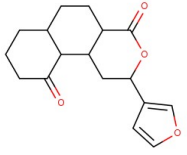
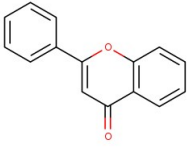


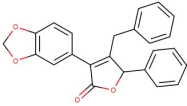

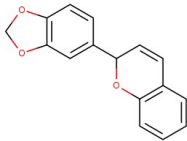


Figure 5: Visualization of derived molecular space (A) Van Krevelen and (B) H/C versus molecular weight diagram. Blue, orange and green spheres represent formulas for FTICR MS compounds, ChEMBL compounds and ChEMBL well-fitted compounds, respectively.

Table 3: Murcko scaffolds distribution for all ChEMBL compounds and ChEMBL compounds with fit_score ≥ 1 (well fitted to BP-Cx-1 bioactivity)

Well fitted ChEMBL compounds			All ChEMBL compounds		
#	Scaffold	%	#	Scaffold	%
1A		11.1	1B		6.3
2A		4.5	2B		5.6
3A		3.7	3B		3.3
4A		3.1	4B		3.0
5A		2.8	5B		2.2

of inflammation and immune response. Examination of molecular composition of BP-Cx-1 with FTICR MS and subsequent identification of possible active components of BP-Cx-1 by searching for their molecular formulas *in silico* in ChEMBL indicate that the said biological activity may stem from multiple polyphenolic components likely contained in BP-Cx-1: flavonoids, saponinins, phenanthrenes etc. *In vitro* and *in silico* target screening yielded not identical, but overlapping lists of proteins. Both methods identified adenosine receptors, dopamine receptor DRD4, glucocorticoid receptor, serotonin receptor 5-HT1, prostaglandin receptors, muscarinic cholinergic receptor, GABAA receptor.

Despite great molecular diversity of polyphenolic compounds and of their targets, biological activities of the known polyphenolic compounds are similar; the most of those activities are the anti-inflammatory and antiproliferative. Even though this class of compounds lacks selectivity towards particular targets, BP-Cx-1 pleiotropic molecular activities may be beneficial in the

treatment of multifactorial disorders, such as diseases associated with chronic inflammation (neurodegenerative diseases, rheumatoid arthritis, autoimmune disorders), cancer, type 2 diabetes.

MATERIALS AND METHODS

BP-Cx-1 composition

BP-Cx-1 was provided by Nobel Ltd as a powder with 95.9% active substance (batch A1) for the characterization analyses and as a sterile 0.42% ammonia solution (batch X112K14A1) for the *in vitro* studies. In the *in vitro* studies BP-Cx-1 was tested at a concentration of 0.0042% (V/V).

BP-Cx-1 characterization

Elemental analysis (C, H, and N) was conducted using Vario El Microcube analyzer. Ash content was measured by manual combustion of the test-agent at 850°C. The content of all elements was calculated on

ash-free basis. Oxygen concentration was calculated as the difference between a mass of the sample and the ash-corrected amount of CHN.

Quantitative ^{13}C solution-state NMR spectra were recorded with Avance NMR spectrometer (Bruker, Germany), operating at 100 MHz carbon-13 frequency. A 50 mg BP-Cx-1 sample was dissolved in 0.6 mL 0.3 N NaOD and transferred into 5 mm NMR tube. ^{13}C NMR spectra were acquired with a 5 mm broadband probe, using CPMG pulse sequence with nuclear Overhauser effect suppression by the INVGATE procedure; the acquisition time and relaxation delay were 0.2 s and 7.8 s, respectively. Quantitative distribution of carbon among the main structural fragments of the test-agent was established under these analytical conditions [85]. The assignments were as follows (in ppm): 5–108, aliphatic C atoms ($\sum C_{\text{Alk}}$) including 50–58, methoxy group ($C_{\text{CH}_3\text{O}}$); 108–165, aromatic C-atoms ($\sum C_{\text{Ar}}$); 165–187, C atoms of carboxylic and ester groups (COO); 187–220, C atoms of quinonic and keto- groups (C=O).

^1H NMR spectra were recorded with Bruker AM-400 at 400 MHz, in standard 5.0 mm NMR tubes. Each sample was analyzed in DMSO- d_6 solution, and then in DMSO- d_6 with addition of deuterated trifluoroacetic acid 20% (wt). Samples were dissolved at 40 mg/mL. Spectra were recorded at 20°C, pulse delay of 3 s, AT = 1 s, NA = 32. Spectra were analyzed in NUTS (Acorn NMR) software. Five integration regions are considered for this work (in ppm): 1.6–3.3, aliphatic protons ($\sum H_{\text{Alk}}$); 3.3–4.3, methoxy group protons ($\sum H_{\text{CH}_3\text{O}}$), 6.2–7.9 ppm, aromatics and vinylics ($\sum H_{\text{Ar}}$); 9.4–11.0, protons in aldehyde groups ($H_{\text{C=O}}$); 12.6–13.5, protons in carboxylic acid groups.

FTICR mass spectra were acquired with FT MS Bruker Apex Ultra with harmonized cell (Bruker Daltonics) equipped with a 7 T superconducting magnet and electrospray ion source (ESI) in negative ionization mode, at the facilities of the Institute of Biomedical Chemistry of RAMS. Sample solution (200 mg/L) was injected into the ESI source using a microliter pump at 90 $\mu\text{L}/\text{h}$ flow-rate, pressure of nebulizer gas of 138 kPa and pressure of the drying gas of 103 kPa. A source heater temperature of 200°C was maintained to ensure rapid dissolution in ionized droplets. The spectra were acquired within a time domain of 4 megawords and 300 scans were accumulated for each spectrum. Resolving power was 530 000 at $m/z = 400$. First, mass-spectrum was externally calibrated using synthetic carboxyl-containing polystyrene standard, as described in [86], and internal calibration was systematically performed against residual peaks of fatty acids [87]. Molecular identifications were performed with the proprietary “Transhumus” software, based on the total mass difference statistics algorithm, designed by A. Grigoriev [88]. The generated CHO formulas were validated by the following constraints (O/C ratio ≤ 1 , H/C ratio ≤ 2 , element counts ($C \leq 120$, $H \leq 200$, $O \leq 60$) and mass accuracy window < 1 ppm).

In vitro study

The study was performed by Eurofins Cerep (France) using standardized Diversity Profile - P9 panel, comprising 53 receptors, 14 ion channels, 5 transporters and 27 enzymes (results from enzyme assays are not disclosed in this article). Binding of BP-Cx-1 to receptors, ion channels and transporters was examined with radioligand binding assays. Assay parameters are presented in the Supplementary File 4.

Binding results are expressed as a percentage inhibition of control specific binding obtained in the presence of BP-Cx-1:

$$100 - \left(\frac{\text{measured specific binding}}{\text{control specific binding}} \times 100\% \right)$$

Specific binding values correspond to values obtained in the presence of BP-Cx-1, less than nonspecific (background) signal. Inhibition effects greater than 50% were considered significant.

ChEMBL data mining

Data processing was carried out using Python 2.7 scripts. Database management was carried out either in InstantJChem 17.2.6.0 [89] or in DataWarrior 4.6.1. [90].

ChEMBL version 23 was used. 3748 formulae identified in FTICR mass spectrum (Supplementary File 5) were extracted as described previously [91]. In brief, all CHO formulae for molecular weight range from 200 to 800 were generated and extracted from MySQL ChEMBL, using *molecule_dictionary.full_molformula* field (Supplementary File 6). The formulae from MS were merged with all formulae extracted from ChEMBL, giving 41834 *compound_properties.molregno* primary keys (Supplementary Files 7). Fields *assays.assay_id* and *assays.description* and *target_dictionary.pref_name* were extracted using *compound_properties.molregno* as a primary key. Fields *target_dictionary.pref_name* and *assays.description* were searched with key substring list (Supplementary File 8), using a python script (Supplementary File 9), and found assays were extracted and saved to a CSV file (Supplementary File 10). For the extracted *assay_id*'s activity data and compounds were retrieved from ChEMBL, using *activities.assay_id* and *compounds_properties.molregno* as primary keys, respectively. Compounds were standardized using ChemAxon Standardizer v.16.8.29.0 [92] (Supplementary File 11). ChEMBL activities were standardized using a python script (Supplementary File 9) and ‘active/inactive’ (1/0) flag for all the compounds was assigned. Two types of activity fingerprints were built using a python script (Supplementary File 9) based on the flag values (Supplementary Files 2, 3). Scaffold distributions were analyzed in DataWarrior 4.6.1.

Network analysis was performed through mapping a list of targets to STRING database (<http://string-db.org/>) which integrates known protein-protein interactions (evidence from curated databases and experiments,

minimum interaction score 0.400). GO molecular functions with highest confidence scores ($p < 1e-06$) were selected from the pathway enrichment results.

Author contributions

Irina V. Perminova and Vladimir N. Anisimov developed the original idea interpreted the data and revised the manuscript. Elena Fedoros, coordinated the research and revised the manuscript. Dmitry I. Osolodkin developed of the data mining protocol and Alexey A. Orlov performed the data mining. Alexander Ya. Zhrebker and Evgeny N. Nikolaev performed FTICR MS measurements, abstracted and analyzed data. Andrey I. Konstantinov, Konstantin A. Krasnov and Ekaterina I. Izotova performed ^{13}C , ^1H NMR and elemental composition study. Ruben N. Karapetian conceived and organized *in vitro* study and Ekaterina A. Gubareva, Mikhail A. Maydin, Andrey V. Panchenko and Margaruta L. Tyndyk analyzed the biological activity data retrieved *in vitro* and *in silico*. Sergey E. Pigarev wrote the paper. All authors contributed to the preparation and reviewed the manuscript.

CONFLICTS OF INTEREST

The authors declare no conflicts of interest.

FUNDING

This work was supported by a grant from the Russian Science Foundation, Grant number 16-15-00142.

REFERENCES

1. Andrei G, Lisco A, Vanpouille C, Introini A, Balestra E, van den Oord J, Cihlar T, Perno CF, Snoeck R, Margolis L, Balzarini J. Topical tenofovir, a microbicide effective against HIV, inhibits herpes simplex virus-2 replication. *Cell Host Microbe*. 2011; 10:379–389. <https://doi.org/10.1016/j.chom.2011.08.015>.
2. Raghuraman A, Tiwari V, Zhao Q, Shukla D, Debnath AK, Desai UR. Viral inhibition studies on sulfated lignin, a chemically modified biopolymer and a potential mimic of heparan sulfate. *Biomacromolecules*. 2007; 8:1759–63. <https://doi.org/10.1021/bm0701651>.
3. Saluja B, Thakkar JN, Li H, Desai UR, Sakagami M. Novel low molecular weight lignins as potential anti-emphysema agents: *In vitro* triple inhibitory activity against elastase, oxidation and inflammation. *Pulm Pharmacol Ther*. 2013; 26:296–304. <https://doi.org/10.1016/j.pupt.2012.12.009>.
4. Barapatre A, Aadil KR, Tiwary BN, Jha H. *In vitro* antioxidant and antidiabetic activities of biomodified lignin from *Acacia nilotica* wood. *Int J Biol Macromol*. 2015; 75:81–89. <https://doi.org/10.1016/j.ijbiomac.2015.01.012>.
5. Hasegawa Y, Kadota Y, Hasegawa C, Kawaminami S. Lignosulfonic acid-induced inhibition of intestinal glucose absorption. *J Nutr Sci Vitaminol (Tokyo)*. 2015; 61:449–454. <https://doi.org/10.3177/jnsv.61.449>.
6. Barapatre A, Meena AS, Mekala S, Das A, Jha H. *In vitro* evaluation of antioxidant and cytotoxic activities of lignin fractions extracted from *Acacia nilotica*. *Int J Biol Macromol*. 2016; 86:443–53. <https://doi.org/10.1016/j.ijbiomac.2016.01.109>.
7. Wang Q, Mu H, Zhang L, Dong D, Zhang W, Duan J. Characterization of two water-soluble lignin metabolites with antiproliferative activities from *Inonotus obliquus*. *Int J Biol Macromol*. 2015; 74:507–14. <https://doi.org/10.1016/j.ijbiomac.2014.12.044>.
8. Perlamutrov YN, Olhovskaya KB, Zakirova SA. Double-blind controlled randomised study of lactulose and lignin hydrolysed combination in complex therapy of atopic dermatitis. *Microb Ecol Health Dis*. 2016; 27:30418. <https://doi.org/10.3402/mehd.v27.30418>.
9. Frangville C, Rutkevicius M, Richter AP, Velev OD, Stoyanov SD, Paunov VN. Fabrication of environmentally biodegradable lignin nanoparticles. *Chemphyschem*. 2012; 13:4235–4243. <https://doi.org/10.1002/cphc.201200537>.
10. Richter A, Brown JS, Bharti B, Wang A, Gangwal S, Houck K, Cohen Hubal EA, Paunov VN, Stoyanov SD, Velev OD. An environmentally benign antimicrobial nanoparticle based on a silver-infused lignin core. *Nat Nanotechnol*. 2015; 10:817–23. <https://doi.org/10.1038/nnano.2015.141>.
11. Pan X, Kadla JF, Ehara K, Gilkes N, Saddler JN. Organosolv ethanol lignin from hybrid poplar as a radical scavenger: relationship between lignin structure, extraction conditions, and antioxidant activity. *J Agric Food Chem*. 2006; 54:5806–5813. <https://doi.org/10.1021/jf0605392>.
12. Qazi SS, Li D, Briens C, Berruti F, Abou-Zaid MM. Antioxidant Activity of the Lignins Derived from Fluidized-Bed Fast Pyrolysis. *Molecules*. 2017; 22:E372. <https://doi.org/10.3390/molecules22030372>.
13. Sun SN, Cao XF, Xu F, Sun RC, Jones GL. Structural features and antioxidant activities of lignins from steam-exploded bamboo (*Phyllostachys pubescens*). *J Agric Food Chem*. 2014; 62:5939–47. <https://doi.org/10.1021/jf5023093>.
14. Shipov V, Pigarev E, Fedoros E, inventors. Benzene Polycarboxylic Acid Compounds and Their Use as Drug. EP patent EP2831151 (B1). 2016.
15. Larsen S, Butthongkomvong K, Manikhas A, Trishkina E, Poddubskaya E, Matrosova M, Srimuninnimit V, Lindkær-Jensen S. BP-C1 in the treatment of patients with stage IV breast cancer: a randomized, double-blind, placebo-controlled multicenter study and an additional open-label treatment phase. *Breast Cancer (Dove Med Press)*. 2014; 6:179–189. <https://doi.org/10.2147/BCTT.S71781>.

16. Panchenko AV, Fedoros EI, Pigarev SE, Bykov VN, Drachev IS, Krayev SY. [Experimental Study of Radioprotective Efficacy of the Polyphenolic Lignin-derived Polymer Composition with Ammonium Molybdate–BP-C2]. [Article in Russian]. *Radiats Biol Radioecol*. 2017; 57:505–511.
17. Fares F, Fares B, Azzam N, Nashashibi M, Nevelsky A, Larsen S, Lindkær-Jensen S. An Innovative Complex of Benzene-Poly-Carboxylic Acid and Molybdenum, for Prevention and Treatment of Radiation Dermatitis. *Med Chem*. 2015; 5:447–451. <https://doi.org/10.4172/2161-0444.1000299>.
18. Ibrahim T, Larsen S, Jensen NHL, Lindkær-Jensen S. BP-C2 improves Functional status, Quality of Life and Corrects Biochemical Imbalances as Adjuvant Therapy to FOLFIRINOX Treatment: A Case of Advanced Inoperable Pancreatic Cancer. *J Clin Case Rep*. 2015; 5:514. <https://doi.org/10.4172/2165-7920.1000514>.
19. Anisimov VN, Popovich IG, Zabezhinski MA, Yurova MN, Tyndyk ML, Anikin IV, Egormin PA, Baldueva IA, Fedoros EI, Pigarev SE, Panchenko AV. Polyphenolic drug composition based on benzenepolycarboxylic acids (BP-C3) increases life span and inhibits spontaneous tumorigenesis in female SHR mice. *Aging (Albany NY)*. 2016; 8:1866–1875. <https://doi.org/10.18632/aging.101024>.
20. Panchenko AV, Fedoros EI, Pigarev SE, Maydin MA, Gubareva EA, Yurova MN, Kireeva GS, Lanskikh GP, Tyndyk ML, Anisimov VN. Effect of the polyphenol composition BP-C3 on haematological and intestinal indicators of 5-fluorouracil toxicity in mice. *Exp Ther Med*. 2018; 15:3124–3132. <https://doi.org/10.3892/etm.2018.5782>.
21. Hertkorn N, Frommberger M, Witt M, Koch BP, Schmitt-Kopplin P, Perdue EM. Natural organic matter and the event horizon of mass spectrometry. *Anal Chem*. 2008; 80:8908–8919. <https://doi.org/10.1021/ac800464g>.
22. Dier TK, Egele K, Fossog V, Hempelmann R, Volmer DA. Enhanced mass defect filtering to simplify and classify complex mixtures of lignin degradation products. *Anal Chem*. 2016; 88:1328–1335. <https://doi.org/10.1021/acs.analchem.5b03790>.
23. Zhrebek A, Kostyukevich Y, Kononikhin A, Kharybin O, Konstantinov AI, Zaitsev KV, Nikolaev E, Perminova IV. Enumeration of carboxyl groups carried on individual components of humic systems using deuteromethylation and Fourier transform mass spectrometry. *Anal Bioanal Chem*. 2017; 409:2477–2488. <https://doi.org/10.1007/s00216-017-0197-x>.
24. Terstappen GC, Schlüpen C, Raggiaschi R, Gaviraghi G. Target deconvolution strategies in drug discovery. *Nat Rev Drug Discov*. 2007; 6:891–903. <https://doi.org/10.1038/nrd2410>.
25. Je nkings JL, Bender A, Davies JW. In silico target fishing: Predicting biological targets from chemical structure. *Drug Discov Today Technol*. 2006; 3:413–421. <https://doi.org/10.1016/j.ddtec.2006.12.008>.
26. Cereto-Massagué A, Ojeda MJ, Valls C, Mulero M, Pujadas G, Garcia-Vallve S. Tools for *in silico* target fishing. *Methods*. 2015; 71:98–103. <https://doi.org/10.1016/j.ymeth.2014.09.006>.
27. Keum J, Yoo S, Lee D, Nam H. Prediction of compound-target interactions of natural products using large-scale drug and protein information. *BMC Bioinformatics*. 2016; 17:219. <https://doi.org/10.1186/s12859-016-1081-y>.
28. Hufsky F, Scheubert K, Böcker S. New kids on the block: novel informatics methods for natural product discovery. *Nat Prod Rep*. 2014; 31:807–817. <https://doi.org/10.1039/c3np70101h>.
29. Hubert J, Nuzillard JM, Renault JH. Dereplication strategies in natural product research: How many tools and methodologies behind the same concept? *Phytochem Rev*. 2017; 16:55–95. <https://doi.org/10.1007/s11101-015-9448-7>.
30. Sakakibara A. A structural model of softwood lignin. *Wood Sci Technol*. 1980; 14:89–100. <https://doi.org/10.1007/BF00584038>.
31. Nebbioso A, Piccolo A, Spiteller M. Limitations of electrospray ionization in the analysis of a heterogeneous mixture of naturally occurring hydrophilic and hydrophobic compounds. *Rapid Commun Mass Spectrom*. 2010; 24:3163–70. <https://doi.org/10.1002/rcm.4749>.
32. Kim S, Kramer RW, Hatcher PG. Graphical method for analysis of ultrahigh-resolution broadband mass spectra of natural organic matter, the van Krevelen diagram. *Anal Chem*. 2003; 75:5336–5344.
33. Hockaday WC, Purcell JM, Marshall AG, Baldock JA, Hatcher PG. Electrospray and photoionization mass spectrometry for the characterization of organic matter in natural waters: a qualitative assessment. *Limnol Oceanogr: Methods*. 2009; 7:81–95. <https://doi.org/10.4319/lom.2009.7.81>.
34. Waggoner DC, Chen H, Willoughby AS, Hatcher PG. Formation of black carbon-like and alicyclic aliphatic compounds by hydroxyl radical initiated degradation of lignin. *Org Geochem*. 2015; 82:69–76. <https://doi.org/10.1016/j.orggeochem.2015.02.007>.
35. Anisimov VN, Larsen S, Lofberg S, Baldueva IA, Malek AV, Kirkegaard Nielsen T, Fedoros EI, Perminova IV, Drobyshev EY, Bykov VN, Panchenko AV, Scherbakov AM, Belyaev AM. Results and prospects of development of new polyphenolic drugs for cancer patients. *Oncotarget*. 2017; 8:100951–100956. <https://doi.org/10.18632/oncotarget.22307>.
36. Grivennikov SI, Greten FR, Karin M. Immunity, inflammation, and cancer. *Cell*. 2010; 140:883–99. <https://doi.org/10.1016/j.cell.2010.01.025>.
37. Chen JJ, Johnson EJ. Targeting the bradykinin B1 receptor to reduce pain. *Expert Opin Ther Targets*. 2007; 11:21–35. <https://doi.org/10.1517/14728222.11.1.21>.
38. Souza DG, Lomez ESL, Pinho V, Pesquero JB, Bader M, Pesquero JL, Teixeira MM. Role of bradykinin B2 and B1

- receptors in the local, remote, and systemic inflammatory responses that follow intestinal ischemia and reperfusion injury. *J Immunol.* 2004; 172:2542–8.
39. Oenema TA, Kolahian S, Nanninga JE, Rieks D, Hiemstra PS, Zuyderduyn S, Halayko AJ, Meurs H, Gosens R. Pro-inflammatory mechanisms of muscarinic receptor stimulation in airway smooth muscle. *Respir Res.* 2010; 11:130. <https://doi.org/10.1186/1465-9921-11-130>.
 40. Thurmond RL, Gelfand EW, Dunford PJ. The role of histamine H1 and H4 receptors in allergic inflammation: the search for new antihistamines. *Nat Rev Drug Discov.* 2008; 7:41–53. <https://doi.org/10.1038/nrd2465>.
 41. Ganesh T. Prostanoid receptor EP2 as a therapeutic target. *J Med Chem.* 2014; 57:4454–65. <https://doi.org/10.1021/jm401431x>.
 42. Hübner S, Dejager L, Libert C, Tuckermann JP. The glucocorticoid receptor in inflammatory processes: transrepression is not enough. *Biol Chem.* 2015; 396:1223–31. <https://doi.org/10.1515/hsz-2015-0106>.
 43. Blackburn MR, Vance CO, Morschl E, Wilson CN. Adenosine Receptors and Inflammation. *Handb Exp Pharmacol.* 2009; 193:215–269. https://doi.org/10.1007/978-3-540-89615-9_8.
 44. Akkari R, Burbiel JC, Hockemeyer J, Müller CE. Recent progress in the development of adenosine receptor ligands as antiinflammatory drugs. *Curr Top Med Chem.* 2006; 6:1375–1399. <https://doi.org/10.2174/15680266106061375>.
 45. Shacter E, Weitzman SA. Chronic inflammation and cancer. *Oncology (Williston Park).* 2002; 16:217–226.
 46. da Costa PLN, Sirois P, Tannock IF, Chammas R. The role of kinin receptors in cancer and therapeutic opportunities. *Cancer Letters.* 2014; 345:27–38. <https://doi.org/10.1016/j.canlet.2013.12.009>.
 47. Herner A, Sauliunaite D, Michalski CW, Erkan M, De Oliveira T, Abiatari I, Kong B, Esposito I, Friess H, Kleeff J. Glutamate increases pancreatic cancer cell invasion and migration via AMPA receptor activation and Kras-MAPK signaling. *International Journal of Cancer.* 2011; 129:2349–2359. <https://doi.org/10.1002/ijc.25898>.
 48. Ribeiro MPC, Custodio JBA, Santos AE. Ionotropic glutamate receptor antagonists and cancer therapy: time to think out of the box? *Cancer Chemotherapy and Pharmacology.* 2017; 79:219–225. <https://doi.org/10.1007/s00280-016-3129-0>.
 49. Banu SK, Lee J, Speights VOJ, Starzinski-Powitz A, Arosh JA. Selective inhibition of prostaglandin E2 receptors EP2 and EP4 induces apoptosis of human endometriotic cells through suppression of ERK1/2, AKT, NFkappaB, and beta-catenin pathways and activation of intrinsic apoptotic mechanisms. *Mol Endocrinol.* 2009; 23:1291–1305. <https://doi.org/10.1210/me.2009-0017>.
 50. Sung YM, He G, Hwang DH, Fischer SM. Overexpression of the prostaglandin E2 receptor EP2 results in enhanced skin tumor development. *Oncogene.* 2006; 25:5507–5516. <https://doi.org/10.1038/sj.onc.1209538>.
 51. Fernández-Martínez AB, Lucio-Cazaña J. Intracellular EP2 prostanoid receptor promotes cancer-related phenotypes in PC3 cells. *Cellular and Molecular Life Sciences.* 2015; 72:3355–3373. <https://doi.org/10.1007/s00018-015-1891-5>.
 52. Allard B, Beavis PA, Darcy PK, Stagg J. Immunosuppressive activities of adenosine in cancer. *Current Opinion in Pharmacology.* 2016; 29:7–16. <https://doi.org/10.1016/j.coph.2016.04.001>.
 53. Allard D, Turcotte M, Stagg J. Targeting A2 adenosine receptors in cancer. *Immunology and Cell Biology.* 2017; 95:333–339. <https://doi.org/10.1038/icb.2017.8>.
 54. Virgilio FD, Adinolfi E. Extracellular purines, purinergic receptors and tumor growth. *Oncogene.* 2017; 36:293–303. <https://doi.org/10.1038/onc.2016.206>.
 55. Dolma S, Selvadurai HJ, Lan X, Lee L, Kushida M, Voisin V, Whetstone H, So M, Aviv T, Park N, Zhu X, Xu C, Head R, et al. Inhibition of Dopamine Receptor D4 Impedes Autophagic Flux, Proliferation, and Survival of Glioblastoma Stem Cells. *Cancer Cell.* 2016; 29:859–873. <https://doi.org/10.1016/j.ccell.2016.05.002>.
 56. Tang J, Li Z, Lu L, Cho CH. β -Adrenergic system, a backstage manipulator regulating tumour progression and drug target in cancer therapy. *Seminars in Cancer Biology.* 2013; 23:533–542. <https://doi.org/10.1016/j.semcancer.2013.08.009>.
 57. Cole SW, Sood AK. Molecular pathways: beta-adrenergic signaling in cancer. *Clin Cancer Res.* 2012; 18:1201–1206. <https://doi.org/10.1158/1078-0432.CCR-11-0641>.
 58. Zhernov YV, Kremb S, Helfer M, Schindler M, Harir M, Mueller C, Hertkorn N, Avvakumova NP, Konstantinov AI, Brack-Werner R, Schmitt-Kopplin P, Perminova IV. Supramolecular Combinations of Fractionated Humic Polyanions as Potent and Cost-Effective Microbicides with Polymodal anti-HIV-Activities and Low Cytotoxicity. *New Journal of Chemistry.* 2017; 41:212–224. <https://doi.org/10.1039/C6NJ00960C>.
 59. Shamsabadipour S, Ghanadian M, Aeedi H, Rahimnejad MR, Mohammadi-Kamalabadi M, Ayatollahi SM, Salimzadehf L. Triterpenes and Steroids from *Euphorbia denticulata* Lam. With Anti-Herpes Symplex Virus Activity. *Iran J Pharm Res.* 2013; 12:759–767.
 60. Hu JM, Chen JJ, Yu H, Zhao YX, Zhou J. Five new compounds from *Dendrobium longicornu*. *Planta Med.* 2008; 74:535–539. <https://doi.org/10.1055/s-2008-1074492>.
 61. Li M, Yunqiao Pu Y, Tschaplinski TJ, Ragauskas AJ. 31P NMR Characterization of Tricin and Its Structurally Similar Flavonoids. *ChemistrySelect.* 2017; 2:3557–3561. <https://doi.org/10.1002/slct.201700735>.
 62. Sjöström E, Alén R, editors. Analytical methods in wood chemistry, pulping, and papermaking. Springer Series in Wood Science. Springer-Verlag Berlin Heidelberg, 2013.
 63. Drummond EM, Harbourne N, Marete E, Martyn D, Jacquier J, O’Riordan D, Gibney ER. Inhibition of proinflammatory biomarkers in thp1 macrophages by

- polyphenols derived from chamomile, meadowsweet and willow bark. *Phytother Res.* 2012; 27:588–594. <https://doi.org/10.1002/ptr.4753>.
64. Lee WJ, Chen YR, Tseng TH. Quercetin induces fasl-related apoptosis, in part, through promotion of histone h3 acetylation in human leukemia hl-60 cells. *Oncol Rep.* 2011; 25:583–591. <https://doi.org/10.3892/or.2010.1097>.
 65. Lin CW, Hou WC, Shen SC, Juan SH, Ko CH, Wang LM, Chen YC. Quercetin inhibition of tumor invasion via suppressing PKC delta/ERK/AP-1-dependent matrix metalloproteinase-9 activation in breast carcinoma cells. *Carcinogenesis.* 2008; 29:1807–1815. <https://doi.org/10.1093/carcin/bgn162>.
 66. Mamani-Matsuda M, Kauss T, Al-Kharrat A, Rambert J, Fawaz F, Thiolat D, Moynet D, Coves S, Malvy D, Mossalayi MD. Therapeutic and preventive properties of quercetin in experimental arthritis correlate with decreased macrophage inflammatory mediators. *Biochem Pharmacol.* 2006; 72:1304–1310. <https://doi.org/10.1016/j.bcp.2006.08.001>.
 67. Comalada M, Ballester I, Bailón E, Sierra S, Xaus J, Gálvez J, de Medina FS, Zarzuelo A.. Inhibition of pro-inflammatory markers in primary bone marrow-derived mouse macrophages by naturally occurring flavonoids: analysis of the structure-activity relationship. *Biochem Pharmacol.* 2006; 72:1010–1021. <https://doi.org/10.1016/j.bcp.2006.07.016>.
 68. Wang J, Zhang Q, Jin S, He D, Zhao S, Liu S. Genistein modulate immune responses in collagen-induced rheumatoid arthritis model. *Maturitas.* 2008; 59:405–412.
 69. Benassayag C, Perrot-Applanat M, Ferre F. Phytoestrogens as modulators of steroid action in target cells. *J. Chromatogr. B Analyt. Technol. Biomed. Life Sci.* 2002; 777:233–248.
 70. Davis TA, Clarke TK, Mog SR, Landauer MR. Subcutaneous administration of genistein prior to lethal irradiation supports multilineage, hematopoietic progenitor cell recovery and survival. *Int J Radiat Biol.* 2007; 83:141–151.
 71. Schaffer S, Halliwell B. Do polyphenols enter the brain and does it matter? Some theoretical and practical considerations. *Genes Nutr.* 2012; 7:99–109. <https://doi.org/10.1007/s12263-011-0255-5>.
 72. Kavvadias D, Sand P, Youdim KA, Qaiser MZ, Rice-Evans C, Baur R, Sigel E, Rausch WD, Riederer P, Schreier P. The flavone hispidulin, a benzodiazepine receptor ligand with positive allosteric properties, traverses the blood–brain barrier and exhibits anticonvulsive effects. *British Journal of Pharmacology.* 2004; 142:811–820. <https://doi.org/10.1038/sj.bjp.0705828>.
 73. Butelman ER, Kreek MJ. Salvinorin A, a kappa-opioid receptor agonist hallucinogen: pharmacology and potential template for novel pharmacotherapeutic agents in neuropsychiatric disorders. *Front Pharmacol.* 2015; 6:190. <https://doi.org/10.3389/fphar.2015.00190>.
 74. Kim JH, Park CY, Lee SJ. Effects of Sun Ginseng on subjective quality of life in cancer patients: A double-blind, placebo-controlled pilot trial. *J Clin Pharm Ther.* 2006; 31:331–334.
 75. Du JR, Long FY, Chen C. Research Progress on Natural Triterpenoid Saponins in the Chemoprevention and Chemotherapy of Cancer. *Enzymes.* 2014; 36:95–130. <https://doi.org/10.1016/B978-0-12-802215-3.00006-9>.
 76. Nag SA, Qin JJ, Wang W, Wang MH, Wang H, Zhang R. Ginsenosides as Anticancer Agents: *In vitro* and *in vivo* Activities, Structure-Activity Relationships, and Molecular Mechanisms of Action. *Front Pharmacol.* 2012; 3:25. <https://doi.org/10.3389/fphar.2012.00025>.
 77. Xu K, Shu Z, Xu QM, Liu YL, Li XR, Wang YL, Yang SL. Cytotoxic activity of Pulsatilla chinensis saponins and their structure-activity relationship. *J Asian Nat Prod Res.* 2013; 15:680–6. <https://doi.org/10.1080/10286020.2013.790901>.
 78. Chen Y, Du Y, Li Y, Wang X, Gao P, Yang G, Fang Y, Meng Y, Zhao X. Panaxadiol Saponin and Dexamethasone Improve Renal Function in Lipopolysaccharide-Induced Mouse Model of Acute Kidney Injury. *PLoS One.* 2015; 10:e0134653. <https://doi.org/10.1371/journal.pone.0134653>.
 79. Dutra RC, Bicca MA, Segat GC, Silva KA, Motta EM, Pianowski LF, Costa R, Calixto JB. The antinociceptive effects of the tetracyclic triterpene euphol in inflammatory and neuropathic pain models: The potential role of PKCε. *Neuroscience.* 2015; 303:126–37. <https://doi.org/10.1016/j.neuroscience.2015.06.051>.
 80. Kovács A, Vasas A, Hohmann J. Natural phenanthrenes and their biological activity. *Phytochemistry.* 2008; 69:1084–110. <https://doi.org/10.1016/j.phytochem.2007.12.005>.
 81. Xue Z, Li S, Wang S, Wang Y, Yang Y, Shi J, He L. Mono-, Bi-, and triphenanthrenes from the tubers of *Cremastra appendiculata*. *J Nat Prod.* 2006; 69:907–913.
 82. Barrajón-Catalán E, Herranz-López M, Joven J, Segura-Carretero A, Alonso-Villaverde C, Menéndez JA, Micol V. Molecular promiscuity of plant polyphenols in the management of age-related diseases: far beyond their antioxidant properties. *Adv Exp Med Biol.* 2014; 824:141–59. https://doi.org/10.1007/978-3-319-07320-0_11.
 83. Martin DA, Bolling BW. A review of the efficacy of dietary polyphenols in experimental models of inflammatory bowel diseases. *Food Funct.* 2015; 6:1773–86. <https://doi.org/10.1039/c5fo00202h>.
 84. Upadhyay S, Dixit M. Role of Polyphenols and Other Phytochemicals on Molecular Signaling. *Oxid Med Cell Longev.* 2015; 2015:504253. <https://doi.org/10.1155/2015/504253>.
 85. Kovalevskii DV, Permin AB, Perminova IV, Petrosyan VS. Conditions for acquiring quantitative ¹³C NMR spectra of humic substances. *Moscow Univ Chem Bull (Vestnik Moskovskogo Universiteta. Khimiya).* 2000; 41:39–42.

86. Zherebker A, Turkova AV, Kostyukevich Y, Kononikhin A, Zaitsev KV, Popov IA, Nikolaev E, Perminova IV. Synthesis of carboxylated styrene polymer for internal calibration of Fourier transform ion cyclotron resonance mass-spectrometry of humic substances. *Eur J Mass Spectrom (Chichester)*. 2017; 23:156–161. <https://doi.org/10.1177/1469066717718963>.
87. Sleighter RL, McKee GA, Liu Z, Hatcher PG. Naturally present fatty acids as internal calibrants for Fourier transform mass spectra of dissolved organic matter. *Limnol Oceanogr Methods*. 2008; 6:246–253.
88. Kunenkov EV, Kononikhin AS, Perminova IV, Hertkorn N, Gaspar A, Schmitt-Kopplin P, Popov IA, Garmash AV, Nikolaev EN. Total Mass Difference Statistics Algorithm: A New Approach to Identification of High-Mass Building Blocks in Electrospray Ionization Fourier Transform Ion Cyclotron Mass Spectrometry Data of Natural Organic Matter. *Anal Chem*. 2009; 81:10106–10115. <https://doi.org/10.1021/ac901476u5>.
89. ChemAxon. Instant JChem was used for structure database management, search and predictionInstant, *J Chem* 17.2.6.0. [Online]. 2017. Available from: <http://www.chemaxon.com>.
90. Sander T, Freyss J, von Korff M, Rufener C. DataWarrior: an open-source program for chemistry aware data visualization and analysis. *J Chem Inf Model*. 2015; 55:460–47. <https://doi.org/10.1021/ci500588j>.
91. Orlov AA, Zherebker AY, Eletskaia AA, Chernikov VS, Kozlovskaya LI, Zhernov YV, Palyulin VA, Osolodkin DI, Perminova IV. Chemoinformatic approach to identification of antiviral components in humic substances. *Chemical Information Bulletin*. 2018; 70:46. http://bulletin.acscinf.org/PDFs/CIB_70-1.pdf.
92. ChemAxon. Standardizer was used for structure canonicalization and transformation, *J Chem* 16.8.29.0. [Online]. 2016. Available from: <http://www.chemaxon.com>.