Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.
Antiviral activity of natural humic substances and shilajit materials against HIV-1: Relation to structure

Yury V. Zhernov, Andrey I. Konstantinov, Alexander Zherebker, Eugene Nikolaev, Alexey Orlov, Mikhail I. Savinykh, Galina V. Kornilaeva, Eduard V. Karamov, Irina V. Perminova

A National Research Center – Institute of Immunology FMBA of Russia, Moscow, 115522, Russia
b Lomonosov Moscow State University, Department of Chemistry, Moscow, 119991, Russia
c Skolkovo Institute of Science and Technology, Skolkovo, Moscow Region, 143026, Russia
d Scientific and Production Company “Sibdalmumiyo” Ltd., Novokuznetsk, Russia
e D.I. Ivanovsky Institute of Virology FSBI «National Research Center for Epidemiology and Microbiology Named After the Honorary Academician N.F. Gamaleya», Moscow, 123098, Russia

ARTICLE INFO

Keywords:
Humic acid
Fulvic acid
Shilajit
Antiviral
HIV
Structure-activity
13C NMR
FTICR MS
ChemBL data-mining

ABSTRACT

Natural products, such as humic substances (HS) and shilajit, are known to possess antiviral activity. Humic-like components are often called as carriers of biological activity of shilajit. The goal of this study was to evaluate anti-HIV activity of well characterized HS isolated from coal, peat, and peloids, and compare it to that of water-soluble organic matter (OM) isolated from different samples of Shilajit. The set of humic materials included 16 samples of different fractional composition: humic acid (HA), hymatomelanic acid (HMA), fulvic acid (FA). The set of shilajit OM included 19 samples of different geographic origin and level of alteration. The HIV-1 p24 antigen assay and cell viability test were used for assessment of antiviral activity. The HIV-1 Bru strain was used to infect CEM-SS cells. The obtained EC50 values varied from 0.37 to 1.4 mg L$^{-1}$ for the humic materials, and from 14 to 142 mg L$^{-1}$ for the shilajit OM. Hence, all humic materials used in this study outcompeted largely the shilajit materials with respect to anti-HIV activity: For the humic materials, the structure-activity relationships revealed strong correlation between the EC50 values and the content of aromatic carbon indicating the most important role of aromatic structures. For shilajit OM, the reverse relationship was obtained indicating the different mechanism of shilajit activity. The FTICRMS molecular assignments were used for ChEMBL data mining in search of the active humic molecules. As potential carriers of antiviral activity were identified aromatic structures with alkyl substituents, terpenoids, N-containing analogs of typical flavonoids, and aza-podophyllotoxins. The conclusion was made that the typical humic materials and Shilajit differ greatly in molecular composition, and the humic materials have substantial preferences as a natural source of antiviral agents as compared to shilajit.

1. Introduction

Fast evolution and mutation of viruses set forward a need for new antiviral agents suitable for treatment of drug-resistant infections. The multiple strains of the influenza viruses show resistance to adamantane derivatives (Dong et al., 2015), the herpes virus family – to acyclovir (Pottage et al., 1995), retroviruses – to azithromycin (Jenninga et al., 2001). Due to a vast genetic diversity, the human immunodeficiency virus (HIV) develops quickly the resistance to monomolecular antiretroviral drugs (Karamov et al., 2018; Moskaleychik et al., 2015). Moreover, the pandemy of COVID-19, which hit already more than 5 million people globally, shows how limited a pool of antiviral drugs is (Dong et al., 2020; Ford et al., 2020; Li and De Clercq, 2020). In this respect, the natural supramolecular systems of biologically active compounds might be of particular value. Humic substances (HS) have been studied for a long time for their antiviral activity (Helbig et al., 1997; Jooen et al., 2003; Kloecking et al., 2002). The different fractions of natural HS and synthetic HS-like materials are active against HIV (Botes et al., 2002; Brucolleri et al., 2013; Kornilaeva et al., 2019; Schneider

* Corresponding author. Leninskie Gory 1-3 119991, Moscow, Russia.
E-mail address: iperm@org.chem.msu.ru (I.V. Perminova).
HS are the products of deep chemical and microbial oxidation of bioprecursors – lignins, tannins, peptides, cellulose, lipids, terpenoids, released from the remnants of plants and animals (Stevenson, 1994). These oxidized biomolecular precursors are assembled into supramolecular multicomponent systems of carboxyl and hydroxyl rich molecules. Given the presence of extended carbon backbone in these molecules, they can be considered as carboxylated polyanions (PAs).

The latter have been intensively studied as components of microbicides aimed at prophylactics of viral infections (Zhernov and Khaitov, 2019; Zhan et al., 2016; Andree-Marobela et al., 2013; Daglia, 2012). Shilajit is another natural biologically active material, which is commonly used in traditional medicine (Agarwal et al., 2007). It is represented by brown to blackish-brown resinous matter found in the different mountain regions and considered to be a complex product of biotransformation of plant residues (Agarwal et al., 2007). Shilajit was reported as a dose-dependent inhibitor against HSV-1, HSV-2, human cytomegalovirus (HCMV), and human immunodeficiency virus (HIV) (Gagné et al., 2015). These are humic-like components, which are suggested to provide for multiple therapeutic properties of Shilajit (Schepekin et al., 2003, 2009). However, the studies on direct comparison of the biological activity of HS from different sources (e.g., soil, peat, coal) and shilajit, as well as on the corresponding structure-activity relationships, are missing.

In our previous study on the anti-HIV activity of HS isolated from peoloids (fresh water bottom sediments), we observed that the strongest inhibiting activity was characteristic of the most hydrophobic fractions of humic (HA) and hymatomelanic acids (HMA) (Zhernov et al., 2017). The most hydrophilic fraction of fulvic acid (FA) had the least activity. The structure-activity analysis revealed direct relationship between antiviral activity and the ratio of aromatic to aliphatic structures (lipophilicity index) of the different humic fractions used, and the inverse relationship with their carboxylic and total acidity. Given these previous results, we expanded largely the set of HS samples on the account of the most hydrophobic natural HS isolated from coal and peat. In addition, we included a broad set of crude shilajit samples to facilitate a direct comparison of biological activity of these two different natural products.

2. Materials and methods

2.1. Humic and shilajit materials used in this study

A set of humic and humic-like materials included 16 fractionated HS samples (HA, FA, HMA, as well as gray humic acid (CHG) from natural sources (coal, peat, peloid), one sample of oxidized lignin, and one sample of synthetic FA. Four samples of coal humic acids (CHA) were isolated by precipitation with acid followed by dialysis from the commercially available samples of potassium humates (Powhumus, Irktst Humate, Sakhalin Humate, Activated Humic Acid) and designated as (CHA-P), (CHA-I), (CHA-S), and CHA-A, respectively. One sample was isolated from lignite of the Baikal region (CHA-G) using alkali extraction (0.3 M NaOH). Peat HS included highmoor humic acid (PHA-T) and fulvic acid (PFA-5); lowmoor non-fractionated HS (PHF-1), and humic acid (PHA-1, PHA-3). Peloid humic materials were extracted from bottom sediments of the Molochka lake in Samara region and included humic acid (PelHA), fulvic acid (PelFA) and hymatomelanic acid (PelHMA). Peloid HS were extracted and fractionated according to the scheme described in our previous publication (Zhernov et al., 2017). The samples set also included veterinary pharmaceutical drug Ligifolium produced from the oxidized lignin, and a fulvic-like fraction of synthetic HS (MHQ-FA) synthesized via oxidative polymerization of phenols as described by (Zherebker et al., 2015). The latter represented a structural analogue of HS with the known molecular features.

Seventeen samples of raw shilajit materials originated from the different geographic regions (the Altai mountains in Russia and Mongolia, the Aldan river basin in Yakutia, Central Asia, the Caucasus). They were homogenised and dissolved in distilled water. Water extracts were centrifuged to separate insoluble part, and dried in a vacuum oven at temperature of 45 °C. This way 17 samples of shilajit materials (SM) were prepared. The set of samples also included Altai “Golden mumiyo” (Sh-Alt, Russia) and Himalayan “Shilajit Dabur” (Sh-Him, India) pharmaceutical shilajit preparations.

The full set of humic and shilajit materials used in this study, along with their CHNO elemental compositions are listed in Table S1 in Supplementary Material.

2.2. Structural characterization of the HS and shilajit materials

$^{13}$C NMR spectra were acquired using a Bruker Avance 400 NMR spectrometer operating at 400 MHz proton frequency. All humic and shilajit solutions were prepared as following: a weight of 45 mg of a sample was dissolved in 0.6 mL of 0.3M NaOD/D2O with isotope purity of 99.7+% (Aldrich). For suppressing the nuclear Overhauser effect, the broadband decoupling from protons was switched off for the relaxation delay. The relaxation time delay was 7.8 s in accordance with the conditions described by (Kovalovsky et al., 2000). The NMR assignments were made after (Herkeorn et al., 2002); carbon of alkyl chains (CHn) – 0–50 ppm; carbon of alkoxide groups CH3O – 50–58 ppm, CH2O – 58–65 ppm, CHO – 65–90 ppm, and OCO – 90–110 ppm; aromatic carbon bonded with H or C atoms (Car) – 110–145 ppm; aromatic carbon...
bonded with oxygen (CarO) – 145–165 ppm; carbon of carboxylic groups (COO) – 165–187 ppm; carbon of carbonyl groups (C=O) – 187–220 ppm. In case of shilajit samples characterized with significant content of nitrogen (see Table S3 in Supplementary Material), the spectral intervals of 50–65 and 165–187 ppm were also attributed, respectively, to CH–N groups of amino acids and to amide structures (CO–N).

2.3. Compilation of the FTICR MS data set for the HS and shilajit materials

For characterization of molecular space of the set of HS samples used in this study the previously published FTICR mass-spectrometric data (peak lists) were used for CHA-G, CHA-P, CHA-S, PelHA, PelHMA, PekFA, PHA-5, PFA-5, MHQ-FA were described (Orlov et al., 2019; Zhernov et al., 2017). Mass-spectrum of the Shilajit sample Sh1 was obtained using 7T FTICR MS Apex Ultra (Bruker Daltonics, Germany) equipped with the harmonized cell (Nikolaev et al., 2011) and electrospray ionization (ESI) source operating in negative ion mode. The spectra were acquired with a time domain of 4 megawords in ESI (−) and 300 scans were accumulated for each spectrum. Resolving power was 530 000 at m/z = 400. Here all spectra were first externally calibrated using the synthesized carboxylated polystyrene as described previously (Zherubek et al., 2017), followed by internal calibration using fatty acids residual signals (Sleight et al., 2008) which yielded accuracy values < 0.5 ppm. Only peaks with a signal to noise (S/N) ratio exceeding 6 were used for formulae assignment.

Molecular compositions were assigned using the open source browser-based application UltraMassExplorer created by Leefmann et al. (http://clockersrv1.awi.de:3838/ume) (Leefmann et al., 2019). The pivot table with all assigned formulae is presented in Supplementary file Table S5. The generated CHONS formulas were validated by setting sensible chemical constraints typical for DOM (O/C ratio ≤ 1, 0.3<H/C ratio ≤ 2.2, element counts (C ≤ 120, H ≤ 200, O≤60, N ≤ 2, S ≤ 1) and mass accuracy window < 0.5 ppm) (Kunenkov et al., 2009; Sleight et al., 2007). Additionally, for HS only formulae presenting in at least two samples were considered. All formulae were further divided into three categories according to a modified aromaticity index (A_{mod}) proposed by (Koch et al., 2016) A_{mod} was calculated according to the following equation (1):

\[ A_{mod} = \frac{1 + C - 0.5O - S - 0.5(N + H)}{C - 0.5O - N - S} \]  

(1)

where C, H, O, N, S are the numbers of atoms in molecular formulae.

2.4. HIV-1 p24 antigen assay and cell viability test

The HIV-1 Bru strain was used in this study to infect CEM-SS cells which were derived from the human lymphoblastoid cell line CEM that expresses high levels of the CD4 antigen (Wain-Hobson et al., 1991). The cells were cultured in the suspension medium RPMI-1640 (Gibco) containing 10% FCS. The CEM-SS cells were cultured at a density of 7.0 × 10⁴ well and infected with HIVbru at the multiplicity of infection (MOI) of 1. All solutions of the HS and shilajit samples were sterilized using 0.22 μm syringe-filters (Merck Millipore Ltd) and stored at −20 °C until use. CEM-SS cells were incubated for 24 h in the absence and presence of HS and SM samples in 96-well plates. The HS and shilajit samples were homogeneous and stable throughout the incubation period in a 96-well plate. As a control, we used the reference HIV-1 inhibitors: the reverse transcriptase inhibitor Azidothymidine (AZT) and the integrase inhibitor Raltegravir (RAL). They were purchased from Sigma-Aldrich. After 3 days following three washings, CEM-SS cells were cultured in 96 well plates for 3 days to measure p24 release. The level of virus replication was detected by ELISA using a «HIV-1 p24 antigen-ELISA-BEST» test system (Vector-Best, Novosibirsk, Russia). The effects on cell viability were monitored by MTT (3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyl tetrazolium bromide) assay, which measures the reduction of yellow MTT to purple formazan by mitochondrial enzymes (Mosmann, 1983).

2.5. Statistical analyses

EC50 and CC50 values were calculated using GraphPad Prism v.5.00 software (GraphPad Software, La Jolla, CA). The approximation for sigmoidal dose-response was used with variable slope by setting the following constraints: fixed top of 100, and fixed bottom of 0. As internal acceptance criteria, R² values were set to >0.9. The therapeutic value of the drug was estimated using the selectivity index (SI). It is calculated as a ratio of 50% cytotoxic concentration (CC50) to 50% effective concentration (EC50). Statistical significance was determined using one-way Anova with the Bonferroni post hoc test (p < 0.0001).

2.6. ChEMBL data mining and analysis

ChEMBL data retrieval was performed similarly to the previously developed pipeline (Orlov et al., 2019; Fedoros et al., 2018). MySQL edition of ChEMBL26 was used for data mining. Briefly, ChEMBL dump file was put into a local MySQL database. All mofregno and canonical_smiles entries were extracted from compound_structures table. Entries of canonical_smiles field containing disconnected fragments (entries containing ‘-‘ symbol) were additionally preprocessed in the following way. All entries were split at ‘-‘ symbol and the largest SMILES substring was kept. Compounds corresponding to these SMILES were further neutralized using RDKit (RDKit: Open-source chemoinformatics; http://www.rdkit.org). rdMolstandardize module and formulae for all found compounds were generated using RDKit. Then the formulae of the samples with the Spearman rank correlation coefficient with EC50 lower or equal to −0.7 (339 formulae) were searched among all formulae from the ChEMBL data base. The data on compound structures (compound_structures), activities (activities), bioassays (assays), etc. were extracted from ChEMBL. To retrieve the activity entries relevant to HIV, case insensitive substring query searching using ‘HIV’ and ‘human immunodeficiency’ strings was performed in assays.description and assays. organism, assays.mc.organism fields. All extracted entries were manually analyzed. Network-based representation was constructed using DataWarrior (v.5.2.1) software (Sander et al., 2015). Stereo-depleted SMILES strings were used for the calculation of DataWarrior’s FragFP fingerprints and network building. Tanimoto index cut-off was set to 0.8. Thus, each network node corresponds to the structure (with removed stereochemistry information) and edge corresponds to structural similarity index above cut-off.

3. Results and discussion

3.1. Antiviral activity and cytotoxicity of the humic and shilajit materials used in this study

The cytotoxicity of the humic and shilajit samples with respect to the CEM-SS cells was assessed using MTT assay. For the majority of the humic and shilajit samples at a concentration of 1000 mg L⁻¹, the mitochondrial activity was above 80%. Ligfolum exhibited low toxicity (CC50 = 708 mg L⁻¹). The FA samples under study at a concentration of 1000 mg L⁻¹ also caused a decrease in mitochondrial activity below 80%, but it was still above 70%. The obtained results are indicative of good safety profile both for humic and shilajit materials used in this study.

The typical dose-inhibition curves against HIV-1 infection for the humic and shilajit materials of different origin and/or fractional composition are shown in Fig. 1. The full set of the EC50 values obtained for the humic and Shilajit samples used in this study is summarized in Table 1. It can be seen that the humic materials were characterized with much higher inhibitory activity with respect to HIV-1 replication in
The typical dose-inhibition curves for the humic and shilajit materials from different sources and/or fractional composition (n = 3; R²>0.9). The samples are highlighted as follows: coal HA (red), peatoid FA (violet), peat HA (orange), peloid HMA (green), non-altered Sh (black), altered Sh (brown), Ligifolum (blue). (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

Table 1
The antiviral activity of the humic and shilajit materials used in this study.

| Humic samples | Source | EC50 (mg L⁻¹) | Shilajit samples | Description | EC50 (mg L⁻¹) |
|---------------|--------|---------------|-----------------|-------------|--------------|
| GHA-G         | Coal   | 0.37 ± 0.03   | Sh1             | Black and rust-colored, Altai | 97.69 ± 0.44 |
| GHA-5         | Coal   | 0.43 ± 0.03   | Sh3             | Black and rust-colored, Aldan | 54.63 ± 0.64 |
| CHG-P         | Coal   | 0.53 ± 0.05   | Sh4             | Blackish-brown, altered, Altai | 103.2 ± 0.2 |
| GHA-P         | Coal   | 0.53 ± 0.03   |                 | Blackish-brown, altered, Altai | 99.23 ± 0.29 |
| GHA-A         | Coal   | 0.69 ± 0.06   | Sh7             | Brown, altered, Altai | 104.2 ± 0.21 |
| GHA-1         | Coal   | 0.72 ± 0.04   | Sh9             | Black, low altered, Altai | 48.84 ± 0.41 |
| PeHMA         | Peloid | 0.45 ± 0.06   | Sh13            | Blackish-brown, altered, Altai | 38.78 ± 0.4 |
| PeHA          | Peloid | 0.58 ± 0.02   | Sh15            | Yellow-brown, altered, Pamir | 96.95 ± 0.64 |
| PeHFA         | Peloid | 1.00 ± 0.06   | Sh18            | Blackish-brown, altered, Altai | 60.81 ± 0.44 |
| Ligifolum     | Oxidized lignin | 0.96 ± 0.11 | Sh19            | Blackish-brown, altered, Central Asia | 26.13 ± 0.21 |
| PHA-5         | Peat   | 0.80 ± 0.03   | Sh20            | Blackish-brown, altered, Altai | 98.85 ± 0.41 |
| PHP-1         | Peat   | 1.13 ± 0.03   | Sh22            | Blackish-brown, altered, Altai | 40.1 ± 0.48 |
| PHP-3         | Peat   | 1.14 ± 0.03   | Sh25            | Black, low altered, Altai | 310.1 ± 0.32 |
| PHP-1         | Peat   | 1.20 ± 0.04   | Sh29            | Black, low altered, Afganistan | 14.03 ± 0.02 |
| PHP-3         | Peat   | 1.22 ± 0.03   | Sh30            | Black, low altered, Afganistan | 54.12 ± 0.36 |
| Fulvic samples | Source | EC50 (mg L⁻¹) | Sh32            | Black, altered, Caucasus | 90.6 ± 0.67 |
| PFA-5         | Peat   | 1.27 ± 0.01   | Sh38            | Black, low altered, Altai | 81.06 ± 0.41 |
| MHQ-Fa        | Synthetic FA | 1.27 ± 0.02 | Sh-Alt          | Pharmaceutical drug, Altai | 44.74 ± 0.22 |
| PeF HA        | Peloid | 1.41 ± 0.04   | Sh-Him          | Pharmaceutical drug, Himalayas | 48.0 ± 0.1 |

3.2. Structure-antiviral activity relationships derived from ¹³C NMR data on the humic and shilajit materials used in this study

In order to reveal the relationship between anti-HIV activity and structural features of the humic and shilajit materials used in this study, their structural composition was determined with a use of quantitative solution-state ¹³C NMR spectroscopy. The ¹³C NMR spectra of the HS samples under study were typical of HS represented by broad, highly overlapped signals (Fig. 3a,b).
The coal HA were characterized with the presence of intense spectral band in the range of aromatic carbon (110–165 ppm), signals of alkyl chains (0–50 ppm), and carboxylic groups (165–185 ppm). The minimum intensity was observed in the range of alkoxide (predominantly, carbohydrate) structures (50–100 ppm). The peat FA were characterized with the presence of an intense signal in the range of carbohydrate structures (60–90 ppm), aromatic and carboxylic groups. The aromatic region was characterized with three distinct peaks at 115, 130 and 156 ppm: the first and second ones are characteristic of aromatic carbon in o- and p-position to hydroxy groups, the third one corresponds to Car-O carbon indicating high degree of oxidation of FA and its enrichment with OH- and COOH- groups.

\(^{13}\)C NMR spectra of shilajit are very different from the HA-related materials (Fig. 3c). The spectra are characterized with the presence of a broad band of alkyl chains (0–50 ppm), sharp signals of carbohydrate and primary/secondary amine structures (65–90 ppm), and carboxylic groups (165–185 ppm). A distinct group of sharp peaks at 127–135 ppm, which were characteristic of all spectra of shilajit, was attributed to aromatic carbon of hippuric acid. The latter is a typical component of shilajit (Agarwal et al., 2007; Konstantinov et al., 2013). A sharp signal at 165 ppm was attributed to amide structures of peptide chains and hippuric acid. Despite the striking difference to coal HA, the \(^{13}\)C NMR spectra of shilajit were rather similar to the spectra of peat FA (Fig. 2b, c). Given that this study used water-extractable portion of shilajit organic matter, it should have much more in common with the acid-soluble fraction of HS - fulvic acid - as compared to the acid-insoluble humic acid. Fulvic acid is characterized with the highest contribution of oxygen-containing functions, which has some similarity with the organic matter of water-soluble fraction of shilajit.

Quantitative structural data for each sample were obtained as a set of integrals of \(^{13}\)C NMR spectral regions assigned to typical structural units of humic and shilajit materials. They are given in Table S2 in the SI. The highest content of aromatic carbon (Car and CarO) from 51 to 61% of the total C was characteristic of humic acids and their fractions isolated from coal and peloid. These samples were characterized with the lowest content of carbohydrate moieties and methoxyl groups (4–15%). On
obtained data are shown in Fig. 4 A and B and Figs. S1 and S2 in Supplementary Material). The discussed above quantitative data on structural group composition of the humic and shilajit samples were used for correlation analysis with the set of the corresponding data on antiviral activity. The obtained data are shown in Fig. 4 A and B and Figs. S1 and S2 in Supplementary Material. It can be seen that the EC50 values of the humic materials used in this study were inversely related to the content of unsubstituted aromatic carbon (Car) \( (r = -0.84; p = 0.00095) \) (Fig. 4 A).

A significant positive relationship was observed for the content of \( \text{CH}_2\text{O} \) groups \( (r = 0.7; p = 0.004) \). This might be indicative of a drop in activity of the HS samples with increasing content of carbohydrates. The correlation analysis for the shilajit samples yielded the opposite trends. The strongest inverse relationship was observed with the content of \( \text{CH}_2\text{O} \ & \text{CH}_2\text{O} \) or CH-N \( (r = -0.63; p = 0.004) \). On contrary to HS, antiviral activity of shilajit was inversely related to the content of aromatic structures reflected in the positive correlation between EC50 and Car \( (r = 0.61; p = 0.006) \) (Fig. 4 B). This might be indicative of the different mechanism of activity inherent within shilajit materials as compared to humic acid-dominated humic materials, e.g. coal, peat, peloids, etc. Given the substantial impact of structural group composition of both humic and shilajit samples on their antiviral activity, it was of particular interest to search for the relationships between molecular composition of HS and shilajit and their antiviral activity. The information on molecular composition is provided by FTICR MS. In this study, we had at our disposal FTICR MS data only for the subset of humic materials. This is a reason why we could conduct the respective studies only on the humic materials.

**3.3. Structure-antiviral activity relationships derived for the FTICR MS data on the HS samples used in this study**

For depiction of molecular composition of humic materials, we used FTICR MS data on nine (out of 16) HS samples used in this study. The amount of empirical formulae, which were present in, at least, 2 samples, ranged from 1902 to 4400. Van Krevelen diagrams for the respective HS samples are shown in Fig. 5. One available spectrum of the shilajit sample (Sh1) is given for comparison.

The dots in Van Krevelen diagrams are highlighted in different colors with respect to the value of aromaticity index (AI\text{mod}). The latter is introduced to evaluate aromaticity of molecular components of complex mixtures (Koch et al., 2006): the values of AI\text{mod} ≥ 0.67 and > 0.5 correspond to condensed aromatic and aromatic compounds, respectively. This representation reveals drastic difference in molecular composition of the samples used in this study: the humic acid samples from coal, peloid, and peat are dominated by condensed aromatic compounds, whereas the fulvic acid samples (and shilajit) are characterized by maximum contribution of highly oxidized unsaturated and saturated compounds with AI below 0.5. These data corroborate well the FTICR MS results reported by Khanna et al. (2008): the authors found very similar molecular space for fulvic acids extracted from Shilajit, which was located above AI = 0.5 line. This characterizes shilajit as very hydrophilic material dominated with carbohydrates and nitrogen-containing compounds. High content of nitrogen (up to 14%) is a remarkable feature of shilajit materials which is very different from N-depleted humic materials (up to 4%) as follows from comparison of the elemental compositions of the sample sets (Table S1).

3.3. Structure-antiviral activity relationships derived for the FTICR MS data on the HS samples used in this study

Of importance is that the first group of samples (all humic acids) is characterized with high anti-HIV activity, whereas the second group (all fulvic acids including shilajit) is characterized with low anti-HIV activity. For revealing molecular composition – anti HIV relationship, the relative intensities of the corresponding peaks in mass-spectra (Supplementary Info Table S5) were used for calculation of Spearman rank correlation with EC50. For better visualization, the data were plotted in van Krevelen diagram with color-coded correlation of molecular constituents (Fig. 6).

It was found that 339 and 201 formulae possessed strong negative
and positive correlation ($|r| > 0.7$ with $p < 0.05$) with EC$_{50}$, respectively. Of interest is that 251 out of 339 molecular components with the strong negative correlation of EC$_{50}$ were CHNO compositions. It can be deduced that the inhibitory activity is mostly related to the presence of relatively reduced aromatic compounds with H/C $< 1$ and O/C $< 0.5$. This is in agreement with the close correlation between the content of aromatic moieties and EC$_{50}$ as it was revealed by the NMR data.

In search of potential structural candidates for the most active components in the humic molecular ensemble, we have mined the data for 339 highly correlated formulae (Spearman’s rank correlation coefficient $\leq -0.7$) in ChEMBL database. We found 792 structures and 7271 activity entries corresponding to the 112 formulae (Supplementary Info Tables S7) and 642 out of which contained N atom and 65 – S atom. The assays entries (55) and the corresponding activity entries (76) related to anti-HIV activity measurement were also retrieved (Supplementary file Table S6). Almost half of the assays included inhibition of the specific HIV protein (mostly, integrase), while the other half was presented by the assessment of antiviral activity in the cell-based assays. Several examples of the retrieved active structures, which are the derivatives of naturally occurring compounds, and corresponding activities, are presented in Table 2. The full data are shown in Table S6 in the Supplementary Material.

Molecular formulae may correspond to a number of isomers. That is why the structural space of the active components is significantly broader as compared to molecular composition space. For deeper analysis of structures suggested for the HS components, the molecular network was plotted. ChEMBL data mining showed the presence of distinct clusters on the structural space network (Fig. 7). These clusters

Fig. 4. The matrix of Spearman correlation coefficients for the EC50 values and the content of carbon in the different chemical environments as measured by $^{13}$C NMR spectroscopy: A) for the humic materials used in this study ($n = 15$); B) for the shilajit materials ($n = 19$).
are produced by the structures with high chemical similarity – with the values of fingerprints T-score exceeding 0.8. In the most cases, the formulae assigned with the similar structures had the similar relationship to the EC50 values.

The largest cluster (19 nodes) is related to ether derivatives of 4′-O-demethylepipodophyllotoxins active against human DNA topoisomerase II (Zhou et al., 1991). Among the other largest clusters produced by the formulae, which are highly correlated to activity of the HS samples, there is the cluster of terpenoids from the Rhizome of Alisma orientale, tested against human Human Carboxylesterase 2 (Mai et al., 2015), and N-containing analogs of typical flavonoid – 3-O-arylmethylgalangin, active against Hepacivirus C reproduction (Lee et al., 2010). The potential structural candidates for the most active components in the humic molecular ensemble are shown in Table S6 in the Supplementary Material.

4. Conclusion

Assessment of anti-HIV activity and structure of two big sets of humic and shilajit materials allowed us to reveal the following trends. For the humic materials used in this study we have discovered two distinct trends in antiviral activity. The first one concerns much higher activity of the humic acids versus fulvic acids: HA > HMA > HFA > FA. The second trend concerns the source or origin: coal > peloid > peat. The both sequences follow the trend in lipophilicity index (a ratio of aromatic to aliphatic carbon) for these materials. This could be indicative of the leading role of the aromatic structures of HS in interaction with HIV. This assumption corroborates well the results of correlation analysis of the obtained EC50 values and structural composition: the closest relationships were observed for $C_{ar} (-0.82)$ and $CH_2O (+0.7)$. This indicates an increase in antiviral activity along with an increase in aromaticity of HS and a decrease along with a contribution of carbohydrate moieties. The obtained results are consistent with our previous

Fig. 5. Van Krevelen diagrams constructed on FTICR MS data of the 9 humic materials and 1 shilajit sample under study. Colors correspond to aromaticity of molecules: black – condensed aromatics ($AI_{mod}$≥0.67); blue – aromatics ($AI_{mod}$≥0.5); red – non-aromatics ($AI_{mod}$<0.5). (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

Fig. 6. Van Krevelen diagram for all formulae present in the set of 9 HS sample under study with color-coded Spearman correlation coefficient with EC50 values. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)
Table 2
Examples of ChEMBL compounds with anti-HIV activity corresponding to the FTICR MS formulae found in the humic materials used in this study.

| Formula     | Structure | $EC_{50}$, $\mu$M | assay description                                      | Ref.                         |
|-------------|-----------|--------------------|-------------------------------------------------------|-----------------------------|
| $C_{24}H_{28}O_6$ | ![Structure](image1.png) | 1.3 | Inhibition of HIV-1 induced CPE in CEM-SS cells       | Kashman et al. (1993)       |
| $C_{22}H_{14}N_2O_7$ | ![Structure](image2.png) | 0.1 | Inhibition of HIV-1 induced CPE in a cell-based assay | Yeo et al. (2005)           |
| $C_{27}H_{21}NO_{10}$ | ![Structure](image3.png) | 12.5 | Inhibition of HIV-1 integrase by electrochemiluminescent-based high-throughput strand transfer assay | Marchand et al. (2008)       |
|             |           | 10.1 | Inhibition of HIV-1 RNase H by FRET high-throughput assay | Marchand et al. (2008)       |

Fig. 7. Network-based representation of ChEMBL chemical space corresponding to the formulae found in the HS samples, which were highly correlated (Spearman’s rank correlation $\leq -0.7$) with $EC_{50}$. A network node represents chemical structure and it is labeled with the corresponding formula. An edge is drawn between a pair of nodes if the Tanimoto similarity index value was not lower than 0.8. Nodes and edges are colored by Spearman’s rank correlation value. Structures presented in Table 2 are shown in green. Clusters of structures discussed in the text are highlighted by orange and their representative structures are shown. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)
findings (Zhernov et al., 2017). The low anti-HIV activity and inverse structural trends found for the shilajit samples point out to different mode of action of shilajit materials as compared to humic substances. This is consistent with very different structural and molecular composition of shilajit as compared to humic materials.

Given the data of our reported investigations (Zhernov et al., 2017), we might suggest that the hydrophobic humic materials prevent fusion and inhibit reverse transcriptase activity of HIV-1, as it was experimentally evidenced in attachment and reverse transcriptase inhibition (RTI) assays conducted in our previous studies. These studies showed that humic materials inhibited HIV-1 reverse transcriptase in a dose-dependent manner similar to AZT. The obtained values of RTI activities were supported by Russian Science Foundation [grant number 19-75-00092].

The studied hydrophobic humic materials lost 50% of their efficacy at 7–8 h post-infection (h.p.i.), which is very close to AZT (7.7 ± 0.2). This also indirectly confirms the presence of RTI-activity in hydrophobic humic materials. With regard to shilajit, we might suggest that the nitrogen-rich humic molecules could interfere with peptide-based attachment and act as nitrogen-fusion inhibitors. This assumption can be indirectly confirmed by the data of the same TOA assay on the nitrogen-rich peloid FA: their profile was similar to fusion inhibitor T20 with a 50% loss of efficacy reached even at 5 ± 0.3 h.p.i. (Zhernov et al., 2017).

In general, the obtained results demonstrate that humic materials represent a rich source of highly potent and polytargeted natural compounds which might be used for prevention and treatment of viral infections. The most promising and straightforward way of humic materials use in clinical practice can be development of anti-HIV-infections. The most promising and straightforward way of humic materials which might be used for prevention and treatment of viral infections. The most promising and straightforward way of humic materials which might be used for prevention and treatment of viral infections.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

Shilajit samples Sh29 and Sh30 were kindly provided by Prof. G.N. Pilipenko (Russian State University for Geological Prospecting). This work was partially supported by the Russian Science Foundation (project 20-63-47070) in the part of structural characterization of the humic materials used in this study, FTICR MS data-mining and chemoinformatics were supported by Russian Science Foundation [grant number 19-75-00092].

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.envres.2020.110312.

Contributors

Y.Z., EK and IP conceived and designed the study. AK, MS and IP provided the HS and shilajit samples. YZ and GK performed the anti-HIV activity assessment. AK performed the 13C NMR characterization and statistical calculations. AZ, AO, and EN performed the FTICR MS and 13C NMR characterization, data-mining and chemometrics of the samples. EN, EK and IP supervised the study. AK, AZ, AO, MS, GK and YZ drafted the manuscript. All authors read and approved the final manuscript.

Funding

This work was partially supported by the Russian Science Foundation (project 20-63-47070) in the part of structural characterization of the humic materials used in this study. FTICR MS data-mining and chemometrics were supported by Russian Science Foundation [grant number 19-75-00092].

References

Agarwal, S.P., Khanna, R., Karmarkar, B., Anwer, M.K., Khar, R.K., 2007. Shilajit: a review. Phyther. Res. 21, 401–405. https://doi.org/10.1002/ptr.2100.

Andrae-Marobela, K., Ghislain, F.W., Okatch, H., Majindia, R.R.T., 2013. Polyphenols: a diverse class of multi-target anti-HIV-1 agents. Curr. Drug Metabol. 14, 392–413. https://doi.org/10.2174/138920121399900095.

Botes, M.E., Dekker, J., van Rensburg, C.E.J., 2002. Phase I trial with oral Oxihumate in HIV-infected patients. Drug Dev. Res. 57, 34–39. https://doi.org/10.1002/ddr.10117.

Brucoleri, A., 2013. Positional adaptability in the design of mutation-resistant nonnucleoside HIV-1 reverse transcriptase inhibitors: a supramolecular perspective. AIDS Res. Hum. Retrovir. 29, 4–12. https://doi.org/10.1089/AID.2012.0141.

Cagno, V., Donaislio, M., Civra, A., Cagliero, C., Rubiolo, P., Lembo, D., 2015. In vitro evaluation of the antiviral properties of Shilajit and investigation of its mechanisms of action. J. Ethnopharmacol. 166, 129–134. https://doi.org/10.1016/j.jep.2015.03.019.

Daglia, M., 2012. Polyphenols as antimicrobial agents. Curr. Opin. Biotechnol. 23, 174–181. https://doi.org/10.1016/j.copbio.2011.08.007.

Dong, E., Du, H., Gardner, L., 2020. An interactive web-based dashboard to track COVID-19 in real time. Lancet Inf. Dis. 20, 533–534. https://doi.org/10.1016/S1473-3099(20)30120-1.

Dong, G., Peng, C., Luo, J., Wang, C., Han, L., Wu, B., Ji, G., He, H., 2015. Adamantane-resistant influenza A viruses in the world (1902-2013): frequency and distribution of M2 gene mutations. PloS One 10, e0119115. https://doi.org/10.1371/journal. pone.0119115.

Fedoros, E.I., Orlov, A.A., Zhrebekker, A., Babureva, E.A., Maydin, M.A., Konstantinov, A. I., Krasnov, K.A., Karapetian, R.N., Izotova, E.I., Pigarev, S.E., Panchenko, A.V., Tyndyk, M.L., O losdikin, D.I., Nikolaev, E.V., Perminova, I.V., Antinov, V.N., 2018. Novel water-soluble lignin derivative BP-Cx-1: identification of components and screening of potential targets in silico and in vitro. Oncotarget 9, 18578–18593. https://doi.org/10.18632/oncotarget.24990.

Ford, N., Vitoria, M., Ranganj, A., Norris, S.L., Calmy, A., Doherty, M., 2020. Systematic review of the efficacy and safety of antiretroviral drugs against SARS, MERS or COVID-19: initial assessment. J. Int. AIDS Soc. 23, e25489 https://doi.org/10.1002/jia2.25489.

Helbig, R., Köckingen, B., Rutzler, P., 1997. Anti-herpes simplex virus type 1 activity of humic acid-like polymers and their o-diphenolic starting compounds. Antivir. Chem. Chemother. 8, 265–273. https://doi.org/10.1089/vir.1996.20790000310.

Hertkorn, N., Permin, A., Perminova, I., Kovalevskii, D., Yudov, M., Petrov, Y., Kemptrup, A., 2002. Comparative analysis of partial structures of a peat humic and humic acid-like polymers and their o-diphenolic starting compounds. Antivir. Chem. Chemther. 8, 265–273. https://doi.org/10.1089/vir.1996.20790000310.

Jeenings, R.E., Keulgen, W., Boucher, C., Sanders, R.W., Berkbooth, B., 2001. Evolution of AZT resistance in HIV-1: the 41-70 intermediate that is not observed in vivo has a replication defect. Virology 283, 294–305. https://doi.org/10.1006/viro.2001.0888.

Jooe, G.K., Dekker, J., Van Rensburg, C.E.J., 2003. Investigation of the immunomodulatory properties of oxihumate. Zeitschrift für Naturforsch. - Sect. C. J. Biosci. 58, 263–267. https://doi.org/10.1515/znc-2003-3-421.

Karamov, E., Efremyan, K., Sinivin, A., Zhernov, Y., Cuevas, M.T., Delgado, E., Sánchez-Martínez, M., Carrera, C., Kornilasova, G., Turgev, A., Bacquet, J., Pérez-Alvarez, L., Thomson, M.M., 2018. HIV-1 genetic diversity in recently diagnosed infections in Moscow: predominance of A FSU , frequent branching in clusters, and circulation of the Iberian subtype G variant. AIDS Res. Hum. Retrovir. 34, 629–634. https://doi.org/10.1089/AID.2018.0052.

Kabashman, Y., Gustafson, K.R., Fuller, R.W., Cardellina, J.H., McMahon, J.B., Currens, M. J., Buckheit, R.W., Hughes, S.H., Cragg, G.M., Boyd, M.R., 1993. HIV inhibitory natural products. Part 7. The calanoides, a novel HIV-inhibitory class of coumarin derivatives from the tropical rainforest tree, Calophyllum lanigerum. [Erratum to document cited in CA 117(11)10801g]. J. Med. Chem. 36, 1110. https://doi.org/10.1021/jm00060a020.

Khranna, R., Witt, M., Anwer, M., Agarwal, S.P., Koch, B.P., 2008. Spectroscopic characterization of fulvic acids extracted from the rock exudate Shilajit. Org. Geochem. 39, 1719–1724.

Klöcking, R., Helbig, B., Schott, G., Schack, M., Wutzler, P., 2002. Anti-HSV-1 activity of synthetic humic acid-like polymers derived from o-diphenolic starting com-

Environmental Research 193 (2021) 110312

https://doi.org/10.1016/j.envres.2020.110312
