

Article

Identification of Antiprotozoal Compounds from *Buxus sempervirens* L. by PLS-Prediction

Lara U. Szabó ¹, Marcel Kaiser ^{2,3}, Pascal Mäser ^{2,3} and Thomas J. Schmidt ^{1,*}

¹ Institute of Pharmaceutical Biology and Phytochemistry (IPBP), PharmaCampus, University of Münster, Corrensstraße 48, D-48149 Münster, Germany; lszabo@uni-muenster.de

² Swiss Tropical and Public Health Institute (Swiss TPH), Socinstrasse 57, CH-4051 Basel, Switzerland; marcel.kaiser@unibas.ch (M.K.); pascal.maeser@swisstph.ch (P.M.)

³ University of Basel, Petersplatz 1, CH-4003 Basel, Switzerland

* Correspondence: thomschm@uni-muenster.de; Tel.: +49-251-83-33378

Abstract: Various *nor*-triterpene alkaloids of *Buxus* (*B.*) *sempervirens* L. have shown remarkable in vitro activity against the causative agents of tropical malaria and East African sleeping sickness. To identify further antiprotozoal compounds of this plant, 20 different fractions of *B. sempervirens* L., exhibiting a wide range of in vitro bioactivity, were analyzed by UHPLC / +ESI-QqTOF-MS/MS. The analytical profiles were investigated by partial least squares regression (PLS) for correlations between the intensity of LC/MS signals, bioactivity and cytotoxicity. The resulting models highlighted several compounds as mainly responsible for the antiprotozoal activity and thus, worthwhile for subsequent isolation. These compounds were dereplicated based on their mass spectra in comparison with isolated compounds recently reported by us and with literature data. Moreover, an estimation of the cytotoxicity of the highlighted compounds was derived from an additional PLS model in order to identify plant constituents with strong selectivity. In conclusion, high levels of antitrypanosomal and antiplasmodial activity were predicted for eight and four compounds, respectively. These include three hitherto unknown constituents of *B. sempervirens* L., presumably new natural products.

Keywords: *Buxus sempervirens* L.; *nor*-cycloartane alkaloids; antiprotozoal activity; multivariate data analysis; partial least squares regression; mass spectrometry; fragmentation pattern



Citation: Szabó, L.U.; Kaiser, M.; Mäser, P.; Schmidt, T.J. Identification of Antiprotozoal Compounds from *Buxus sempervirens* L. by PLS-Prediction. *Molecules* **2021**, *26*, 6181. <https://doi.org/10.3390/molecules26206181>

Academic Editors: Vera Muccilli, Monica Scognamiglio and Julien Wist

Received: 11 September 2021

Accepted: 8 October 2021

Published: 13 October 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Unicellular eukaryotic (“protozoan”) parasites cause enormous morbidity and mortality [1]. One of the most widespread and serious protozoan diseases is malaria.

Malaria is caused by protozoans of the genus *Plasmodium* (*P.*), which occur in erythrocytes and in liver parenchyma cells depending on their developmental stage. The disease is transmitted by insect vectors, female *Anopheles* mosquitos. Several human-pathogenic *Plasmodium* species (*P. falciparum*, *P. vivax*, *P. ovale*, *P. malariae*, *P. knowlesi* and *P. cynomolgi*) cause different clinical manifestations of the disease. The World Health Organization (WHO) estimated the annual malaria incidence at 229 million with a mortality of 409,000 cases in 2019 [2]. The human-pathogenic species *P. falciparum* (*Pf*) is responsible for the majority of deaths. The increasing development of resistance, especially of *Pf*, complicates the therapy and makes the development of new antimalarial drugs inevitable [3].

Protozoans of the genus *Trypanosoma* cause very severe infections considered neglected tropical diseases, namely Chagas disease in Latin America (caused by *T. cruzi*) and human African trypanosomiasis (HAT, “sleeping sickness”, caused by *T. brucei*) in sub-Saharan Africa. Although these infections are widespread and lead to severe and long-lasting illnesses, trypanosomiasis still do not receive sufficient attention in drug research and development. This is mainly because they affect the poorest part of the population in already low-income countries. *Trypanosoma brucei rhodesiense* (*Tbr*) is the causative agent of

East African sleeping sickness and is transmitted by the Tsetse fly (*Glossina* sp.) [4]. There are only limited therapeutic options for this medical condition, which are often associated with side effects [5,6]. Untreated, the disease has a lethality rate of 100%. Especially in view of the spread of vectors due to climate change and globalization, research into innovative and well-tolerated antiprotozoal drugs is highly relevant and future-oriented.

In our previous study, *nor*-triterpene alkaloids of the *nor*-cycloartane type from *B. semperivirens* L. displayed promising in vitro antiprotozoal activity and selectivity [7]. In the mentioned study, 25 alkaloids were isolated in a systematic manner from the aerial parts of *B. semperivirens* and tested for their antiprotozoal activity; several constituents with high activity against *Pf* and/or *Tbr* were isolated besides compounds with medium or low levels of activity. However, the total alkaloid fraction was found to be a highly complex mixture that might contain further alkaloids with promising activity that have not been isolated so far. Therefore, in order to identify possible further alkaloids with high antiprotozoal activity besides those already isolated, a multivariate data analysis (MVDA) approach was chosen to pinpoint the compounds responsible for the growth inhibition of the protozoa from a large amount of chromatographic/mass spectrometric data. Partial least squares regression (PLS) was used to correlate the complex UHPLC/+ESI-QqTOF-MS/MS (henceforth termed "LC/MS") data of 20 different alkaloid fractions from a *B. semperivirens* L. leaf extract with their antiprotozoal bioactivity data. PLS regression has already been used in previous work by our group as well as others to make accurate predictions of constituents with antitrypanosomal [8–10] or antiplasmodial [11] activity, respectively, in plant extracts and fractions thereof. In this work, we present PLS models that highlight, besides those already isolated, further alkaloids of *B. semperivirens* L. with high predicted activity against *Pf* and *Tbr*. Furthermore, a corresponding model was generated for cytotoxicity against mammalian cells, to obtain some information on selectivity.

2. Results and Discussion

2.1. Antiprotozoal Activity of Crude Extract and Fractions from *B. semperivirens* L.

The alkaloids (ALOF) of *B. semperivirens* L. leaves were enriched from crude dichloromethane extract (GBUS) by acid-base extraction and subsequently separated by CPC into 20 different fractions [12]. The samples were tested in vitro against *Pf* and *Tbr* as well as against mammalian control cells (L6 rat skeletal myoblasts) to evaluate the cytotoxicity (Cytotox.) of the fractions (Table 1).

The dichloromethane extract (GBUS) and the two subfractions LNB (lipophilic and neutral constituents) and ALOF (alkaloids) showed clear differences in their antiprotozoal activity in the in vitro assays (Table 1). The alkaloid fraction very clearly displayed the highest activity against both pathogens, *Pf* and *Tbr*, as well as considerable antiplasmodial selectivity with a selectivity index (SI) of 34. The antiplasmodial activity of the alkaloid fraction was most pronounced with an IC_{50} value of 0.99 $\mu\text{g}/\text{mL}$. The growth inhibition against *Tbr* with an IC_{50} value of 5 $\mu\text{g}/\text{mL}$ was weaker but still interesting. The strong increase of activity and selectivity of the alkaloid rich fraction (ALOF) in comparison with the crude extract and, particularly, with the lipophilic and neutral fraction (LNB) implied that the alkaloids are mainly responsible for the antiprotozoal activity of the dichloromethane extract of *B. semperivirens* L. Therefore, the alkaloid fraction was fractionated by CPC for further study.

The 20 CPC fractions displayed a wide range of activities (Table 1) in the antiprotozoal in vitro tests while the cytotoxicity against L6 rat-skeletal myoblasts, also tested as a marker for selectivity, was generally rather low and did not vary to a significant degree, with IC_{50} values ranging from 12 $\mu\text{g}/\text{mL}$ in case of fraction 4 to 55 $\mu\text{g}/\text{mL}$ in case of fraction 20.

Regarding activity against *Tbr*, fraction 2 was the most active with an IC_{50} value of 0.63 $\mu\text{g}/\text{mL}$. Fraction 3 and fraction 12 also exhibited IC_{50} values below 1 $\mu\text{g}/\text{mL}$. Fraction 18 was the least active against *Tbr* with an IC_{50} value of 39.3 $\mu\text{g}/\text{mL}$.

Several of the fractions (fractions 2, 3, 4, 5, 7–13) displayed a high level of antiplasmodial activity against *Pf* with IC_{50} values below 1 $\mu\text{g}/\text{mL}$. Fraction 2 was the most potent

fraction also in this case, with an IC_{50} value of $0.35 \mu\text{g/mL}$. In addition, this fraction had a high SI of 43 against *Pf*. For the other mentioned active fractions, the concentration range of antiplasmodial IC_{50} values was also distinctly below that of cytotoxicity (all $SI > 10$). This suggests a selective inhibitory effect of the *Buxus*-alkaloids, as opposed to an unspecific cytotoxic effect.

Table 1. In vitro antiprotozoal and cytotoxic activity of the crude dichloromethane (DCM) extract and fractions from *B. sempervirens* L. leaves.

Sample	<i>Pf</i>	<i>Tbr</i>	Cytotox.	SI <i>Pf</i>	SI <i>Tbr</i>
GBUS	7.9 ± 0.06	14.4 ± 2.7	56.4 ± 1.7	7	4
LNB	16 ± 6.1	15 ± 5	54 ± 4	3	4
ALOF	0.99 ± 0.03	5 ± 1.4	34 ± 11.7	34	7
F1	1.5 ± 0.2	1.98 ± 0.5	48	32	24
F2	0.35 ± 0.01	0.63 ± 0.03	15	43	24
F3	0.55 ± 0.04	0.92 ± 0.06	14	26	15
F4	0.92 ± 0.03	1.4 ± 0.3	12	13	9
F5	0.99 ± 0.004	1.4 ± 0.04	13	13	9
F6	1.5 ± 0.4	1.7 ± 0.7	14	9	8
F7	0.75 ± 0.04	1.8 ± 0.6	15	20	8
F8	0.76 ± 0.07	1 ± 0.3	16	21	16
F9	$0.62 \pm 0.35^*$	2.4 ± 1.41	15	24	6
F10	0.95 ± 0.19	9.8 ± 5.3	39	41	4
F11	0.65 ± 0.25	1.8 ± 0.72	15	23	8
F12	$0.5 \pm 0.23^*$	0.8 ± 0.04	15	30	19
F13	$0.49 \pm 0.21^*$	2.6 ± 0.9	15	31	6
F14	$1.18 \pm 0.54^*$	12 ± 1	41	35	3
F15	1.0 ± 0.26	12 ± 0.1	46	46	4
F16	1.82 ± 0.53	2.7 ± 0.2	37	20	14
F17	3.0 ± 0.5	11 ± 1	44	15	4
F18	7.11 ± 1.8	39.3 ± 0.7	48	7	1
F19	2.67 ± 0.2	14.4 ± 1.7	47	18	3
F20	1.9 ± 0.7	13.5 ± 1	55	29	4
Chloroquine	0.003 ± 0.001			3	
Melarsoprol		0.004 ± 0.001			3
Podophyllotoxin			0.01 ± 0.004		

GBUS: crude dichloromethane extract of *B. sempervirens* L. leaves; LNB: lipophilic and neutral fraction; ALOF: alkaloid fraction; SI: selectivity index. All IC_{50} values are expressed in $\mu\text{g/mL}$. While $n = 3$ was reported as the mean value from three independent measurements with the standard deviation, the other values were determined with $n = 2$ as the mean value from two independent measurements with the fluctuation range and $n = 1$ for the cytotoxicity data.

2.2. PLS Modeling

Correlations between analytical fingerprint data of complex mixtures, such as extracts and fractions and their bioactivity data can be investigated by PLS regression and the resulting models used to identify constituents mainly responsible for the biological effect under study [8–10]. In order to apply such an approach to the present set of fractions, they were all analyzed by LC/MS (Figure S1, Supplementary Materials) and the fingerprint data were processed using the software DataAnalysis 4.1 and ProfileAnalysis 2.1 (Bruker Daltonik GmbH, Bremen, Germany). The resulting data matrix (“bucket table”) of variables representing MS signal intensity within defined retention time and m/z value intervals consisted of $1420 [m/z: t_R]$ variables \times 20 analyses. These data constituted the X-matrix (independent variables) while the respective negative logarithmic IC_{50} values (pIC_{50}) of the fractions were treated as the Y-matrix (dependent variables) to generate PLS regression models in SIMCA 16.0.1 (Sartorius Stedim Data Analytics AB, Umeå, Sweden). Three different PLS models were calculated, namely, for the activity against *Pf* and *Tbr* as well as for cytotoxic activity against L6 rat skeletal myoblasts used as mammalian control cells to assess selectivity. High values for R^2 (coefficient of determination of values predicted during model calibration) and, particularly, for Q^2 (coefficient of determination of values

predicted in leave-one-out cross validation) obtained with few PLS components in each case indicated that stable models with good internal predictive power had been generated (Table 2).

Table 2. Statistical characteristics of the three different PLS models.

PLS Model	PLS Components	R ²	Q ² *
<i>Pf</i>	2	0.9	0.82
<i>Tbr</i>	4	0.98	0.89
Cytotox.	3	0.96	0.83

* leave-one-out cross validation.

2.2.1. PLS-Prediction of Antiplasmodial Compounds from *B. semperivirens* L.

The PLS model for antiplasmodial activity consisted of two PLS components and yielded R² and Q² values of 0.9 and 0.82, respectively (Table 2). Scores and loadings plots are presented in Figure 1. In the scores plot of the PLS, the different fractions from *B. semperivirens* L. are grouped along their PLS components, which form the axes of the coordinate system. The influence of individual X variables (“buckets”) on the samples’ positions in the scores plot is reflected in the loadings plot. For instance, if a fraction is located in the upper right quadrant of the scores plot (high values on both, the first and the second PLS component), the buckets responsible for the position of this fraction are found in this area of the loadings plot. At the same time, these variables are closest to the position of the Y variable, which means that they are most strongly correlated with activity.

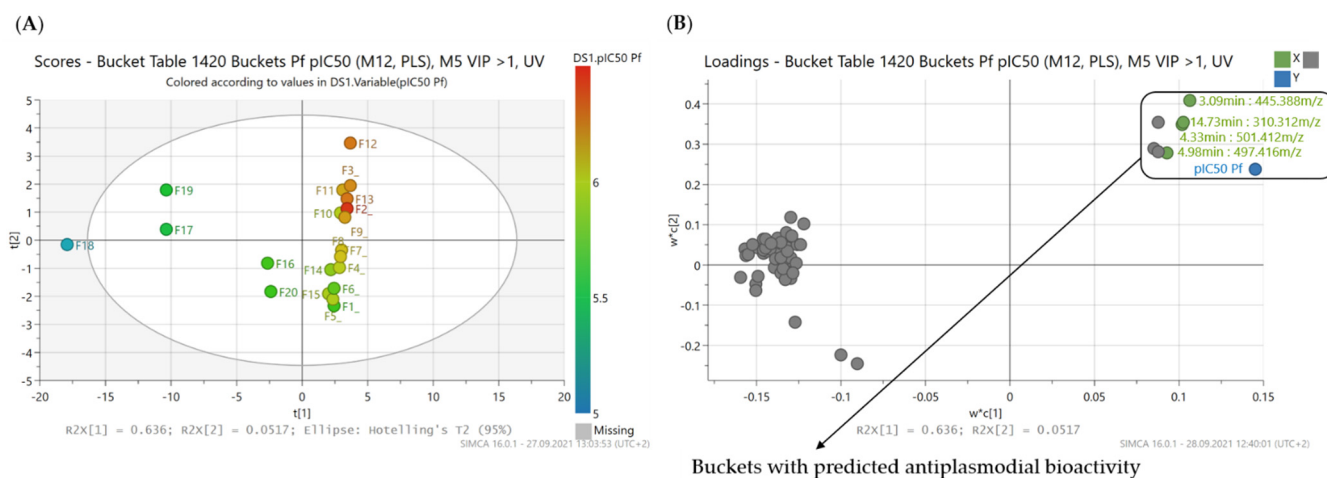


Figure 1. PLS model with prediction of antiplasmodial compounds from *B. semperivirens* L. (A) Scores plot of the second vs. the first PLS component of the PLS model. Coloration indicates the bioactivity against *Pf* of the fractions ranging from a pIC₅₀ value of 5.15 (F18, blue) to a value of 6.46 (F2, red); (B) corresponding loadings plot to (A). The buckets marked with a box in the upper right quadrant are predicted as compounds with high antiplasmodial activity by the PLS model (Buckets marked and labelled in green are detailed in Figure 2).

2.2.2. Identification of the Antiplasmodial Compounds Highlighted by the PLS Model

In the loadings plot, four bucket variables are emphasized as mainly important for the antiplasmodial activity of the alkaloid fractions (Figure 2). In order to assign and identify these compounds, the chromatographic and mass spectral data of the LC/MS measurements were evaluated (Table 3). The +ESI-QqTOF MS/MS spectra corresponding to the highlighted variables were examined in particular for the occurrence of prominent core fragments, which initially provide information about the substitution present and the occurrence of double bonds in the structures of the compounds. In previous studies, characteristic core fragments of Buxaceae alkaloids have already been identified for the amino-/amidosteroids in *Pachysandra* [8] and *Sarcococca* [13]. For the *nor*-triterpene alka-

oids of the *nor*-cycloartane type, occurring in the genus *Buxus*, a few prominent fragments were also discovered, which can be used for structural assignment [14]. These characteristic diagnostic fragments were supplemented in the present work by further ones as shown in Figure 3.

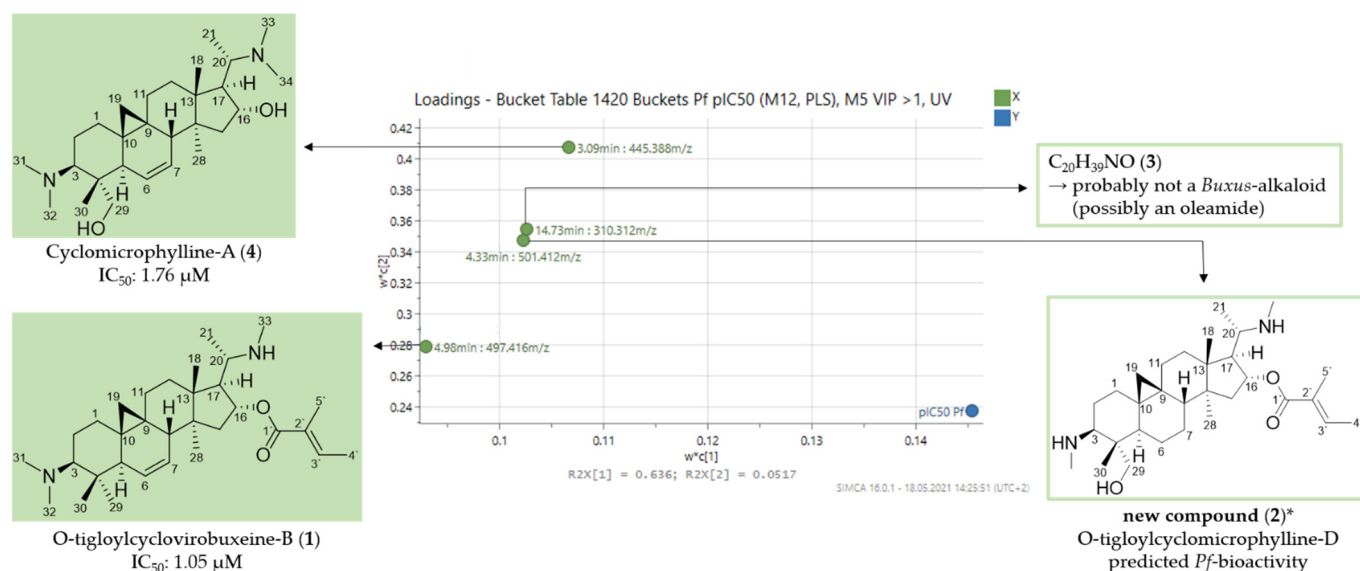


Figure 2. Expansion of region of the loadings plot (Figure 1) with the buckets that, according to the PLS model, have a high impact on the antiplasmodial activity of *B. semperivirens* L. * Structural assignment based only on mass spectral data; stereochemistry tentatively assumed in analogy with known compounds [7].

Table 3. LC/MS-characteristics of the compounds with predicted antiplasmodial activity (compare Figure 2).

Bucket	Adduct Ions	Structural Formula (DBE)	Core Fragment(s) <i>m/z</i>	Identified <i>Buxus</i> -Alkaloid
4.98 min: 497.416 <i>m/z</i> (1)	$[M + H]^+ < [M + 2H]^{2+}$	C ₃₂ H ₅₂ N ₂ O ₂ (8)	321; 295	O-tigloylcyclovirobuxeine-B
4.33 min: 501.412 <i>m/z</i> (2)	$[M + H]^+ < [M + 2H]^{2+}$	C ₃₁ H ₅₂ N ₂ O ₃ (7)	323	new compound
14.73 min: 310.312 <i>m/z</i> (3)	$[M + H]^+$	C ₂₀ H ₃₉ NO (2)	n.i.	n.i.; possibly an oleamide structure
3.09 min: 445.388 <i>m/z</i> (4)	$[M + H]^+ < [M + 2H]^{2+}$	C ₂₈ H ₄₈ N ₂ O ₂ (6)	321; 295	cyclomicrophylline-A

DBE: double bond equivalent; n.i.: not identified.

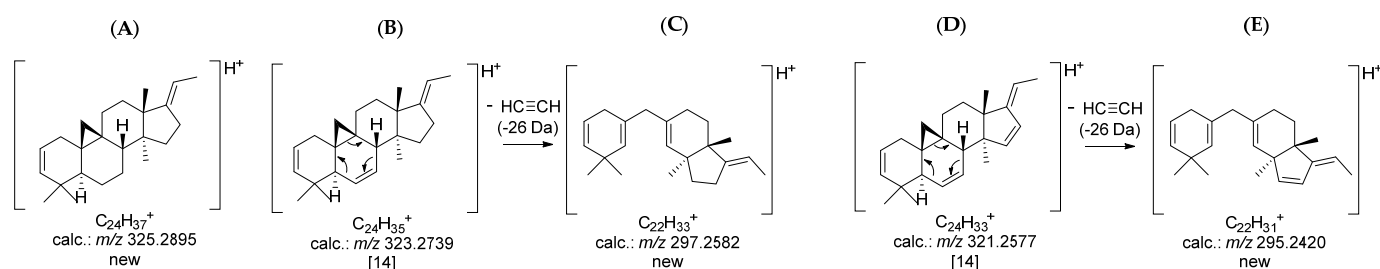


Figure 3. Prominent core fragments of the investigated *nor*-triterpene alkaloids of the *nor*-cycloartane type from *B. semperivirens* L. (A) core fragment *m/z* 325 for fully saturated C-3/C-20 diamines; (B) core fragment *m/z* 323 for C-3/C-20 diamines possessing one additional substitution or double bond; (C) core fragment *m/z* 297 for C-3/C-20 diamines with a Δ^6 ; (D) core fragment *m/z* 321 for C-3/C-20 diamines doubly substituted or unsaturated; (E) core fragment *m/z* 295 for C-3/C-20 diamines with a Δ^6 and one further substitution or unsaturation. (The postulated positions of unsaturation are based on the already known structures of compounds 1 and 4).

Based on the MS chromatograms and the fragmentation pattern, as well as by identity of the retention times with samples of the authentic, isolated alkaloids [7], com-

pounds **1** (4.98 min: 497.416 m/z) (Figures S2 and S3, Supplementary Materials) and **4** (3.09 min: 445.388 m/z) (Figures S9–S11, Supplementary Materials) were identified as *O*-tigloylcyclovirobuxeine-B (**1**) and cyclomicrophylline-A (**4**), respectively, both isolated in our previous study and proven to be highly active against *Pf* in the in vitro assay (IC_{50} : 1.05 μ M (**1**); IC_{50} : 1.76 μ M (**4**)) [7]. The finding of these two compounds with already proven antiplasmodial activity can be taken as evidence for the validity of the calculated PLS model.

The occurrence of a core fragment at m/z 323 allowed assignment of compound **2** (4.33 min: 501.412 m/z) as a C-3/C-20 diamine with an additional single substitution or unsaturation in the core skeleton (Figure 3). The fragment m/z 401 indicated the neutral loss of a C_5 -acid moiety (100 Da; a tiglic acid ester is assumed in analogy to compound **1** from the $[M + H]^+$ ion of **2**). The presence of two monomethylated amino groups at C-3 and C-20 was evident from the dual neutral losses of 31 Da (m/z 385 \rightarrow 354 \rightarrow 323) (Figures S4 and S5, Supplementary Materials). The detailed fragmentation pathway is reported in Figure S6, Supplementary Materials. The structure, thus determined from the mass spectral data (Figure 2), could not be found in the literature, so that we assume, to the best of our knowledge, that compound **2** is a new natural product. It should be noted that the stereochemistry of **2**, which is not accessible on grounds of the mass spectra alone, is assumed in analogy with related compounds of known configuration (e.g., reported in [7]). The generic name of the new compound, *O*-tigloylcyclomicrophylline-D, was chosen based on the existing classification for *Buxus*-alkaloids [15,16].

Compound **3** (14.73 min: 310.312 m/z) displayed a completely different chromatographic and mass spectral behavior from the *Buxus*-alkaloids (+ESI-QqTOF MS/MS spectra see Figures S7 and S8, Supplementary Materials). Furthermore, the molecular formula $C_{20}H_{39}NO$ with only two double bond equivalents is not compatible with a triterpenoid alkaloid and does not match any other known constituent from *B. sempervirens* L. According to the data, it could be an oleamide structure, but this was not investigated further in this study. The identification and clarification of this compound's contribution to the antiplasmodial activity must await its isolation in subsequent studies.

2.2.3. PLS-Prediction of Antitrypanosomal Compounds from *B. sempervirens* L.

The PLS model for antitrypanosomal activity is composed of four PLS components and shows good statistical performance with R^2 and Q^2 values of 0.98 and 0.89, respectively (Table 2). The resulting scores and loadings plots are depicted in Figure 4.

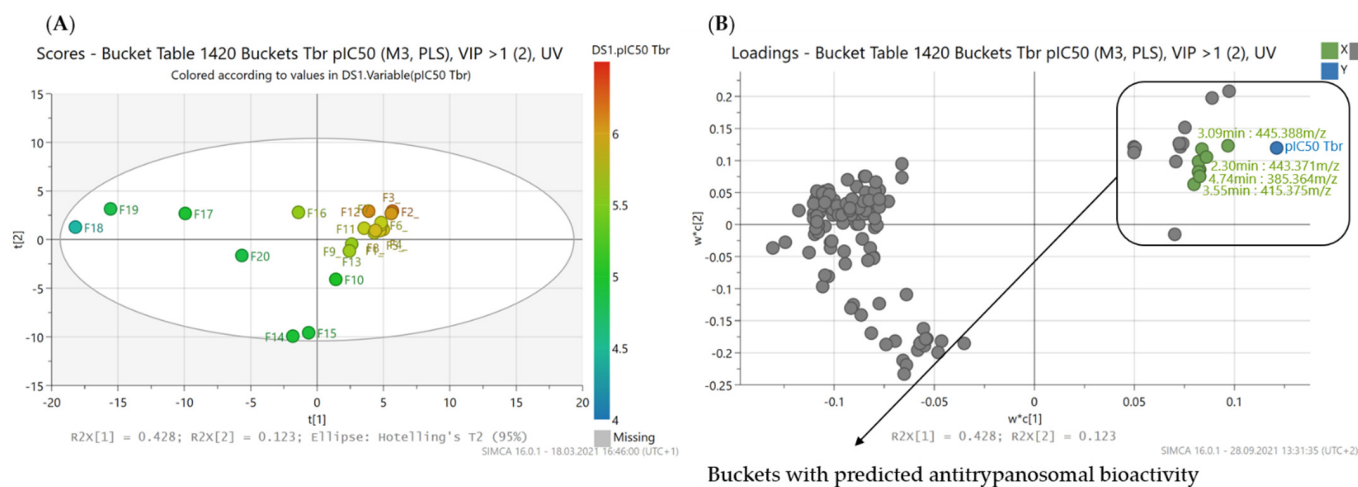


Figure 4. PLS model with prediction of antitrypanosomal compounds from *B. sempervirens* L. (A) Scores plot of the second vs. the first PLS component of the PLS model. Coloration indicates the bioactivity against *Tbr* of the fractions ranging from a pIC_{50} value of 4.41 (F18, blue) to a value of 6.2 (F2, red); (B) corresponding loadings plot to (A). The buckets marked with a box in the upper right quadrant are predicted to possess high antitrypanosomal activity by the PLS model (Buckets marked and labeled in green are detailed in Figure 5).

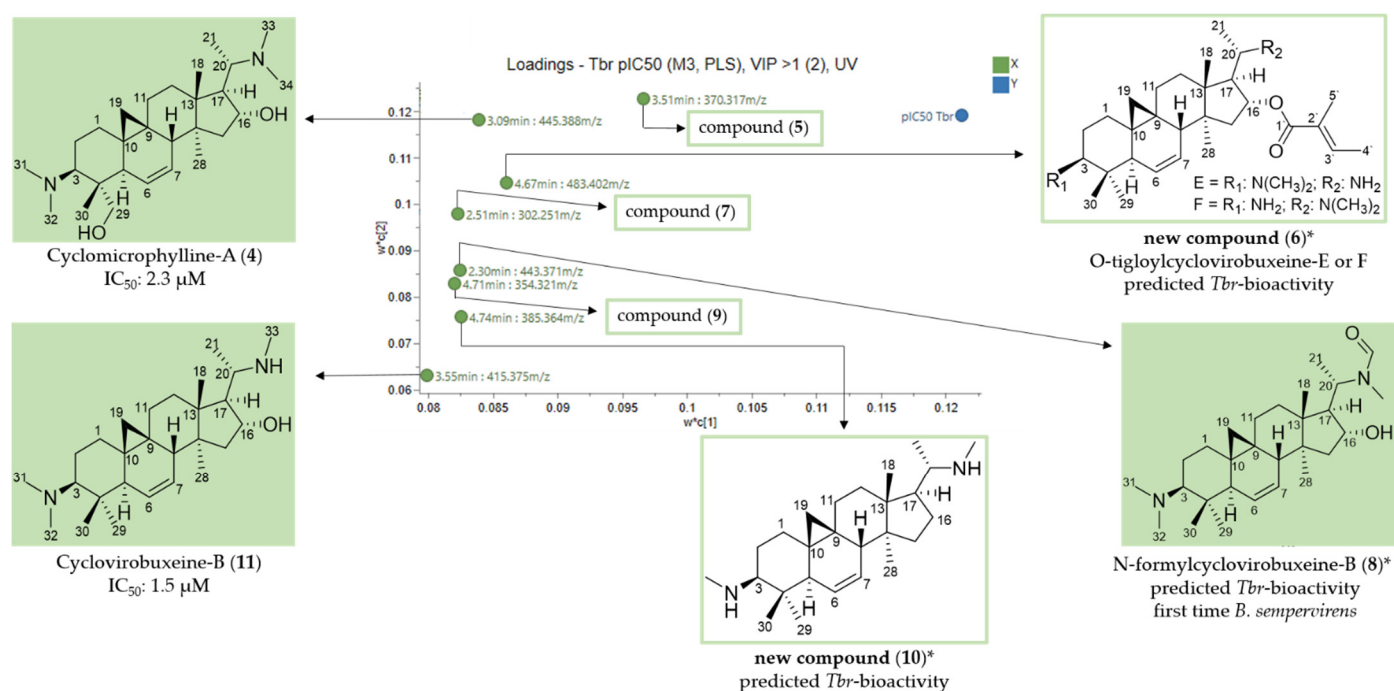


Figure 5. Expansion of the region of the loadings plot (Figure 4) with the buckets that, according to the PLS model, have a high impact on the antitrypanosomal activity of *B. semperivirens* L. * Assignment based only on mass spectral data; stereo-chemistry tentatively assumed in analogy with known compounds [7].

2.2.4. Identification of the Antitrypanosomal Compounds Highlighted by the PLS Model

In the loadings plot of the PLS model, eight buckets were located in the upper right quadrant and thus predicted to be mainly responsible for the antitrypanosomal activity (Figure 5). The data of the LC/MS analysis of the underlying compounds are listed in Table 4.

Table 4. LC/MS-characteristics of the compounds with predicted antitrypanosomal activity (compare Figure 5).

Bucket	Adduct Ions	Structural Formula (DBE)	Core Fragment(s) m/z	Identified <i>Buxus</i> -Alkaloid
3.51 min: 370.317 m/z (5)	$[M + H]^+$	C ₂₅ H ₃₉ NO (7)	n.i.	n.i.
3.09 min: 445.388 m/z (4)	$[M + H]^+ < [M + 2H]^{2+}$	C ₂₈ H ₄₈ N ₂ O ₂ (6)	321; 295	cyclocymophylline-A
4.67 min: 483.402 m/z (6)	$[M + H]^+ < [M + 2H]^{2+}$	C ₃₁ H ₅₀ N ₂ O ₂ (8)	321; 295	new compound
2.51 min: 302.251 m/z (7)	$[M + H]^+ < [M + 2H]^{2+}$	C ₂₀ H ₃₁ NO (6)	n.i.	n.i.
2.30 min: 443.371 m/z (8)	$[M + H]^+ < [M + 2H]^{2+}$	C ₂₈ H ₄₆ N ₂ O ₂ (7)	321; 295	<i>N</i> -formylcyclovirobuxeine-B
4.71 min: 354.321 m/z (9)	$[M + H]^+ < [M + 2H]^{2+}$	C ₂₅ H ₃₉ N (7)	n.i.	n.i.
4.74 min: 385.364 m/z (10)	$[M + H]^+ < [M + 2H]^{2+}$	C ₂₆ H ₄₄ N ₂ (6)	323; 297	new compound
3.55 min: 415.375 m/z (11)	$[M + H]^+ < [M + 2H]^{2+}$	C ₂₇ H ₄₆ N ₂ O (6)	321; 295	cyclovirobuxeine-B

DBE: double bond equivalent; n.i.: not identified.

Cyclocymophylline-A (4) (3.09 min: 445.388 m/z) was also found as an active compound in the model against *Pf* (see Section 2.2.2 above, Figure 2, Table 3) and had already demonstrated activity in the *in vitro* assay against *Tbr* with an IC_{50} value of 2.3 μM [7]. In addition, compound 11 (3.55 min: 415.375 m/z) could be assigned (Figures S28–S30, Supplementary Materials) to cyclovirobuxeine-B, which had also been isolated and displayed promising antitrypanosomal activity (IC_{50} : 1.5 μM) in our previous study [7]. These findings, as in the case of the antiplasmodial model above, corroborate the validity of the calculated model.

Compound 8 (2.30 min: 443.371 m/z) exhibited analytical data analogous to compound 11 (Figures S20–S22, Supplementary Materials). The structural formula C₂₈H₄₆N₂O₂

derived from the $[M + H]^+$ ion (m/z 443) differed in the occurrence of additional oxygen and carbon atoms in comparison to **11**. According to these mass spectral data compound **8** could be assigned to the known *Buxus*-alkaloid *N*-formylcyclovirobuxeine-B. This alkaloid has already been isolated from *Buxus malayana* Ridl. [17]. Antitrypanosomal activity of this compound has not been reported so far.

Compound **6** (4.67 min: 483.402 m/z) showed a characteristic fragmentation pattern (Figures S14–S17, Supplementary Materials) almost identical to compound **1**. In contrast to *O*-tigloylcyclovirobuxeine-B (**1**), the presence of a primary amino group could be concluded in the structure of **6** from the neutral loss of NH_3 observed in the fragment at m/z 466 $[483-NH_3]^+$. The detailed fragmentation pathway is reported in Figure S17, Supplementary Materials. Based on the fragmentation, it could not be determined whether the primary amino group is present at position 3 or 20. To the best of our knowledge, neither structure has been described in the literature so that **6** certainly represents a new natural product. In accordance with the systematic naming of such alkaloids [15,16], this compound must be either *O*-tigloylcyclovirobuxeine-E or F (E = R₁: N(CH₃)₂; R₂: NH₂; F = R₁: NH₂; R₂: N(CH₃)₂).

The core fragments m/z 323 and 297 in the +ESI-QqTOF MS/MS spectrum of compound **10** (4.74 min: 385.364 m/z) (Figures S25–S27, Supplementary Materials) indicated a C-3/C-20 diamine with a double bond between C-6 and C-7 (Figure 3). The fragments m/z 354 $[M + H-CH_3NH_2]^+$ and 323 $[354-CH_3NH_2]^+$ indicated a substitution with two monomethylamino groups. The resulting structure (Figure 5) is, to the best of our knowledge, so far unknown.

Compounds **5** (3.51 min: 370.317 m/z) (Figures S12 and S13, Supplementary Materials), **7** (2.51 min: 302.251 m/z) (Figures S18 and S19, Supplementary Materials) and **9** (4.71 min: 354.321 m/z) (Figures S23 and S24, Supplementary Materials) did not display the characteristic mass spectral behavior of typical *Buxus*-alkaloids so that they could not be structurally elucidated based on their fragmentation. Their identification, hence, as in the case of **3**, must await further studies.

2.2.5. PLS-Prediction of Cytotoxic Compounds from *B. sempervirens* L.

Although the cytotoxic activity of the alkaloid fractions with good activity against the protozoan parasites was usually rather low and their selectivity indices high, especially in the case of antiplasmodial activity, it was nevertheless interesting to create a PLS model for the cytotoxicity data of the alkaloid fractions. In this case, a model consisting of three PLS components resulted, which yielded good statistical quality with an R^2 of 0.96 and a Q^2 of 0.83 (Table 2). The scores and loadings plots (Figure 6) showed the distribution of fractions and buckets according to their effect on cytotoxicity, respectively.

Of the 11 compounds with predicted antiprotozoal activity, the cytotoxicity model included compounds **7** (2.51 min: 302.251 m/z), **9** (4.71 min: 354.321 m/z), **10** (4.74 min: 385.364 m/z) and **11** (3.55 min: 415.375 m/z), which were identified as the four compounds with the greatest impact on cytotoxicity (Figure 7).

This might suggest that the antitrypanosomal activity of these four compounds could be due to an unspecific cytotoxic effect rather than a selective impact on the parasites. However, the cytotoxicity of compound **11** (3.55 min: 415.375 m/z) against the same L6 cell line was already tested in our previous study [7] and it was much lower than the antitrypanosomal activity (IC₅₀ value against *Tbr* of 1.5 μ M vs. 35.5 μ M for cytotoxicity against L6 cells; SI: 24). With some caution, this result may probably be extrapolated to the other three compounds dominating the PLS model for cytotoxicity. Thus, it can be concluded that even if these four ingredients have a significant influence on the cytotoxicity of the boxwood leaves, they may still possess a selective antiprotozoal effect since the overall cytotoxicity is rather low.

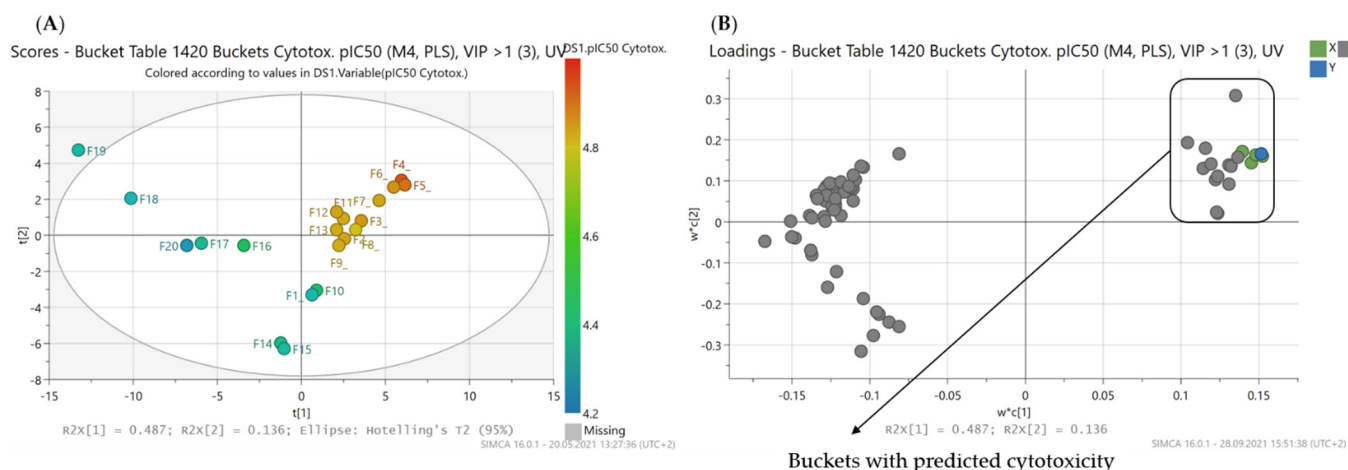


Figure 6. PLS model with a prediction of cytotoxic compounds from *B. sempervirens* L. (A) Scores plot of the second vs. the first PLS component of the PLS model. Coloration indicates the cytotoxicity of the fractions ranging from a pIC_{50} value of 4.26 (F20, blue) to a value of 4.94 (F4, red); (B) corresponding loadings plot to (A). The buckets marked with a box in the upper right quadrant are predicted as compounds mainly responsible for the low overall cytotoxic activity of the *B. sempervirens* extract (Buckets marked in green are detailed in Figure 7).

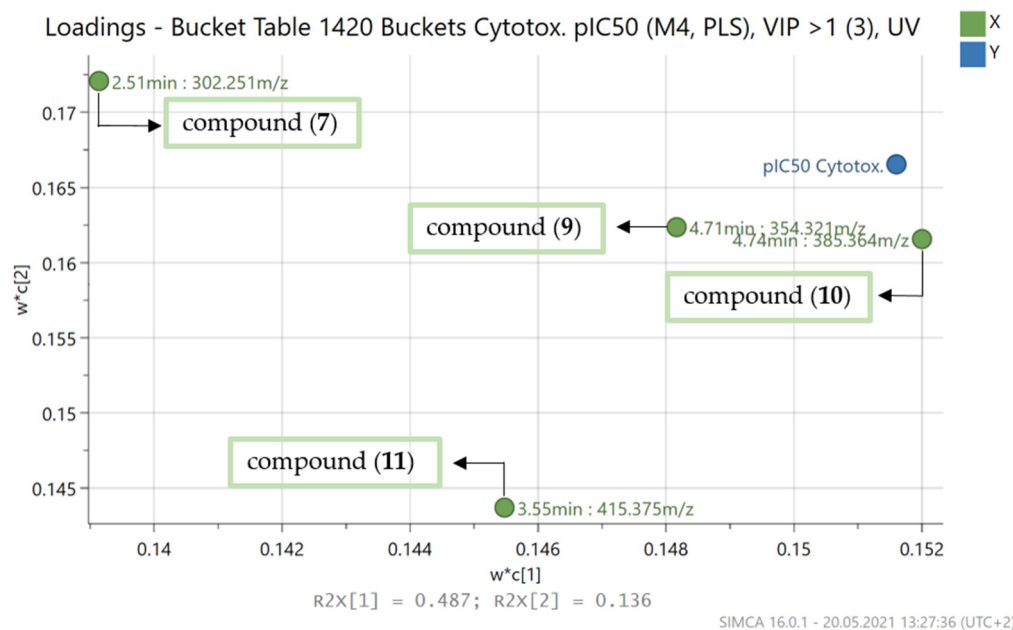


Figure 7. Expansion of region of the loadings plot (Figure 6) with the buckets that, according to the PLS model, have a high impact on the cytotoxicity of *B. sempervirens* L.

3. Materials and Methods

3.1. Plant Material

The collection, identification and other details of the plant material used in this study were reported previously [12].

3.2. Extraction and Fractionation of *B. sempervirens* L. Leaves

The extraction of plant material, the acid-base extraction to obtain the alkaloid fraction (ALOF) and the separation of ALOF into 20 different fractions by centrifugal partition chromatography (CPC) was described previously [12].

3.3. UHPLC/+ESI-QqTOF-Mass Spectrometry

LC/MS measurements for multivariate data analysis: The previously described method and parameters were used to obtain these data [12]. The 20 fractions were analyzed in a continuous sequence to obtain optimal comparability of retention and intensity data as needed for multivariate data analysis. The sample concentration of all fractions was 2 mg/mL.

LC/MS measurements to identify constituents of particular interest: To investigate in more detail the characteristic MS fragmentation of several constituents with predicted antiprotozoal activity (i.e., after analysis of the PLS models), a separate experiment was performed in MRM mode, for which the respective $[M + H]^+$ ion (deviation $m/z \pm 10$) was selected beforehand. The collision energy was 40 eV. Based on these spectra, fragmentation pathways were reconstructed (Figures S6, S11, S17, S22, S27 and S30, Supplementary Materials). The mass deviation of the fragments taken into consideration was <40 ppm in all cases.

3.4. Pretreatment of LC/MS Data

The raw LC/MS data were converted into molecular features using the find molecular features function of the software Data Analysis 4.1 (Bruker Daltonik GmbH, Bremen, Germany). The parameters were set as follows: signal/noise threshold: 2, correlation coefficient threshold: 0.7, minimum compound length: 10 spectra, smoothing width: 1, additional smoothing: yes, chemistry: positive adducts $M + H$, $M + NH_4$, $M + Na$, $M + K$, $M - H_2O + H$, $2M + H$, $2M + Na$, $2M + NH_4$, $2M + K$, $2M + CH_3CN + H$, $2M + CH_3CN + Na$, spectrum type: line spectra only, background subtraction: none and add fragment spectra: no.

For further processing, the obtained molecular features were imported into the software ProfileAnalysis 2.1 (Bruker Daltonik GmbH, Bremen, Germany). A descriptor table ("bucket table") was created in which each bucket value corresponds to mass signal intensity at a given m/z value and a given retention time [min]. The following parameters were applied: data selection and processing: find molecular features, retention time range: 1–15 min, mass range: 100–1500 m/z , advanced bucketing: width of retention time window 0.1 min; width of m/z window 30 ppm, split buckets with multiple compounds: yes, bucket filter: value count of bucket $\geq 10\%$, allow empty group attributes: yes, bucket value transformation: none and display bucket values in table: yes. The resulting bucket table consisted of 1420 buckets and served as a data matrix for the calculation of the following statistical models.

3.5. PLS Modeling

The bucket table (1420 bucket variables \times 20 analyses) was imported to the software SIMCA 16.0.1 (Sartorius Stedim Data Analytics AB, Umeå, Sweden) and constituted the X-matrix (independent variables) of the PLS model. The bioactivity data of all fractions were transformed to logarithmic scale ($pIC_{50} = -\log IC_{50} [\mu M]$) and used as the Y-matrix (dependent variables). Three different PLS models were generated with the activity data against *Pf* and *Tbr*, respectively, as well as the cytotoxicity values against mammalian cells (L6 cell line). All variables were scaled to unit variance (scaling factor: standard deviation) in order to assign equal weights to every variable in the model. The leave-one-out cross-validation was used to fully cross-validate the resulting models.

Buckets with a value ≥ 0.09 for $w^*c(1)$ (combined loading weight vector of X and Y for the first PLS component) and a $w^*c(2) \leq 0.41$ (combined loading weight vector of X and Y for the second PLS component) ($pIC_{50} Pf$ 0.145381 $w^*c(1)$ and 0.237369 $w^*c(2)$) were analyzed in the loadings plot of the PLS model with prediction of antiplasmodial compounds (Figures 1 and 2). In the antitrypanosomal PLS model buckets with a $w^*c(1) \geq 0.08$ and a $w^*c(2) \leq 0.13$ ($pIC_{50} Tbr$ 0.121195 $w^*c(1)$ and 0.119258 $w^*c(2)$) were evaluated (Figures 4 and 5).

3.6. In Vitro Bioassays

In vitro assays for the bioactivity of the 20 CPC fractions against *Tbr* (bloodstream trypomastigotes, STIB 900 strain) and *Pf* (intraerythrocytic forms, NF54 strain), and cytotoxicity tests against mammalian cells (L6-cell line from rat-skeletal myoblasts), were performed according to established standard protocols [18].

4. Conclusions

In conclusion, PLS models with good statistical performance were obtained with the LC/MS-data of 20 alkaloid fractions from the leaf extract of *B. sempervirens* L. for their antiplasmodial and antitrypanosomal activity as well as their cytotoxicity against mammalian cells. Using these models, antiprotozoal activity was predicted for eleven constituents. Of the four compounds responsible for the antiplasmodial activity, compound **1** and **4** have already been tested for their growth inhibition against *Pf* in the in vitro assay and exhibited promising IC₅₀ values (IC₅₀: 1.05 μM (**1**); IC₅₀: 1.76 μM (**4**)). Compound **2** represents a previously undescribed natural product. This new *Buxus*-alkaloid is of interest for subsequent isolation and testing, especially because it is not predicted by the respective PLS model to significantly contribute to cytotoxicity.

Among the eight compounds highlighted by the PLS model for antitrypanosomal activity, compound **4** and **11** have already been tested against *Tbr* and displayed reasonable activity (IC₅₀: 2.3 μM (**4**); IC₅₀: 1.5 μM (**11**)). Moreover, two new natural products, compound **6** and **10**, were predicted to be effective against *Tbr*. Compound **8** could be dereplicated as *N*-formylcyclovirobuxeine-B which, to the best of our knowledge, is described for the first time as a constituent of *B. sempervirens* L. Since they did not significantly contribute to the PLS model for cytotoxicity, compounds **6** and **8** are the most interesting compounds for further studies in terms of antitrypanosomal *Buxus*-alkaloids.

The present study thus represents a new, impressive example showing that the chosen PLS approach is an auspicious means of identifying active natural products in a time- and resource-saving manner that can rationally guide the subsequent isolation of such compounds. Since the concentrations of compounds **2**, **6**, **8** and **10** in their respective fractions were quite low, their isolation will have to start from a larger amount of plant material which can now be performed in a target-oriented manner.

Supplementary Materials: The following are available online, Figure S1: UHPLC/+ESI-QqTOF-MS/MS chromatogram of fraction 1–20, Figure S2: +ESI-QqTOF MS spectrum of compound **1**, Figure S3: +ESI-QqTOF MS/MS spectrum of compound **1**, Figure S4: +ESI-QqTOF MS spectrum of compound **2**, Figure S5: +ESI-QqTOF MS/MS spectrum of compound **2**, Figure S6: Possible fragmentation pathway of the [M + H]⁺ ion of compound **2**, Figure S7: +ESI-QqTOF MS/MS spectrum of compound **3**, Figure S8: +ESI-QqTOF MS spectrum of compound **3**, Figure S9: +ESI-QqTOF MS spectrum of compound **4**, Figure S10: +ESI-QqTOF MS/MS spectrum of compound **4**, Figure S11: Possible fragmentation pathway of the [M + H]⁺ ion of compound **4**, Figure S12: +ESI-QqTOF MS spectrum of compound **5**, Figure S13: +ESI-QqTOF MS/MS spectrum of compound **5**, Figure S14: +ESI-QqTOF MS spectrum of compound **6**, Figure S15: +ESI-QqTOF MS/MS spectrum of compound **6**, Figure S16: Detailed +ESI-QqTOF MS/MS spectrum of compound **6**, Figure S17: Possible fragmentation pathway of the [M + H]⁺ ion of compound **6**, Figure S18: +ESI-QqTOF MS spectrum of compound **7**, Figure S19: +ESI-QqTOF MS/MS spectrum of compound **7**, Figure S20: +ESI-QqTOF MS spectrum of compound **8**, Figure S21: +ESI-QqTOF MS/MS spectrum of compound **8**, Figure S22: Possible fragmentation pathway of the [M + H]⁺ ion of compound **8**, Figure S23: +ESI-QqTOF MS spectrum of compound **9**, Figure S24: +ESI-QqTOF MS/MS spectrum of compound **9**, Figure S25: +ESI-QqTOF MS spectrum of compound **10**, Figure S26: +ESI-QqTOF MS/MS spectrum of compound **10**, Figure S27: Possible fragmentation pathway of the [M + H]⁺ ion of compound **10**, Figure S28: +ESI-QqTOF MS spectrum of compound **11**, Figure S29: +ESI-QqTOF MS/MS spectrum of compound **11**, Figure S30: Possible fragmentation pathway of the [M + H]⁺ ion of compound **11**. The bucket tables (X and Y variable data) of all three PLS models are available online.

Author Contributions: Conceptualization, T.J.S.; software, L.U.S. and T.J.S.; investigation, L.U.S. and M.K.; resources, T.J.S. and P.M.; data curation, L.U.S., T.J.S. and M.K.; writing—original draft

preparation, L.U.S. and T.J.S.; writing—review and editing, L.U.S., T.J.S., M.K. and P.M.; supervision, T.J.S.; project administration, T.J.S. and P.M.; funding acquisition, L.U.S. and T.J.S. All authors have read and agreed to the published version of the manuscript.

Funding: This research received external funding in the form of a doctoral fellowship for L. U. Szabó from Apothekerstiftung Westfalen-Lippe.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data is contained within the article and Supplementary Materials. The underlying raw data are available from the corresponding author on request.

Acknowledgments: L.U.S. is most grateful for financial support by Apothekerstiftung Westfalen-Lippe in the form of a doctoral fellowship. Thanks are due to J. Sendker (Münster) for support with the UHPLC/+ESI-QqTOF-MS/MS analyses. We thank M. Cal, S. Keller-Maerki and R. Rocchetti for assistance with the parasite and cytotoxicity assays. This work is part of the activities of the Research Network Natural Products against Neglected Diseases (ResNet NPND, see www.resnetnpnd.org, accessed on 11 October 2021).

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

Sample Availability: Samples of the compounds **1**, **4** and **11** are available from the authors.

References

1. Ignatius, R.; Burchard, G.-D. Protozoen. In *Medizinische Mikrobiologie und Infektiologie*; Suerbaum, S., Burchard, G.-D., Kaufmann, S.H., Schulz, T.F., Eds.; Springer: Berlin/Heidelberg, Germany, 2016; pp. 653–675.
2. World Health Organization. World Malaria Report 2020. Available online: <https://www.who.int/publications/i/item/9789240015791> (accessed on 30 August 2021).
3. Ignatius, R.; Cramer, J.P. Allgemeine Parasitologie. In *Medizinische Mikrobiologie und Infektiologie*; Suerbaum, S., Burchard, G.-D., Kaufmann, S.H., Schulz, T.F., Eds.; Springer: Berlin/Heidelberg, Germany, 2016; pp. 651–652.
4. Brun, R.; Blum, J.; Chappuis, F.; Burri, C. Human African trypanosomiasis. *Lancet* **2010**, *375*, 148–159. [[CrossRef](#)]
5. Lee, S.-M.; Kim, M.-S.; Hayat, F.; Shin, D. Recent Advances in the Discovery of Novel Antiprotozoal Agents. *Molecules* **2019**, *24*, 3886. [[CrossRef](#)] [[PubMed](#)]
6. Njoroge, M.; Njuguna, N.M.; Mutai, P.; Ongarora, D.S.B.; Smith, P.W.; Chibale, K. Recent approaches to chemical discovery and development against malaria and the neglected tropical diseases human African trypanosomiasis and schistosomiasis. *Chem. Rev.* **2014**, *114*, 11138–11163. [[CrossRef](#)] [[PubMed](#)]
7. Szabó, L.U.; Kaiser, M.; Mäser, P.; Schmidt, T.J. Antiprotozoal Nor-Triterpene Alkaloids from *Buxus sempervirens* L. *Antibiotics* **2021**, *10*, 696. [[CrossRef](#)] [[PubMed](#)]
8. Flittner, D.; Kaiser, M.; Mäser, P.; Lopes, N.P.; Schmidt, T.J. The Alkaloid-Enriched Fraction of *Pachysandra terminalis* (Buxaceae) Shows Prominent Activity against *Trypanosoma brucei rhodesiense*. *Molecules* **2021**, *26*, 591. [[CrossRef](#)] [[PubMed](#)]
9. Llurba Montesino, N.; Kaiser, M.; Mäser, P.; Schmidt, T.J. *Salvia officinalis* L.: Antitrypanosomal Activity and Active Constituents against *Trypanosoma brucei rhodesiense*. *Molecules* **2021**, *26*, 3226. [[CrossRef](#)] [[PubMed](#)]
10. Ellendorff, T.; Brun, R.; Kaiser, M.; Sendker, J.; Schmidt, T.J. PLS-Prediction and Confirmation of Hydrojuglone Glucoside as the Antitrypanosomal Constituent of *Juglans* Spp. *Molecules* **2015**, *20*, 10082–10094. [[CrossRef](#)] [[PubMed](#)]
11. Bailey, N.J.C.; Wang, Y.; Sampson, J.; Davis, W.; Whitcombe, I.; Hylands, P.J.; Croft, S.L.; Holmes, E. Prediction of anti-plasmodial activity of *Artemisia annua* extracts: Application of 1H NMR spectroscopy and chemometrics. *J. Pharm. Biomed. Anal.* **2004**, *35*, 117–126. [[CrossRef](#)] [[PubMed](#)]
12. Szabó, L.U.; Schmidt, T.J. Target-Guided Isolation of O-tigloylcyclovirobuxeine-B from *Buxus sempervirens* L. by Centrifugal Partition Chromatography. *Molecules* **2020**, *25*, 4804. [[CrossRef](#)]
13. Musharraf, S.G.; Goher, M.; Ali, A.; Adhikari, A.; Choudhary, M.I.; Atta-ur-Rahman. Rapid characterization and identification of steroidal alkaloids in *Sarcococca coriacea* using liquid chromatography coupled with electrospray ionization quadrupole time-of-flight mass spectrometry. *Steroids* **2012**, *77*, 138–148. [[CrossRef](#)] [[PubMed](#)]
14. Musharraf, S.G.; Goher, M.; Shahnaz, S.; Choudhary, M.I.; Atta-ur-Rahman. Structure-fragmentation relationship and rapid dereplication of *Buxus* steroidal alkaloids by electrospray ionization-quadrupole time-of-flight mass spectrometry. *RCM* **2013**, *27*, 169–178. [[PubMed](#)]
15. Černý, V.; Šorm, F. Steroid Alkaloids: Alkaloids of Apocynaceae and Buxaceae. In *The Alkaloids: Chemistry and Physiology*; Academic Press: New York, NY, USA, 1967; pp. 305–426.
16. Blaschek, W.; Hager, H. *Hagers Enzyklopädie der Arzneistoffe und Drogen*; Wiss. Verl.-Ges: Stuttgart, Germany, 2007.

17. Khuong-Huu-Laine, F.; Magdeleine, M.J. Steroidal alkaloids. XCVI. Buxaceae family. Alkaloids of *Buxus malayana*: Cycloprotopuxine C, cyclovirobuxine C, *N*-formyl cyclovirobuxine B, Ovanillyl cyclovirobuxine B, and cyclobuxophyllinine M. *Ann. Pharm. Fr.* **1970**, *28*, 211.
18. Bernal, F.A.; Kaiser, M.; Wunsch, B.; Schmidt, T.J. Structure-Activity Relationships of Cinnamate Ester Analogues as Potent Antiprotozoal Agents. *Chem. Med. Chem.* **2020**, *15*, 68–78. [[CrossRef](#)] [[PubMed](#)]