Lipophilic Guanylhydrazone Analogues As Promising Trypanocidal Agents: An Extended SAR Study

Vasiliki Pardali\textsuperscript{a}, Erofili Giannakopoulou\textsuperscript{a}, Dimitrios-Ilias Balourdas\textsuperscript{a}, Vassilios Myrianthopoulos\textsuperscript{a}, Martin C. Taylor\textsuperscript{b}, Marina Šekutor\textsuperscript{c}, Kata Mlinarić-Majerski\textsuperscript{c}, John M. Kelly\textsuperscript{b}, and Grigoris Zoidis\textsuperscript{a,*}

\textsuperscript{a}School of Health Sciences, Faculty of Pharmacy, Department of Pharmaceutical Chemistry, National and Kapodistrian University of Athens, Panepistimiopolis-Zografou, GR-15771 Athens, Greece; \textsuperscript{b}Department of Infection Biology, London School of Hygiene and Tropical Medicine, Keppel Street, London WC1E 7HT, UK; \textsuperscript{c}Department of Organic Chemistry and Biochemistry, Rudjer Bošković Institute, Bijenička cesta 54, 10 000 Zagreb, Croatia

\*Corresponding author:
Prof. Grigoris Zoidis (zoidis@pharm.uoa.gr)

Keywords: Adamantane, S-adenosylmethionine decarboxylase (AdoMetDC), guanylhydrazones, structure-activity relationships, trypanocidal agents, Kernel-based Partial Least Squares regression, SZmap, Hydration analysis, Docking-scoring calculations

Abstract

In this report, we extend the SAR analysis of a number of lipophilic guanylhydrazone analogues with respect to in vitro growth inhibition of \textit{Trypanosoma brucei} and \textit{Trypanosoma cruzi}. Sleeping sickness and Chagas disease, caused by the tropical parasites \textit{T. brucei} and \textit{T. cruzi}, constitute a significant socioeconomic burden in low-income countries of sub-Saharan Africa and Latin America, respectively. Drug development is under-funded. Moreover, current treatments are outdated and difficult to administer, while drug resistance is an emerging concern. The synthesis of adamantane-based compounds that have potential as antitrypanosomal agents, is extensively reviewed. The critical role of the adamantane ring was further investigated by synthesizing and testing a number of novel lipophilic guanylhydrazones. The introduction of hydrophobic bulky substituents onto the adamantane ring generated the most active analogues, illustrating the synergistic effect of the lipophilic character of the C1 side chain and guanylhydrazone moiety on trypanocidal activity. The \textit{n}-decyl C1-substituted compound \textbf{G8} (\(R = C_{10}H_{21}\)) proved to be the most potent adamantane derivative against \textit{T. brucei} with activity in the nanomolar range (\(EC_{50}=90\) nM). Molecular
Simulations were also performed to better understand the structure-activity relationships between the studied guanylhydrazone analogues and their potential enzyme target.

1. Introduction

Neglected tropical diseases (NTDs) constitute a diverse group of chronic bacterial, protozoan, fungal, viral and helminth infections [1,2] in 149 countries worldwide [3]. It is estimated that more than a billion people — one-sixth of the world’s population, are affected by NTDs in tropical and subtropical regions. These infections are considered to be one of the leading causes of high morbidity and mortality in underdeveloped countries [4]. Leishmaniasis, Human African Trypanosomiasis (HAT), Chagas disease and malaria are the most serious and prominent protozoan infections, accounting for the highest death toll among the existing NTDs annually [5,6]. In this review, we will focus on the two trypanosomiasis, HAT and Chagas disease, caused by protozoan parasites of the genus *Trypanosoma*.

1.1 Human African trypanosomiasis

Human African trypanosomiasis (HAT), also known as sleeping sickness, is a vector-borne disease that affects poor populations in the African continent. It is caused by infection with the protozoan parasite *Trypanosoma brucei* and is transmitted by the bite of tsetse flies of the genus *Glossina* [7]. The disease is found in 36 countries of sub-Saharan Africa and about 65 million people are at risk of developing the infection [8].

Two morphologically identical subspecies of *T. brucei* are responsible for the onset of the disease in humans. They display a different host range, exhibit dissimilar virulence profiles, occur in separate geographical regions, and display distinct epidemiological characteristics and clinical manifestations [9]. *Trypanosoma brucei rhodesiense* causes an acute and rapidly progressing form of the disease, known as *rhodesiense* HAT (rHAT), which is localized to 13 countries in eastern and southern Africa [10]. It is a zoonotic disease [11], that is transmitted by tsetse flies of the *Glossina morsitans* group.
and passes directly or indirectly from animals or insects to humans. *Trypanosoma brucei gambiense* leads to a slow-progressing and chronic form of the disease, Gambian sleeping sickness or *gambiense* HAT (gHAT), which is found predominantly in 24 countries of western and central Africa [10]. The *palpalis* group is the vector of *T. b. gambiense*, with the main carriers being *Glossina palpalis palpalis* and *Glossina palpalis gambiensis* [13]. gHAT is an anthropoontic disease which is transmitted predominantly from person to person, although pigs and some wild animals have been reported as smaller reservoirs [11]. Infections from *T. b. gambiense* are the most common and represent over 97% of current cases [9].

Three severe outbreaks of the disease have been recorded in Africa. The first occurred between 1896 and 1906 with 800,000 deaths. The second began in 1920, peaking in 1930 with 70,000 registered cases, and ended in the late 1940s. The most recent widespread outbreak began in the 1970s and lasted until the end of the 1990s [14,15]. Incidence data from the World Health Organization (WHO) for 2009 showed a total of 9,878 new cases, with a sharp drop off observed in 2014. Nowadays, only the Democratic Republic of the Congo (DRC) reports an average of more than 1,000 new cases per year, comprising over 70% of the total recorded *T. b. gambiense* cases. However, the real number of people affected by this disease is estimated to be much higher due to under-reporting [8,16,17].

Tsetse flies are hematophagous, cyclical vectors of HAT [18]. Congenital transmission has also been demonstrated in newborn children within the first days of life and in older children born in nonendemic regions from infected mothers [19-22]. Transmission through sexual contact has been described once when the infection was passed from a male with confirmed gHAT to a woman who had never been to an endemic country [21-23]. Furthermore, mechanical transmission has also been reported *via* other blood-sucking insects, whereas accidental infections may occur from needle stick injuries in laboratories, or through blood transfusion [18,23].

HAT is clinically defined by two distinct stages. During the early hemolymphatic stage, trypanosomes spread and proliferate in the blood and the lymph. In the second or late meningoencephalitic stage, parasites are able to cross the blood-brain barrier (BBB) and invade the central nervous system (CNS) [16]. Although the progression of each stage involves different clinical features, it is often difficult to distinguish reliably between the two stages and they merge into each other, especially upon infection with *T. b. rhodesiense*. rHAT progresses rapidly and parasite invasion of the CNS occurs in a few weeks, causing death within months of infection. On the other hand, the neurological stage of gHAT generally starts several months or even years after infection, and the disease can take years or decades before becoming lethal [9,24].

After the bite from a tsetse fly, a chancre also known as a trypanome – a local skin reaction, develops at the feeding point. This occurs in about 20% of patients with rHAT, but is rarely observed in gHAT cases [10,11,25]. After an initial asymptomatic period that lasts one to three weeks in *rhodesiense*
infection, and usually some weeks longer in gHAT, patients develop nonspecific symptoms such as fever with chills, fatigue, headache, arthralgia, weight loss and a general sense of discomfort and unease. Splenomegaly, or less frequently hepatomegaly, has been observed, while enlargement of the posterior cervical lymph nodes, known as Winterbottom’s sign, is common in Gambian sleeping sickness [10,11,16]. The meningoencephalitic stage involves symptoms from almost all regions of the nervous system. These clinical manifestations are categorized by psychiatric, motor, sensory disturbances and sleep abnormalities, the main features that gave the disease its name. Deep and painful hyperesthesia is the leading sensory symptom, known as Kerandel’s sign. The disease causes deregulation of the circadian clock characterized by daytime somnolence followed by nocturnal insomnia. As the disease progresses, patients display an uncontrollable desire to sleep and alterations of sleep structure. During the final stage, cerebral edema, coma and systemic organ failure are observed, leading to sudden death [10,11,16,26-28].

The first issue in HAT diagnosis is to identify the disease as soon as possible by confirming the presence of parasites in the peripheral blood, or in other infected tissues such as the lymph nodes. Subsequently, the key step constitutes the accurate staging of the disease [29]. The non-specific signs and symptoms are insufficient for an accurate diagnosis. In rHAT, identification of trypanosomes is feasible because of the high levels of parasitemia, by microscopic examination of blood, or by a lymph node aspirate. The presence of the trypanosomal chancre is also typical of infection with T. b. rhodesiense. In contrast, gambiense infection is typified by cyclical parasitemia, reflecting a process of antigenic variation by the parasite, and parasitological diagnosis has poor sensitivity [7,24]. However, rapid and reliable serological tests (e.g., Card Agglutination Test for Trypanosomiasis (CATT), HAT Sero-K-SeT, SD Bioline HAT), based on the detection of specific antibodies, can be used for mass population screening [16,30,31]. Identification of the late stage of both types of HAT is based on the examination of cerebrospinal fluid (CSF) collected by lumbar puncture. WHO has defined 5 white blood cells (WBC) per μL as a threshold, or the presence of trypanosomes in the CSF, or both, as indicators that the infection has progressed to the CNS-stage. In gambiense infections, 20 cells per μL is used as the cutoff point [30]. Concentrations of immunoglobulin M (IgM) in the CSF, neopterin and combined panels of cytokines and chemokines, released after the tissue damage, have also been evaluated as biomarkers of HAT, and they reflect the severity of neurological signs [32,33].

Recognition of the causative T. brucei sub-species and correct staging of the disease are critical for choosing the appropriate chemotherapeutic regimen [10,34]. Drugs used for the first-stage disease are not able to cure second-stage disease, since therapy against the late stage requires drugs that travel across the BBB. Treatment of the first stage of sleeping sickness is limited to two drugs, pentamidine and suramin [35].
Pentamidine isethionate (Fig. (1)) is an aromatic diaminidine and was registered as the drug of choice for the treatment of infection caused by *T. b. gambiense* in 1940. As a positively charged molecule, pentamidine interacts electrostatically with cellular polyanions. The drug seems to bind to the minor groove of DNA and to be involved in the disruption of both kinetoplast (kDNA) and nuclear DNA. kDNA is thought to be the primary target with drug activity linked to inhibition of the mitochondrial topoisomerase II. It has also been suggested that diamidines may act through other mechanisms, such as inhibition of the enzyme S-adenosylmethionine decarboxylase (AdoMetDC), or by deregulation of oxidative phosphorylation. [34,36-40].

Suramin (Fig. (1)) is a polysulphonated naphthylamine and has been used as a standard treatment for the hemolymphatic stage of rHAT. Suramin appears to bind to glycolytic enzymes through electrostatic interactions and inhibit them, thus interrupting glycolysis. In a recent study, dysregulation in cytokinesis was observed following suramin treatment, while parasites with more than two nuclei were abundant, indicating a possible new mode of action. Anaphylactic shock, renal failure, skin lesions, neurological complications, such as peripheral neuropathy, and bone marrow toxicity have been reported as the main adverse reactions [34,36,37,39,40].

Melarsoprol (Fig. (1)) is a trivalent arsenical and the only effective drug for the second stage of both forms of sleeping sickness. It was first used in 1949. Melarsoprol is a prodrug which is converted, in trypanosomes, into the active form melarsen oxide. The active metabolite seems to engage in stable interactions with sulphhydryl groups and has a high affinity for thiols such as trypanothione, affecting the redox (reduction–oxidation) balance of the parasites. It also interacts with a variety of glycolytic and pentose phosphate pathway enzymes, and thus, inhibits ATP synthesis. The use of melarsoprol is linked with a high frequency of severe and life-threatening adverse effects, with an overall mortality rate of 5%. This drug is very toxic, and patients may develop a severe post-treatment reactive encephalopathy (PTRE) with mortality rates that climb to 50%. [9,34,36,37,39,40].

Eflornithine (Fig. (1)) was originally registered as an anticancer drug, but in the 1980s, it was found to be an effective therapeutic solution for the second stage of gHAT. It is a specific and irreversible inhibitor of ornithine decarboxylase (ODC), the first enzyme involved in the biosynthesis of polyamines, which are essential for cell growth and proliferation. The main adverse effects reported are fever, pruritus, hypertension, bone marrow suppression, gastrointestinal symptoms and rarely, seizures. [34,36,37,39,40].

Combination regimens of melarsoprol, eflornithine and nifurtimox were assessed for the treatment of HAT. Nifurtimox is registered to treat patients with American trypanosomiasis (below). Nifurtimox-eflornithine combination therapy (NECT) (Fig. (1)) was found to be the most effective and now constitutes the first-line therapy for the late stage of *gambiense* infection. Nifurtimox is a prodrug that
is bioactivated through reduction of the nitro group, yielding nitro-anion radicals and other active metabolites. NECT is based on the synergistic actions of the two drugs; nifurtimox leads to oxidative stress and widespread damage to macromolecules, while eflornithine reduces the levels of trypanothione, the key redox-protecting metabolite. Administration includes oral nifurtimox with a daily dose of 15 mg/kg, three times a day for 10 days, and eflornithine infusions of 400 mg/kg per day given every 12 hours for 7 days. NECT produced a marked improvement in terms of administration (56 infusions reduced to 14), high cure rates, low mortality rates and less adverse side effects, especially compared to melarsoprol. Shorter hospitalization is also required, and the cost of the therapeutic schedule has been significantly reduced. NECT-treated patients may be affected by musculoskeletal and abdominal pain, tremors, vomiting and nausea, headache and gastrointestinal disorders [36,41-43]. Two new drugs are currently undergoing clinical trials, fexinidazole and acoziborole (Fig. 1). Fexinidazole is a prodrug that is rapidly metabolized in vivo—through two different pathways, cytochrome P450 and a monooxygenase—into two metabolites, a sulfoxide and a sulfone, both of which exhibit trypanocidal activity. It is the only orally administered drug that is effective against both stages of HAT caused by *T. b. gambiense* and *T. b. rhodesiense*. This drug has been deemed broadly acceptable due to the easy administration (one oral dose per day for 10 days), its safety profile, its reduced cost in terms of synthesis, and the fact that there is no requirement for disease staging. Fexinidazole has received marketing authorization in the DCR for the treatment of gHAT, and has successfully entered phase 3 and phase 2 clinical trials in patients infected with gHAT of any stage and chronic Chagas disease, respectively [44-47]. Acoziborole (SCYX-7158) is another oral drug candidate that has been tested as a treatment for *T. b. gambiense* and *T. b. rhodesiense* infections. After phase 1 clinical trial, completed in 2015, it was confirmed that the drug is able to access the CNS, so may have potential against both disease stages. Phase IIb/III clinical trials have been ongoing in the DCR since 2016 [46-48]. A series of amidine-containing compounds have also been studied for their effect on *T. brucei* and some of them were evaluated in animal models and/or in clinical trials. However, the trials were discontinued due to high toxicity.
**Fig. (1).** Chemotherapeutic strategy against sleeping sickness and Chagas disease. Highlighted regions are reported to play a major role in the mechanism of action.
1.2 American trypanosomiasis

Chagas disease, or American trypanosomiasis, is a systemic, life-threatening zoonotic infection [49], caused by the protozoan parasite *Trypanosoma cruzi*. The disease is endemic in 21 continental Latin American countries, with an estimated 6-8 million people infected, mainly in poor rural areas, resulting in over 10,000 deaths annually. Chagas disease is also becoming a global public health problem. The migration of chronically infected individuals has spread the disease beyond its natural geographical boundaries, particularly to the USA and Europe [50]. The main route of transmission to humans is via the blood-sucking insect vector. Infection begins after mucous membranes, conjunctivae, or breaks in the skin become contaminated with parasites contained in the faeces of various species of the triatomine bug. However, other routes of transmission also exist; these include blood transfusion, organ transplantation, and vertical transmission from mother to infant [51,52]. Numerous outbreaks of the disease related to oral transmission, via food or liquid products contaminated with *T. cruzi*, have also been reported [49]. *T. cruzi* is able to invade almost all nucleated mammalian cells and has a wide host range.

Progression of the disease is characterized by two distinct clinical phases: an acute phase, and a chronic phase that develops over many years. The acute phase has a duration of 4-8 weeks and occurs shortly after the initial infection. This first stage is asymptomatic in 60 to 70% of patients. In the others, symptoms include inflammation at the inoculation site, fever, enlarged lymph nodes, splenomegaly and the highly characteristic unilateral palpebral oedema (Romaña sign). The infection is considered to be life-long, with 30% to 40% of patients subsequently developing chronic stage pathology, often 20 to 30 years after the acute infection. Symptoms can include cardiac disorders, mainly cardiomyopathy and arrhythmias, as well as gastrointestinal and/or neurological alterations that can result in death [51,53].

Currently, two drugs are available for the treatment of Chagas disease; nifurtimox (see above) and the most commonly used, benznidazole (BZ) (Fig. (1)). Recent findings suggest that BZ is responsible for the induction of potentially lethal double-stranded breaks in parasite DNA. This mechanism of action is based on reactive metabolites that are generated following nitroreductase mediated drug activation, which can damage DNA and other macromolecules, and result in the incorporation of oxidized nucleotides during DNA replication [54]. However, both drugs are only partially effective and relatively toxic, which, along with the prolonged administration required, limit their use [55].

There are no vaccines available against these two parasitic diseases. Vaccine development remains a daunting challenge, particularly in the case of HAT, where antigenic variation results in frequently changing surface glycoproteins that facilitate immune evasion. Thus, the major ways to prevent disease transmission are vector control and effective treatment. New and improved therapeutic options for both sleeping sickness and Chagas disease have been limited. The available drugs have
been in use for many years and are associated with poor efficacy, emergence of resistance, difficulties with administration, and severe adverse effects. Despite the progress made in the development of fexinidazole and, to a lesser extent, of acoziborole, both diseases continue to compromise public health in some of the poorest regions of the world. More efforts, focused on the discovery of new classes of therapeutics, are required [46,56,57].

1.3 The role of polyamines in trypanosomatids

The polyamine biosynthetic pathway is highly conserved across eukaryotic organisms and is crucial for the proliferation of trypanosomatids. The polyamines putrescine, spermidine and spermine are low molecular weight cationic organic amines that are essential for gene expression and cell growth, and play a crucial role in diverse cellular processes such as the biosynthesis of nucleic acids and proteins, cell proliferation and differentiation. The building blocks for the synthesis of polyamines (Fig. (2)) are the amino acids arginine and methionine. L-arginine is the substrate for the formation of L-ornithine, an intermediate of the urea cycle, upon hydrolysis catalyzed by arginase. L-ornithine is further decarboxylated to the diamine putrescine by ODC, the main regulatory enzyme in this pathway. S-adenosylmethionine synthase (AdoMet synthase) mediates the formation of S-adenosylmethionine (AdoMet) from methionine and ATP. Then, S-adenosylmethionine decarboxylase (AdoMetDC), a pyruvoyl-dependent enzyme, catalyzes the decarboxylation of AdoMet to generate an essential precursor (dcAdoMet) for spermidine biosynthesis. Subsequently, spermidine synthase transfers an aminopropyl group donated by the decarboxylated AdoMet to putrescine, forming spermidine. In mammalian cells and T. cruzi, addition of a second aminopropyl moiety leads to the synthesis of spermine via spermine synthase. Unique to trypanosomatid organisms, trypanothione, an essential redox cofactor, is required for defense against oxidative damage. Synthesis of this thiol is initiated by glutathionylspermidine synthase, which conjugates one glutathione molecule to spermidine, forming N1-glutathionylspermidine (Gsp). In the next step, a second glutathione is added to Gsp by trypanothione synthase, to generate trypanothione [58-60].
In general, the enzymes involved in polyamine biosynthesis differ between mammalian cells and parasites in terms of their structural features, or their presence or absence from the host. Therefore, they emerge as potential targets for the development of trypanocidal agents. Of these enzymes, ODC and AdoMetDC, which regulate polyamine homeostasis, constitute druggable targets [61-63]. As previously described, ODC is inhibited irreversibly by eflornithine, and other approaches toward the inhibition of ODC in *T. brucei* are being considered. *T. cruzi* lacks this enzyme [64]. However, both
parasites do express AdoMetDC. Inhibitors of AdoMetDC have been reported that have potent activity against *T. brucei* and *T. cruzi in vitro*, as well as in mouse models [65-71]. The trypanosomatid AdoMetDC differs from the respective human enzyme, as it possesses a unique regulatory mechanism. The enzyme is allosterically activated by formation of a heterodimer between the active subunit and a catalytically inactive paralogous pseudenzyme, referred to as a prozyme. Heterodimerization increases enzyme activity by three orders of magnitude. This novel regulation of the polyamine pathway in trypanosomatids suggests that selective inhibition of *T. brucei* AdoMetDC may be feasible, supporting the value of the target to combat HAT [72].

1.4 Drug design

**Adamantane**

The adamantane scaffold is a polycyclic cage moiety, characterized by high symmetry and a unique geometry of three fused cyclohexanes in a chair-like conformation. It is a rigid and practically strain-free structure that has been applied as a valuable structural subunit in medicinal chemistry for many years [73]. The adamantane moiety can either be incorporated into existing drugs to improve their pharmacokinetic and pharmacodynamic profile, or used as a starting compound for the development of novel therapeutic agents. Therefore, it has been widely used in drug discovery strategies. More specifically, it has been demonstrated that the addition of the adamantyl moiety into a molecule enhances its overall lipophilicity and subsequently increases cell membrane permeability of the drug candidate [73-75]. Another example of adamantane’s usefulness in medicinal chemistry is the increased affinity for lipophilic receptors that it confers, presumably due to greater hydrophobic interactions, giving rise to higher selectivity. Adamantane can be used to add steric bulk to a compound and to protect nearby functional groups from metabolic cleavage, thus offering greater stability and prolonged duration of action [73,75,76]. Finally, due to its rigidity and bulkiness, this cage scaffold can accommodate β-cyclodextrins (β-CD), resulting in inclusion complexes, or act mechanically as a blocker for several ion channels [77,78]. Because of all the above, the adamantyl subunit constitutes a privileged building block that provides drug-like properties to lead compounds. Amantadine, rimantadine, memantine, tromantadine, adapalene, vildagliptin and saxagliptin are seven examples of adamantane-containing marketed drugs (Fig. (3)) that are available today. Currently approved uses for the above-mentioned therapeutics range from viral infections (Influenza A, *Herpes simplex*), and neurodegenerative disorders (Parkinson’s disease, Alzheimer’s disease), to acne vulgaris and type II diabetes mellitus. It therefore comes as no surprise that many adamantyl compounds are under development, yielding a plethora of pharmacologically active analogues that
have been studied for various medical conditions, including iron overload disease, cancer, CNS diseases, inflammation, hypertension, hyperglycemia, malaria and tuberculosis [74-76,78-81].

Fig. (3). Clinically approved adamantane derivatives and their pharmacological profiles.

**Adamantane-based guanylhydrazone analogues**
Over the years, we have synthesized a variety of structurally diverse adamantane derivatives that target influenza A virus, *T. brucei* and *T. cruzi* [96-103]. In continuation of our studies, we were interested in developing further active adamantane-based compounds that could inhibit the proliferation of the above-mentioned trypanosomatids. Intrigued by the *in vitro* and *in vivo* trypanocidal properties of several guanylhydrazone analogues, that could act as potential AdoMetDC inhibitors [104-112], and following a molecular hybridization approach, we designed derivatives of the adamantane core bearing the iminoguanidine moiety in position 2 of the cage (Fig. (4)) [113].
Fig. (4). Structures of adamantane-substituted guanylhydrazones G1 – G25 described in the present study [112-114].

From this perspective, the design strategy was aimed at combining the increased lipophilicity of the adamantane cage structure with the trypanocidal properties activity of the guanylhydrazone moiety to generate compounds with better antiparasitic activity. The general approach for this class of compounds was to use the adamantane component as a lipid carrier for the BBB to counterbalance the potential transport limitations conferred by the conjugated polar iminoguanidine side chain. Thus, improving their BBB permeability enables enhanced CNS penetration and increased concentration in brain tissue, which is of crucial importance for the treatment of the second stage HAT. Finally, the incorporation of the guanylhydrazone group into this class of compounds should have the added advantage of structural similarity with the guanidine pharmacophore of the licensed trypanocidal drug pentamidine. To study the impact of changing the guanylhydrazone position and to evaluate the significance of the adamantane scaffold, we proceeded with the synthesis of novel guanylhydrazone derivatives (Series 5, Fig. (5)). In analogues G26 and G27, the iminoguanidine group was shifted to the methylene carbon of the benzyl substituent, leading to the idea of synthesis of the fused adamantane-cyclopentane-benzene guanylhydrazone analogue G28. Finally, replacement of the
adamantane cage structure by the tricyclic scaffolds of fluorene and dibenzocycloheptene, and the bendable 4-phenyl cyclohexane, afforded led to the analogues G29, G30 and G31, respectively.

![Chemical structures](image)

**Fig. (5).** Structures of novel guanylhydrazone derivatives G26 – G31, synthesized in the present study.

Here, we report the synthesis and structural characterization of novel guanylhydrazones (G26 – G31, Fig. (5)), with their subsequent in vitro biological evaluation against bloodstream form *T. brucei* and *T. cruzi* epimastigotes. We also report, for the first time, the results of the trypanocidal activity of previously synthesized adamantane guanylhydrazones (G1 – G5, Fig. (4)) that have been used as butyrylcholinesterase inhibitors. Finally, we provide an overview on the chemistry and pharmacological activity of adamantane-based guanylhydrazone derivatives (G6 – G25, Fig. (4)), synthesized in the recent past, by our research group. Yet another objective of the study was to evaluate the effect of structural modifications and steric parameters on antiparasitic activity and to correlate this activity with possible interactions in the active site of AdoMetDC, using docking-scoring calculations and Quantitative Structure-Activity Relationship (QSAR) studies.
2. Experimental

2.1 General Experimental methods and techniques

Melting points were determined using a Büchi capillary apparatus and are uncorrected. NMR experiments were performed to elucidate the structure and determine the purity of the newly synthesized compounds. $^1$H NMR and 2D NMR spectra (COSY, HSQC, HSQC-DEPT and HMBC) were recorded on a Bruker Ultrashield™ Plus Avance III 600 spectrometer (600.11 MHz, $^1$H NMR). $^{13}$C NMR spectra were recorded on a Bruker Ultrashield™ Plus Avance III 600 spectrometer (150.9 MHz, $^{13}$C NMR). Chemical shifts $\delta$ (delta) are reported in parts per million (ppm) downfield from the NMR solvent, with the tetramethylsilane or solvent (DMSO-$d_6$) as internal standard. Data processing including Fourier transformation, baseline correction, phasing, peak peaking and integrations were performed using MestReNova software v.12.0.0. Splitting patterns are designated as follows: s, singlet; br s, broad singlet; d, doublet; t, triplet; q, quartet; dd, doublet of doublets; dt, doublet of triplets; dq, doublet of quartets; td, triplet of doublets; tt, triplet of triplets; ttd, triplet of triplets of doublets; qd, quartet of doublets; quin, quintet; complex m, complex multiplet. Coupling constants ($J$) are expressed in units of Hertz (Hz). The spectra were recorded at 305 K (32 °C) unless otherwise specified. The solvent used to obtain the spectra was deuterated dimethyl sulfoxide, DMSO-$d_6$ (quin, 2.50 ppm, $^1$H NMR; septet, 39.52 ppm, $^{13}$C NMR). Elemental analyses (C, H, N) were performed by the Service Central de Microanalyse at CNRS (France), and were within ± 0.4% of the theoretical values. Elemental analysis results for the tested compounds correspond to >95% purity. The commercial reagents were purchased from Alfa Aesar, Sigma-Aldrich, and Merck, and were used without further purification. Organic solvents used were of the highest purity.

2.2 Synthetic procedure for the preparation of analogues G26 – G31

Aminoguanidine bicarbonate (1.5 mmol) was suspended in 18 mL of absolute ethanol, and it was progressively dissolved by dropwise acidification with concentrated hydrochloric acid, until pH was adjusted to 2. Subsequently, the appropriate ketone (1.5 mmol) was added and the mixture was refluxed under stirring for 5 h at 105 – 110 °C. After completion of the reaction time, ethanol was evaporated under reduced pressure and the residue was treated with dry Et$_2$O. The resulting product was cooled at 5 °C for 1 h and filtered off in vacuo. The precipitate was washed with dry Et$_2$O and dried to obtain the desired analogues G26 – G31 as crystals.

Tricyclo[3.3.1.1$^{3,7}$]decan-1-yl(phenyl)methylene)guanylhydrazone hydrochloride (G26)
Product **G26** was obtained using 500 mg (2.08 mmol) of 1-adamantyl phenyl ketone and 283 mg (2.08 mmol) aminoguanidine bicarbonate in 25 mL of absolute EtOH, resulting in 690 mg of white crystalline solid. Yield: almost quantitative; mp >250 °C (from EtOH/dry Et2O); 1H NMR (600.11 MHz, DMSO-d6) δ (ppm) [1.57 (dq, 3H, J=12.2 Hz, Jz=2.1 Hz), 1.66 (dt, 3H, Ji=12.5 Hz, Jz=3.1 Hz), H4-Ad, H6-Ad, H10-Ad], 1.75 (d, 3H, J=3.0 Hz, H2-Ad, H8-Ad, H9-Ad), 1.96 (~quin, 3H, J=3.2 Hz, H3-Ad, H5-Ad, H7-Ad), 7.09 (dd, 2H, Ji=8.1 Hz, Jz=1.5 Hz, Hz, Hz), 7.32-8.01 (br s, 3H, NH(NH)NH2), 7.47-7.54 (complex m, 3H, H3, H5, H5), 9.25 (s, 1H, NH(NH)NH2); 13C NMR (150.9 MHz, DMSO-d6) δ (ppm) 27.6 (C3-Ad, C5-Ad, C7-Ad), 36.0 (C4-Ad, C6-Ad, C10-Ad), 39.4 (C2-Ad, C8-Ad, C9-Ad), 127.6 (C2, C6), 129.0 (C3, C5), 129.1 (C4), 131.3 (C1), 155.7 (NH(NH)NH2), 164.6 (C=N). Anal. calcd. For C18H25ClN4, C, 64.95; H, 7.57; N, 16.83. Found: C, 65.00; H, 7.50; N, 16.79.

**Tricyclo[3.3.1.1k7]decan-2-yl(phenyl)methyleneguanylhydrazone hydrochloride (G27)**

Product **G27** was obtained using 500 mg (2.08 mmol) of 2-adamantyl phenyl ketone and 283 mg (2.08 mmol) aminoguanidine bicarbonate in 25 mL of absolute EtOH, resulting in 680 mg of white crystalline solid. Yield: almost quantitative; mp >250 °C (from EtOH/dry Et2O); 1H NMR (600.11 MHz, DMSO-d6) δ (ppm) 1.45 (d, 2H, Ji=12.3 Hz, H4c-Ad, H9c-Ad), 1.66 (d, 2H, Ji=3.1 Hz, H6-Ad), 1.69-1.78 (complex m, 4H, H5-Ad, H10-Ad), 1.79 (~quin, 1H, Ji=3.1 Hz, H5-Ad), 1.86 (~quin, 1H, J=3.0 Hz, H7-Ad), 1.93 (dd, 2H, Ji=12.8 Hz, Jz=2.8 Hz, H4a-Ad, H9a-Ad), 2.04 (q, 2H, Ji=2.9 Hz, H1-Ad, H3-Ad), 2.92 (s, 1H, H2-Ad), 7.09-7.29 (br s, 3H, NH(NH)NH2), 7.28 (dd, 2H, Ji=8.2 Hz, Jz=1.4 Hz, H2, H4), 7.47 (td, 1H, Ji=7.4 Hz, Jz=1.4 Hz, Hz), 7.51 (~dd, 2H, Ji=8.1, 6.6 Hz, Jz=1.5 Hz, H3, H5), 9.74 (s, 1H, NH(NH)NH2); 13C NMR (150.9 MHz, DMSO-d6) δ (ppm) 27.2 (C3-Ad), 27.3 (C7-Ad), 28.6 (C1-Ad, C3-Ad), 32.0 (C4-Ad, C6-Ad), 37.1 (C10-Ad), 37.9 (C9-Ad, C10-Ad), 50.1 (C2-Ad), 126.9 (C2, C6), 129.3 (C2, C5), 129.4 (C4), 155.8 (NH(NH)NH2), 159.5 (C=N). Anal. calcd. For C18H25ClN4, C, 64.95; H, 7.57; N, 16.83. Found: C, 64.91; H, 7.63; N, 16.85.

**Adamantane[1,2-a]inden-11-ylidene)guanylhydrazone hydrochloride (G28)**

Product **G28** was obtained using 500 mg (2.10 mmol) of adamantane[1,2-a]inden-11-one and 286 mg (2.10 mmol) aminoguanidine bicarbonate in 25 mL of absolute EtOH, resulting in 575 mg of white crystalline solid. Yield: 83%; mp 230 – 232 °C (from EtOH/dry Et2O); 1H NMR (600.11 MHz, DMSO-d6) δ (ppm) 1.51 (d, 2H, Ji=12.5 Hz, H9c-Ad, H13c-Ad), 1.61-2.08 (complex m, 7H, H5, H7, H8, H12), 2.31 (t, 4H, Ji=11.2 Hz, H6, H9a, H10, H13a), 2.81 (s, 1H, H10a), 7.29 (t, 1H, Ji=7.2 Hz, H3), 7.48 (t, 1H, Ji=7.4 Hz, H2), 7.83 (d, 1H, Ji=7.7 Hz, H1), 7.84-8.11 (br s, 3H, NH(NH)NH2), 8.22 (d, 1H, Ji=7.6 Hz, H4), 10.33 (s, 1H, NH(NH)NH2); 13C NMR (150.9 MHz, DMSO-d6) δ (ppm) 27.0 (C8), 27.1 (C6), 27.6 (C10), 32.2 (C7), 36.3 (C9), 36.5 (C12), 42.3 (C4b), 46.7 (C13), 47.8 (C5), 82.6
(C_{10a}), 122.2 (C_4), 126.0 (C_2), 128.0 (C_3), 133.8 (C_1), 136.1 (C_{4a}), 156.5 (C=N), 158.2 (NHC(=NH)NH_2), 162.5 (C_{11a}). Anal. calcd. For C_{18}H_{23}ClN_4, C, 65.34; H, 7.01; N, 16.93. Found: C, 65.40; H, 7.08; N, 16.88.

2-(9H-Fluoren-9-ylidene)guanylhydrazone hydrochloride (G29)
Product G29 was obtained using 500 mg (2.77 mmol) of 9H-fluoren-9-one and 377 mg (2.77 mmol) aminoguanidine bicarbonate in 33 mL of absolute EtOH, resulting in 695 mg of yellow crystalline solid. Yield: 92%; mp 238 – 240 °C (from EtOH/dry Et_2O); ^1H NMR (600.11 MHz, DMSO-^d_6) δ (ppm) [7.35 (td, 1H, J=7.4 Hz, J_z=0.9 Hz), 7.39 (td, 1H, J=7.6 Hz, J_z=1.2 Hz), H_2, H_7], [7.44 (td, 1H, J=7.5 Hz, J_z=1.2 Hz), 7.54 (t, 1H, J=7.5 Hz), H_3, H_6], [7.79 (d, 1H, J=7.5 Hz), 7.88 (d, 1H, J=7.5 Hz), H_4, H_5], 8.11 [(d, 1H, J=7.5 Hz), 8.76 (d, 1H, J=7.7 Hz), H_1, H_5], 8.23-8.67 (br s, 3H, NHC(NH)NH_2), 11.45 (s, 1H, NHC(NH)NH_2); ^13C NMR (150.9 MHz, DMSO-^d_6) δ (ppm) 120.2, 120.9 (C_4, C_5), 122.7, 127.4 (C_1, C_8), 128.2, 128.3 (C_2, C_7), 130.7, 131.9 (C_3, C_6), 136.3, 139.1 (C_{8a}/C_{9a}), 141.8 (C_{4a}, C_{4b}), 148.2 (C=N), 157.3 (NHC(=NH)NH_2). Anal. calcd. For C_{14}H_{13}ClN_4, C, 61.65; H, 4.80; N, 20.54. Found: C, 61.70; H, 4.89; N, 20.49.

2-(10,11-Dihydro-5H-dibenzo[a,d][7]annulen-5-ylidene)guanylhydrazone hydrochloride (G30)
Product G30 was obtained using 1.0 g (4.80 mmol) of 10,11-dihydro-5H-dibenzo[a,d][7]annulen-5-one (dibenzosuberone) and 653 mg (4.80 mmol) aminoguanidine bicarbonate in 58 mL of absolute EtOH, resulting in 405 mg of beige crystalline solid. Yield: 28%; mp 155-157 °C (from EtOH/dry Et_2O); ^1H NMR (600.11 MHz, DMSO-^d_6) δ (ppm) [2.87 (dd, 2H, J=17.8 Hz, J_z=10.0 Hz), 3.06 (d, 1H, J=13.6 Hz), 3.26 (d, 1H, J=17.2 Hz), H_{10}, H_{11}], 7.13 (d, 1H, J=7.7 Hz, H_1 or H_9), [7.22 (t, 1H, J=7.4 Hz), 7.23 (t, 1H, J=7.1 Hz), H_3, H_7], 7.29 (d, 1H, J=7.4 Hz, H_4 or H_8), 7.33 (t, 1H, J=7.3 Hz, H_2 or H_5), 7.41 (d, 1H, J=7.3 Hz, H_1 or H_9), 7.43 (t, 1H, J=7.8 Hz, H_2 or H_8), 7.55-8.13 (br s, 3H, NHC(NH)NH_2), 7.75 (d, 1H, J=7.7 Hz, H_4/H_6), 10.13 (s, 1H, NHC(NH)NH_2); ^13C NMR (150.9 MHz, DMSO-^d_6) δ (ppm) 31.1, 33.5 (C_{10}, C_{11}), 126.1, 126.6 (C_3, C_7), 127.3, 128.8 (C_2, C_8), 129.3, 129.6 (C_4, C_6), 130.6, 130.8 (C_1, C_9), 133.4, 135.2 (C_{4a}, C_{5a}), 138.4, 139.4 (C_{9a}, C_{11a}), 156.2 (C=N), 159.1 (NHC(=NH)NH_2). Anal. calcd. For C_{16}H_{17}ClN_4, C, 63.89; H, 5.70; N, 18.63. Found: C, 63.93; H, 5.82; N, 18.66.

2-(4-Phenylcyclohexylidene)guanylhydrazone hydrochloride (G31)
Product G31 was obtained using 500 mg (2.87 mmol) of 4-phenylcyclohexan-1-one and 391 mg (2.87 mmol) aminoguanidine bicarbonate in 34 mL of absolute EtOH, resulting in 760 mg of white crystalline solid. Yield: almost quantitative; mp 218 – 220 °C (from EtOH/dry Et_2O); ^1H NMR (600.11 MHz, DMSO-^d_6) δ (ppm) 1.59 (qd, 1H, J=12.9 Hz, J_z=4.1 Hz, H_{2a} or H_{6a}), 1.64 (qd, 1H,
\[ J_1 = 12.7 \text{ Hz}, J_2 = 4.3 \text{ Hz}, H_{3\alpha} \text{ or } H_{5e}, \]

\[ 1.98 \text{ (ttd, } 2\text{H, } J_1 = 9.3 \text{ Hz, } J_2 = 6.3, 5.8 \text{ Hz, } J_3 = 3.3 \text{ Hz, } H_{2a} \text{ or } H_{6a}, H_{3\alpha} \text{ or } H_{5e}), \]

\[ 2.14 \text{ (td, } 1\text{H, } J_1 = 14.1 \text{ Hz, } J_2 = 5.1 \text{ Hz, } H_{2e} \text{ or } H_{5e}), \]

\[ 2.42 \text{ (td, } 1\text{H, } J_1 = 13.5 \text{ Hz, } J_2 = 3.5 \text{ Hz, } H_4), \]

\[ 2.85 \text{ (tt, } 1\text{H, } J_1 = 12.2 \text{ Hz, } J_2 = 3.5 \text{ Hz, } H_6), \]

\[ 3.15 \text{ (dq, } 1\text{H, } J_1 = 14.7 \text{ Hz, } J_2 = 5.1 \text{ Hz, } H_{2e} \text{ or } H_{6e}), \]

\[ 7.25 \text{ (d, } 2\text{H, } J = 7.4 \text{ Hz, } H_2, H_6), \]

\[ 7.28 \text{ (t, } 2\text{H, } J = 7.5 \text{ Hz, } H_3, H_5), \]

\[ 7.48 - 7.95 \text{ (br s, } 3\text{H, } NHC(=N)NH_2), \]

\[ 8.84 \text{ (s, } 0.25\text{H, } NHC(=N)NH_2) \]

\[ 11.34 \text{ (s, } 0.75\text{H, } NHC(=N)NH_2); \]

\[ ^{13}\text{C NMR (150.9 MHz, DMSO-}d_6\text{) } \delta \text{ (ppm) 27.3, 32.5 (C}_2, C_6), 33.8, 34.3 (C}_3, C_5), 42.2 (C}_4), 126.1 (C}_4), 126.6 (C}_2, C_6), 128.3 (C}_3, C_5), 145.5 (C}_1), 156.2 (NHC(=N)NH_2), 158.8 (C=N), 159.2 (C=N). \]

2.3 Trypanocidal assays

For compounds G1 – G25: A bloodstream form of *T. brucei brucei* (strain 427, derivative 221) was cultured in 25 mL flasks at 37 °C in modified Iscove’s medium (pH 7.4) [115]. To assess activity, parasites were grown for 3 days in the presence of test compounds and the EC\textsubscript{50} and EC\textsubscript{90} values were determined (concentrations that inhibit growth by 50% and 90%, respectively). In these experiments, the densities of untreated cultures increased from 0.25 × 10\textsuperscript{4} to 4 × 10\textsuperscript{6} cells mL\textsuperscript{-1}. After determination of cell densities at each drug concentration with a hemocytometer (Weber Scientific International Ltd.), drug sensitivity was expressed as a percentage of growth of control cells. Data presented are the mean of triplicate experiments ± SEM.

For compounds G26 – G31: Bloodstream form *T. brucei* (strain 427, derivative 221) were cultured in modified Iscove’s medium, as outlined previously [116]. Eight-point potency curves were performed in 96-well plates (200 μL volumes), and the compound concentrations that inhibited growth by 50% (EC\textsubscript{50}) and 90% (EC\textsubscript{90}) were determined. Parasites were first diluted to 2.5 × 10\textsuperscript{4} mL\textsuperscript{-1}, compounds were added at range of concentrations, and the plates incubated at 37 °C. Resazurin was added after 48 h, and the plates incubated for a further 16 h. *T. cruzi* epimastigotes (strain CL Brener) were cultured as previously described [117]. Trypanocidal activity was determined in microtiter plates as outlined above, with the following modifications. Experiments were initiated by seeding the parasites at 2.5 × 10\textsuperscript{5} mL\textsuperscript{-1}, and after the addition of test compounds, cultured at 28 °C for 4 days. Resazurin was added, the plates were incubated for a further 2 days, and then assessed as above. Fluorescence intensities were determined using a BMG FLUOstar Omega (excitation 545 nm, emission 590 nm). Data were analyzed using Graph Pad Prism 7 software. Values are expressed as EC\textsubscript{50} ± SD and are the average of three independent replicates.

2.4 *In vitro* cytotoxicity assays on rat skeletal myoblast L6 cells
Cytotoxicity against L6 cells was assessed using microtiter plates. Briefly, cells were seeded in triplicate at $1 \times 10^4 \text{mL}^{-1}$ in growth medium containing different compound concentrations. The plates were incubated for 6 days at 37 °C and resazurin was then added to each well. After a further 8 h of incubation, the fluorescence was determined using a BMG FLUOstar Omega plate reader.

2.5 Computer-aided drug design

The discovery of new drugs is undoubtedly an expensive and time-consuming multi-step process. However, it is accelerated through our increased understanding of the fundamental principles of protein-ligand interactions, in addition to the integration of a computer-aided drug design (CADD) strategy. CADD strategies involve application of computer modelling technologies, which ultimately aid in selection and filtering of candidate compounds, generation of hypotheses on their mode of action, and also facilitate their further development [118,119]. Over the last decades, due to advancements in computer technology, CADD has been considered a rapidly growing field and plays a crucial role in drug development projects, from hit identification to lead optimization of novel, potentially therapeutically valuable small molecules [120-123]. CADD strategies can be mainly classified into three broad categories; the ligand-based (LBDD) and the structure-based (SBDD) approaches, as well as fragment-based drug discovery (FBDD), a newly emerging lead discovery method [124,125]. LBDD approaches utilize structure and activity data from a set of known active and inactive compounds in order to retrieve other potential molecular scaffolds based on similarity measures, common pharmacophores or descriptor values. LBDD methods include, amongst others, similarity and substructure searching, quantitative structure-activity relationships (QSAR), as well as pharmacophore and three-dimensional (3D) shape matching [126,127]. Contrarily, SBDD approaches are based on direct calculation of protein-ligand interactions utilizing the 3D structure of the biological target, determined through X-ray crystallography and NMR, or through homology modeling, and include ligand docking-based scoring, molecular dynamics and de novo design [128]. Notably, both of the above mentioned knowledge-driven approaches can be synergistically integrated to improve the drug design process [129]. In this current study, we are mainly focusing on QSAR analysis via 2D fingerprint similarity and docking-based scoring, one methodology from each CADD approach.

2.6 Docking-based scoring

Introduction
The docking process involves the prediction of ligand conformation and posing, as well as binding interactions within the targeted protein site. This is a rather challenging task by itself, since even relatively simple organic molecules may contain many conformational degrees of freedom, or show a diversity of binding modes. Thorough sampling of ligand orientation within the binding domain of the target is provided through stochastic search techniques such as the Monte Carlo algorithm [130]. Rapid and accurate assessment of protein-ligand complexes is pivotal for docking approaches, since a typical ligand docking experiment generates hundreds to thousands of ligand conformations [131]. Thus, an efficient scoring function is essential to rank order different docked poses for each ligand, and ultimately to discriminate the valid binding mode predictions from the invalid ones [132].

Evaluation of docked poses is achieved by four distinct classes of scoring functions that are currently applied in computational drug design: force-field-based, empirical, knowledge-based and a rather new approach, machine learning-based scoring functions [133], with the first being the scoring function of our choice.

Protocol
Prior to any docking experiment, protein and ligand databases need to be prepared. The crystal structure of *T. brucei* AdoMetDC/prozyme heterodimer (PDB structure: 5tvf) was computationally cleaved to include only AdoMetDC in complex with the bound ligand, CGP. Then, Protein Preparation Wizard [134] workflow was executed, as previously thoroughly described in detail previously [135], generating a structurally sound model ensuring accuracy on all downstream modeling simulations. Preparation of the guanylhydrazone analogues was performed using the LigPrep tool [136]. LigPrep generates precise, energy minimized 3D molecular structures with correct protonation states, tautomerism and enumeration of all possible stereoisomers. Docking grid preparations and docking calculations were accomplished as previously described [135]. Two different grids were generated using no scaling for protein van der Waals (VdW) radii and an extended ligand centering box of $10 \times 10 \times 10 \, \text{Å}$ and $15 \times 15 \times 15 \, \text{Å}$, respectively, to ensure proper fitting of the large aliphatic chains of some guanylhydrazone analogues. In this study, the Glide algorithm [137] was utilized, which is considered to be one of the most efficient and accurate virtual screening tools [138] for structure-based virtual screening of the guanylhydrazone analogues inside the AdoMetDC binding site. All docking calculations were performed by using VdW scaling of 0.8 for ligand atoms, a maximum of 3 poses per ligand and extra-precision (XP), a rather extensive, accurate and time- and CPU-demanding mode [139].

2.7 QSAR analysis
Introduction

Although QSAR was first developed more than 55 years ago [140], it still remains one of the most promising research areas in medicinal chemistry and cheminformatics. QSAR aims to reveal a quantitative correlation between the structural features of chemical entries and their many different properties, such as continuous (pIC_{50}, pEC_{50}, Ki, etc.), categorical/binary (active, inactive, toxic, nontoxic, etc.) and biological/toxicological. To this end, the main strategy involves the generation of a statistical linear or non-linear regression model that matches experimental values to a diverse set of molecular descriptors determined from the structure of the compounds [141]. Chemical descriptors constitute the basis of QSAR modeling and reflect a plethora of chemical structure representation. These molecular descriptors have been categorized into different classes, including constitutional, geometrical, topological, quantum chemical and so on, with 2D (two-dimensional) being the most popular among drug developers. Molecular descriptors compose the basis of QSAR modeling and represent different levels of chemical structure representation ranging from single physico-chemical properties (1D QSAR) to three-dimensional structures (3D QSAR), and even more intricate, multidimensional approaches like simulation of induced fit and solvation scenarios (5D and 6D QSAR) [142]. The success of QSAR analysis depends on the quality of the input data, selection of appropriate descriptors and statistical methods to validate the developed model. QSAR analysis has yielded quite a few success stories, regarding including currently approved drugs [143-145]. In the framework of this study, the Canvas cheminformatics software package was utilized [146], in an attempt to connect ones and zeros with reliable SAR predictions. Canvas provides a wide range of applications for structural and data analysis, including 2D and 3D fingerprints, similarity searching, as well as building regression and classification models.

Protocol

Preparation of the guanylhydrazone analogues was performed with Ligprep, as previously described [135], with the exception of generating only the most probable ionization and tautomeric state of each analogue, chosen on the basis of energetic considerations. Then, prepared compounds were introduced into the Canvas interface and the complete set of 498 physicochemical, topological, Ligfilter and Qikprop descriptors were determined for all entries. A total of nine fingerprints were calculated using default atom typing and scaling parameters. Seven 2D fingerprints, (linear, radial, dendritic, MOLPRINT 2D, atom pairs, atom triplets, and topological torsions) and two 3D pharmacophore fingerprints (2p and 3p pharmacophore), with Tanimoto similarity were used for ranking. A detailed description of each fingerprint has been realized established by Sastry and colleagues [147]. Statistical regression models were then generated, utilizing three diverse and well-established statistical methods, namely Multiple Linear Regression (MLR), Partial Least Squares
regression (PLS) and Kernel-based Partial Least Squares regression (KPLS), in order to successfully and efficiently correlate the cell-based in vitro activity of the guanylhydrazone analogues on against T. brucei to the aforementioned molecular descriptor and 2D/3D fingerprints. The initial approach included the implementation of all guanylhydrazone analogues in the training set. As expected, the compounds with none or marginal activity against T. brucei (G12, G14, G15 and G20) seemed to impede the generation of an accurate and reliable model. Indicatively, standard deviation (SD), as well as the coefficient of determination (R^2), resulted from MLR analysis, and proved that the model containing all analogues was quite questionable (SD = 1.526; R^2 = 0.5720). Therefore, in order to improve the predictivity of the model, only active compounds were utilized. A training set of 27 active compounds were analysed through MLR, PLS and KPLS regression functions with a 20% of random test set assignment for model evaluation. The option of five best subsets was used for MLR, attempting to find the highest-ranking model containing said number of independent variables, utilizing a simulated annealing Monte Carlo technique. For each MLR model generated, the option of “best subsets” was chosen, attempting to find the highest-ranking models containing five independent variables. Default settings were used for the simulated annealing Monte Carlo technique. Selection of the best MLR model was performed through careful evaluation of each model, considering taking into account intercorrelation between selected independent variables. For both PLS and KPLS, four factors were used. Hierarchical clustering of the guanylhydrazone analogues was attempted, based on the 2D fingerprint of atom pairs. Analogues were divided into three clusters, which were then statistically processed as previously described. Note that this approach did not yield any meaningful results.

3. Results and Discussion

3.1 Chemistry

Synthesis of Series 1 compounds
For the synthesis of the starting ketones 3 and 4 (Scheme 1), adamantan-2-one 1 was first converted to 4-oxahomoadamantan-5-one 2 via a selenium dioxide mediated oxidation. Lactone 2 was heated
with a 50% sulfuric acid and yielded a mixture consisting of the axial and the equatorial isomer of 3, in an approximate ratio of 5 : 1. Oxidation of the latter with chromic acid afforded the adamantane dione 4 [148].

Scheme 1. Synthetic route to the starting compounds 3 and 4 and synthesis of mono and bis(guanylhydrazone) adamantanes G1, G2 and G4. Reagents and conditions: (a) SeO₂, H₂O₂, t-BuOH, 96%; (b) H₂SO₄ (50%), 5 h, 90 °C; (c) H₂CrO₄; (d) H₂NNHC(=NH)NH₂ · HCl, EtOH, reflux, 5 h, 99% (G1), 97% (G2) and 94% (G4) [114,148].

The synthetic route for the formation of starting compounds 9 and 10 is described in Scheme 2. Oxygen-containing adamantane analogues are provided by functionalization of the adamantane core with the use of oxidizing agents. Initially the adamantane hydroxylation was performed upon treatment with fuming sulfuric acid (H₂O₇S₂). The products of the oxidative procedure are three diols obtained in different ratio. In fuming acid with a SO₃ percentage of 20%, 1,4-adamantane diol 6 is the main product in an overall yield of 45 – 50%. 1,3-Diol 7 and 2,6-diol 8 were also formed with 15 – 20% and 5 – 10% yields, respectively. The mixture of diols 6, 7 and 8 was subjected to oxidation with chromic acid to give 5-hydroxyadamantan-2-one/1-hydroxyadamantan-4-one 9 and adamantane-2,6-dione 10 [149].
Scheme 2. Synthetic route to the starting compounds 9 and 10 and synthesis of mono and bis(guanylhydrazone) adamantanes G3 and G5, respectively. Reagents and conditions: (a) 20% SO\(_3\) (10 °C), 20 min, 25 °C then 100 °C, 2 h; (b) CrO\(_3\), H\(_2\)O, 30 min, 70 °C; (c) H\(_2\)NNHC(=NH)NH\(_2\)·HCl, EtOH, reflux, 5 h, 98% (G3) and 99% (G5) [114,149].

Adamantyl guanylhydrazones G1 – G5 used in this study were synthesized according to the previously published procedure by condensing the corresponding adamantanone with aminoguanidine hydrochloride (Schemes 1 and 2) almost quantitatively. Note that compounds G2 and G4 (Scheme 1) have been isolated as a mixture of diastereoisomers and derivative G2 has three possible diastereoisomers while derivative G4 has four possible diastereoisomers. However, for guanylhydrazones G1 (Scheme 1), G3 and G5 (Scheme 2), only one diastereomeric form is possible due to molecular symmetry [114].

**Synthesis of Series 2 compounds**

The synthesis of 1,2-disubstituted adamantane derivatives was the key step for the preparation of analogues described in this study and was achieved by rearrangement of the protoadamantane framework [150]. The starting compound, protoadamantanone 13, was synthesized by Majerski’s group for the first time in 1979. As shown in Scheme 3, the commercially available adamantan-1-ol 11 reacted with lead tetraacetate and iodine in benzene as a solvent, and the iodo ketone 12 that was generated was refluxed with potassium hydroxide to provide protoadamantanone 13 in yields ranging from 71% to 82% [151]. The amount of the unreacted starting alcohol was decreased by heating the
mixture composed of adamantanol 11, lead tetraacetate and iodine for 2 h at 75 – 76 °C. The target compound 13 was then produced, almost pure, in excellent yields [152].

![Scheme 3](image)

**Scheme 3.** Synthesis of protoadamantanone. Reagents and conditions: (a) Pb(OAc)_4, I_2, C_6H_6, 70 – 75 °C (b) KOH, MeOH, reflux [151].

As depicted in Scheme 4, synthesis of the starting ketones 18a – d [112] proceeded with formation of the acetylenic alcohols 14a – d. More specifically, protoadamantanone 13 reacted with the appropriate lithiated acetylide, which was prepared in situ by addition of a substituted acetylene to a solution of butyllithium, to give the desired compounds 14a – d as a mixture of endo/exo isomers (1:1) in an overall yield of 82 – 95%. Catalytic hydrogenolysis of the latter with Adams’ catalyst yielded the corresponding saturated isomeric alcohols 15a – d. The mixture was refluxed with formic acid and the formate esters immediately underwent an intramolecular metathesis to give analogues 16a – d. Subsequent saponification of the intermediates 16a – d in the presence of an aqueous sodium hydroxide solution led to alcohols 17a – d in excellent yields which underwent Jones oxidation to form the target 1-alkyl-adamantan-2-ones 18a – d almost quantitatively [112].

For the synthesis of 1-alkyl-adamantane 2-guanylhydrazones G6 – G9, the starting ketones 18a – d were refluxed with aminoguanidine bicarbonate in the presence of concentrated hydrochloric acid (pH ~ 2) and ethanol as a solvent. The target compounds G6 – G9 were obtained in an overall yield of 85 to 94% [112].
Scheme 4. Synthetic route to the starting compounds 1-alkyl-adamantan-2-ones 18a – d and synthesis of 1-alkyl-adamantane 2-guanylhydrazones G6 – G9. Reagents and conditions: (a) R–C≡CH, n-BuLi, dry THF, 1 h, 0 °C then 20 h, rt, 82 – 95%; (b) H2O; (c) H2/PtO2, EtOH, 4 h, rt, 55 psi; (d) HCOOH, reflux, 30 min; (e) NaOH, EtOH/H2O, reflux, 2 h, 91 – 96%; (f) CrO3, aq. H2SO4 (8N), acetone, 15 °C then rt, 1 h, 95 – 98%; (g) H2NNHC(=NH)NH2 · HCO3, conc. HCl (pH ~ 2), EtOH, reflux, 5 h, 85 – 94% [112].

Thiosemicarbazones G10 and G11 (Scheme 5), possessing a n-hexyl and a n-tetradecyl chain, respectively, were obtained by condensation of thiosemicarbazide with the structurally related ketones 18a and 18d [112].

Scheme 5. Synthesis of the 1-alkyl-adamantane 2-thiosemicarbazones G10 and G11. Reagents and conditions: (a) H2NNHC(=S)NH2, EtOH, reflux, 5 h, 82% (G10) and 89% (G11) [112].

Synthesis of Series 3 compounds

Synthesis of the adamantane diol 25 (Scheme 6) [152] starts with a Reformatsky reaction between protoadamantanone 13 and bromoacetic acid ethyl ester in the presence of zinc metal, as shown in
Scheme 6. Upon saponification with sodium hydroxide in an EtOH/H₂O solution, the hydroxyacid 19 was obtained in poor yields along with some starting ketone. Alternatively, when protoadamantanone 13 was used as starting material and treated with LiCH₂COOCH₃, followed by hydrolysis of the intermediate ester 20 under basic conditions, the desired compound 19 was afforded with an improved yield of 98%. LiCH₂COOCH₃ was prepared by lithium bis(trimethylsilyl)amide and methyl ethanoate in one pot. Then, the hydroxyacid 19 provided a mixture of the respective ester 21 as a major product and the lactone 22 in traces, under reflux with formic acid. Subsequent saponification of the mixture and oxidation of the intermediate 23 with a solution of chromium trioxide in diluted sulfuric acid generated the ketoacid 24. Finally, lithium aluminum hydride catalyzed reduction of the latter gave rise to the desired diol 25 [152].

1-Alkylxyadamantane 2-guanylhydrazones G12 – G15 and 1-benzyloxyadamantane 2-guanylhydrazone G20 were prepared as outlined in Scheme 6. The intermediate 1-alkoxy alcohols 26a – d were synthesized via alkylation with the appropriate alkyl bromide in the presence of sodium hydride in N,N-dimethylformamide. Under similar reaction conditions, benzyl chloride was used as an alkylating agent for the formation of the benzyl substituted alkoxy alcohol 26e. The aforementioned intermediates were oxidized to the respective ketones 27a – e by chromium trioxide in an aqueous sulfuric acid solution. Treatment of the latter with aminoguanidine bicarbonate, upon pH adjustment to 2 with hydrochloric acid, yielded the guanylhydrazones G12 – G15 and G20 [113].
Scheme 6. Synthesis of the starting adamantane diol 25 and synthetic procedure for the preparation of the 1-alkyloxy and 1-benzyloxyadamantane 2-guanylhydrazones G12 – G15 and G20, respectively. Reagents and conditions: (a) i) BrCH₂COOEt, Zn, C₆H₆, reflux; ii) NaOH, EtOH/H₂O, 90 °C, 2.5 h, (HCl), 25%; (b) [(CH₃)₃Si]₂NLi, CH₃COOCH₃, dry THF, −75 °C, 20 min and −60 °C, 30 min; (c) NaOH, EtOH/H₂O, 90 °C, 2.5 h, (HCl), 98%; (d) HCOOH, reflux, 30 min; (e) NaOH, EtOH/H₂O, 90 °C, 2.5 h, (HCl), 87%; (f) CrO₃, aq. H₂SO₄ (1M), acetone, 15 °C then rt, 24h, 92%; (g) LiAlH₄, dry Et₂O, reflux; (h) For compounds 26a – d: i) NaH, dry DMF, 20 min, rt; ii) RBr, dry DMF, 16 h, 105 °C, 69 – 87%; For compound 26e: i) NaH, dry DMF, 20 min, rt; ii) PhCH₂Cl, dry DMF, 2 h, rt; iii) 1 h, 60 °C; iv) 16 h, rt, 82%; (i) CrO₃, aq. H₂SO₄ (8N), acetone, 15 °C then rt, 1h, quantitative yield; (j) H₂NNHC(=NH)NH₂ · HCO₃, conc. HCl (pH ~ 2), EtOH, reflux, 6 h, 56 – 89% [113,152].

For the synthesis of the (butoxyethoxy)ethyl substituted adamantane 2-guanylylhydrazone G17 (Scheme 7) [113], the adamantane diol 25 was first treated with sodium hydride and the resulting alkoxide anion was subsequently reacted with ethyl bromoacetate. The intermediate ethanoate was saponified, and upon acidification with concentrated hydrochloric acid, the hydroxy acid 28 was produced with an overall yield of 85%. The latter was reduced with lithium aluminum hydride to the respective diol 29, which was alkylated with 1-bromobutane in the presence of sodium hydride in poor yields. Jones oxidation of the alcohol 30, followed by coupling of the precursor ketone 31 with aminoguanidine bicarbonate, as previously described, led to the desired compound G17 [113].

Scheme 7. Synthetic procedure for the preparation of 1-(butoxyethoxy)ethyladamantane 2-guanylhydrazone G17. Reagents and conditions: (a) i) NaH, dry DMF, 20 min, rt; ii) BrCH₂COOCH₂CH₃, dry DMF, 8 h, 110 °C; (b) NaOH, EtOH/H₂O, 2 h, 100 °C, conc. HCl, 85%; (c) LiAlH₄, dry THF, reflux, 8 h, 96%; (d) i) NaH, dry DMF, 20 min, rt; ii) CH₃CH₂CH₂CH₂Br, dry
DMF, 15 h, 105 – 110 °C, 18%; (e) CrO₃, aq. H₂SO₄ (8N), acetone, 15 °C then rt, 30 min, 89%; (f) H₂NNHC(=NH)NH₂ · HCO₃, conc. HCl (pH ~ 2), EtOH, reflux, 6 h, 79% [113].

As depicted in Scheme 8 [113], a five-step synthetic route was applied for the preparation of the guanylhydrazone G18. Starting from diol 25, the ethyl ester 32 was obtained by alkylation with ethyl 5-bromovalerate. Then, conversion of 32 to diol 33 was carried out by reduction mediated by lithium aluminum hydride. Subsequent alkylation of 33, to the corresponding methoxy congener 34 was achieved with methyl iodide and sodium hydride as a base. Employment of the above-mentioned procedure, oxidation of 34 with Jones reagent and condensation of the intermediate ketone 35 with aminoguanidine bicarbonate afforded the analogue G18 in 69% yield [113].

Scheme 8. Synthetic procedure for the preparation of 1-(methoxypentyl)oxyethyladamantano 2-guanylhydrazone G18. Reagents and conditions: (a) i) NaH, dry DMF, 30 min, rt; ii) Br(CH₂)₄COOCH₂CH₃, dry DMF, 15 h, 105 °C, 32%; (b) LiAlH₄, dry Et₂O, reflux, 15 h, 88%; (c) i) NaH, dry DMF, 20 min, rt; ii) CH₃I, dry DMF, 24 h, rt, 40%; (d) CrO₃, aq. H₂SO₄ (8N), acetone, 15 °C then rt, 30 min, 89%; (e) H₂NNHC(=NH)NH₂ · HCO₃, conc. HCl (pH ~ 2), EtOH, reflux, 5 h, 69% [113].

The synthetic route to 1-benzyloxyadamantane 2-guanylhydrazone G19 [113] is depicted in Scheme 9. For the preparation of the adamantane diol 37 [100], protoadamantanone 13 was treated with dimethylsulfonium methylide to give a mixture of endo/exo epimers of the epoxide 36 in a 1:15 ratio. Then, oxirane ring opening under acidic conditions in acetone produced the target compound 37. Alternatively, the cleavage of oxirane rings is pronounced in the presence of formic acid. Reaction of epoxide 36 with formic acid, upon heating, yielded the respective formate diester, which was further saponified to provide diol 37 [100].
The intermediate 1-alkoxy alcohol 38 was synthesized via alkylation of 37 with benzyl chloride, using sodium hydride as a base and N,N-dimethylformamide as solvent. Alcohol 38 was oxidized to the corresponding ketone 39 by chromium trioxide in an aqueous sulfuric acid solution. The guanyldydrazone G19 was prepared by reacting the precursor ketone 39 with aminoguanidine bicarbonate in the presence of hydrochloric acid with an overall yield of 65% [113].

Scheme 9. Synthesis of the starting adamantane diol 37 and synthetic procedure for the preparation of 1-benzyloxyadamantane 2-guanylhydrazone G19. Reagents and conditions: (a) (CH$_3$)$_2$S=CH$_2$, DMSO, 3 h, 25 °C and 8 h, 55 °C; (b) H$_2$SO$_4$/H$_2$O (0.085 M), acetone, T °C or (c) HCOOH, 30 min, T °C; (d) NaOH, EtOH/H$_2$O, T °C; (e) i) NaH, dry DMF, 20 min, rt; ii) PhCH$_2$Cl, dry DMF, 2 h, rt; iii) 1 h, 60 °C; iv) 16 h, rt, 80%; (f) CrO$_3$, aq. H$_2$SO$_4$ (8N), acetone, 15 °C then rt, 1h, quantitative yield; (g) H$_2$NNHC(=NH)NH$_2$ · HCO$_3$, conc. HCl (pH ~ 2), EtOH, reflux, 6 h, 65% [100,113].

The key intermediate 43 [152] was obtained by a four-step synthetic procedure as depicted in Scheme 10. Protoadamantanone 13 undergoes a typical Grignard reaction with the dimagnesium derivative of propargylic alcohol to form the acetylenic diol 40, which was hydrogenated in the presence of platinum (IV) oxide as a catalyst, to the respective saturated diol 41. The diol 43 was produced in good yields upon heating of diol 41 with formic acid and subsequent saponification of the intermediate diester 42 [152].

Compound 43 was alkylated with 1-bromohexane and benzyl chloride to give the alcohols 44a and 44b, respectively. Employment of the previously described experimental protocol for the preparation of guanylhydrazones G12 – G15 and G20 (Scheme 6), Jones oxidation of the intermediates 44a and 44b, and subsequent condensation of ketones 45a and 45b with aminoguanidine bicarbonate, led to the desired analogues G16 and G21 in 94% and 53% yields, respectively [113].
**Scheme 10.** Synthesis of the starting adamantane diol 43 and synthetic procedure for the preparation of 1-alkyloxy and 1-benzyloxy propylandamantane 2-guanylhydrazones G16 and G21. Reagents and conditions: (a) BrMg–C≡C–CH₂OMgBr, dry THF, reflux, 6 h and 24 h, 20 °C; ii) NH₄Cl, H₂O, 82%; (b) H₂/PtO₂, 97%; (c) HCOOH, reflux, 30 min, 100%; (d) NaOH, EtOH/H₂O, reflux, 2 h, 81%; (e) for compound 44a: i) NaH, dry DMF, 20 min, rt; ii) C₆H₁₃Br, dry DMF, 16 h, 105 °C, 78%; for compound 44b: i) NaH, dry DMF, 20 min, rt; ii) PhCH₂Cl, dry DMF, 2 h, rt; iii) 1 h, 60 °C; iv) 16 h, rt, 85%; (f) CrO₃, aq. H₂SO₄ (8N), acetone, 15 °C then rt, 1 h, quantitative yield; (g) H₂NNHC(=NH)NH₂·HCO₃, conc. HCl (pH ~ 2), EtOH, reflux, 6 h, 94% (G16) and 53% (G21) [113,152].

**Synthesis of Series 4 compounds**

The preparation of 1-phenyl substituted adamantan-2-one 49a [153] was achieved following a synthetic procedure similar to that of 1-alkyl-adamantan-2-ones 18a – d (Scheme 4). Starting from protoadamantanone 13, the tertiary alcohol 46a was obtained upon treatment with phenyl lithium to facilitate the formation of the C–C bond. Employment of the previously described methodology, heating with formic acid to produce the formate 47a, saponification of the latter and oxidation with Jones reagent of the intermediate alcohol 48a, gave the desired ketone 49a in good yields [153]. The starting compound 1-benzylandamant-2-one 49b was synthesized as described above; with the only exception being that benzylmagnesium bromide reacted with protoadamantanone 13 instead of phenyl lithium. Following the similar procedure for the formation of guanylhydrazones G6 – G9 (Scheme 4), 1-phenyl adamantan 2-guanylhydrazone G22 and 1-benzylandamantane 2-guanylhydrazone G23 were obtained almost quantitatively and in 96% and 78% yields, respectively [112].
Scheme 11. Synthetic route to the starting compounds 1-phenyl and 1-benzyl adamantan-2-ones 49a and 49b, respectively and synthetic route for the preparation of 1-phenyl and 1-benzyladamantane 2-guanylhydrazones G22 and G23, respectively. Reagents and conditions: (a) CsHsLi in [CH$_3$(CH$_2$)$_3$]$_2$O (2M), PhCH$_2$MgBr, dry THF, 1 h, 20 °C, 96%; (b) H$_2$O; (c) HCOOH, reflux, 30 min; (d) NaOH, EtOH/H$_2$O, T °C, 2 h, 96%; (e) CrO$_3$, aq. H$_2$SO$_4$ (8N), acetone, 15 °C then rt, 1h; (f) H$_2$NNHC(=NH)NH$_2$ · HCO$_3$, conc. HCl (pH ~ 2), EtOH, reflux, 5 h, 96% (G22) and 78% (G23) [112,153].

The synthesis of 1-(benzyloxy)adamantane 2-guanylhydrazones G24 and G25 [113] was initiated with conversion of protoadamantanone 13 to the tertiary alcohols 50a and 50b, upon treatment with the appropriate organomagnesium bromide via a Grignard reaction. According to the procedure described in Scheme 12, the intermediates 50a and 50b were esterified with formic acid, and upon self-metathesis, the respective formates (not presented in Scheme 12) were obtained. Saponification of the latter afforded the corresponding alcohols which underwent Jones oxidation. The resulting ketones 51a and 51b were coupled with aminoguanidine bicarbonate in an ethanolic solution under acidic conditions. The desired guanylhydrazones G24 and G25 were isolated as hydrochloride salts in overall yields of 79% and 66%, respectively [113].
Scheme 12. Synthetic procedure for the preparation of 1-(benzyloxy)adamantane 2-guanylhydrazones G24 and G25. Reagents and conditions: (a) XCeH₄MgBr, dry THF; (b) HCOOH, reflux, 30 min; (c) NaOH, EtOH/H₂O, 2 h, T °C; (d) CrO₃, aq. H₂SO₄ (8N), acetone, 15 °C then rt, 30 min, 93% (51a) and 91% (51b); (e) H₂NNHC(=NH)NH₂ · HCO₃, conc. HCl (pH ~ 2), EtOH, reflux, 5 h, 79% (G24) and 66% (G25) [113].

Synthesis of Series 5 compounds

For the preparation of starting ketone 53, adamantanecarbonyl chloride 52 was allowed to react with phenylmagnesium bromide in anhydrous THF and the target compound 53 was obtained in 67% yield [154]. The novel guanylhydrazone G26 was prepared following the synthetic protocol described above for the formation of analogues G6 – G9 (Scheme 4) [112].

Scheme 13. Synthesis of 1-adamantyl methyl guanylhydrazone G26. Reagents and conditions: (a) PhMgBr, dry THF (−78 °C), 2 h, rt, Ar, 67%; (b) H₂NNHC(=NH)NH₂ · HCO₃, conc. HCl (pH ~ 2), EtOH, reflux, 5 h, almost quantitatively [112,154].

The synthetic route to 2-adamantyl phenyl ketone 57 (Scheme 14) involved a typical Horner-Wadsworth-Emmons reaction of 2-adamantanone 1 with (methoxymethyl)triphenylphosphonium chloride, in the presence of n-BuLi as a base, to deprotonate the latter under anhydrous conditions to give the unsaturated ether 54. Acid-promoted hydrolysis of the enol ether 54 with a 70% perchloric
acid generated the aldehyde 55 in good yields. Then, the latter reacted with phenylmagnesium bromide to give the respective alcohol 56. A pyridinium chlorochromate (PCC) catalyzed oxidation of 56 yielded ketone 57, almost quantitatively [155]. The reaction of equimolar quantities of the precursor 57 with aminoguanidine bicarbonate under the same conditions previously mentioned, produced the novel guanylhydrazone G27 almost quantitatively [112].

Scheme 14. Synthesis of 2-adamantyl methyl guanylhydrazone G27. Reagents and conditions: (a) CH₃OCH₂P(C₆H₅)₃Cl, n-BuLi, dry Et₂O (0 °C), 1 h then 15 h, rt, N₂; (b) HClO₄ 70%/H₂O, 78%; (c) i) PhMgBr, dry THF (0 °C), 5 h, rt; ii) NH₄Cl, H₂O, 89%; (d) [C₅H₅NH][CrO₃Cl]³⁻, CH₃COONa, CH₂Cl₂, 2 h, 95%; (e) H₂NNHC(=NH)NH₂ · HCO₃, conc. HCl (pH ~ 2), EtOH, reflux, 5 h, almost quantitatively [112,155].

The formation of guanylhydrazone G28 is presented in Scheme 15. The synthesis of the intermediate epoxide 58 was achieved via the Corey-Chaykovsky reaction which involves the base-mediated addition of trimethylsulfoxonium iodide to 1-phenyl adamantan-2-one 49a. Sodium hydride was used as a base and the reaction was performed in anhydrous dimethyl sulfoxide. Afterwards, the oxirane ring was cleaved by boron trifluoride diethyl etherate and subsequent rearrangement provided the aldehyde 59. Upon Jones oxidation, carboxylic acid 60 was produced, and it reacted with thionyl chloride to give the respective chloride. Finally, treatment with aluminum chloride in carbon disulphide drives cyclization of the latter to ketone 61, almost quantitatively [156]. Following the general procedure presented in Scheme 4, the novel guanylhydrazone G28 was obtained upon treatment with aminoguanidine bicarbonate under acidic conditions, with an overall yield of 83% [112].
Scheme 15. Synthesis of guanylhydrazone G28. Reagents and conditions: (a) (CH$_3$)$_3$SO$^+$, NaH, dry DMSO, 30 min, 20 °C then 2 h, 55 – 58 °C; (b) BF$_3$ · Et$_2$O, dry C$_6$H$_6$; (c) CrO$_3$, aq. H$_2$SO$_4$ (8N), acetone, 15 °C then rt, overnight; (d) SOCl$_2$, 15 min, 65 °C; (e) AlCl$_3$, CS$_2$, 30 min, rt then reflux, 30 min, quantitatively; (f) H$_2$NNHC(=NH)NH$_2$ · HCO$_3$, conc. HCl (pH ~ 2), EtOH, reflux, 5 h, 83% [112,156].

As previously described [112], the newly synthesized guanylhydrazones G29 – G31 were prepared by condensation of the commercially available precursors 62a – c with a suspension of aminoguanidine hydrochloride. The latter was prepared in turn from a suspension of aminoguanidine bicarbonate and an excess of concentrated hydrochloric acid upon pH adjustment to 2. The desired compounds were produced almost quantitatively, with the only exception being G30, which was isolated in an overall yield of 28%.

Scheme 16. Synthesis of novel guanylhydrazones G29 – G31. Reagents and conditions: (a) H$_2$NNHC(=NH)NH$_2$ · HCO$_3$, conc. HCl (pH ~ 2), EtOH, reflux, 5 h, 28 – 98%.

3.2 Biological evaluation
Series 1-5 analogues (G1 – G31) were assessed based on their ability to inhibit proliferation of cultured bloodstream form *T. brucei in vitro*. Series 5 analogues (G26 – G31) were also tested against *T. cruzi* epimastigotes *in vitro*. The EC<sub>50</sub> and EC<sub>90</sub> values for each compound are presented in Tables 1 – 5. As shown, 20 (G3, G8 – G11, G13, G16 – G19, G21, G22 – G25, G26 – G29 and G31) out of the 31 tested compounds had promising EC<sub>50</sub> values in the low micromolar to sub micromolar range against *T. brucei*. Most of the compounds were evaluated as hydrochloride salts, with the only exception being analogues G10 and G11, which were tested as free bases.

<table>
<thead>
<tr>
<th>Cmpd</th>
<th>Structure</th>
<th><em>T. brucei</em> EC&lt;sub&gt;50&lt;/sub&gt; (μM)</th>
<th><em>T. brucei</em> EC&lt;sub&gt;90&lt;/sub&gt; (μM)</th>
</tr>
</thead>
<tbody>
<tr>
<td>G1</td>
<td><img src="image" alt="Structure" /></td>
<td>56.9 ± 3.3</td>
<td>88.6 ± 13.1</td>
</tr>
<tr>
<td>G2</td>
<td><img src="image" alt="Structure" /></td>
<td>167 ± 6</td>
<td>214 ± 4</td>
</tr>
<tr>
<td>G3</td>
<td><img src="image" alt="Structure" /></td>
<td>9.67 ± 1.21</td>
<td>21.7 ± 3.9</td>
</tr>
<tr>
<td>G4</td>
<td><img src="image" alt="Structure" /></td>
<td>&gt;300</td>
<td>&gt;300</td>
</tr>
<tr>
<td>G5</td>
<td><img src="image" alt="Structure" /></td>
<td>&gt;300</td>
<td>&gt;300</td>
</tr>
</tbody>
</table>

* Concentrations required to inhibit growth of *T. brucei* by 50% and 90%, respectively. EC<sub>50</sub> and EC<sub>90</sub> data are the mean of triplicate experiments ± SEM.

As shown in Table 1, the 2-(guanylhydrazone)adamantane G1 exhibited moderate inhibitory activity in the micromolar range, with an EC<sub>50</sub> value of 56.9 μM. Introduction of a second guanylhydrazone group at position 4 of the adamantane ring resulted in a 3-fold decrease in the activity of G2 (EC<sub>50</sub> = 167 μM). It is of great interest that shifting the pharmacophore moiety from position 4 to position 6 of the adamantane cage increased the activity to the low micromolar range. The bis(guanylhydrazone) analogue G3 was 17 times more potent (EC<sub>50</sub> G3 = 9.67 μM) than the structurally related compound
G2. Incorporation of the OH functional group at position 4 of mono(guanylhydrazone) analogue G1 led to a substantial loss of activity for the G4 (EC$_{50}$ = >300 μM). Changing the position of the hydroxyl group to the vicinal carbon atom yielded compound G5, which also showed no activity at the highest concentration tested (~300 μM).

<table>
<thead>
<tr>
<th>Cmpd</th>
<th>Structure</th>
<th>T. brucei EC$_{50}$ (μM)</th>
<th>T. brucei EC$_{90}$ (μM)</th>
</tr>
</thead>
<tbody>
<tr>
<td>G6</td>
<td><img src="image" alt="" /></td>
<td>&gt;6.0</td>
<td>-</td>
</tr>
<tr>
<td>G7</td>
<td><img src="image" alt="" /></td>
<td>&gt;6.0</td>
<td>-</td>
</tr>
<tr>
<td>G8</td>
<td><img src="image" alt="" /></td>
<td>0.09 ± 0.02</td>
<td>0.11 ± 0.00</td>
</tr>
<tr>
<td>G9</td>
<td><img src="image" alt="" /></td>
<td>0.84 ± 0.50</td>
<td>1.30 ± 0.50</td>
</tr>
<tr>
<td>G10</td>
<td><img src="image" alt="" /></td>
<td>2.81 ± 0.09</td>
<td>5.64 ± 0.54</td>
</tr>
<tr>
<td>G11</td>
<td><img src="image" alt="" /></td>
<td>0.42 ± 0.04</td>
<td>0.54 ± 0.02</td>
</tr>
</tbody>
</table>

*a Concentrations required to inhibit growth of T. brucei by 50% and 90%, respectively. EC$_{50}$ and EC$_{90}$ data are the mean of triplicate experiments ± SEM.

Table 2: Activity of the 1-alkyl-2-guanylhydrazone adamantanes G6 – G9 and the 1-alkyl-2-thiosemicarbazone adamantanes G10 and G11 against cultured bloodstream form T. brucei (pH 7.4) in vitro.

Compounds that constitute Series 2, characterized by a lipophilic tail in position 1 of the adamantane core, were more potent than those of Series 1. In particular, the EC$_{50}$ values were at low micromolar to submicromolar levels (2.81 μM to 0.09 μM), as indicated in Table 2. The least amount of inhibitory activity was displayed by 1-hexyl- and 1-octyl substituted adamantane guanylhydrazones G6 and G7, respectively, whose median effective concentration was above 6 μM. 1-Decyl substitution provided the most potent analogue (G8) of this series, with an EC$_{50}$ value of 0.09 μM. Compared to its unsubstituted congener G1 (EC$_{50}$ = 56.9 μM), G8 was 632-fold more active. Larger alkyl groups
seem to reduce trypanocidal activity, given that the 1-tetradecyl substituted analogue G9 (EC$_{50}$ = 0.84 μM) was 9 times less effective than G8. It is noteworthy that replacement of the C=NH of guanylhydrazone moiety by the C=S functional group, which resulted in thiosemicarbazones G10 and G11, enhanced their activity compared to the structurally related analogues G6 and G9 (EC$_{50}$ G10 = 2.81 μM; EC$_{50}$ G11 = 0.42 μM).

Table 3 Activity of the 1-alkyloxy-2-guanylhydrazone adamantanes G12 – G18 and the 1-benzyloxy-2-guanylhydrazone adamantanes G19 – G21 against cultured bloodstream form T. brucei (pH 7.4) in vitro.

<table>
<thead>
<tr>
<th>Cmpd</th>
<th>Structure</th>
<th>T. brucei EC$_{50}$ (μM) $^a$</th>
<th>T. brucei EC$_{90}$ (μM) $^a$</th>
</tr>
</thead>
<tbody>
<tr>
<td>G12</td>
<td><img src="image1" alt="Structure" /></td>
<td>Marginal activity</td>
<td>Marginal activity</td>
</tr>
<tr>
<td>G13</td>
<td><img src="image2" alt="Structure" /></td>
<td>1.17 ± 0.13</td>
<td>1.61 ± 0.26</td>
</tr>
<tr>
<td>G14</td>
<td><img src="image3" alt="Structure" /></td>
<td>Marginal activity</td>
<td>Marginal activity</td>
</tr>
<tr>
<td>G15</td>
<td><img src="image4" alt="Structure" /></td>
<td>Marginal activity</td>
<td>Marginal activity</td>
</tr>
<tr>
<td>G16</td>
<td><img src="image5" alt="Structure" /></td>
<td>0.70 ± 0.07</td>
<td>0.98 ± 0.37</td>
</tr>
<tr>
<td>G17</td>
<td><img src="image6" alt="Structure" /></td>
<td>3.64 ± 0.07</td>
<td>4.94 ± 0.05</td>
</tr>
<tr>
<td>G18</td>
<td><img src="image7" alt="Structure" /></td>
<td>3.62 ± 0.28</td>
<td>4.91 ± 0.36</td>
</tr>
<tr>
<td>G19</td>
<td><img src="image8" alt="Structure" /></td>
<td>1.84 ± 0.16</td>
<td>2.42 ± 0.08</td>
</tr>
<tr>
<td>G20</td>
<td><img src="image9" alt="Structure" /></td>
<td>Marginal activity</td>
<td>Marginal activity</td>
</tr>
</tbody>
</table>
Table 3 presents the results of the biological evaluation of the Series 3 analogues, 1-alkyloxy and 1-benzyloxy adamantane 2-guanylhydrazones (G12 – G21). Most of them exhibited good trypanocidal activity, with potency in the range of 0.70-3.64 μM. The 1-hexyloxyethyl substituted adamantane G12, had a marginal activity towards T. brucei. Increasing the length of the alkyloxy chain of G12 by only one carbon atom, resulting in derivative G13 (heptyloxyethyl analogue), appears to have had a favorable effect on trypanocidal activity, resulting in an EC50 of 1.17 μM. Somewhat surprisingly, further extension of the C1 alkylloxy side chain of G13 by 1 or 3 carbon atoms (octyloxyethyl analogue G14 and decyloxyethyl analogue G15, respectively) caused a reduction in activity. These biological results prompted the synthesis of homologue G16, whose trypanocidal activity was improved compared with G13, by shifting the oxygen atom from position 3 to position 4 of the side chain. Compound G16 was the most potent of the mono(oxygenated) alkyl substituted adamantanes of this series, with an EC50 value of 0.70 μM. The bis(oxygenated) ten-membered side chain substitution in analogue G17 led to a slight decrease in activity, by 3 times with respect to the congener G13 (EC50 G17 = 3.64 μM; EC50 G13 = 1.17 μM). Increasing the distance between the two oxygen atoms (from 2 to 5 carbons) gave rise to compound G18, which was equipotent to G17 (approximately 3.6 μM). The same trend was observed for the three benzyl ethers G19 – G21, which were either active in the low micromolar range, or inactive. More specifically, the benzyloxyethyl analogue G20 showed a marginal activity. On the other hand, incorporation of a methylene and a 3-carbon propylene linker between the oxygen atom and the adamantane component, in analogues G19 and G21, had an opposite effect. The resulting EC50s were 1.84 μM and 1.78 μM, respectively.

Table 4 Activity of the aromatic-substituted 2-guanylhydrazone adamantanes G22 – G25 against cultured bloodstream form T. brucei (pH 7.4) in vitro.

<table>
<thead>
<tr>
<th>Cmpd</th>
<th>Structure</th>
<th>T. brucei EC50 (μM) a</th>
<th>T. brucei EC90 (μM) a</th>
</tr>
</thead>
<tbody>
<tr>
<td>G21</td>
<td><img src="image" alt="Structure" /></td>
<td>1.78 ± 0.17</td>
<td>3.88 ± 0.33</td>
</tr>
</tbody>
</table>

*Concentrations required to inhibit growth of T. brucei by 50% and 90%, respectively. EC50 and EC90 data are the mean of triplicate experiments ± SEM.*
Activity of the aromatic 1-phenyl- or 1-benzyl-substituted 2-guanylhydrazone adamantanes G22 – G25 of Series 4 against cultured bloodstream form T. brucei (pH 7.4) in vitro, is presented in Table 4. Aromatic substituted analogues G22 – G25 were active in the low micromolar range, with EC_{50} values between 0.41 μM and 2.12 μM. More specifically, phenyl substitution on the C-1 atom of the parent compound G1 led to a 27-fold increase in potency of analogue G22 (EC_{50} = 2.12 μM). Increasing the distance between the adamantane scaffold and the aromatic ring with a methylene group in compound G23 appears to conferred a beneficial effect on trypanocidal activity. The EC_{50} value of G23 (EC_{50} = 0.41 μM) was 5 times lower than that of G22. Introduction of the fluorine electron withdrawing group, in the para-position of the aromatic nucleus in analogue G24, slightly decreased activity (EC_{50} = 1.22 μM). The presence of the methoxy electron donating group at the same position had a similar effect. The EC_{50} of G25 (2.06 μM) was increased by 1.7- and 5-fold compared to analogues G24 and G23, respectively.

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>G22</td>
<td><img src="image" alt="Image" /></td>
<td>2.12 ± 0.09</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2.81 ± 0.03</td>
</tr>
<tr>
<td>G23</td>
<td><img src="image" alt="Image" /></td>
<td>0.41 ± 0.06</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.47 ± 0.01</td>
</tr>
<tr>
<td>G24</td>
<td><img src="image" alt="Image" /></td>
<td>1.22 ± 0.17</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2.48 ± 0.17</td>
</tr>
<tr>
<td>G25</td>
<td><img src="image" alt="Image" /></td>
<td>2.06 ± 0.13</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4.08 ± 0.13</td>
</tr>
</tbody>
</table>

* Concentrations required to inhibit growth of T. brucei by 50% and 90%, respectively. EC_{50} and EC_{90} data are the mean of triplicate experiments ± SEM.

**Table 5** Activity of the novel guanylhydrazone analogues G26 – G31 tested in vitro against cultured bloodstream form T. brucei (pH 7.4) and T. cruzi epimastigotes.

Cytotoxicity against cultured rat skeletal myoblast L6 cells and selectivity indices.
<table>
<thead>
<tr>
<th>Cm pd</th>
<th>Structure</th>
<th>$T. \text{brucei}$ EC$_{50}$ (μM)</th>
<th>$T. \text{brucei}$ EC$_{90}$ (μM)</th>
<th>SI</th>
<th>$T. \text{cruzi}$ EC$_{50}$ (μM)</th>
<th>$T. \text{cruzi}$ EC$_{90}$ (μM)</th>
<th>SI</th>
<th>L6 cells CC$_{50}$ (μM)</th>
</tr>
</thead>
<tbody>
<tr>
<td>G26</td>
<td><img src="image" alt="Structure G26" /></td>
<td>1.62 ± 0.06</td>
<td>2.16 ± 0.02</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>G27</td>
<td><img src="image" alt="Structure G27" /></td>
<td>1.35 ± 0.09</td>
<td>1.71 ± 0.02</td>
<td>3.4</td>
<td>5.40 ± 1.17</td>
<td>11.2 ± 1.0</td>
<td>0.9</td>
<td>4.65 ± 0.54</td>
</tr>
<tr>
<td>G28</td>
<td><img src="image" alt="Structure G28" /></td>
<td>1.99 ± 0.06</td>
<td>2.72 ± 0.12</td>
<td>6.5</td>
<td>3.47 ± 0.15</td>
<td>4.90 ± 0.12</td>
<td>3.7</td>
<td>12.9 ± 0.3</td>
</tr>
<tr>
<td>G29</td>
<td><img src="image" alt="Structure G29" /></td>
<td>6.88 ± 0.72</td>
<td>13.1 ± 0.7</td>
<td>3.7</td>
<td>6.14 ± 0.33</td>
<td>11.6 ± 0.3</td>
<td>4.2</td>
<td>25.5 ± 0.6</td>
</tr>
<tr>
<td>G30</td>
<td><img src="image" alt="Structure G30" /></td>
<td>&gt;30</td>
<td>-</td>
<td>-</td>
<td>&gt;30</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>G31</td>
<td><img src="image" alt="Structure G31" /></td>
<td>14.3 ± 1.2</td>
<td>21.1 ± 0.6</td>
<td>13</td>
<td>&gt;30</td>
<td>-</td>
<td>-</td>
<td>186 ± 11</td>
</tr>
</tbody>
</table>

$^a$ Concentrations required to inhibit growth of $T. \text{brucei}$ and $T. \text{cruzi}$ by 50% and 90%, respectively. EC$_{50}$ and EC$_{90}$ data are the mean of triplicate experiments ± SEM. $^b$ Selectivity indices were calculated as the ratio of the CC$_{50}$ for L6 cells to EC$_{50}$ for $T. \text{brucei}$ or $T. \text{cruzi}$, respectively. $^c$ Cytotoxicity was determined by establishing the concentration required to inhibit growth of cultured L6 cells by 50% (CC$_{50}$). Data are the mean of triplicate experiments ± SEM.

The novel analogues G26 – G31 were tested for their trypanocidal potency against both cultured bloodstream form $T. \text{brucei}$ and $T. \text{cruzi}$ epimastigotes in vitro. The resulting EC$_{50}$ and EC$_{90}$ values are presented in Table 5. Cytotoxicity values (CC$_{50}$) and selectivity indices (SIs) were also determined. The compounds bearing an adamantane core (G26 – G28) were the most active, being in the low micromolar range, while the replacement of the adamantane scaffold led to compounds with decreased potency, in a higher micromolar level, with the exception of compound G30, which was inactive. More specifically, compounds G26 – G29 and G31 exhibited considerable activity against $T. \text{brucei}$ with EC$_{50}$ values ranging from 1.35 to 14.3 μM. With the G26 analogue, the guanylhydrazone group was shifted in the methylene carbon of the benzyl substituent, with respect to the parent compound G23. This structural modification resulted in a higher EC$_{50}$ value by 4 times (EC$_{50}$ G26 = 1.62 μM; EC$_{50}$ G23 = 0.41 μM). The respective 2-substituted analogue G27 was
equipotent to G26 (EC₅₀ G27 = 1.35 µM). Ring closure by formation of the C1 – C’2 bond of G27 resulted in the adamantane-cyclopentane-benzene fused analogue G28, which exhibited a similar inhibitory activity (EC₅₀ G28 = 1.99 µM). Replacing the adamantane component with a second benzene ring lowered the activity observed for the analogue G29 by 3.5 fold (EC₅₀ G29 = 6.88 µM), while changing the linker cyclopentane ring to the bulkier cycloheptane in compound G30 abolished the activity. Finally, the use of a completely different and non-rigid scaffold, like the one of such as the 4-phenyl cyclohexane (G31), seems to compromise disfavor activity compared to the structurally related analogue G29, as G31 was 2-fold less effective.

Most of the newly synthesized compounds were also tested against T. cruzi epimastigotes. Derivatives G27, G28 and G29 were the most potent, with EC₅₀ values being 5.40 µM, 3.47 µM and 6.14 µM, respectively. On the other hand, G30 and G31 showed no inhibition in vitro at 30 µM, which was the highest concentration tested. The cytotoxicity of the target compounds against mammalian cells was determined using the rat skeletal myoblast L6 cell line. The CC₅₀ values ranged between 4.65 for analogue G27 and 186 µM for G31. Moreover, selectivity indices observed for T. brucei were generally higher than those observed for T. cruzi, as shown in Table 5.

3.3 Computational studies

S-Adenosylmethionine Decarboxylase

Recently, Volkov and colleagues [157] elucidated the crystal structure of the T. brucei AdoMetDC/prozyme heterodimer in complex with a carbonimidic dihydrazide derivative well-known for its AdoMetDC inhibition activity, namely CGP 40215 (CGP) [108]. Structural data from the binding site of the protein-ligand complex reveal that CGP forms several interactions, including a bidentate H-bond with E266 and monodentate H-bonds with the backbone carbonyl of L83, the side chain hydroxyls of S249 and Y243, and with C100. Additionally, CGP is further stabilized through π-π stacking between Y243 and F28, as well as cation-π interactions with F28 (Fig. (6)). Due to the structural similarity of CGP with our guanylhydrazone analogues, it was thus hypothesized that both will share a common binding pattern, rendering these as interesting and novel in drug development for AdoMetDC. In order to test this hypothesis, in silico docking experiments were conducted. Furthermore, guanylhydrazone compound characterisation and model validation was obtained through an extended SAR and QSAR analysis.
Fig. (6). Superposition of CGP and docking-fitted guanylhydrazone analogue G8 within the AdoMetDC active site. The computationally cleaved three-dimensional crystal structure of *T. brucei* AdoMetDC (PDB: 5TVF) is shown in “ribbon” representation, colored in pink. Selected important residues contributing to various interactions are highlighted in “stick” representation. The pink “bean”-like shape on the right, symbolizes the hydrophobic cavity of AdoMetDC oriented towards the back of the binding site. CGP, the compound co-crystallized with AdoMetDC, is colored in green and the nanomolar range *T. brucei* inhibitor, guanylhydrazone analogue G8, is shown in light blue. Intriguingly, the positions of benzene and adamantane groups of CGP and G8, respectively, coincide and both compounds share common binding patterns with E266, as illustrated. CGP is further stabilized through two more hydrogen bonds with S249 and L83. Observed π-π stacking between Y243/F28 and cation-π interactions with F28, for CGP are not depicted in this figure.

**Results of Docking-based scoring**

Docking results shed light on the binding pattern of the guanylhydrazone analogues, which exhibit high resemblance compared to the crystal structure of the AdoMetDC and CGP complex, as shown in Fig. (6). Generally, all analogues form mono- and/or bidentate H-bonds through their guanyl group, either with E266 or with C246, and a minority of them with L83 or E85. Remarkably, the aliphatic side chains (that substitute the adamantane group) of some analogues stay inside the active site of
AdoMetDC in top-ranked poses scored within the broader $15 \times 15 \times 15$ Å generated grid. However, this was not the case with the $10 \times 10 \times 10$ Å grid, which seemingly impelled steric clashes with all lipophilic side chains due to space confinement. Therefore, the novel analogues $G26 – G28$ (Fig. (7A)) ranked on top even though their $T. brucei$ activity proves otherwise. Due to this observation, all subsequent studies were performed within the broader grid.

**Fig. (7).** Docking results for several guanylhydrazone analogues. A) Guanylhydrazone analogues $G26$ (pale green), $G27$ (pale yellow) and $G28$ (pale red), three novel low micromolar $T. brucei$ inhibitors, share identical polar interaction with E266 and position within the AdoMetDC active site, hence their minor activity variance. B) The binding profile of dietheric guanylhydrazone analogue $G18$ (copper yellow). Formation of an intramolecular H-bond probably does not occur, due to stabilization via hydrogen bond with C100. Similar interactions
are observed for dietheric analogue G17. C) Superposition of guanylhydrazone analogues: G12 (white), G16 (orange), G13 (purple), G14 (dark grey) and G15 (light grey). Active analogues G16 and G13 exhibit indistinguishable binding patterns, in contrast to the inactive analogue G12, which despite forming an H-bond with E266 of AdoMetDC, its large aliphatic group bolts towards the bulk of the solvent. Inactive analogues G14 and G15 score amongst the bottom ranks in docking due to the lack of polar interactions with AdoMetDC, in addition to the “false” orientation within the active site; the hydrophilic guanylhydrazone group faces towards the solvent and the hydrophobic adamantane group proximates polar E266.

Note that the substitution of the adamantane group at the β-position of the etheric oxygen (G12, G14, G15 and G20) eliminates the activity when compared to the α- and γ- position (G19 and G16, G21 respectively), with the exception of analogue G13 and the dietheric analogues G17 and G18. This unexpected bioactivity result was further scrutinized by using molecular mechanics simulations, which yielded a high possibility of intramolecular H-bond formation between the etheric oxygen and the guanyl-group nitrogen, hence reduction in the in vitro activity. However, this was not applicable to the analogue G13. Furthermore, docking structural data show that the aliphatic chain of G12 is placed outside of the AdoMetDC active site, resulting in the a lack of activity. Additionally, as illustrated in Fig. (7C), analogues G14 and G15 exhibit a dissimilar binding pattern to that previously described, with the guanylhydrazone group oriented towards the bulk of the solvent. Thus, no polar interactions with the protein are made and the activity becomes marginal. In contrast, dietheric analogues G17 and G18 exhibit T. brucei inhibitory activity, probably due to the second etheric oxygen interacting with AdoMetDC via an H-bond with the backbone of Cys100 (Fig. (7B)).

Docking-based structural data also failed to provide a rationale regarding the SAR analysis on slight side-chain differentiations among guanylhydrazone analogues. For instance, amongst the analogues G6 – G9 (of which G8, with a 10-carbon aliphatic chain, is the most potent guanylhydrazone inhibitor), analogue G9, with a 14C substitution, displays slightly diminished activity compared to G8. This outcome, could be termed, paraphrasing Schönerr and Cernak [158], as the “magic ethyl effect”. It is therefore possible that reliance on solely docking-based approaches could be insufficient to interpret the SAR. As an alternative, a thermodynamic-driven approach was therefore implemented.

The thermodynamics of binding is influenced by multiple factors, including hydrogen bonding, hydrophobic interactions, protein and ligand rigidification, as well as the solvation and desolvation effect [159]. Therefore, avoiding a linearly correlated rule of thumb approach is vital when attempting to interpret binding affinity data. Even though H-bonds in theory favour enthalpy, hydrophobic
interactions favor entropy and conformational changes are unfavorable for entropy. One needs to consider that enthalpy and entropy of an intricate molecular system are vaguely interconnected. The theory of Enthalpy–Entropy Compensation (EEC) [160,161], which states that the enthalpy gain of a rigid protein-ligand complex is compensated by an entropy loss caused by the decrease of the available states for the system, provides partial explanation. However, as Wienen-Schmidt and colleagues showed, this is not always the case. Highest flexibility ligands can display, counterintuitively, the entropically most favored binding. Thus, the loss of degrees of freedom of the ligand does not necessarily monopolize the thermodynamic profile [162]. Adequate explanation for this unexpected thermodynamic outcome could be offered by the role of water molecules, which also strongly affect protein and ligand stability through their dipolar nature and the entropic gain associated with the hydrophobic effect [163]. Water molecules of the protein-ligand complex can be classified on the basis of their energetics into either high energy, unstable water molecules trapped inside hydrophobic cavities or stable water molecule that are stabilized through extensive interactions with the protein over the corresponding solvent accessible surface [164]. Computational tools such as SZmap [165] from Openeye, have been used to incorporate the water network in drug design. SZmap (solvent-Zap-mapping) is capable of quantitatively predicting water positions and their energetics, and is potentially a useful tool for identifying significant favorable or unfavorable regions of solvent thermodynamics in the binding site [166]. The hydration analysis via SZmap for the active site of AdoMetDC was performed as previously described [167] and the results were visually inspected.

Predictably, SZmap solvation mapping analysis yielded interesting results that compensated for the limitations of a merely docking-scoring approach. The binding site of AdoMetDC is rich in both stable and unstable water molecules, which through their displacement influence binding thermodynamics. Since all analogues are comprised of bulky hydrophobic groups, such as aliphatic hydrocarbons, adamantane and benzene moieties, water molecules are inevitably displaced. With the introduction of the polar guanylhydrazone group, which ideally contributes to favorable enthalpy via structural water molecule displacement, elucidating the complete thermodynamic binding profile becomes complex. Specifically, for G8, a nanomolar-range T. brucei inhibitor, SZmap data showed that the adamantane group as well as the 10C side chain displace unstable, high-energy water molecules (Fig. (8)), thus contributing to the increase of entropy.
**Fig. (8).** SZmap-generated hydration analysis of the active site of AdoMetDC with docking-fitted G8 analogue. The red spheres define the high energy water molecules, while gold spheres depict structural water molecules. Selected amino acids are illustrated for viewing reference. The inhibitor is oriented so as to position both hydrophobic adamantane and 10-carbon aliphatic groups towards unstable water molecules, thus permitting their displacement and leading to affinity enhancement due to favorable entropic contributions. The polar guanylhydrazone group thereby provokes the displacement of three highly stable water molecules, from the hydration site adjacent to E266, and the enthalpic effect is pronounced.

In contrast, the guanylhydrazone group displaces low energy structural water molecules and therefore adds to the enthalpic effect. Hydration data for analogue G9 show a contiguity, with the exception of the adamantane group, which does not displace water molecules in this context. Apparently, this guanylhydrazone analogue poorly docks within the active site of AdoMetDC. Therefore, a displacement of three stable water molecules by the adamantane group would occur, resulting in enthalpic cost, which conceivably eliminates the activity against *T. brucei*. Analogous interpretation of the displacement of high/low energy water molecules by polar or hydrophobic groups can also be made for other guanylhydrazone analogues of similar structure, but different activity (G10 and G11;
G24 and G25; G2 and G3). These differences could be related to EEC and should be further investigated.

**Results of QSAR analysis**

All statistical methods comprised of only the active analogues yielded substantial model fidelity, based on their statistical properties, as illustrated in Table 6.

<table>
<thead>
<tr>
<th>Statistical Regression Model</th>
<th>Training set</th>
<th>Test set</th>
<th>R²</th>
<th>Q²</th>
<th>SD</th>
<th>RMSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>MLR</td>
<td>22</td>
<td>5</td>
<td>0.9339</td>
<td>0.8270</td>
<td>0.2510</td>
<td>0.3569</td>
</tr>
<tr>
<td>PLS</td>
<td>22</td>
<td>5</td>
<td>0.8890</td>
<td>0.3573</td>
<td>0.3350</td>
<td>0.5466</td>
</tr>
<tr>
<td>KPLS</td>
<td>22</td>
<td>5</td>
<td>0.9662</td>
<td>0.9105</td>
<td>0.1752</td>
<td>0.2922</td>
</tr>
</tbody>
</table>

Interestingly, the most accurate regression model was generated through KPLS (R² = 0.9662; Q² = 0.9105). PLS had less predictive accuracy (R² = 0.8890; Q² = 0.3575). MLR analysis also produced an acceptable model (R² = 0.9339; Q² = 0.8270). However, it was challenged by notable intercorrelation within the Canvas-derived optimal subset of descriptors, thus rendering the respective MLR model as one of limited predicting capacity. Out of the molecular descriptors identified as significant by the statistical analysis, the most meaningful were the following; average vertex distance degree, Balaban distance connectivity index, conformation independent QPlogS, Gutman molecular topological and topological charge index of order 9. Details description of these descriptors can be acquired from the Handbook of Molecular Descriptors, by Todeschini and colleagues [168]. Fundamentally, molecular descriptors, due to the involvement of mathematics, usually lack direct structural interpretation of a molecule. Since chemical information is transformed through various procedures into sole numeric values, deducing tangible characteristics about the structure of a compound is rather difficult. That said, it is important to construct a reliable and highly predictable model, regardless of its ability to interpret structural features. Such a model would enable rapid and accurate activity predictions, even for novel compounds currently non-synthesized, and provide proof-of-concept for further drug development approaches.

Experimental and predicted EC50 values of active guanlyhydrazone analogues for *T. brucei* are shown in Table 7.
Table 7. Statistical analysis comparison for the active guanylhydrazone analogues with Multiple Linear Regression (MLR), Partial Least Squares regression (PLS) and Kernel-based Partial Least Squares regression (KPLS). Experimental and predicted *T. brucei* inhibitory activity for only active guanylhydrazone analogues is presented. IC$_{50}$ (μM) values are logarithmically converted and the absolute error is calculated. Error values are color-coded. As the absolute error approximates the numeric value of one, activity shifts one order of magnitude.

<table>
<thead>
<tr>
<th>Active Compound</th>
<th>Activity (-logIC$_{50}$, M)</th>
<th>MLR</th>
<th>PLS</th>
<th>KPLS</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Pred(Activity) (-logIC$_{50}$, M)</td>
<td>Abs Error</td>
<td>Pred(Activity) (-logIC$_{50}$, M)</td>
<td>Abs Error</td>
</tr>
<tr>
<td>G1</td>
<td>4.245</td>
<td>0.184</td>
<td>4.3413</td>
<td>0.096</td>
</tr>
<tr>
<td>G2</td>
<td>3.777</td>
<td>0.076</td>
<td>3.8435</td>
<td>0.066</td>
</tr>
<tr>
<td>G3</td>
<td>5.015</td>
<td>0.128</td>
<td>3.9843</td>
<td>1.030</td>
</tr>
<tr>
<td>G4</td>
<td>3.523</td>
<td>0.071</td>
<td>3.4518</td>
<td>0.071</td>
</tr>
<tr>
<td>G5</td>
<td>3.523</td>
<td>0.451</td>
<td>3.4956</td>
<td>0.027</td>
</tr>
<tr>
<td>G6</td>
<td>5.222</td>
<td>0.219</td>
<td>5.5995</td>
<td>0.378</td>
</tr>
<tr>
<td>G7</td>
<td>5.222</td>
<td>0.721</td>
<td>5.9604</td>
<td>0.738</td>
</tr>
<tr>
<td>G8</td>
<td>7.046</td>
<td>0.375</td>
<td>6.3055</td>
<td>0.741</td>
</tr>
<tr>
<td>G9</td>
<td>6.076</td>
<td>0.122</td>
<td>6.7108</td>
<td>0.635</td>
</tr>
<tr>
<td>G10</td>
<td>5.551</td>
<td>0.000</td>
<td>5.3923</td>
<td>0.159</td>
</tr>
<tr>
<td>G11</td>
<td>6.377</td>
<td>0.169</td>
<td>6.4279</td>
<td>0.051</td>
</tr>
<tr>
<td>G13</td>
<td>5.932</td>
<td>0.130</td>
<td>5.8398</td>
<td>0.092</td>
</tr>
<tr>
<td>G16</td>
<td>6.155</td>
<td>0.093</td>
<td>5.8456</td>
<td>0.309</td>
</tr>
<tr>
<td>G17</td>
<td>5.439</td>
<td>0.011</td>
<td>5.4164</td>
<td>0.023</td>
</tr>
<tr>
<td>G18</td>
<td>5.441</td>
<td>0.009</td>
<td>5.3975</td>
<td>0.043</td>
</tr>
<tr>
<td>G19</td>
<td>5.735</td>
<td>0.196</td>
<td>5.8350</td>
<td>0.100</td>
</tr>
<tr>
<td>G21</td>
<td>5.75</td>
<td>0.057</td>
<td>6.0994</td>
<td>0.349</td>
</tr>
<tr>
<td>G22</td>
<td>5.674</td>
<td>0.032</td>
<td>5.7052</td>
<td>0.031</td>
</tr>
<tr>
<td>G23</td>
<td>6.387</td>
<td>0.254</td>
<td>5.9330</td>
<td>0.454</td>
</tr>
<tr>
<td>G24</td>
<td>5.914</td>
<td>0.016</td>
<td>5.9442</td>
<td>0.030</td>
</tr>
<tr>
<td>G25</td>
<td>5.686</td>
<td>0.117</td>
<td>5.8503</td>
<td>0.164</td>
</tr>
<tr>
<td>G26</td>
<td>5.790</td>
<td>0.209</td>
<td>5.8412</td>
<td>0.051</td>
</tr>
<tr>
<td>G27</td>
<td>5.870</td>
<td>0.255</td>
<td>5.9097</td>
<td>0.040</td>
</tr>
<tr>
<td></td>
<td>G28</td>
<td></td>
<td>G29</td>
<td></td>
</tr>
<tr>
<td>---</td>
<td>-------</td>
<td>---</td>
<td>-------</td>
<td>---</td>
</tr>
<tr>
<td></td>
<td>5.701</td>
<td></td>
<td>5.6679</td>
<td></td>
</tr>
<tr>
<td></td>
<td>5.6679</td>
<td></td>
<td>4.8028</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.033</td>
<td></td>
<td>0.179</td>
<td></td>
</tr>
<tr>
<td></td>
<td>5.7783</td>
<td></td>
<td>4.7354</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.077</td>
<td></td>
<td>0.360</td>
<td></td>
</tr>
<tr>
<td></td>
<td>5.590</td>
<td></td>
<td>4.8898</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.111</td>
<td></td>
<td>0.080</td>
<td></td>
</tr>
</tbody>
</table>

As can be observed, all three regression models yielded medium to relatively high absolute error values for analogues G7 and G8. A possible explanation for this observation can be, again, related to the effect of the solvent. Software implementing the ligand-based drug discovery approach, such as Canvas, are exclusively focused on the ligand structure and can thus seriously neglect the exact thermodynamic effect that individual water molecules might have on the binding affinity of small molecules that can displace them. The displacement of water molecules affects each compound differently, hence the almost absolute agreement for analogues G2, G17, G18, G22, G24 among all statistical approaches. Generated statistical data should be considered as a stepping-stone for further investigation.

**Conclusion**

In summary, the data obtained from this extended SAR study illustrate: (i) The synergistic effect of the lipophilic character of the C1 substitution and C2 guanylhydrazone moiety on trypanocidal activity. (ii) The critical role of the adamantane moiety for trypanocidal activity. Replacement of the adamantane cage structure by the tricyclic scaffolds of fluorene and dibenzocycloheptene, and the bendable 4-phenyl cyclohexane decreases activity against *T. brucei*. (iii) The great importance of the nature and length of the second substitution to antiparasitic activity. Replacement of the lipophilic alkyl chain at of the adamantane-based guanylhydrazones with a more polar component (alkyloxy, hydroxyl, aminoguanidine) results in a substantial reduction in trypanocidal properties. Finally, a number of diverse approaches were undertaken to understand the interplay between the structure of the promising bioactive compounds and the activity against their putative protozoan target, AdoMetDC. Chemoinformatics-based models afforded a series of molecular descriptors that might account for the observed bioactivity and sustain the design of new derivatives. In addition, molecular docking simulations, enhanced by a hydration analysis, derived interesting thermodynamic insights into the role of structural factors, such as intramolecular hydrogen bonding or displacement of energetically unfavorable water molecules on the observed binding affinity. The design of new compounds with improved bioactivity against AdoMetDC could greatly benefit from such rational approaches. Efforts towards this direction are currently underway.
LIST OF ABBREVIATIONS
3D : Three-dimensional
AdoMet : S-adenosylmethionine
AdoMetDC : S-adenosylmethionine decarboxylase
BBB : blood-brain barrier
BZ : benznidazole
CADD : Computer-aided drug design
CATT : card agglutination test for trypanosomiasis
CNS : central nervous system
CSF : cerebrospinal fluid
CC50 : half maximal cytotoxic concentration, concentration required to cause cytotoxic effect by 50% or to inhibit growth of cultured L6 cells by 50%
CGP : derivative CGP 40215 (AdoMetDC inhibitor)
DMF : N,N-dimethylformamide
DMSO : dimethyl sulfoxide
DRC : Democratic Republic of Congo
EC50 : half maximal effective concentration, concentration required to give half-maximal response or concentration required to inhibit growth of T. brucei and T. cruzi by 50%
EEC : Enthalpy−Entropy Compensation
Et2O : diethyl ether
EtOH : ethanol
FBDD : Fragment-based drug discovery
gHAT : gambiense HAT
Gsp : glutathionylspermidine
HAT : Human African Trypanosomiasis
LBDD : Ligand-based drug design
IgM : Immunoglobulin M
IM : intramuscular
IV : intravenous
KPLS : Kernel-based Partial Least Squares
MLR : Multiple Linear Regression
Acknowledgements:
G.Z. would like to thank Empirikon Foundation for the financial support. V.P. would like to thank the State Scholarships Foundation of Greece for providing her a Ph.D. fellowship (MIS-5000432, NSRF 2014–2020). K.M.M. and M.Š. thank the Croatian Ministry of Science, Education and Sports for the financial support of this study (grant number 098-0982933-2911).

Conflict of Interest
There are no conflicts to declare.

Funding:
Croatian Ministry of Science, Education and Sports (grant number 098-0982933-2911).

Consent for Publication:
All the authors consent for this publication.

REFERENCES


[43] Kansiime F, Adibaku S, Wamboga C. A multicentre, randomised, non-inferiority clinical trial comparing a nifurtimox-eflornithine combination to standard eflornithine monotherapy for late stage


[165] SZMAP, 1.2.1 OpenEye Scientific Software Inc. , Santa Fe, NM, USA, 2015.

